

Modern Search Engines: A Journey of Technological Advancements, Privacy, and User-Centered Improvements for Information Need

Ayesha Khan ¹, Sarah Bukhari ^{1*}, Umar Khattab ¹, Abubakar Kamal ¹

¹Department of Information and communication technology, Bahauddin Zakariya University, Multan, Pakistan

Keywords: Search Engine, Information Retrieval, Precision, Recall, Privacy, Algorithmic Bias.

Journal Info:
Submitted: May 05, 2025
Accepted: September 05, 2025
Published: September 20, 2025

Abstract

Search Engines (SEs) become increasingly important for locating, organizing, and retrieving relevant information in a world of massive and complex digital information. However, contemporary SEs are poor at striking a balance between precision and comprehensiveness, particularly when conducting exploratory or specific information queries. Therefore, the study aims to find out which search engine most effectively satisfied the needs of various users with a lower amount of time and effort. In order to solve this issue, ten such as SEs consisting of Google, Bing, Yahoo, DuckDuckGo, Swisscows, Qwant, Ecosia, Ask.com, WolframAlpha, and StartPage, were tested using two user-effort-sensitive measures: time to relevant result and number of relevant results. The study employed a lot of standard search queries and placed emphasis on the performance of each engine. Ultimately, we established that Google (83%) and StartPage (80%) delivered the most suitable results, while Ask.com (26%) and Swisscows (43%) performed less. In conclusion, contemporary SEs are getting better, but current ones still contain privacy problems, omitted results, and unbalanced performance. The future research ought to be dedicated to enhancing ranking techniques, better search systems, and functionality which will benefit even more people, such as non-English users.

*Correspondence author email address: drsarahbukhari@gmail.com

DOI: [10.21015/vtse.v13i3.2137](https://doi.org/10.21015/vtse.v13i3.2137)

1 Introduction

SEs become increasingly important for locating, organizing, and retrieving relevant information in a world of massive and complex digital information. Users' demands for speed, relevance, privacy, and

increasingly complex interactions drive technology development behind SEs. Web SEs exist because no one has the time to manually type in URLs for the URLs of many websites to learn something specific. By using a web search engine (SE), the user does not have



This work is licensed under a Creative Commons Attribution 3.0 License.

to memorize a growing array of URL names. An online SE allows the user to input a query for which the user needs information and then offers all possible results tied to that query [1].

Lots of people use the web every day to look for particular data, which is an essential part of the endless digital world. Instead of actually putting in the URL, almost half of all website visitors find their way there through a SE [2]. Recent progress in Natural Language Understanding (NLU) is driving fast-paced advances in Information Retrieval (IR), largely owed to fine-tuning deep language models (LMs) for document ranking. While remarkably effective, the ranking models based on these LMs increase computational cost by orders of magnitude over prior approaches, particularly as they must feed each query-document pair through a massive neural network to compute a single relevance score [3].

Web SEs are tools that users use to search through the vast and growing repository of information available on the internet. Since most digital IR involves SEs, learning about how they work is an important part of user education as well as for developers. Even traditional performance metrics (advanced precision, recall and response time) are often used to describe the accuracy and relevancy of SE results. These metrics can be useful, but are typically system-based and do not consider user effort to find relevant information. Because the original SE, Veronica, is a text SE, Web SEs also use file transfer protocol (FTP) files [4].

Nowadays the first SE that was used for searching FTP files in the 1990's was Archie and Veronica was the first based on texts sort machine to be released [5].

By highlighting the user's experience more especially, the amount of time and activities needed to get pertinent results the idea of user-effort-sensitive evaluation criteria offers a fresh viewpoint. By evaluating the user's dedication, we may gain a better understanding of how effective a SE is in real-life circumstances, meeting the demands and difficulties that users have while looking for facts online. The user experience is a key factor in evaluating a SE's quality in a world where time is a valuable resource. A Search Engine Optimized (SEO)-designed website has

a greater likelihood of appearing higher on a search engine results page (SERP). The SE with the greatest user base is Google, however there are other SEs including Bing, Yahoo, DuckDuckGo, and others use their own unique web page crawling algorithms to produce the top results for searches [6].

The SEs, such as Google, have become increasingly crucial in how people locate and access information in their daily lives. They serve as a facilitating structure, influencing how information is filtered and presented while also making it accessible to the wider public [7]. There are several popular such SEs. Google, Bing, Yahoo, and others are popular SEs, with Google accounting for more than 80% of all queries.

Most consumers do not go beyond the very first webpage of searches, therefore Internet SE Ads and SEOs are more effective ways to improve the SE's web page ranking. Natural Language Processing (NLP) methods have greatly enhanced SEs in recent years, transforming the area of information retrieving. Additionally, although typical SEs may occasionally interpret keywords, they are still unable to understand the significance of the user concern [8]. These engines use complex algorithms and artificial intelligence (AI) to improve search accuracy; Google's PageRank and RankBrain, for example, establish high standards for ranking and relevancy.

Ecosia has established out a distinct market by concentrating on the environmental effect and sponsoring worldwide reforestation programs with ad a profit. Furthermore, specialized SEs like WolframAlpha cater to specific needs, such as scientific or mathematical inquiries, by providing precise, data-driven answers rather than conventional web pages. This broad ecology of SEs reflects users' expanding expectations, as they seek platforms that correspond with their beliefs, ethical goals, and unique informational requirements. As a result, user-effort-sensitive evaluation metrics provide a more thorough and accurate picture of how SEs actually function from the viewpoint of the user.

Leading SEs such as Google, Bing, and Yahoo dominate the market by leveraging complex algorithms like Google's PageRank, Bing's BM25, and Yahoo's integration with Bing's infrastructure, to provide

personalized and highly relevant results. These algorithms utilize learning algorithms and advanced automated language processing techniques to interpret complex queries and rank content by authority and relevance. In contrast, privacy-oriented SEs like DuckDuckGo and Qwant are run on non-personalized ranking algorithms with no tracking of users ensuring privacy. Another example is Ecosia, which has used Bing's search algorithms with the objective of partnering with initiatives to protect Earth's resources by reforesting old forests with search ad revenues.

Products dedicated to scientific computation are used exclusively by WolframAlpha, a highly computational SE with answers based on facts and not website rankings. This research makes an exhaustive comparison of SEs. It questions how their algorithmic basis, information retrieval methods, and security measures for privacy match their actual users' needs. The use of Boolean operators such as AND, OR, and NOT is another useful way to further break down search results.

SE have evolved from mere indexing tools to intelligent systems that leverage algorithms like PageRank and bidirectional encoder representations from transformers (BERT) to better contextualize search as well as improve user experience. Since this section is about SEs, it contains lots of studies focusing on improvement to individual engines' performance, algorithms, or SEO best practices. But on the other hand, only small work has been done of comparative evaluation of two or more SEs and how good they are at meeting different user needs. Aiming at filling this gap, the next section reports on related work that covers various aspects of SE technologies, ranking algorithms and optimization techniques and forms the basis of the comparison which we will do in this article.

2 Related Work

Such SEs are divided into crawler-based, human-powered, or hybrid SEs. Crawler-based SEs (in this case Google) use computerized spiders to file web pages and human-powered SEs use manual submission and indexing techniques (e. g. as in Yahoo's older

directory). Ask. com, and many other hybrid SEs use a combination of the above to increase accuracy of their results.

