DESIGNING THE INTELLIGENT LIBRARY SEARCH SERVICE

FEDERATED SEARCHING AND DISCOVERY SERVICES

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, SEMANTICS AND LIBRARY APP

-KAUSTUBH MUKHOPADHYAY,

SR. ENGINEER, EBSCO

WHAT IS FEDERATED SEARCHING?

- A search mechanism that retrieves information from multiple resources simultaneously.
- A query is made by user which is then passed on to the respective search engines of individual resources.
- Each search engine makes the same query on its own database/s.
- Top results from all these individual searches are aggregated and brought to display for end user.
- There are various ways by which the federated search connects with the individual platform searches. For e.g. through APIs, RSS feeds, MySQL queries etc.
- Federated Searches operate real time and depend on the resources they are interacting with for data storage and retrieval, they themselves do not store data anywhere

HOW DOES IT WORK?

USER INTERFACE

Gets user query, Displays end results

Request

FEDERATED SEARCH BACK END

Passes on query to all resources, aggregates results from each and sends to UI

Response

RESOURCE 1

API/RSS FEED ETC. RETURNS RESULTS ON THE BASIS OF QUERY

RESOURCE 2

API/RSS FEED ETC. RETURNS RESULTS ON THE BASIS OF QUERY

RESOURCE 3

AND SO ON…

ADVANTAGES AND DISADVANTAGES

What you gain

- ○+ A single interface to search multiple platforms
- ○+ Usually a low cost solution, can be implemented using open source technologies as well
- ○+ Good solution if low number of resources are to be searched
- ○+ Real time, so mostly gives the latest and updated results

What you lose

- Complete dependency on individual platforms' live API/RSS feeds (won't work if any one goes down)
- Not a good solution for large number of resources, too many request/responses, delays
- Data sources may provide different formats, deciding relevancy of each platform live is a challenge
- -Result list populates based on time taken to fetch the results not necessarily relevance

WHAT IS DISCOVERY SERVICE?

- A search tool that enables users to search multiple resources through a single interface and a common database
- The main difference between a federated search and discovery service is that a discovery service would have its own knowledgebase comprising of pre harvested results from all the different resources
- Discovery Services usually gather metadata from different resources and create a unified index
- Inbuilt algorithms to match parameters and **create uniformity across different kinds of metadata formats**
- Once users enter search query, the entire index is searched to give the most relevant results
- Discovery Services depend on the willingness of the individual resources to share their entire metadata and let it be pre harvested

HOW DOES IT WORK?

User Interface

Search query, Display results

Discovery

Back end queries central index and brings out results based on relevance ranking algorithm

Central Index

Data harvested from all resources are indexed

- Resource 1
- Resource 2
- Resource 3
- Resource 4

ADVANTAGES AND DISADVANTAGES

What you gain

○+ A single interface to search multiple platforms

○+ Data searched across a central index, no dependency on individual platforms real time, fast and reliable

○+ Having a central index enables the ability to get relevant results when searched,

○+ A good solution if large number of platforms/resources are to be searched

- Creating a central index is usually a huge task, a good discovery service usually may have some cost involved

- A central index may be difficult to maintain and updating all resources may be a challenge

- Not all resources/platforms may give instant updates to central index which may cause some delays

