

Improved Online Federated Search Technique for Systematic Literature Review Process

Ms.Mohini Arote
Dept. of Computer Engg,
D. Y. Patil College of Engg, Akurdi,
Pune, India
mohiniarote@gmail.com

Mrs. D. A. Phalke
Dept. of Computer Engg,
D. Y. Patil College of Engg, Akurdi,
Pune, India
a_dhanashree@rediffmail.com

Abstract— The process of federated search is nothing but the method for finding multiple text information's at the same time. This technique is also known as distributed or federated information retrieval. In this process, user queries are submitted to the collections subset those are nearly returning the queries relevant answers. Returned results of selected collections are then integrated and merged into a single list. Federated search technique is most frequently adopted as compared to centralized search methods at many places. In this paper, federated search tool is being discussed on behalf of its use in academic research studies. Evidence-based research, systematic literature review (SLR) now day's process widely used method. SLR is important for Federated search tool and process data could change the outcome of any SLR. This may bias the research. Thus the SLR process techniques are introduced to improve first, and then the well-known software engineering database automatically in order to provide integrated search mechanism improved SLR based Federated search tool. In this paper, the experimental evaluation of this approach is presented and its showing the efficiency of the proposed method like Google is well aware that the current search engine to compare its performance against.

Keywords— *Federated Search and Systematic Literature Review, Crawler, evidence based research, Google, Search process.*

I. Introduction

Federated search systems alternatively called meta search systems aim to search a collection of databases from one interface and present one set of results, thereby reducing the amount of time and energy that a researcher must invest in learning and using individual database interfaces[7]. Are arranged as federated search, perception, discovery is an ideal way to make the process easier, in practice they often idling, less advanced search refinement, and poor results from several sources, including the integration suffers from some weaknesses [6] [8]. Many problems stem primarily from a lack of consistencies between database systems. However, despite such common weaknesses, federated search systems can provide a relatively quick and simple mechanism for conducting a broad search of multiple resources in one step [9].

In this paper, the performance of federated search engine is aiming to improve in terms of time, reliability, and accuracy of returning results. This can be done by improving the process of Systematic Literature Review (SLR) in software engineering. The method that has been more widely used for different analysis in the field of software engineering is Systematic Literature Review (SLR). "A systematic literature review & evaluate all available research relevant to a particular research question, subject area, or a means of interpreting the phenomenon of interest. Systematic review," Trustworthy rigorous and auditable method using is to present a fair assessment of a research topic. The presenting approach is focusing to improve the SLR process and by then automated federated search tool. In the next sections, the second Federated search tool more than literature reviews, the third point SLR Limitations, section IV introduced the proposed framework for Federated search SRL, presenting the results and discussion section V and section VI presenting Conclusion and Future Work .

II. Literature Review

In this section, the time will taken to review federated search method. In many places, the end user federated search multiple databases stored by different companies (sometimes broadcast search, distributed search, cross- search & met searching is known as parallel search) is a relatively recent development. For this several discrete communication service is possible, proprietary database when multiple database search with a search of the concept goes back at least to 1966 [6]. However, in contrast to the databases searched by current federated search products, the Dialog databases were (1) stored by a single company in a single location and (2) usually searched for an end-user by a librarian due to both the fee structure and the proprietary command-driven nature of the search interface. Roger K. Summit's article of 1971 on Dialog's user interface and Stanley Elman's various articles on the cost-benefit of Dialog examined this forerunner to federated searching [7].