Various features such as multimedia support, privacy preferences, personalization, and audio or visual search are present in modern SEs. For example, Google uses artificial intelligence algorithms (e. g. BERT) to understand queries with context, while privacy-oriented SEs like DuckDuckGo help protect privacy. When evaluating SEs, precision (precision of the search result) and recall (recall of a result), are crucial elements. Modern SEs have been greatly improved by innovations such as natural language searches and real-time indexing, whereas traditional SEs previously struggled to achieve high precision and recall (because existing algorithms are static).

But there are issues with SEs such as flawed ranking algorithms, fears about data privacy, and difficulty balancing recall and precision, especially in exploratory searches. The development of SE features and algorithms has greatly improved their usefulness, and they remain important resources for retrieving information even with these issues.

2.1 Classification of Search Engines

SEs are classed according to their operational methods and priority regions, as follows:

1. Crawler-based SE.
2. Human-powered SE.
3. Hybrid SE.

Figure 1 illustrates the three main types of SEs: crawler-based, human-powered, and hybrid systems. This classification illustrates the evolution of search technology, where crawler-based SEs dominate today's web, while hybrid models combine automation with human input to refine results.

The three primary phases of a crawler-based SE are as follows:

1. **Crawling:** It involves the use of a crawler, also known as a spider, searching the global web and visiting appropriate websites in order to obtain possibly corresponding search terms that are part of the query being searched. The crawler

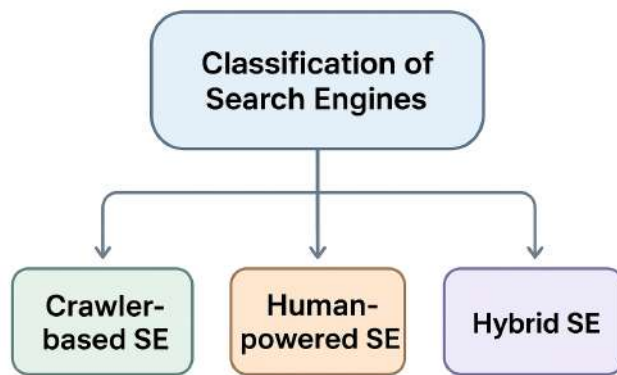


Figure 1. Three main categories of SEs

then gathers the material (photos, hyperlinks, as well as other information) of each of these websites and creates a copy of them for further analysis.

2. **Indexing and Relevance Estimation:** In this phase, the SE uses the retrieved website content to build a directory and gives a rating based on different web page ranking algorithms.
3. **Retrieving the outcome:** As the title implies, it basically returns the results to the user in the browser in the order of most significant relevance. Google, Bing, Yahoo, and others are examples of crawler-based SEs.

This involves the web page creator sending the directory with a list of potential terms together with their type, which is later thoroughly examined and arranged into the relevant category [5]. Hybrid SE makes use of both the crawler-based and manual indexing to produce their search result. As these needs have expanded, so has the variety of SEs, each employing unique algorithms to cater to specific types of queries and user priorities [9].

With the explosion of digitized and online information, tools and processes for accessing them have had to evolve rapidly, resulting in many advancements in a short period of time when put in the context of the history of human information production, storage, and dissemination. While the goals of these advancements for IA tools and technologies have been around

retrieval, filtering, and accessibility of information, the recent focus has shifted more toward what might be considered the generation of information [10].

[11] suggested Dense Passage Retrieval (DPR), a dual-encoder model that learns dense representations for both questions and passages, to enhance open-domain question answering by displacing conventional sparse techniques like BM25. The model was tested on five QA benchmarks using positive and negative passage pairs with in-batch negatives for training, and it consistently exceeded BM25 by 9-19% in top-20 retrieval accuracy. Their findings showed that DPR not only boosted end-to-end QA performance but also achieved high efficiency through Facebook AI Similarity Search (FAISS) indexing, proving dense retrieval to be a practical and effective alternative to keyword-based search.

[12] conducted a global survey of 386 software professionals to explore the use and challenges of software testing techniques in industry. They found that unit and system testing are widely used, while automated testing faces issues like tool usability and high maintenance. The study highlights a gap between academic testing methods and industry practices, calling for more practical, industry-aligned research to improve real-world testing effectiveness.

[13] featured the ground-breaking PageRank algorithm from Google, an important advancement in SE innovation. Using links as a measure of authenticity and validity, they developed an algorithm designed to rank web pages according to their link structure. By analyzing the quantity and quality of backlinks to a page, PageRank improved the relevance of search results significantly. Through experiments, they demonstrated that PageRank could accurately rank pages in a way that corresponded with user expectations, setting the stage for Google's rapid growth and establishing PageRank as a core standard for web search relevance.

Figure 2 shows the PageRank algorithm as a network of nodes and links, where authority is distributed through interconnections. SEO was analyzed in the context of the entire online ad mix by them. Using case studies and SEO analytics, he showed that SEO

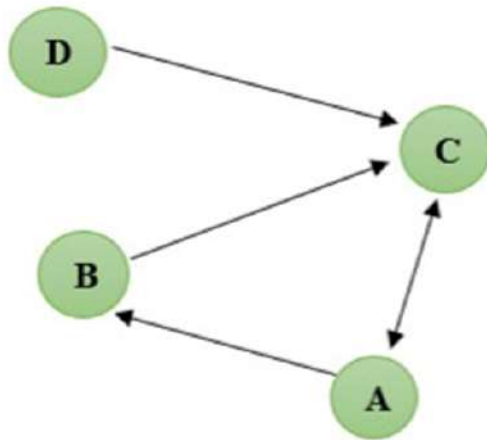


Figure 2. Page Rank Concept (Nodes and Links).

works best when integrated with content strategy, user experience, and social media, rather than being treated as a stand-alone tactic. The study found that SEO delivers long-term, cost-effective visibility, especially when aligned with broader marketing goals [14].

[15] they focused on SEO for Single-Page Applications (SPAs) as opposed to Multi-Page Applications (MPAs). They constructed three prototypes of a travel website on Flask (MPA), React (SPA), and Next.js (hybrid), and evaluated them with SEO tools. As per their results, Next.js supplemented with client-side interactivity and React-only SPAs on the server-side fared better in indexing than SPAs with only server-side rendering. They concluded that hybrid approaches can make SPAs as SEO friendly as MPAs.

[16] evaluated the difficulty in finding appropriate information about SEs as a result of the rapid growth of online data. They pointed out that consumers often get a lot of results with much of it being irrelevant to them and this makes it hard for them to know which pages are the most relevant. To address this, the authors explained that SEO can be used to improve a site's rankings by employing some of the following strategies. The study examined several different SEO methods and concluded that some of the best practices were on-page optimization as well

as off-page optimization. Techniques on-page (using relevant keywords and providing quality content) and offline (building backlinks and participating in online communities) were identified as important for optimizing a website's visibility and accessibility. The study also examined different algorithms used to rank web pages and found that those allow SEs to return more pertinent results to users. As part of future work, the authors plan to use a better, more accurate SEO system to further improve webpage ranking in SE results.

[4] presents a review of page ranking algorithms (PCAs) used in SEO to address the problems caused by growing number of online pages as well as search results with too many irrelevant or incomplete information. The main research topics covered are Genetic Algorithm, Particle Swarm Optimization, and distance ranked algorithm for ranking the sites. The distance ranked algorithm is also discussed because it ranked the sites according to their proximity to the website. The results show that Page Ranking algorithms (SP) can refine search results on more than one factor such as user input, domain-specific data and vector space models, namely. We conclude that while the approaches are improved in SEO it is necessary further researches to further improve page ranking and reduce the computing load on the SEs.

