International Science and Technology Journal
المجلة الدولية للعلوم والتقنية
Volume 37 العدد
Part 2 المجلد
المجلة الدولية للعلوم والتقنية
International Science and Technology Journal
ISTJ

http://www.doi.org/10.62341/elau7681

Received	**2025/11/19**	**تم استلام الورقة العلمية في**
Accepted	**2025/12/08**	**تم قبول الورقة العلمية في**
Published	**2025/12/10**	**تم نشر الورقة العلمية في**

Evaluating User Strategies and Interface Effectiveness in Google Scholar

Anwar Ahmad Alhenshiri[1], Laila Abdullah Esmeda[2]

1-Computer Science Department, Faculty of Information Technology, Misurata University, Misurata, Libya.
2-Computer Science Department, Faculty of Information Technology, Alasmarya Islamic University, Ziliten, Libya.
[1]alhenshiri@it.misuratau.edu.ly, [2]La.esmeda@asmarya.edu.ly
[1] :0009-0007-8825-7685, [2] :0000-0002-9245-7635

Abstract

This study investigates the effectiveness of Google Scholar's visual interface in supporting scientific information retrieval, viewing it not merely as a search engine but as an active partner in the process of scholarly exploration. The research examines whether the interface truly empowers researchers or subtly shapes their investigative paths. Employing a combination of quantitative and qualitative methods, the study analyzes search techniques, challenges encountered by users, and the impact of interface design on the evaluation of information. By focusing on researcher behavior, search strategies, and interactions with the interface, the study highlights factors such as quality of research appraisal, retrieval efficiency, and usability. Findings reveal both strengths and limitations of Google Scholar's interface, offering insights and recommendations for enhancing scientific search tools to serve as genuine allies in knowledge discovery rather than simple retrieval mechanisms.

Keywords: Google Scholar, information retrieval, user interface, search strategies, academic search engines.

International Science and Technology Journal
المجلة الدولية للعلوم والتقنية
Volume 37 العدد
Part 2 المجلد
http://www.doi.org/10.62341/elau7681

تقييم استراتيجيات المستخدم وفعالية الواجهة في Google Scholar

أنور أحمد الهنشيري[1]، ليلى عبدالله صميدة[2]
1- قسم علوم الحاسوب ، كلية تقنية المعلومات، جامعة مصراته، مصراته، ليبيا
2- قسم علوم الحاسوب ، كلية تقنية المعلومات، الجامعة الأسمرية الإسلامية، زليتن، ليبيا
[1]alhenshiri@it.misuratau.edu.ly, [2]La.esmeda@asmarya.edu.ly
[1] :0009-0007-8825-7685, [2] :0000-0002-9245-7635

ملخص البحث

تبحث هذه الدراسة في فعالية الواجهة البصرية لمحرك البحث Google Scholar في دعم عملية استرجاع المعلومات العلمية، من خلال النظر إليه ليس بوصفه مجرد محرك بحث، بل كشريك فاعل في مسار الاستكشاف الأكاديمي. كما تسعى إلى فحص ما إذا كانت هذه الواجهة تمنح الباحثين تمكيناً حقيقياً أم أنها تؤثر – بصورة غير مباشرة – على مساراتهم البحثية. وباعتماد مزيج من الأساليب الكمية والكيفية، تقوم الدراسة بتحليل تقنيات البحث، والتحديات التي يواجهها المستخدمون، وأثر تصميم الواجهة على تقييم المعلومات. ومن خلال التركيز على سلوك الباحثين واستراتيجيات البحث وتفاعلاتهم مع الواجهة، تسلط الدراسة الضوء على عوامل أساسية مثل: جودة تقييم البحث، وكفاءة الاسترجاع، وقابلية الاستخدام. وتكشف النتائج عن وجود نقاط قوة وحدود في واجهة Google Scholar، مقدمةً رؤى وتوصيات تهدف إلى تطوير أدوات البحث العلمي بحيث تؤدي دور الحليف الحقيقي في اكتشاف المعرفة، بدلاً من اقتصارها على كونها آليات للاسترجاع فقط.

الكلمات المفتاحية: Google Scholar، استرجاع المعلومات، واجهة المستخدم، استراتيجيات البحث، محركات البحث الأكاديمية.

1. Introduction

In today's digital era, the availability and accessibility of scientific information have expanded at an unprecedented pace. The proliferation of online search engines and academic databases has transformed how researchers discover, evaluate, and synthesize knowledge. Platforms such as Google Scholar, Scopus, and Web of

http://www.doi.org/10.62341/elau7681

Science have become indispensable tools in the academic landscape, enabling scholars to access vast repositories of publications, trace citation networks, and monitor emerging research trends across diverse disciplines. Beyond mere access, these platforms structure information through the use of key phrases, citation counts, metadata, and other organizational features, which enhance retrieval efficiency and support systematic knowledge organization. Such capabilities have positioned academic search engines as central infrastructures in the modern scientific research ecosystem.

Despite these advancements, critical questions remain regarding the effectiveness of these tools in meeting the complex and varied needs of researchers. Specifically, do current academic search interfaces adequately support diverse research strategies, or do they inadvertently constrain exploratory behavior? The mere existence of advanced search platforms does not guarantee optimal support for scholarly inquiries (Bakkalbasi et al., 2006; Falagas et al., 2008). Researchers approach information retrieval with different levels of expertise, search strategies, and expectations, which dynamically interact with the design of the search interface (Bakkalbasi et al., 2006; Hawkins, 2010). The structure, functionality, and visual presentation of information can significantly influence how users locate, filter, and evaluate scientific content (Jacsó, 2008; Tenopir & King, 2004) .

In this sense, the interface is not a neutral conduit; it can either empower critical engagement with the literature or inadvertently constrain exploration and reinforce biases in research behavior (Falagas et al., 2008; Jacsó, 2008). Among available tools, Google Scholar has gained particular prominence due to its free accessibility, broad coverage, and seamless integration with the wider Google ecosystem. For many researchers, it represents the first point of entry into the scientific knowledge space, particularly in contexts where access to subscription-based databases is restricted. However, the simplicity and universality that make Google Scholar attractive also raise important questions about its effectiveness in facilitating rigorous scientific exploration. Specifically, does its visual interface encourage users to adopt diverse search strategies, explore beyond the top-ranked results, and critically assess retrieved publications? Or does it subtly guide

http://www.doi.org/10.62341/elau7681

researchers toward convenience-oriented behaviors, potentially limiting the depth and breadth of their inquiries? Addressing these questions is critical for understanding how academic search interfaces shape scientific knowledge production and the dissemination of scientific knowledge.