Desirability and / or the difficulty of making a strong federated search tool, on one or more specific federated search implementation (2) report, (1) issue: The majority of today's

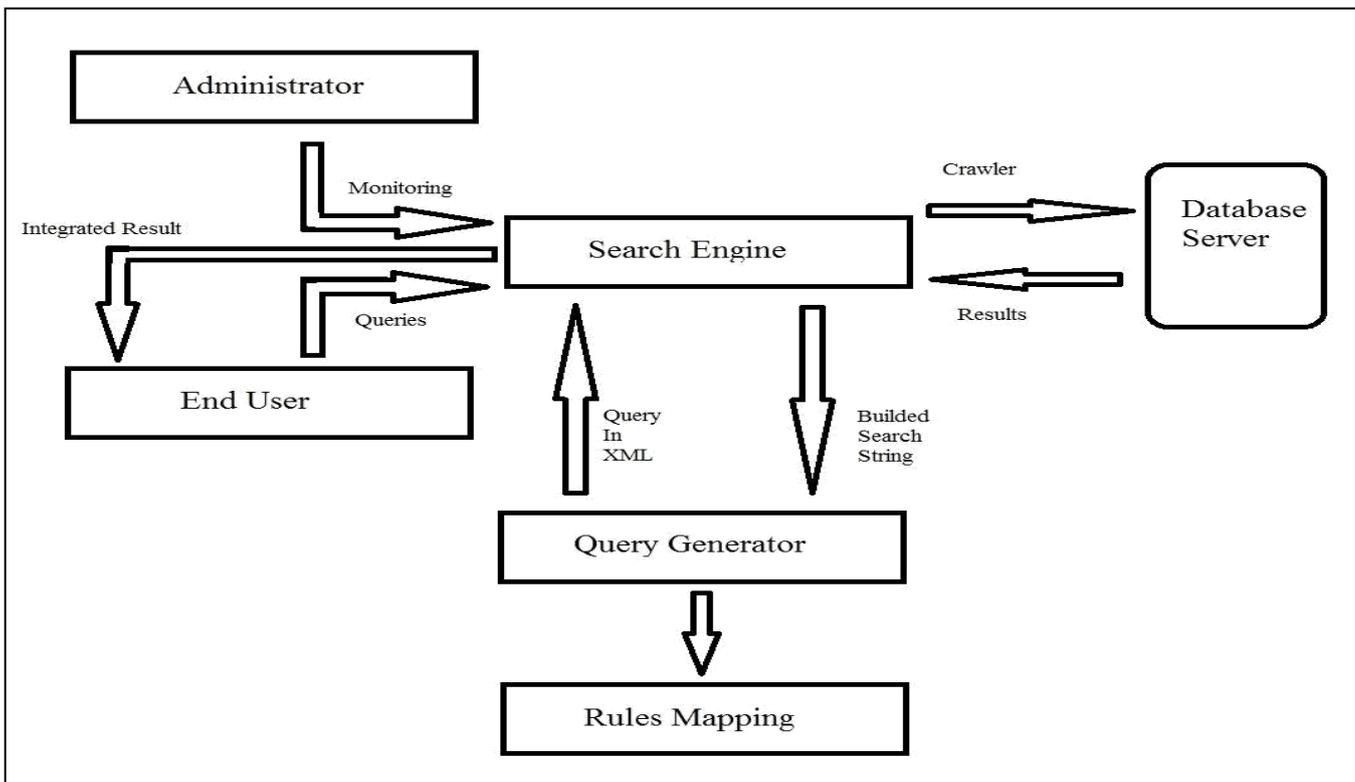


Fig. 1. Proposed System Architecture for Federated Search

federated search technology articles about four fall into the categories currently on the market for a second and / or Google Scholar, or a subject about implementing federated search tool (4) sequences of the federated search products) comparisons. These articles are theoretical, actual, or are comparable, because they contain little data based on quantitative research [8].

Literature about federated search clears too many and include reasonable assumptions. Serials Review column, "A Bow assumptions about literature, including federated search, is very clear and reasonable. Serials Review column, " One Box Challenge: The advantage is that the federal discovery process, " Allen provides a federal search Scherlen edited by contributions from five academic library, Federated search with the assumptions made about the recent example's challenge, with contributions from five academic library, edited by Alan Scherlen Research process, " assumptions about the benefits of federated search provides a recent example. Introduction editorial column states? Federal find some aspects of the research will certainly easier , but it would make it better , " Hampton contributor to Marian , "met searching an advantage is clear - a simple interface for multiple sources " money Pugh " ' e Zsearch to search multiple databases at once provides a quick and easy way . ' " writes Frank Cervone , " as simple as point of federated search is to search in : West Virginia University of federated search " minimum conditions " quote possible ... "federated search , then , research easier , providing a simple interface is assumed to require minimal training [9].

