

# **An Overview of the Relevance Ranking Based Open Source Resource Discovery Interface (VuFind) for Library and Information Services**

**Gautam Mukhopadhyay**

*Librarian, Chandrapur College, Burdwan, West Bengal, India.*

## **Abstract**

This article gives an overview of the resource discovery interface tool with special reference to VuFind. It describes the meaning of some essential components (for example, Discovery Layer, Relevance Ranking) for considering an open source resource discovery interface as an outright replacement of traditional Online Public Access Catalogue (OPAC). It also highlights very briefly the key features of relevance recommendation supported discovery services e.g. EBSCO, Worldcat etc.

**Keyword(s)** Relevance Ranking, Discovery Interface, VuFind

**Abbreviation(s)** RDF (Resource Description Framework), API (Application Programming Interface), ILS (Integrated Library System)

## **1. INTRODUCTION**

The growing needs of library materials have changed the mode of access and retrieval of information resources. Various library activities on the web environment may have a trend to replace the old ones. Global issues and prevalent situations have expedited rapidly the user expectations. Information seekers are not happy with the query results of the existing search databases. Librarians and information providers browse many databases for their information. Library and information scientists are now exploring

information by developing the searching tools for retrieving information having relevance based ordering among the search results. They attempt to design an interface tool with which multiple database systems can be integrated and users' community can interact with prevailing systems. Online Public Access Catalogue (OPAC) is being replaced with a newly evolved discovery service. Information researchers are to combine MARC 21 format of metadata with newly structured metadata formats like RDF metadata model. The mode of access has been changed dramatically with change of type and form of resources. Now, Open Source Software e.g. VuFind ensures web interface and accessibility standards and allows to integrate multiple library systems with its powerful centralized indexing mechanism. This paper presents VuFind as a discovery interface tool with its key feature relevance ranking based faceted searching system.

## **2. OBJECTIVES**

The main objectives of this study are:

- to define some inevitable components in terms of relevance ranking search in discovery interface tool,
- to give a basic idea on the community supported flexible search engine VuFind, and
- to assess whether a discovery interface is a total replacement to the existing OPAC or not,

## **3. SEARCHING, BROWSING AND DISCOVERY**

Searching means looking over closely or thoroughly to find something such as to examine specific information. Normally, we perform searching if we know what we are looking for. Search engines retrieve meaningful results but a user can never know what other things he/she missed out on. Especially when we are passing through the knowledge era where generation of information resources is growing exponentially. Search is generally approached to a definite and goal oriented track towards foreordained query results.

Searching refers to looking for known terms while browsing is to look over casually through an aggregate of things for topic of interest or to access a network by a browser. Browsing denotes scanning over the world wide web via links from one site to another or one page to another.

Discovery usually denotes "the disclosure of pertinent facts or documents..." Discovery is much more seamless and allows content recommendation. Discovery searching retrieves relatively new content or new results.

#### **4. CONCEPTS OF DISCOVERY LAYER, DISCOVERY INTERFACE AND DISCOVERY SERVICE**

Discovery layer is a software or third party component for libraries while a layer indicates multi-tier architecture or a set of linking technology which includes logical structuring mechanism for the components that construct the software solution. This is actually a unified and more structured view of outputs via search interface for users' community to search for information. Discovery layer provides a web search engine that can interact with a searchable metadata index of library holdings.

Discovery interface is a software to software interacting tool whereas interface acts as a software intermediary. It is a web based search interface which is a key to an integrated library system (ILS) that holds metadata of different formats such as MARC, XML and RDF. This web search engine not only allows merely a library catalogue or a single library database but it may index metadata for library users of multiple ILSs or library catalogues, websites, web contents and digital repositories.

A discovery service is an online searching interface that furnishes a seamless interaction tool for searching both local collection and web resources. Such as EBSCO Discovery Service. This is an all inclusive or complete search solution that makes thorough research more flexible and powerful.

#### **5. RELEVANCE RANKING SEARCH**

This is a search mechanism to query for information and the query results essentially include two main features, one is relevancy and the other is ranking. This is a process of searching which determines the possibility of relevance between the search terms and results. Here those resources are likely to be most relevant will come first in the query results. The possibility of achieving the good ordering of search results depends only on how far a text of a web document matches a query term. Search result relevance mainly denotes the how relevant a content of a web page is in terms of a specific query topic.

#### **6. KEY FEATURES OF DISCOVERY SERVICES**

Discovery services are also known as web-scale discovery services. The web-scale discovery services started emerging from 2009 (Sonawane, 2017). The discovery services brought a revolution in library and information systems and services. This provides a combined ability to perform a search among the vast resources subscribed by a library (Thompson, 2014). The index based discovery services include Primo, EBSCO Discovery Services, OCLC Worldcat, Summon etc. There are federated search engines which include Knimbus and WebFeat. Among the many key features of discovery services, the central index is worth mentioning. Central index denotes the metadata as well as the full text content converted into a single unified search (Bowen, 2008). EDS allows the content recommendation with greater relevant results to a query of researchers. Single search box, relevance ranking and facets are the other key

features of the central index based searching.

Summon unified indexing can be used in a comprehensive content indexed by a library system to produce better results from a single search. Relevance ranking can be applied in a simple or direct content neutral way with indexed databases optimized for discovery. The combination of dynamic (shows how well the search matches a record) and static rank (shows the importance of each record's characteristics to relevancy overall) represents that the most relevant results always come first for superior user-friendly searching.