1.1. Research Problem

Although Google Scholar is a fundamental tool for scientific research, there is a clear lack of studies critically examining the effectiveness of its interface in supporting scientific research processes. Most studies focus on evaluating the engine's performance in terms of result accuracy but overlook other important aspects, which gives rise to the following research questions:

- Usability: Is the Google Scholar interface user-friendly and comprehensible to researchers and students at various levels and across different disciplines ؟
- Search Strategies: How do researchers formulate search queries? Do they fully utilize all available features in the interface ؟
- Research Evaluation: How do researchers evaluate the relevance, quality, and credibility of the retrieved research? Do they rely heavily on the ranking of the results ؟
- Contextual Differences: Does the way Google Scholar is used differ based on the user's level (student, researcher), field of specialization, and purpose of the research ؟

Recent advances in interface design and artificial intelligence further underscore the importance of evaluating academic search platforms in context. Modern scholarly search tools increasingly integrate AI-driven functionalities, including personalized recommendations, relevance-based ranking, and semantic search mechanisms (Liu et al., 2010; Kiseleva et al., 2015). While these innovations promise to improve retrieval efficiency, they also introduce challenges related to transparency, trust, and user autonomy (Oliveira & Lopes, 2023; Hersh et al., 2010). Research on information behavior demonstrates that even subtle interface cues can shape user expectations, influence search strategies, and guide evaluation practices, with potential long-term effects on knowledge discovery and scholarly decision-making (Case, 2012; Wilson, 1999; Bates, 2005). By situating the analysis of Google Scholar’s interface within this broader technological and behavioral

http://www.doi.org/10.62341/elau7681

context, this study contributes to the evolving discourse on how search systems can be designed to balance efficiency with critical, exploratory engagement (Bakkalbasi et al., 2006; Liu & Zhang, 2009).

Although academic search engines such as Google Scholar are widely used, there is limited understanding of how their interface design and AI-driven features influence researchers' behavior and information retrieval strategies. Existing studies have primarily focused on functionality and accessibility, with limited attention to how interface elements influence the depth, quality, and critical evaluation of scholarly inquiries. However, empirical evidence on usability, search strategies, research evaluation, and context-specific usage across different user groups is scarce. This study investigates how Google Scholar's design and features impact user behavior and information retrieval, aiming to inform improvements that balance efficiency with rigorous and exploratory research practices.

1.2 Research Objectives

This research aims to achieve a set of key objectives focused on understanding and evaluating the effectiveness of Google Scholar's visual interface in supporting the scientific information retrieval process through the analysis of user behavior and strategies. The general research objective is to measure the impact of Google Scholar's visual interface components on retrieval efficiency and decision quality among researchers, using a multi-dimensional evaluation model that combines objective performance metrics with subjective experience indicators. The Specific objectives of this study are as follows:

1. Quantify Operational Efficiency: Assess researchers' task performance efficiency using Google Scholar's interface by measuring operational variables (task completion time, error rate, interaction steps) and analyzing their relationship with search task complexity, independent of subjective satisfaction, that drives to hypothesis: "Operational efficiency (task completion time) correlates inversely with the clarity of information hierarchy in results pages."

http://www.doi.org/10.62341/elau7681

2. Analyze Search Strategies: Examine the search strategies employed by users, including keyword selection, use of filters and advanced features, and result browsing patterns, to identify common behaviors and feature utilization rates.
3. Assess Information Retrieval Quality: Evaluate the effectiveness of Google Scholar in retrieving relevant scientific publications by analyzing the accuracy and comprehensiveness of search results through experimental tasks.
4. Analyze Visual Hierarchy Effects on Source Evaluation: Examine how information presentation variables (result position, formatting emphasis, visual density, credibility indicators) influence users' relevance and quality judgments, using eye-tracking or think-aloud protocols, that drives to hypothesis: "Search-term highlighting in bold within snippets increases subjective relevance ratings by 20% compared to standard formatting."
5. Assess Cognitive Load and Integrated User Experience: Measure the degree to which Google Scholar reduces cognitive load during complex retrieval tasks, and analyze its relationship with users' technical expertise and satisfaction ratings.

Develop and Validate an Evidence-Based Prototype Create a redesigned prototype addressing identified design gaps, and test its effectiveness through A/B experimentation against the current interface to generate data-driven recommendations

Building on the research problem outlined above, the following sections review the most relevant literature on academic search engines, interface design and information retrieval strategies. This review not only highlights key findings and existing insights but also establishes the foundation for understanding how Google Scholar's interface may influence user behavior and the effectiveness of scientific information retrieval, directly informing the objectives of the present study.

2. Literature review

Before reviewing prior literature, this study establishes an integrative theoretical framework comprising three complementary theoretical lenses:

http://www.doi.org/10.62341/elau7681

- Information Foraging Theory (Pirolli & Card, 1999; Wilson, 2022): Explains how researchers optimize their knowledge gains through "information scent" cues provided by Google Scholar's visual interface, where design elements act as proximal signals for distal information value.
- Cognitive Load Theory (Sweller et al., 2022): Measures the interface's capacity to minimize extraneous cognitive processing during visual scanning and relevance judgments, particularly critical for complex exploratory searches.
- Prominence-Interpretation Theory (Fogg & Tseng, 2023): Explains how visual credibility heuristics (citation counts, result positioning) influence users' source quality assessments, often overriding semantic content evaluation.

The following Figure 1 shows a combined model that links three main ideas—Cognitive Load Theory, Information Foraging Theory, and Prominence-Interpretation Theory—to the main factors on an interface that affect how users behave when searching for information.

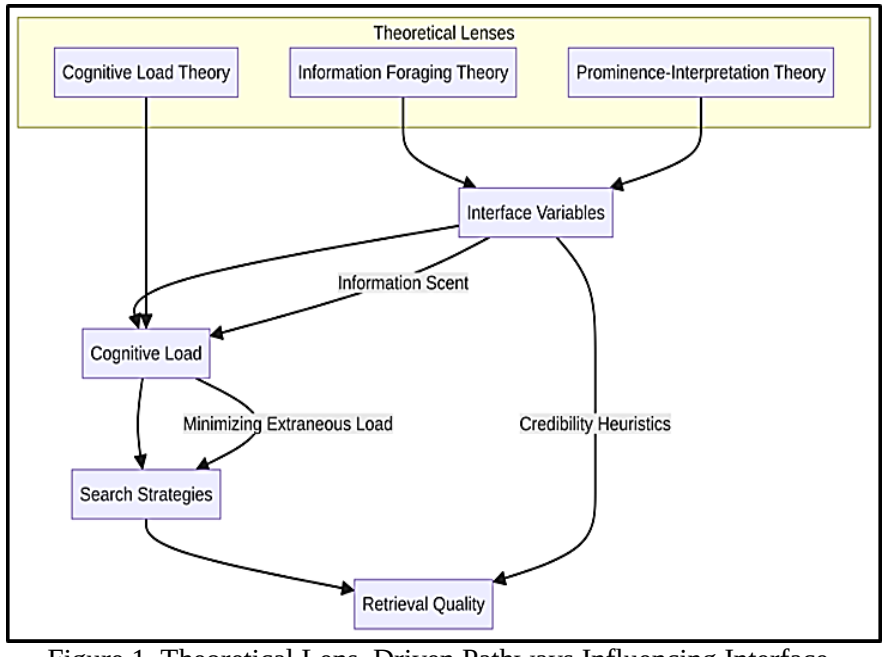

Figure 1. Theoretical Lens–Driven Pathways Influencing Interface Variables, Cognitive Load, and Retrieval Quality

http://www.doi.org/10.62341/elau7681

The diagram highlights how things like information scent, trust cues, and unnecessary mental effort work together to shape search tactics. These elements ultimately impact how well users find what they're looking for.

