The value of information education.
The value of Dialog.

INTRODUCTION to DIALOG for INFORMATION PROFESSIONALS

THOMSON™

Table of Contents

INTRODUCTION

This section describes

- The general objectives, intended audience, and setting for this workbook
- Recommendations for accompanying materials for students
- Supporting materials available for instructors

What Is Dialog?

Dialog is a provider of information in electronic form, commonly known as an online information service. The major brand-name services offered by Dialog are Dialog® and DataStar™, which have different user interfaces and a different mix of information sources. For decades Dialog was one of only a handful of online information services available to librarians and researchers. For many librarians, "Dialog" is synonymous with classic Dialog, Dialog's flagship service. In fact, there are several different interfaces available from Dialog, including DialogLink software with enhanced search features and output options.

The widespread commercial adoption of the World Wide Web in the mid-1990s changed the online information landscape radically. As a consequence, most of Dialog's current user interfaces are Web-based. For many years researchers have had an ever-expanding and constantly changing choice of online resources available. Nevertheless, after more than 30 years of partnership with librarians, Dialog remains the premier online information service for professional researchers.

The purpose of *Introduction to Dialog for Information Professionals* is to introduce concepts and develop skills essential for a professional-level knowledge of online searching. To accomplish these goals, DialogClassic is used for illustration and practice.

What Is DialogClassic?

DialogClassic is the oldest and most familiar online information service offered by Dialog. For those accustomed to using search engines and portals on the Web, a text-only, command-driven interface may seem strange, even archaic. Dialog was designed for information professionals, not for casual or end users. Few assumptions are made in the Dialog system about how a search is to be done and what information sources are to be searched. The success of a search depends largely on the knowledge and skill of the searcher.

To use Dialog successfully, a searcher is expected to:

- be familiar with the scope of the online data or content,

- choose the appropriate information sources based on an understanding of their unique attributes (indexing, data structure), and to

- craft a search using the features of the Dialog search engine with as much art and skill as possible.

DialogClassic is particularly well suited for learning the fundamentals of online searching for several reasons:

- **Uncluttered interface**
  The absence of advertisements and extraneous links enables students to focus on the search process without distractions. The text-only interfaces also result in greater speed in conducting searches online.

- **Emphasis on search planning and preparation**
  The organization of DialogClassic requires that students prepare a search in advance by selecting the sources to search and the strategy they will use. The process forces students to consider each step in the search process and consider alternate approaches.

- **Transparency**
  Students can see a search in progress, evaluate, and correct the search strategy in an iterative process. Database-specific features are also easily viewed online. Students learn how data can be structured online and view and compare database features, such as subject indexing, thesauri, and editorial and data normalization practices.

- **Variety of sources**
  The variety of online sources available on DialogClassic, both in subject matter and format, is unique. Students are able to explore the characteristics of data types (bibliographic, fulltext, directory/reference, and numeric) and to develop greater familiarity with information sources and communication patterns within specific disciplines, industries and subject areas.

Data on DialogClassic is organized into separate files or databases—nearly 600! Few databases are actually created by Dialog; rather, each database is licensed to Dialog by an information provider. Most information providers are publishers. Their newspapers, journals, company directories, et cetera, are usually produced and sold in print and in electronic formats to their own subscribers. Licensing the information to an information aggregator such as Dialog is a way for publishers to increase their readership and revenues. Most of the information on Dialog is copyrighted by the publishers. Dialog pays royalties for the use of the information. The usage fees charged by Dialog cover the operating costs of the service as well as the payment of royalties to the information providers.

Why Use Dialog?

Most students begin a course in online searching having experience using Web search engines. As a result, they may have difficulty understanding why anyone would pay to use a commercial online information service when so much is available on the Web for free. Most Web-based search engines trade search precision and control of the search process in exchange for ease of use. There is nothing wrong with this trade-off for many kinds of searches, particularly where limited or consumer-oriented information is needed. The benefits and convenience that Dialog offers the professional researcher are many and explain why thousands of organizations are willing to pay to use Dialog. Major differences between Dialog and typical Web searching are outlined below.

- **Professional quality information**
  The data available from Dialog is professionally produced. That means the information has been edited so that it adheres to a minimum standard of accuracy. Many sources available from Dialog are created by and for professionals: scientists, scholars, business managers, market researchers, stock brokers, attorneys, as well as many others—including librarians. The recognized authority of publishers such as Dun & Bradstreet, Standard & Poor's, *Financial Times*, Economist Intelligence Unit, Chemical Abstracts Service, and thousands of others means Dialog is a trusted source for professional researchers. Make sure any "expert" offering information online truly is a credible authority.

- **Breadth & depth of content**
  Dialog is unique in the vast array of information covering virtually every subject. Thousands of serial publications are indexed in Dialog, and many of these are in fulltext. In addition, Dialog includes the complete text of many reference works and specialty publications, such as market research and brokerage house reports, patents, trademarks, and copyright registrations. Dialog is also unique in depth of information. Archival data is available for many databases back to the 1960s and 1950s—some even offer data back to the eighteenth century! Be wary of exaggerated claims about content. It is important to understand what an online service provider means when it claims to offer "5,000 fulltext sources." Is a journal counted as one source, or is each issue or individual article counted as a separate source? Is a fulltext source a three-page government pamphlet, or a multi-volume scientific reference?

- **Selectable content**
  A key attribute of Dialog is the ability to choose which data sources will be included in a search. In other words, the searcher can use her knowledge of specific information sources to control the scope of a search. Features to control what is being searched on Dialog permit the searcher to design a search that can scan the entire data collection, some 12 terabytes, or pinpoint a single data element in a specific record. Knowledge of what sources have already been searched will make the research process more efficient by eliminating redundant searches.

- **One-stop convenience**
  Dialog is often referred to as a data "supermarket." This is a useful analogy for explaining the convenience of going to a single supplier for information that may be available in many places. It is also a fact that no single source can be absolutely comprehensive, whether that source is a reference work, a search engine, or a single database. MEDLINE®, for example, is available from the National Library of Medicine's Web site for free. It does not cover all the medical literature, however. It is only one of several important sources for medical researchers: EMBASE®, BIOSIS® Previews, SciSearch® and many others, include biomedical journals not covered in MEDLINE. A search of every source, each with a different interface and record format, would be very inefficient. Then trying to integrate the results, including eliminating any duplicate references, would be a time-consuming and tedious task. Dialog simplifies the process by collecting all the major sources in a discipline so that these can be searched together. Duplicate detection and elimination are Dialog features.

