CHAPTER 3
INTRODUCTION TO ONLINE DATABASES

Online database is a record of logically related information, recorded in computer files in a uniform form to facilitate easy and efficient retrieving of data by means of internet or communication networks. The emergence of online databases has introduced a new era for information searching and retrieval from traditional retrieval systems to network based IR systems. Gray (1976) defined online database as “a collection of bibliographic information stored in a computer or a central computer and retrospective search is made possible remotely from a distance computer through the use of softwares”. Basically, it is an integrated collection of bibliographic information arranged logically in a far located central computer with which users can communicate via telecommunication networks. Online database is a product of information retrieval services provided by producers or vendors of online databases in which searches are carried out by means of a local computer that interacts with far-off systems containing information contents. Soyizwapi (2005) stated that electronic databases can be accessed anywhere from remote location and users do not need to go to the library to get information. Online databases involve searching of remotely located information through interactive computer and communications networks irrespective of any physical or geographical location. Today’s online databases typically work on time-sharing and real-time modes to enable users to access information simultaneously by communicating host computer directly.

Online databases include information that is organized and represented in a logical manner. Structurally, online database made up of records, further divided into a number of fields (author, title, etc.) for categorizing, searching and retrieving information. Initially, online databases were used to provide bibliographic information, later they started to comprise full-text information to provide actual information, since then they were called as full-text online databases. These are the vast and continuously updated records of information in the form of abstract, full-text references or citations on general or specific field of knowledge. Thus, these are the online sources that provide information on different fields of study with greater ease, accuracy and speed.
According to Sinh and Nhung (2012), “online database is an important part of
the information resources provided by universities in many countries. This is a
valuable and up do date source of information, therefore it is indispensable to the
learning, teaching and investigating for the university communities. They become
accustom to the use of the online database to search the information and develop
certain searching skills”. It is a vital source to locate different types of electronic
documents, such as books, scholarly journals, theses, reviews, official documents and
so on. These are the collection of logically arranged information sources that enable a
user to search information on a particular topic, article, or book by using different
descriptive elements, i.e., keyword, author, title, subject and date of publication etc.
However, some of these databases contain full-text information, while others contain
only abstracts and indexes; and citations of published literature. Online databases can
be general, multidisciplinary or subject specific in scope. The availability of these
databases enables users to search information in a variety of forms and formats, such
as HTML, PDF, RTF, etc. In online databases, contents are arranged in such a way
that it can be easily searched and retrieved without the help of any intermediaries.
These are excellent tools for in-depth study, providing a range of consistent and
scholarly literature from well known online database producers or publishers.

Online databases provide access to those digitalized resources that are not
generally available on the World Wide Web. These databases make use of computers
and communications to offer various types of information for serving the academic
and research community globally. Furthermore, the information provided by online
databases is scholarly in nature and users may possibly need to pay an amount of
money to subscribe it. Each database varies from the other in terms of their features
and functionality, types of documents (books, journal articles, theses, etc.) included,
coverage of contents, language and date of contents, searching capabilities and
interfaces. Some of the well known online databases in the field of education and
research are: JSTOR, Lexus-Nexus, Medline, ProQuest, ScienceDirect, Compendex,
ERIC, Cambridge University Press, Project Muse, Project Euclid, SIAM, MLA
(English), WorldCat and so forth. Thus, these are the excellent sources for those who
are looking for authentic and credible sources of information.

3.1. CONCEPT OF ONLINE DATABASES

Information retrieval is concerned with all the activities that related to the
organization of information, its processing and providing access to the recorded
information in different forms and formats. The practice of Information Retrieval has been a part of human technological development since the time of writing. Information retrieval (IR) systems were developed to retrieve relevant data from the huge collection of information in response to users’ queries. These systems are the results of many research and development activities that were carried out more than 50 years ago. Before 1960s, the major function of IRs was to organize information. The examples of earliest IR systems included the tools or organization schemes of ancient archives and libraries, viz. ‘Sumerian archives’ or the ‘Pinakes’ developed by Callimachus for the library of Alexandria. With the evolution time and techniques many IR systems evolved. However, the strong need for managing the increasing amount of information in the field of business and scientific development gave impetus for the development of automated IR systems. Consequently, the concept of automated retrieval system was developed for carried out automated searches for recorded collections of information. However in 1945, the concept of using machine-based systems for storage and retrieval of information was got recognition after a writing by Vannevar Bush (Larson, 2010). Hence, initial IR systems allowed people to interact with an information system in which information is recorded to find relevant information of all forms, i.e., text, images, audio video recordings to fulfill their specific needs. A typical information retrieval system functions in a two ways, on the one side it analyse the contents, on other the side, it evaluates the user queries to match them with each other and retrieve the relevant information from an organized collection of documents according to users’ needs.

