Annals of Library and Information Studies
Vol. 66, September 2019, pp. 117-122

Retrieval performance of select search engines in the field of physical sciences

S.M. Shafi[a] and Sabah Ali[b]

[a]UGC Emeritus Fellow, Department of Library and Information Science, University of Kashmir, Jammu and Kashmir, Email: smshafi@kashmiruniversity.ac.in

[b]Department of Library and Information Science, University of Kashmir, Jammu and Kahsmir, Email: sanazworld.ali@gmail.com

Received: 04 July 2019; accepted: 18 September 2019

The study aims to provide a systematic evaluation of the search engines on the basis of two information retrieval parameters (precision and relative recall) with reference to physical sciences. It employed 'Web of Science' to identify data (one to three word queries) of highly ranked authors who have contributed to the discipline of physical sciences. The three English language search engines (Google, Yahoo and Bing) were selected on the basis of ranking of 'Alexa' (Actionable Analytics for the Web). The study reveals that in all (one, two and three word) queries 'Google' obtained highest precision and relative recall followed by 'Yahoo' and 'Bing'. It further shows that 'Google' and 'Yahoo' achieve the highest 'precision' and 'relative recall' due to their wide coverage. Bing once prominent one, however lags behind in retrieval effectiveness.

Keywords: Search engines; Retrieval effectiveness; Precision and recall

Introduction

A large volume of information is available through the internet and search engines are used for retrieving information from the web by applying various indexing techniques, searching behavior, algorithm etc.[1]. However, owing to massive information on the internet, search engines many times are not able to provide the most relevant information to users.

Kim and Carvalho (2011) reveal that search engines are the best tools available for finding information on the web and have similar intensity of fluctuation among top results but most of these variations are not permanent[2]. Chowdhury and Soboroff (2002) state that search engines vary in their searching interface but some of the search engines are generally quite similar in terms of effectiveness and there exists a significant gap among the most excellent and poorer[3]. Deka and Lahkar (2010) focus on use of different techniques and indexing algorithm in search engines to find and present web information for users and therefore the results displayed by search engines vary from one another[4]. The performance of search engines can be improved if these manipulate an amplified knowledge of user behavior to simply recognize the basic target of searchers. These can determine the primary goals of users based on their queries and other interactions by employing algorithms and interfaces with the assistance of designers of web search engines that can help users to accomplish their searching goals[5].

On the other hand, search engines provide essential access to the web equally to those who have something to communicate and recommend including those who contribute some information to web viz; information providers, writers, authors etc. as well as to those who desire to listen and discover including people who searches for information viz., information users[6].

Kumar (2012) states that users make use of search engines to achieve their information needs as most of the users view them as a basic tool for retrieving information where users get the results from them. However, most of the users are not aware about the search strategies offered by search engines to assist their users in order to get the relevant results. Meanwhile, it is understood that a number of users in the current era look for the information while using search engines frequently and only a very small percentage of users gaze for the information on the web rarely[7].

One of the best tools available for seeking online information is a search engine which acts as a platform where the people can search for any kind of information[8]. It has been revealed that most of the users make use of Google while looking for the information

as it provides better interface, features and ease of use to the users as compared to any other search engine available. Search engines have become vital creators of knowledge where knowledge is not only disseminated to its users but also manufactured by ranking system of search engines[9]. A number of search engines available nowadays have different searching mechanism, interface, features, techniques, coverage of the web, algorithm, indexing and ranking methods etc. which make them different from one another[3].

Technical mechanism including crawling, indexing and ranking algorithms are the important parameters which help search engines to provide eminence to established, prosperous and dominant sites[6]. However, effectiveness is an important measure which helps a researcher to identify the ability of different search engines to retrieve better results while taking into account factors including, "interface design" "result appearance" and "relevance of the hits"[10]. Lewandowski (2012) highlights that a retrieval performance of search engines can be enhanced or improved by applying various quality factors like "index quality, quality of the results, quality of search features and search engine usability"[11].

Kumar and Prakash (2009) point out that a variation lies in search aptitude, user interface and also in the quality of information among two search engines viz; Google and Yahoo. However, both these search engines retrieve more relevant sites when comparing with other search engines. It is noticed that Google makes use of web graph and link structure for more inclusiveness and consistency[12]. Lewandowski (2011) found that the performance of two search engines viz; Google and Yahoo is better as compared to any other search engine e.g. MSN[13].