What you lose

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN THESE SERVICES

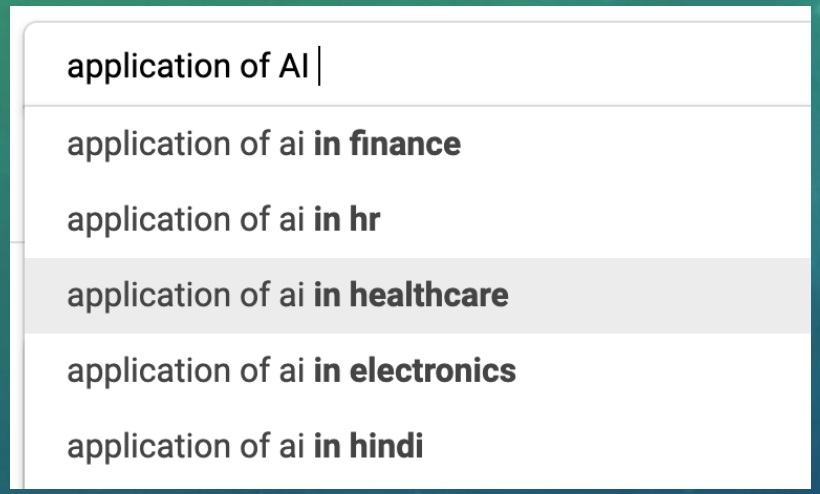
- Artificial intelligence aims towards making a system smart/intelligent and to be able to perform cognitive functions that can be associated with the human mind.
- Machine Learning is a part of AI and involves the mechanism of the system training on it's own and learning how to give better output.

Application in Discovery and Federated Searches

- It is important for Discovery and Federated Searches to be intelligent enough to understand the user intention while searching. The system should be able to provide features like autocomplete, Did you mean, auto correct and present related search subject areas so users can filter their results better.
- The system should be able to learn and understand individual search preferences, store them and provide recommendations, alerts to users based on their search preferences . This is something that is still being worked on for most Discovery and Federated Searches.
- Discovery systems should be capable of maintaining the central index properly and be able to normalize data efficiently from different resources without having much data loss while making sure that the relevancy algorithm is optimally performing. They should be able to classify and put records in appropriate categories.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN THESE SERVICES - EXAMPLES

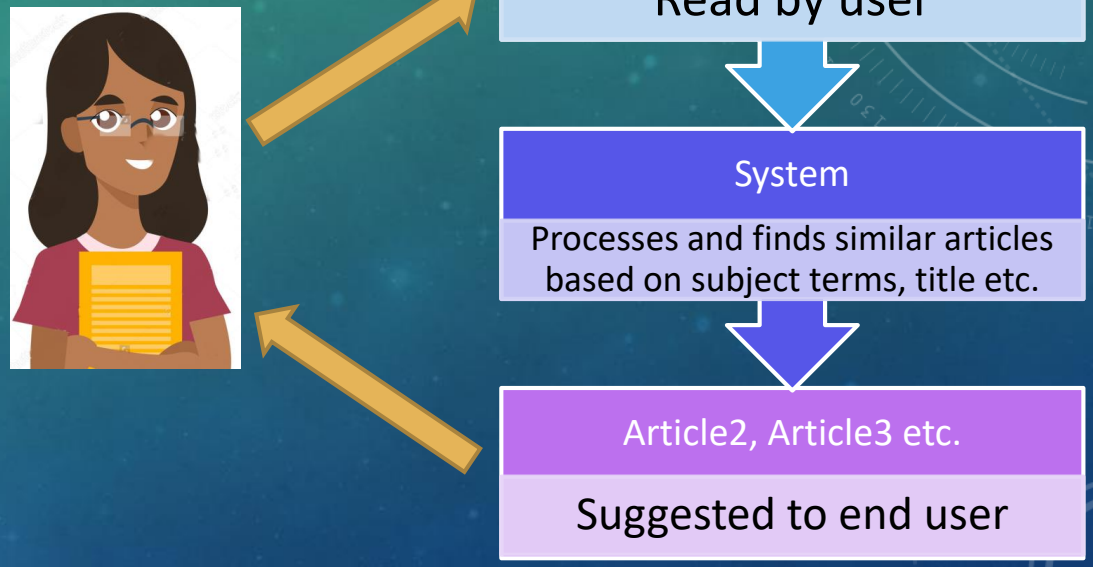
Autocomplete (Predicting user intention)



- Predicting User intention is a key feature in modern day Search Engines
- Most Autocomplete features try to finish the keywords based on algorithms to match spelling of the words
- More advanced versions try to predict based on previous user searches
- One good feature would be to have autocompletes based on the subject area of the user
- Existing solutions include implementation of Autocomplete for the Journals Subscribed by library to locate them faster