Generally, the term automated IR systems were used in computer-based systems for different types of computer methods have been used for information storage and retrieval, accommodating significant growth of both information and its needs. These IR Systems stored data in computer files, accept requests, search the files against the request and provide information that is relevant for the requester. These systems were mainly designed to store large number of record and to facilitate its quick and easy access in response to the request. Later on, the hasty developments in information and communication technologies (ICTs) such as, emergence of internet and World Wide Web (www) technologies have brought major changes and improvements in the way the information was collected, stored, retrieved, and distributed. The introduction of online Information Retrieval systems is one of the major outcomes of application ICTs in information retrieval. These new systems
facilitated searching of remotely located records of information with the help of ICTs. According to Xie (2010), “Online IR systems can be characterized as IR systems that allow remote access with searches conducted in real time. Users generally search information from four types of online IR systems: online databases, online public access catalogs (OPACs), Web search engines and digital libraries”. While, online IR systems are also known as professional online systems, which are also abbreviated as online systems or online databases (Chu, 2009). Therefore, an online database is widely recognized as a typical online IR system.

Online database is one of the products of information and communication technologies involving searching of information from remote located databases via computers and communication networks. The database can be accessed through some intermediaries or vendors by means of online networks. Before the 1980s, online databases were mostly concerned with the retrieval of bibliographic information from recorded contents of information. Afterwards, these databases began to be include as numeric information, then full-text information was incorporated in these systems. The main purpose of an online database is to retrieve the information that fully or partially matches with the user’s query. Presently, the online databases comprise of abstracts, citations, full texts, statistical information in the form of journals, magazines, dictionaries, government documents, financial reports, audio and video contents and so on.

3.2. HISTORICAL OVERVIEW OF ONLINE DATABASES

Online database systems developed as a result of ICT applications in the process of information retrieval, and over the past five decades these have undergone several changes in their searching and retrieval capabilities. Online databases have the long history of developments that fall into several periods of time as stated in the Online Information System (1977) and Information Storage and Retrieval (1829).

Before the 1940s the information retrieval systems were purely of manual type, such as indexes and catalogues. These systems included pre-coordinate and non-manipulative retrieval devices. While, during the period of 1940s brought about an important development in the history of information retrieval, i.e., the invention of post-coordinate and manipulative retrieval systems, though these systems were entirely manually operated. These included Peek-a-Boo or Optical Co-Incidence Systems introduced by Batten and Cordonnier, Edge-Notched Cards System developed by Mooers and the Uniterm System of Taube. These early post coordinate
systems were the predecessors of modern computer-based systems, actually the two fundamental forms of file organization i.e., term entry and item entry were introduced by Batten and Mooers in the 1940s, became the base element in the designing and development of modern information retrieval systems.

In the 1950s, Punched Card Data Processing Systems based on the earlier forms of mechanization was introduced in the late 1950s, techniques of automatic indexing were developed by Hans Peter Luhn. During 1950s, the first computer based retrieval systems were introduced as the immediate predecessors of the computer-based systems of the 1960s.

The period of 1960s was identified as the era of computer retrieval with an off-line, batch-processing, tape-oriented model. In the early 1960s, a number of information centers and government agencies in the United States started to use these information retrieval techniques for developing their operations, e.g., the National Aeronautics and Space Administration (NASA) used computer systems to publish its index in Aerospace reports and the National Library of Medicine (NLM) used these techniques to publish Index Medicus (index to medical literature). However, the major off-line systems were developed during the 1960s, but the widespread transformation of operations from off-line to the on-line mode was observed in 1970s.

In the late 1960s and early 1970s, different types of reference databases were made available through different online national database services including, MEDLINE service of the National Library of Medicine, the bibliographic services of OCLC and RLIN and many other commercial and governmental services. Presently, these online IRS are globally used along with hundreds of reference databases accessible through them. Since 1975, on-line systems were used more frequently for providing information services. Some of the online databases developed to provide access to a wide range of information including, Chemical Abstracts Service (CHEMCON), Engineering Index (COMPENDEX), the National Technical Information Service, the Science Citation Index, ERIC (the Educational Resources Information Center), and CAIN (the machine-readable data base of the National Agricultural Library).

The growth of online databases can also be interpreted in terms of number of database producers, providers or vendors. During the 1970s online service industry had continued to grow. Online search services, or vendors, are those organizations that provided value-added services to databases with several search capabilities.
Chapter 3

Introduction to Online Databases

online service providers obtained databases through licensing agreements and loaded these on its own computer to provide subscriptions of them on payment basis to institutions and libraries. The major online service providers included; the Dialog (System Development Corporation, SDC), ORBIT (Lockheed Information Systems), Mead Data Central and Bibliographic Retrieval Service (BRS). Several major online service providers are now providing access to a number of online databases that became the central interest of libraries.

In addition to this, the Text Retrieval Conference (TREC) series of experiments and conferences have also played an important role in the development of information retrieval research and development activities (Chowdury, 2010). The TREC (Text Retrieval Conference) project was launched in 1992, under the sponsorship of the U.S. National Institute of Standards and Technology and the U.S. Advanced Research Projects Agency, has conducted many constant and dedicated attempts on interactive searching. The purpose of the TREC project was to provide a forum to compare a variety of approaches to information representation and retrieval (Sutcliffe, 2010). It proposed general framework for the investigation of interactive information retrieval, evaluation and comparison of the performance of interactive IR systems (Dumais & Belkin, 2005).