- **Highly structured and consistent data tagging**
  As you will learn, the indexing and structure of data—or lack of it—can help or hinder your search. On Dialog, similar data is structured and tagged in the same way.

- **Powerful search engine**
  The Dialog search engine includes a sophisticated array of commands developed to meet the needs of searchers. In addition to tools for precision retrieval, Dialog offers options for data analysis and extraction and customized output. Search results may be sorted in a variety of ways, including by statistical relevance as well as chronological or alphabetical on selected data.

- **Many value-added features & services**
  Meeting the needs of professional researchers does not begin and end with an online search. Managing search results to meet the needs of individuals, departments and organizations is a key role for today's information professionals. Ancillary services are needed to assist in managing and distributing electronic information within the organization, such as: setting up customized data feeds or current awareness services for specific users; distributing data via e-mail or intranet; managing digital copyrights; providing document delivery services; even direct support through consulting. All these services are available to Dialog customers.

- **Extensive documentation**
  Unlike most Web-based information services, Dialog provides extensive documentation on the content and features of its services. The *Dialog Database Catalog*, *Dialog Pocket Guide*, and Bluesheets are some titles you will become familiar with as you learn DialogClassic. They represent just a small sample of the extensive documentation available in print, online, and via the Web.

- **Training and support**
  Extensive documentation is just one kind of support available to Dialog customers and not often found in Web-based services. Other kinds of customer support and services also distinguish Dialog from most Web search engines and portals: help desk support via telephone; an account representative who will maintain contact with the Dialog customer and help deal with questions and problems; and training opportunities.

- **Commitment to the profession**
  Unlike many Web-only services, Dialog will still be providing quality information and services to its customers for decades to come. Dialog's 30+ years of success is testimony to the value that professional researchers place on the breadth and depth of its content, comprehensive indexing, and its feature-rich search engine. Dialog has achieved this success by partnering with information professionals in a variety of ways. Ongoing support for professional development has been a feature of Dialog through such programs as Quantum[2], the Graduate Education Program, and sponsorship of library association events. Of particular interest is the Roger K. Summit Scholarship offered by Dialog to outstanding students in library and information science.

Dialog History

Dialog has a unique relationship with the information professional community. It is one of the oldest online information services. In 1964, with the space age in full swing, Roger Summit was assigned to manage an obscure program for the Lockheed Missile and Space Company. One year later, the Information Sciences Laboratory demonstrated its first interactive information retrieval service.
In 1972, Summit convinced Lockheed management of the commercial application of this powerful new technology, and a company was born, with Roger Summit as its leader. Dialog began offering the first publicly available online research service, which led to the advent of the online information industry.

Since its inception, Dialog has evolved through various corporate owners. In 1981, Dialog became a subsidiary of Lockheed, moved its offices to Mountain View, California—in the heart of Silicon Valley—and was renamed Dialog Information Services. Seven years later, in 1989, the company was sold to Knight-Ridder, and became Knight-Ridder Information, Inc. In 1997, it was purchased by the U.K.-based M.A.I.D., and the combined business was renamed The Dialog Corporation. Company headquarters were moved in 1997 to the current location in Cary, North Carolina.

In May 2000, The Thomson Corporation acquired Dialog in a move that partnered the world's foremost global e-information solutions company with the company that started online information retrieval more than 30 years ago.

Graduate Education Program

The Graduate Education Program (GEP) represents Dialog's commitment to the continued development of information professionals throughout the world. The program has been offered to educational institutions since 1976. GEP provides a cost effective way for professional library schools to train students in the use of the Dialog search system. It offers many resources: access to the Dialog service for teaching purposes, student and instructor materials, Dialog search aids, and documentation.
In addition, The Roger K. Summit Scholarship, awarded annually by Dialog, was established to honor Dr. Roger K. Summit, the founder of Dialog, for his outstanding contributions to the field of information science. The $5,000 award is made to a graduate student in library and information science who has demonstrated outstanding interest or performance in electronic information services.

About This Workbook

Audience and Setting

This *Introduction* is intended primarily for use by graduate students in library and information science or other research-intensive disciplines.

Although the workbook is intended to be self-sufficient in describing the operation of the Dialog search system, available documentation to be consulted as needed is:

- *Dialog Pocket Guide,* a summary of Dialog system commands and capabilities.

- *Dialog Database Catalog.*

- *Successful Searching with Dialog Command Language.*

- Dialog Bluesheets for information specific to individual databases.

These materials are available on the Dialog Web site.

Conventions Used

This workbook uses a few typographical conventions to make it easy to follow the search examples provided.

- **Bold underlined text** is used for commands that the searcher types on the keyboard. Lowercase letters are used in the illustrations, but commands may be entered using upper or lowercase.

- Ellipses (**...**) are used where the system output has been shortened to save space on the page.

- Capital letters within the text are used to indicate a command name or a specific key on the keyboard. For example, 'the SELECT command,' or 'type the password and press ENTER.'

Since many kinds of computers and terminals can be used to access Dialog, the key labels on your equipment may be different.

Format and Screen Displays

Dialog continues to develop new features and improve existing system features and functions. As a result, some screen displays and features may have changed by the time you use this *Introduction.*

About the Exercises

It is recommended that a minimum of four hours of online time be allowed for each student in order to complete the required exercises. The instructor may choose to use only selected parts of the exercises in order to reduce the amount of time required.

Instructions to Students

Prepare for the online sessions by using available documentation and preparing worksheets for the search exercises. Perform the exercises in three stages:

1. Activities to complete before connecting.

2. Searches and commands to practice while online.

3. Tasks to complete after the search.

Follow these tips when working on the exercises:

• Use the Graduate Education Program (GEP) user numbers and passwords assigned to you by your instructor.

• Use the Dialog Bluesheets and other database documentation to prepare for the online exercises.

• Have your worksheets for the online exercises and publications, such as Bluesheets or the Pocket Guide, available. Do not go online if you are unprepared.

• As part of each exercise, annotate your search results and return them with the completed exercise to your instructor.

• Keep a log sheet of your online searches on which you record the databases, time, and costs for each exercise.

