*data*

MDPI

*Review*

# Beyond the List: A Framework for the Design of Next-Generation MEDLINE Search Tools

**Vladimir Zhurov [1,*], Kamran Sedig [1,2] and Mostafa Milani [1]**

[1] Faculty of Science, Department of Computer Science, University of Western Ontario, Room 355, Middlesex College, London, ON N6A 5B7, Canada; sedig@uwo.ca (K.S.); mostafa.milani@uwo.ca (M.M.)
[2] Faculty of Information and Media Studies, University of Western Ontario, FIMS & Nursing Building, Room 2050, London, ON N6A 5B9, Canada
* Correspondence: vladimir.zhurov@uwo.ca

**Abstract**

Despite the critical importance of biomedical databases like MEDLINE, users are often hampered by search tools with stagnant designs that fail to support complex exploratory tasks. To address this limitation, we synthesized research from visual analytics and related fields to propose a new design framework for non-traditional search interfaces. This framework was built upon seven core principle: visualization, interaction, machine learning, ontology, triaging, progressive disclosure, and evolutionary design. For each principle, we detail its rationale and demonstrate how its integration can transcend the limitations of conventional search tools. We contend that by leveraging this framework, designers can create more powerful and effective search tools that empower users to navigate complex information landscapes.

**Keywords:** MEDLINE; information retrieval; exploratory search; search interfaces; design framework; visual analytics; interaction design

check for updates

Academic Editor: Han Woo Park

Received: 12 September 2025
Revised: 15 October 2025
Accepted: 18 October 2025
Published: 21 October 2025

**Citation:** Zhurov, V.; Sedig, K.; Milani, M. Beyond the List: A Framework for the Design of Next-Generation MEDLINE Search Tools. *Data* **2025**, *10*, 167. https://doi.org/10.3390/data10100167

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

MEDLINE [1] is a critical biomedical database containing over 31 million records from all areas of healthcare, maintained by the U.S. National Library of Medicine. Researchers, clinicians, and patients (henceforth, users) rely on MEDLINE to conduct research and stay current in the medical domain. Given MEDLINE's scale and importance, the tools used to search it must effectively support users in finding relevant information. This paper addresses the shortcomings of traditional search tools by establishing foundational principles for the design of more effective, non-traditional interfaces for MEDLINE.

To use MEDLINE, users engage in complex search tasks, typically via traditional tools such as PubMed [2]. However, these tools frequently fail to leverage users' prior knowledge or provide sufficient contextual cues, hindering effective query formulation and results assessment [3]. For example, Herceg et al. [4] found that users routinely struggled to understand the search domain, articulate their information-seeking goals, and assess the relevance of results. When a tool fails to facilitate these activities, its effectiveness is significantly compromised, leaving users overwhelmed by information and unable to make rapid relevance judgments [4]. In short, the design and capabilities of a search tool directly dictate a user's ability to complete their tasks successfully.

As an illustrative case, PubMed, the primary tool for biomedical literature searches, is widely used but has notable limitations [5]. PubMed requires the knowledge of specific

medical vocabularies and often returns an unmanageably large number of results [6]. For instance, Salvador-Oliván et al. [7] crafted an optimally detailed PubMed query (over 1000 characters long) to retrieve systematic reviews with high precision, but even this expert query returned more than 45,000 results. In practice, most users do not construct such elaborate queries, tending instead to use only a few keywords over very few iterative attempts [8]. Furthermore, Morshed and Hayden [9] observed that for quick clinical searches, PubMed's success rate was worse than that of almost all other tools, outperforming only a generic Google search. These issues make it evident that PubMed and similar traditional search tools are ill-equipped to fully support many of their users' information retrieval needs.

While some prior research has investigated search tool design, it often yields fixed lists of features or requirements that assume a traditional interface. For example, Gusenbauer and Haddaway [10] enumerated desired features for academic search systems but derived their requirements largely from existing search paradigms. Because such efforts extrapolate from traditional designs, they risk perpetuating the very limitations they aim to address. A clear gap exists in research that explores fundamentally different solutions to the inherent problems of traditional search.

One area of research that offers novel solutions is visual analytics, particularly through tools for searching and triaging large document sets [6,11,12]. These tools integrate interactive visualizations and advanced data analytics to empower users in complex, data-driven tasks [13]. They enable a richer exploration, analysis, and organization of document collections than is possible with a static list of results [14–17]. Alongside visual analytics, technologies such as ontologies [18] and large language models (LLMs) [19] have also been studied to inform more effective search tool design. It is from these lines of research that we can identify principles to guide a new approach.

Building on this foundation, this paper identifies and discusses seven key design principles for developing next-generation search tools for MEDLINE. These principles were selected because they directly address the core limitations consistently identified in the literature evaluating traditional search interfaces, MEDLINE search, and large-scale document analysis. While not exhaustive, these seven principles—visualization, interaction, machine learning, ontology, triaging, progressive disclosure, and evolutionary design—provide a robust framework for conceptualizing more effective search tools. This paper examines how each principle influences the search process and helps transcend the shortcomings of conventional designs, thereby presenting practical considerations for creating more powerful, non-traditional MEDLINE search tools.

The rest of this paper is structured as follows. We first provide background on traditional search paradigms before introducing the seven design principles. Subsequently, each principle is discussed in detail, with examples from prior work illustrating its implications. We then examine the principles holistically, highlighting their interconnectivity and trade-offs. The paper concludes with a summary of the findings.

## 2. Background

To understand the need for a new design framework, it is essential to first analyze the functional limitations of traditional search tools. This section begins by deconstructing the search process into its core sub-tasks: query building, parsing, and analysis; afterward, it examines how conventional interfaces like PubMed support them. This analysis reveals the critical shortcomings that motivated the introduction of the seven design principles that form the foundation of our proposed framework.

*Traditional Search Tools*

Traditional search tools, such as Google and PubMed, rely heavily on keyword matching and fixed algorithms. While effective at processing vast datasets, they typically return lists of results without conveying a deeper contextual understanding of the query or user intent. Their interfaces have converged on a standard design comprising a few common elements: a text search bar for query input (Figure 1a); filter options for refining results (Figure 1b); a list of titles or snippets to display results (Figure 1c); pagination controls (Figure 1d); and occasionally a simple visualization such as PubMed's timeline bar chart (Figure 1e).

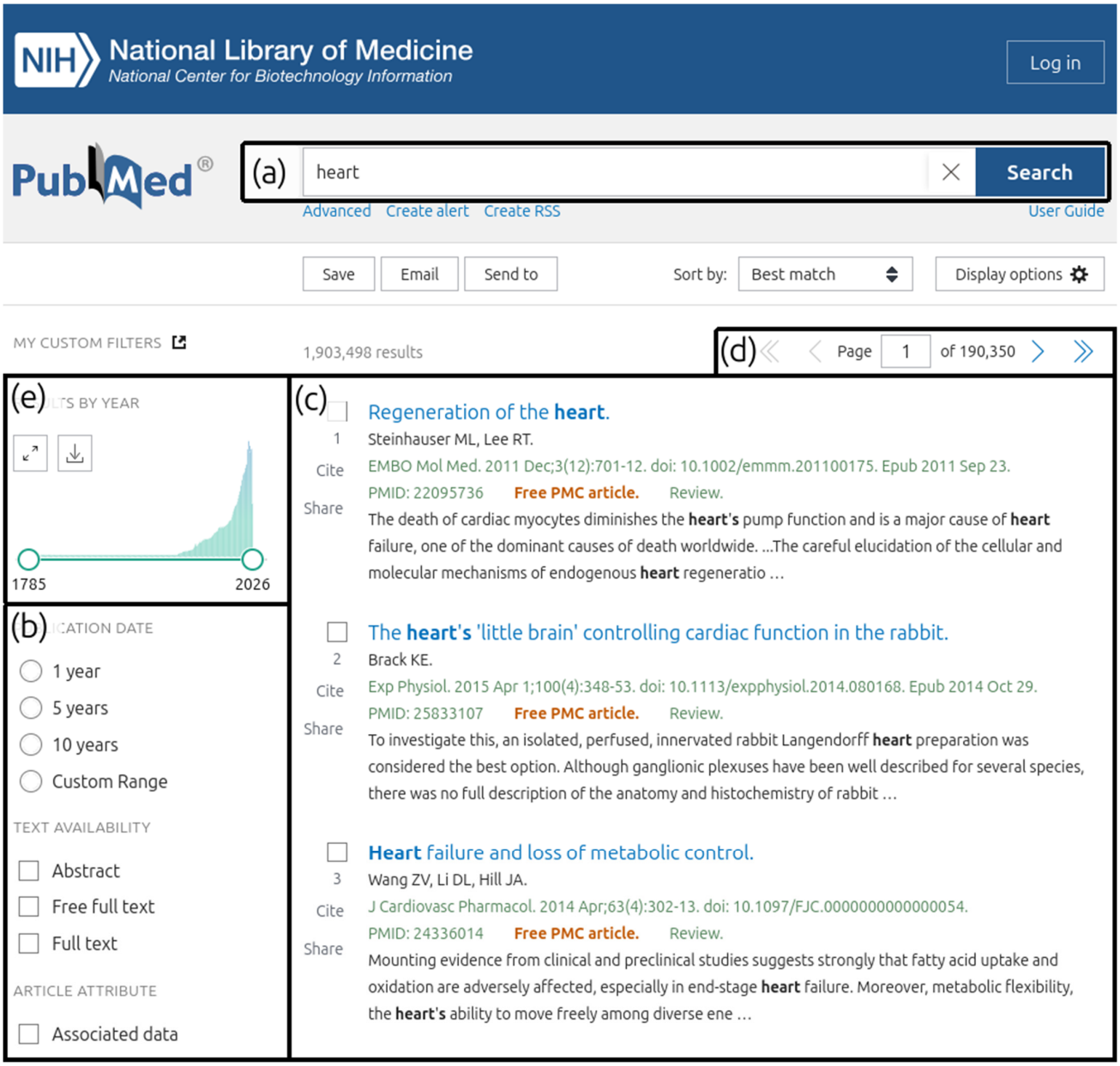



**Figure 1.** An annotated screenshot of PubMed, a traditional search tool, divided into five sections: text search bar (**a**); filter options, partial (**b**); result list, partial (**c**); page buttons (**d**); and timeline bar chart (**e**). Source: Unannotated image generated on 6 June 2025, using the public web portal provided by The National Library of Medicine, https://pubmed.ncbi.nlm.nih.gov/?term=heart (accessed on 6 June 2025).

