



The comprehensive study of Information Retrieval System

¹ Mr. J. Albert Irudaya Raj, ²Dr. R. Kadher Farook, ³Mr. R. A. Vinoth Kumar,

¹Assistant Professor, Department of Information Technology & Management
Arul Anandar College, Karumathur, Madurai – 625 514.

² Assistant Professor, Department of Information Technology & Management
Arul Anandar College, Karumathur, Madurai – 625 514.

¹Assistant Professor, Department of Information Technology & Management
Arul Anandar College, Karumathur, Madurai – 625 514.

Abstract

Information retrieval (IR) is the field of computer science that deals with the processing of documents containing free text, so that they can be rapidly retrieved based on keywords specified in a user's query. Information retrieval is the science of searching for information in a document, searching for documents themselves, and also searching for the metadata that describes data, and for databases of texts, images or sounds.

Index Terms: Learning Information Retrieval, Intelligent Search, IR types, Web IR, HTA.

I. INTRODUCTION

Information retrieval is **intended to support people who are actively seeking or searching for information**, as in Internet searching. Information retrieval typically assumes a static or relatively static database against which people search.

Advanced internet technologies providing services like e-mail, social networking, online banking, online shopping etc., have made day-to-day activities simple and convenient. Increasing dependency on the internet, convenience, and decreasing cost of electronic devices have resulted in frequent use of online services. However, increased indulgence over the internet has also accelerated the pace of digital crimes. The increase in number and complexity of digital crimes has caught the attention of forensic investigators.

A. FEATURES OF INFORMATION RETRIEVAL:

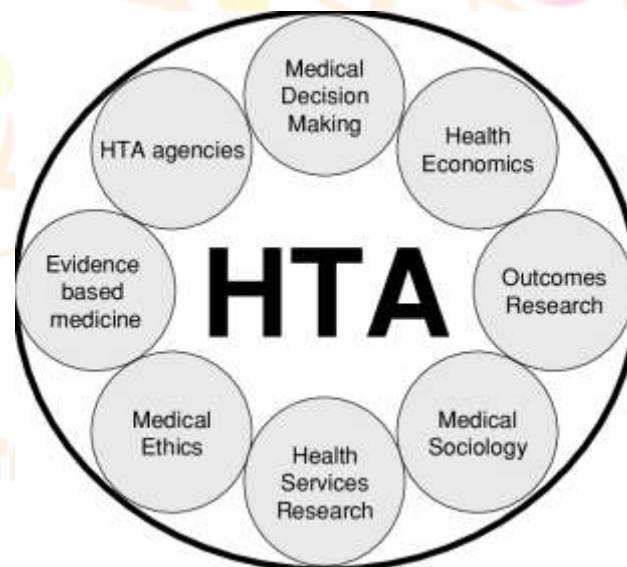
Twelve other characteristics of IR models are identified: search intermediary, domain knowledge, relevance feedback, natural language interface, graphical query language, conceptual queries, full-text IR, field searching, fuzzy queries, hypertext integration, machine learning, and ranked output.

B. WEB INFORMATION RETRIEVAL

Web Information Retrieval models are ways of integrating many sources of evidence about documents, such as the links, the structure of the document, the actual content of the document, the quality of the document, etc. so that an effective Web search engine can be achieved. State-of-the-art approaches to retrieving information employ two generic techniques: (1) matching words in the query against the database index (key-word searching) and (2) traversing the database with the aid of hypertext or hypermedia links.

II. Literature Review – 1 : MEDICAL INFORMATION RETREIVAL

Information retrieval (IR) in **Health Technology Assessment** (HTA) calls for transparency and reproducibility, but common practice in the documentation and presentation of this process is inadequate in fulfilling this demand. The main objective is to promote good IR practice by presenting the conceptualization of retrieval and transcription readable to non-information specialists, and reporting of effectively processed search strategies. This kind of information is usually housed in database management systems and contains several types of resources like images (from radiology, pathology and other areas), genomics (gene sequencing, protein characterization and others), citations (which link scientific literature) and Evidence Based Medicine (EBM) resources.



Methods: Here used comprehensive database search (04/2010) to synthesize the current state-of-the-art. We then developed graphical and tabular presentation methods and tested their feasibility on existing research questions and defined recommendations.

Results: No generally accepted standard of reporting of IR in HTA exists. We, therefore, developed templates for presenting the retrieval conceptualization, database selection, and additional hand-searching as well as for presenting search histories of complex and lengthy search strategies. No single template fits all conceptualizations, but some can be applied to most processes. Database interface providers report queries as entered, not as they are actually processed. The huge difference between entered and processed query is shown in “Details.” Quality control and evaluation of search strategies using a validated tool such as the PRESS checklist is suboptimal when only entry-query based search histories are applied.

Conclusions: Moving toward an internationally accepted IR reporting standard calls for advances in common reporting practices. Comprehensive, process-based reporting and presentation would make IR more understandable to others than information specialists and facilitate quality control.