• For the first exercise, spell out the commands instead of using abbreviations or symbols, in order to make the results easier to read. As you become more proficient, use abbreviations to save time.

Where to Get Help

Dialog provides many print and online tools to help you learn and use DialogClassic. To help you get started learning DialogClassic, here are a few key resources.

• **The Training Center**
  Training materials, including online courses and tip sheets, are collected in the Training Center on the Dialog Web site.

• **Principal Dialog References**
  o *Dialog Database Catalog* – a quick guide to the databases available from Dialog.
  o *Dialog Pocket Guide* – a handy reference to DialogClassic commands and features.
  o *Successful Searching with Dialog Command Language* – the complete reference to DialogClassic with detailed explanations of DialogClassic commands and features.

• **KnowledgeBase**
  KnowledgeBase is an online finding tool you can use to search through the Dialog Web site for information on topics of interest to you. KnowledgeBase is located under "Support" on the Web site toolbar.

BASIC COMMANDS FOR ONLINE SEARCHING USING DIALOGCLASSIC

Objectives

In this section, you learn how to:

- Conduct a simple search using four basic Dialog commands
- Search with word roots and words in a phrase
- Combine terms using logical operators
- Display the search results in a variety of formats
- Display search terms in context

Features

- BEGIN command to start a search in a specific database
- SELECT command to search with words or terms
- Logical operators OR and AND
- Proximity operators (W) and (N)
- TYPE command and pre-defined display formats
- KWIC format
- LOGOFF command to disconnect from Dialog and see the session charges

Background

DialogClassic has features that make it easy for the searcher to:

- Access and disconnect from the service

- Specify the database(s) to be searched

- Enter commands to determine the number of records in a database that contain the terms of interest

- Search word roots or terms of variable length

- Connect relevant search terms using the logical operators AND, OR, NOT, and proximity operators

- Display records online or have them delivered by mail, ftp, or email

To connect to DialogClassic and retrieve information, you must have a computer with a connection to the Internet or use a modem with communication software. There are various interfaces that can be used to access Dialog. However, the version of Dialog in this *Introduction*, known as DialogClassic, is a plain ASCII interface. The examples will be in DialogClassic for the sake of clarity.

Dialog Sample Search

The following example is a brief search through DialogClassic using four basic commands on the Dialog search system. Searcher input is shown in **<u>bold underlined</u>** typeface.

```
Dialog level 05.12.03D
Last logoff:  04aug06 12:10:28
Logon file405 07aug06 11:21:39

            *** ANNOUNCEMENTS ***
                     ***
NEW FILES RELEASED
***EMCare (File 45)
***Trademarkscan – South Korea (File 655)
***Regulatory Affairs Journals (File 183)
***Index Chemicus (File 302)
***Inspec (File 202)

RESUMED UPDATING
***File 141, Reader's Guide Abstracts
                     ***
RELOADS COMPLETED
***File 11, PsycInfo
***File 516, D&B--Dun's Market Identifiers
***File 523, D&B European Dun's Market Identifiers
***File 531, American Business Directory
*** The 2005 reload of the CLAIMS files (Files 340, 341, 942) is now available online.
                     ***
Chemical Structure Searching now available in Prous Science Drug
Data Report (F452), Prous Science Drugs of the Future (F453),
IMS R&D Focus (F445/955), Pharmaprojects (F128/928), Beilstein
Facts (F390), Derwent Chemistry Resource (F355) and Index Chemicus (File 302).

                   *** DIALOG HOMEBASE(SM) Main Menu ***

 Information:
  1.  Announcements (new files, reloads, etc.)
  2.  Database, Rates, & Command Descriptions
  3.  Help in Choosing Databases for Your Topic
  4.  Customer Services (telephone assistance, training, seminars, etc.)
  5.  Product Descriptions

 Connections:
  6.  DIALOG(R) Document Delivery
  7.  Data Star(R)

    (c) 2003 Dialog, a Thomson business.     All rights reserved.

     /H = Help          /L = Logoff          /NOMENU = Command Mode
```

```
Enter an option number to view information or to connect to an online
 service.  Enter a BEGIN command plus a file number to search a database
(e.g., B1 for ERIC).

?BEGIN 1
      07aug06 11:35:47 User300065 Session D793.1
             $0.00    0.290 DialUnits FileHomeBase
    $0.00  Estimated cost FileHomeBase
    $0.11  INTERNET
    $0.11  Estimated cost this search
    $0.11  Estimated total session cost 0.290 DialUnits

[File 1] ERIC 1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

      Set  Items  Description
      ---  -----  -----------
?SELECT LIBRARY(W)SCHOOL? AND DISTANCE(W)EDUCATION
        45993   LIBRARY
       432425   SCHOOL?
         1567   LIBRARY(W)SCHOOL?
        13982   DISTANCE
       823715   EDUCATION   (PROCESS OF IMPARTING OR OBTAINING KNOWLEDGE,...)
         9906   DISTANCE(W)EDUCATION
S1        40   SELECT LIBRARY(W)SCHOOL? AND DISTANCE(W)EDUCATION

?TYPE S1/8/1-3

1/8/1
ERIC
(c) format only 2006 Dialog. All rights reserved.
01100165 ERIC Number: EJ629765   Clearinghouse Number: IR543722
Investigating the Evaluation Procedures for a Distance Learning Undergraduate Degree in LIS.
2000 (20000000)
Descriptors: *Distance Education; *Evaluation Methods; Foreign Countries; Higher Education;
*Information Science Education; *Library Education; Library Schools; Library Science;
Undergraduate Study
Identifiers: *University of Wales Aberystwyth


1/8/2
ERIC
(c) format only 2006 Dialog. All rights reserved.
01086013 ERIC Number: EJ623512   Clearinghouse Number: IR543064
Reflective Practice: Revisiting Pedagogy and Reaffirming Core Value.
2000 (20000000)
Descriptors: Accreditation (Institutions); Administrator Education; Competence; Computer
Literacy; Conferences; Core Curriculum; Distance Education; Economic Factors; Educational
Trends; Foreign Countries; Futures (of Society); *Information Science Education; *Library
Schools; Practicums; Professional Continuing Education; Professional Development
Identifiers: *Australia; *Canada


1/8/3
ERIC
(c) format only 2006 Dialog. All rights reserved.
01078266 ERIC Number: ED450802   Clearinghouse Number: IR058070
Distance Education in Library and Information Science Education: Trends and Issues.
2000 (20000000)
Descriptors: Computer Assisted Instruction; *Distance Education; Higher Education; *Information
Science; Journal Articles; *Library Education; Library Research; *Library Schools; *Library
Science; Program Evaluation; *Trend Analysis

?T S1/5/4

DIALOG(R)File 1:ERIC
(c) format only 2006 Dialog. All rights reserved.

01078266 ERIC NO.: ED450802 CLEARINGHOUSE NO.: IR058070
Distance Education in Library and Information Science Education: Trends and Issues.
  Zepp, Diana
  59pp.
  2000 (20000000)
NOTES: Master of Library and Information Science, Research Paper, Kent State University.
EDRS Price MF01/PC03 Plus Postage.
LANGUAGE: English
DOCUMENT TYPE: 40 (Dissertations/Theses)
RECORD TYPE: ABSTRACT
COUNTRY OF PUBLICATION: U.S.; Ohio
JOURNAL ANNOUNCEMENT: RIEAUG2001
```