A complete search task can be conceptualized as three sequential sub-tasks: **query building, result parsing,** and **document analysis**. Initially, users must build a query to communicate their information need. In traditional interfaces, this is confined to typing keywords into a search bar and applying pre-set filters. While alternative query-building mechanisms exist, such as hierarchical tree structures [20] or spatial layouts [21], they are not common. Next, users must parse the results to identify potentially relevant documents. This step involves scanning result lists and using supplementary cues, like PubMed's timeline, to triage the information. Researchers have proposed more advanced representations to aid parsing including graphical overviews [17,22] and spatial layouts [23,24]. Finally, users analyze selected documents to determine whether they contain the desired information. Traditional tools offer minimal support for analysis, typically redirecting the user to the full document on a separate page. While methods for in-tool analysis like word cloud visualizations [25] or automatic summaries [26] have been suggested, they are rarely integrated into mainstream tools.

This functional breakdown reveals a significant imbalance; traditional search tools primarily support the parsing sub-task while offering minimal assistance for query building or in-depth analysis. They seldom leverage users' prior domain knowledge or accommodate their cognitive needs during the search process [3,6]. The consequences are well-documented; when faced with a large number of results, users tend to examine only the first page, potentially missing important information [8]. Although extensive research has been conducted to address these issues, from evaluating existing tools [10] to developing entirely novel semantic or visual interfaces [17,27,28], the prevailing design paradigm remains largely unchanged. This persistent gap between user needs and tool capability demonstrates that a principled, framework-driven approach to search interface design is necessary.

## 3. A Framework for the Design of MEDLINE Search Tools

Having mentioned the seven principles of the design framework, this section demonstrates how they can be systematically applied to address the specific challenges of searching MEDLINE. For each principle, we first identify a key limitation inherent in traditional MEDLINE search tools. We then analyze how the targeted application of that principle can overcome this limitation, with a focus on supporting the core search sub-tasks of query building, parsing, and analysis. Finally, we ground this analysis in examples from prior work. This is followed by a discussion of the interconnections among the principles and the trade-offs that may arise between them. By systematically applying this framework, while acknowledging that it is a preliminary, non-exhaustive model, designers can create search tools that are demonstrably more usable and effective.

### *3.1. Visualization*

As the primary medium through which users engage with search tools, visualization is a foundational design principle. Visualizations are graphical representations of information that can facilitate faster perceptual processing than text alone [29]. By compactly encoding vast amounts of information in a given space [6,29], they allow for the nonlinear exploration of data, bypassing the sequential constraints of a traditional results list [30,31]. When made interactive, these visual representations enrich the exchange of information between user and tool, providing superior support for complex cognitive tasks like searches [32].

In practice, visualizations are implemented in search interfaces in several ways. Some systems use visual markers to augment a traditional results list, indicating key document properties at a glance [12,27]. Others supplement the list with adjacent charts or graphs such as PubMed's timeline (Figure 1e) or the EEEvis system's data plots [17]. Recent studies

have applied LLMs to organize search results [33] or to generate charts and graphs [34]. More advanced systems abandon the traditional list entirely, replacing it with a fully visual display for representing and navigating search results [12,35]. This flexibility allows designers to present information in ways that are better aligned with a user's task.

Crucially, visualization can directly support each sub-task of the search process. For query building, visual interfaces have been used to enable alternative forms of query formulation [20,21,36]. To facilitate the parsing of large result sets, various visual approaches have been shown to improve comparison and triage [11,23,27]. Finally, to aid document analysis, techniques like integrated word clouds have provided novel representations of document content [25,37]. Across numerous studies, the effective integration of visualization has been demonstrated to improve the users' search effectiveness and overall experience [6,11,12,27,38].

#### 3.1.1. The Problem: Textual Overload in MEDLINE Search

Users access MEDLINE through representations of its documents, and in traditional tools like PubMed, these representations are overwhelmingly textual. While a title and abstract are more useful than a title alone [39], relying on text creates a significant bottleneck. MEDLINE's sheer scale means that even well-crafted queries can return thousands of results, far more than a user can reasonably parse by reading sequentially [40]. Traditional interfaces offer few visual aids to manage this complexity; the simple timeline chart in PubMed, for example, is a minor component that provides only a single dimension of information (publication date). This forces users into a laborious process of skimming text, making it difficult to spot broader patterns, compare documents efficiently, or quickly assess the overall landscape of the results.

#### 3.1.2. The Solution: Applying Visualization to the Search Workflow

Visualization addresses this challenge by leveraging human perceptual strengths to process large amounts of information in parallel. It can be applied across the entire search workflow: for the parsing sub-task, visualizations can compactly represent thousands of documents, their properties, and their relationships [13,41]. Instead of reading line by line, users can interpret heatmaps, scatter plots, or network graphs to quickly identify clusters, outliers, and distributions in the data, thus accelerating the triage process [17,42]. For analysis, visualizations can reveal patterns within and between documents that are difficult to discern from text. A co-authorship network can instantly reveal key researchers in a field, while a heatmap can show the distribution of query terms within a set of documents, aiding relevance judgments. Visualization can even support query building. By interacting with a visual representation (for instance, selecting a cluster of documents in a scatterplot), a user is effectively performing a visual query to filter the dataset, a more fluid and exploratory process than typing keywords.

#### 3.1.3. Examples in Practice

The following examples from the research literature illustrate these principles.

- **DG-Viz** [35] is a visual analytics tool that uses multiple coordinated visualizations to display large collections of patient records. This allows users to observe data distributions across many documents at once. By leveraging visualization, DG-Viz can show far more records simultaneously than a list-based interface. Instead of reading each document, users can quickly select a region or cluster in a visualization that meets certain criteria (Figure 2A), greatly enhancing the parsing and exploratory querying process.

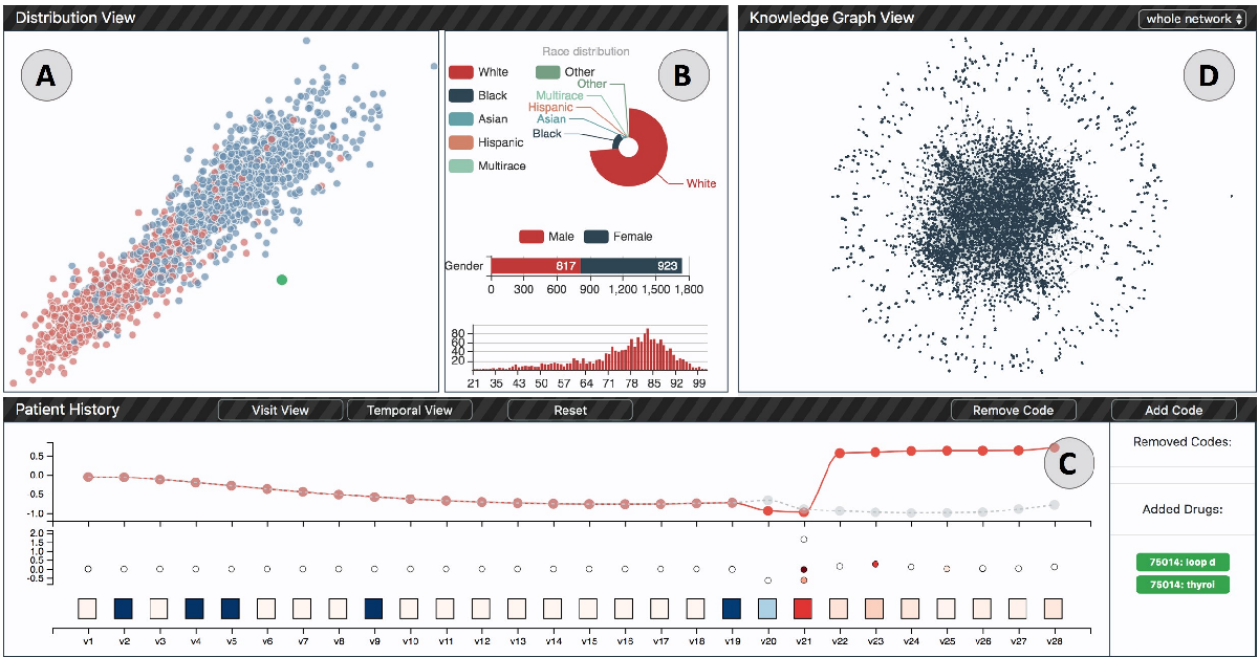



**Figure 2.** DG-Viz is a visual analytics tool with visualizations to present patient records: (**A**) patient distribution view; (**B**) patients demographic charts; (**C**) patient history across visits; and (**D**) knowledge graph of medical codes. Source: Reprinted from the Journal of Medical Internet Research, 22 (9): e20645, Li, R.; Yin, C.; Yang, S.; Qian, B.; Zhang, P. Marrying medical domain knowledge with deep learning on electronic health records: a deep visual analytics approach, Copyright (2020), with permission from Ping Zhang. https://www.jmir.org/2020/9/e20645/ (accessed on 18 September 2024).

- **OVERT-MED** [6] is a tool designed specifically for search and triage in MEDLINE. It represents search results using a heatmap where each document is a horizontal bar and color saturation indicates the presence of query terms (Figure 3). This visual encoding allows users to perform a rapid perceptual scan to identify which results are most relevant to their query, facilitating rapid parsing and comparison without reading any text.