It is currently implemented as others Federated search. These problems pointed to the problems inherent with all citations can be seen before the slower unavailable to de-duplication true ranking relevancy database is impossible and true, waiting for results back. Rochkind (2007) comments: "the current library met search usually request at the point, in parallel, to seek multiple source repositories at once depends on and then merge the results." the election created a federated Federator's weaknesses to vendors all stem search this way. If libraries compiled and indexed the metadata from all the third-party database subscriptions and made the data searchable, true de-duplication and relevancy ranking would be possible. Because the system to deliver results on different database vendors will not have to wait in addition, results can be returned very rapidly.

Imperfect as it is, federated search is still significant potential to save time by requiring fewer benefits. It also automatically the appropriate database sailaitas that a federated search engine would not have been found otherwise students they will likely find the contents, serendipitous discovery. Students are searching for a particular subject advantage of databases that may not be aware. Graduates are satisfied with their information needs, from traditional search to Favorites and if our study, a federated search using saved a lot of time by determining how to test these assumptions yielded a federated high compared to low-quality or non-Federated search results searching [9].

III. Limitations

All four federated search portals such as science listed below all available resources, relevance, topic/author/date/etc. to rank search results by one side of the menu with advanced search forms as very similar interface and features. Groups, search, email alerts and to make choices, and opens directly into the search results that are the result of the selected third-party sites.

- For each question, related documentation criteria most likely are a subset of the collection: Although the concept of federated search portals still such as major challenges is suffering from below. This collection contains a selection problem.
- Select the appropriate collection; Federated search systems content of each collection requires some knowledge about how to represent a problematic store.
- Returned results for the user selected from the collection prior to the final presentation are merged. The final step is to merge the resulting problem.

IV. Proposed Federated Search Method

A. Problem Definition

There are two main issues exist in search process is reviewed:

1) The database search results are directly dependent on search strings, and search strings are designed to be adequate care. In addition, some databases do not support e.g. ACM combination of fields, and many areas within the search string in order to specify, based on the search order should be used.

2) The search string, for example, ACM Springer Link offer cannot be used in each database has its own search string structure. Consequently, the search string must be constructed manually, one by one and you want to be checked within the database. ACM cannot be used in Springer Link. Consequently, the search string must be manually constructed one after another within the desired database to be checked.

Driven by these issues, the discovery process will be tedious and prone to error perseverance. Any loss of data is also important to provide a more reliable search process, can cause bias in the research. Organized and SE have been reported despite the extensive number of SLRs, community still like the lack of a mechanism which mitigates these problems.

B. Proposed Approach

Therefore to overcome above stated limitations, in this paper the new federated search approach is presented which will improve the overall performance of federated search engine. Below is our proposed framework which is an extension of recently presented tool in [1]. For this approach, you can also refer tables of well know databases, define in the

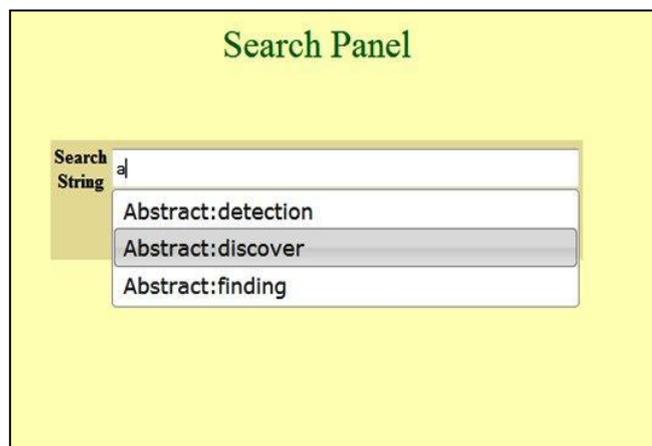


Fig. 2. Search Panel for End users

paper [1]. This approach is briefly summarized in the above architecture.

