

Bahari Varzaneh, Hosein (2023). Comparing the Performance of Information Retrieval of Semantic and Keyword Search Engines Based on Phrase Search. *Journal of Knowledge-Research Studies*, 1 (2), 100-114.

DOI: 10.22034/JKRS.2022.51351.1014

URL: [https://jkrs.tabrizu.ac.ir/article\\_15085.html](https://jkrs.tabrizu.ac.ir/article_15085.html)

The paper is an open access and licensed under the Creative Commons CC BY NC license.



## Comparing the Performance of Information Retrieval of Semantic and Keyword Search Engines Based on Phrase Search

Hosein Bahari Varzaneh\*<sup>1</sup>

Received: April, 21, 2022;

Accepted: June, 17, 2022

### Abstract

**Purpose:** The aim of the research is to compare the performance of Information Retrieval of semantic and keyword search engines based on phrase search (simple & complex).

**Methodology:** The present applied and semi-experimental research community includes all active search engines on the web. Research samples were selected based on stratified random sampling and purposive sampling. The data collection tool of two researcher-made checklists includes ten simple and complex phrase queries.

**Findings:** Bing and Cluuz (with similar precision of 53%), DuckDuckGo, and Yahoo were the most accurate in searching for simple phrases, respectively. Bing, DuckDuckGo, Yahoo, and Cluuz were the most accurate in their search for complex terms, respectively. In general, Bing, DuckDuckGo, Cluuz and Yahoo have the highest precision, respectively. Also, the average total precision of keyword search engines is more heightened than semantic search engines.

**Conclusion:** The Bing keyword search engine performs better than the other three semantic search engines and other keywords. Semantic search engines claim to have more capabilities in retrieving relevant information than keyword search engines. But in this study, it was found that Cluuz and DuckDuckGo do not excel in search terms over keyword search engines. These tools did not perform as well as semantic web search engines, and it seems that they have a long way to go to become real semantic search engines. And to achieve this, it is necessary to use the facilities, tools, modules, and emerging technologies of the new age, such as machine learning, deep learning, combining these modules with pervasive techniques, data mining, etc.

**Value:** So far, not been compared the phrase search performance in the sample semantic and keyword search engines. And in this regard, the researcher has tried to achieve an actual result with an exact Survey.

**Keywords:** *Information Retrieval, Phrase Search, Performance, Keyword Search Engines, Semantic Search Engines*

---

1. Master, Department of Knowledge and Information Science, Faculty of Education and Psychology, University of Isfahan, Isfahan, Iran (Corresponding Author) [bahari\\_2062@yahoo.com](mailto:bahari_2062@yahoo.com)

---

## Extended Abstract

### Introduction

The emergence of the web and the pervasiveness of information production among the general public has made the field of information retrieval an important issue (Nadi Ravandi & Hariri, 1394, 17-18). However, successful data retrieval from the web has its limitations. Like a. Content without proper structure for displaying information, b. Ambiguity in information due to the weak link between them, c. Inability to understand the information machines provide due to the lack of a global format, and d. The lack of automatic transfer of information leads to the retrieval of many irrelevant records (Sudeepthi et al., 2012). Although there are many search engines today, the main challenge of these search engines is the inability to retrieve information intelligently (Gayar, Mekky, & Atwan, 2015). The information retrieved by these search engines is associated with high retrieval and low precision, with or without retrieval (Fatima, Luca, & Wilson, 2014). Semantic web technology is a solution to this problem. Semantic Web Achievement is Semantic Search Engines (S.S.E.); these tools claim to have advantages over keyword search engines: a. unlike keyword search engines (K.S.E.) that use the page ranking algorithm; S.S.E use ontology (Malve, & Chawan, 2015); B. They work according to a semantic approach (Sahu, Mahapatra & Balabantaray, 2015) and analyze the user's desire and identify the real and hidden meaning of the user's question; J. using logical reasoning, they retrieve more accurate results (Dorri, 2014).

### Purpose

Internet users do not know much about the performance of S.S.E and keywords. Also, it is not clear how these S.S.E and keywords perform in retrieving information for phrase queries (simple and complex). The purpose of this research is to identify a successful search engine and compare the performance of these two tools in this regard.

