Course Name: Information Retrieval System

Course Code: 2CSDE53

Roll No. : 20BCE204

Practical 1

Aim: Study various information retrieval systems and prepare a write-up by comparing various information retrieval systems based on different characteristics. Prepare a report (writeup) with the following detail for each information retrieval system.

- Identify the domain of the given information retrieval system
- The data source
- Objective
- Characteristics
- Conclusion and Observation

1. Lucene
   Domain: Text search

   Data source: Various types of text-based documents, such as PDFs, Word documents, and HTML pages.

   Objective: To provide fast and efficient search capabilities for text-based data.

   Characteristics: Uses an inverted index to quickly locate relevant documents for a given query. Supports Boolean and proximity operators for advanced query formulation. Can be integrated with other applications to provide search functionality.

2. UpToDate
   Domain: Clinical decision support

Data source: A team of over 5,700 physician authors, editors, and peer reviewers.
Objective: To provide users with evidence-based, current clinical information.
Characteristics: Includes over 30 specialties, and more than 12,000 topics. Provides advanced search options and filters, such as by publication date, author, and journal. Allows for the creation of saved searches and alerts for new literature on a specific topic.

3. Wolfram Alpha

Domain: Knowledge engine
Data source: Curated and structured data from various sources
Objective: To provide accurate and comprehensive answers to a wide range of questions
Characteristics: Wolfram Alpha uses a proprietary algorithm that takes into account a wide range of structured data to provide answers to user queries. It can answer a wide variety of questions, including mathematical, scientific, historical, and geographic questions.

4. Solr

Domain: Enterprise search
Data source: Databases, data warehouses, and other structured data sources.
Objective: To provide a scalable and customizable search platform for enterprise applications.
Characteristics: Solr is built on the Apache Lucene search library, which allows for advanced search functionality such as faceted search, full-text search, and geospatial search. It also has a RESTful API that allows for easy integration with other applications.

5. Social Media Applications (Instagram,LinkedIn)

Domain: Entertainment, Educational
Data Source : This kind of application ,uses information which is based on user activities such as what kind of post are most reacted by the user, what kind of post are saved by user etc.
Objective: This application (IRS) 24/7 records the user activities and helps in searching the content and recommend the related content.
Characteristics : This application keeps user up to date regarding the latest trend (in Instagram) and keeps user informed regarding the opportunity available for him/her in corporate world (in LinkedIn).It also helps recruiting companies to know which employee can be employed to their company.

- Conclusion and Observation :-
  IRS in today's time have become need of the hour.
  • Today we have more information on websites
  • Also, many choices for a user call for relevant results.
  • User wants more personalized users.
  • So, to deal with choices, IRS becomes a necessity.
  • Also with the advancement in technology, it has become increasingly
  feasible for developers to exploit data in a useful manner by using IRS and
  other technology