[5] defines the importance of SEs like Google, Bing, MSN and Yahoo in searching huge amounts of online content. SEs use algorithms to categorize and rank websites, and it also emphasizes the importance of SEO for long-term optimization, thus optimizing a website over time, increasing its visibility, traffic, and online presence. Different free and paid SEO tools are discussed for keyword research, optimization performance evaluation, and so on. Most previous studies used only a specific algorithm or one platform to evaluate the SEs; however, this study suggests that newer, more user-oriented evaluations of SEs are needed to better reflect the results and behavior of users searching. To address this gap, the following section describes the methodology used in this study to evaluate ten mainly SEs using different query types: The evaluations are based on precision and recall. By

measuring SE accuracy against precision, we aim to tell which SE delivers rich and relevant results.

3 Methodology

In some geographic information spaces, expert sites are more informative than hub sites because they are more specific. SEs are methods of reaching content sites and not sources of information [17]. The ranking algorithms used to evaluate SE performance are two related user-effort-sensitive metrics: the number of relevant results obtained and the time taken to get relevant results. Ten SEs chosen in this study were selected for various reasons such as visual adaptability, privacy focus, and specialized search capabilities. A set of queries with the highest information demands from various users (fact-based, exploratory, research-focused, transactional, navigational, and multimedia) were selected.

The research design in figure 3 demonstrates how queries were categorized into navigational, transactional, informational, and exploratory types and tested across ten SEs. This structured process ensured that relevancy, speed, and user-effort were consistently evaluated. These categories follow [18], who identified them as the most common user intents were selected. We then statistically evaluated the data to explore how SEs address different information demands from various users to reduce user effort. By analyzing the function and functioning of SEs, we aim to learn how SEs meet information needs from different stakeholders with the greatest amounts of ease. The analysis method uses a variety of approaches to examine SEs' operating mechanisms, algorithms, and user-interaction mechanisms in order to gain insights into how they serve different information needs.

The technical base of a SE is its architecture, which consists of web crawlers, indexing systems, and ranking algorithms. Web crawlers, also commonly called spiders, are computer programs that continuously crawl the web in order to collect information about websites. Rank pages based on certain criteria, such as how often they are updated, and the strength of backlinks from other pages. These SEs also use protocols, such as robots.txt, to protect website

privacy settings. An inverted index that links terms with their locations on multiple pages stores the collected data and makes it easier for fast and effective searches. Various investigations have been conducted for improving search methods to help users obtain accurate and useful information. Some studies have improved the search results or efficiency by analyzing the purpose of the Internet search [19].

3.1 Selection of Search Engines

Because there are so many SEs available on the Internet, it can be challenging for someone to decide which ones would best meet their information needs. Building a metasearch engine on top of the SEs is a popular remedy. The metasearch engine forwards user queries to the underlying SEs that are most likely to provide the requested documents for the query [20]. We compare a number of SEs, including Google, Bing, Yahoo, DuckDuckGo, Qwant, Ecosia, Swisscows, Ask.com, WolframAlpha, and StartPage. By include both well-known and obscure SEs, the target was to offer a thorough grasp of how different search algorithms work. The results aim to encourage the adoption of the most efficient SEs by helping consumers comprehend the benefits of each one. While this study solely examines online databases, some SEs also index content from chat groups and Gopher networks. We did not include SEs such as CUSI that merely gather information without offering new insights [6].

3.2 Research Design

Figure 3 The flowchart demonstrates how queries were categorized and tested across SEs to evaluate relevancy, speed, and user-effort performance.

3.3 Identifying the Research Scope

A key part of any study is the development of the research scope which defines how the research criteria, objectives, and content to examine should be addressed. For this study, ten popular SEs and major information retrieval systems, such as Google, Bing, Yahoo, DuckDuckGo, Swisscows, Qwant, Ecosia, Ask.com, WolframAlpha, and StartPage, were evaluated through an evaluation of the functionality of these SEs. For this study, a query was selected to represent

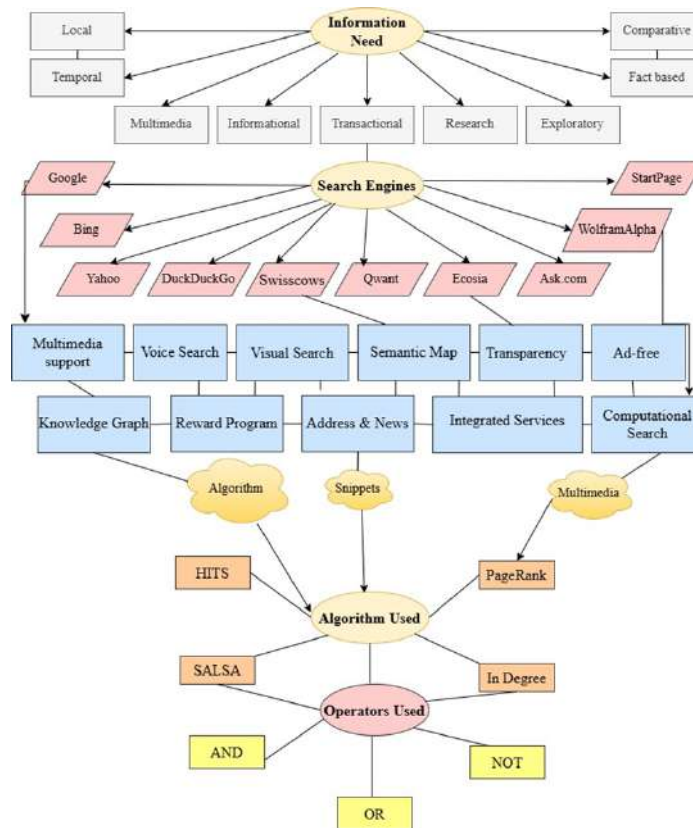


Figure 3. Research design outlining the comparative evaluation of ten SEs using user-effort-sensitive measures.

a wide range of information needs typically expressed by users in web searches. The selection process was designed to ensure that the queries included certain types of search types: navigational, informational, transactional, research-oriented, exploratory, fact-based, and multimedia. The sets of queries were chosen based on frequency of use in the past year so that they are representative of common and current user interest. Informational, transactional, and navigational search intents influence a SE advertising campaign’s performance. Additionally, SE advertisement results with an organic result placed on the same SE results page where an organic result first appeared is considered which provides for a balanced approach to research topics related to SE performance [18]. The specific purpose of the research scope is to understand how these SEs address the different information needs of users, which can be classified as the following:

Navigational Needs: These questions are specifi-

cally aimed at finding a specific website or page. Say a user asking for “Twitter login” and expects to be directed to the login page for Twitter. The research looks at SEs ability to prioritize queries and provide pertinent information quickly.

Informational Needs: Users with information and explanation needs regarding a specific topic of interest (e.g. “what is climate change?”) are considered informative. The study investigated how well SEs fulfilled these kinds of queries and how accurately they delivered them.

Transactional Needs: These are action-oriented questions, based around tasks like downloading, buying or signing up. A good example of a transactional query is “Buy iPhone 15”. The study evaluates how well SEs can offer relevant items/links, including the use of ads, and how best to prioritize e-commerce platforms.