3.2.1. Timeline of developments in online IRS

Online retrieval systems had been developed as a means to facilitate searching of bibliographic information. The historical developments of online databases are illustrated here were mainly focused on the developments that occurred during different periods of time:

1. The first experiments in on-line information retrieval conducted by Kessler at MIT in 1964 brought the Kessler’s experimental system in physics, known as the TIP (Technical Information Project). It was the first on-line system for bibliographic searching, which included nonconventional methods of searching.

2. In 1964, the Medical Literature Analysis and Retrieval System (MEDLARS), one batch retrospective search service of the National Library of Medicine (NLM) was made available to the general public.

3. SDC Information Services (Systems Development Corporation) demonstrated the first interactive online system, ‘Protosynthex’ developed by Robert Simons and John Olney, in 1960.
4. In late 1964, the Bibliographic Organization for Library Display (BOLD), a system for browsing of literature citations on magnetic tapes was developed by Harold Borko, H. P. Burnaugh and W. H. Moore at SDC. It was one of the first systems capable of displaying an online thesaurus.

5. In late 1967 NLM (National Library of Medicine) experimented with SDC's Online Retrieval of Bibliographic Information Timeshared (ORBIT) retrieval language to search NLM’S database of 10,000 citations on neurology.

6. In 1970 NLM introduced MEDLINE as a free database of more than 400,000 citations. While, in 1972 NLM started to use TYMNET, the first public telecommunication network to access to MEDLINE.

7. In 1962 Roger K. Summit designed the DIALOG language for the Lockheed Information Sciences Laboratory.

8. RECON (Remote Console), the first large-scale, on-line, retrieval system of the National Aeronautics and Space Administration (NASA) became fully operational in 1969. Now, the RECON, a NASA database is recognized as an international system for on-line search in Europe through the European Space Agency.

9. In 1970s, ‘The New York Times Information Bank’ an important on-line system became operational for providing access to current awareness information from The Times and other selected sources.

10. In 1972, Lockheed's information service became commercially available as DIALOG information retrieval Service with two bibliographic databases of scientific and technical information.

11. By 1985, DIALOG had become the most comprehensive online information service in the world, with more than 200 separate databases in business and economics, chemical, patent and trademark information, science and technology, medicine and the biosciences, news and current events, education, directories, energy and the environment, law and government, computer science and microcomputers, books, the social sciences, and the humanities.

12. Mead was mainly associated with the development of full-text online information services primarily for law and legal research. In 1972, this
service introduced commercially as LEXIS an outgrowth of OBAR (Ohio Bar Automated Research) and in 1980 NEXIS, a full text information service for news and current events; and MEDIS was introduced in 1985 for medicine.

13. BRS (Bibliographic Retrieval services) expanded by 1985 to include 73 separate databases in the life sciences, medicine and pharmacology; the physical and applied sciences; education; the social sciences and humanities; and business.

14. In 1907, the American Chemical Society (ACS) published Chemical Abstracts as an index to chemical literature.

15. In 1980, CAS introduced CAS ONLINE as an online dictionary of chemical substances and it was expanded in late 1983 to include the database of Chemical Abstracts citations dating back to 1967. CAS made available several million of unlicensed abstracts for citations in the database available through CAS ONLINE.

16. In the late 1970s a consortium of British producers of scientific, technical, and patent databases created INFOLINE. In 1981, INFOLINE was purchased by Pergamon Press and developed as Pergamon INFOLINE. Since then, the online databases have continuously growing in numbers and scope. Today millions of online databases are available in different fields of education and research to cater to diverse needs of different types of users and these may be subject specific, interdisciplinary or multidisciplinary. Some of the well-known online databases are: Medline, ScienceDirect, BIOSIS, JSTOR, Annual Reviews, Emeradinsight, Cambridge University Press, Oxford University Press, Chemical Abstract, Web of Science, Scopus and LEXIS and Taylor & Francis.

3.3. DEFINITIONS

- ODLIS (2012) defines databases as, “A large, regularly updated file of digitized information (bibliographic records, abstracts, full-text documents, directory entries, images, statistics etc.) related to a specific subject or field, consisting of records of uniform format organized for ease and speed of search and retrieval and managed with the aid of database management system (DBMS) software. Content is created by the database producer (for example, the American Psychological Association), which usually publishes a print version (Psychological Abstracts) and leases the
content to one or more database vendors (EBSCO, OCLC etc.) that provide electronic access to the data after it has been converted to machine-readable form (PsycINFO), usually on CD-ROM or online via the Internet, using proprietary search software.