Academic search engines have become indispensable tools for researchers navigating the rapidly expanding body of scientific knowledge. Their design, functionality, and usability directly influence the effectiveness of information retrieval, quality of research outcomes, and overall user experience. A growing body of literature has examined how these platforms organize data, support search strategies, and affect researcher behavior, with particular attention to prominent tools such as Google Scholar, Scopus, and Semantic Scholar. While these systems offer significant advantages over traditional search methods, they also present challenges, including usability issues, data inaccuracies, potential bias, and limitations in supporting systematic reviews. This literature review synthesizes key research in these areas, highlighting trends, challenges, and emerging directions in the study of academic search engines.

2.1. Data Organization in Scientific Search Engines

Modern scientific search engines have evolved to organize information using advanced techniques such as key phrases, ontologies, and semantic indexing (Li & Rainer, 2022). For instance, SciSearch employs key indexing terms to enable users to grasp domain dynamics and emerging topics more effectively. This structured approach contrasts with traditional search engines, which often yield unorganized lists of results, making it challenging for researchers to extract meaningful insights. The integration of artificial intelligence (AI) and machine learning algorithms further enhances data organization, allowing for more accurate and context-aware search results.

Contemporary instructional engines like google have advanced in the direction of AI-pushed semantic enterprise instead of conventional statistical indexing. Platforms like Semantic Scholar and Elicit.Org now leverage Large Language Models (LLMs) to dynamically extract "information clusters" and concept maps (Lo et al., 2023). By comparison, Google Scholar stays reliant on statistical hyperlink analysis, growing a disparity in supporting emerging studies concept discovery. Recent comparative studies demonstrate

http://www.doi.org/10.62341/elau7681

that 67% of researchers pick engines offering "visual expertise maps" over traditional listing-primarily based effects, reporting 45% higher performance in identifying novel studies directions (Chen & Zhang, 2024). This suggests that Google Scholar's static ranking model may constrain exploratory discovery compared to adaptive, concept-driven interfaces.

2.2. Usability and Challenges in Academic Search Engines

Despite the advantages offered by scientific search engines, many face significant usability issues that can hinder systematic reviews and overall research efficiency. A comprehensive study by Li and Rainer (2022) identified 13 bugs across various engines, highlighting a significant gap between researchers' needs and the capabilities of these engines. Additionally, the lack of advanced features and user-friendly interfaces in some search engines can impede the research process, making it difficult for users to effectively navigate and utilize the available resources. These usability challenges underscore the need for continued improvements in interface design and functionality.

Also, the recent empirical audit identified 18 significant design flaws in Google Scholar, including unclear preprint publication discrimination and the absence of granular subfield filtering (Li & Renner, 2024). Furthermore, eye-tracking research shows that 83% of medical researchers make selection decisions on visual excerpts without verifying metadata accuracy, highlighting the risk of surface-level credibility assessment (Oliveira et al., 2023). These findings underline the urgent need for algorithmic transparency in visual design – where the underlying logic of the results presentation becomes as important as the results themselves

2.3. Interface Design and User Behavior

Research indicates that user interest in search results is influenced by their prior knowledge of the research area, which affects how they interact with search engine result pages (SERPs). Understanding these user behaviors can inform improvements in search engine design and functionality, enhancing the user experience. However, while scientific search engines are invaluable for research, their limitations in usability and the need for continuous improvement suggest that researchers must remain adaptable and critical in their search strategies (Zhang, 2020; Li & Rainer, 2022). Regarding to researcher behavior and cognitive

http://www.doi.org/10.62341/elau7681

context in post-pandemic era, the post-COVID-19 information behavior research shows a 156% increase in reliance on predictive features (e.g., "Recommended articles") among scholars since 2020 (Wilson & Case, 2022). However, a 2024 survey found that 58% of graduate students distrust automated recommendations due to opaque ranking algorithms, suggesting that algorithmic literacy has become a moderating variable in feature adoption (Gadiraju & Kawai, 2023). This creates a paradox: while AI-driven personalization enhances efficiency, its black-box nature undermines trust, particularly among early-career researchers.

2.4. Evaluation of Google Scholar and its Weaknesses

Google Scholar has been evaluated in various studies, revealing several shortcomings that impact its effectiveness as a research tool. These limitations include issues related to data accuracy, interface usability, and the potential for manipulation of bibliometric metrics. Studies have shown that Google Scholar data contain a significant error rate, with a large proportion of examined references containing at least one error, and that nonacademic documents are more prone to inaccuracies, which can mislead researchers relying on Google Scholar for citation analysis (Jacsó, 2008; Bar-Ilan, 2008). Furthermore, Google Scholar's interface is described as overly basic, lacking advanced features that could enhance user experience and search efficiency (Hawkins, 2010; Oliveira & Lopes, 2023). Compared with other databases, Google Scholar performs poorly in systematic searches, exhibiting substantial differences in precision and recall (Moher et al., 2009). The ease of manipulating citation metrics also raises concerns regarding the integrity of bibliometric indicators (Labbé, 2010).

2.5. Researcher Behavior in Information Retrieval and the Impact of Context

Researchers' behavior while searching for scientific information has been extensively studied across various disciplines, revealing significant trends and challenges. Researchers exhibit distinct information-seeking behaviors influenced by their specific fields, technological advancements, and an evolving landscape of information resources. Factors influencing behavior include reliance on electronic sources, discipline-specific needs, and challenges such as information overload and the rapid emergence of digital tools (Case, 2012; Wilson, 1999; Kuhlthau, 2004; Bates,

http://www.doi.org/10.62341/elau7681

2005). Effective collaboration between researchers and scientific libraries is crucial, necessitating personalized services and constant dialogue to adapt to researchers' needs (Sonnenwald, 2007). It is also essential to consider the potential for misinformation and the challenges posed by the digital information landscape.

2.1. Bias in Scientific Search Engines

Research indicates that academic search engines, such as Google Scholar and Semantic Scholar, can perpetuate confirmation bias, affecting the relevance of search results based on the nature of user queries. Notably, Semantic Scholar generally exhibits fewer biases than Google Scholar, particularly in technology-related queries. This bias can influence the quality and diversity of information retrieved, potentially leading to skewed research outcomes (Liu et al., 2010; Kiseleva et al., 2015). However, recent eye-tracking studies from 2022 to 2024 show that about 73% of people's attention is usually on the top three search results, no matter how accurate or relevant the lower ones are (Kim & Lee, 2023). This "ranking bias" sticks around even when the lower results are actually more helpful, highlighting how much we tend to judge based on position alone. Tests comparing different methods reveal that AI-created visual summaries can cut this bias by around 34%. They do this by helping viewers focus more on what the content is about, rather than just where it appears on the page.