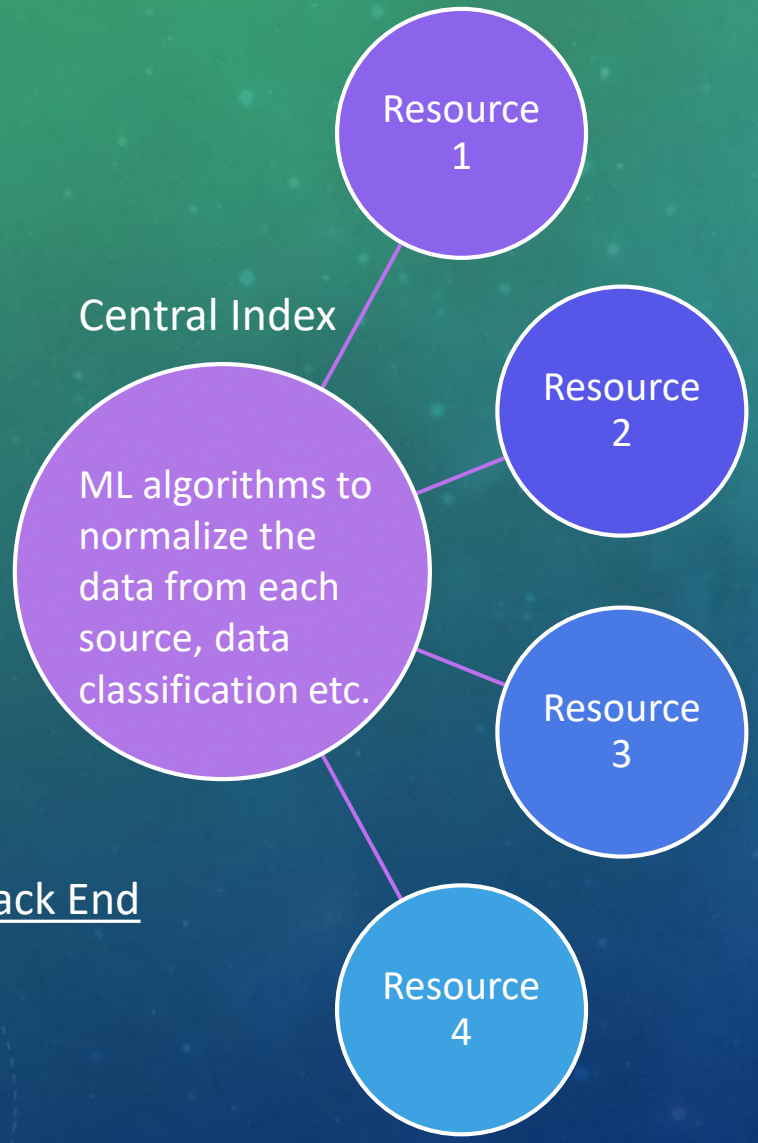
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN THESE SERVICES - EXAMPLES

- Recommending resources based on user interest/history is a good application of ML relevant for Discovery and Federated Searches
- System tries to learn from each search made by user as to what are the interest areas
- Based on the searches made it will try to find similar resources and suggest them to the end user at certain time intervals
- Any new content coming into the system can also be classified accordingly and suggested to relevant end user and sent as alerts