- **Admin Panel:** The system administrator can define some research categories and subcategories which will be managed by their own domain experts. These people are responsible for defining available research keywords and their corresponding synonyms for each category. They also present potential biases in the research process; students may be directed to stop the search process. For example, the software reconfiguration of software development is a subcategory, such as domain experts Runtime defines certain keywords, automatically changes such as restructuring, development will be included in searches for synonyms. These keywords keyword "Options" are stored in a database called.
- **Search Panel:** Provides an interface to the database in order to search within this particular advanced search. This is an area code in some offers. Full text, title, abstract, and dates as well as the Boolean operator to combine searched words using the ability to search through a limited set of features, the ability to search. Search criteria are specified by the user; they are sent to the query generator.
- **Query Generator:** This part corresponds to the database are selected by the user who is responsible for generating search strings. Questions governance code in order to build a similar structure XML files. These rules for the target database within a sea CH string are used to map each database field called uses. Here, the context of discovery is also considered synonyms available for each keyword. Finally, these questions are being deployed to the target database are designed to IEEE, ACM, IEEEExplore, and etc.
- **Crawler:** As soon as the queries are prepared, they are sent to the target databases. Then, a crawler starts downloading the search results provided by each database. Finally, all data are gathered into a single database.

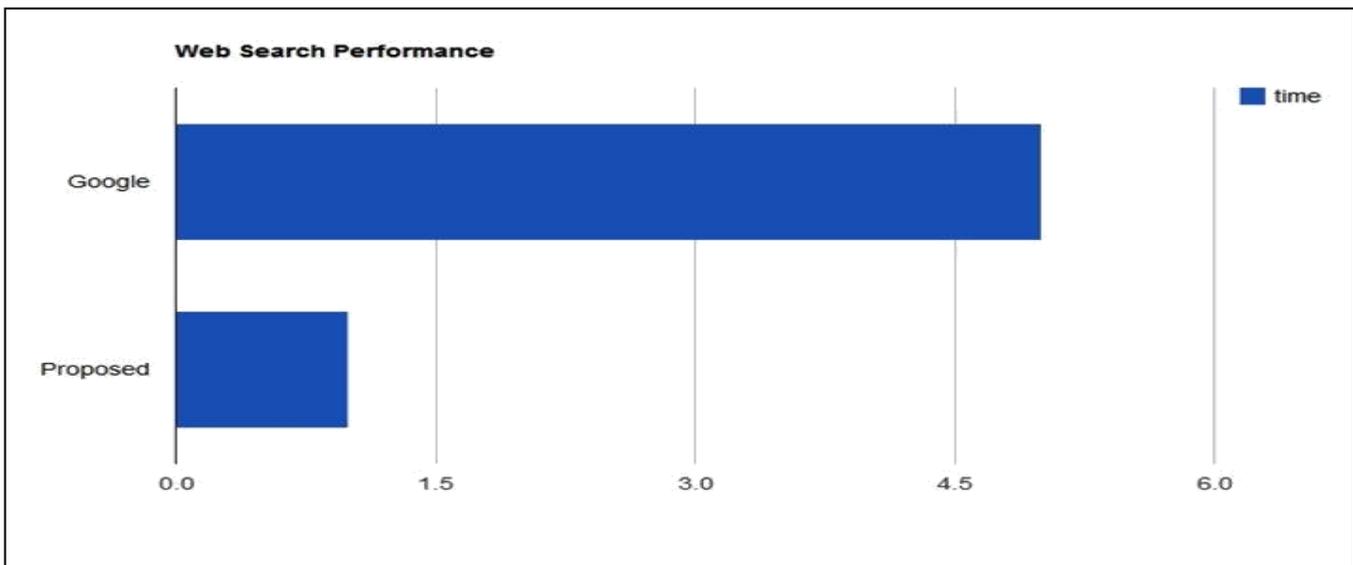


Fig. 3. Performance Analysis in between proposed approach and Google

v. Results and Discussion

In this section, the practical work done for above proposed approach is discussed. Above Fig. 2 and 3 are showing the same.