Research-Oriented Needs: Such queries require deeper, more extensive resources; such as scholarly articles or scholarly research, or expert opinion. For

example, someone might want to find “Impact of artificial intelligence on healthcare”, for example, and they need a variety of trustworthy sources for that. Such demands can be met by SEs like Google Scholar or WolframAlpha (Google offers scholarly literature that are highly credible).

Exploratory Needs: This type of question requires more in-depth research when people are unknown to a subject and wish to learn more about it. An example is “Blockchain applications in finance.” SEs’ ability to assist users by providing relevant queries, recommendations, or thorough summaries is measured by the assessment.

Fact-Based Needs: Examples such as “Pakistan’s current population” and “Today’s Bitcoin price” center on gathering particular data. With snippets or structured data frequently used to display important information, the study assesses how well SEs deliver precise, dependable, and timely results.

Multimedia Needs: Given the increasing demand for non-textual content, the study also assesses SEs’ multimedia capabilities. Searches like “Solar energy videos” or “Images of Amazon forest” are used to evaluate how well SEs retrieve relevant, high-quality visual or audio content.

Additionally, given the growing concerns about data gathering and user targeting, the study looks at how SEs handle privacy-related issues. Privacy-focused SEs such as DuckDuckGo, Swisscows, and StartPage are compared to popular engines such as Google and Bing to determine the best mix of user customization and anonymity. Knowing into how SEs efficiently arrange and access data can be gained from analyzes of the technical structure of the SEs, including their crawling strategies, indexing processes, and evaluating algorithms. For example, DuckDuckGo maintains privacy by not recording your search query requests, while Google’s PageRank algorithm prioritizes links above all other metrics. By precisely defining search terms, SEs and review criteria, the depth ensures a thorough and objective analysis. By helping consumers choose the SEs that suit their needs best, this process sets the stage for insightful results.

3.4 Significance in Search Engine Evaluation

Additionally, given the growing concerns about data gathering and user targeting, the study looks at how SEs handle privacy-related issues. Privacy-focused SEs such as DuckDuckGo, Swisscows, and StartPage are compared to popular engines such as Google and Bing to determine the best mix of user customization and anonymity. Knowing into how SEs efficiently arrange and access data can be gained from analyzes of the technical structure of the SEs, including their crawling strategies, indexing processes, and evaluating algorithms. For example, DuckDuckGo maintains privacy by not recording your search query requests, while Google’s PageRank algorithm prioritizes links above all other metrics. By precisely defining search terms, SEs and review criteria, the depth ensures a thorough and objective analysis. By helping consumers choose the SEs that suit their needs best, this process sets the stage for insightful results.

3.5 Indexing

To efficiently retrieve relevant data from large datasets indexing is an important process in SEs. The content of online pages is analyzed and stored in an index (a structured database that stores terms and keywords along with the documents they occur in after being crawled by bots or spiders). Because of that SE’s is able to deliver results to clients in a fast and efficient manner. An inverted index links keywords to the text in which they appear instead of arranging pages in chronological order.

Example: Subsequent to crawling the web pages, their content is analyzed and the index is created such that documents are stored in a structured manner with keywords and phrases as links to the corresponding documents. This process first analyzes the document to determine if it contains the words “climate”, “change” and “effects”. If the document does, it then proceeds to create a mapping for it such as:

climate—Document 1, Document 3, Document 5
change—Document 1, Document 2, Document 4
effects—Document 2, Document 3, Document 6

This allows for a more SE friendly retrieval process

which in turn enables SEs to deliver results in a more timely and efficient manner.

3.6 Data Collection Process

Each SE was evaluated by a set of volunteers who ran searches using prepared questions. To prevent any inconsistency and eliminate any factor that might have influenced the results, the tests were conducted in the same environments with the same browser and hardware settings. All SEs were queried, and the time and number of clicks required to obtain suitable results were recorded. To avoid any error, automatic monitoring systems and human evaluators both measured this data. Number of clicks and time to a relevant result were used to rank the results for each SE. The top-performing SEs in user effort were then identified by comparing these rankings with other SEs.

3.7 Analysis

The collected data was analyzed using basic statistical techniques to understand the variability in the number of relevant results across and response time of SEs. The results were further analyzed in terms of:

- The efficiency of each SE for diverse categories of queries.
- How the performance of SEs with varying focus areas differed from more general-purpose engines like Google and Bing.
- The relationship between efficiency and relevance.

3.8 Relevance of Data Retrieval and Indexing

SEs use sophisticated indexing strategies to produce results quickly. Regardless of the complexity of the query, fast indexing guarantees that it will be answered in milliseconds. Even quicker information retrieval is made possible by features like caching. Instead of only matching keywords, more advanced techniques like semantic indexing enable SEs to understand the context and purpose of a query. For example, a search for "renewable energy" may return results for solar, wind, and geothermal energy, even if those terms are not explicitly included in the indexed text. As the web grows, SEs must scale their retrieval

and indexing capabilities to handle the increasing volume of data. Distributed systems, such as Google's Bigtable, enable parallel indexing of massive datasets. Privacy-focused SEs like DuckDuckGo and StartPage implement minimal indexing to avoid tracking user behavior, relying on external engines for their data. The methodology described above provides a structured framework for evaluating each SE's effectiveness in addressing various information needs. By applying precision and recall metrics to the top ten results of different query types, we obtain a detailed performance profile of each engine. The following section presents and interprets the results of this evaluation, highlighting the strengths and limitations observed in the output of each SE.

4 Research Results

From analyzing the performance of the ten SEs: Google, Bing, Yahoo, DuckDuckGo, Swiss-cows, Qwant, Ecosia, Ask.com, WolframAlpha, and StartPage, we observed that there is high volatility across all of them regarding information needs and retrieval usability. The needs were defined as navigational, informational, transactional, research, exploratory, fact-based, and multimedia. Each SE's performance evaluation report is based on important issues such as relevancy, accuracy, search speed, privacy, and multimedia features which pointed out their particular strengths and weaknesses.

These results are determined by the sources of data and the modes of information retrieval employed by each SE. Understanding the user's needs, whether transactional, exploratory, informative, or navigational, is the initial step needed to extracting results from a SE. The first ten pages are retrieved from the SE based on a formulated and executed search query stemming from this requirement. After the evaluation of relevancy of these pages to the query, the efficiency and reliability of the SE are evaluated through precision and recall approaches.

Emphasis on relevance results ratio illustrates precision, which is the number of relevant results retrieved out of all results presented. Conversely, recall computes the corresponding retrieved results

ratio out of desired relevant results that ensures the completeness of the search. Applying a blend of these metrics gives insight into the efficacy and precision of the SE in meeting the need of the user.

The scientific approach makes it possible to compare various SEs in detail based on how they strike a balance between precision and recall, thereby how well they retrieve relevant and complete results. Precision and Recall are simple measures used in the measurement of functionality of SEs, namely significance and completeness of search results. Desirable online search assessment features are listed in [23] features 1-7 and [24] features 8-1.

1. Searches should be prompted by actual information of demands for online users.
2. The main searcher's information requirement should be completely and contextually recorded and sent to the search intermediary if one is used.
3. To estimate the functioning of SEs, a significant number of searches are required.
4. Considerations should be given to majority of major SEs.
5. The best way to utilize each SE's unique capabilities should be used i.e., the queries supplied to SEs may change.
6. According to [24], individuals seeking information should rely on unbiased judges to determine what is relevant.
7. To analyze SE performance, utilize statistical tests, known retrieval measures, and eliminate bias towards SEs throughout studies.
8. The scope of information demands from which conclusions are to be drawn should be reflected in the search subjects.
9. Judges results based on the type of question provided e.g., online answers.
10. Documents should be presented in web browser-like format, including visible images and links.
11. Dead links should be considered invalid replies.