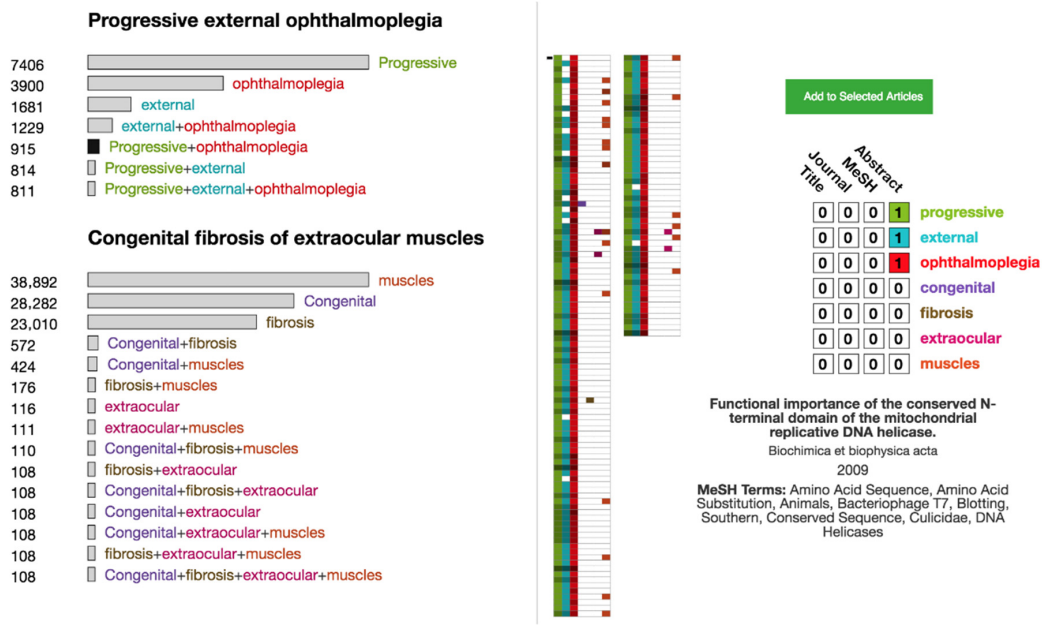



**Figure 3.** A screenshot of two panels of OVERT-MED: an interactive tool that represents search results with visualizations.

- **EEEvis** [17] is another tool designed for MEDLINE search that, in conjunction with a standard list view, provides multiple visualizations. Its co-authorship network (Figure 4) is particularly useful for analysis, as it shows the prevalence and connections between authors in the search results. This visualization allows users to quickly identify key researchers and collaborative groups, a task that would otherwise require significant time and manual effort.

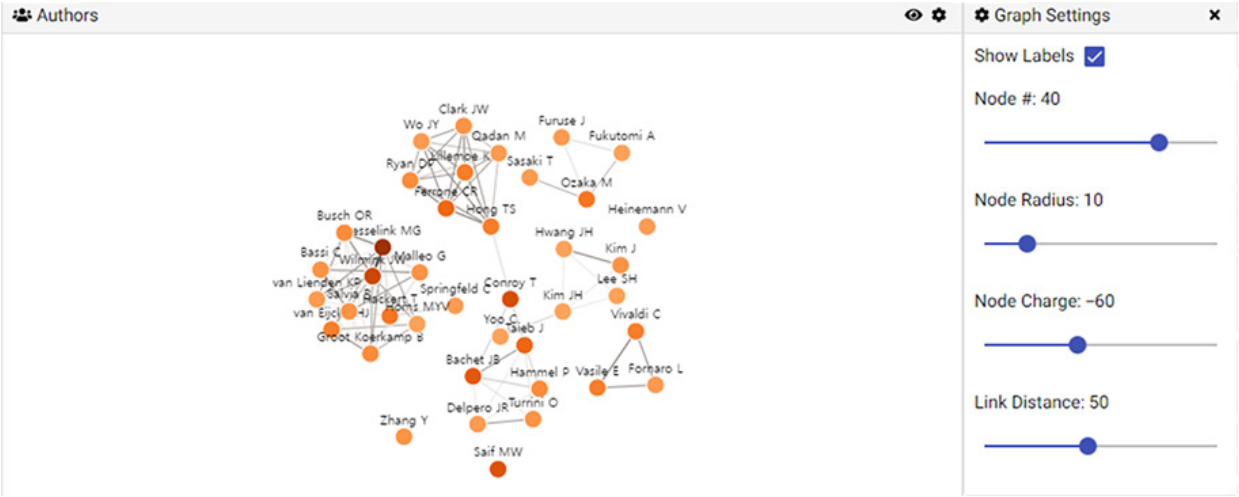



**Figure 4.** An image of EEEvis' co-authorship network and controls. Source: Cropped from Figure 5 in [17], licensed under CC BY 4.0.

These examples demonstrate how integrating visualizations can overcome the limitations of purely text-based interfaces, especially when dealing with large and complex information spaces like MEDLINE. By leveraging human perceptual strengths, visualizations enable users to parse and analyze results more quickly and effectively. Deciding what information to visualize and how is therefore critical considerations in the design of next-generation MEDLINE search tools.

### *3.2. Interaction*

Interaction is the mechanism that enables the exchange of information between the user and the system [32,43]. More than just a means of control, the design of a tool's interactive elements fundamentally shapes the users' cognitive processes during search. Common interactions include annotating results, comparing items, filtering subsets of data, and selecting or bookmarking items of interest [32]. Without a rich set of interactive capabilities, users cannot effectively articulate their needs to a search tool or manipulate the information it returns.

Each stage of the search task is fundamentally dependent on interaction. For query building, user interaction is often limited to typing keywords, with little support for iterative refinement. For parsing results, users must interactively sift through documents to judge relevance. While traditional interfaces allow users to click through pages, they offer few aids for comparison such as viewing results side-by-side. Richer interactions like zooming into a data subset or highlighting matching terms across documents could significantly improve this process. Finally, while document analysis inherently requires interaction (e.g., scrolling, highlighting, note-taking), traditional search tools provide almost no analysis-supporting features within the search interface itself.

Beyond these pragmatic functions, interaction is deeply entwined with what are known as epistemic activities: actions undertaken to support higher-order cognitive tasks like sensemaking and knowledge creation [32,44]. As users search, they engage in these

deeper cognitive processes, with interaction acting as the scaffolding that supports them. Annotation provides a clear example; while its pragmatic effect is to mark a document, its epistemic effect is to engage the user in categorization, comparison, and reflection [32]. It follows that supporting a meaningful breadth of interactions allows users to engage with information more deeply and accomplish their goals more effectively [45].

Because interaction underpins the entire search process, it is a uniquely central principle in this framework. Nearly all of the other principles discussed in this paper, from visualization to machine learning, are expressed to the user and made useful through interaction. Therefore, interaction should be seen not merely as one component among seven, but as a core, unifying element in the design of any effective search tool.

#### 3.2.1. The Problem: The Interaction Deficit in MEDLINE Search

While interaction is fundamental to any search tool, traditional interfaces like PubMed offer only a minimal set of capabilities. Users are typically limited to typing keyword queries, clicking a few filter checkboxes, and paginating through a static list. These tools provide little feedback to help users formulate better queries; the number of results returned is often the only clue to a query's specificity. Crucially, they lack features for manipulating the result set. If a user wishes to compare two promising articles, they must open them separately and assess them manually. PubMed offers no interactive tool to compare documents side-by-side or to group and rearrange results. This rigid design puts the entire cognitive load of synthesis and analysis onto the user [46]. By severely constraining what users can do, this interaction deficit hinders all three sub-tasks: it limits the complexity of query building, makes parsing a laborious manual process, and offers almost no support for deeper analysis.

#### 3.2.2. The Solution: Building a Richer Interactive Environment

Expanding the set of interactions can directly address these limitations and empower users throughout the search workflow. Rather than a single approach, a richer interactive environment can be built by incorporating several key ideas:

1. **Provide proactive feedback**. Interfaces can be designed to signal the expected outcome of an action before it is taken, a concept known as sensitivity encoding [47]. In a search context, the tool could show a real-time estimate of the result count as a user types a query. This immediate feedback helps users gauge a query's restrictiveness and adjust it on the fly, reducing trial-and-error.
2. **Empower user control and manipulation**. Search should not be a one-way street. Richer interactions allow users to actively organize, annotate, and manipulate the result space. The ability to visually group similar documents, mark items as "seen" or "relevant", hide irrelevant results, and rearrange items based on personal criteria allows users to externalize their mental model and manage the information in a way that suits their specific task [44].
3. **Support complex querying and analysis**. Beyond a simple text box, advanced interfaces can allow users to build queries visually or facet-by-facet. For analysis, interactions like selecting a document to see its relationship to all others, or highlighting a term to see its distribution across the entire result set, enable a much deeper engagement with the information than simply clicking a link.

#### 3.2.3. Examples in Practice

- **OVERT-MED** [6] exemplifies the principle of proactive feedback. As a user builds a query, the system displays a real-time estimate of the result count. This simple but powerful interaction gives the user immediate feedback on their query building,

helping them avoid queries that are either too broad or too narrow before they even run the search.

- **NameClarifier** [14], a tool for disambiguating homonymous and synonymous author names in document sets, demonstrates the power of user control. It provides interactive features that let users group search results by inferred author identity, eliminate irrelevant groups, and iteratively refine how the system resolves ambiguity. By supporting these additional interactions, NameClarifier enhances the user's ability to parse and make sense of complex, ambiguous results.
- ChatRetriever [48] uses LLMs to allow users to engage in conversational search. Unlike traditional search, ChatRetriever allows users to articulate their information seeking needs through natural dialogue rather than formal query syntax. The system preserves dialogue context, enabling users to refine and extend their search as the discussion progresses. By leveraging this accumulated context, users' prior interactions inform the system, empowering their ability to communicate and clarify their needs.

These examples show that moving beyond the minimal interactions of traditional tools can make the search process more efficient, transparent, and user-friendly. By consciously considering and expanding the interactive possibilities, designers can create tools that actively support the user's cognitive and epistemic needs.

### *3.3. Machine Learning*

Machine learning (ML), the creation of computational models that learn from data, offers a powerful way to make search tools more adaptive and intelligent. By identifying patterns in large datasets, ML models can automate and enhance key aspects of the search process, moving beyond static, one-size-fits-all algorithms. These models can be trained with or without explicit human-provided labels (i.e., supervised or unsupervised learning) and can even improve iteratively by learning from user interactions (i.e., reinforcement learning) [49,50]. For search tools, the integration of ML is not just a backend enhancement; it is a design principle that directly shapes the user's ability to find relevant information.

ML techniques can be applied to support users across the entire search workflow. For query building, models can power query expansion algorithms that suggest relevant terms [51] or help reformulate queries for greater precision [52]. For parsing results, ML is critical for intelligent ranking, such as PubMed's "Best Match" feature, which prioritizes results based on learned criteria rather than simple chronology [53]. Other models can automatically cluster results into thematic groups, helping users make sense of large, undifferentiated lists [15]. Finally, for document analysis, ML drives features like automatic text summarization [54] and the extraction of key entities from documents [55]. Given the rapid advancement of this field, leveraging ML is a pivotal strategy for creating modern search tools that can understand user needs and manage information overload.

