

1996

Upgrading search strategy skills: unit one: developing search strategies for online searching: a continuing professional development course for library and information service managers

Gulten S. Wagner
Edith Cowan University

Follow this and additional works at: <https://ro.ecu.edu.au/ecuworks>



Part of the [Library and Information Science Commons](#)

Wagner, G., (1996). *Upgrading search strategy skills: unit one: developing search strategies for online searching: a continuing professional development course for library and information service managers*. Perth, Australia: Edith Cowan University.

This Book is posted at Research Online.
<https://ro.ecu.edu.au/ecuworks/6754>

Edith Cowan University

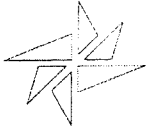
Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement.
- A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.



AUSTRALIAN
LIBRARY AND
INFORMATION
ASSOCIATION



EDITH COWAN
UNIVERSITY
PERTH WESTERN AUSTRALIA

Upgrading Search Strategy Skills

Unit One

Developing Search Strategies for Online Searching

A Continuing Professional Development Course
for Library and Information Service Managers

Dr Gülten S. Wagner
Edith Cowan University

Copyright © 1996

Edith Cowan University

All rights reserved. Except for the quotation of short passages for the purpose of criticism and review, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without written permission of Edith Cowan University.

ISBN 0-7298-0281-7

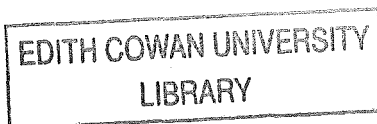


Table of Contents

INTRODUCTION	5
Unit Outcomes	5
Library Competency Standards	6
Legal and Ethical Dimensions of Online Search Strategy Development	6
01 WHAT IS A SEARCH STRATEGY	7
Objectives.....	7
Competency Standards	7
Introduction	8
The Search Process	9
Translating the Information Request - Identification of Key Concepts	9
Boolean Logical Connectors	11
Summary	19
02 WHAT IS AN ONLINE SEARCH?	21
Objectives.....	21
Competency Standards	21
Introduction - Seven Stages of an Online Search	22
Types of Online Databases	25
Creating a Database	27
Controlled Vocabulary	28
Using Multiple Terms to Define a Single Concept.....	30
Precision and Recall	30
Strategies for Narrowing and Broadening a Search	31
Truncation	32
Proximity or Contextual Operators	32
Evaluation of Search Strategy Development Results	34
Summary	34

03 ONLINE SEARCH TECHNIQUES	35
Objectives	35
Competency Standards	35
Introduction	36
Search Strategy Construction Styles	36
More Complex Search Techniques	40
Summary	43
03 Hands-On Tutorials - Using CARL's UnCover	45
Objectives	45
Competency Standards	45
Introduction	45
CARL's UnCover	45
Signing on to CARL's UnCover	46
UnCover Search Features	46
Summary	52
A Select Bibliography	53
Appendix 1	54
Appendix 2	55
Appendix 3	63

INTRODUCTION

This unit provides information professionals with some skills and knowledge in the retrieval and provision of access to required information that is available online. The unit consists of four Modules. Students seeking accreditation are required to complete the assignment (Appendix 3) and submit it for assessment.

Unit Outcomes

By the end of this unit, participants will:

- understand the basic concepts involved in developing a search strategy in online an information retrieval environment;
- be able to apply concepts of Boolean logical operators in constructing a search strategy;
- determine which elements in the search strategy are essential key concepts;
- be familiar with various approaches used in developing a search strategy;
- develop an online search strategy to meet an information need;
- appreciate the advantages of vocabulary control in online search strategy development;
- understand the concepts of precision and recall;
- appreciate the legal and ethical requirements of online searching;
- apply evaluation techniques in assessing the outcome of a search.

Library Competency Standards

This course addresses mainly the competencies explained in Unit 49 (Level 6) *Provide clients with specialist and complex access to information* and Unit 17 (Level 4) *Contribute to client access to information* of the Library Competency Standards developed by Arts Training Australia. However, the competencies developed by Unit One are far reaching. Aspects of the following competency units are affected to some extent:

- Unit 16 (Level 4) Obtain material from remote sources for clients
- Unit 27 (Level 4) Provide research assistance
- Unit 38 (Level 5) Improve accessibility of information
- Unit 45 (Level 5) Collect, analyse and evaluate information for research
- Unit 57 (Level 6) Provide specialist technological support for information access; and
- Unit 65 (Level 7) Manage information access

Legal and Ethical Dimensions of Online Search Strategy Development

It almost goes without saying that search strategy development in an online information retrieval environment has to take place within the parameters of copyright and intellectual property legislation. Although these two very important dimensions fall outside of the immediate instructional focus of this course, it is useful to point out that all online search strategy development activities for this course should be conducted within the legal and ethical boundaries of accessing database systems.

01

WHAT IS A SEARCH STRATEGY?

Objectives

By the end of this module you should be able to:

- define search strategy;
- identify the essential decision points in formulating a search strategy;
- understand the basic components of a search strategy;
- identify key concepts in an information request; and
- apply Boolean Logical operators to combine key concepts of a search statement.

Competency Standards

This part of Unit One provides background information for issues raised by units 49 and 17 of the Library Competency Standards developed by Arts Training Australia.

Introduction

Information retrieval systems are complex systems requiring specialised training. The process of translating an information request into the language of a computer in an online, CD-ROM or network environment is called a search strategy development. Information retrieval systems operate on the principle of a set of matching functions taking place at high speed. The skill and knowledge of the searcher are the main determinants in obtaining the best results from an information retrieval system. A good searcher should have a variety of competencies. These fall into four main categories:

1. General principles of searching such as a general knowledge of information retrieval; human communication skills relating to reference interviews and problem solving; search strategy formulation skills, and, evaluation of search results.
2. Subject knowledge, including the structure of knowledge in the disciplines in which searches are conducted. Knowledge of the database content and construction is also required.
3. System related skills - such as awareness of search software protocols, correct entry forms of search terms and establishing logical combinations on the system.
4. Technical skills in the use of computers and telecommunication technology.

The main objective of this course is to develop search strategy formulation skills in three information retrieval environments: online, CD-ROM and World Wide Web.

The Search Process

Regardless of the information retrieval environment - online, CD-ROM or World Wide Web - conducting a search contains several distinct stages such as clarifying the search request, deciding on information sources, translating the information request by identifying its key concepts, conducting the search and evaluating the results. In this Module the emphasis will be on the translation of an information request into the language of online database systems and Boolean logic.

Translating the Information Request - Identification of Key Concepts

The search strategy performs two important tasks: it explicitly identifies the concepts in an information request and it also explicitly states the relationships between those concepts. The latter task is made possible with Boolean logical operators.

The first step is to identify correctly the concepts in an information request. For example the statement of "I would like to find some information on the New Zealand Cricket Academy" is a simple search request containing only one concept: the New Zealand Cricket Academy. No other concepts are implied. Whereas, a request of "I am looking for information on breeding alligators" contains two concepts. Breeding and alligators need to appear in the same article. This is an AND relationship. An information request for "Articles on company intelligence or corporate intelligence" presents an OR relationship between concepts. Articles either on company intelligence or corporate intelligence will satisfy this search request.

Traditionally a standard search strategy form is used by searchers in developing a search strategy. All of the essential components of such a form are provided in Appendix 1. (This form can be copied to make additional forms, as a new form is necessary for each search request). Use this form to develop your search strategies.

ACTIVITY I

Indicate the number of concepts in the following information requests:

1. Find some articles on pesticide caused eggshell thinning in birds.
2. Articles on the efforts to re-establish sea otter populations.
3. Find publications on making facilities accessible to the handicapped.
4. Find articles on drug abuse or drug addiction in primary and secondary schools.
5. Find information on video games.

Boolean Logical Connectors

Search strategy construction in the online, CD-ROM or the World Wide Web environment depends on Boolean logical operators. Boolean logic using AND, OR and NOT operators were first introduced by George Boole, the mathematician in 1847. Ironically, Boolean logic remains the main information retrieval tool since its introduction nearly one hundred and fifty years ago, in spite of advancements in information retrieval science. It represents symbolic relationships between concepts or sets. Relationships between concepts can also be viewed in pictorial form using Venn diagrams (See Figures 1, 2, 3 and 4).