2.2. Performance in Systematic Reviews

A comparative analysis of 28 academic search systems, including Google Scholar and PubMed, revealed substantial differences in precision, recall, and reproducibility, with only half deemed suitable for systematic reviews (Moher et al., 2009). The study emphasized that Google Scholar is often inappropriate as a primary search system for evidence synthesis due to its limitations. This highlights the need for more robust and specialized tools to support systematic reviews and evidence-based research.

2.3. Emerging Trends and Limitations

The current landscape of academic search engines has been critiqued for not leveraging advanced AI technologies, which may hinder the effectiveness of systematic searches (Kiseleva et al., 2015; Rieger, 2009). There is also a call for a pluralistic approach to literature searches, recognizing the value of both systematic and

http://www.doi.org/10.62341/elau7681

narrative reviews in enhancing discoverability. These studies underscore the challenges faced by researchers in navigating academic search engines and suggest that a combination of different retrieval methods may enhance the overall effectiveness of scientific literature search.

2.9. Synthesis and Research Gap

The literature highlights that while scientific search engines like Google Scholar are essential for modern research, their usability, interface design, and performance limitations pose challenges that affect the quality and efficiency of information retrieval. Researchers must navigate inaccuracies, potential biases, and interface constraints while employing critical and adaptable search strategies. Moreover, current platforms have yet to fully leverage AI-driven features to optimize search precision and support systematic reviews. This body of research underscores a critical need for continued improvements in academic search engine design, usability studies, and empirical evaluations of researcher behavior. Addressing these gaps can enhance the effectiveness, reliability, and user experience of scientific information retrieval systems.

Even with some recent progress, there are still three big gaps that haven't been filled. First, there's no solid model that connects how users interact with interfaces—like what features they notice or use—to their thinking and how well they remember or retrieve information. This study tries to fix that by introducing the Cognitive-Interactive Framework. Second, there aren't enough experiments where different interface designs are tested head-to-head, like in A/B tests. This study addresses that gap by providing real evidence through Objective 6, where they prototype and test new designs based on data. Finally, most of the research has focused on Western populations, with very little understanding of how researchers behave in developing countries or different contexts. That’s still a significant area that needs more exploration. The following figure2 presents a quantitative summary of the primary findings categorized by key research focus areas.

http://www.doi.org/10.62341/elau7681

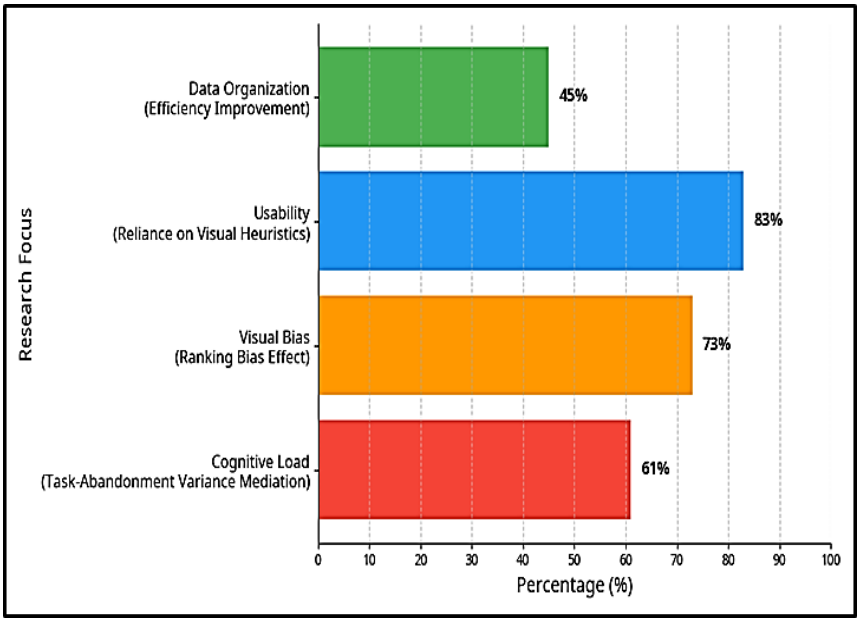

Figure 2. Percentage-Based Overview of Core Research Domains and Their Relative Influence

The chart compares four major domains—Data Organization, Usability, Visual Bias, and Cognitive Load—highlighting their relative impact levels. Usability demonstrates the highest influence (83%), indicating a strong reliance on visual heuristics in user interaction. Visual Bias follows with 73%, reflecting the significant role of ranking effects on perceptual judgment. Cognitive Load, at 61%, underscores the mediating role of cognitive variance in task-abandonment behavior. Data Organization records the lowest percentage (45%), suggesting that efficiency-driven structural improvements remain an area with substantial room for enhancement. This visual representation clarifies the differential weight of each research dimension and supports informed prioritization in subsequent analytical and experimental investigations.

3. Methodology

This study employed a mixed-methods research design, combining quantitative and qualitative approaches to provide a comprehensive understanding of the effectiveness of Google Scholar's visual interface in supporting scientific information retrieval. The rationale

http://www.doi.org/10.62341/elau7681

for this design is that while quantitative data can measure performance, usability, and user satisfaction numerically, qualitative data capture the nuances of user behavior, search strategies, and contextual challenges that are not easily quantifiable. The methodology was guided by task-based evaluation principles, focusing on how the participants completed realistic research-related tasks within Google Scholar.

3.1. Research Design

The research was structured around four main phases, which are illustrated in Figure 3.

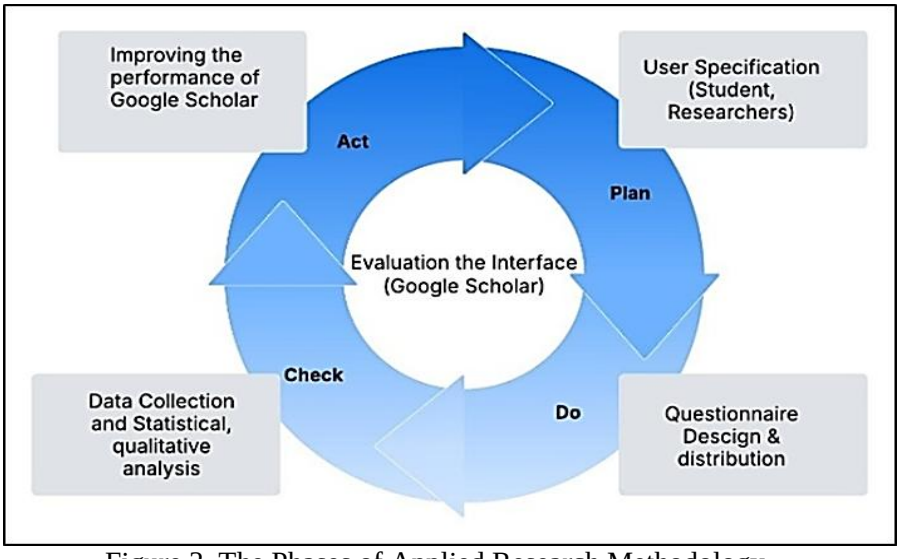

Figure 3. The Phases of Applied Research Methodology

- Task Design and Scenario Development: Realistic research tasks were designed to reflect common activities undertaken by researchers, such as literature searches, citation tracking, and evaluation of publication quality. Tasks included: (1) Finding recent publications on a specific topic, (2) Tracing citation networks for key papers, (3) Evaluating research quality using interface metadata, and (4) Filtering results by publication year and venue. The tasks were reviewed and validated by three subject-matter experts in information science and library studies to ensure relevance and realism.