- According to the Macmillan Dictionary of information Technology (2014), “Database is a collection of interrelated stored data so that it may be accessed by users with simple user friendly dialogues”.
- The chambers Science and Technology Dictionary (1999) defines database as “A collection of structured data independent of any particular application”.
- The Law dictionary (2012), defines online databases as “A web-based filing system used to store information or records, accessible by using web scripts. Use often requires a paid subscription”.
- According to the University Santo Tomas (2008), “Online databases are collections of computerized information or data such as articles, books, graphics and multimedia that can be searched to find information. Databases can be general or subject based in form of abstracts and or full text”.

3.4. CHARACTERISTICS

i. Organized Collection

In an online database, contents are arranged logically to facilitate easy access and retrieval. Documents in such types of systems are organized in a suitable manner for carrying out easy and fast retrieval of information.

ii. Credibility

In online databases, recorded contents of information are reviewed by subject experts and publishers to maintain credibility and authenticity of the resources. The contents of database are finely evaluated in terms of their accuracy and credibility.

iii. Usability

The well defined organization of information contents and search capabilities of online databases allow users to search and retrieve results more efficiently and effectively.

iv. Conversational

Searches in online databases are conducted as a two-way communication between the searcher and the system, in which each get a chance to communicate with each other. Therefore, online databases are referred to as interactive or conversational.
v. **Expert System**

Online databases are characterized as expert systems that provide information on specialized areas of knowledge, e.g. a nuclear database gathers specific information on nuclear sciences from experts or specialized associations and provides particulars of the nuclear sciences.

vi. **Controlled Vocabulary**

To support searching, online databases usually have their own controlled vocabulary. Controlled vocabulary is largely used for information presentation and retrieval, though the keyword searching is also supported by almost all the online systems.

vii. **Permanence**

Published documents, such as journals, reviews and books, etc. in online databases do not change frequently. These information documents remain in databases for a long time in the form of archives to again retrieve the information.

viii. **Up-to-Date**

All the online databases comprise current information on its concerned areas. Online database providers or publishers are regularly updating their contents by adding new information to provide current and copyrighted scholarly materials.

ix. **Real-time**

Real time in online database operations implies that the remote terminals respond quickly to the user's search processes. Remote terminals receive data, search the information, and return the results more frequently to be utilized by the users in ongoing activities.

x. **Time-sharing**

On-line time-sharing implies the sharing of machine processing time among a number of terminals. It means that the processing time of the computer is equally shared amongst independent activities. By time sharing many users can search the information simultaneously either from the same place or variant locations.

### 3.5. **TYPES OF ONLINE DATABASES**

Online databases are divided into the two main categories; Reference and Sources databases, on the basis of information incorporated in them. According to Chowdhury (2010), online databases are categorized under the two major divisions include, reference databases and source databases. Reference databases direct seekers towards the source of information, while source databases include actual
information itself. On basis of content, scope and the information incorporated, online databases are grouped under the following categories:

- On the Basis of Information Incorporated
- On the Basis of Scope of Data
- On the Basis of Contents
- On the Basis of Providers

**I. On the Basis of Information Incorporated**

Based on the information included in online databases, this category comprises the most common forms of online databases, which are grouped into following four types:

a. Full-text Online Databases
b. Reference Databases
c. Numeric Databases
d. Multimedia Databases

**a. Full-text Online Databases**

A full-text database is a compilation of documents or other information in the form of database in which complete text of each referenced document is available for online viewing, printing, or downloading. In addition to, text documents, images are often included as graphs, maps, photos, and diagrams. Full text online databases are comprising full text information of the publications that are basically either print or online in origin. It includes large files of text such as, all the paragraphs of a journal article or all the chapters from a book along with abstract or citations of the text files incorporated in them. According to Bandyopadhyay (1999), Full text online databases are now being using as effective and important sources of periodical literature that are not usually available in local collection of libraries.

Examples: JSTOR, Emeraldinsight and Wiley Online Library etc.

**b. Reference Databases**

The references databases include terms descriptive of content on which retrieval is based and some databases also provide abstract to give brief description of original documents and in these databases the retrieval of information is mostly based on the words appeared in abstract. The retrieved data helps requester to identify where an original source can then be found (Information Storage and Retrieval, 1829). The reference databases provide bibliographic descriptions to published literature.
Chapter 3

Introduction to Online Databases

provides abstract, references or citations to documents. These databases are divided into two main categories:

- **Bibliographic Databases**

  Bibliographic database is one of the most important forms of reference databases. These databases are widely used as reference tools and provide citations or references, abstracts and index to published literature. Online bibliographic databases provide quick information about publications, which may or may not be available in library’s own collection. These are the excellent means to access information, rather than merely an item of information in the collection of any library. Bibliographic databases contain elements of bibliographic description that used to describe books, journals, documents, and other publications or portions. Rice (1985) stated that these databases are typically online bibliographic files, are the online equivalent of print abstracts and indexing services and mostly used for bibliographic verification of literature.

  Examples: LISA, Indian Citation Index, Scopus and Web of Science etc.

- **Referral databases**

  Referral databases direct users towards the particulars of actual source of information i.e. name of a person or institution. It offers references to information, such as names, address, specialization of persons, institutions, information systems, and so forth.