The performance of our proposed federated approach for SLR against the existing performance of Google is presented in Fig. 3. It is shown in this result that proposed approach is faster and reliable as compared existing methods.

vi. Conclusion and Future Work

In this paper, our main aim was to present the automated and efficient federated search engine in order to overcome the limitations existing federated search methods. In the paper, an introduction and literature review of federated search is presented. In addition to this, some of the federated search portals along with their limitations is discussed. The federated search method is presented in this paper and evaluated practically, as well. This approach is presented in order to facilitate the SLR search process. This proposed method is resulted into bridge the gap between the spread of databases in SE as well as integrated search required by SLR. From the results, it is clear that this proposed method is more robust and efficient as compared to exist methods. For the future work, real environment deployment of this method can be suggest and also evaluates its performances.

References

[1] Mohammad Ghafari, Mortaza Saleh, Touraj Ebrahimi, "A Federated Search Approach To Facilitate Systematic Literature Review In

Software Engineering", *International Journal of Software Engineering & Applications (IJSEA)*, Vol.3, No.2, March 2012.

[2] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3," Keele University and University of Durham, Technical report EBSE-2007-01, 2007.

[3] M. Petticrew and H. Roberts, *Systematic reviews in the social sciences: a practical guide*. Malden, MA: Blackwell Pub., 2005.

[4] T. Dybå and T. Dingsøyr, "Strength of evidence in systematic reviews in software engineering," 2008, pp. 178–187.

[5] B. Kitchenham, "Procedures for Performing Systematic Reviews," Software Engineering Group; National ICT Australia Ltd., Keele; Everleigh, Technical report Keele University Technical Report TR/SE-0401; NICTA Technical Report 0400011T.1, 2004.

[6] C. Jeffrey Belliston, Jared L. Howland, and Brian C. Roberts, "Undergraduate Use of Federated Searching: A Survey of Preferences and Perceptions of Value-Added Functionality," *College & Research Libraries* 68, no. 6 (2007): 472–87.

[7] Electronic Publishing Initiative at Columbia, "The Electronic Publishing Initiative at Columbia (EPIC) Online Survey of College Students: Executive Summary," September 2004, www.epic.columbia.edu/eval/eval04frame.html (accessed Dec. 12, 2008).

[8] Jillian R. Griffiths and Peter Brophy, "Student Searching Behavior and the Web: Use of Academic Resources and Google," *Library Trends* 53, no. 4 (2005): 545.

[9] Helen Laurence and William Miller, *Academic Research on the Internet: Options for Scholars and Libraries* (New York: Haworth, 2000): 1.

[10] Budgen D., Bailey J., Turner M., Kitchenham B., Brereton P., Charters S.; "Cross-domain investigation of empirical practices", *Institution of Engineering and Technology*, Vol. 3, pp. 410-21, 2009.

[11] MacDonnell S., Sheppard M., Kitchenham B., Mendes E.; "How reliable are systematic reviews in empirical software engineering?", *IEEE Transactions on Software Engineering*, Vol. 38, pp. 676-687, 2009.

[12] Silva F., Santos A., Soares S., França A., Monteiro C., Maciel F.; "Six years of systematic literature reviews in software engineering: An updated tertiary study", *Information and Software Technology*, Vol. 53, pp. 899–913, 2011.

[13] T. Dyba, T. Dingsøyr, "Strength of evidence in systematic reviews in software engineering", *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, pp. 178-187, 2008.