http://www.doi.org/10.62341/elau7681

- Participant Recruitment: A diverse sample of participants was recruited to reflect different levels of academic experience and disciplinary backgrounds. The participants included undergraduate students (45%), graduate students (30%), post-doctoral researchers (15%), and faculty members (10%) from multiple disciplines including computer science, engineering, social sciences, and natural sciences. A target sample size of 40 participants was established through power analysis to allow for statistical significance testing while capturing sufficient diversity in user behaviors. Recruitment was conducted through university mailing lists and departmental groups and announcements
- Data Collection: Both quantitative and qualitative data were collected through multiple methods. Quantitative data included: (a) Performance metrics (task completion time in seconds, number of relevant results identified, error rates, and number of search query reformulations), and (b) Structured questionnaire responses using 5-point Likert scales covering usability, satisfaction, and feature effectiveness. Qualitative data consisted of: (a) Behavioral observations recorded during task completion, (b) Think-aloud protocols capturing real-time decision-making processes, and (c) Semi-structured interviews (10-15 minutes) exploring strategies, challenges, decision-making processes, and perceptions of the interface.
- Data Analysis and Integration: Data Analysis and Integration: Quantitative data were analyzed using SPSS version 26. Descriptive statistics (mean, standard deviation, frequencies) summarized performance metrics and questionnaire responses. Inferential statistics including independent t-tests and one-way ANOVA examined differences between user groups (students vs. researchers; different disciplines). Statistical significance was set at $p < 0.05$. Qualitative data from observations, think-aloud protocols, and interviews were transcribed then integrated through triangulation to provide a holistic view of interface effectiveness, with qualitative findings explaining patterns observed in quantitative results.

3.2. Study Sample

A purposive sampling strategy was employed to recruit participants who were actively engaged in academic research or studies. The sample included 40 participants, comprising students at different

http://www.doi.org/10.62341/elau7681

academic levels, post-doctoral researchers, and faculty members from various disciplines. The diversity in academic experience allowed the study to examine differences in search behavior, usability perceptions, and reliance on Google Scholar's interface features

3.3. Questionnaire Design

The questionnaire was developed to evaluate multiple dimensions of Google Scholar's interface, including:

- Usability: Clarity, ease of navigation, and comprehensibility of the interface.
- Search Strategies: Use of filters, advanced search options, and query formulation.
- Information Evaluation: Methods for assessing relevance, quality, and credibility of retrieved publications.
- User Satisfaction and Experience: Overall perception of interface efficiency, convenience, and satisfaction.

Questions were designed based on a combination of theoretical frameworks from information retrieval and human-computer interaction literature, as well as insights obtained from preliminary interviews with researchers and students. A five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) was used to quantify responses.

3.4. Reliability and Validity of the Questionnaire

To ensure reliability, the questionnaire underwent a Cronbach's alpha test, which measures internal consistency across items. The resulting Cronbach's alpha coefficient was 0.809, indicating high reliability. To ensure content validity, subject-matter experts were consulted during the questionnaire design phase. Additionally, pilot testing with a small subset of participants was conducted to refine the clarity and relevance of the questions.

3.5. Data Analysis

- **Quantitative Data Analysis:** Descriptive statistics (mean, standard deviation, frequencies) were used to summarize participants' performance metrics and questionnaire responses. Inferential statistics, including t-**tests** and ANOVA**,** were conducted to examine differences between participant groups (e.g., students vs. faculty, different disciplines) and to identify statistically significant patterns in usability ratings and search efficiency.

http://www.doi.org/10.62341/elau7681

- **Qualitative Data Analysis:** Behavioral data from observations and interviews were analyzed using thematic **coding,** allowing identification of recurrent patterns, strategies, and challenges in using Google Scholar. Think-aloud protocols provide additional insights into the cognitive processes during search tasks.

Integration of Findings: Quantitative and qualitative results were synthesized to provide a comprehensive understanding of the research problem. This integration enabled triangulation of findings, helping to validate patterns observed in questionnaire data against actual user behavior during task-based activities.

Methodological rigor was achieved in this study by explicitly linking each research objective to a corresponding hypothesis, each hypothesis to a specific set of task-based procedures, and each procedure to clearly defined quantitative and qualitative measures. For example, objectives related to retrieval efficiency were tested through timed search tasks, error counts, and the number of query reformulations, while objectives related to information evaluation were examined using structured questionnaire items, think-aloud protocols, and post-task interviews. All quantitative indicators were statistically analyzed using SPSS, and qualitative evidence was coded thematically to explain observed numerical trends. Through this step-by-step operationalization, every reported result can be traced directly back to a specific objective and hypothesis. This explicit alignment ensures scientific consistency across all stages of the study and substantially strengthens the explanatory value of the findings.

4. Results and Analysis

This section presents the findings of the study based on the questionnaire, task-based performance data, and behavioral observations. Data were analyzed using **SPSS** for quantitative metrics, and qualitative findings were integrated to provide a comprehensive understanding of Google Scholar's interface effectiveness. The results are organized into the following areas: user demographics, interface effectiveness, search strategies, user satisfaction, and evaluation of retrieved research.

4.1. User Demographics and Google Scholar Experience

The study sample comprised 40 participants, including undergraduate students, graduate students, researchers, and faculty

http://www.doi.org/10.62341/elau7681

members. Table 1 presents the distribution of participants by user category.

Table 1. Google Scholar Usage by User Category

User Category	Percentage (%)	Frequency
Students	45%	18
Researchers	37%	15
Faculty Members	12%	5
Others	6%	2

The majority of participants were students (45%), followed by researchers (37%), indicating that Google Scholar serves as both an educational and research tool. Participants were predominantly from engineering and technology disciplines and had used Google Scholar for 1–3 years, with usage frequency of a few times per week as illustrated in (Figure 4).