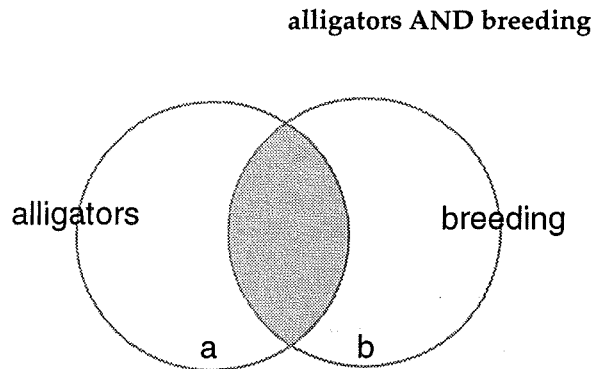
AND logical operator, or connector, requires that the identified search terms that it connects must be present in all documents. In other words, every keyword connected by an AND logical operator must be present for a record to qualify.

OR logical operator, or connector, will retrieve documents containing either one or both of the search terms it connects. In other words, the occurrence of any one of the keywords connected by an OR logical operator will qualify the record to be retrieved.

NOT logical operator, or connector, will eliminate any documents containing keywords connected by the NOT operator.

Examples of Boolean logical connectors:

1. "I am looking for information on breeding alligators" search statement has two concepts: alligators and breeding as indicated in the following Venn diagram:

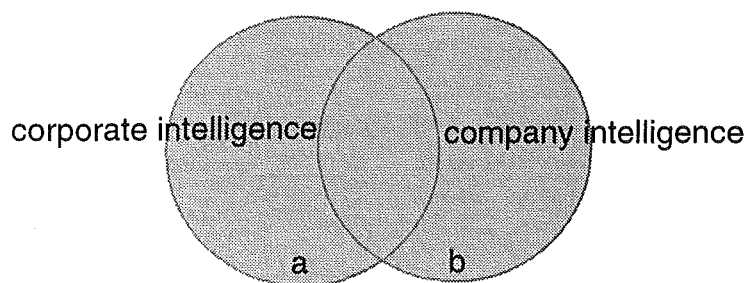
**Figure 1**

The shaded area of the Venn diagram represents items retrieved using the logical connector AND

2. "Articles on corporate intelligence or company intelligence" search statement is represented in Figure 2 below:

corporate intelligence OR company intelligence

The shaded area indicates the retrieved articles by this search statement.

**Figure 2**

Alternatively, this statement can also be expressed as follows:

intelligence AND (corporate OR company)

In this example parentheses are used to clarify relationships between concepts.

The Venn diagram for this alternative expression is slightly different as shown in the figure below:

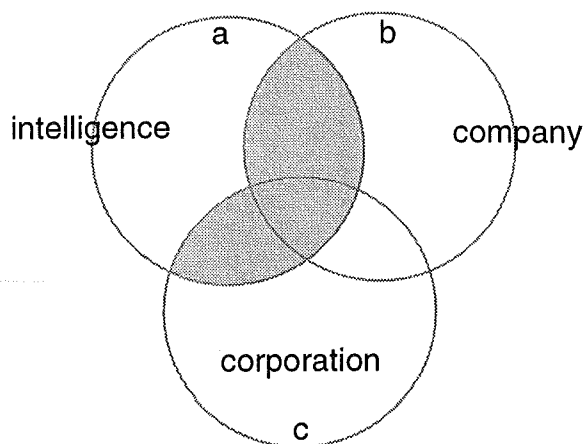


Figure 3

- 3. "Articles on company intelligence but not computers" search statement is represented in Figure 4 below:

company intelligence BUT NOT computers

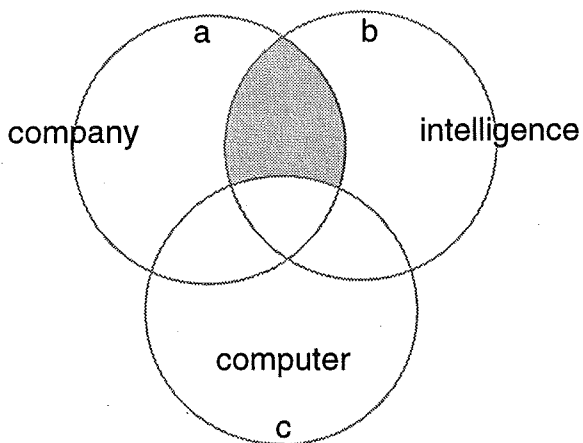


Figure 4

Boolean Commands and Their Order of Execution

It is important to know the order of processing in an online system. For example in the Dialog system the processing order is:

Parentheses ()

NOT

AND

OR

This understanding is essential for building correct logical expressions.

Nesting of Boolean Logical Operators

Nesting is the term used for the facility that allows a searcher to use several Boolean operators in the same sentence without creating any confusion. Nesting is possible by means of parentheses. Parentheses help to indicate the order of execution of logical connectors as they are first executed by the system. It is then possible to separate ORs from ANDs:

(a OR b) AND c

In this expression either terms **a** or **b** or both will be ANDed with the term **c**. Whereas this expression will be interpreted differently if it is expressed without the parentheses, thus giving a different result:

a OR b AND c

First terms **b** and **c** will be ANDed according to the rules of execution in the Dialog system. The computer will then retrieve records containing either **a** or the intersection of **b** and **c**; or both **a** and the intersection of **b** and **c**.

Consider these two expressions:

(sun OR solar) AND energy
and
sun OR Solar AND energy

In the first example, records containing the concepts of sun or solar energy will be retrieved. The second expression is not correct as the computer will combine solar AND energy, then it will retrieve records with only the sun concept.

ACTIVITY 2

State if these search statements are correct or incorrect:

Search Tip:

Note that **Parentheses ()** are processed first, followed by **ANDs** then by **ORs**.

1. (dose OR dosage) AND radiation

Explain why?

2. radwaste OR radioactive AND waste

Explain why?

3. dose OR dosage AND radiation

Explain why?

Stop Words

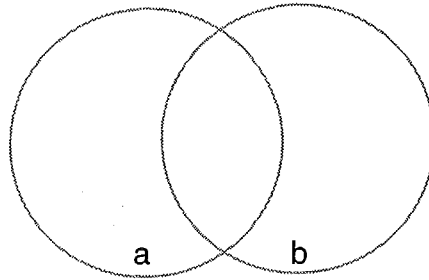
In a retrieval system although all of the words occurring in records are indexed, some very commonly occurring words are not. These are known as stop words. Their number is surprisingly low, for example there are only nine in the Dialog system:

an	for	the
and	from	to
by	of	with

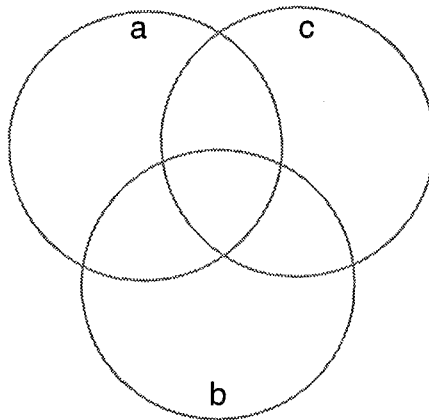
ACTIVITY 3

Express the following statements by shading appropriate sections of Venn diagrams:

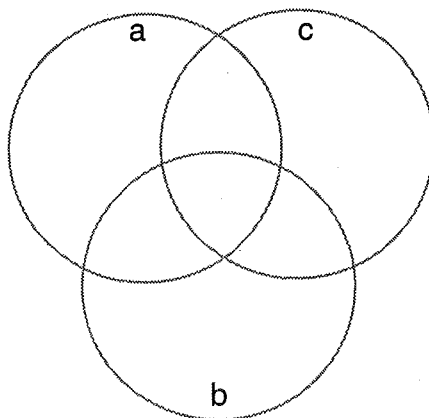
1. $a \text{ OR } b$



2. $(a \text{ AND } b) \text{ NOT } c$

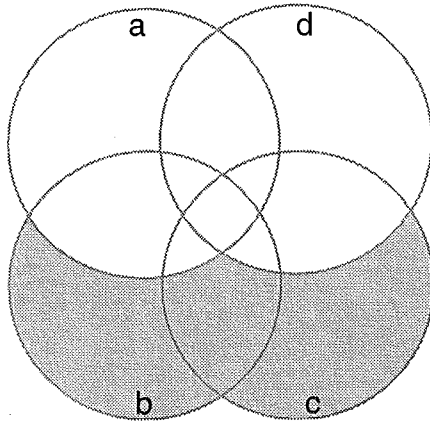


3. $(a \text{ OR } b) \text{ AND } c$

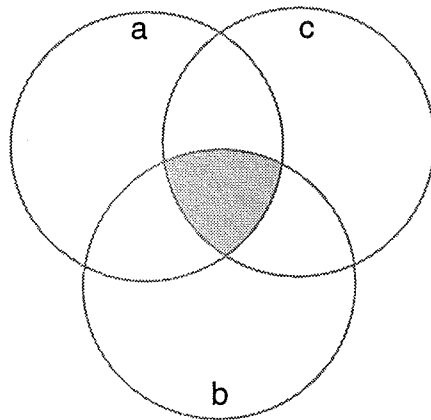


Express the following Venn diagrams as search statements using a,b, c or d. Retrieved sections are shaded.

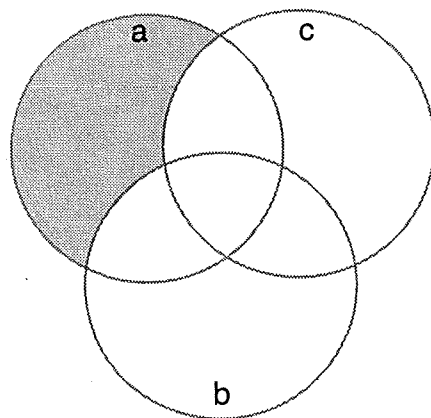
4.



5.



6.



ACTIVITY 4

Express the search statements below by using Venn diagrams.

1. Find some recent articles on pesticide caused eggshell thinning in birds.
2. Articles on the efforts to re-establish sea otter populations.
3. Find publications on making facilities accessible to the handicapped.
4. Find articles on drug abuse or drug addiction in primary and secondary schools.
5. Find information on video games.

Summary

This Module establishes the basics of online searching. It provides the ground for two essential aspects of online searching: Boolean logic and translating an information request to the language of computers. These two dimensions of searching constitute the heart of online information retrieval.

02 WHAT IS AN ONLINE SEARCH?

Objectives

By the end of this module you should be able to:

- understand the basic concepts involved in developing an online search strategy;
- identify key concepts in a search request;
- be familiar with strategies used for narrowing and broadening a search;
- understand the concepts of precision and recall;
- appreciate the role of controlled vocabulary in online information retrieval.

Competency Standards

This part of Unit One provides background information for issues raised by units 49 and 17 of the Library Competency Standards developed by Arts Training Australia.

Unit 17

Element	Performance Criteria
1	2,3
2	2,3
3	2
4	2

Unit 49

Element	Performance Criteria
1	1,2,3,4
2	1,2,3,4
3	1,3,4,5,6

Introduction - Seven Stages of an Online Search

An online search is a set of strategies that are used to retrieve relevant information or materials from online systems. Normally an online search involves seven stages, some take place offline such as the reference interview. Most writers such as Quint (1992: 9-12) agree on the following seven stages of an online search:

- Reference Interview
- Strategic Overview
- Database Selection
- Search Strategy Formulation
- Seeking Feedback on Reviewing Results
- Presenting Final Search Results
- Evaluation of Search Strategies Results

The Reference Interview

The Reference Interview is unquestionably the most critical point of the online search process. It is at this point that the information need is clarified. All tools of interpersonal communication - verbal and non-verbal come into play to elicit the real information need. Three components of the process are: the requestor; the request and expectations of the requestor.

The Requestor

- Who is the requestor?
- What is the requestor's background vis-a-vis the search? (Is the requestor an expert in the area of the request?)
- What output format is required? (How the results will be sent, by facsimile, on a disk or over the Internet?)

The Search Request

- What is the question? What does the requestor already know regarding this request? (Key sources, key authors/ organisations / citations, etc., will be immensely useful in structuring the search strategy).
- How soon does the requestor need the results?
- How much will be spent for the search?
- What will the requestor do with the search results? (Information on this may be useful in manipulating data for company reports).

The Requestor's Expectations

- What is the requestor's knowledge of the search process?
- What are the requestor's intentions in using the retrieved data? (Does the requestor understand the legal limits of using the data, etc.)?

(Based on Quint, 1992)

Strategic Overview

This stage, which is ideally placed between the reference interview and the database selection stages, offers an opportunity to plan the search process. The searcher needs to make several rather important decisions. Does it warrant an online search - or can the request be satisfied by a manually conducted "offline" search? How will the quality control issues be addressed? In other words, is the requestor an expert to judge the quality of the retrieved data? Or, will it be necessary to go to other sources to evaluate the search results?

Database Selection

This stage makes great demands on the search intermediary. Knowledge of vendor systems and the range of databases they provide, as well as the knowledge of databases covering specific subjects, their value-added features and costs, will play an important role in database selection.

Search Strategy Formulation

Search Strategy Formulation involves three basic and two other related approaches to the search. Donald T. Hawkins and Robert Wager in their seminal article "Online bibliographic search strategy development" (1982) named five approaches to developing a search strategy. They are:

- Building blocks approach
- Citation pearl growing approach
- Successive fractions approach to searching
- Most specific facet first
- Lowest posting facet first.

These will be discussed later in more detail in this module.

Seeking Feedback on Reviewing Results

This is an important interactive communication stage in which the searcher needs to use all of his/her interviewing skills to obtain the necessary feedback on the initial search results. This step may require some amendments to the search strategy.

Presenting Final Search Results

This stage involves some value-added processes. Presenting the requestor with a ripped off print-out computer sheet at the conclusion of a successful session falls definitely short of meeting quality requirements in most information environments. Information brokerage companies have always provided search results with many value-added services, such as adding annotations, sorting records, etc., since the earliest days of online services, in the 1970s.

Evaluating Search Strategy Results

A user's satisfaction with the search results is the ultimate test of an online search session. Verbal or written feedback should be obtained from the user to determine the efficacy of the search. If the results are not entirely satisfactory, problem areas should be clarified with the user before a follow-up search is conducted.

In this Course, the focus is placed on the fourth stage, search strategy formulation.

Types of Online Databases

There are two main types of databases:

- Reference databases; and
- Source databases.

Reference databases point the user to source documents which contain the desired data. These databases are also known as bibliographic databases. They contain records with full bibliographical details (i.e., author, title and source; often accompanied with an abstract, controlled vocabulary terms or descriptors, identifiers and classification codes). Citation databases (i.e., Science Citation Index, Social Sciences Citation Index, etc., also fall into the reference databases category).

Source Databases contain the entire data sought without needing to refer to other sources. Five types of source databases exist: numeric, textual-numeric, full-text databases, dictionary and directory.

- **Numeric and Textual Numeric Databases** may contain census, banking or stock market information. Statistical abstracts and tables provide enough source-type information within the records. Some will have enough textual information to qualify them as textual-numeric databases. Normally data in numeric databases can be manipulated to obtain a desired statistical output.
- **Full-text Databases** contain an entire document. Newspaper and legal databases (e.g., Boston Globe, Business Dateline, Reuters; Lexis and Westlaw) fall into this category.

More and more bibliographic databases are becoming available in full-text version; in addition to providing fully indexed records of these documents. The provision of the full-text version of the document is often offered as a record display option, such as format 9 in the Dialog system.

- **Dictionary Databases** may contain classification codes, registration numbers, molecular formulae, chemical ring systems - such as in the Chemical Abstracts database or a dictionary of taxonomic codes as it is supplied in the Biocodes Database.
- **Directory Databases** provide non-bibliographic records. They supply information about organisations and companies. There are a wide range of directory databases, some are devoted to grants, some to public law notices and Presidential proclamations, as in Fedreg database. In the print trade bibliographies also can be considered as directory databases.