Among the recent advances in ML is the development of LLMs—complex language models trained on large-scale text corpora. Such models can process and generate natural language and generalize across a wide range of tasks [56]. Their adaptability has led to applications in document summarization [57,58], feature extraction [59], and query construction [60]. While many of these capabilities can also be achieved with smaller models, LLMs are particularly effective at leveraging contextual information embedded within user prompts [58,61–63]. Moreover, because LLMs accept natural language input, they enable conversational systems through which users can more effectively communicate their information-seeking needs [48]. These capabilities have broad implications for the design of modern search tools, as they can influence each of the three major search subtasks—query building, result parsing, and document analysis.

#### 3.3.1. The Problem: The Black Box, Usability Gap, and Hallucinations

While ML offers powerful techniques for handling MEDLINE's scale, its integration into search tools presents two major risks. The first is the "black-box" effect. When users do not understand how an algorithm works, they can lose trust in its results. PubMed's "Best Match" ranking, for instance, is powered by an ML model, but the logic is hidden from the user [64]. This opacity can impede both querying and parsing, as users cannot form a mental model of how the system will respond to their input. The second risk is the usability gap. Many ML components are not designed for general users and may require specialized knowledge to operate [65]. Expecting clinicians or biomedical researchers to be ML experts is unreasonable. If an ML-powered feature is confusing or difficult to control, it undermines the user's sense of agency and can hurt the search experience more than it helps [49].

The increased usability of LLMs, however, comes with additional risks. The concerns associated with their black-box nature are magnified and accompanied by a new, major risk—hallucinations. Hallucinations are outputs consisting of unfaithful or nonsensical text [66]. "Unfaithful" responses often appear indistinguishable from correct ones, and without verification, they undermine the perceived reliability of generated outputs. This uncertainty is further compounded by the scale of the training data, which may itself be erroneous or biased. For example, in the medical domain, there are significant concerns regarding reliability and privacy [67,68]. Therefore, designers of search tools that integrate LLMs must find ways to address the users' needs for transparency, reliability, and trust.

#### 3.3.2. The Solution: User-Centered Machine Learning Integration

To be effective, ML must be integrated in a way that is both transparent and user-centered. This involves shifting the focus of ML application in two key ways:

1. **From keywords to intent**. Instead of relying solely on keyword matching, ML can enable semantic search, which seeks to understand the concept behind a query [51]. If a user searches for "rough skin wound", a semantic system can infer the user means "abrasion" and retrieve relevant documents even if they do not contain the exact keywords. This approach dramatically improves query building, especially for non-expert users who may not know the precise terminology [5].
2. **From opaque automation to human-in-the-loop**. Rather than having ML models operate as a hidden black box, they can be designed to work collaboratively with the user. Techniques like active learning use human feedback to iteratively refine the model's behavior. This keeps the user in control, builds trust, and leverages both human domain expertise and machine processing power to facilitate the parsing and analysis of large result sets.

#### 3.3.3. Examples in Practice

- **LitSuggest** [69] exemplifies the shift from keywords to intent. Instead of keywords, users provide sets of "positive" (relevant) and "negative" (irrelevant) documents. An ML model learns from these examples to recommend new documents. This allows users to perform abstract queries like "find more like this", fundamentally changing the query building process from a lexical task to a conceptual one.
- **ASReview** [15] is a powerful example of a "human-in-the-loop" system for systematic reviews. The tool uses active learning to assist with the parsing of thousands of articles. The user labels one document at a time as relevant or not, and the ML model instantly uses that feedback to re-rank the remaining documents, pushing the most likely relevant ones to the top. The user remains in complete control and can

clearly see how their judgments guide the machine's behavior, avoiding the black-box problem entirely.

- ALMANAC [63] is a retrieval-augmented generation framework for clinical information retrieval. To ground LLMs in factual context, ALMANAC augments user requests with a curated corpus of medical documents. Rather than relying on an LLM's potentially unreliable training data, the framework leverages a data repository of trusted sources. Furthermore, ALMANAC requires the LLM to annotate its response with citations to the provided sources, allowing users to verify their authenticity and evaluate the quality of the response. This approach improves the factual reliability of responses compared with ungrounded LLMs. Users of tools that implement frameworks like ALMANAC retain the language processing power of LLMs while reducing their exposure to hallucinations.

When integrated thoughtfully, ML can substantially enhance the search process. However, it is crucial that designers prioritize user trust and control. By focusing on user intent and collaborative human-in-the-loop systems, ML can be a potent factor in designing modern MEDLINE search tools that are both powerful and transparent.

### *3.4. Ontology*

An ontology is an expert-curated, formal representation of knowledge for a specific domain. Typically structured as a graph or hierarchy, it defines the standardized concepts, their attributes, and the relationships between them [70]. Relationships can be hierarchical (a muscle cell is a type of cell), instantiative (this specific cell is an instance of a muscle cell), or associative (this cell is part of that tissue) [70]. By encoding domain knowledge in a computable format (e.g., OWL, OBO), ontologies provide the semantic scaffolding needed for more intelligent information systems.

The biomedical field, in particular, has produced several widely used ontologies. The Human Phenotype Ontology (HPO), for example, defines terms for phenotypic abnormalities and is used in tools for genetic disease research [71]. Gene Ontology (GO) provides a controlled vocabulary for gene attributes that enhances bioinformatics search [72,73]. Crucially, the Medical Subject Headings (MeSH) thesaurus, which underpins the indexing of MEDLINE, functions as a de facto ontology, and its hierarchical structure is often leveraged to improve search performance [74].

This structured knowledge can be leveraged to support every stage of the search process. For **query building**, ontologies serve as a bridge between a user's vocabulary and the system's formal terminology. They enable powerful features like query expansion, which automatically adds synonyms and related terms to improve recall [75], and term disambiguation, which clarifies user intent [76]. For **parsing results**, ontologies support more meaningful retrieval by allowing documents to be indexed with standardized, semantic concepts instead of just keywords [77]. This allows for a more sophisticated filtering and organization of results. Finally, for **document analysis**, ontologies facilitate information extraction such as automatically identifying and linking all mentions of specific medical concepts within a body of text [78]. By providing this semantic backbone, ontologies are a key principle for designing more context-aware and effective search tools.

#### 3.4.1. The Problem: The Vocabulary Mismatch

A fundamental challenge in search is the "vocabulary mismatch"—the frequent misalignment between the terms a user provides and the vocabulary used within the document collection [79]. Traditional search tools do little to resolve this. They typically assume a user's keywords are precise and sufficient, offering at most spellcheck or auto-complete. This forces the user to guess the system's preferred terminology. In a specialized domain

like biomedicine, this is a significant barrier. A layperson searching for "heart attack" may miss seminal articles indexed under "myocardial infarction". This problem persists even for experts; one study found that over 90% of researchers conducting literature reviews missed relevant keywords, leading to incomplete searches [80]. This gap in understanding is a primary reason why traditional keyword search often fails to meet user needs.

#### 3.4.2. The Solution: The Ontology as a Semantic Bridge

An expert-curated ontology can act as a "semantic bridge" to resolve this vocabulary mismatch. This principle can be applied in two primary ways:

1. **To enhance user queries**. The most common application is for query expansion. An ontology can automatically augment a user's query with synonyms and related concepts [75]. A search for "lungs" could be expanded to include "pulmonary" and "respiration", retrieving a more comprehensive set of documents and relieving the user of the burden of brainstorming every possible term. This directly improves the query building sub-task.
2. **To enhance system understanding**. Ontologies can also be used on the back end to make the system itself "smarter". By indexing documents based on a structured ontology rather than just keywords, the system gains a much richer, more accurate understanding of its own content. This, in turn, can improve the performance of machine learning models used for ranking, recommendation, or entity recognition [18].

#### 3.4.3. Examples in Practice

Several tools demonstrate the power of ontology-driven query enhancement. **ONSTI** [12] uses ontologies to perform query expansion and translation, allowing users to choose an ontology that best fits their subject area. **G-Bean** [81] also uses ontologies to expand initial queries in MEDLINE. Furthermore, G-Bean allows users to mark documents of interest, after which it extracts key biomedical concepts from them and uses the ontology to formulate a new query to find related documents, thus aiding the **parsing** and discovery process.

Beyond query enhancing applications, the use of ontologies to augment the back end of systems has become increasingly common. Recent work has demonstrated that integrating ontology concepts and relations into LLM prompts can provide semantic grounding, reducing hallucinations and improving performance [61,62]. Such ontology-guided frameworks are particularly useful for domain-specific search contexts such as MEDLINE.

By acting as a mediating layer, ontologies address the critical communication gap that plagues traditional search. They allow for a more robust and conceptually rich dialogue between the user and the system. Given that formulating an effective query is the vital first step in any successful search, integrating ontologies is a key consideration for the design of modern MEDLINE search tools.

### *3.5. Triaging*

In information retrieval, triaging refers to the process of rapidly assessing a set of documents to determine which are worth further attention [82]. Analogous to medical triage, it is an iterative process that occurs at multiple levels. Users perform high-level triage early on, broadly filtering and grouping documents without deep examination. They then perform low-level triage, examining the details of individual candidates more closely [82,83]. This cycle of broad filtering followed by specific examination is central to how users navigate large information spaces. A well-designed search tool must therefore provide explicit support for this triaging process to help users parse results and find relevant information effectively.

Supporting triage requires a deliberate application of the other principles in this framework. For example, many tools facilitate triage by organizing results into tiers of decreasing abstraction, allowing users to move from coarse to fine-grained assessment; this relies on principles of progressive disclosure and interaction design [12]. Richer document summaries, which help users judge relevance more quickly, can be automatically generated using machine learning [84]. LLMs, in particular, have been increasingly used to generate such summaries, though issues such as hallucinations and factual accuracy remain concerns [57,58]. Other techniques integrate multiple principles at once: visual indicators that highlight the most relevant documents are often the output of a machine learning model, expressed through effective visualization [85]. Similarly, semantic zooming, which allows users to seamlessly move between overviews and detailed views, is a powerful interaction technique built upon a foundation of visualization [86]. Ultimately, designing for triage means using a full suite of principles to reduce cognitive load and accelerate the discovery of relevant information.