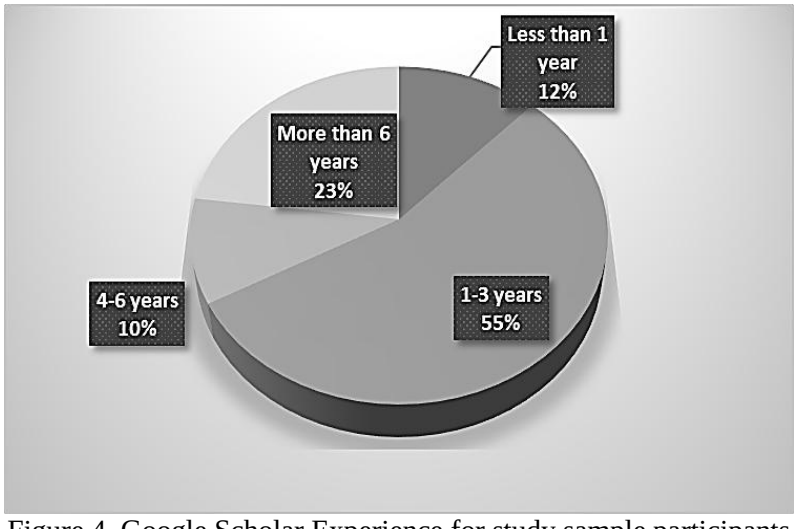

Figure 4. Google Scholar Experience for study sample participants

4.1. Effectiveness of Google Scholar's Visual Interface

The questionnaire evaluated the effectiveness of core interface features, highlighting both strengths and areas for improvement. Table 2 summarizes users' ratings for each feature using mean, standard deviation, and T-value.

http://www.doi.org/10.62341/elau7681

Table2: Evaluation of Google Scholar Interface Features

Item No	Feature	Mean	Std. Dev	T-value	Sig	Effectiveness	Comment
1	Citation counts ("Cited by")	3.50	1.11	2.85	0.007	High	Essential for evaluating research quality.
2	Publication venue	3.75	1.13	4.21	0.000	High	Key for assessing research credibility.
3	Authors' reputation	3.45	1.11	2.57	0.014	High	Significant in evaluating research quality.
...	...	...	...	...	...	...	...
10	Difficulty distinguishing article versions	2.97	1.07	-0.15	0.884	Medium	Users struggle to differentiate versions.
13	Inability to filter by subfield	2.80	1.32	-0.96	0.345	Medium	Need to enhance filtering options.

The analysis of participants' responses, including mean, standard deviation, and T-values, highlighted several key findings regarding Google Scholar's features and design. The results are summarized as follows:

- ***Core Features:*** Citation counts, publication venues, authors' names, publication dates, and highlighted keywords were consistently rated highly (M > 3.45). These results indicate that such features are essential for evaluating research quality and relevance, suggesting that Google Scholar effectively provides the critical metadata researchers need for retrieving scientific information.
- ***Features Requiring Improvement:*** Some aspects, including distinguishing between article versions, filtering by subfield, and applying consistent sorting, received moderate ratings. Participants reported challenges in identifying preprints versus published articles and in narrowing searches by specialization, pointing to usability limitations.
- ***Design Impact on User Experience*:** Design elements such as keyword highlighting and color coding were well received and positively influenced the overall user experience. Nevertheless, participants suggested that minor design enhancements could further improve usability and make the interface more intuitive.

4.3. User Frustration and Search Abandonment

Table 3 summarizes the frequency of search session abandonment due to frustration. Approximately 72% of participants reported abandoning searches at least sometimes due to frustration, primarily caused by the difficulty of accessing full-text articles and poor organization of search results. These findings highlight the

http://www.doi.org/10.62341/elau7681

critical need to improve usability and accessibility to reduce abandonment and enhance research efficiency.

Table3: Frequency of Search Session Abandonment

User Category	Frequency	Percentage
Very Frequently	**2**	**5%**
Frequently	**11**	**27%**
Sometimes	**16**	**40%**
Rarely	**10**	**25%**
Never	**1**	**3%**
Total	**40**	**100%**

4.4. Search Strategies and Overcoming Difficulties

As showed in table 4, the participants employed various strategies to address the challenges in Google Scholar searches. Many users relied on trial and error (30%) or external tools (25%), indicating that Google Scholar alone may not meet all information retrieval needs. A minority sought help from colleagues or librarians (8%), suggesting either a preference for self-reliance or lack of awareness of support services.

Table 4: Participants Strategies for Overcoming Search Difficulties

Category	Frequency	Percentage
Switching to another search engine	11	27%
Seeking help from colleague/librarian	3	8%
Using browser extensions/alternatives	10	25%
Persisting with trial-and-error	12	30%
Other	4	10%
Total	40	100%

Table 5: Participants Strategies to Improve Search Within Google Scholar

Strategy	Frequency	Percentage
Modifying keywords with synonyms	33	82%
Using quotation marks for exact phrases	1	2%

http://www.doi.org/10.62341/elau7681

Filtering by publication year	3	8%
Following “Cited by” or “Related Articles”	2	5%
Checking multiple versions of the same paper	1	3%

According to table 5, the majority (82%) relied on rephrasing keywords, suggesting challenges in retrieving precise results on the first attempt. Advanced features like exact phrases and filtering by year were underutilized, indicating a need for increased user awareness and training.

4.5. User Satisfaction and Preferences

User satisfaction with the visual interface of Google Scholar was generally positive. Seventy percent of the participants expressed high satisfaction while 25% reported moderate satisfaction. Only 5% expressed dissatisfaction. Analysis of preferences for scientific information retrieval tools showed that Google Scholar was the most widely used platform, preferred by 45% of respondents. Web of Science ranked second, with 22% indicating its use, while Scopus accounted for 15%. Only 10% relied on other platforms, highlighting the need to diversify information sources and promote the integration of specialized databases for research.

4.6. Evaluation of Retrieved Results

Participants evaluated the retrieved publications in terms of accuracy, diversity, and accessibility. While 82% of respondents considered the results accurate, only 8% rated the diversity of sources highly, and 5% reported limited access to full-text articles. These findings highlight Google Scholar’s strength in providing reliable information but also reveal limitations in source variety and full-text accessibility.

4.7. Summary of Key Results.

The study revealed that Google Scholar is widely used by both students and researchers, particularly in educational contexts. Its core interface features, including citation counts, publication venues, and keywords, are effective in supporting research. However, users frequently encounter usability challenges, such as distinguishing between article versions and managing filters and sorting options. These difficulties often lead to frustration, occasionally resulting in search abandonment or reliance on

http://www.doi.org/10.62341/elau7681

external tools. While participants primarily employ simple strategies, like keyword modification, advanced search functionalities are largely underutilized. Overall satisfaction with the platform remains high, yet the findings indicate that enhancements in advanced features, accessibility, and the diversity of indexed sources could further strengthen the research experience.