  Examples: Ulrich's Periodicals Directory and Electronic Yellow Pages

- **Numeric Databases**

  A numeric database is a computer-readable collection of data that are primarily numeric in nature. These are also known as the fact sources and non bibliographic databases and are mostly used for supporting business or financial research. These databases include organized numerical data along with brief textual description and provide it access in the form of statistics, demographic and financial reports, stock market quotations, chemical and physical properties, and chemical nomenclature and graphic structures, etc. These are the files of primarily statistical information from which a user can extract specific forms of data.

  Examples: COMPUSTAT and ProQuest Statistical Insight

- **Multimedia Databases**

  A multimedia database is a collection of related multimedia data objects of different types. A multimedia database contains various data types such as images,
sound recordings, video recordings, signals, graphics together with text data etc. Multimedia databases host different types of media file, such as .txt for documents, .jpg used for images, .swf deals with videos, .mp3 use for generating audio files etc. These databases involve the activities related to acquisition, generation, storage, processing and transmission of multimedia data over networks. The databases primarily provide access to art prints, animations, photos, audio clips, videos and others multimedia contents.

Examples: Artstor and Academic Video Online

II. On the Basis of Scope of Data

Online databases can be classified by the scope of information contained in them.

a. General interest Databases

Such types of databases provide broad range of information on different subject and disciplines. General interest databases comprise of information that is more general in nature, like current news and opinion, social and political affairs, cultural, educational, health and on public issues.

Example: Academic Search Complete (EBSCO) and Encyclopaedia Britannica

b. Discipline Specific Databases

Discipline-based databases are somewhat more specific than general interest databases. These provide information on several related areas. If information is not found in general interest databases then it is better to search in such types of databases.

Examples: SocINDEX and PAIS (Public Affairs Information Service)

c. Subject Specific Databases

These databases are well suited for in-depth research and study on a particular topic. These databases provide information from professional publications and scholarly journals. The subject specific databases are devoted to only one subject. Searches in subject specific databases are more comprehensive in nature to provide access to more scholarly articles.

Example: Historical abstracts and PsycINFO

III. On the Basis of Contents

Online databases can be organized according to the type of documents they possessed.
Chapter 3  

Introduction to Online Databases

a. Article Database

An article database allows a person to search across thousands of various journals and magazines to locate an article on any specified subject. Articles databases mostly provide full text of articles, but sometimes they also provide abstracts of articles. Some online databases only present citations instead of full text article or an abstract that helps to locate the original article.

Example: Google Scholar and Annual Reviews

b. Theses/Dissertation Database

Theses/dissertations databases are developed to maximize the visibility and availability of research output and to provide opportunities for further research. Such databases enable searching for dissertations and theses through a single access point, which presents an extensive and authentic collection of millions of research works in full-text. These databases are the record of doctoral theses or dissertations awarded by Higher Education institutions.

Example: ProQuest Dissertations and Theses and EThOS- Electronic Theses Online Service

c. Citation Database

Citation databases are index of citations of published literature. It enables to locate bibliographic citations for journal articles and track articles in a specific subject. It allows users to track which current documents cite which previous documents. Many citation databases include index of journal articles along with its abstracts. By searching with keywords that might appear in an article, users can retrieve citations of an article.

Example: Scopus and Web of science

d. Audio / video Database

An Audio video database is a collection of Audio Video materials such as digital audio and video data and Audio video activities.

Example: Audiovisual Database of Spoken American English and Academic Video Online

e. Online Catalogue Databases

An online catalogue database is a bibliographic database that describes the books, periodicals, and electronic resources, etc. that are available in the library. Online catalogues are those online databases that enable searchers to search for
documents by author, title, subject heading, keyword, call number, or government documents number available in a particular library.

Examples: IndCAT and WorldCat

f. **Dictionary Databases**

In dictionary databases, likewise of directory catalogue each record identifies something. The purpose of dictionary databases is to provide a measure of control in the use of bibliographic databases.

Examples: Oxford English Dictionary and Chemical substance Dictionary

g. **Directory Databases.**

Directory databases offer the information of published directories or serve a purpose similar to that of published directories without having published equivalents. These are not full-text databases although they may represent the complete text of a publication in machine-readable form; nor are they numeric in nature.

Examples: Electronic Yellow Pages and Encyclopedia of Associations

h. **Indexing & Abstracting Databases**

These databases provide brief summary of publications along with descriptors as access points to documents. Such databases provide clues to the relevance and location of the publication.

Example: SocINDEX and Educational Research Abstracts Online (ERA)

**IV. On the Basis of Providers**

Based on their providers, online databases can be classified into the following categories:

a. **Publisher /Commercial Databases**

Publisher databases are produced by online commercial service provider that sell their data to the clients and deliver information through the telecommunication networks. These databases are commercial electronic information services that people access through the Web or internet. Anyone can download or electronically copy of the information contents from the Internet anywhere in the world through database’s home page. To access such types of databases, searchers need to have an authorization number and password provided by the publishers.