A Sample Online Record from a Bibliographic Database

Title: Use of Personal Space in State and Public Libraries: A Review of Research.
Author: Wagner, Gulten S.
Source: **Research in Space (A Hypothetical Journal)**, Vol.1, No. 1
Year: 1996
No. of pages: 42p.
Language: English
Type: Review; Research
Country Code: Australia; Western Australia

The author argues that the study of use of personal space is an important aspect of understanding the total communication process in State, Public and Academic libraries, this document reviews the literature on personal space research that has been conducted over the last 20 years. Following a general review of the roots of such research, the numerous studies that are described are presented in eight categories as they focus on: (1) seating preferences in relation to the area of the room and proximity to other people; (2) the correlation between spatial proximity and certain attitudes; (3) use of study carrels; (4) the appropriateness of three-dimensional models of personal space; (5) dyadic boundary adjustments; (6) development of models of personal space (the stress model and the systems model); (7) cultural, gender, and age differences with respect to the use of space; and (8) space utilization between dominant versus subordinate personalities, and acquaintances versus strangers. Most of the studies reviewed were conducted in State, Public and University libraries in Australia.

Major Descriptors: *Interpersonal Communication; *Personal Space;
 *Library Facilities; *Library Research; *Use Studies

Descriptors: Public Libraries; State Libraries; University Libraries;
 Environmental Research

Identifiers: Australia; Western Australia

An Online Record Structure From a Directory Database

Books in Print (File No. 470) in the Dialog system is an example of a directory database. It provides:

Full title
Full author
ISBN
Status information - i.e., Active record or New (with date information)
Publisher
Publication date: (month/year)
Number of pages
Binding: i.e., Trade cloth and Retail Price
Notes: - i.e., illustrated
Date in file:
Library of Congress Headings
Review(s): - if any reviews are published the following is provided:
A short abstract of the review article
Source information
Journal announcement information

Creating a Database

In building a computer based database, records are manipulated to create inverted files. Two main types of indexes exist: the basic index and special indexes. The basic index is an alphanumeric list of all words, with the exception of stop words, taken from the title, abstract and subject indexing fields of a record. In other words the basic index contains:

- keywords in titles
- keywords in abstracts
- descriptors
- identifiers
- keywords contained in descriptor phrases
- keywords contained in identifier phrases

Each keyword is tagged for identification with the accession number of the record; name of the field where it belongs such as title or abstract; even its exact location in the field is noted. Inverted files present these words/keywords in alphabetical order. Thus searching a database takes place in these files. For example, proximity of terms can be established because each word/keyword carries a virtual address as part of its identification.

In addition to basic indexes, databases normally contain special indexes such as author, document type, language, publication year, etc. Information for additional indexes come from non-subject fields. A field prefix code consisting of a two character abbreviation followed by an equal sign is used to signify their type - such as:

AN= Accession number

AU= Author

DT=Document Type -(Such as DT=Journal Article; DT=Research Report)

JN= Journal Name

LA=English

PY=Publication Year

Unfortunately databases often use non-standard coding, e.g., **PY= Publication Year** as well as **PD=Publication Date** are used.

Controlled Vocabulary

Database creators frequently develop a list of subject terms that are selected to represent concepts. These are usually made available online, as well as in a print thesaurus such as the *ERIC Thesaurus of ERIC Descriptors*.

There are many types of controlled vocabularies in databases. In fact a single database often contains several types. The most common types of controlled vocabularies are classification codes, descriptors and subject headings.

An example taken from the *ERIC Thesaurus of ERIC Descriptors* below illustrates how indexers define concepts by establishing hierarchical relationships between them. For example the concept of "information needs" has fifteen related terms or descriptors:

Information needs

UF (It is a Used For symbol)	Data Needs
BT (A Broader Term for Information Needs)	Needs
RT (15 Related Terms for Information Needs)	Access to Information Evaluation Needs Hotlines (Public) Information Management Information Networks Information Seeking Information Services Information Sources Information Systems Information Utilisation Online Searching Research Needs Search Strategies Selective Dissemination of Information User Needs (Information)

In this example a narrower term (NT) does not appear. According to *ERIC Thesaurus of ERIC Descriptors* "librarians" is a narrower term (NT) for "Informations Scientists".

ACTIVITY 5

Design a search strategy for a request on "Information needs" incorporating some of the related concepts provided by the *ERIC Thesaurus of ERIC Descriptors* as listed above:

Search request:

Could you please find me some articles on the information needs of teachers?

(Use a copy of the Online Search Strategy Development form provided in Appendix 1. **NB:** This form is an abbreviated one. A fully developed form would contain other headings such as: client details, client provided - keywords, relevant titles; limitation parameters by date, language, geography, cost and others. To view a typical online

search request form see a standard textbook such as the most recent edition of *Online Searching Principles and Practice* by R.J. Hartley, E.M. Keen, J.A. Large and L.A. Tedd (London: Bowker-Saur).

Search Strategy:

Using Multiple Terms to Define a Single Concept

It is often necessary to use synonyms or alternative terms to define a single term, as computers can only match terms letter by letter; they cannot search "concepts". Furthermore, indexers and authors are likely to use descriptors that are different from a user's. Using multiple terms increases the chances of retrieving relevant records from databases. For example, to search the concept of **oil spillage** in the Indian Ocean it would be useful to include oil leaks and leakage as well as oil spills for the sake of comprehensiveness:

oil spillage indian ocean
 spills
 leakage
 leaks

Variation of terminology such as different spellings of words, e.g., **aluminum** or **aluminium**; acronyms and abbreviations, e.g., **AIDS** and word segmentation, e.g., **database** or **data base** should be considered in developing a search strategy.

Precision and Recall

The objectives of each search are different. Some search objectives call for a narrow search with very few but highly relevant records. Others will require a high number of records. In search strategy terminology this dichotomy is expressed as high precision versus high recall. High precision or high recall can be achieved by narrowing or

broadening the search parameters. Maximising the retrieval of relevant records while minimising retrieval of irrelevant records is described as maximising both recall and precision. However, there is often a negative correlation between achieving these objectives. In other words, if the precision is very high the recall will be lower; alternatively, if the recall is very high the precision will tend to be lower. The formulae below illustrate the ratios between recall and precision:

$$\text{Recall ratio} = \frac{\text{Relevant Records retrieved}}{\text{Total relevant records in database}}$$

$$\text{Precision ratio} = \frac{\text{Relevant Records retrieved}}{\text{Total records retrieved}}$$

Strategies for Narrowing and Broadening a Search

Strategies for narrowing a search include the use of:

- Descriptors
- More AND logical operators
- Specific terminology
- Field-specific searching, i.e., limited to title (ti), description (de) abstract (ab) fields

Example: If too many records are retrieved by the search statement:
search (w) strategy

try to limit it by restricting it to title or description fields only:
search (w) strategy / ti,de

Strategies for broadening a search include the use of:

- Truncation
- More OR logical operators and fewer AND operators
- Full text searching
- Broader terms
- Expanding facility to seek alternative spellings or to see how concept terms appear in the database index
- Looser proximity operators between terms such as (1-7w) or (1-7n)
- Eliminating non essential concepts

Example: If too few records are retrieved by the search statement:
 information (w) literacy
 try to broaden it by adding a similar concept
 (information OR computer) AND literacy

NB: The use of proximity operators such as (w) or (n) will be described in more detail later in this module.

Truncation

Computers have the capacity to mask the stem or a portion of a word and to retrieve any character(s) in that masked space. This is called truncation. Most retrieval systems provide a truncation feature. Unfortunately truncation symbols vary in different online systems - such as a question mark (Dialog), a hash #, a colon (BRS) or a dollar sign (\$) (Ozline). Truncation is a short cut expression of OR logic as it combines a variety of terms, any of which is acceptable, without spelling all the variations of it. The most common type of truncation is applied to the stem of a word.