#### 3.5.1. The Problem: The Burden of Manual Triage

As the volume of information in MEDLINE grows, triaging search results becomes an overwhelming task. Traditional tools like PubMed offer little assistance, forcing users to manually sift through long, linear lists of documents. This process is not only time-consuming [87] but also ineffective. Studies show that users of list-based interfaces focus their attention on only the first few results, potentially missing relevant documents buried on subsequent pages [8]. The linear format encourages a slow, one-by-one review, which is ill-suited for the rapid, nonlinear way users need to explore large result sets (e.g., by skipping around, grouping, and eliminating large chunks at once) [88]. This lack of support places the entire burden of triage on the user, leading to frustration, wasted effort, and incomplete searches.

#### 3.5.2. The Solution: Designing for the Triage Cycle

An effective search tool must support the iterative nature of triage. This can be conceptualized as designing for the triage cycle, which involves two distinct but interconnected levels of activity [82]:

1. **High-level triage**: This occurs during the initial parsing of results. The goal is to obtain a broad overview and quickly filter or group large sets of documents. This requires features that allow users to see the forest, not just the trees.
2. **Low-level triage**: This occurs closer to the analysis phase. Here, the user examines the details of a smaller set of promising candidates to make final relevance judgments. This requires features that allow for focused comparison and inspection.

The key is to support the fluid movement between these two levels, allowing users to zoom out to see the big picture and zoom in to examine specifics, repeating the cycle as needed.

#### 3.5.3. Examples in Practice

The need for better triage support is so significant that entire systems like Rayyan [87] and Abstrackr [89] exist solely to help users filter documents for reviews. Other integrated tools also showcase excellent triage support:

- **VisualQUEST** [11] was designed specifically to facilitate the triage cycle. Its interface is split into two linked views: a high-level view for grouping documents by topic or similarity, and a low-level view for showing detailed snippets of selected groups (Figure 5). This design directly maps to the two levels of triage, allowing users to efficiently switch between broad filtering and detailed examination.

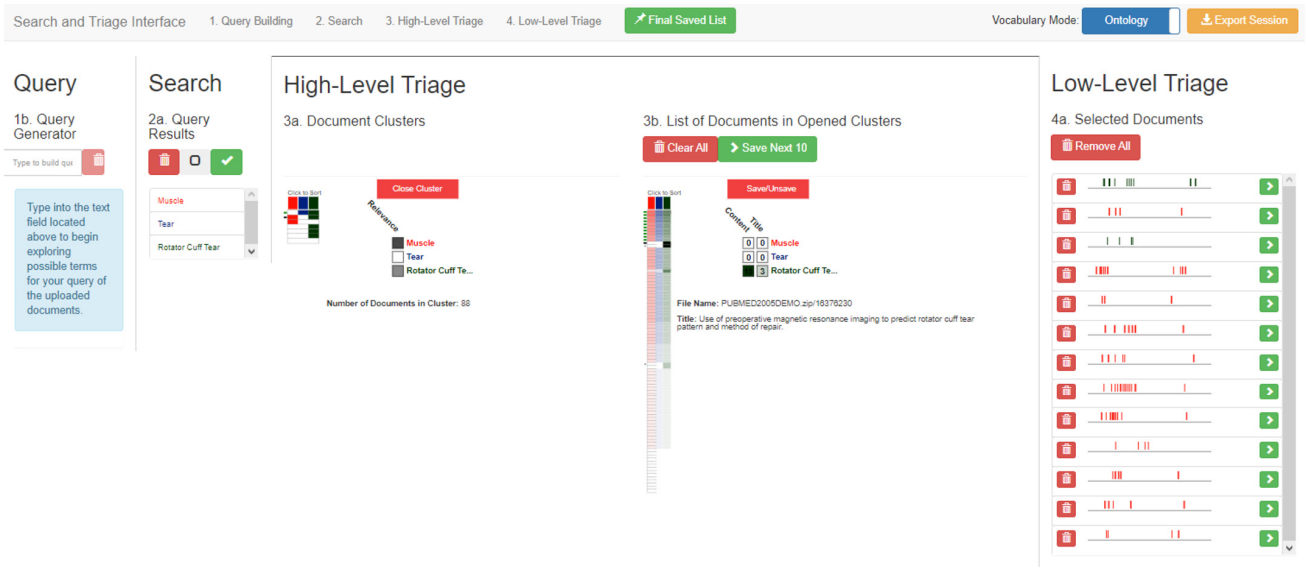



**Figure 5.** VisualQUEST: a tool that supports the users' triaging tasks with dedicated subviews.

- **DocFlow** [16], a system for systematic reviews, supports triage through a customizable "pipeline". Users can build a multi-step process to filter documents (e.g., first by keyword, then by topic), which supports high-level triage. The system also provides detailed visualizations like scatter plots to explore smaller subsets, supporting low-level triage. By giving users interactive control over the entire filtering process, DocFlow empowers a more systematic and transparent triage workflow.

By actively supporting the triage cycle, search tools can help users make more accurate relevance judgments, efficiently organize which documents to read, and manage the information overload inherent in searching MEDLINE.

### *3.6. Progressive Disclosure*

Progressive disclosure is a design technique that manages complexity by incrementally revealing information to users as needed. The goal is to reduce cognitive load by initially showing only essential information and deferring details until a user explicitly requests them. Effective progressive disclosure meets three requirements: information is revealed on demand, it is organized hierarchically from general to specific, and the system maintains context to avoid redundancy [90]. For search tools that handle vast amounts of data and offer numerous features, this principle is essential for preventing information overload.

Progressive disclosure is not a single feature, but an interaction strategy that can be implemented in many ways, often leveraging dynamic visualization. For example, semantic zooming is a technique where the visual representation changes based on the user's focus; a high-level overview of documents can smoothly transform into a detailed view of a single document as the user zooms or selects it [91]. A related technique is linked highlighting, where multiple, synchronized views are maintained. Interacting with an item in one view (e.g., selecting a document) causes corresponding highlights or updates in another view (e.g., a details panel). Both techniques use interaction to allow users to control the level of detail they see, thus managing interface complexity. As discussed earlier, the summaries generated by LLMs can complement such techniques by providing content for high-level overviews [58]. Likewise, LLM-based feature extraction can populate secondary views with details [59].

Numerous studies have confirmed the value of this approach, examining how users interact with interfaces that progressively reveal information and how these techniques impact specific tasks [90,92,93]. By applying progressive disclosure, a search tool can

enhance both parsing and analysis. Instead of displaying a flat list of 10,000 results, a tool can first present high-level thematic groups, allowing the user to drill down into areas of interest. Similarly, during analysis, detailed metadata about a document can remain hidden until requested, keeping the interface clean and focused. In essence, progressive disclosure empowers users by allowing them to navigate complex information landscapes in a staged, context-aware, and user-driven manner.

#### 3.6.1. The Problem: Interface and Information Overload

Supporting the multi-stage search process often requires tools with many features, leading to complex interfaces. Simultaneously, MEDLINE documents contain a wealth of useful metadata (abstracts, MeSH terms, author information). Traditional search tools struggle with this dual complexity. They tend to either hide advanced features behind a single "advanced search" button or, more often, display all available information for every result at once. When a search yields hundreds of results, this approach quickly overwhelms the user with a wall of text and options, making it difficult to focus on the task at hand [94]. The core challenge is how to make rich data and powerful features available without creating a cluttered and confusing interface.

#### 3.6.2. The Solution: Aligning the Interface with the Search Stage

Progressive disclosure solves this problem by incrementally revealing information and functionality in a way that aligns with the user's workflow. The principle is to introduce the right tools and information at the right time. For example, during the initial parsing stage, the interface can show a high-level overview of the results (e.g., thematic clusters or a summary visualization) while hiding granular details. Once a user selects a document or group to focus on, the interface can then transition to the analysis stage, revealing detailed information like the abstract or related items for that specific selection. By dynamically adjusting to the user's focus, progressive disclosure reduces clutter and helps guide the user through the information space step-by-step.

#### 3.6.3. Examples in Practice

- **VisualQUEST** [11], a tool for literature search, is designed around this principle. Its interface components are tied to different stages of the search process (querying, high-level triage, low-level detail). As the user drills down from a broad overview to specific documents, the corresponding interface panels expand to show more detail while others recede into the background. This ensures that the user's attention is always focused on the relevant tools for their current task.
- **JARVIS** [95] demonstrates how progressive disclosure can be used to manage recommendations. Instead of showing a long list of suggested queries upfront, JARVIS uses an ontology to gradually reveal related recommendations as the user searches and shows interest in certain topics. This avoids overwhelming the user while also providing contextual guidance at the moment it is most useful.

The need for progressive disclosure is a direct consequence of the scale of MEDLINE and the complexity of modern search tools. The vast amount of information cannot be presented to users all at once; it must be introduced incrementally and meaningfully. By considering what a user needs to see at each stage of their search—and deferring what they do not—designers can use progressive disclosure to create interfaces that are both powerful and clean, ensuring that advanced features enhance, rather than hinder, the user's journey.

### *3.7. Evolutionary Design*

The design of a search tool must satisfy needs arising from many sources: the users (with their varying expertise and goals), the data (its size and structure), and the tasks themselves (e.g., exploratory search vs. simple lookup) [96]. Given this complexity, rigid, linear design models are often insufficient. Instead, iterative processes that allow for continuous learning and refinement become crucial [97]. One such approach is evolutionary design. Here, evolutionary design refers to an iterative, user-centered design process, distinct from computational approaches such as genetic algorithms or evolutionary computation.

Evolutionary design is an iterative model that involves repeatedly refining a design through cycles of development and user feedback. A formalization of this process describes a four-stage cycle: **formulation**, **realization**, **validation**, and **refinement** [97]. During formulation, designers establish requirements informed by prior research and user studies. In the realization stage, a working prototype is created. This prototype is then put through validation via user testing and usability evaluation. Finally, in the refinement stage, the findings from validation are used to improve the design, starting the cycle anew. This process continues until the design robustly satisfies all requirements.

#### 3.7.1. The Problem: The Limits of Upfront Design

When designing tools for complex domains like MEDLINE, it is rarely possible to determine an optimal solution from the outset. The requirements are often fuzzy and multifaceted; supporting an activity like triage, for instance, involves qualitative user needs that are difficult to fully quantify in advance [88]. Traditional, linear design processes, which rely on establishing fixed requirements upfront, are ill-suited for this ambiguity. They cannot easily accommodate the discovery of new user needs or the creative exploration required to build a truly innovative and effective tool.

