ONLINE AND CD-ROM INFORMATION RETRIEVAL
INTRODUCTION

• GROWTH OF PUBLISHED LITERATURE IN THE 1960S RESULTED IN THE DELAY OF PRIMARY PUBLICATION AS WELL AS SECONDARY PUBLICATIONS, LIKE ABSTRACTS AND INDEXES WHICH REFER TO THEM.

• THIS LED TO THE USE OF COMPUTERS IN THE PRINTING PROCESS TO SPEED UP THE PRODUCTION OF SECONDARY PUBLICATIONS.

• THE RESULTING BIBLIOGRAPHIC REFERENCE MATERIAL WAS STORED IN THE FORM OF STRUCTURED DATABASES IN THE HARD DISC AND MEMORIES OF COMPUTERS.

• THE STORED INFORMATION IS EITHER PRINTED OR ACCESSED DIRECTLY VIA A COMPUTER TERMINAL.

• THE INFORMATION CAN ALSO BE STORED IN OPTICAL MEMORIES SUCH AS CD-ROM.
ONLINE INFORMATION RETRIEVAL

- ONLINE INFORMATION RETRIEVAL IS THE ACQUISITION OF INFORMATION FROM A REMOTE COMPUTER SYSTEM OF DATABASES VIA A COMPUTER TERMINAL. USERS ACCESS THESE DATABASES THROUGH AN ONLINE SEARCH PROVIDER KNOWN AS A VENDOR.

- WITH THE ADVENT OF THE INTERNET, THESE DATABASES CAN NOW BE SEARCHED THROUGH THE WORLD WIDE WEB BY USING THEIR URLS TO ACCESS THEIR WEB PAGES.
DEVELOPMENT OF ONLINE SEARCHING

- **MEDLINE WAS THE FIRST MAJOR ONLINE DIAL-UP SERVICE FOLLOWED BY DIALOG IN 1972.**
- **BY 1975 THERE WERE ABOUT 300 PUBLIC ACCESS DATABASES INITIALLY MADE OF BIBLIOGRAPHIC OR REFERENCE DATABASES**
- **RECENTLY MOST DATABASES PROVIDE FULL-TEXT DOCUMENTS INCLUDING GRAPHICS AND PICTURES**
## DEVELOPMENT OF ONLINE SEARCHING

Rowley (cited in Chowdhury, 2010) identifies 3 generations of online searching:

   1. Bibliographic databases with
   2. Slow transmission speeds.

2. **Second Generation (throughout the 1980s)** — Both
   1. Bibliographic/full-text databases,
   2. Used PCs as work stations,
   3. Medium transmission speed

3. **Third Generation (beginning of the 1990s)** —
   1. Used multimedia PCs,
   2. Higher transmission speed,
   3. Improved user interfaces,

4. **Fourth Generation (1990s onward)** —
   1. Web-based online search services,
   2. Fast and easy access online,
   3. Multiple search and retrieval facilities (e.g. OVID, DIALOG, OCLC, FIRSTSEARCH)
COMPONENTS OF AN ONLINE SEARCH SERVICE

• INFORMATION PROVIDERS/DATABASE PRODUCERS – MAKE DATABASE AVAILABLE ONLINE

• SEARCH SERVICE PROVIDER/VENDOR- PROVIDES ACCESS TO DATABASE AND SOFTWARE FOR CONDUCTING SEARCH

• COMMUNICATION LINKS – TO CONNECTS THE USER WITH HOST DATABASE (NOW VIA THE INTERNET)

• LOCAL WORK STATION- LINKS USER TO THE SERVICE
BASIC STEPS IN AN ONLINE SEARCH

1. Study the search topic and develop a clear understanding of the information required
2. Access an online service by subscription/license agreement
3. Log on to the service provider using the requisite log-in information
4. Select the right database for your search
5. Formulate your query using appropriate search terms, know the nature, structure and content of the selected database and use multiple search strategies to get the required information.
6. Select appropriate format for display of retrieved records
CD-ROM DATABASES

INTRODUCTION

• CD-ROM was produced in the mid 1980s and used for storage of textual data in place of magnetic storage media because of the limitation of the latter. For example, the presence of dust can lead to loss of information from the magnetic storage media. However, dust/scratch on a CD-ROM only causes a slight loss of resolution because the data is recorded in the entire medium.