For example **gene?** as a truncated expression will retrieve the following:

genetic
genetical
genetically
genetics

Another very common use of truncation occurs in author searching. This facility helps to determine an author's name and its various usage. (Authors often use a variety of styles in presenting their names - fully spelled, with some or all of the initials provided. For example some articles by Gulden S. Wagner have been signed as Gulden Wagner, G. S. Wagner and Gulden S. Wagner over the years. In Dialog's databases this problem is solved by means of truncation: **au=wagner, g?**). In some cases an internal truncation, e.g., **organi?ation** will be useful in covering Australian, British and American spellings of the word.

Proximity or Contextual Operators

The facility to nominate positioning of the words and phrases in records is an added advantage in developing a search strategy. Contextual or proximity operators position terms or concepts in records. If you are trying to locate information on information

dissemination, the **AND** logical operator will not retrieve the most relevant items. What is needed is a proximity or a contextual operator to indicate that these two terms should be located next to each other in a prescribed order. Tools for this operation are called proximity operators. The exact terminology depends on the vendor system used. For example in the Dialog system the **WITH** operator (**w**) as shown in the example below is a facility for adjacency; **nW** indicates that up to **n** (i.e., $n=1-7$) words may intervene between them, read from left to right) :

information(w)dissemination

will retrieve these words when they occur adjacent to each other, in this specified order. Whereas, **NEAR** (**n**) will retrieve the two terms that must occur next to each other but in any order (i.e., read from left to right or from right to left). Intervening words are allowed between the terms with a number placed in front of the **WITH** or **NEAR** - e.g., (**1w**) or (**1n**) as explained above. For example:

information(1n)dissemination

will retrieve the two terms that must occur next to each other but in any order:

information dissemination

as well as:

dissemination of information

The intervening words facility is especially useful in dealing with stop words such as in: **management(1w)objectives** for "management by objectives" and **wash(1w)wear** for "wash and wear". In this case (**1w**) is masking "and" so that it will not be confused with the Boolean logical operator **AND**.

Some proximity operators specify positions of terms in specified fields such as:

term(f)term = will retrieve terms occurring in the **same field**

term(s)term = will retrieve terms appearing in the **same sentence**

Other systems such as Ozline and BRS use **ADJ**, **SAME** and **WITH** to provide various proximity positioning between words.

Evaluation of Search Strategy Development Results

Evaluation of search strategy development results is crucial for the success of the online information retrieval process. The results of an online search need to be assessed not only in terms of its success in meeting a client's information needs but also for the validity and reliability of the information provided.

Summary

This second Module investigates some of the most important features of online searching such as reference interviews, types of online databases, database features, precision and recall, truncation and narrowing and broadening a search. The evaluation process is identified as an important component of search strategy development.

References

- Quint, B. (1992). Inside a searcher's mind: the seven stages of an online search: part 1 & 2. *In Online search strategies & Tips: the best of online & database '88-'89*. Compiled by N. J. Garman. Wilton, CT: Eight Bit Books. pp. 9-22
- Hawkins, D.T. and Wagers, R, (1982). Online bibliographic search strategy development. *Online*, 6 (3) May p. 12-19.

03 ONLINE SEARCH TECHNIQUES

Objectives

By the end of this module, you should be able to:

- understand the basic models of search strategy construction;
- understand the advantages and disadvantages of five different approaches to search strategy development; and
- apply some of the more complex techniques in developing online search strategies.

Competency Standards

This part of Unit One provides background information for issues raised by units 49 and 17 of the Library Competency Standards developed by Arts Training Australia.

Unit 17

Element	Performance Criteria
1	2
2	3

Unit 49

Element	Performance Criteria
2	1,3,4
3	1,2,3,5,6

Introduction

In this module various online searching techniques will be discussed, starting with search construction styles. More complex searching methods such as: expanding; prefix and suffix searching; accessing multiple databases all at once will also be covered.

Search Strategy Construction Styles

It is very important to point out that choice of search strategy development styles depend entirely on the nature of the client's information needs. Not all search requests require access to online systems, some queries may be answered by looking them up in a print directory or encyclopedia or a CD-ROM database. In these cases, the client may require some assistance in the use of these resources.

There are five basic models of constructing search strategies based on a correct assessment of the client's information needs. They are:

1. Building Blocks Approach
2. Citation Pearl Growing Approach
3. Successive Fractions Approach to Searching
4. Most Specific Facet First
5. Lowest Posting Facet First

Building Blocks Approach

The Building Blocks search strategy development method provides a modular approach to the search process. Each aspect of a search is developed separately, as if each is a discrete search on its own. The final logical statement assembles these subsearches. Many searchers prefer this style. It is basically a method of dividing a problem into sub-problems and conducting them as mini searches.

Building Blocks Search Strategy Example:

Search Topic: Documents are required on the identification of athletically gifted children.

<u>ATHLETICALLY</u>	<u>GIFTED</u>	<u>IDENTIFICATION</u>
ATHLETIC	GIFTED	IDENTIFICATION
ATHLETICAL	TALENT	IDENTIFIED
ATHLETICALLY	TALENTED	IDENTIFIES
ATHLETICS	TALENTS	IDENTIFY
SPORT		IDENTIFYING
SPORTING		
SPORTS		

Search strategy

Building Block One

- 1 ATHLETIC?
- 2 SPORT?
- 3 COMBINE 1 OR 2

Building Block Two

- 4 GIFTED
- 5 TALENT?
- 6 COMBINE 4 OR 5

Building Block Three

- 7 IDENTIF?

Final BuildingBlock

- 8 COMBINE 3 AND 6 AND 7

The Building Blocks Approach has several advantages:

1. It provides a clearer trace of search history.
2. It is logical, easy to follow and to review.
3. It provides a logical progression of the search.

On the other hand the Building Blocks Approach has the following disadvantages:

1. It often uses more online time than necessary.
2. Its highly structured approach means that it needs to be pre-planned and cannot take advantage of the Interactiveness and spontaneity of online searching. The logical flow of the Building Blocks Approach may discourage more effective interaction with an online system.

Citation Pearl Growing Approach (or Spiral Search Development Technique)

This method uses the interactive capabilities of online searching. The search starts with a known highly relevant and pertinent citation. Appropriate indexing terms are searched to reveal other relevant keywords. The process of reviewing index terms from these additional records and incorporating them into the search can be continuous, this is how "citation pearl growing" or "spiral technique of searching" derives its name. A brief search is conducted first to retrieve at least one relevant record.

ATHLETIC AND GIFTED AND IDENTIFICATION

Descriptors of the first retrieved record are reviewed for possible inclusion of other descriptors or terms.

RECOGNITION and SPORTS concepts can be incorporated as a result of this review:

(ATHLETIC? OR SPORT?) AND (IDENTIF? OR RECOGNIT?) AND GIFTED

Thus, the known article provides additional relevant keywords.

Citation Pearl Growing Approach is a truly interactive method of searching but it may also take longer than necessary time online.

Successive Fractions Approach to Searching

The search starts with a large set on a general topic, then specific concepts are intersected until the search statement is reduced to a desirable size.

The search above can be limited to EJ (Index to Journals in Education) sub-database of ERIC database, for example for dividing the larger database. It may be fractured further by limiting to ERIC's Clearing House that specialises in the indexing of "Exceptional Children" with the CH=EC facility.

- 1 IDENTIF?
- 2 SPORT?
- 3 TALENT?
- 4 COMBINE 1 AND 2 AND 3
- 5 CH=EC [This code stands for ERIC's Exceptional Children Clearinghouse]
- 6 COMBINE 4 AND 5

One of the most important advantages of this strategy is that it allows the search to be completed at an earlier point than the building blocks approach. The process of fractioning eliminates back tracking. The "NOT" Boolean operator can also be used as a device of "successive fractions", as follows:

Topic: A search for articles on obesity in children, but do not include babies.

- 1 OBESITY
- 2 CHILD?
- 3 INFANT?
- 4 COMBINE 1 AND 2
- 5 COMBINE 4 NOT 3

Most Specific Facet First

This method saves time and costs as the specific term is searched first in a multi faceted search. The next term or concept can be combined with the most specific term. If the postings (i.e., the number of retrieved records) remain low, the searcher may elect not to impose any further restrictions.