Google remains the top general search engine because of its enormous authority on the web panorama and thus users optimize their web pages to enhance and increase the rankings of the pages on Google[14]. It is rather difficult for the search engines to provide enough information when the length of query is shorter and thus users receive low quality result list[15]. Users of search engines do not make use of sophisticated search features offered by the search engines and thus lack the relevance among the results. Further, these also reveal that a large number of users utilize only a small number of search queries and thus scrutinize only a small number of Webpages[16,17].

In order to categorize queries as per the needs of different users, authors reported about an automatic method to categorize different Arabic queries in three types viz; navigational, informational, and transactional[18]. Broder (2002) reveals that informational and navigational queries are well treated by the latest search engines that provide most relevant results to its users[19].

This study has attempted to evaluate the selected search engines, taking into account different types of web queries (simple to complex) in physical sciences and compares their retrieval effectiveness in terms of precision and relative recall.

Objective of the study

- To evaluate and estimate retrieval effectiveness (Precision &Relative Recall) of select search engines.

Methodology

Alexa (Actionable Analytics for the Web) has listed top 500 sites viz: search engines, portals, directories, social networking sites, networking tools etc. The study, however, has confined itself to three highest ranking general search engines viz; Google, Yahoo and Bing in English language. Further, the study has made use of 'Web of Science'(WOS) as a source for collecting data (web queries) from highest ranked authors in the field of physical sciences. Fifteen terms related to physical sciences were obtained from Web of Science (Table 1). All the terms were later classified into three categories viz; "one-word term, two-word term and three-word term"

Table 1 — Terms identified for the study

S. no.	Type	Terms	Number (%)
1.	One-Word Term (simple)	Solvent Isotope Valence Density Velocity	5 (33.33)
2.	Two-Word Term (compound)	Chemical Reaction Atomic Mass Activation Energy Potential Energy Thermal Expansion	5 (33.33)
3.	Three-Word Term (complex)	Balanced Chemical Equation Average Atomic Mass Double Displacement Reactions Expansion Combustion Engine Passive Solar Heating	5 (33.33)
	Total		15 (100)

**The figures in the parenthesis indicate the percentage.*

The selected terms were queried in the identified search engines. The first twenty results were evaluated for the purpose of estimation of precision and relative recall for the respective search engines. The scale and formula, more or less, adopted earlier by Shafi and Rather (2005) was used with a few minor modifications for the calculation of precision and relative recall[20].

The "Average precision" is the value obtained for set of top documents existing from each relevant document retrieved. The following scoring was given to different web pages(documents) to estimate the precision of search engines for the selected queries.

a. If a web page is significantly related to the subject matter of the search query, it is grouped as "more relevant" and given a score of 2.
b. If a web page includes some relevant ideas of the subject matter of the search query, it is grouped as "less relevant" and given a score of 1.
c. If a web page is not associated to the subject matter of the search query, it is grouped as "irrelevant" and given a score of 0.

If a web page consists of a whole series of links, rather than the information required, then it was categorized as 'links' and given a score of 0.5, if inspection of one or two of the links proved to be useful.

The formula for estimation of precision and related recall of selected search engine for each of the search queries adopted is:

$$\text{Precision} = \frac{\text{Sum of the scores retrieved by a search engine}}{\text{Total number of sites selected for evaluation}}$$

$$\text{Relative recall} = \frac{\text{Total sites retrieved by a particular search engine}}{\text{Sum of sites retrieved by three search engines}}$$

Test environment

A total of 15 queries, distributed equally among three categories, were identified from top authors listed in Web of Science (WOS) in the field of Physical Sciences from 1st May to 20th June 2017. (Table 1)

Analysis

Relevance and mean precision: simple terms

For simple terms (one word), Google gave the highest relevant results followed by Yahoo whereas Bing retrieved most irrelevant results. Google had the highest Mean precision (1.13) followed by Yahoo (1.07) and Bing (0.98). Google had the highest precision (1.25) for a single term while both Yahoo and Bing yielded highest values of 1.15 and 1.07 respectively for the terms. Thus simple terms have a disadvantage of poor precision and require to refinement by the users to reach to a better precision in achieving their retrieval goals (Tables 2 & 2 .1).