Example: Oxford University Press and Taylor & Francis
b. **Institutional Databases**

These types of online databases are developed by the professional associations or institutions to increase the knowledge of their concerned areas among the people. These associations mainly work for promoting research and developments in their working areas in the broadest manner. They develop different types of information sources to increase dissemination of knowledge related to their concerned areas, including books, journals, reports, and databases etc. The online databases produced by them are mainly subject specific in nature.

Example: PsycInfo (American Psychological Association) and EconLit (American Economic Association)

c. **Aggregator Databases**

Aggregator databases are defined as the service providers that make available contents, licensed by several publishers and is offered in packages at a single price to libraries. These offer extensive depth and breadth of contents of information along with effective features and functionalities.

Examples: ProQuest and EBSCO's

### 3.6. SEARCH STRATEGIES

Search strategies are a combination of different methods used for searching documents and can be identified by types of search methods used and dimensions related to searches such as purpose, type of resources to be searched, methods for searching. According to Bates (1996), “A search strategy is a plan for the whole search, while a tactic is a move made to further a search”. Search strategies are the demonstration of patterns used in search processes and for searching the information.

Database searching allows users to search information from an organized collection of records. Users can search through by using different moves and tactics to get efficient results. In online databases, search strategies can be classified in two types, Search and Browsing. These are the main strategies employed by users while communicating with online databases. Browsing needs more interactions with online databases than logic based search strategies use in searching.

#### 3.6.1. **Search Methods**

Search Methods are defined as methods that assist users in constructing their queries that include two types of searches, viz., basic and advanced searches. All online systems possess these two search methods, which further provide different retrieval techniques, such as Keyword search, Boolean operators, truncation, phrase
searching, proximity search and so on for more effective and efficient information retrieval.

I. Basic or Simple search

It is a commonly used search method in database searching. It is also referred to as keyword searching, as keywords are important words that are used as index to the information in a database. In basic or keyword searching, online database searches the whole document to locate words and phrases defined by the users. Thus, it is useful, when user do not know the exact title or author name and want to link terms from different parts of a record, such as title, abstract, author etc.

II. Advanced search

This search provides the following search options:

a. Boolean operators:

In database searching, Boolean operators are used to narrow or broaden the search. The most useful logical connectors include, AND, OR, NOT. Amongst which AND is used to narrow the search, OR to broaden the scope of search, whereas NOT is used to eliminate unwanted terms from the search. Boolean operators are considered as a common technique for advanced searching. For example:

- Online AND Database
- Academic database OR Library database
- Electronic database NOT CD-ROM

b. Combined search

This method provides the opportunity to combine two or more Boolean operators in the same search statement. Combined search allows users to search databases by adding different logical operators at the end of each search string.

Example: Online AND Databases NOT Journals

c. Phrase searching

Phrase search retrieves exact words in the same order mentioned by a searcher. This feature produces that a result, in which exact phrase is retrieved within a document or any specified field (title, abstract etc.) of documents. It retrieves documents with exact terms adjacent to each other within the same document. Each database provider allows phrase searching as exact word search by using different command, for examples; the symbols (“ ”, { }) are used for “Online databases” to retrieve all the words in a phrase in an exact order.
d. Field specific searches

Field specific searches facilitate users to limit their search terms to a specific field(s) of documents including author, title, abstracts, subjects and volume, etc. Thus it allows users to search within specified fields of recorded documents.

e. Limiters

While searching databases, users are provided the facility to limit their searches by using different types of elements. Limiters constrain a users’ query within a specified requirement by imposing limit using specific elements, i.e., date, type, etc. Almost all the online databases have limiters to control the results by using specified parameters; this retrieval feature is helpful in eliminating records that are outside the scope of defined limits. Users can narrow their searches limiting; by date, language, title of article, author(s) of article, subject and limit to a particular journal.

f. Truncation and wildcard

Truncation and wildcard search techniques are used to retrieve variations of a word. In truncation user can enter a base word (prefix) to retrieve all the words beginning with that base term. Generally, these techniques are used to truncate or shorten a word to retrieve singular, plural and its variant spellings. The symbols (*, $, ?) are used to represent truncation in some online databases, while in others these are used to represent wildcard searching. Examples for searching variants of the word politics: poli* (policy, politics, political, politically, etc.)

g. Proximity search

A proximity search is used for fixing the distance between two words in the same sentence or paragraph of a document. It sets search terms that occur in the same order as defined by the searcher. Commands used for proximity search include; ADJ, WITH and/or WITHIN and NEAR. For example, the statement STRUCTURE (WITH7) DATABASE indicates that the word STRUCTURE must not be separated from the word DATABASE by more than seven intervening words.

h. Stemming (related terms)

Stemming technique allows searching of all the related variants of a term. It uses the base words of the search term as the stem to retrieve all the related terms as its stem variations. This retrieval feature is presented by many of the online databases but their implementation and interpretation may differ from one another, such as in EI Village and Ovid dollar sign ($) is required before the search terms, whereas the
databases, ACM Digital Library and Emerald automatically stemmed the search terms.

i. **Term Boosting**

The online database provides the relevance level of matching descriptions based on the terms found. The higher the boost factor, the more relevant the term will be. Term boosting is the ability used to assign higher weightage to specific words in a query. It provides facility to users to control the significance of a description by boosting its term. A symbol of caret, “^”, is used at the end of the term to boost its relevance in retrieved contents. For example: electronic database^4 journal. This reflects that the term electronic database is four times more relevant in a description than journal.