#### 3.7.2. The Solution: Embracing an Iterative, User-Centered Process

By treating the design process not as a linear path but as a cycle of continuous improvement, evolutionary design addresses these challenges. This approach acknowledges that understanding of the problem deepens as one builds and tests prototypes. This iterative loop allows designers to remain flexible, blending broad exploration of ideas with deep dives into promising concepts, ensuring that the final product is shaped by user feedback rather than initial assumptions.

#### 3.7.3. Examples in Practice

The development of **Open MS BioScreen** [98], a tool for visualizing the data of patient with multiple sclerosis, provides an excellent case study. The designers used an evolutionary process, involving both patients and clinicians at every stage. They began with interviews to gather requirements, built initial prototypes based on that feedback, and then repeatedly tested and refined those prototypes with the end-users. Through these iterations, the designers discovered subtle requirements and usability issues that were not obvious at the start, leading to a final product that successfully met the distinct needs of both user groups.

This process is just as relevant for established tools. **PubMed** itself has undergone years of iterative design changes informed by usability studies and user feedback [99]. This demonstrates that even the most widely used tools must continuously evolve to better serve their users. For a new MEDLINE search tool, this would mean prototyping specific features from this framework—a novel visualization or a new interaction—and testing them with researchers to see what truly helps them in their work.

While the choice of a design process does not alter a tool's functions directly, it profoundly influences the final quality and usability of the tool. Evolutionary design, by building in cycles of user feedback, ensures that the complex interplay of the other six principles is tuned to the actual user needs. This iterative loop is what allows designers to discover the right visualizations, refine interactions, and apply machine learning methods, such as LLMs, in ways that are genuinely helpful. Therefore, evolutionary design is not an optional add-on but the foundational principle that enables the successful implementation of all others, ensuring that complex search tools are ultimately effective, usable, and aligned with their users.

### *3.8. Interconnections and Trade-Offs*

While presented separately for clarity, the seven principles are deeply interconnected. Each principle both augments and restricts the others. This section considers them collectively, highlighting some of their interconnections and the trade-offs that emerge from their combined application. Across the examples, each principle leads to others and introduces costs that must be addressed through careful, iterative design.

Visualization and progressive disclosure shape how users perceive and manage information. Visualizations increase informational richness at the cost of consuming limited screen space and contributing to cognitive load. Progressive disclosure counterbalances this by limiting the presented information to only what is needed. However, this requires interaction to transition between states smoothly and maintain context. If the information being presented requires processing, then machine learning may be employed, raising the computational costs.

These computational demands become especially relevant when dealing with large document sets, where triaging becomes necessary. This requires interactions to perform the triage, visualizations to represent the documents, and progressive disclosure to incrementally show information. However, this integration introduces interface complexity, requiring users to learn new interactions and interpret potentially complex visualizations. Designers must balance this complexity against the tools' usability.

In specialized domains such as medicine, ontologies provide formal representations that can guide retrieval and interpretation. Integrating ontologies into machine learning systems grounds them semantically, reducing hallucinations. However, the ontology's structure and coverage constrain the flexibility of the tool. Aggressive usage of a restrictive ontology can hinder performance on unexpected documents.

Finally, evolutionary design acts as a method to account for the emergent complexity of the relationships between the principles, users, and data. Iterative prototyping and user feedback are essential to calibrate competing priorities that arise during design.

## 4. Conclusions

The vast and vital resource of MEDLINE is constrained by the tools we use to access it. As this paper has argued, the prevailing paradigm of traditional search—characterized by linear result lists, minimal interaction, and a significant vocabulary gap between user and system—is increasingly inadequate for the complex exploratory tasks that define modern biomedical research. By adhering to these stagnant designs, we perpetuate a critical bottleneck that impedes discovery.

In response, this paper has proposed a framework of seven interconnected principles to guide the design of the next generation of MEDLINE search tools. This is not a menu of optional features, but a call for an integrated design philosophy where these principles work in concert. Visualization and interaction combine to create a rich, manipulable canvas for exploration. This canvas is infused with intelligence by machine learning and given

semantic structure by ontologies, enabling a deeper dialogue between the user and the data. The user's journey through this complex information space is carefully managed by the principles of triaging and progressive disclosure, which reduce the cognitive load and align the interface with the user's immediate task. Finally, the entire system is shaped and honed through an evolutionary design process, ensuring that the final product is not merely powerful, but is robustly aligned with the real-world needs and workflows of its users.

While the framework describes a conceptual foundation for next-generation search tools, it does not prescribe specific evaluation criteria. Any such metric would depend on specific contexts, tasks, users, and document sets and require empirical studies to evaluate. Consequently, no universal set of evaluation metrics has been proposed.

When considered together, these principles provide a roadmap for moving beyond simple information retrieval toward the creation of true instruments for discovery. By systematically incorporating them into the design process, developers can create search tools for MEDLINE and similar complex domains that are more powerful, intuitive, and genuinely supportive of the nuanced work of researchers, clinicians, and patients. The future of biomedical discovery depends not only on the data we collect, but on our ability to see, navigate, and understand it. Building better tools is a critical step toward that future.

**Author Contributions:** Conceptualization, V.Z. and K.S.; Methodology, V.Z. and K.S.; Writing—original draft preparation, V.Z., M.M., and K.S.; Writing—review and editing, V.Z. and K.S.; Supervision K.S. and M.M.; Funding acquisition, K.S. and M.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), grant number RGPIN-2023-04735; the NSERC Discovery Launch Supplement, grant number DGECR-2021-00447; and the NSERC Discovery Grants, grant number RGPIN-2021-04120.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| LLM | Large language model |
| ML | Machine learning |
| MeSH | Medical Subject Headings |

## References

1. MEDLINE. Available online: https://www.nlm.nih.gov/medline/medline_home.html (accessed on 6 June 2025).
2. PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/ (accessed on 6 June 2025).
3. Wall, E.; Blaha, L.M.; Franklin, L.; Endert, A. Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics. In Proceedings of the 2017 IEEE Conference on Visual Analytics Science and Technology (VAST), Phoenix, AZ, USA, 3–6 October 2017; IEEE: New York, NY, USA, 2017; pp. 104–115.
4. Herceg, P.M.; Allison, T.B.; Belvin, R.S.; Tzoukermann, E. Collaborative Exploratory Search for Information Filtering and Large-Scale Information Triage. *J. Assoc. Inf. Sci. Technol.* **2018**, *69*, 395–409. [CrossRef]
5. Jin, Q.; Leaman, R.; Lu, Z. PubMed and Beyond: Biomedical Literature Search in the Age of Artificial Intelligence. *eBioMedicine* **2024**, *100*, 104988. [CrossRef] [PubMed]
6. Demelo, J.; Parsons, P.; Sedig, K. Ontology-Driven Search and Triage: Design of a Web-Based Visual Interface for MEDLINE. *JMIR Med. Inform.* **2017**, *5*, e6918. [CrossRef] [PubMed]
7. Salvador-Oliván, J.A.; Marco-Cuenca, G.; Arquero-Avilés, R. Development of an Efficient Search Filter to Retrieve Systematic Reviews from PubMed. *J. Med. Libr. Assoc.* **2021**, *109*, 561. [CrossRef] [PubMed]
8. Islamaj Dogan, R.; Murray, G.C.; Névéol, A.; Lu, Z. Understanding PubMed® User Search Behavior Through Log Analysis. *Database* **2009**, *2009*, bap018. [CrossRef]