Relevance and mean precision: compound terms

The overall highest yield for most relevant terms for two word-terms (compound) was retrieved by Google followed by Yahoo. Bing got a higher score for compound term as compared to simple terms. However, the mean precision for two word terms reveals that Google had highest precision (1.16) followed by Yahoo (1.08) and Bing (1.05). In Google, the search query "*Thermal Expansion"* got the highest precision of 1.27 and Yahoo also had the highest

Table 2.1 — Mean precision : simple terms

Term	Precision (Google)	Precision (Yahoo)	Precision (Bing)
Solvent	1.15	1.07	0.90
Isotope	1	1	0.92
Valence	1.15	1.07	1.05
Density	1.1	1.05	0.97
Velocity	1.25	1.15	1.07
Overall mean precision	1.13	1.07	0.98

Table 2 — Relevance status: simple terms (n=20)

Term	Most relevant			Less Relevant			Irrelevant			Links		
	G	Y	B	G	Y	B	G	Y	B	G	Y	B
Solvent	8	7	6	5	6	7	3	4	5	4	3	2
Isotope	7	7	6	4	4	5	5	5	6	4	4	3
Valence	9	8	7	3	4	6	4	5	5	4	3	2
Density	8	7	6	4	5	6	4	4	5	4	4	3
Velocity	9	8	7	5	5	6	2	3	4	4	4	3
	41	37	32	21	24	30	18	21	25	20	18	13

Note: G-Google, Y-Yahoo, B-Bing

precision of 1.15 for "*Thermal Expansion"*. Bing had the highest precision of 1.12 for "*Chemical Reaction"* (Tables 3 & 3.1). This suggests that variations exist even among matching / tagging of the compound terms in different engines, although Google gave a better performance.

Relevance and mean precision: complex terms

The three word terms (complex phrases) in the field show that most relevant results are available with Google and sizeable variation does not exist in other engines in retrieving irrelevant or even less relevant results (Table 4). The mean precision for such complex terms expose that 'Google' has the highest precision (1.22)followed by 'Yahoo'(1.15) and Bing goes down to1.08.The search term "*Double Displacement Reaction"* attains the highest precision of 1.32 through Google. However, in Yahoo the highest precision (1.22) is attained for a term "*Balanced Chemical Equation"* and Bing gets a better precision (1.12)for two queries namely "*Balanced Chemical Equation"* and "*Expansion Combustion Engine"* (Table 4.1). The whole estimation establishes that the complex terms accessed through advanced techniques helps to have better relevant retrieval in almost all search engines and definitely better precision.

Relative recall: simple terms

The relative recall of one word simple terms reveal that such queries yield very high results in each search engine but relative recall for different terms varies even within each search engine. The overall relative recall of Google is the highest among the search engines and very low in the Bing. One term gets relative recall of 0.72, higher than the Mean value for the Google and most of the terms in Bing get even as less relative recall as 0.1. (Table 5)

Relative recall: compound terms

The mean recall for two-word term is comparatively very low for Google but higher for Bing when compared with the one-word term. The overall relative recall for two word queries reveal that Google has the highest relative recall (0.44) followed

Table 3.1 — Mean precision: compound terms

Term	Precision (Google)	Precision (Yahoo)	Precision (Bing)
Chemical Reaction	1.15	1.12	1.12
Atomic Mass	1.1	1.02	1
Activation Energy	1.15	1.05	1.07
Potential Energy	1.17	1.1	1.02
Thermal Expansion	1.27	1.15	1.07
Overall Mean Precision	1.16	1.08	1.05

Table 4.1 — Mean precision of complex terms

Term	Precision (Google)	Precision (Yahoo)	Precision (Bing)
Balanced chemical equation	1.25	1.22	1.12
Average atomic mass	1.2	1.1	1.02
Double displacement reactions	1.32	1.2	1.1
Expansion combustion engine	1.2	1.15	1.12
Passive solar heating	1.15	1.1	1.05
Overall mean precision	1.22	1.15	1.08

Table 3 — Relevance status: compound terms (n=20)

Term	Most Relevant			Less Relevant			Irrelevant			Links		
	G	Y	B	G	Y	B	G	Y	B	G	Y	B
Chemical Reaction	9	8	8	4	5	5	5	4	4	2	3	3
Atomic Mass	8	7	6	4	4	6	4	4	4	4	5	4
Activation Energy	8	7	7	5	5	6	3	4	4	4	4	3
Potential Energy	9	8	7	4	4	5	4	4	5	3	4	3
Thermal Expansion	10	9	8	4	3	4	3	4	5	3	4	3
Grand total	44	39	36	21	21	26	19	20	22	16	20	16

Note: G-Google, Y-Yahoo, B-Bing

Table 4 — Relevance status of complex terms (n=20)