Economies of scale can be achieved, by the Most Specific Facet First, if this is the most appropriate method for a search problem. This approach would be suitable for larger databases, such as Chemical Abstracts, BIOSIS, etc.

Lowest Postings Facet First

This approach is somewhat similar to the previous one. Concepts with lower postings are entered first. If the combined results come up with very low posts, no further restrictions would be needed. For example, for searching cases of malpractice involving breast implants in a legal database, concepts of implants would be entered before legal aspects. In searching medical databases this order would be reversed.

Advantages and disadvantages of this method are the same as the Most Specific Facet First discussed above.

More Complex Search Techniques

In this section some of the more complex searching features will be discussed under three headings:

- Prefix Searching and Expanding
- Suffix searching
- Multiple database searching

Activity segments will provide opportunities for practising these techniques in search strategy construction.

Prefix Searching and Expanding

Many of the additional indexes are identified with prefixes which are usually two letter codes such as au=author, cs=corporate source field, dt=document type and rn=report number. It would be necessary to expand the prefix code to establish its exact format. For example,

dt (a prefix for document type) needs to be expanded to discover its exact usage in a given database. The expand command placed after a system prompt in Dialog system as shown below will reveal what document types are provided:

?e dt=

Ref	Items (Retrieved)	Index term
E1		
E2		
E3		
E4		
E5	5000	DT=CONFERENCE ARTICLE
E6	712500	DT=CONFERENCE PAPER
E7	12700	DT=CONFERENCE PROCEEDINGS
E8	99	DT=DISSERTATION

Expanding with an Author Prefix code is usually expressed with the prefix au (the shortened expression for author). The author prefix is always expanded to check the exact format used in that database. Beware of punctuation variations between databases even in the use of commas after a surname. Let us check the author format for Albert Einstein's name in an hypothetical database in the Dialog system.

?e au=einstein a

Ref	Items (Retrieved)	Index term
E1	34	AU=EINSTEIL A
E2	23	AU= EINSTEIL ALAN
E3	79	AU=EINSTEIN A
E4	224	AU=EINSTEIN ALBERT
etc.		

In the example above both e3 and e4 need to be selected as possible variations of Albert Einstein's name.

Suffix Searching

Suffix searching is a facility to limit a search concept to specific datafields according to Basic and Additional Indexes available for a given database. For example a concept can be limited to title fields only for the sake of increased relevance:

information(w)superhighway/ti

Or two or more fields at the same time such as by titles (ti), descriptors (de) and by language (eng):

reactor?/de,ti,eng

In the Dialog system Bluesheets provide listings of prefixes and suffixes for each database with examples for their correct format in search strategy construction.

Multiple Database Searching

Searching multiple databases at the same time provides economies of scale in searching a concept or several concepts simultaneously in ERIC, Library and Information Abstracts (LISA) and International Dissertation Abstracts. However, a word of caution is necessary here. As suffixes and prefixes, controlled vocabulary conventions, and other indexing facilities, differ from database to database, such features should be checked before attempting multiple database searching (or one-search, as it is sometimes called).

Dialog's Dialindex provides a value-added subject searching capacity by clusters of subject groupings. For example SOCSCI encompasses all social sciences databases.

ACTIVITY 6

Develop the following search strategies.

1. Find articles on estrogen; limit these to human subjects only.
2. Locate a few very highly relevant references to blood pressure and stress.

Summary

Online Search Techniques examines some of the main approaches to search strategy development, namely: building blocks approach, citation pearl growing, successive fractions, most specific aspect first and lowest posting aspect first. This section also examines some of the more complex search techniques such as expanding, prefix and suffix searching and limiting searches to specific fields.

04 HANDS-ON TUTORIALS - USING CARL'S UNCOVER

Objectives

By the end of this module you should be able to:

- develop an online search to meet an information need;
- apply necessary database dependent search facilities;
- evaluate the search results.

Competency Standards

This part of Module One provides online hands-on practice for some of the competencies required by units 49 and 17 of the Library Competency Standards developed by Arts Training Australia.

Introduction

In this section CARL's UnCover will be used to provide an opportunity for some hands-on practice in developing a search strategy.

CARL's UnCover

CARL'S UnCover has been providing an index to periodical articles in over 14,000 journals since 1988.

For this unit we will be using the Open Access version which provides an excellent database for searching for journal articles and can be used without an account or password.

UnCover is also available using a password. Selected articles can then be faxed to the requestor. The cost of articles include the copyright charge, the service and facsimile surcharge fees.

Signing on to CARL's UnCover

Step 1.

If you have Telnet access from your organisation, you can Telnet to: **database.carl.org (or 192.54.61.76)** Select your terminal type, such as VT100, depending on your computer system).

This unit uses the Telnet version of UnCover as it has facilities for nominating fields which will develop skills for searching bibliographic data bases.

Step 2.

Choose: **UnCover (Article Access and Delivery)**

Step 3.

Press the ENTER key 4 separate times to reach the main menu. A password or a profile number **is not** required to access CARL's UnCover Open Access catalogue.

Step 4.

To logoff after completing a search , type **//exit**

UnCover Search Features

The UnCover database offers Boolean, name, journal title and article summary search facilities.

w for **word** or **topic** (subject) searching. Words from articles, journal titles, subtitles and article summaries are searchable with a word search. Also, a word search may retrieve the names of people discussed or referred to in the articles.

n for **name** or author searching

b for browsing **journal** titles, then by typing **e** it is possible to browse through recent issues of journals

t command makes it possible to nominate a journal for browsing - for example for browsing some of the recent issues of *American Libraries*:

```
//tamerican libraries
```

Key Commands Used to Access UnCover

// Symbol indicates a **command** function.

A command has three parts as shown in the example below:

```
//winformation superhighway
```

1. two slashes (//) signal the existence of a command;
2. **w** indicates that it is a **word** or a **topic** (subject) search; and
3. the third part (information highway) tells the system which **words** or **topics** to search.

Other UnCover Search Commands

//h A useful command for displaying the search history

q s Used for search short cuts - this eliminates some of menu options for more experienced users.

Boolean Searching

The UnCover retrieval system supports Boolean searching with its AND, OR, AND NOT and BUT NOT operators.

AND Boolean Logic

AND Boolean Operator with WORD or TOPIC (subject) search:

Search Example:

Find documents on adults and literacy.

```
//wadults and literacy
```

OR Boolean Logic

OR Boolean Operator with WORD or TOPIC (subject) search:

Search Examples:

Locate articles on literacy or numeracy.

```
//wliteracy or numeracy
```

Locate documents on genetic counselling.

```
//w(genetic counselling) or (gene therapy)
```

AND NOT or BUT NOT

According to UnCover these logical operators perform the same function.

Search Example:

Locate articles dealing with water found in desert areas, with the exception of the Australian continent.

```
//wwater and desert but not australia
```

Nesting of Boolean Logical Operators

Locate articles on children's literacy or numeracy, but do not include any articles dealing with computing skills.

//children and (literacy or numeracy) but not computing

Other Search Features

More UnCover Search Features

This section provides additional information on the retrieval facilities of the UnCover database.

Displaying listing of citations

Type **d** to display the listing of citations; then press **ENTER** to return to the listing

To start a New Search

To conduct a new search type **q (quit)**

Limiting a Search

To **LIMIT** a search: at the prompt of **NEW WORDS** enter additional words/terms.

Truncation

UnCover facilitates right-hand truncation with an asterisk (*) symbol:

//wsearch*

will retrieve all the words sharing this stem such as searchable, searched, searches, searching, searches, etc.

Search Example:

Locate articles on teenage crime or violence.

```
//wteenage* (crime* or violence)
```

Searching for Authors or Names

Names in UnCover are usually authors listed in Table 3 of Contents of journals. For example,

```
//njames martin
```

looks for articles by James Martin.

A name and word search may be used to search for an author in combination with words from article or journal titles in a single search statement such as:

```
//njames martin/chemistry
```

```
//nspiller/aids
```

```
//neco/semiotics
```

Initials in Names

Initials and full spellings of the name may be covered with the **OR** Boolean operator:

```
//n (robert or r or bobby or b) and kennedy
```

```
//n(john or jack) and blacksmith
```

Joint Authors

The command of **//j** searches for joint authors of articles:

```
//j{s fowell} and {p levy}
```

A word of caution is necessary here, Boolean searches on names work better when retrieved results are less than 500. Therefore it is advisable to refine the search to retrieve a smaller set.