BEGIN to access
a database

SELECT to enter
search terms

TYPE records in a
short format

TYPE a complete
record

```
   This study measured current trends in distance education in the United States within Library
and Information Science programs. The study was conducted, for the period 1989 to 1998, through
a content analysis of journal articles from the "Library Literature" database, and through a
content analysis of graduate catalogs from American Library Association (ALA) accredited
library schools. Of 128 journal articles analyzed, 86% were non-research articles, with the
main topic of discussion being distance education at a specific library program. The remaining
14% of the articles were dedicated to research, with 44% of that research pertaining to the
study of distance education students. A significant finding was that slightly more than half
(56%) of the 128 articles were published in the
                            .
                            .    lines deleted
                            .
Appendixes include a list of ALA accredited library programs and coding sheets. (Contains 28
references.) (AEF)

DESCRIPTORS: Computer Assisted Instruction; *Distance Education; Higher Education; *Information
Science; Journal Articles; *Library Education; Library Research; *Library Schools; *Library
Science; Program Evaluation; *Trend Analysis

?logoff

COST
        07aug06 11:37:47 User311518 Session D293.2
               $1.80    0.515 DialUnits File1
                  $0.80  1 Type(s) in Format  5
                  $0.00  4 Type(s) in Format  8
               $0.80  5 Types
       $2.60  Estimated cost File1
       $0.53  INTERNET
       $3.13  Estimated cost this search
       $3.15  Estimated total session cost   0.703 DialUnits
```

> LOGOFF to disconnect. Final cost estimate given.

System Logon Message

At the beginning of each DialogClassic search session, Dialog displays a ***logon message*** consisting of brief messages and system news, as shown in the sample search. The logon message announces new or reloaded databases and the availability of new features added to the system. If you have already been online earlier in the day, the full logon message will not display; you will only see any warning messages about temporary problems.

After the logon message displays, you are automatically placed in the Dialog Homebase[SM] (File HOME) database. To get more detailed information about recent announcements, system features, or training programs, you can enter one of the option numbers listed on the Homebase menu.

When you establish a Dialog account, you are asked to specify which file you want to have as your ***default file***, the file you are connected to when you logon. Dialog Homebase is the default file for most customers and for the Graduate Education Program. A customer can change the default file by contacting the Customer Administration Department at Dialog.

At the end of the logon message, Dialog displays the ***system prompt***, a question mark (?), which indicates that the system is ready to receive a command. The first command is usually a BEGIN command that specifies which database you want to search.

BEGIN Command: Accessing a Database

Most Dialog databases are identified by file numbers; menu-driven databases, such as Dialog Homebase, are identified by acronyms, e.g., HOME. To begin a search in a specific database, enter the command BEGIN (which can be abbreviated to B) followed by the file number of the database to be searched. Databases and their file numbers are listed in the *Dialog Database Catalog* and on the web. To access the ABI/INFORM® database in File 15, enter BEGIN 15, as shown below.

```
?begin 15

     03aug06 15:44:25 User311518 Session D289.1
          $0.00     0.321 DialUnits FileHomeBase
    $0.00   Estimated cost FileHomeBase
    $0.05   INTERNET
    $0.05   Estimated cost this search
    $0.05   Estimated total session cost   0.321 DialUnits

File 15:ABI/Inform(R) 1971-2006/Aug 03
      (c) 2006 ProQuest Info&Learning

     Set  Items  Description
     ---  -----  -----------
?
```

Dialog databases may also be searched as a group, using a file acronym or combination of file numbers. For example, BEGIN TELECOM or BEGIN 26,144,115,165. For now, you will be searching one database at a time. You can search more than one topic in the same database without re-entering BEGIN, but since BEGIN clears your workspace, you may find it helpful to BEGIN before each new search topic. Remember that you lose all previous work when you type the BEGIN command.

SELECT Command: Entering Search Terms

The SELECT command, abbreviated S, retrieves records containing the specified search terms and stores them in a numbered *set*.

The system's response to every SELECT command appears below the column headings that display after a BEGIN command, as shown below.

```
File 1:ERIC  1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

     Set  Items  Description
     ---  -----  -----------
?select malnutrition

    S1     379  MALNUTRITION
```

S1 is the *Set* number that Dialog assigns to the first set of records retrieved. Under the column heading *Items*, the number 379 indicates the number of records (sometimes called "hits" or "results") in the database that include the word MALNUTRITION in the Basic Index of searchable terms (the Basic Index is discussed in section 3). The exact terms you entered in the SELECT command are always repeated under the column headed *Description*.

Online search systems are very literal and the SELECT command requires the system to find a character-by-character match on the search terms. Unless truncation (covered later in this section) is used, SELECT does not retrieve terms that differ at all in spelling, punctuation, or special characters. Empty sets (zero items) are created when an exact match is not found for a SELECTed term. In the example below, an empty set is created due to a misspelled term.

```
?s malntrition

    S2      0  MALNTRITION
```

Any single word, number, or alphanumeric expression can be entered as a search term in a SELECT command, with the exception of nine non-searchable words called *stop words*. The stop words are:

AN	FOR	THE
AND	FROM	TO
BY	OF	WITH

Dialog also has words that are reserved for system commands, which can be searched if enclosed in quotation marks. The *reserved words* are:

FILES	SELECT	R1, R2, etc.
NOT	STEPS	S1, S2, etc.
OR	E1, E2, etc.	