III. **Expert search**

Expert searching implies the application of a range of advanced search skills and knowledge to get more specific information. It allows users to perform more complex and complicated searches in multiple sources simultaneously, which include all sources or journals, books or reference works, etc. With expert search users can enable to prioritize their search terms to access the most appropriate item of information. The online databases ScienceDirect and MEDLINE are providing this option for searching.

IV. **Citation Locators/ Trackers**

Citation Locators enable users to locate any article by entering its details from an article reference or citation. It searches throughout the different fields of a citation including author name, article title, journal title, volume/issue information, year, etc. In case, if an exact match to the citation is not found, then it will retrieve closely matched articles or information. Some of the online databases provide the citation locator or tracker to find, check, and track citations. Citation tracker/locator offers citations to users to provide an overview of how many times a chosen document has been cited in other works and the number of documents that cited it.

3.6.2. **Browsing**

Browsing is another type of search strategy. It is a process of seeking searching of information by skimming and scanning of contents mainly in leisure. Browsing is exploration of information with a specific goal but without any planned search strategy. It is a technique to look through documents in an informal manner, to search information without any defined purpose, to look and learn new information.
and to obtain an overview of the information offered by online databases. Browsing is a technique of information retrieval where the initial attempt of searching is usually undefined.

While, browsing users do not need to define specific terms as required in the searching, so, it requires less intellect than searching. Irrespective of searching, browsing can be done without any training and practice. Hence, it is a rewarding exercise that is mainly based on serendipity in finding some useful information unexpectedly, because it is carried out unstructurally. Browsing can be done by using different components of a document, the common browsing options are:

- ‘Browse by title of journals’, allows users to browse alphabetical list of journals available in that particular online database.
- ‘Browse by keywords’, facilitates users to browse different types of collections (journals, books, reviews and reports etc.) available in online databases by using keywords.
- ‘Browse by subjects’, feature allows users to browse contents by subjects.

However, all the online databases are providing this retrieval capability, but differ in implementation of the components used for browsing.

3.7. OTHER FEATURES

Online databases also offer some more retrieval features other than basic, which are discussed as follow:

- **Links**
  
  This feature helps users in getting full-text items or articles available in other databases or e-journals of the same publisher.

- **Report**
  
  Online databases generate Electronic holding reports of an institution. It provides a list of content subscribed by the institution.

- **Login/ Register**
  
  Allow personalization of contents and features of online databases by creating personal account. By this facility users can customized their usage by save searches, subscribe different types of alerts (table of content alerts, favourite topic or journals alerts etc.) through e-mail and create a list of favourite journals, books and topic. Users can also do purchasing of any item of database through their account.
Modify Search

It allows users to modify or change search strategy to get better result. It is done by adding or removing the elements of content for which search is conducted.

Export Data

Users can generate bibliographic data of the article by using different citation manager formats including EndNote, CSV, BibTex and TSV.

News and Updates

Provide information regarding the new launches and acquisitions made by database providers.

3.8. ONLINE DATABASES IN THE ARTS AND SOCIAL SCIENCES

Online databases are now became an important source in online collection of libraries. Libraries are acquiring online databases through different sources, i.e., vendors or intermediaries. In India, UGC INFONET Consortium is the central source of acquiring online databases to the libraries. UGC INFONET Consortium is initiated by the UGC (University Grant Commission). The consortium acts as an intermediary to provide access to scholarly online resources to the academic libraries from reputed publishers, aggregators and society. According to Madhusudhan (2008), consortium covers almost all areas of knowledge, such as Arts, Humanities, Social Sciences, Computer Sciences, Life Sciences, Physical Sciences, Chemical Sciences, Mathematics and Statistics etc.

The libraries under study are acquiring an adequate number of online databases in different fields of learning from UGC INFONET Consortium. Online databases that comprises of literature on the Arts and Social Sciences are listed below:

I. Annual Reviews

Annual Reviews provides researchers, professors, and scientific professionals with a definitive academic resource in 37 scientific disciplines, also covering some areas of social sciences. Annual Reviews provides primary research literature and identifying the principal contributions in the field. It provides access to 33 full text journals and archival access is provided up to 10 years back issues.

II. Cambridge University Press (CUP)

Cambridge University Press is a publisher online database. It is dedicated to advance learning, knowledge and research worldwide, the database currently publishes over 220 peer-reviewed academic journals for the global market, containing the latest research from a broad sweep of subject areas. The CUP database also
publishes on behalf of over 100 learned and professional societies. UGC INFONET Digital Library Consortium provides access to 224 Cambridge University Press journals with back files since 1997.