9. Morshed, T.; Hayden, S. Google Versus PubMed: Comparison of Google and PubMed's Search Tools for Answering Clinical Questions in the Emergency Department. *Ann. Emerg. Med.* **2020**, *75*, 408–415. [CrossRef]
10. Gusenbauer, M.; Haddaway, N.R. Which Academic Search Systems Are Suitable for Systematic Reviews or Meta-Analyses? Evaluating Retrieval Qualities of Google Scholar, PubMed, and 26 Other Resources. *Res. Synth. Methods* **2020**, *11*, 181–217. [CrossRef]
11. Demelo, J.; Sedig, K. Interfaces for Searching and Triaging Large Document Sets: An Ontology-Supported Visual Analytics Approach. *Information* **2021**, *13*, 8. [CrossRef]
12. Demelo, J.; Sedig, K. Design of Generalized Search Interfaces for Health Informatics. *Information* **2021**, *12*, 317. [CrossRef]
13. Cui, W. Visual Analytics: A Comprehensive Overview. *IEEE Access* **2019**, *7*, 81555–81573. [CrossRef]
14. Shen, Q.; Wu, T.; Yang, H.; Wu, Y.; Qu, H.; Cui, W. Nameclarifier: A Visual Analytics System for Author Name Disambiguation. *IEEE Trans. Vis. Comput. Graph.* **2016**, *23*, 141–150. [CrossRef] [PubMed]
15. Van De Schoot, R.; De Bruin, J.; Schram, R.; Zahedi, P.; De Boer, J.; Weijdema, F.; Kramer, B.; Huijts, M.; Hoogerwerf, M.; Ferdinands, G.; et al. An Open Source Machine Learning Framework for Efficient and Transparent Systematic Reviews. *Nat. Mach. Intell.* **2021**, *3*, 125–133. [CrossRef]
16. Qiu, R.; Tu, Y.; Wang, Y.-S.; Yen, P.-Y.; Shen, H.-W. DocFlow: A Visual Analytics System for Question-Based Document Retrieval and Categorization. *IEEE Trans. Vis. Comput. Graph.* **2022**, *30*, 1533–1548. [CrossRef] [PubMed]
17. Lee, J.-C.; Lee, B.J.; Park, C.; Song, H.; Ock, C.-Y.; Sung, H.; Woo, S.; Youn, Y.; Jung, K.; Jung, J.H.; et al. Efficacy Improvement in Searching MEDLINE Database Using a Novel PubMed Visual Analytic System: EEEvis. *PLoS ONE* **2023**, *18*, e0281422. [CrossRef] [PubMed]
18. Xu, J.; Kim, S.; Song, M.; Jeong, M.; Kim, D.; Kang, J.; Rousseau, J.F.; Li, X.; Xu, W.; Torvik, V.I.; et al. Building a PubMed Knowledge Graph. *Sci. Data* **2020**, *7*, 205. [CrossRef]
19. Zhu, Y.; Yuan, H.; Wang, S.; Liu, J.; Liu, W.; Deng, C.; Chen, H.; Liu, Z.; Dou, Z.; Wen, J.-R. Large Language Models for Information Retrieval: A Survey. *arXiv* **2023**, arXiv:2308.07107. [CrossRef]
20. Russell-Rose, T.; Shokraneh, F. Designing the Structured Search Experience: Rethinking the Query-Builder Paradigm. *Weav. J. Libr. User Exp.* **2020**, *3*. [CrossRef]
21. Nitsche, M.; Nürnberger, A. QUEST: Querying Complex Information by Direct Manipulation. In *Human Interface and the Management of Information: Information and Interaction Design, 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21–26, 2013, Proceedings, Part I*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 240–249.
22. Nowell, L.T.; France, R.K.; Hix, D.; Heath, L.S.; Fox, E.A. Visualizing Search Results: Some Alternatives to Query-Document Similarity. In Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Zurich, Switzerland, 18–22 August 1996; pp. 67–75.
23. Peltonen, J.; Belorustceva, K.; Ruotsalo, T. Topic-Relevance Map: Visualization for Improving Search Result Comprehension. In Proceedings of the 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 13–16 March 2017; pp. 611–622.
24. Nguyen, T.; Zhang, J. A Novel Visualization Model for Web Search Results. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12*, 981–988. [CrossRef]
25. Heimerl, F.; Lohmann, S.; Lange, S.; Ertl, T. Word Cloud Explorer: Text Analytics Based on Word Clouds. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014; IEEE: New York, NY, USA, 2014; pp. 1833–1842.
26. Mendoza, M.; Bonilla, S.; Noguera, C.; Cobos, C.; León, E. Extractive Single-Document Summarization Based on Genetic Operators and Guided Local Search. *Expert Syst. Appl.* **2014**, *41*, 4158–4169. [CrossRef]
27. Hearst, M.A. Tilebars: Visualization of Term Distribution Information in Full Text Information Access. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 7–11 May 1995; pp. 59–66.
28. Uren, V.; Lei, Y.; Lopez, V.; Liu, H.; Motta, E.; Giordanino, M. The Usability of Semantic Search Tools: A Review. *Knowl. Eng. Rev.* **2007**, *22*, 361–377. [CrossRef]
29. McCallum, A.; Nigam, K.; Rennie, J.; Seymore, K. A Machine Learning Approach to Building Domain-Specific Search Engines. In Proceedings of the 16th International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 31 July–6 August 1999; Volume 99, pp. 662–667.
30. Larkin, J.H.; Simon, H.A. Why a Diagram Is (Sometimes) Worth Ten Thousand Words. *Cogn. Sci.* **1987**, *11*, 65–100. [CrossRef]
31. Scaife, M.; Rogers, Y. External Cognition: How Do Graphical Representations Work? *Int. J. Hum.-Comput. Stud.* **1996**, *45*, 185–213. [CrossRef]
32. Sedig, K.; Parsons, P. Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach. *AIS Trans. Hum.-Comput. Interact.* **2013**, *5*, 84–133. [CrossRef]
33. Zhuang, S.; Zhuang, H.; Koopman, B.; Zuccon, G. A Setwise Approach for Effective and Highly Efficient Zero-Shot Ranking with Large Language Models. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, Washington, DC, USA, 14–18 July 2024; pp. 38–47.

34. Wu, Y.; Wan, Y.; Zhang, H.; Sui, Y.; Wei, W.; Zhao, W.; Xu, G.; Jin, H. Automated Data Visualization from Natural Language via Large Language Models: An Exploratory Study. *Proc. ACM Manag. Data* **2024**, *2*, 1–28. [CrossRef]
35. Li, R.; Yin, C.; Yang, S.; Qian, B.; Zhang, P. Marrying Medical Domain Knowledge with Deep Learning on Electronic Health Records: A Deep Visual Analytics Approach. *J. Med. Internet Res.* **2020**, *22*, e20645. [CrossRef]
36. Scells, H.; Zuccon, G. Searchrefiner: A Query Visualisation and Understanding Tool for Systematic Reviews. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, 22–26 October 2018; pp. 1939–1942.
37. Clarkson, E.; Desai, K.; Foley, J. Resultmaps: Visualization for Search Interfaces. *IEEE Trans. Vis. Comput. Graph.* **2009**, *15*, 1057–1064. [CrossRef]
38. Görg, C.; Liu, Z.; Stasko, J. Reflections on the Evolution of the Jigsaw Visual Analytics System. *Inf. Vis.* **2014**, *13*, 336–345. [CrossRef]
39. Liu, Y.-H.; Thomas, P.; Gedeon, T.; Rusnachenko, N. Search Interfaces for Biomedical Searching: How Do Gaze, User Perception, Search Behaviour and Search Performance Relate? In Proceedings of the 2022 Conference on Human Information Interaction and Retrieval, Regensburg, Germany, 14–18 March 2022; pp. 78–89.
40. Aula, A.; Khan, R.M.; Guan, Z. How Does Search Behavior Change as Search Becomes More Difficult? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 35–44.
41. Stolper, C.D.; Perer, A.; Gotz, D. Progressive Visual Analytics: User-Driven Visual Exploration of in-Progress Analytics. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 1653–1662. [CrossRef]
42. Shao, L.; Silva, N.; Eggeling, E.; Schreck, T. Visual Exploration of Large Scatter Plot Matrices by Pattern Recommendation Based on Eye Tracking. In Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics, Limassol, Cyprus, 13 March 2017; pp. 9–16.
43. Ola, O.; Sedig, K. The Challenge of Big Data in Public Health: An Opportunity for Visual Analytics. *Online J. Public Health Inform.* **2014**, *5*, 223. [PubMed]
44. Fast, K.V.; Sedig, K. Interaction and the Epistemic Potential of Digital Libraries. *Int. J. Digit. Libr.* **2010**, *11*, 169–207. [CrossRef]
45. Tenner, E. The Design of Everyday Things by Donald Norman. *Technol. Cult.* **2015**, *56*, 785–787. [CrossRef]
46. Sedig, K.; Parsons, P.; Dittmer, M.; Ola, O. Beyond Information Access: Support for Complex Cognitive Activities in Public Health Informatics Tools. *Online J. Public Health Inform.* **2012**, *4*, 1–23. [CrossRef] [PubMed]
47. Spence, R. Sensitivity Encoding to Support Information Space Navigation: A Design Guideline. *Inf. Vis.* **2002**, *1*, 120–129. [CrossRef]
48. Mao, K.; Deng, C.; Chen, H.; Mo, F.; Liu, Z.; Sakai, T.; Dou, Z. ChatRetriever: Adapting Large Language Models for Generalized and Robust Conversational Dense Retrieval. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, Miami, FL, USA, 12–16 November 2024; pp. 1227–1240.
49. Endert, A.; Ribarsky, W.; Turkay, C.; Wong, B.W.; Nabney, I.; Blanco, I.D.; Rossi, F. The State of the Art in Integrating Machine Learning into Visual Analytics. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2017; Volume 36, pp. 458–486.
50. Portugal, I.; Alencar, P.; Cowan, D. The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review. *Expert Syst. Appl.* **2018**, *97*, 205–227. [CrossRef]
51. Fang, F.; Zhang, B.-W.; Yin, X.-C. Semantic Sequential Query Expansion for Biomedical Article Search. *IEEE Access* **2018**, *6*, 45448–45457. [CrossRef]
52. Aphinyanaphongs, Y.; Aliferis, C.F. Prospective Validation of Text Categorization Filters for Identifying High-Quality, Content-Specific Articles in MEDLINE. In *AMIA Annual Symposium Proceedings*; American Medical Informatics Association: Washington, DC, USA, 2006; Volume 2006, p. 6.
53. Fiorini, N.; Canese, K.; Starchenko, G.; Kireev, E.; Kim, W.; Miller, V.; Osipov, M.; Kholodov, M.; Ismagilov, R.; Mohan, S.; et al. Best Match: New Relevance Search for PubMed. *PLoS Biol.* **2018**, *16*, e2005343. [CrossRef]
54. Ma, C.; Zhang, W.E.; Guo, M.; Wang, H.; Sheng, Q.Z. Multi-Document Summarization via Deep Learning Techniques: A Survey. *ACM Comput. Surv.* **2022**, *55*, 1–37. [CrossRef]
55. Khalid, S.; Khalil, T.; Nasreen, S. A Survey of Feature Selection and Feature Extraction Techniques in Machine Learning. In Proceedings of the 2014 Science and Information Conference, London, UK, 27–29 August 2014; IEEE: New York, NY, USA, 2014; pp. 372–378.
56. Naveed, H.; Khan, A.U.; Qiu, S.; Saqib, M.; Anwar, S.; Usman, M.; Akhtar, N.; Barnes, N.; Mian, A. A Comprehensive Overview of Large Language Models. *ACM Trans. Intell. Syst. Technol.* **2025**, *16*, 1–72. [CrossRef]
57. Tang, L.; Sun, Z.; Idnay, B.; Nestor, J.G.; Soroush, A.; Elias, P.A.; Xu, Z.; Ding, Y.; Durrett, G.; Rousseau, J.F.; et al. Evaluating Large Language Models on Medical Evidence Summarization. *NPJ Digit. Med.* **2023**, *6*, 158. [CrossRef] [PubMed]