### Methodology

The statistical population of this applied and quasi-experimental research includes all active search engines on the web. First, based on a random sampling of search engines, it was divided into two categories: S.S.E and keyword. Then, in these two categories of search engines, based on the superiority and use of search engines in previous research works and their popularity with users, they were selected from S.S.E, DuckDuckGo, Cluuz, and K.S.E, Yahoo, and Bing. The data collection tool has created two researcher checklists that include 10 query samples for phrase search (simple and complex). The relevance of the first 20 results retrieved for each query was examined, and the barrier of these tools was calculated using the precision coefficient. Finally, for each search engine, an average of the total obstacles was obtained, based on which their performance values were compared.



Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

Serial Number 2

---

## Findings

Findings obtained to achieve the objectives of the research, according to Figure 1, show that Bing and Cluuz search engines with the same precision of 53% achieved the highest precision in searching for type 1 expressions, which are considered simple expressions. Next to them is the DuckDuckGo search engine, and Yahoo's search engine showed the least precision in this study, with a minimal difference after DuckDuckGo. In addition, this chart shows that the Bing search engine has shown the highest level of precision in searching for Type 2 expressions, which are considered complex expressions. Then there are DuckDuckGo, Yahoo, and Cluuz search engines, respectively. Overall, the meanings obtained from the phrase search in these search engines showed that Bing, DuckDuckGo, Cluuz, and Yahoo retrieved the most relevant records, respectively.

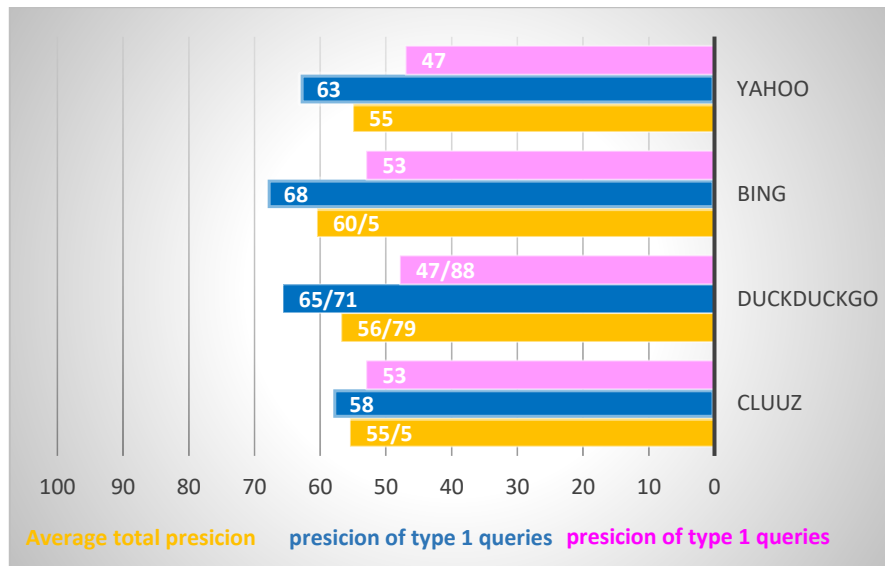


Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

Serial Number 2



**Figure 1. The average precision of search engines**

The average total precision of S.S.E and K.S.E. is 57.75% and 56.145%, respectively. Based on this, the sample S.S.E has higher average precision than the average total accuracy obtained by K.S.E.

---

## Conclusion

In simple term search, the Bing search engine performed best. This search engine can make decisions and provide immediate answers with multilingual support, and the type of results presented in it is free text and links (Negi & Kumar, 2014). Its unique feature is the "community" that allows learning and sharing resources with other people who use Bing (Dutta & Bansal, 2016). After Bing, DuckDuckGo had the best performance. DuckDuckGo goes to retrieve information with instant answers (Nandy et al., 2016) and results in categorized summaries. These results are based on many sources, including Yahoo, Wikipedia, Wolfram Alpha, and its crawler web. It has multilingual support, and the primary approach in it is cluster search and natural language processing (Negi & Kumar, 2014). Next to DuckDuckGo is a cluster whose results include cluster diagrams, images extracted from recovered results, link tags, and semantic graphs. It is possible to display the results related to the user search on the left side of the results page. Yahoo also showed the weakest performance by recovering the lowest number of related records. Yahoo, a type of search engine (Davarpanah, 1390, 96), has a hierarchical structure, and its main content is focused on financial news (Dutta and Bansal, 2016). The search results are provided along with a brief description of the retrieved websites (Davarpanah, 1390, 96), and there is a possibility of selective search based on it (Rathee, et al., 2013). In searching for Type 2 queries, which are complex, Bing performed best. It is followed by DuckDuckGo, Yahoo, and Cluuz, respectively. In searching for the two types of sample phrases, the Bing search engine performed better than Yahoo, and the dock search engine performed better than Cluuz. In general, the meanings obtained from the phrase search in these search engines showed that Bing was more powerful than the two S.S.E DuckDuckGo, Cluuz, and Yahoo search engines. This indicates that DuckDuckGo and Cluuz 's S.S.E, although they use semantic web standards, have a weak expression compared to the Bing search engine. The Bing search engine appears to be meaningful. Therefore, it is recommended that users visit the Bing search engine for a phrase search. The findings from the present study are consistent with the results of Singh and Sharan (2013), Ahmad Khan, Sangroha, & Rahman (2014), and Nagi, & Kumar (2014). But the results of Mala & Lobiyal Research (2016) are inconsistent with the results of the present study. The number of inquiries selected for performance evaluation appears to be the number of records reviewed by the queries, the searches in search engines, the new features and tools used by the search engines, and most importantly busy time and time. And the web-based non–webmaster engines perform strong



Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

---

Serial Number 2

---

---

or poor at different times. The present study is also consistent with some of the previous research and some of them inconsistent.

In a general conclusion, it can be said that although S.S.E claims to be a: Understanding users' queries; B. Realizing what the meaning of the user's inquiry and purpose is; c. Responding to long and complex queries; d. Adapting the concepts, generalization, text mining on web pages; e. Comparing links by RDF and provide graph patterns; f. Benefiting from the rules of inference, logic, and natural language system (Ramachandran, A., & Sujatha, 2011), But in this study, it was found that the DuckDuckGo and Cluuz, when searching for a phrase, are not significantly superior to K.S.E, and only a few hundredths of a few hundred are careful. Therefore, these search engines did not show an acceptable performance as web search engines. They seem to have a long path to becoming real S.S.E. To achieve this, it is necessary to use the facilities, tools, modules, and emerging technologies of the new age from machine learning, deep learning, and integrating these modules with text techniques, data mining, and so on.

### Value

So far, in the S.S.E and k.S.E., the data recovery performance has not been reviewed and compared through a phrase search. In this regard, the researcher has tried to achieve an actual result with a careful examination to introduce the powerful search engine in this type of search to users.

### Extended Abstract References

Khan, J.A., Sangroha, D., Ahmad, M., & Rahman, M.T. (2014). A performance evaluation of semantic based search engines and keyword based search engines. 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 168-173.

Aldabbagh, M. S. M. (2012). An Approach for Evaluation of Semantic Performance of Search Engines: Google, Yahoo, Msn and Hakia. *Journal of Education Science*, 25,124-130.

Aliyu, F. M., & Yahaya, Y. I. (2021). An Investigation of the Accuracy of Knowledge Graph-base Search Engines: Google knowledge Graph, Bing Satori and Wolfram Alpha. *International Journal of Scientific & Engineering Research*, 12(1), 11-15.

Azizan, A., Abu Bakar, Z., Abd Rahman, N., Masrom, S. & Khairuddin, N. (2018). A Comparative Evaluation of Search Engines on Finding Specific Domain Information on the Web. *International Journal of Engineering & Technology*, 7 (4), 1-4.

Choudhury, N. (2014). World Wide Web and Its Journey from Web 1.0 to Web 4.0. Nupur Choudhury/ (IJCSIT). *International Journal of Computer Science and Information Technologies*,5(6),8096-80100.

Davarpanah, M.R. (2011). *Scientific Information Seeking in Printed and Electronic*. Tehran: Dabizash.



Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

Serial Number 2

---

- Dorri, R. (2015). Comparison and Evaluation of Semantic Search Engines. *Iranian Journal of Information processing and Management*, 30 (2) ,467-490. [In Persian].
- Dutta, M., & Bansal, K. L. (2016). A Review Paper on Various Search Engines (Google, Yahoo, Altavista, Ask and Bing). *International Journal on Recent and Innovation Trends in Computing and Communication*, 4 (8),190- 195.
- Fatima, A., Luca, C., & Wilson, G. (2014). New Framework for Semantic Search Engine. *International Conference on Computer Modelling and Simulation*. UKSim-AMSS,445-450. DOI:10.1109/UKSim.2014.114
- Gayar, M.M.E., Mekky N., & Atwan, A. (2015). Efficient Proposed Framework for Semantic Search Engine using New Semantic Ranking Algorithm. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 6(8),136- 143.
- Goutam, R. K., & Dwivedi, S. K. (2012). Performance Evaluation of search engines via user efforts Measures. *International Journal of Computer Science Issues*, 9(2),432- 437.
- Guha, R., McCool, R. ,& Miller, E. (2003). *Semantic Search*. Retrieved 23 October, 2017 from: <https://dl.acm.org/citation.cfm?id=775250>
- Signorini, A., & Imielinski, T. (2009). If You Ask Nicely, I will Answer: Semantic Search and Today's Search Engines. 2009 IEEE International Conference on Semantic Computing, 184-191.
- Kousha, K. (2001). *Internet Tools: Principles Skills and Facilities for Searching on World Wide Web*. Tehran: ketabdar.
- Kumar, D., & Mishra, R. (2015). Deep Web Performance Enhance on Search Engine. 2015 International Conference on Soft Computing Techniques and Implementations- (ICSCTI), Hybrid Intelligence. 137- 140.
- Lancaster, F. W. (2002). indexing and abstracting( A.Gilvari Trans.). Tehran: chapar.
- Mala, V., & Lobiyal, D. K. (2016). Semantic and keyword based web techniques in information retrieval. 2016 International Conference on Computing, Communication and Automation (ICCCA). <https://doi.org/10.1109/cca.2016.7813724>
- Malve, A., & Chawan, P. M. (2015). A Comparative Study of Keyword and Semantic based Search Engine. *International Journal of Innovative Research in Science, Engineering and Technology*, 11(4),11156-11161.
- Nadi Ravandi, S., & Hariri, N. (2016.). *Information Retrieval Systems* .Tehran: Ketabdar.
- Nandy, D., Gohel, H., Kalyani, F. ,& Khakhariya, M. (2016). DuckDuckGo vs. Google search engine comparative analysis. *International Journal of Science Research and Technology*, 2(1), 37-41.
- Negi, Y. S., & Kumar, S. (2014). A Comparative Analysis of Keyword-and Semantic-Based Search Engines. In *Intelligent Computing, Networking, and Informatics*(pp.727-736).



Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

Serial Number 2

- 
- Nowruzzi, M., & Taherian, M. (2012). Semantic web. Tehran: fors.
- Ramachandran, A., & Sujatha, R. (2011). Semantic search engine: A survey. *Int. J. Comp. Tech. Appl*, 2 (6),1806-1811.
- Rashid, J., & Wasif Nisar, M. (2016). A Study on Semantic Searching, Semantic Search Engines and Technologies Used for Semantic Search Engines. *I.J. Information Technology and Computer Science*, 10(10),82- 89.
- Rathee, D., Atri, J., Rathee, J., & Nath, R. (2013). A Framework to Evaluate Search Engines. *IJCSC*, 4 (2), 184-190.
- Sahu, S. K., Mahapatra, and, D.P., & Balabantaray, R.C. (2015).Comparative Study of Different Search Engines in Context of Features and Semantics, *International Journal of All Research Education and Scientific Methods*, 14(1), 15-26.
- Sayed, A., & Al Muqrishi, A. (2017). IBRI-CASANTO: Ontology-based semantic search engine, *Egyptian Informatics Journal*,18 (3), 181- 192.
- Sheela, A.C., & Jayakumar, C.T. (2019). Comparative Study of Syntactic Search Engine and Semantic Search Engine: A Survey. 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 1, 1-4.
- Shiri, A. A. (1994). Evaluate recall and precision of information retrieval systems. *Research on Information Science and Public Libraries*, 4 (1,2), 38-45. [In Persian].
- Singh, J., & Sharan, A. (2013). A Comparative Study between Keyword and Semantic Based Search Engines. *International Conference on Cloud, Big Data and Trust 2013*: 130- 134.
- Solomou, G. & Koutsomitropoulos, D. (2015). Towards an evaluation of semantic searching in digital repositories: a DSpace case-study. *Electronic library and information systems*, 49 (1), 63-90.
- Sudeepthi, G., Anuradha, G., Surendra, M.,& Babu, P. (2012). A Survey on Semantic Web Search Engine. *IJCSI International Journal of Computer Science Issues*, 9(2), 241-245.
- Tümer, D., Shah, M.A., & Bitirim, Y. (2009). An Empirical Evaluation on Semantic Search Performance of Keyword-Based and Semantic Search Engines: Google, Yahoo, Msn and Hakia. 2009 Fourth International Conference on Internet Monitoring and Protection, 51-55.
- Yadav, A., & Maheshwari, S. (2017). Building of Intellectual Web Search Engines through Semantic Web. *International Journal of Latest Trends in Engineering and Technology*, 7 (3),384-391.



Journal of  
Knowledge-Research Studies  
(JKRS)

Vol 1

Issue 2

Serial Number 2

---