Term	Most Relevant			Less Relevant			Irrelevant			Links		
	G	Y	B	G	Y	B	G	Y	B	G	Y	B
Balanced chemical equation	10	9	8	4	5	5	4	3	4	2	3	3
Average atomic mass	9	8	7	4	4	5	3	4	5	4	4	3
Double displacement reactions	11	9	8	3	4	4	3	3	4	3	4	4
Expansion combustion engine	9	8	8	4	5	5	3	3	4	4	4	3
Passive solar heating	9	8	7	3	4	5	4	4	4	4	4	4
G. Total	48	42	38	18	22	24	17	17	21	17	19	17

Note: G-Google, Y-Yahoo, B-Bing

Table 5 — Retrieval status and relative recall: simple terms

Term	Google		Yahoo		Bing	
	Sites retrieved	Relative Recall	Sites Retrieved	Relative Recall	Sites retrieved	Relative Recall
Solvent	8,400,000	0.58	45,400,000	0.30	16,765,566	0.11
Isotope	1,440,000	0.57	483,000	0.19	585,256	0.23
Valence	3,600,000	0.46	41,200,000	0.52	637,855	0.01
Density	462,000,000	0.72	167,000,000	0.26	4,545,556	0.01
Velocity	50,400,000	0.52	40,200,000	0.41	5,454,455	0.05
G. Total	638,540,000	0.66	294,283,000	0.30	27,988,688	0.03

Table 6 — Retrieval status and relative recall: compound terms

Term	Google		Yahoo		Bing	
	Sites retrieved	Relative Recall	Sites Retrieved	Relative Recall	Sites retrieved	Relative Recall
Chemical Reaction	7,510,000	0.04	7,460,000	0.04	16,656.8767	0.91
Atomic Mass	782,000,000	0.88	98,700,000	0.11	4,552.265	0.01
Activation Energy	25,800,000	0.14	25,800,000	0.14	12,589,6325	0.70
Potential Energy	35,500,000	0.06	458,000,000	0.83	5,4562,565	0.09
Thermal Expansion	700,000	0.27	31,800,000	0.12	152,654,555	0.59
G. Total	921,010,000	0.44	621,760,000	0.30	504,234,477	0.24

Table 7 — Retrieval status and relative recall: complex terms

Term	Google		Yahoo		Bing	
	Sites retrieved	Relative Recall	Sites Retrieved	Relative Recall	Sites retrieved	Relative Recall
Balanced chemical equation	46,000,000	0.78	127,000,000	0.21	1,275,607	0.00
Average atomic mass	8,270,000	0.23	12,000,000	0.33	15,433,656	0.43
Double displacement reactions	12,200,000	0.08	2,960,000	0.02	12,454,5677	0.89
Expansion combustion engine	186,000	0.06	190,000	0.06	254,6767	0.87
Passive solar heating	863,000	0.10	2,940,000	0.35	4,567,677	0.54
Total	489,519,000	0.62	145,090,000	0.18	14,836,9384	0.18

by Yahoo (0.30) and Bing (0.24). However, in case of Google, the search query *chemical reaction* has the highest relative recall (0.88) and least relative recall (0.06) for search query *potential energy*. The highest relative recall (0.83) is for search query *potential energy* and lowest relative recall (0.04) for the query *chemical reaction* is maintained by Yahoo. Bing, on the other hand has highest relative recall (0.70) for the query *activation energy*. (Table 6).

Relative recall: complex terms

The overall relative recall for three word queries reveals that Google has the highest relative recall (0.62) followed by Yahoo and Bing each obtaining 0.18. While comparing with one-word sample queries, it is still lower but better as compared with two word queries. However, in case of Google, the search query *balanced chemical equation* bears the highest relative recall (0.78) and least (0.06) for search query *expansion combustion engine*. Yahoo, on the other hand, has the highest relative recall (0.35) for search query *passive solar heating* and lowest relative recall (0.02) for search query *double displacement reactions*. Bing had highest relative recall of 0.89 for query *double displacement reactions* (Table 7). The relative recall, thus, improves with addition of more terms perhaps refining the broader ones with qualifiers etc.

Conclusion

It is concluded that all three search engines perform better in terms of mean precision for queries with more words. For one, two and three word queries, Google maintains highest precision and relative recall in the field of physical science followed by Yahoo and Bing. However, this conclusion is limited by many parameters which include the queries. Further research is needed to include a larger and more diverse sample of queries with different levels of domain expertise and familiarity with information retrieval systems.

References

1 Xie H, Online IR system evaluation: online databases versus Web search engines. *Online Information Review,* 28 (3) (2004) 211-219. Available at https://doi.org/10.1108/14684520410543652.