So, for example, to search for the R2 Call Management System, you would type **select "r2"**
A word containing punctuation—such as a hyphen or apostrophe—cannot be SELECTed with the punctuation intact; punctuation must be replaced with the (W) operator. For example, instead of searching for the word E-MAIL, you would search for E(W)MAIL. (In fact, you should probably also look for the word EMAIL, as hyphenated words often appear in print as single words without the hyphen.)

Phrases, or multiple-word search terms, such as PUBLIC LIBRARIES, must be SELECTed using the (W) operator, unless the term is a *descriptor*, *identifier*, or *subject heading*. Descriptors are subject-related terms taken from a thesaurus or controlled vocabulary and assigned to a record to reflect the article's major subject areas. Identifiers are additional subject terms that are *not* drawn from a controlled vocabulary list. Descriptors, identifiers, and subject headings are usually assigned by indexers working for the information provider. The terms may be single words or multiple-word phrases. If you are not sure whether or not a phrase is a descriptor, identifier, or subject heading, use the EXPAND command, described later, to see how the subject term is used in the database you are searching.

If you SELECT a complete descriptor or identifier phrase, including spaces and any punctuation, the system only retrieves items with that phrase in the Descriptor or Identifier fields. This is the only instance in which you can successfully search on a phrase in the Basic Index without using a proximity operator such as (W). The Basic Index is the default group of fields that are searched if no particular field is specified. See the Bluesheets for a list of fields in the Basic Index in different databases.

Bluesheets are written guides for every database on the Dialog® service. They contain detailed instructions on search techniques for the special features of each database, including file description, subject coverage, date range, update frequency, sources of the data, and the origin of the information. On the Bluesheet you will also find a sample record that shows what you can expect to obtain when you perform a search in the database. Bluesheets are explained in more detail in a later section.

```
?s public libraries

     S3   6129  PUBLIC LIBRARIES  (LIBRARIES FREELY AVAILABLE TO ALL, THAT SERV...)

?s bilingual education act 1968

     S4    373  BILINGUAL EDUCATION ACT 1968
```

Note that the search term listed under the *Description* column shows annotations the database producer includes for clarification of the meaning of the term.

You can enter more than one search term in a SELECT command when using the logical operators AND, OR, and NOT. For example:

```
?s bilingualism and canada

           5364  BILINGUALISM
          26423  CANADA
     S5     396  BILINGUALISM AND CANADA
```

Logical operators are discussed later in this section.

You can use set numbers in a SELECT command to include the records from a previous search set in a new set. The set number must be preceded by an S, for example, S3.

```
?s s3 and internet

          4903  S3
          7682  INTERNET
    S4     236  S3 AND INTERNET
```

SELECT STEPS Command

A variation of the SELECT command is the command SELECT STEPS, abbreviated SS. SELECT STEPS causes a set number to be assigned to each term in the command.

```
?select steps bilingualism and canada

    S7     5364  BILINGUALISM
    S8    26423  CANADA
    S9      396  BILINGUALISM AND CANADA
```

Always enter a space following a SELECT or SELECT STEPS command or their abbreviations. Otherwise, Dialog looks for some form of the command, and retrieves whatever character string follows, sometimes with amusing results:

```
?selectric
    S10       18  RIC

?ssympathy
    S11        0  YMPATHY
```

Note: SELECT (S) rather than SELECT SETS (SS) is the preferred command for most of the examples in this workbook. It generates fewer sets to keep track of and can save processing time in lengthy searches and in large files. SELECT also encourages novice searchers to think in terms of "concept" groups and makes it easier to visualize groups of synonyms. The best search strategy creates one set for each concept.

SELECT STEPS (SS) enhances search flexibility by assigning set numbers to each term, allowing you to re-combine sets in alternate configurations, alter strategies, and correct typographical errors more readily.

For example:

```
?ss online and information and retieval

    S12    16240  ONLINE
    S13   201866  INFORMATION
    S14        1  RETIEVAL
    S15        0  ONLINE AND INFORMATION AND RETIEVAL
```

RETRIEVAL was misspelled as RETIEVAL in the above search. However, the sets for ONLINE and INFORMATION are valid, since the terms were spelled correctly. To be efficient, the next command could be:

```
?ss s12 and s13 and retrieval

         16240  S12
        201866  S13
    S16   10278  RETRIEVAL
    S17    2921  S12 AND S13 AND RETRIEVAL
```

When you want to mix and match your search terms individually, having separate sets for each term is helpful. For other searches, especially when you are searching for specific phrases, you will have no need for separate sets for each term and the SELECT command will be preferable. For example, if you were searching for the Distance Learning Project, you would find set S7 or set S11 of the following search useful whereas you would probably not want to have your search cluttered with the intermediate sets S8, S9, and S10.

```
?select distance(w)learning(w)project

        13982  DISTANCE
       247237  LEARNING  (PROCESS OF ACQUIRING KNOWLEDGE, ATTITUDES, O...)
        90432  PROJECT
    S7     42  DISTANCE(W)LEARNING(W)PROJECT

?select sets distance(w)learning(w)project

    S8  13982  DISTANCE
    S9 247237  LEARNING  (PROCESS OF ACQUIRING KNOWLEDGE, ATTITUDES, O...)
   S10  90432  PROJECT
   S11     42  DISTANCE(W)LEARNING(W)PROJECT
```

As you plan your search, weigh the merits of both the SELECT and SELECT STEPS commands and choose the version most suitable for the length and type of search to be conducted.

► **Dialog System Capacities**

TIP

2000 characters	per SELECT command
60 files	per BEGIN command
999 search sets	per BEGIN command
17,000,000 retrieved records	per BEGIN command
20,000,000 postings	per SELECT command

Truncating to Search with Word Roots

To retrieve variations in word endings, place a question mark at the end of a word "root." This is often called "open" truncation because the system searches for all words containing the root word, regardless of how many characters they contain. In the example below, all words beginning with the letters COMPUT are retrieved – COMPUTER, COMPUTING, COMPUTERIZED, COMPUTATIONAL, etc.

```
?s comput?

    S1  79495  COMPUT?
```

Truncation can be a mixed blessing and it is wise to avoid short word roots that could retrieve many different words. Consider the word BEE and its plural BEES. If you truncate BEE? to get both, you also retrieve BEEF, BEEFALO, BEER, BEEN, BEETHOVEN, and many other words. A good general rule is to truncate word roots having at least four characters or to use ***controlled length truncation***.