58. Van Veen, D.; Van Uden, C.; Blankemeier, L.; Delbrouck, J.-B.; Aali, A.; Bluethgen, C.; Pareek, A.; Polacin, M.; Reis, E.P.; Seehofnerová, A.; et al. Adapted Large Language Models Can Outperform Medical Experts in Clinical Text Summarization. *Nat. Med.* **2024**, *30*, 1134–1142. [CrossRef]
59. Ntinopoulos, V.; Biefer, H.R.C.; Tudorache, I.; Papadopoulos, N.; Odavic, D.; Risteski, P.; Haeussler, A.; Dzemali, O. Large Language Models for Data Extraction from Unstructured and Semi-Structured Electronic Health Records: A Multiple Model Performance Evaluation. *BMJ Health Care Inform.* **2025**, *32*, e101139. [CrossRef]
60. Jagerman, R.; Zhuang, H.; Qin, Z.; Wang, X.; Bendersky, M. Query Expansion by Prompting Large Language Models. *arXiv* **2023**, arXiv:2305.03653. [CrossRef]
61. Agrawal, G.; Kumarage, T.; Alghamdi, Z.; Liu, H. Can Knowledge Graphs Reduce Hallucinations in LLMs?: A Survey. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*; Association for Computational Linguistics: Mexico City, Mexico, 2024; pp. 3947–3960.
62. Tan, X.; Wang, X.; Liu, Q.; Xu, X.; Yuan, X.; Zhang, W. Paths-over-Graph: Knowledge Graph Empowered Large Language Model Reasoning. In Proceedings of the ACM on Web Conference 2025, Sydney, Australia, 28 April–2 May 2025; pp. 3505–3522.
63. Zakka, C.; Shad, R.; Chaurasia, A.; Dalal, A.R.; Kim, J.L.; Moor, M.; Fong, R.; Phillips, C.; Alexander, K.; Ashley, E.; et al. Almanac—Retrieval-Augmented Language Models for Clinical Medicine. *Nejm Ai* **2024**, *1*, AIoa2300068. [CrossRef]
64. Kiester, L.; Turp, C. Artificial Intelligence Behind the Scenes: PubMed's Best Match Algorithm. *J. Med. Libr. Assoc.* **2022**, *110*, 15. [CrossRef]
65. Cierco Jimenez, R.; Lee, T.; Rosillo, N.; Cordova, R.; Cree, I.A.; Gonzalez, A.; Indave Ruiz, B.I. Machine Learning Computational Tools to Assist the Performance of Systematic Reviews: A Mapping Review. *BMC Med. Res. Methodol.* **2022**, *22*, 322. [CrossRef]
66. Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y.J.; Madotto, A.; Fung, P. Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.* **2023**, *55*, 1–38. [CrossRef]
67. Hu, J.-M.; Liu, F.-C.; Chu, C.-M.; Chang, Y.-T. Health Care Trainees' and Professionals' Perceptions of ChatGPT in Improving Medical Knowledge Training: Rapid Survey Study. *J. Med. Internet Res.* **2023**, *25*, e49385. [CrossRef] [PubMed]
68. Spotnitz, M.; Idnay, B.; Gordon, E.R.; Shyu, R.; Zhang, G.; Liu, C.; Cimino, J.J.; Weng, C. A Survey of Clinicians' Views of the Utility of Large Language Models. *Appl. Clin. Inform.* **2024**, *15*, 306–312. [CrossRef] [PubMed]
69. Allot, A.; Lee, K.; Chen, Q.; Luo, L.; Lu, Z. LitSuggest: A Web-Based System for Literature Recommendation and Curation Using Machine Learning. *Nucleic Acids Res.* **2021**, *49*, W352–W358. [CrossRef]
70. Arp, R.; Smith, B.; Spear, A.D. *Building Ontologies with Basic Formal Ontology*; The MIT Press: Cambridge, MA, USA, 2015.
71. Gargano, M.A.; Matentzoglu, N.; Coleman, B.; Addo-Lartey, E.B.; Anagnostopoulos, A.V.; Anderton, J.; Avillach, P.; Bagley, A.M.; Bakštein, E.; Balhoff, J.P.; et al. The Human Phenotype Ontology in 2024: Phenotypes Around the World. *Nucleic Acids Res.* **2024**, *52*, D1333–D1346. [CrossRef] [PubMed]
72. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]
73. Doms, A.; Schroeder, M. GoPubMed: Exploring PubMed with the Gene Ontology. *Nucleic Acids Res.* **2005**, *33*, W783–W786. [CrossRef]
74. Trieschnigg, D.; Pezik, P.; Lee, V.; De Jong, F.; Kraaij, W.; Rebholz-Schuhmann, D. MeSH up: Effective MeSH Text Classification for Improved Document Retrieval. *Bioinformatics* **2009**, *25*, 1412–1418. [CrossRef] [PubMed]
75. Bhogal, J.; MacFarlane, A.; Smith, P. A Review of Ontology Based Query Expansion. *Inf. Process. Manag.* **2007**, *43*, 866–886. [CrossRef]
76. Gracia, J.; Trillo, R.; Espinoza, M.; Mena, E. Querying the Web: A Multiontology Disambiguation Method. In Proceedings of the 6th International Conference on Web Engineering, Palo Alto, CA, USA, 11–14 July 2006; pp. 241–248.
77. Asim, M.N.; Wasim, M.; Khan, M.U.G.; Mahmood, N.; Mahmood, W. The Use of Ontology in Retrieval: A Study on Textual, Multilingual, and Multimedia Retrieval. *IEEE Access* **2019**, *7*, 21662–21686. [CrossRef]
78. de Silva, N.; Dou, D.; Huang, J. Discovering Inconsistencies in Pubmed Abstracts Through Ontology-Based Information Extraction. In Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Boston, MA, USA, 20–23 August 2017; pp. 362–371.
79. Furnas, G.W.; Landauer, T.K.; Gomez, L.M.; Dumais, S.T. The Vocabulary Problem in Human-System Communication. *Commun. ACM* **1987**, *30*, 964–971. [CrossRef]
80. Salvador-Oliván, J.A.; Marco-Cuenca, G.; Arquero-Avilés, R. Errors in Search Strategies Used in Systematic Reviews and Their Effects on Information Retrieval. *J. Med. Libr. Assoc.* **2019**, *107*, 210. [CrossRef]
81. Wang, J.Z.; Zhang, Y.; Dong, L.; Li, L.; Srimani, P.K.; Yu, P.S. G-Bean: An Ontology-Graph Based Web Tool for Biomedical Literature Retrieval. *BMC Bioinform.* **2014**, *15*, S1. [CrossRef]
82. Loizides, F.; Buchanan, G. An Empirical Study of User Navigation During Document Triage. In *Research and Advanced Technology for Digital Libraries: 13th European Conference. ECDL 2009, Corfu, Greece, September 27–October 2, 2009, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 138–149.

83. Loizides, F.; Buchanan, G. Towards a Framework for Human (Manual) Information Retrieval. In *Multidisciplinary Information Retrieval: 6th Information Retrieval Facility Conference, IRFC 2013, Limassol, Cyprus, October 7–9, 2013, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 87–98.
84. Jonker, D.; Wright, W.; Schroh, D.; Proulx, P.; Cort, B. Information Triage with TRIST. In Proceedings of the 2005 Intelligence Analysis Conference, Washington, DC, USA, 2–6 May 2005; pp. 2–4.
85. Macskassy, S.A.; Provost, F. Intelligent Information Triage. In Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New Orleans, LA, USA, 9–13 September 2001; pp. 318–326.
86. Buchanan, G.; Owen, T. Improving Skim Reading for Document Triage. In Proceedings of the Second International Symposium on Information Interaction in Context, London, UK, 14–17 October 2008; pp. 83–88.
87. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A Web and Mobile App for Systematic Reviews. *Syst. Rev.* **2016**, *5*, 210. [CrossRef]
88. Badi, R.; Bae, S.; Moore, J.M.; Meintanis, K.; Zacchi, A.; Hsieh, H.; Shipman, F.; Marshall, C.C. Recognizing User Interest and Document Value from Reading and Organizing Activities in Document Triage. In Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney, Australia, 29 January–1 February 2006; pp. 218–225.
89. Rathbone, J.; Hoffmann, T.; Glasziou, P. Faster Title and Abstract Screening? Evaluating Abstrackr, a Semi-Automated Online Screening Program for Systematic Reviewers. *Syst. Rev.* **2015**, *4*, 80. [CrossRef]
90. Springer, A.; Whittaker, S. Progressive Disclosure: When, Why, and How Do Users Want Algorithmic Transparency Information? *ACM Trans. Interact. Intell. Syst. (TiiS)* **2020**, *10*, 1–32. [CrossRef]
91. Chuang, J.; Ramage, D.; Manning, C.; Heer, J. Interpretation and Trust: Designing Model-Driven Visualizations for Text Analysis. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 443–452.
92. Phan, D.; Paepcke, A.; Winograd, T. Progressive Multiples for Communication-Minded Visualization. In Proceedings of the Proceedings of Graphics Interface 2007, Montreal, QC, Canada, 28–30 May 2007; pp. 225–232.
93. Springer, A.; Whittaker, S. Progressive Disclosure: Empirically Motivated Approaches to Designing Effective Transparency. In Proceedings of the 24th International Conference on Intelligent User Interfaces, Marina del Ray, CA, USA, 17–20 March 2019; pp. 107–120.
94. Oulasvirta, A.; Hukkinen, J.P.; Schwartz, B. When More Is Less: The Paradox of Choice in Search Engine Use. In Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Boston, MA, USA, 19–23 July 2009; pp. 516–523.
95. Ribeiro, D.S.; de Sousa, A.G.; de Almeida, R.B.; Thompson Furtado, P.H.; Côrtes Vieira Lopes, H.; Barbosa, S.D.J. Exploring Ontology-Based Information Through the Progressive Disclosure of Visual Answers to Related Queries. In *Human Interface and the Management of Information. Designing Information: Thematic Area, HIMI 2020, Held as Part of the 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 104–124.
96. Stouffs, R.; Rafiq, Y. Generative and Evolutionary Design Exploration. *AI EDAM* **2015**, *29*, 329–331. [CrossRef]
97. Guerrero-García, J. Evolutionary Design of User Interfaces for Workflow Information Systems. *Sci. Comput. Program.* **2014**, *86*, 89–102. [CrossRef]
98. Schleimer, E.; Pearce, J.; Barnecut, A.; Rowles, W.; Lizee, A.; Klein, A.; Block, V.J.; Santaniello, A.; Renschen, A.; Gomez, R.; et al. A Precision Medicine Tool for Patients with Multiple Sclerosis (the Open MS BioScreen): Human-Centered Design and Development. *J. Med. Internet Res.* **2020**, *22*, e15605. [CrossRef] [PubMed]
99. Fiorini, N.; Canese, K.; Bryzgunov, R.; Radetska, I.; Gindulyte, A.; Latterner, M.; Miller, V.; Osipov, M.; Kholodov, M.; Starchenko, G.; et al. PubMed Labs: An Experimental System for Improving Biomedical Literature Search. *Database* **2018**, *2018*, bay094. [CrossRef] [PubMed]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.