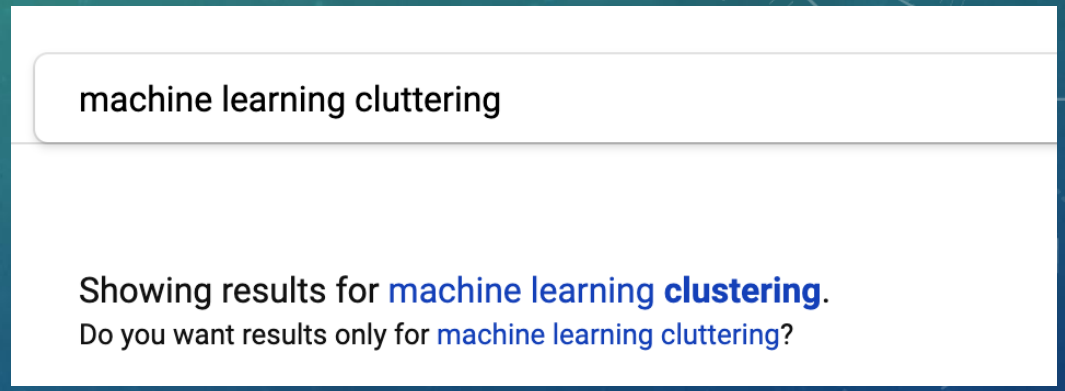


Recommendation System (Giving relevant suggestions to end users)

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN THESE SERVICES - EXAMPLES



Discovery Back End



Autocorrect (Understanding user intent)

SEMANTIC SEARCHING

- Semantic search aims at not only finding the keyword being searched for but also performs the search on the basis of the meaning, context of the search being made.
- This also involves Machine Learning where the system tries to learn from user's queries and provide more accurate results.
- For e.g., a search for 'Which is the highest mountain?' should ideally look for the answer i.e. 'Mount Everest' rather than search for keyword like 'What', 'mountain' etc.
- They should be able to map the abbreviated terms: for e.g. adhd should automatically map to Attention Deficit Hyperactivity Disorder and then search results accordingly rather than just 'adhd'.
- Some systems have already started incorporating such techniques, however, there is still plenty of work to be done in this field.

SEMANTIC SEARCHING - EXAMPLE EXPLAINED

User queries for phrases e.g. 'Which is the best journal on Machine Learning'

Semantic search instead of looking for the keywords will evaluate the meaning of the search

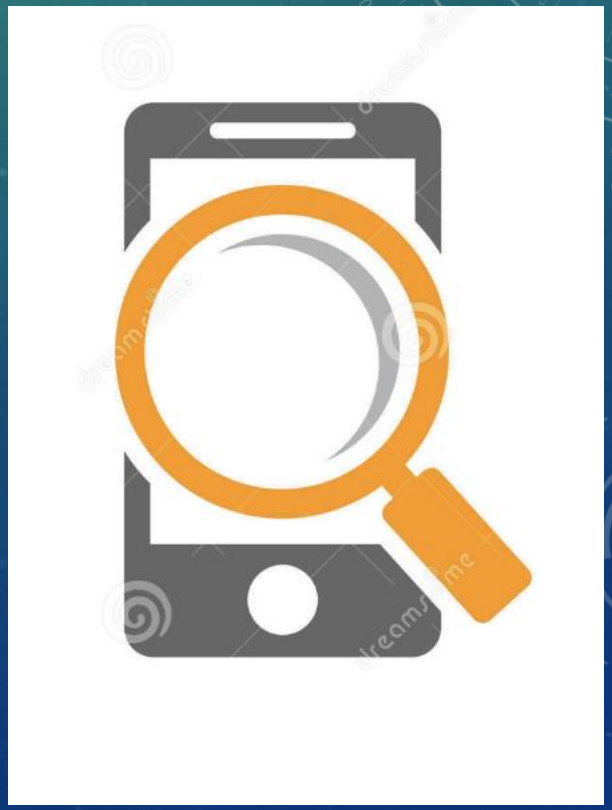
Results: Returns journal with highest relevancy in the area of Machine Learning

PUBLICATION FINDER (ATOZ)

- An application capable of building an AtoZ browse and search functionality, comprising of mainly journals, eBooks and other online resources.
- This is different from Discovery and Federated Search, since this is focused on building a listing of these resources and does not really go into depth of article level search.
- Usually built directly by obtaining list from respective library, not having too much dependency on the resource platforms.
- Useful in providing quick information to users on the different kinds of subscriptions available with the Library.
- Usually offered complementary to Discover/Federated Search services.