Help Screens

A complete list of HELP screens are available in UnCover by entering ?H2 command. Here is a list of help screen commands.

```
?a ?a2 help for advance searching
?b ?b2 b3 help for levels of Boolean searching
?j help for joint name searching
?h ?h2 help for help screens(?h2 - provides advanced help
features)
?i help for right hand truncation
?m help for marking articles - mainly for ordering
purposes
?n help for name searching
?o help for ordering
?t help for journal title searching
?w help for word searching
?y help information on combined word/name searching
```

ACTIVITY 7

Develop the following search strategies using CARL's UnCover database:

1. Find out what is written on censorship and the Internet regarding children.
2. Search the UnCover database online using its various search facilities such as:
word/topic search
name search
joint name search
truncation
combination of name and topic search
multiple Boolean logical operators
browsing a journal's content pages
3. Write down comments on these online search results.

Summary

The Internet is an immense resource of information retrieval. Search strategy development exercises can be applied to any online information retrieval system. In this section CARL's UnCover database is used.

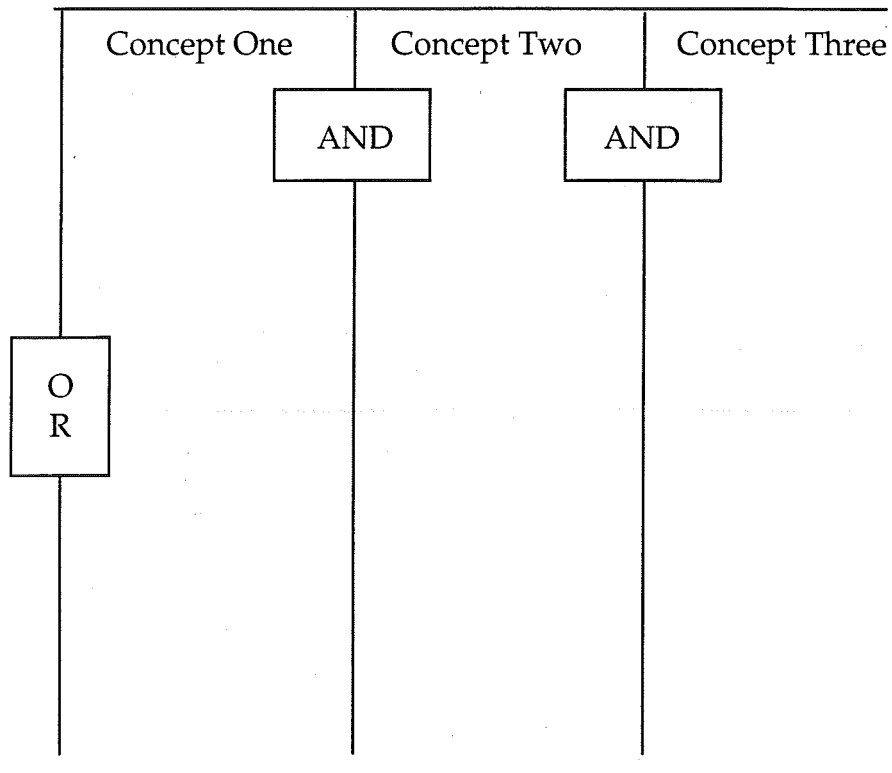
A Select Bibliography

- Bacig, M. (1994). Date searching on DIALOG. *Online*, 18 (2) Mar p. 49-50,52,54-55.
- Basch, R.(1994) Secrets of the super searchers: the accumulated wisdom of 23 of the world's top online searchers. *Online*, 18 (3) May p.100-1.
- Bates, M. E.(1995). Search strategies for Dialog's view fee. *Online*, 19 (1) Jan/Feb p.22-31.
- Cahn, P. (1994). Testing database quality. *Database*, 17 (1) Feb p.23-26, 28-30.
- Germain, J. M.(1994). ERIC Goes Internet. *Online Access*, 9 (1) Jan-Feb p. 50-52,54,58-60.
- Iivonen, M.(1995). Consistency in the selection of search concepts and search terms. *Information Processing & Management*, 31 (2) Mar-Apr p.173-90.
- Jack, R. F.(1995). The LIMIT command. *Searcher*, 3 (7) Jul/ Aug p.20-22.
- Morse, A. and M. L. Pao (1994). A decision tree for end user searching in a large interactive legal database. *Proceedings of the 15th National Online Meeting 1994. Edited by Martha E. Williams, Learned Information, Inc., New York, 10-12 May 1994. Medford, New Jersey: Learned Information, Inc. p.335-43.*
- Morton, D. (1994). Refresher course: expanding your outlook. *Online*, 18 (3) May p.77-9.
- Ojala, M.(1995). Troubleshooting your search: whatever can go wrong, will go wrong. *Online*, 19 (6) Nov/Dec p.59-61.
- Ojala, M.(1995). People who need people: extending the never-ending search for people. *Online*, 19 (1) Jan/Feb p.36-40.
- Pappas, M. L.(1995). Analytical searching. *School Library Media Activities Monthly* 12 (4) Dec p. 35-37.
- Shuman, B. A.(1994). Cases in online search strategy. *Education for Information* 12 (2) Jun p.285-7.
- Tillotson, J. (1995). Is keyword searching the answer? *College & Research Libraries*; 56 (3) May p. 199-206.
- Valenza, J. (1995). Steps to simplifying student searches. *Technology Connection*, 2 (4) Jun p. 21-22.

Appendix I: Online Search Strategy Development Form

Online Search Strategy Form

Search request (expressed in a sentence):



Search strategy

APPENDIX TWO

Answers

ACTIVITY I

Indicate the number of concepts in the following information requests:

- Find some recent articles on pesticide caused eggshell thinning in birds.

Concept 1

Concept 2

Concept 3

eggshell

thin

pesticide

egg shell

thick

ddt

insecticide

dieldrin

etc.

- Articles on the efforts to re-establish sea otter populations.

Concept 1

Concept 2

Concept 3

sea otter
sea otters

population

- Find publications on making building facilities accessible to the handicapped.

Concept 1

Concept 2

Concept 3

handicapped

building
buildings

blind

shop
shops

disabled	store stores
etc.	hotel hotels
	restaurant restaurants
	etc.

4. Find articles on drug abuse or drug addiction in primary and secondary schools.

<u>Concept 1</u>	<u>Concept 2</u>	<u>Concept 3</u>
drug drugs	primary school primary schools	
marijuana	elementary school elementary schools	
substance abuse	secondary school secondary schools	
etc.	high school high schools	

5. Find information on video games.

<u>Concept 1</u>	<u>Concept 2</u>	<u>Concept 3</u>
video video	game games	
virtual reality		
multimedia		

ACTIVITY 2

First state if these search statements are correct or incorrect, then explain why.

Search Tip:

Note that **parentheses ()** are processed first, followed by **ANDs** then by **ORs**.

1. (dose **OR** dosage) **AND** radiation Correct!

Explain why?

Parentheses are processed first, thus the combined set is simply **anded** with the radiation concept.

2. radwaste **OR** radioactive **AND** waste Correct!

Explain why?

ANDs are processed before **ORs**, thus the logical expression is maintained.

3. dose **OR** dosage **AND** radiation Wrong!!

Explain why?

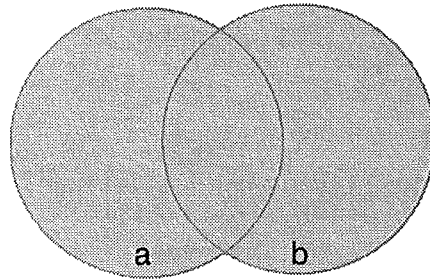
ANDs are processed before **ORs**, thus the logical meaning is spoiled in this case.

Explain how you can remedy this situation.