4.1 Precision in Search Results

The accuracy of the results that are returned and the proportion of results that are relevant to the query are

the primary concerns of precision. High accuracy indicates that the user may benefit from the majority of the documents that are retrieved.

Formula:

$$\text{Precision} = \frac{\text{Number of Relevant Results Retrieved}}{\text{Total Number of Results Retrieved}}$$

4.2 Recall in Search Results

Recall quantifies how many of the relevant documents that were available were retrieved by the SE, indicating how comprehensive the search results were. Fewer pertinent papers will be overlooked when memory is high.

Formula:

$$\text{Recall} = \frac{\text{Number of Relevant Results Retrieved}}{\text{Total Number of Relevant Results Available}}$$

4.2.1 Relevance Points of Top 10 Pages of Google

$$\text{Precision} = \frac{25}{30} = 0.83 \approx 83\%$$

$$\text{Recall} = \frac{25}{50} = 0.50 = 50\%$$

Result	Relevance	Points
Result 1	Highly Relevant	3
Result 2	Highly Relevant	3
Result 3	Relevant	2
Result 4	Highly Relevant	3
Result 5	Highly Relevant	3
Result 6	Relevant	2
Result 7	Relevant	2
Result 8	Highly Relevant	3
Result 9	Seem Relevant	1
Result 10	Highly Relevant	3
Total		25/30

Accuracy for Google: $25/30 = 0.83$ or 83%

4.2.2 Relevance Points of Top 10 pages of Bing

$$\text{Precision} = \frac{20}{30} = 0.66 \approx 66\%$$

$$\text{Recall} = \frac{20}{50} = 0.40 = 40\%$$

Accuracy for Bing: $20/30 = 0.66$ or 66%

Result	Relevance	Score
Result 1	Highly Relevant	3
Result 2	Highly Relevant	3
Result 3	Relevant	2
Result 4	Relevant	2
Result 5	Seem Relevant	1
Result 6	Relevant	2
Result 7	Relevant	2
Result 8	Seem Relevant	1
Result 9	Highly Relevant	3
Result 10	Seem Relevant	1
Total		20/30

4.2.3 Relevance Points of Top 10 Pages of Yahoo

Result	Relevance	Score
Result 1	Relevant	2
Result 2	Relevant	2
Result 3	Seem Relevant	1
Result 4	Irrelevant	0
Result 5	Relevant	2
Result 6	Seem Relevant	1
Result 7	Highly Relevant	3
Result 8	Relevant	2
Result 9	Seem Relevant	1
Result 10	Irrelevant	0
Total		14/30

$$\text{Precision} = \frac{14}{30} = 0.46 = 46\%$$

$$\text{Recall} = \frac{14}{50} = 0.28 = 28\%$$

Accuracy for Yahoo: 14/30 = 0.46 or 46%

4.2.4 Relevance Points of Top 10 Pages of DuckDuckGo

Result	Relevance	Score
Result 1	Highly Relevant	3
Result 2	Relevant	2
Result 3	Seem Relevant	1
Result 4	Relevant	2
Result 5	Seem Relevant	1
Result 6	Relevant	2
Result 7	Relevant	2
Result 8	Seem Relevant	1
Result 9	Highly Relevant	3
Result 10	Relevant	2
Total		19/30

$$\text{Precision} = \frac{19}{30} = 0.63 = 63\%$$

$$\text{Recall} = \frac{19}{50} = 0.38 = 38\%$$

Accuracy for DuckDuckGo: 19/30 = 0.63 or 63%

4.2.5 Relevance Points of Top 10 Pages of Swisscows

Result	Relevance	Score
Result 1	Relevant	2
Result 2	Seem Relevant	1
Result 3	Irrelevant	0
Result 4	Relevant	2
Result 5	Seem Relevant	1
Result 6	Seem Relevant	1
Result 7	Relevant	2
Result 8	Irrelevant	0
Result 9	Seem Relevant	1
Result 10	Highly Relevant	3
Total		13/30

$$\text{Precision} = \frac{13}{30} = 0.43 = 43\%$$

$$\text{Recall} = \frac{13}{50} = 0.26 = 26\%$$

Accuracy for Swisscows: 13/30 = 0.43 or 43%

4.2.6 Relevance Points of Top 10 Pages of Qwant

Result	Relevance	Score
Result 1	Relevant	2
Result 2	Highly Relevant	3
Result 3	Seem Relevant	1
Result 4	Relevant	2
Result 5	Seem Relevant	1
Result 6	Seem Relevant	1
Result 7	Highly Relevant	3
Result 8	Relevant	2
Result 9	Relevant	2
Result 10	Seem Relevant	1
Total		18/30

$$\text{Precision} = \frac{18}{30} = 0.60 = 60\%$$

$$\text{Recall} = \frac{18}{50} = 0.36 = 36\%$$

Accuracy for Qwant: 18/30 = 0.6 or 60%

Result	Relevance	Score
Result 1	Highly Relevant	3
Result 2	Highly Relevant	3
Result 3	Highly Relevant	3
Result 4	Relevant	2
Result 5	Relevant	2
Result 6	Relevant	2
Result 7	Seem Relevant	1
Result 8	Relevant	2
Result 9	Seem Relevant	1
Result 10	Highly Relevant	3
Total		22/30

Result	Relevance	Score
Result 1	Highly Relevant	3
Result 2	Highly Relevant	3
Result 3	Highly Relevant	3
Result 4	Seem Relevant	1
Result 5	Seem Relevant	1
Result 6	Relevant	2
Result 7	Irrelevant	0
Result 8	Seem Relevant	1
Result 9	Highly Relevant	3
Result 10	Relevant	2
Total		19/30

4.2.7 Relevance Points of Top 10 Pages of Ecosia

$$\text{Precision} = \frac{22}{30} = 0.73 = 73\%$$

$$\text{Recall} = \frac{22}{50} = 0.44 = 44\%$$

Accuracy for Ecosia: $22/30 = 0.73$ or 73%

4.2.8 Relevance Points of Top 10 Pages of Ask.com

Result	Relevance	Score
Result 1	Seem Relevant	1
Result 2	Irrelevant	0
Result 3	Irrelevant	0
Result 4	Seem Relevant	1
Result 5	Seem Relevant	1
Result 6	Relevant	2
Result 7	Irrelevant	0
Result 8	Seem Relevant	1
Result 9	Irrelevant	0
Result 10	Relevant	2
Total		8/30

$$\text{Precision} = \frac{8}{30} = 0.26 = 26\%$$

$$\text{Recall} = \frac{8}{50} = 0.16 = 16\%$$

Accuracy for Ask.com: $8/30 = 0.26$ or 26%

4.2.9 Relevance Points of Top 10 Pages of WolframAlpha

$$\text{Precision} = \frac{19}{30} = 0.63 = 63\%$$

$$\text{Recall} = \frac{19}{50} = 0.38 = 38\%$$

Accuracy for WolframAlpha: $19/30 = 0.63$ or 63%

4.2.10 Relevance Points of Top 10 Pages of StartPage

Result	Relevance	Score
Result 1	Highly Relevant	3
Result 2	Highly Relevant	3
Result 3	Relevant	3
Result 4	Highly Relevant	1
Result 5	Relevant	1
Result 6	Relevant	2
Result 7	Highly Relevant	0
Result 8	Seem Relevant	1
Result 9	Relevant	3
Result 10	Highly Relevant	2
Total		24/30

$$\text{Precision} = \frac{24}{30} = 0.80 = 80\%$$

$$\text{Recall} = \frac{24}{50} = 0.48 = 48\%$$

Accuracy for StartPage: $24/30 = 0.80$ or 80%

Table 1. Percentage of relevant results retrieved by each SE, showing Google’s dominance in accuracy

Search Engines	Relevancy	Percentage
Google	25	83%
StartPage	24	80%
Ecosia	22	73%
Bing	20	66%
DuckDuckGo	19	63%
WolframAlpha	19	63%
Qwant	18	60%
Yahoo	14	46%
Ask.com	13	43%
Swisscows	8	26%

Table 1 Google achieved the highest relevancy rate (83%), followed by StartPage (80%), while Ask.com returned the lowest (26%), confirming strong variations in effectiveness.