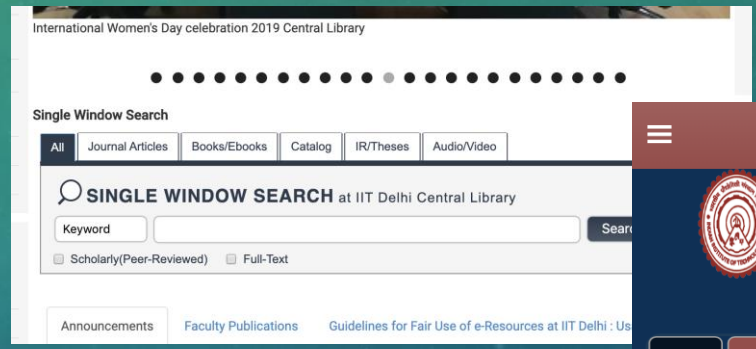
LIBRARY APP

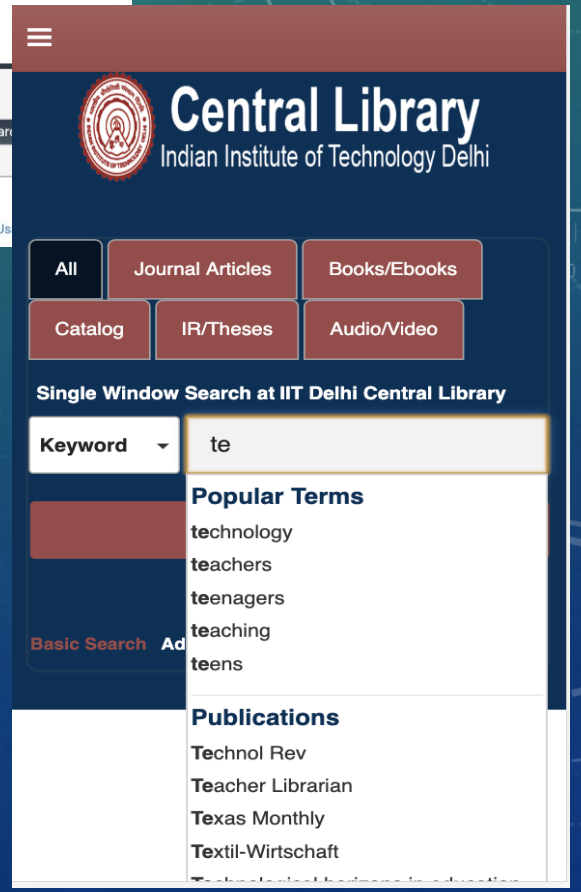
- Having a mobile based library app for these services is the need of the hour and helps the modern generation researchers stay connected with the Library and its resource subscriptions through mobile.
- It is also now good practice to make applications mobile compatible and hence important to make sure that these search services have responsive interfaces.
- Having an app is also useful for the library to be able to promote their own brand value and popularize these search services amongst the users.
- Overall promotes greater connectivity and helps in better utilization of resources subscribed by the library.