2 Kim J, and Carvalho V R, An analysis of time-instability in web search results. In *European Conference on Information*

Retrieval, (2011). pp. 466-478. Springer, Berlin, Heidelberg. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.414.1497&rep=rep1&type=pdf.

3 Chowdhury A and Soboroff I, Automatic evaluation of world wide web search services. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*,(2002) pp. 421-422. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.7192&rep=rep1&type=pdf

4 Deka S K and Lahkar N, Performance evaluation and comparison of the five most used search engines in retrieving web resources, *Online Information Review*, 34(5)(2010) 757-771. Available at https://doi.org/10.1108/ 14684521011084609.

5 Jansen B J, Booth D L and Spink A, Determining the user intent of web search engine queries, *In Proceedings of the 16th international conference on World Wide Web*(2007). pp. 1149-1150. Available at http://www.bernardjjansen.com/uploads/2/4/1/8/24188166/jansen_user_intent_search_queries.pdf

6 Introna L D and Nissenbaum H, Shaping the Web: Why the politics of search engines matters, *The Information Society*, 16(3) (2000)169-185. Available at https://doi.org/10.1080/01972240050133634.

7 Kumar S, The impact of demographic characteristics of users on patterns of usage on search engines and OPAC, *Library Review*, 61(3) (2012) 172-187. Available at https://doi.org/10.1108/00242531211259300.

8 Kaur M, Bhatia N and Singh S, Web search engines evaluation based on features and end-user experience, *International Journal of Enterprise Computing and Business Systems*, 1(2)(2011) 1-19. Available at http://www.ijecbs.com/July2011/47.pdf

9 Van Dijck J, Search engines and the production of academic knowledge, International *journal of cultural studies*, 13(6)(2010) 574-592. Available at https://doi.org/10.1177/1367877910376582.

10 Oppenheim C, Morris A, McKnight C and Lowley S, The evaluation of WWW search engines, *Journal of Documentation*, 56(2)(2000) 190-211. Available at https://doi.org/10.1108/00220410010803810.

11 Lewandowski D, A framework for evaluating the retrieval effectiveness of search engines. In Next generation search engines: *Advanced models for information retrieval*,(2012) pp. 456-479. DOI: 10.4018/978-1-4666-0330-1.ch020.

12 Kumar B S and Prakash J N, Precision and relative recall of search engines: A comparative study of Google and Yahoo, *Singapore Journal of Library & Information Management*, 38(1)(2009) 124-137. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.465.6275&rep=rep1&type=pdf

13 Lewandowski D, The retrieval effectiveness of search engines on navigational queries. In *Aslib Proceedings*, 63(4) (2011) 354-363. Available at https://doi.org/10.1108/00012531111148949.

14 Bar Ilan J, Methods for measuring search engine performance over time, *Journal of the American Society for Information Science and Technology*, 53(4)(2002) 308-319 Available at https://doi.org/10.1002/asi.10047.

15 Liu Y, Zhang M, Ru L and Ma S, Automatic query type identification based on click through information. In: Ng H T, Leong M K, Kan M Y, Ji D. (eds).

16 Information Retrieval Technology. *AIR Lecture Notes in Computer Science*, 4182 (2006) pp.593-600. Springer, Berlin, Heidelberg. Available at https://doi.org/10.1007/11880592_51.

17 Moukdad H and Large A, Users' perceptions of the Web as revealed by transaction log analysis, *Online Information Review*, 25(6) (2001) 349-359. Available at https://doi.org/10.1108/EUM0000000006534.

18 Spink A, Wolfram D, Jansen M B and Saracevic T, Searching the web: The public and their queries, *Journal of the American Society for Information Science and Technology*, 52(3) (2001) 226-234. Available at https://doi.org/10.1002/10974571(2000)9999:9999<: AID-ASI1591>3.0.CO;2-R.

19 Maabreh M A, Al-Kabi M N and Alsmadi I M, Query classification and study of university students' search trends, *Program*, 46(2) (2012) 220-241. Available at https://doi.org/10.1108/00330331211221855.

20 Broder A, A taxonomy of web search, In *ACM Sigir Forum*, 36(2) (2002) 3-10. Available at doi>10.1145/792550.792552.

21 Shafi S M and Rather R A, Precision and recall of five search engines for retrieval of scholarly information in the field of biotechnology, *Webology*, 2(2)(2005) 42-47. Available at http://www.webology.org/2005/v2n2/a12.html