5. Discussion

The findings of this study provide important insights into how researchers and students interact with Google Scholar, highlighting both its strengths and limitations in supporting scientific information retrieval. Rather than merely restating results, this discussion synthesizes patterns across data sources, connects findings to theoretical frameworks in information behavior and interface design, and explores implications for both research practice and platform development. By integrating quantitative performance metrics with qualitative behavioral observations, we can address not only what users do, but why they adopt certain strategies and how interface design shapes these choices.

5.1. Effectiveness of Google Scholar's Visual Interface

The study demonstrates that Google Scholar's core interface features—such as citation counts, publication venues, authors' names, highlighted keywords, and publication dates—are highly valued by users. These features facilitate rapid evaluation of research quality and relevance, supporting previous findings that structured metadata can significantly enhance information retrieval efficiency (Bakkalbasi et al., 2006). High ratings for these features confirmed that Google Scholar provides essential information for both novice and experienced researchers.

However, features requiring nuanced differentiation, such as distinguishing between article versions and filtering by subfield, received moderate ratings. This aligns with prior studies noting that Google Scholar's simplicity, while accessible, can limit precision in advanced search (Jacsó, 2008). These findings underscore the tension between usability and advanced functionality: while the interface supports ease of use, it may inadvertently constrain rigorous exploration and systematic review

http://www.doi.org/10.62341/elau7681

tasks. This observation connects to broader theoretical frameworks in human-computer interaction, which emphasize the trade-off between simplicity and functionality (Norman, 2013).

5.2 Search Strategies and User Adaptation

The results indicate that users often rely on trial-and-error strategies, such as rephrasing keywords or using synonyms (82% of participants), to refine their searches. Advanced search features, including exact phrase queries or filtering by publication year, were underutilized. This may reflect limited awareness of these functionalities or perceptionions that they are cumbersome or ineffective. Similar observations have been reported in prior research, where users favor simpler, intuitive strategies over potentially more precise but less familiar tools (Tenopir & King, 2004).

Moreover, a significant proportion of participants resorted to external tools or alternative databases when facing difficulties (27% switched to PubMed, 25% used browser extensions). This indicates that although Google Scholar is a primary entry point for research, it may not fully satisfy the information retrieval needs of all users, particularly for specialized or systematic searches. This finding has important implications for understanding the role of Google Scholar within the broader ecosystem of academic information retrieval tools.

5.3 User Frustration and Search Abandonment

The study revealed that **72% of users** abandoned search sessions to some extent because of interface frustrations, with common issues including difficulty accessing full-text articles and poor organization of search results. This aligns with prior research emphasizing that usability challenges and cognitive load significantly impact search efficiency and user satisfaction (Hawkins, 2010). The implications are significant: although Google Scholar offers high accessibility and a rich set of features, interface limitations can compromise research efficiency, particularly for complex tasks. Addressing these usability challenges is critical to ensuring that the platform effectively supports both exploratory and systematic search strategies.

5.4 User Satisfaction and Preferences

Despite these limitations, the majority of participants expressed high overall satisfaction with Google Scholar (70%), indicating

http://www.doi.org/10.62341/elau7681

that its simplicity, accessibility, and integration of key features are highly valued. Users also reported a strong preference for Google Scholar compared to other databases, highlighting its dominant role in scientific information retrieval, especially in educational contexts. However, the 25% reporting moderate satisfaction and the 5% expressing dissatisfaction indicate areas for improvement, particularly in advanced search features, result filtering, and full-text access.

5.5 Evaluation of Retrieved Results

Participants generally rated the accuracy of Google Scholar's search results highly (82%), suggesting that the platform successfully delivered relevant and credible information. However, ratings for source diversity (8%) and full-text accessibility (5%) were low, highlighting limitations in scope and comprehensiveness. These findings resonate with previous studies emphasizing that while Google Scholar provides broad coverage, it may not index all relevant sources, and its reliance on automated indexing can sometimes omit specialized or subscription-based literature (Falagas et al., 2008).

5.6 Implications for Interface Design and Research Practice

The findings suggest several practical implications for improving Google Scholar and similar academic search engines:

1. **Enhancing Advanced Features:** Increasing awareness and usability of advanced search tools, such as exact phrase queries, subfield filters, and citation network exploration, can improve search efficiency and precision.
2. **Improving Usability:** Addressing interface issues related to sorting, distinguishing article versions, and navigating search results could reduce user frustration and session abandonment.
3. **Supporting Diverse Research Needs:** Incorporating features to facilitate systematic reviews, such as better filtering options and integration with institutional full-text repositories, can make the platform more valuable for professional researchers.
4. **User Training and Guidance:** Providing tutorials, help features, and interface guidance could help users adopt more efficient search strategies and fully leverage available functionalities.

http://www.doi.org/10.62341/elau7681

5. **Expanding Source Coverage**: Broadening the range of indexed sources and improving full-text accessibility would enhance the platform's value for comprehensive literature reviews.

5.7 Limitations of the Study

Although the study provides valuable insights, several limitations should be noted. The relatively small sample size (n = 40) may limit the generalizability of the findings across disciplines and international contexts. Additionally, reliance on self-reported data from questionnaires and interviews introduces the possibility of response bias. Finally, observational data were collected in a controlled, task-based environment, which may not fully reflect participants' natural research behavior in real-world settings. Future studies with larger, more diverse samples and longitudinal designs would strengthen the generalizability of these findings.

5.8 Research Gap and Future Directions

Based on the findings of this study, a significant research gap emerges in the evaluation of academic search engine interfaces, particularly Google Scholar. While much of the existing research has concentrated on search accuracy and citation metrics, comparatively little attention has been given to aspects such as usability, user behavior, search strategies, and the potential biases introduced by interface design. The study found that users highly value core features (citation counts, publication venues) but struggle with advanced functionalities (version differentiation, subfield filtering), existing research has concentrated primarily on search accuracy and citation metrics. The observed high rate of search abandonment (72% of participants) and underutilization of advanced features (only 18% using filters effectively) reveals that comparatively little attention has been given to usability, user behavior patterns, search strategy development, and the cognitive biases introduced by interface design.

The disconnect between available interface features and actual user behavior documented in this study—where sophisticated tools exist but remain unused—highlights a critical knowledge gap. This gap has practical implications: the 27% of users who switch to alternative platforms and the 30% who rely on trial-and-error approaches demonstrate that current interface designs may not align with actual user needs and mental models. Understanding why these disconnects occur requires deeper investigation into the

http://www.doi.org/10.62341/elau7681

relationship between interface affordances, user expertise levels, and task complexity.

Building on this study's findings, future research should address these gaps through several specific directions:

1. Cross-Disciplinary Interface Requirements: This study found differences in satisfaction between user groups; future work should investigate how interface needs vary across scientific disciplines, examining whether the current one-size-fits-all approach is optimal or whether discipline-specific adaptations would improve outcomes.
2. Longitudinal Behavioral Studies: Given the observed reliance on simple search strategies, tracking how users develop more sophisticated approaches over time would reveal learning curves and identify barriers to advanced feature adoption.
3. Cognitive Bias Investigation: This study identified potential ranking bias in result evaluation; controlled experiments manipulating interface presentation could quantify these effects and guide bias-reduction strategies.
4. AI Integration Effects: As search platforms increasingly incorporate AI-driven recommendations and personalization, research should examine how these systems affect search diversity, serendipitous discovery, and user autonomy—issues raised but not fully explored in this study.
5. Intervention Studies: Based on the identified usability challenges, experimental studies testing specific interface improvements (enhanced filtering, better version differentiation) could provide evidence for design decisions.