CASE STUDY: SINGLE WINDOW SEARCH AT IIT DELHI

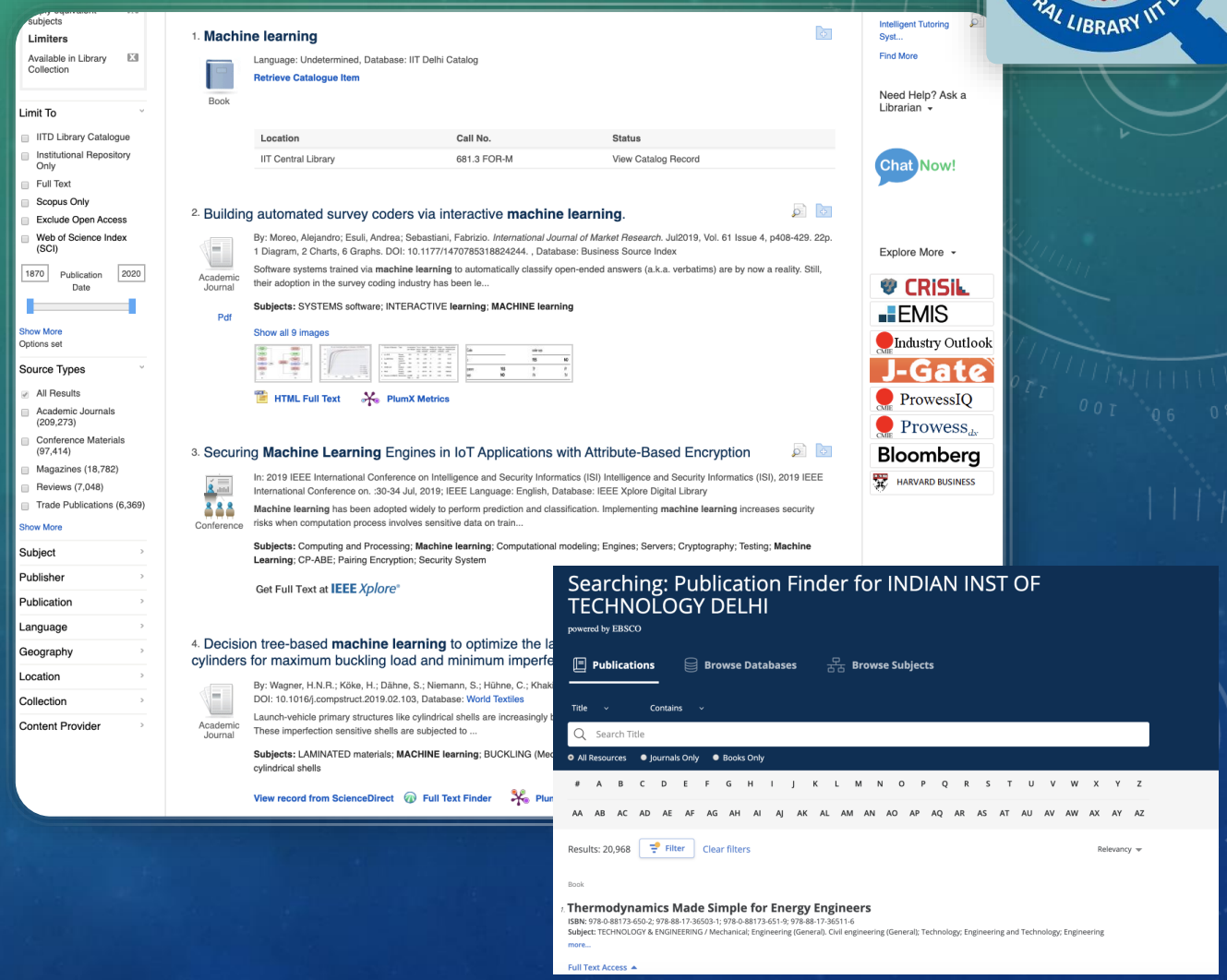
- One good example of a Discovery implementation is Single Window Search by IIT Delhi Central Library using the EBSCO Discovery Service. The Discovery service was implemented under the guidance of Dr. Nabi Hasan and team.
- SWS has been able to successfully index and integrate most of their content including: Catalog, Institutional Repository and Online Academic Journals, eBooks, etc.
- SWS has a dedicated mobile app and web interface and is clearly highlighted with a search box in the library website.
- The Discovery service has features like autocomplete and autocorrect, other features to improve functionalities are continuously being worked upon.
- Users can efficiently segregate results based on their choice and even set up alerts to know about latest content from a particular journal, interest area etc.





CASE STUDY: SINGLE WINDOW SEARCH AT IIT DELHI

- SWS is efficiently able to index and bring out relevant results from all different resources as can be seen in the screenshot (from Catalog, IEEE, ScienceDirect etc.)
- This leads to greater usage of unused resources or resources that many users may not be even aware of.
- SWS also has a dedicated link resolver (Full Text Finder) that tries to find the most applicable Full Text of articles from their metadata.
- An AtoZ Publication Finder is also available as part of the EBSCO Discovery Service.
- Dedicated Chat and Ask a Librarian features to provide users constant support.



Thank You