Open Truncation

Open truncation is the broadest form of truncation. To retrieve all words beginning with the same word root, put a question mark at the end of the root:

select path?	retrieves	path	all words that begin with **path**
		paths	
S1 3145 PATH?		pathos	
		pathway	
		pathways	
		pathology	
		pathologists …	

Controlled Length Truncation

Controlled length truncation specifies a maximum number of characters that may follow the word root:

select path? ?	retrieves	path	retrieves the root and up to one additional
S2 32 path? ?		paths	character
select path??	retrieves	path	retrieves up to two additional characters
S3 35 PATH??		paths	
		pathos	
select path???	retrieves	path	retrieves up to three additional characters
S4 85 PATH???		paths	
		pathos	
		pathway	

You can truncate a descriptor or identifier phrase by placing a question mark anywhere in the phrase. This can be useful for retrieving similar descriptors. For example, ERIC uses the descriptor *Minority Groups*, as well as *Minority Group Children*, *Minority Group Influences*, and *Minority Group Teachers*. All of these descriptors are retrieved by the SELECT command: SELECT MINORITY GROUP?

Embedded Character Truncation

Embedded character truncation is particularly helpful when searching on words that may have variant British and American spellings:

select organi?ation	retrieves	organization
		organisation
select fib??board	retrieves	fiberboard
		fibreboard

This truncation feature is also useful for searching for unusual plural forms:

select wom?n	retrieves	woman
		women

Embedded character truncation cannot be used to retrieve variant spellings where the number of characters varies. That is, each truncation character used internally represents a specific character; for example, SELECT LAB??R retrieves LABOUR but not LABOR.

Using (W) and (N) to SELECT Words near Each Other

Dialog provides many different options for searching words that are adjacent or near each other. The following is a brief introduction to two of the most commonly used proximity operators.

To search an exact phrase, use the (W) operator between words in a phrase. For example, the sample search at the beginning of this section showed the command SELECT LIBRARY(W)SCIENCE?. This retrieves records where the two words appear next to each other and in this exact order. The search will retrieve records that have the phrase LIBRARY SCIENCE or LIBRARY SCIENCES. It will not retrieve records with the phrase SCIENCE LIBRARY nor SCIENCES LIBRARY.

To search a phrase and retrieve results without regard to the order of the words, use the (N) operator between words. For example, the command SELECT HEART(N)DISEASE? retrieves records with the phrases: HEART DISEASE, HEART DISEASES, and DISEASED HEART.

Logical Operators OR and AND

An essential aspect of online searching is how you define the logical relationships among the terms being searched. Three kinds of logical relationships are defined on Dialog: OR, AND, and NOT. **Logical operators** are also known as **Boolean operators**.

OR Logic

The OR operator groups search terms inclusively into a single set, when any of the terms is acceptable. Thus, the retrieval of LIBRAR? OR INFORMATION(W)CENTER? consists of all records containing LIBRARY, LIBRARIAN, LIBRARIES, INFORMATION CENTER, INFORMATION CENTRALIZED, and so on. Records need not have both search terms; either one or the other or both will meet the search requirement. The following Venn diagram illustrates that all of the hits on the terms in the shaded areas are retrieved, including the records with both terms.

LIBRAR? ⟶   ⟵ INFORMATION (W)CENTER?

LIBRAR? OR INFORMATION(W)CENTER?
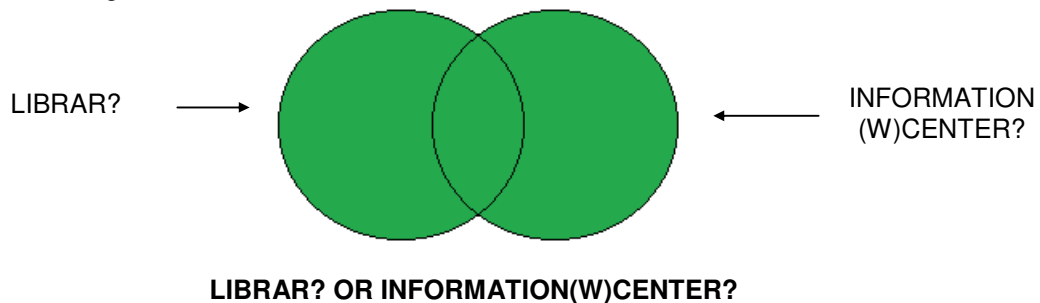
```
?s librar? or information(w)center?

        52714   LIBRAR?
       201866   INFORMATION
        79109   CENTER?
         3960   INFORMATION(W)CENTER?
S1      55302   S LIBRAR? OR INFORMATION(W)CENTER?
```

OR increases the number of records retrieved. The OR operator is often used to search for all of the records that contain synonymous or alternative terms. If more than one of the terms occurs within a given record, the retrieved set includes the record only once.

AND Logic

The AND operator retrieves records where two or more search terms or groups of search terms must both occur in the same record. SELECT LIBRARIES AND AUTOMATION retrieves only those records containing both of these words, as represented by the shaded area in the Venn diagram below.

LIBRARIES ⟶   ⟵ AUTOMATION

LIBRARIES AND AUTOMATION
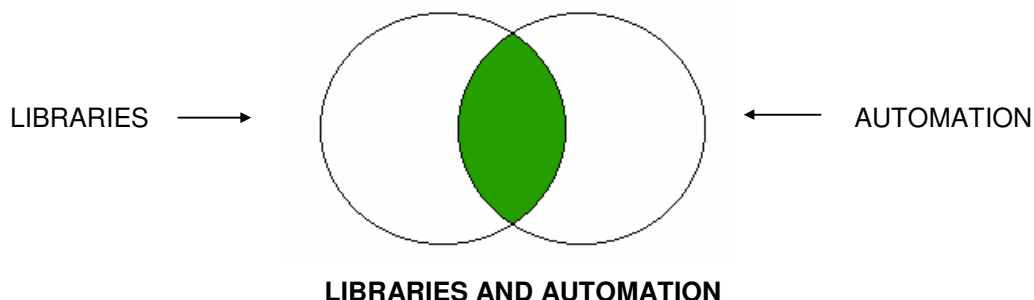
```
?s libraries and automation

        33332  LIBRARIES
         6193  AUTOMATION
   S2    3283  LIBRARIES AND AUTOMATION
```

AND decreases the number of records retrieved, since it requires that all search terms be present for retrieval. Each time you use AND in a search, you are restricting the number of records retrieved.