EBSCO Discovery Service provides the researchers with the most relevant results for their search queries. EDS also ensures a powerful search strategy across the library's rich and diverse collection of information resources. With the robust central indexing, EBSCO discovery service API offers the library's ILS interface. Through the library's user interface user can access multiple integrated database systems with the local content of full-text information resources.

Library collections and databases are accessible by OCLC Worldcat Discovery Service via remote access from a subscriber's site. These databases are also available through a centrally indexed metadata from various database providers. Worldcat first queries its central index and then displays results matching the search terms. The relevance algorithm is applied to a single combined set of records that is a combination of centrally indexed databases and the remote databases. Then ranking is performed in the discovery tool for returning search results.

ExLibris Primo library discovery service offers a personalised ranking which gives a different ranking of results for the same query searching according to the topic of interest of the users. It can be applied to a diverse aspect of disciplines. It ensures an adjusted ranking for local collections with a greater flexibility.

## **7. DISCOVERY INTERFACE TOOL (VuFind)**

VuFind is a community supported highly customisable, configurable and extensible open source discovery interface. "Typically referred to as a 'discovery layer', it provides a single point of access to the library's local and licensed collections while providing a seamless, consistent, and locally customizable user experience. VuFind combines easy, convenient access to the scholarly materials that users need with a modern, feature-rich interface that they have come to expect." It is a flexible search engine originally developed by Villanova University in 2010. It is written in PHP codebase and backed by Apache Solr centralized indexing and also Laminas components. VuFind is structured on API based powerful discovery services such as EBSCO Discovery Service, OCLC Worldcat and Summon. VuFind supports third party systems such as Refworks, EndNote etc. and allows importing and exporting of varied metadata content formats like MARC, XML etc. In fact, this discovery interface facilitates citation management just as zotero reference management software. It can be integrated with multiple database systems such as Koha, Alma and FOLIO. VuFind offers a unified central indexing and allows single platform access to the repository of institutional

resources and beyond. What is utmost important in VuFind tool is that it ensures powerful community interaction support through which easy users' participation can be enabled.

## **8. TRADITIONAL SEARCHING VS. RELEVANCE RANKING SEARCH**

Search technique of looking for terms on traditional OPAC is based on exact title and subject headings including an alphabetical list of results. "Relevance ranking usually is best for searches that are not "either/or" types of searches. For example, in most traditional title searches, the result is either the library has the books, or it does not...To better understand relevance rankings, a very simplified description of the typical process for assigning relevance ranking to a web page is in order. In this example, the program that is being used examines the contents of the page, going through the text, counting how many times a word appears on that page. The ranking program then takes that information and uses that word count as a relevance value. This value is an attempt to measure how relevant the page is if that word appears in a search query. It stores the word, the relevance value generated, and the URL of the page in the search engine's index. This process is repeated over and over as more pages are found" (Childers, 2006). Traditional OPAC does not allow searching across the multiple databases whereas a single search interface in the discovery interface tool like VuFind allows combined database searching. Traditional OPAC does not enable relevance ranking, conversely "alphabetizing makes for easy lookups, but ranking is better for human interest" (White, 2007). Hence, the present day library and information centres felt necessary to integrate web search engine technology into their existing library catalogues via discovery layer (backed up by the toolkit e.g. Apache Solr, Apache Tika, APIs) so as to make OPAC more interactive and user-friendly.

As the discovery layer covers a remarkably greater scope than a traditional search for library catalogue, it can be treated as a total replacement to an existing OPAC. The leading open source resource discovery interface tool Vufind makes this replacement possible using the powerful search engine Apache Solr and a comprehensive discovery product ExLibris Primo.

## **9. CONCLUSION**

Varied requirements of the library users in the diverse fields of study accelerated the rapid progress of search engine technology. Researchers usually do not prefer to move from one database to another for their specific need of information. They need to have relevant and reliable resource materials for their topic of interest within the shortest possible time. Discovery interface tools provide an integrated multiple or combined library databases with relevancy ranked results via a centralised indexing system for satisfying precise needs of information users. Despite some challenges and threats, contemporary project planning and continuous updation of resource discovery software for better community support ensures integration of the existing library database system

with the robust interactive online search interface mechanism "web-scale discovery services".

## REFERENCES

1. Sonawane, CS. (2017). Library discovery system: an integrated approach to resource discovery. *Informatics Studies*, 4(3), 27-38.
2. Thompson, J. (2014). Implementing web-scale discovery services: a practical guide for librarians (Practical Guides for Librarians Series, 9). Lanham, MD: Rowman & Littlefield.
3. Bowen, J. (2008). Metadata to support next-generation library resource discovery: lessons from the extensible catalog, Phase 1. *Information Technology and Libraries*, 27(2), 6-19.
4. Childers, S. (2006). An introduction to relevance ranking systems. *Faculty Publications, UNL Libraries*, 117.
5. White, HD. (2007). Combining bibliometrics, information retrieval, and relevance ranking, Part 2: some implications of information science. *Journal of the American Society for Information Science and Technology*, 58(4), 583-605.
6. <https://foss4lib.org/decision-support/discovery-layer-ssm/definition>
7. <https://ebSCO.com/e/products-and-services/saas-and-discovery/eds>
8. <https://exlibrisgroup.com/products/summon-library-discovery/relevance-ranking/>
9. <https://oclc.org/en/worldcat-discovery/features.html>
10. <https://exlibrisgroup.com/products/primo-discovery-service/relevance-ranking/>
11. <https://vufind.org/docs/vufind-os-community-2020.pptx>