• A CD-ROM file management system allows users to view the disc as a collection of files. A complete CD-ROM file management system has 3 main components:
  • The structure/logical format of data
  • Software for writing the data in that format
  • Software for reading and translating the logical format for use called the destination software
## Differences Between CD-ROM and Online Databases

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Online Databases</th>
<th>CD-ROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>Several gigabytes of data can be searched per time</td>
<td>Has limited storage capacity which can be increased by use of networked CD-ROMs</td>
</tr>
<tr>
<td>Currency</td>
<td>Can be updated in real time</td>
<td>There is always a time lag. Updates take a while to reach users.</td>
</tr>
<tr>
<td>Costs</td>
<td>Often pay as you go which may be expensive for frequent and long search sessions</td>
<td>Cheaper because price is fixed and can be used many times</td>
</tr>
<tr>
<td>Response time</td>
<td>Depends on vendor computer and performance of Internet</td>
<td>Depends on local computers and networks</td>
</tr>
<tr>
<td>Cross-searching</td>
<td>Possible and depends on vendor</td>
<td>Possible and depends on networks</td>
</tr>
<tr>
<td>Equipment</td>
<td>Depends largely on the reliability of Internet infrastructure</td>
<td>Can run a minimum computer configuration although sometimes not very efficient</td>
</tr>
</tbody>
</table>
FACTORS TO CONSIDER IN CHOOSING BETWEEN A CD-ROM AND ONLINE DATABASE

- DATABASE CONTENT BASED ON USER NEEDS
- SOFTWARE RETRIEVAL CAPABILITIES – SOME CD-ROMS REQUIRE SEVERAL STEPS FOR SEARCHING AND PRINTING COMPARED TO AN ONLINE VERSION FOR E.G.
  - ONLINE VERSION OF SCIENCE CITATION INDEX REQUIRES 13 STEPS FOR AN AUTHOR SEARCH ON CD-ROM EDITION.
  - ONLINE VERSION REQUIRES JUST 4 STEPS
- RESPONSE TIME
- INDEXING- INCREASES SPEED
- DATABASE SEGMENTATION – LARGE ONLINE DATABASES REQUIRE SEVERAL CD-ROMS FOR PUBLICATION (I.E. SEGMENTATION). THUS A USER HAS TO SWAP DISCS TO COMPLETE A SEARCH ON THE WHOLE DATABASE
COMMON SEARCH FEATURES AVAILABLE IN CD-ROM DATABASES

- KEYWORD SEARCH
- PHRASE SEARCH - USE OF MULTIPLE WORD PHRASES
- BOOLEAN SEARCH - AND, OR, NOT
- TRUNCATION
- INDEX AND THESAURUS SUPPORT
- PROXIMITY SEARCH
- FIELD SPECIFIC SEARCH
- FREE-TEXT SEARCH
- COMBINING SEARCH SETS AND SEARCH REFINEMENTS
- LIMITING OR RANGE SEARCH
- SEARCHING THROUGH RETRIEVED RECORDS
FACTORS TO BE CONSIDERED IN THE SELECTION AND EVALUATION OF ONLINE SEARCH SERVICES

- DATABASES OFFERED
- SEARCH AND RETRIEVAL FACILITIES
- SEARCH INTERFACE, PROVISION OF SIMPLE AND ADVANCED USER INTERFACES.
- COST
- TIME REQUIRED FOR SEARCHING
- DATABASE STRUCTURE AND RECORD FORMATS
- CROSS-DATABASE SEARCHING FACILITY
- COMMUNICATION FACILITIES
- SUPPORT SERVICES
- CURRENT AWARENESS SERVICES, SELECTIVE DISSEMINATION OF INFORMATION SERVICES
ASSIGNMENT

• READ AND MAKE NOTES ON THE SEARCH FEATURES OF CD-ROM DATABASES (CHOWDHURY, PP. 340-342)