NOT Logic

The NOT operator prevents records that include unacceptable or irrelevant search terms from being retrieved. The expression LIBRARIES NOT SCHOOLS retrieves records that contain the term LIBRARIES but excludes those records that also contain the term SCHOOLS, as shown in the following Venn diagram:

LIBRARIES NOT SCHOOLS
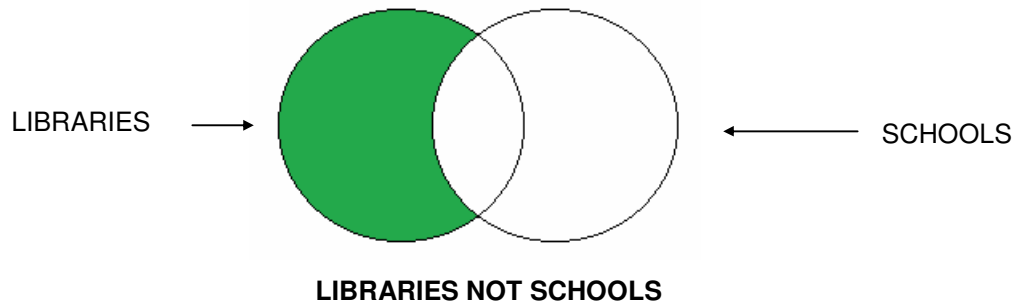
```
?s libraries not schools

        33332  LIBRARIES
       210346  SCHOOLS
   S3   29442  LIBRARIES NOT SCHOOLS
```

The NOT operator is useful for excluding search terms that are known in advance to be irrelevant to the search topic. Use NOT carefully, since you may unintentionally eliminate useful records. For example, the strategy ENERGY NOT NUCLEAR would eliminate a record entitled "Alternatives to Nuclear Energy." Controlled vocabulary phrases can often be used to eliminate unwanted concepts more precisely. For example, to retrieve records about libraries, but avoid retrieving records about library schools, you could use the ERIC descriptor *Library Schools* following NOT:

```
?s librar? not library schools

        52714  LIBRAR?
         1215  LIBRARY SCHOOLS   (PROFESSIONAL SCHOOLS, DEPARTMENTS, OR DIVISI...)
   S4   51499  LIBRAR? NOT LIBRARY SCHOOLS
```

You can effectively use NOT logic to exclude specific non-subject parameters, such as the language of the original document, in order to refine your search. Non-subject parameters are searched with prefixes (for example, LA= for Language. The example below searches on the descriptor *Bilingual Education* but eliminates documents that are written in Spanish:

```
?s bilingual education not la=spanish

         7959  BILINGUAL EDUCATION   (ENCOURAGEMENT OF BILINGUALISM THROUGH THE TE...)
         1888  LA=SPANISH
   S5    7650  BILINGUAL EDUCATION NOT LA=SPANISH
```

Using Multiple Logical Operators: 'Nesting'

You can use any combination of logical operators within the same SELECT command, but you may want to use parentheses to specify the order of processing; this technique is called **nested logic**. Unless you specify otherwise, Dialog first processes any NOT logic contained in a SELECT statement, then it

processes any AND logic, and finally, any OR logic. To vary this order of processing, put parentheses around the group of terms and operators that you want to be processed first.

Say you wanted to search in the ERIC file on the topic CURRICULUM DEVELOPMENT, focusing specifically on APHASIA or DYSLEXIA. In the first example below, the parentheses cause Dialog to correctly combine APHASIA OR DYSLEXIA before applying the AND operator to include CURRICULUM DEVELOPMENT in the search. In the second example, the lack of parentheses causes Dialog to retrieve *all* occurrences of APHASIA, as well as records containing both DYSLEXIA and CURRICULUM DEVELOPMENT, which was not what the search topic called for.

```
File 1:ERIC  1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

     Set  Items  Description                                       Correct Logic
     ---  -----  -----------
?s (aphasia or dyslexia) and curriculum development

          572  APHASIA  (IMPAIRMENT OF THE ABILITY TO COMPREHEND OR P...)
         1460  DYSLEXIA  (IMPAIRMENT IN THE ABILITY TO READ DESPITE AD...)
        46073  CURRICULUM DEVELOPMENT  (ACTIVITIES SUCH AS CONCEPTUALIZING, PLANNING...)
     S1    14  (APHASIA OR DYSLEXIA) AND CURRICULUM DEVELOPMENT

?s aphasia or dyslexia and curriculum development                 Wrong Logic

          453  APHASIA  (IMPAIRMENT OF THE ABILITY TO COMPREHEND OR P...)
         1180  DYSLEXIA  (IMPAIRMENT IN THE ABILITY TO READ DESPITE AD...)
        43170  CURRICULUM DEVELOPMENT  (ACTIVITIES SUCH AS CONCEPTUALIZING, PLANNING...)
     S2   467  APHASIA OR DYSLEXIA AND CURRICULUM DEVELOPMENT
```

The correct search can be diagrammed as follows:
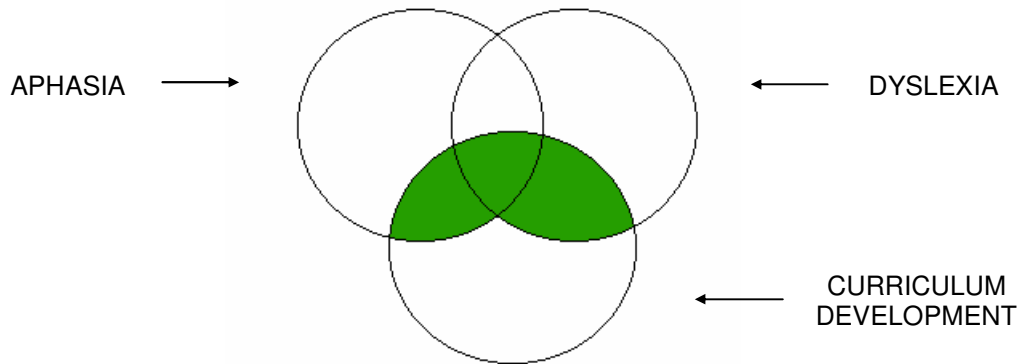
You can use more than one set of parentheses in a SELECT command:

```
?s (aphasia or dyslexia) and (fail? or underachiev?)

          572  APHASIA (IMPAIRMENT OF THE ABILITY TO COMPREHEND OR P...)
         1460  DYSLEXIA (IMPAIRMENT IN THE ABILITY TO READ DESPITE AD...)
        24486  FAIL?
         2459  UNDERACHIEV?
     S3   149  (APHASIA OR DYSLEXIA) AND (FAIL? OR UNDERACHIEV?)
```