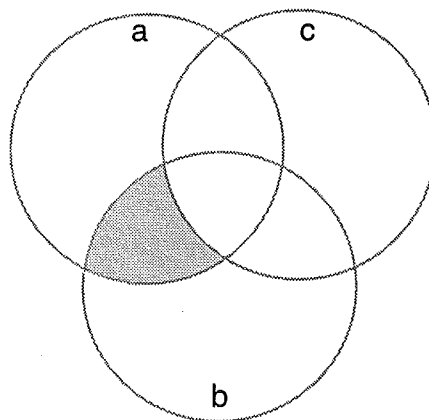
ACTIVITY 3

Express the following statements by shading appropriate sections of Venn diagrams:

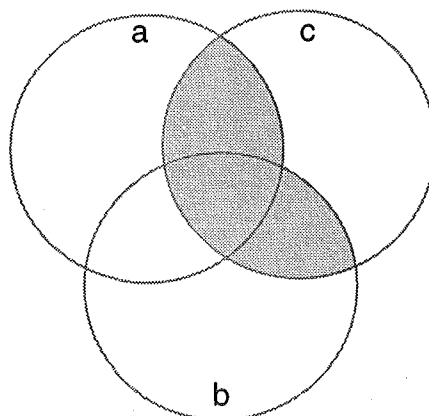
1. **a OR b**



2. **(a AND b) NOT c**

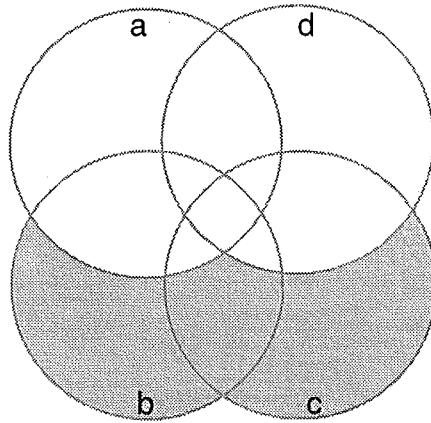


3. **(a OR b) AND c**

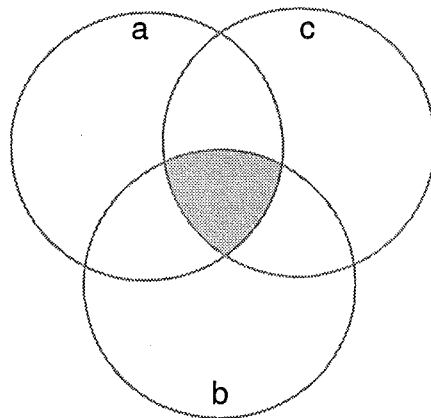


Express the following Venn diagrams as search statements using a, b, c, or d. Retrieved sections are shaded.

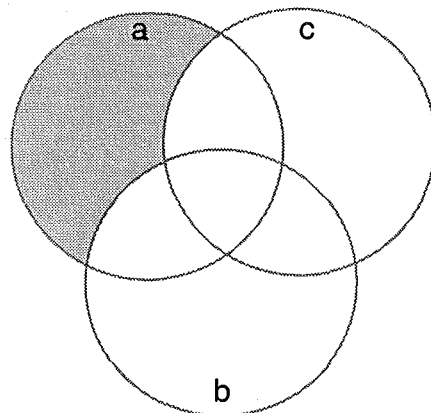
4. (b OR c) NOT (a OR d)



5. a AND b AND c



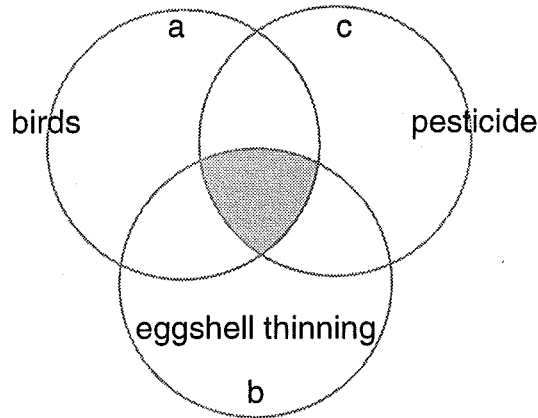
6. a NOT (b OR c)



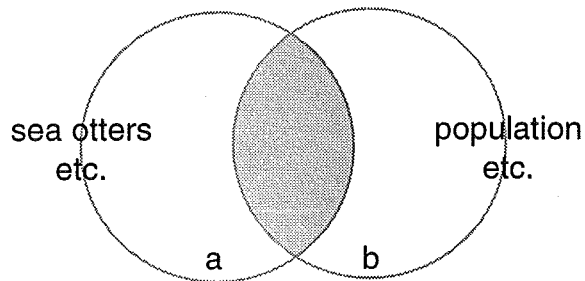
ACTIVITY 4

Express search statements below by using Venn diagrams.

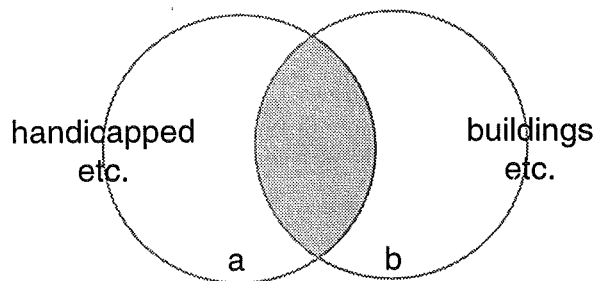
1. Find some recent articles on pesticide caused eggshell thinning in birds.



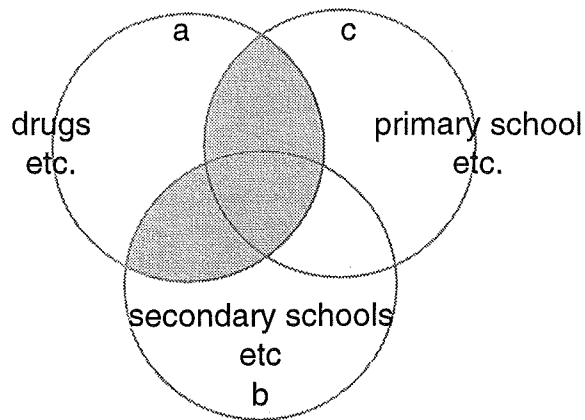
2. Articles on the efforts to re-establish sea otter populations.



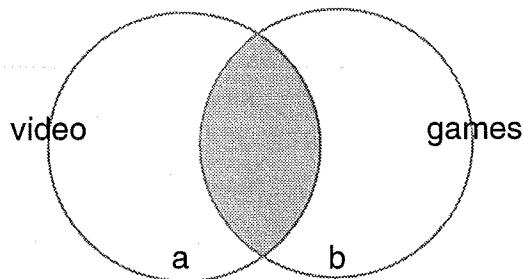
3. Find publications on making building facilities accessible to the handicapped.



- Find articles on drug abuse or drug addiction in primary and secondary schools.



- Find information on video games.



ACTIVITY 5

Design a search strategy for a request on “Information needs” incorporating some of the related concepts provided in the text (just above Activity 5 in the text).

Search request:

Could you please find me some articles on the information needs of teachers? (Use a copy of the Online Search Strategy Development form provided in Appendix 1).

Search Strategy:

(Teacher **or** teachers **or** instructor **or** instructors) **and** (information needs **or** research needs)

ACTIVITY 6

1. Find articles on estrogen; limit these to human subjects only.

estrogen/human

2. Locate a few very highly relevant references to blood pressure and stress.

blood(w)pressure/ti,de and stress?/ti,de

ACTIVITY 7

Develop the first search strategy using CARL's UnCover database:

1. Find out what is written on censorship and the Internet regarding children.

//wcensor* and (internet or net or superhighway) and (child* or juvenile*)

2. AND 3. These tasks will contribute to your Assessable Assignment which is described in Appendix 3.

Appendix 3

Assessable Assignment

The assessable assignment for Unit One: Developing Search Strategies for Online Searching has two sections: Hands-on tutorial exercises and the evaluation of these exercises.

1. Submit the printed results of any three of the hands-on tutorial exercises provided in the Module Four of this unit.
2. Accompany the search results with an evaluation of these searches in the context of the information retrieval facilities of this UnCover database.

The length of the evaluation section should be approximately 500 words.