III. Literature Review – 2 : WEB INFORMATION RETREIVAL

The concept of **Web-based Information Retrieval Support Systems (WIRSS)** is introduced. The needs for WIRSS are shown by a detailed case study of existing research article indexing and citation analysis systems, such as Current Content, DBLP. The objective of WIRSS is to build new and effective research tools for scientists to access, explore and use information on the Web, which may lead to improved research productivity and quality.

The analysis of existing systems and products leads to an important conclusion. The needs for the design and implementation of new generation systems that explore additional structures and provide more functionalities are obvious. We suggest the term Web-based information retrieval support systems (WIRSS) for such a study. WIRSS are designed with the objective to provide the necessary utilities, tools, and languages that support a user to perform various tasks in finding useful information and knowledge. They can be designed as an integrated systems combining existing systems. Information retrieval support systems, Web browsers, and Web search engines extend the basic search functionalities of data retrieval systems exemplified by a database system. They provide basic functionalities to assist a user in the context of libraries and in the early stage of the Web. A user may need to perform many different tasks when finding useful information. The new tasks include understanding, analysis, organization, and discovery, in addition to the conventional tasks of search and browsing. WIRSS is actually a natural evolution from information retrieval systems (IRS). The evolution from data retrieval systems to information retrieval systems and from information retrieval systems to information retrieval support systems were discussed in details in

IV. Literature Review – 3 : COMPUTER FORENSICS INFORMATION RETREIVAL

A Comprehensive Survey on Computer Forensics: State-of-the-Art, Tools, Techniques, Challenges, and Future Directions. The surveys reported in the past are limited to facilitate forensic investigators to pick a suitable forensic tool. Some previous research works such as focused more on providing an overview of digital forensics methodologies, finding errors in toolkits, and research directions but did not provide any guideline to investigators for intelligent selection of appropriate toolkit for evidence analysis.

With the alarmingly increasing rate of cybercrimes worldwide, there is a dire need to combat cybercrimes timely and effectively. Cyberattacks on computing machines leave certain artifacts on target device storage that can reveal the identity and behavior of cyber-criminals if processed and analyzed intelligently. Forensic agencies and law enforcement departments use several digital forensic toolkits, both commercial and open-source, to examine digital evidence. The proposed research survey focuses on identifying the current state-of-the-art digital forensics concepts in existing research, sheds light on research gaps, presents a detailed introduction of different computer forensic domains and forensic toolkits used for computer forensics in the current era. The proposed survey also presents a comparative analysis based on the tool's characteristics to facilitate investigators in tool selection during the forensics process. Finally, the proposed survey identifies and derives current challenges and future research directions in computer forensics.

V. Conclusion:

We had three different Literature Review of Information Retrieval (Medical Information Retrieval, Web Information Retrieval and Computer Forensics). In Literature Review – I, the use of Information Retrieval in Medical Field such as Virtual Health Assistant, Advanced Treatment and etc. In Literature Review – II, (Web Information Retrieval), it is used for Information Retrieval Supports from Web Browsers & Search Engines. In Literature Review – III Specifically focus on data retrieving in the digital crime such as Cyber Attack, Cyber Crimes and etc.

In all three of them common Problem addressing is advanced and easy report generation the tools are missing. In future we can improve the following things in Information retrieval System. They are

- Finding the advanced and enhanced tools for report Generation.
- While Searching information, finding out the missing information, Fake id & automatic searching.
- We can use IRS with Data mining for Automatic Report Generation.
- We can Combine Machine Learning, Deep learning with IRS for Automatic Searching.
- Image processing is used with Medical IR.

VI. References

1. **Introduction to Information Retrieval** “Christopher D. Manning” Stanford University “Prabhakar Raghavan” Yahoo! Research “Hinrich Schutze” University of Stuttgart - Cambridge University Press 2008.
2. Frommholz, I., Mayr, P., Cabanac, G., Verberne, S. (2022). **Bibliometric-enhanced Information Retrieval:** 12th International BIR Workshop (BIR 2022). Lecture Notes in Computer Science, vol 13186. Springer, Cham. https://doi.org/10.1007/978-3-030-99739-7_70
3. **Web-based Information Retrieval Support Systems: building research tools for scientists in the new information age** “J.T. Yao”, “Y.Y. Yao” Department of Computer Science, University of Regina Regina, Saskatchewan, CANADA S4S 0A2.
4. H. Chen, Sherrylynne S. Fuller, and Charles P. Friedman. *Medical Informatics: Knowledge Management and Data Mining in Biomedicine* (Integrated Series in Information Systems). Springer, June 2005.
5. **A Review on recent research in information retrieval** by S. Ibrihicha,* , A.Oussousb, O. Ibrihicha , M. Esghir, Published by Elsevier B.V, DOI : <https://doi.org/10.1016/j.procs.2022.03.106>
6. **Reporting and presenting information retrieval processes:** The need for optimizing common practice in health technology assessment Christina Johanna Niederstadt.Medical Consulting Service Article in International Journal of Technology Assessment in Health Care · October 2010 DOI: <https://doi.org/10.1017/S0266462310001066>