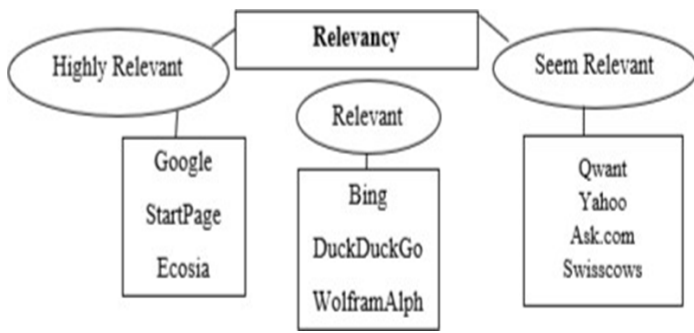


Figure 4. Comparative relevancy scores of ten SEs, highlighting Google and StartPage as top performers.



Figure 5. Graphical comparison of SEs relevancy showing clear gaps between high-performing and low-performing engines

Figure 4, 5 The graph highlights the performance gap, with Google and StartPage consistently outperforming privacy-oriented engines like Swisscows and Ask.com.

4.3 Comparison of Results with Information Needs

By using powerful algorithms like PageRank and domain authority, Google and Bing were able to provide accurate and reliable connections to their intended locations. Due to their priority for user confidentiality, privacy-focused engines such as DuckDuckGo and StartPage fared well but provided somewhat less customization. On the other hand, Yahoo and Ask.com sometimes delivered irrelevant or outdated results, indicating a less effective focus on navigational queries. Google, Bing, and Qwant provided comprehensive and reliable information, enhanced with snippets and knowledge panels, which

improved user experience. Although in the case of tackling complex issues Swisscows is recognized for family friendly browsing, it offered limited scope data which impacted its precision.

Google and Bing during transactional search queries performed exceptionally well with regards to displaying shopping ads and product listings in an orderly manner. Ecosia also performed reasonably well, although their products pertained to sustainability. DuckDuckGo and Startpage due to their privacy regulations put less emphasis on commercial links making them unsuitable for transactional inquiries.

Google Scholar and Bing offered comprehensive scholarly research papers proficiently. Despite being great at supplying computational data, WolframAlpha lacked the wider breadth required for general study. Less successful were Yahoo and Ask.com, which frequently returned irrelevant or out-of-date material. For broad, exploratory questions, SEs such as Google and DuckDuckGo enabled plenty of different viewpoints and related searches. Additionally, Qwant did well by offering impartial, well-rounded findings.

However, Swisscows and Ecosia, while dependable in specific areas, tended to return more limited results, which restricted their use for wider searches. Particularly useful for analytical and scholarly investigations, WolframAlpha distinguished itself with its accuracy and data-driven methodology. Google and Bing excelled at utilizing organized data to produce broad highlighted snippets, enhancing search outcomes. However, Yahoo and Ask.com battled with reliability, often offering limited or less trusted data.

In multifaceted search, Google and Bing both distinguished themselves by providing excellent photos, videos, and animations.

4.4 Comparison of Results with Information Sources

The consistency and variety of SEs data sources have a significant impact on how well they perform overall. As shown in Table 2, with the help of strong crawling and indexing systems, Google and Bing provided the most thorough and reliable results, guaranteeing a more efficient search experience. DuckDuckGo and Startpage relied on Bing and Google’s indexing,

Table 2. Accuracy rates of SEs based on user relevance judgments.

SEs	Navigational	Informational	Transactional	Research	Multimedia
Google	YES	YES	YES	YES	YES
Bing	YES	YES	YES	YES	YES
Yahoo	YES	YES	YES	YES	YES
DuckDuckGo	YES	YES	Limited	Limited	YES
Swisscows	YES	YES	Limited	Limited	YES
Qwant	YES	YES	YES	YES	YES
Ecosia	YES	YES	YES	YES	Limited
Ask.com	YES	YES	NO	YES	Limited
WolframAlpha	NO	NO	YES	YES	NO
StartPage	YES	YES	YES	YES	YES

respectively, ensuring relevance but limiting query depth. WolframAlpha's reliance on curated datasets ensured unparalleled accuracy for fact-based and computational queries, but its results lacked the general-purpose versatility of Google or Bing. Swisscows and Qwant, while privacy-focused, indexed fewer sources, leading to limitations in breadth and diversity. Ecosia's integration of Bing's infrastructure, combined with eco-focused content prioritization, offered a unique, sustainability-driven perspective.

4.5 Comparison of Results with Literature Review

The results of this study, which assess ten widely used SEs using precision and recall metrics across various user intents, offer a broader and more user-centered view compared to earlier literature. Prior research has largely focused on specific aspects, such as Google's PageRank algorithm for link-based ranking [14], the ethical and unethical use of SEO techniques [21], and user engagement factors like page speed, bounce rate, and session duration [22]. While these works contributed to improving individual SE's performance, they often lacked a direct, comparative analysis from the user's perspective. In contrast, this study evaluates real-time performance and relevance across multiple engines Google, Bing, Yahoo, DuckDuckGo, Swisscows, Qwant, Ecosia, Ask.com, WolframAlpha, and StartPage based on their ability to fulfill different types of queries: navigational, informational, transactional, exploratory, and more. Unlike [17] and

[4], which investigated ranking algorithms and optimization methods such as Genetic Algorithms

or Ontology-Based Ranking, this research emphasizes overall user satisfaction through relevance and diversity of results. Privacy-conscious engines like DuckDuckGo also showed notable performance in maintaining confidentiality without compromising relevance, aligning with concerns highlighted in [6]. This comparative approach fills a gap in existing research by evaluating multiple engines side-by-side, offering practical insights for everyday users rather than focusing on algorithmic or developer-centric improvements. It adds to the current literature by translating technical performance into user impact, bridging the gap between technical innovation and user experience.

Table 3. Average time taken by SEs to return relevant results, reflecting user-effort sensitivity

SEs	Response Time
Google	Under 1 sec
Bing	1-2 sec
Yahoo	1-2 sec
DuckDuckGo	Under 2 sec
Swisscows	2-3 sec
Qwant	2-4 sec
Ecosia	2-3 sec
Ask.com	2-4 sec
WolframAlpha	Under 2 sec
StartPage	1-3 sec

Table 3 Google and StartPage returned relevant results in the shortest time, while Ask.com required considerably longer, demonstrating efficiency gaps in query handling.

4.6 Search Engine Features

A SE is a software program that helps people find information on the web by using keywords or phrases.