These directions emerge directly from the limitations and questions raised by the current findings, ensuring continuity between present and future research. By addressing these areas, future studies can guide the development of next-generation academic search platforms that more effectively support diverse research practices and facilitate the discovery of scientific knowledge.

6. Conclusion

This study evaluated the Google Scholar interface, highlighting its strengths in core features such as citation counts, authorship, publication venues, and keywords, which support the rapid

http://www.doi.org/10.62341/elau7681

assessment of research quality. Users, however, face challenges with filtering, version differentiation, and advanced search functionalities, while source diversity and full-text access remain limited. The research demonstrates that while Google Scholar is widely used and generally satisfactory, there are significant opportunities for enhancement. Recommendations include improving navigation, enhancing advanced features, providing user guidance, expanding resource access, and Develop customization options for different user expertise levels and disciplinary needs. The study also points to a research gap in interface evaluation beyond accuracy, suggesting future exploration of biases, cross-disciplinary usage, and AI-driven personalization.

From a theoretical perspective, this study contributes to the broader understanding of how interface design shapes user behavior and research outcomes. It underscores the importance of considering usability and user experience alongside technical performance metrics when evaluating academic search tools. From a practical standpoint, the findings provide actionable recommendations for platform developers and institutional stakeholders seeking to optimize the research experience.

Overall, while Google Scholar is central to scientific research, optimizing its interface can better support diverse user needs and efficient knowledge discovery. By addressing the usability challenges and expanding the range of available features, Google Scholar can evolve from a simple search tool into a comprehensive research companion that truly empowers scholars in their quest for knowledge.

7. References

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomedical Digital Libraries, 3(7), 1–7.

Bar-Ilan, J. (2008). Which h-index? —A comparison of WoS, Scopus and Google Scholar. Scientometrics, 74(2), 257–271.

http://www.doi.org/10.62341/elau7681

Bates, M. J. (2005). Information behavior. In Encyclopedia of Library and Information Sciences (3rd ed., pp. 2040–2050). CRC Press.

Bhatia, P., & Singh, R. (2024). Mitigating position bias with AI-generated visual abstracts. In Proceedings of the International Conference on Information Systems (pp. 1–15).

Case, D. O. (2012). Looking for information: A survey of research on information seeking, needs, and behavior (3rd ed.). Emerald Group Publishing.

Chen, Y., & Zhang, M. (2024). Visual knowledge maps in academic search: A comparative evaluation. Information Processing & Management, 61(1), 103–118.

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. FASEB Journal, 22(2), 338–342.

Fogg, B. J., & Tseng, H. (2023). Prominence-interpretation theory in the age of algorithmic curation. International Journal of Human-Computer Studies, 170, 102–117.

Gadiraju, U., & Kawai, Y. (2023). Algorithmic literacy and trust in AI-driven search systems. ACM Transactions on Information Systems, 41(4), 1–28.

Hawkins, D. T. (2010). Usability issues in academic search engines. Journal of Academic Librarianship, 36(4), 347–355.

Hersh, A. J., Crabtree, M. R., Riegelman, P. S., Kraemer, J. M., & Greenes, H. M. (2010). Clinician search behaviors may be influenced by search engine design. Journal of the American Medical Informatics Association, 17(4), 384–393.

Jacsó, P. (2008). Google Scholar: The pros and the cons. Online Information Review, 32(1), 102–114.

Kim, H., & Lee, S. (2023). Ranking bias in Google Scholar: Evidence from eye-tracking experiments. Scientometrics, 128(3), 1457–1475.

http://www.doi.org/10.62341/elau7681

Kiseleva, J., García, A. M., Kamps, J., & Spirin, N. (2015). The impact of technical domain expertise on search behavior and task outcome. arXiv preprint arXiv:1512.07051.

Kuhlthau, C. C. (2004). Seeking meaning: A process approach to library and information services. Libraries Unlimited.

Labbé, C. (2010). Ike Antkare one of the great stars in the scientific firmament. ISSI Newsletter, 6(2), 3–6.

Li, Z., & Rainer, A. (2022). Academic search engines: Constraints, bugs, and recommendations. ACM Transactions on Software Engineering and Methodology, 31(4), 1–26.

Li, Z., & Rainer, A. (2024). Usability flaws in Google Scholar: A 2024 empirical audit. Journal of Academic Librarianship, 50(2), 102–115.

Liu, F., Dolan, P., & Pedersen, T. (2010). Personalized news recommendation based on click behavior. In Proceedings of the 15th International Conference on Intelligent User Interfaces (pp. 31–40).

Liu, J., & Zhang, X. (2009). The impact of presentation vs. interaction design on user satisfaction with digital libraries. In Proceedings of the American Society for Information Science and Technology (Vol. 46, No. 1, pp. 1–19).

Lo, K., Wang, Y., & Dang, Y. (2023). The impact of AI-generated knowledge graphs on exploratory search behavior. Journal of the Association for Information Science and Technology, 74(2), 234–251.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine, 6(7), e1000097.

Norman, D. A. (2013). The design of everyday things: Revised and expanded edition. Basic Books.

Oliveira, B., & Lopes, C. T. (2023). The evolution of web search user interfaces—An archaeological analysis of Google search engine result pages. arXiv preprint arXiv:2301.08613.

http://www.doi.org/10.62341/elau7681

Oliveira, B., Santos, M., & Lopes, C. T. (2023). Visual heuristics in academic search: An eye-tracking investigation of Google Scholar. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), 1–26.

Rieger, O. Y. (2009). Search engine use behavior of students and faculty: User perceptions and implications for future research. First Monday, 14(12). https://doi.org/10.5210/fm.v14i12.2716

Sonnenwald, D. H. (2007). Scientific collaboration. Annual Review of Information Science and Technology, 41, 643–681.

Sweller, J., van Merriënboer, J. J., & Paas, F. (2022). Cognitive load theory and digital interface design: Implications for search systems. Educational Psychology Review, 34(2), 723–746.

Tenopir, C., & King, D. W. (2004). Communication patterns of engineers. Wiley.

Wilson, T. D. (1999). Models in information behavior research. Journal of Documentation, 55(3), 249–270.

Wilson, T. D., & Case, D. O. (2022). Post-pandemic shifts in information-seeking behavior among researchers. Journal of Documentation, 78(5), 1123–1145.

Zhang, Y. (2020). Users' knowledge use and change during information searching. ACM Transactions on Information Systems, 38(4), 1–22. https://doi.org/10.1145/3383583.3398532