**Course: Advanced Bioinformatics**

**Module title: Search Engines in Bioinformatics**

**Module no. : 174**

Search Engine Categories:

1. Which perform the query only in the fields of citations
2. Which perform the query in the full text article
3. Which further process retrieved citations to organize them and/or to retrieve further information.

2ndCateogry:

QUERTLE

eTBLAST: Search for citations; Medline & full-text articles; PubMed Central. Text-similarity search by comparing documents in target DB with input text. Finds documents best match keywords extracted from query by analyzing. It is based on the idea of word alignment.

**QUERTLE: It is a s**emantic search engine where queries are based on the meaning and the context of documents.

3rd Category:

GoPubMed:

* extracts Gene Ontology terms from the retrieved abstracts
* supplies the user with the relevant ontology for browsing
* Indexes PubMed search results with ontological background knowledge

XploreMed

* Filters PubMed results according to main MeSH categories
* Extracts topic keywords & their co occurrences, with the goal of extracting abstracts

EBIMed

* IR+ IE from Medline
* Analyzes retrieved Medline abstracts to highlight associations
* Results are displayed in tables, and all terms are linked to their entries in biomedical DBs.

iHOP

* Genes and proteins as hyperlinks b/w sentences and abstracts
* Information in PubMed to navigable resources
* Navigation along gene network allows a stepwise & controlled exploration of the information space

To find out scientific resources, as journals & conference proceedings

Systems have been developed to retrieve scientific publications

**CiteSeer:** well-known automatic generator of digital libraries of scientific literature

CiteSeer was a public search engine and digital library for scientific and academic papers, primarily in the fields of computer and information science that has been replaced by CiteSeerX. Many consider it to be the first academic paper search engine. It became public in 1998 and had many new features unavailable in academic search engines at that time. These included:

* Autonomous Citation Indexing automatically created a citation index that can be used for literature search and evaluation.
* Citation statistics and related documents were computed for all articles cited in the database, not just the indexed articles.
* Reference linking allowing browsing of the database using citation links.
* Citation context showed the context of citations to a given paper, allowing a researcher to quickly and easily see what other researchers have to say about an article of interest.
* Related documents were shown using citation and word based measures and an active and continuously updated bibliography is shown for each document.

1st Cateogry:

* RefMed :
* MedlineRanker
* iPubMed

RefMed: relevance feedback retrieval system for PubMed. Finding related articles from the PubMed (a large biomedical literature repository) is challenging because it is hard to express the user's specific relevance in the given query interface and a keyword query typically retrieves many results. Biomedical researchers spend a critical amount of time (e.g., often more than several days) in the literature search process. This paper proposes RefMed, a novel search system for PubMed, which supports relevance ranking by enabling relevance feedback on PubMed. RefMed first returns initial result documents for a user's keyword query as in PubMed. The user then makes relevance judgments on some of the resultant documents while browsing them. Once the user "pushes the feedback", the system induces a relevance function using RankSVM and ranks the results according to the function. To realize the ad-hoc relevance retrieval on PubMed, RefMed "tightly" integrates RankSVM within RDBMS and runs the rank learning and process on the fly with a response time of a few minutes.Our qualitative experiments with biomedical researchers show that RefMed substantially reduces the amount of effort required to search related PubMed articles. RefMed is accessible at "http://dm.postech.ac.kr/refmed".

The MedlineRanker web server is dedicated to scientists interested to rank the biomedical literature according to a selected topic. The query page allows to search for any biomedical topic. The web server is fast enough to process thousands of scientific abstracts from the PubMed database in few second. Search engines for Medline. Learns most discriminative words by comparing a set of abstracts provided by user with the Medline. Ranks abstracts according to learned discriminative words.

iPubMed: Search engine for Medline Implements the search-as-you-type paradigm. Has the main advantage to provide results on the fly, which allows users to dynamically modify their query.