Table 4. Distinctive features of major SEs, including Bang search, Knowledge Graph, and AI-driven enhancements

Features	Description
Algorithm	Google uses sophisticated algorithms, including as PageRank, to rank results according to user activity, authority, and relevancy.
Snippets	In the search results, Google offers rich snippets, such as definitions, photos, videos, and fast replies.
Knowledge Graph	Google displays pertinent data related to the search query in a special panel located on the page's right side.
Voice Search	Google enables speech commands for users to search.
Personalization	Google search outcomes are tailored according on surfing history, location, and user activity.
Search Filters	Google allows for picture, news, video, book, map, commerce, and more filtering options.
Multimedia Support	Google offers a thorough visual search for news articles, videos, and photos.
Snippets	Bing provides comprehensive snippets with connections to pertinent media, articles, and definitions.
Rewards Program	Bing users may earn points by utilizing Bing's special rewards program, which can then be redeemed for services and gift cards.
Multimedia Search	Bing's excellent picture and video search is well-known.
Search Filters	Bing lets you filter by maps, photos, videos, and news.
Visual Search	Bing users can upload or capture photographs to use in reverse image search.
Yahoo Answers and News	Offers Yahoo Answers as a source of user-generated information, and Yahoo News for curated and aggregated news articles.
Integrated Services	Yahoo seamlessly creates a user experience by integrating its many services, including Yahoo Mail, Yahoo Finance, Yahoo Sports, and its SE.
Trending Topics	Yahoo showcases the most searched topics directly on the homepage.
Privacy Focus	Being a privacy-first tool, DuckDuckGo garners popularity since users' activities are not tracked and search results are not tailored to them.
Bang Search	This feature allows users to search directly on other websites by using commands like "! w" for Wikipedia or "! yt" for YouTube.
Instant Answers	DuckDuckGo provides quick answers at the top of the page without requiring users to click on a link.
Qwant Music	Special feature to search for music-related content like artists, albums, and lyrics.
Transparency	Ecosia provides transparency reports on how revenue from searches is used to fund environmental projects.
Web Search	Originally concentrating on Question and Answers, Ask.com now uses conventional online search using algorithms that are comparable to those of other SEs.
Computational Search	WolframAlpha focuses on providing computational answers, calculations, and data-driven insights rather than traditional web search results.
Clean Interface	Startpage meets the needs of privacy conscious users and combines the accuracy of Google's results with anonymity, providing an uncluttered user interface.
Ad-Free Interface	The SE places minimal ads and does not track users for targeted advertising.
Family-Friendly	Swisscows is designed to filter out explicit content, making it a family-friendly SE.
Privacy Focus	Like DuckDuckGo, Swisscows does not track users or store their personal data.

Table 4 This comparison shows that advanced features like Google's Knowledge Graph and DuckDuckGo's Bang Search enhance usability, while engines without such innovations remain basic.

Table 5 The overall evaluation ranks Google and StartPage highest, demonstrating balanced performance across relevancy, speed, and features, while Ask.com scored lowest across all criteria.

The data uncovered that each search engine has its own way of relevance, speed, and privacy handling different types of queries. These differences were analyzed in the next section by comparing them with the available literature, examining emerging trends, and discussing implications for the users and developers.

5 Discussion Summary

SE's innovation has evolved because of user needs and the challenges of speed, relevance, and personalization. Early SEs, Archie and Veronica, were primitive due to their simple indexing and keyword matching techniques, which severely limited their scope of retrieving complex and relevant results. Modern SEs such as Google and Bing use highly sophisticated

algorithms like PageRank and BERT to process even the most intricate user requests. Even with these advancements, striking a balance between recall and accuracy remains a challenge. For research-focus searches, high accuracy deeply impacts recall.

Privacy-first engines like Swisscows and DuckDuckGo prioritize ethical data management, but such a lack of tailoring may inhibit answering ambiguous queries. Google and other AI SEs deliver highly relevant results, but often relies on massive amounts of user data, creating ethics and privacy risks. Another big concern is discrimination in results ranking. PageRank and participation rates are two examples of ways how SE algorithms may unknowingly favor particular viewpoints and thus affected the overall consensus.

While technologies like voice search and multimedia have improved the user experience, these models fail to support irregular types of language. There has been considerable improvement when comparing traditional SE features with modern SE features. Semantic search and AI-driven learning are just two examples of very advanced technology that can today success-

Table 5. Overall comparative evaluation of ten search engines across relevancy, speed, and feature richness

Property	Google	Bing	Yahoo	DuckDuckGo	Swisscows	Qwant	Ecosia	Ask.com	WolframAlpha	StartPage
Personalized Results	YES	YES	YES	NO	NO	NO	YES	NO	NO	YES
Ads Displayed	YES	YES	YES	NO	NO	YES	YES	YES	NO	NO
Privacy-Focused	NO	NO	NO	YES	YES	YES	YES	NO	YES	YES
Global Reach	YES	YES	YES	YES	YES	YES	YES	YES	Limited	YES
Rich Media	YES	YES	YES	NO	NO	YES	YES	YES	NO	YES
Q&A-style Results	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
AI-Powered Suggestions	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES

fully answer an array of information requests than did simple keyword indexing technology, which lacked objective ranking capability. Problem-solving in the domains of high-level queries, unbiased ranking, and fair access remains issues.

The argument then continues to describe how tested SEs work with users (in reality) and whether and the results differ from earlier studies. This analogy has numerous practical uses, but it will also bring along such concern as fair access, unbiased ranking, and error-filled searches. Subsequent to that analysis of those data, the conclusion has a summary of the key findings and the likely worth of this comparison to SE innovation.

Google's superior performance (83%) can be attributed to RankBrain and BERT, which capture semantic meaning, while Bing relies heavily on BM25, favoring keyword-based matching. Privacy engines like DuckDuckGo lack such advanced ranking, explaining their lower recall scores.

6 Limitations

A good study showing some visions. Nevertheless, there are some limitations. The evaluation was done only on desktop computers and with a single browser on this browser that means that some aspects of SE performance may not be the same experience on mobile devices or other pcs. Future studies should therefore include mobile devices, tablets, and multiple browsers to provide a more representative evaluation.

The queries used were chosen pre-screening based on English language and their chosen nominated answers. This means that the questions do not represent the full range of queries people search for in different languages and in real world settings. Future stud-

ies should test multilingual datasets to make results more globally applicable. Personal factors like search history, location and user preferences were not considered even though it is obvious that they have effect on SE performance and in practice they affect more than one aspect of SE behavior. It was found that searches for popular culture and Ecommerce were almost half of all SE queries whereas half of the queries were also entered with the website in mind . To address this, future research should incorporate personalization factors such as search history, geolocation, and user profiles to better simulate real-world conditions.

The other limitation of the study is that the results are judged manually which may lead to personal bias or enigma. This can be remedied by employing multiple evaluators and calculating inter-rater reliability (e.g, Cohen's kappa) or by using automated evaluation tools. The other limitation of the study is that the results are judged manually which may lead to personal bias or enigma. Also, the study did not measure how SE results may change over time or improve with repeated use.

A further limitation is that not all of the newer technologies like voice search, visual search and AI interfaces can be incorporated even though they are increasingly common in today's search behavior. This limitation could be addressed by including emerging technologies such as AI-based conversational search, voice-based assistants, and image/video search in future evaluations. While this study provides a comprehensive picture of SE performance across a variety of criterion, various limitations may have applied to the results; this latter section describes these limitations and provides suggestions for future research for

improving reliability and depth in SE evaluation.