III. Emerald

Emerald is a publisher based journal online database, linking research and practice to the benefit of society. The database covers nearly 300 journals and over 2,350 books and book series volumes in business and management, library and information sciences, engineering and materials science. As the leading publisher database for LIS research, Emerald's Library and Information Studies publications provide comprehensive and quality coverage in all areas of this field. Spanning a range of topics such as collection building to library finances, to document supply and inter-library lending, this is an essential resource for information professionals, librarians, educators, students and researchers around the world. Under UGC INFONET e-journals consortium access is made available for 29 e-journals from Library and Information Science full text database and archival access is varies from journal to journal (mostly 2001- onwards).

IV. JSTOR

JSTOR (Journal Storage) is full-text database for scholarship, established in 1995 as digital archives. The majority of content in the archive is journal literature, though inclusion of other materials such as conference proceedings, transactions, pamphlets, monographs, manuscripts, and other materials is continuously growing. At present, there are 2,000 journals, including previous titles, as well as other content available. New titles and other materials are being added regularly. It provides full text searches of almost.

V. Oxford University Press (OUP)

Oxford University Press provides access to Oxford Journals. It publishes well over 230 academic and research journals covering a broad range of subject areas. OUP database covers Life Sciences, Mathematics & Physical Sciences, Medicine, Social Sciences, Humanities and Law and include some of the most authoritative journals in these fields. Through UGC INFONET consortium, 198 Oxford University Press journals are available with back files since 1998 to the members libraries.

VI. Project MUSE

Project MUSE is an excellent full-text online database, provides affordable and user-friendly online access to a comprehensive selection of prestigious humanities
and social sciences journals. MUSE's online journal collections support a diverse array of research needs at academic, public, special and school libraries worldwide. It’s journals are heavily indexed and peer-reviewed, with critically acclaimed articles by the most respected scholars in their fields. MUSE is also the sole source of complete, full-text versions of titles from many of the world's leading university presses and scholarly societies. Currently, MUSE provides full-text access to current content from over 400 titles representing nearly 100 not-for-profit publishers.

VII. ScienceDirect

ScienceDirect is the world's renowned multidisciplinary online database, publishes over 2,000 journals as well as books and secondary databases. There are currently more than 9.5 million articles or chapters, a content base that is growing at a rate of almost 0.5 million additions per year. It offers subject coverage broadly includes all aspects of Physical Sciences and Engineering, Life Sciences, Health Sciences, Social Sciences and Humanities. It includes over 2,000 peer-reviewed journals and over 8,057,764 articles. It includes over hundreds of book series, handbooks and reference works and collection contain 4 million articles prior to 1995, and 2.75 million articles from after 1994.

VIII. Taylor and Francis

It is the oldest commercial journals publisher in the world, by providing access to its collection through online it comes under the category of a commercial online database. It provides access to more than 1100 journals and around 1,800 new books that enable the customers and end-users to perform their jobs efficiently, continue their education, and help contribute to the advancement of their chosen profession. It is a widely known online source among researchers, students, academics and increasingly professionals. UGC INFONET Consortium access more than 1365 journals with archival access to 1998 onwards issues.

IX. Web of Science

The Web of Science provides access to three major databases in Sciences, Social Sciences, Arts and Humanities. It provides seamless access to information from the world’s most influential, highly-utilized scholarly literature across a broad range of topics which covers more than 12200 of the most prestigious, high impact research journals in the world going back to 1898. With web of Science users can also navigate to electronic full-text journal articles. It also provides a unique search
method called cited reference searching. With it, users can navigate forward, backward, and through the literature, searching all disciplines and time spans to uncover all the information relevant to their research.

X. Wiley Online

Wiley Online is the online database of an international scientific, technical, medical and scholarly publishing business of John Wiley & Sons that provides literature in every major academic and professional field. Wiley Online is of the world's foremost academic and professional database. It provides access to more than 1,400 scholarly peer-reviewed journals and an extensive collection of books with global appeal in the life and physical sciences, medicine and allied health, engineering, humanities and social sciences. The UGC consortium accesses 908 journals from Blackwell publishing with back files since 1997.

3.9. SUMMARY

Online databases have developed as benchmark systems of information retrieval, mainly designed to retrieve documents required by the users. It works on the principle of providing right information to the right user instantly. These types of online systems designed with the aims of providing scholarly information in one or more areas of education and research. Online databases deal with a variety of information in variant forms and formats, comprising textual information, bibliographic information, numeric and multimedia information including text, audio, images and video. The online databases are acting as a bridge between the generators of information and the seekers of that information.

The chapter provides a conceptual overview of online databases and its development. It discusses definitions, characteristics and types of online databases. In addition, search strategies used in online database searches are also discussed in detail in this chapter. The prominent online databases in the field of Social Sciences and Arts are also mentioned in this chapter.
REFERENCES