7 Future Work

Future research could extend the evaluation by incorporating more SEs from an overall dataset and those of a larger range, such as SEs designed for a given domain (e.g. academic, medical, legal) or larger sets of domains. Further tests should be conducted using more sophisticated user interaction metrics (e.g. click-through rate, dwell time, and user satisfaction surveys) to analyze the real-world effectiveness of the SEs.

Testing SEs in different languages and under more diverse network conditions could also help improve their validity globally. In addition, search technology is constantly evolving (e.g. artificial intelligence, voice-based interfaces) and further work should examine the impact of these advances on information retrieval, personalization, and privacy. Finally, the use of automated tools to conduct large-scale comparative testing could greatly increase the efficiency and accuracy of evaluations.

Overall, the results reflect the strengths and weaknesses we observed with the evaluated SEs. These findings provide useful insight into the capabilities of the evaluated SEs and areas of improvement. Nonetheless, given the broad range of information retrieval environments, there are several opportunities for further advanced and much wider analysis. Hence, the next section will outline future directions that can build on this recent work and address the limitations of this study.

8 Conclusion

Modern SEs accuracy/speed and user-centric features have improved the ability to find information and the results of AI and natural language processing and real-time indexing have greatly impacted search quality and quantity. As shown in Table 1, Google outperformed others with 83% relevancy, followed by StartPage (80%).

Nevertheless, problems such as bias in the algorithm, data privacy, and recall gaps still exist, and addressing these issues requires an intelligent strategy to optimize recall without affecting precision,

especially for exploratory and specific queries, developing transparent and explainable algorithms can reduce bias and increase user trust, sophisticated features should prioritize privacy by minimizing user tracking and expanding support for non-standard languages and inclusive interfaces will further improve accessibility.

In conclusion, while SEs have transformed information retrieval, their continued evolution must align with ethical considerations, inclusivity, and user-centric innovation. These directions will pave the way for more equitable, transparent, and efficient search technologies.

Author Contributions

Ayesha Khan: Conceptual Framework, Theoretical Writing, Editing, Proofreading, Visualization and Graphics, Result Analysis **Dr. Sarah Bukhari:** Supervision, Writing- Reviewing and Editing, Conceptualization **Umar Khattab:** Technical Writing. **Abubakar Kamal:** Experimental Evaluation.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain any studies with human participants or animals performed by any of the authors.

References

- [1] R. C. Balabantaray, "Evaluation of web search engine based on ranking of results and its features," *International Journal of Information and Communication Technology*, vol. 10, 2017.
- [2] B. Kumar, K. S. M., P. Reddy, and T. Prathima, "Advancements in search engines," *Grenze International Journal of Engineering & Technology (GIJET)*, vol. 10, no. 2, pt. 2, p. 873, 2024.
- [3] O. Khattab and M. Zaharia, "Colbert: Efficient and effective passage search via contextualized late interaction over BERT," in *Proc. 43rd Int. ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR)*, 2020, pp. 39–48.
- [4] A. Ullah, N. M. Nawi, E. Sutoyo, A. Shazad, S. N. Khan, and M. Aamir, "Search engine optimization algorithms

- for page ranking: Comparative study," *International Journal of Integrated Engineering*, vol. 10, pp. 19–25, 2018.
- [5] U. Khattab, S. Bukhari, A. Khan, A. Kamal, and C. Author, "A multidimensional comparison of search engines and ChatGPT in terms of information needs." [Online]. Available: <https://doi.org/10.56979/901/2025>
- [6] D. Sharma, R. Shukla, A. K. Giri, and S. Kumar, "A brief review on search engine optimization," in *Proc. 9th Int. Conf. Cloud Computing, Data Science and Engineering (Confluence)*, 2019.
- [7] F. Almukhtar, N. Mahmood, and S. Kareem, "Search engine optimization: A review," *Applied Computer Science*, vol. 17, pp. 69–79, 2021.
- [8] J. Rowland, S. López-Asensio, A. Bagci, A. Delicado, and A. Prades, "Shaping information and knowledge on climate change technologies: A cross-country qualitative analysis of carbon capture and storage results on Google search," *Journal of the Association for Information Science and Technology*, vol. 75, pp. 625–639, Sep. 2024.
- [9] S. S. S. B, S. M, S. K, K. G, and A. R, "Natural language processing techniques for information retrieval enhancing search engines with semantic understanding," *ITM Web of Conferences*, vol. 76, p. 5013, Sep. 2025. [Online]. Available: <https://www.itm-conferences.org/10.1051/itmconf/20257605013>
- [10] N. Khokale, "Overview of search engine optimization," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 2019.
- [11] C. Shah and E. M. Bender, "Envisioning information access systems: What makes for good tools and a healthy web?" *ACM Transactions on the Web*, vol. 18, Sep. 2024.
- [12] V. Karpukhin, B. Oğuz, S. Min, P. Lewis, L. Wu, S. Edunov, D. Chen, and W. tau Yih, "Dense passage retrieval for open-domain question answering," Sep. 2020. [Online]. Available: <http://arxiv.org/abs/2004.04906>
- [13] L. Page and S. Brin, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks*, vol. 30, 1998.
- [14] N. Makrydakakis, "SEO mix 6 O's model and categorization of search engine marketing factors for websites ranking on search engine result pages," *International Journal of Research in Marketing Management and Sales*, vol. 6, pp. 18–32, Jan. 2024.
- [15] K. Kowalczyk and T. Szandala, "Enhancing SEO in single-page web applications in contrast with multi-page applications," *IEEE Access*, vol. 12, pp. 11597–11614, 2024.
- [16] S. Gupta, N. Agrawal, and S. Gupta, "A review on search engine optimization: Basics," *International Journal of Hybrid Information Technology*, vol. 9, pp. 381–390, Sep. 2016.
- [17] L. ling Hsu and Z. Walter, "Search engine or content website? A local information seeking classification model based on consumer characteristics and website perceptions," *International Journal of Human-Computer Interaction*, vol. 31, pp. 263–276, Sep. 2015.
- [18] C. D. Schultz, "Informational, transactional, and navigational need of information: Relevance of search intention in search engine advertising," *Information Retrieval Journal*, vol. 23, pp. 117–135, Sep. 2020.
- [19] L. Wang, J. Wang, M. Wang, Y. Li, Y. Liang, and D. Xu, "Using internet search engines to obtain medical information: A comparative study," *Journal of Medical Internet Research*, vol. 14, 2012.
- [20] K.-L. Liu, A. Santoso, C. Yu, and W. Meng, "Discovering the representative of a search engine," in *Proc. 10th Int. Conf. Information and Knowledge Management (CIKM)*, 2001, pp. 577–579. doi: 10.1145/502585.502696.
- [21] M. Gordon and P. Pathak, "Finding information on the World Wide Web: The retrieval effectiveness of search engines," *Information Processing and Management*, vol. 35, 1999.
- [22] D. Hawking, N. Craswell, P. Bailey, and K. Griffiths, "Measuring search engine quality," *Information Retrieval*, vol. 4, 2001.
- [23] G. Egri and C. Bayrak, "The role of search engine optimization on keeping the user on the site," *Procedia Computer Science*, vol. 36, pp. 335–342, 2014.
- [24] V. Waller, "Not just information: Who searches for what on the search engine Google?" Feb. 2011.