You can also specify more than one level of parentheses:

```
?s (visual environment or visual acuity) and (classroom or library) and reading

          226  VISUAL ENVIRONMENT
          223  VISUAL ACUITY (CLEARNESS OR KEENNESS OF VISION AS MEASURED ...)
       107120  CLASSROOM
        45993  LIBRARY
        98691  READING  ((NOTE: USE A MORE SPECIFIC TERM IF POSSIBLE))
     S1     7  (VISUAL ENVIRONMENT OR VISUAL ACUITY) AND (CLASSROOM OR LIBRARY) AND READING
```

(Note that in the above example, we searched specifically for the descriptors VISUAL ENVIRONMENT or VISUAL ACUITY, not for those phrases anywhere in the documents.)

It's often a good idea to use parentheses to clarify the logic of your SELECT statements, even if they are not strictly necessary for correct processing. For example, the SELECT statement above might be clearer if you add another set of parentheses around the combination of (VISUAL ENVIRONMENT OR VISUAL ACUITY) AND (CLASSROOM OR LIBRARY), even though this is not necessary, since AND is processed before OR:

```
?s ((visual environment or visual acuity) and (classroom or library)) and reading

          226  VISUAL ENVIRONMENT
          223  VISUAL ACUITY  (CLEARNESS OR KEENNESS OF VISION AS MEASURED ...)
       107120  CLASSROOM
        45993  LIBRARY
        98691  READING  ((NOTE: USE A MORE SPECIFIC TERM IF POSSIBLE))
    S5      7  ((VISUAL ENVIRONMENT OR VISUAL ACUITY) AND (CLASSROOM OR LIBRARY)) AND READING
```

When using nested logic, be sure that the number of left parentheses is the same as the number of right parentheses. Dialog displays an error message if there are unmatched parentheses in a SELECT command.

►

TIP There's a fine line between using nested logic to do an efficient search and limiting yourself too much by a search that's too narrow. Unless you are very sure that you have exactly the right search strategy (and that doesn't happen very often), put each set of synonyms in a separate search statement, then combine the sets. This gives you the flexibility to mix and match the various search concepts without having to retype the entire search each time.

The above search might be constructed most effectively as follows:

```
Set     Items   Description
S1        445   VISUAL ENVIRONMENT OR VISUAL ACUITY
S2     150865   CLASSROOM OR LIBRARY
S3      98691   READING  ((NOTE: USE A MORE SPECIFIC TERM IF POSSIBLE))
S4          7   S1 AND S2 AND S3
```

And while Select Steps allows you to correct typos, it can also result in an unmanageable number of search sets, particularly if you use adjacency operators. Note what the above search would look like if we used Select Sets and adjacency operators.

```
?  ss VISUAL(w)ENVIRONMENT OR VISUAL(W)ACUITY
S1      29413   VISUAL
S2      91774   ENVIRONMENT (SURROUNDING CONDITIONS, FORCES, OR FACTORS ...)
S3        251   VISUAL(W)ENVIRONMENT
S4      29413   VISUAL
S5        425   ACUITY
S6        306   VISUAL(W)ACUITY
S7        551   SS VISUAL(W)ENVIRONMENT OR VISUAL(W)ACUITY

?  ss classroom or library
S8     107120   CLASSROOM
S9      45993   LIBRARY
S10    150865   SS CLASSROOM OR LIBRARY

?  s reading
S11     98691   READING ((NOTE: USE A MORE SPECIFIC TERM IF POSSIBLE))

?  s s7 and s10 and s11
          551   S7
```

Displaying a List of Current Sets

If you forget what sets you've created, you can see a list of them by entering the DISPLAY SETS command, abbreviated DS, at any time during a search:

```
?display sets

Set    Items   Description
S1     84836   DANGER? OR HAZARD? OR RISK? OR ACCIDENT?
S2    244831   EARTHQUAKE? OR SEISM?
S3     30096   S1 AND S2
```

The DISPLAY SETS command provides a list of your current sets at any time during an online session. This includes all sets created since your last BEGIN command. If you want to see a partial list of your current sets, you can enter a range, e.g., DISPLAY SETS S15-S22.

If your search is interrupted by an accidental disconnect, do **not** enter a BEGIN command after reconnecting to the system. When the *Reconnected* message appears with the file number, use the DISPLAY SETS (DS) command to view the set history and determine where the search was interrupted, then continue with the search strategy.

```
>>> Cost Estimate prior to Disconnect, information only
>>>        19jul06 16:07:42 User303192 Session D63.2
>>>            $0.96    0.016 Hrs File89
>>>      $0.96  Estimated cost File89
>>>      $0.10  INTERNET
>>>      $1.06  Estimated cost this search
>>>      $1.08  Estimated total session cost   0.020 Hrs.
>>>
Reconnected in file 89 19jul06 16:07:50

File 89:GeoRef  1785-2006/Jul B2
(c) 2006 American Geological Institute. All rights reserved.

     Set  Items  Description
     ---  -----  -----------
?ds

Set     Items   Description
S1      84836   DANGER? OR HAZARD? OR RISK? OR ACCIDENT?
S2     244831   EARTHQUAKE? OR SEISM?
S3      30096   1 AND 2
```

TYPE Command: Displaying Records Online

Records retrieved by a search can be displayed online at your computer or sent to your Internet email account.

Note: This last option (specified using the PRINT command rather than the TYPE command) is not available to Graduate Education Program accounts.

Reviewing results online lets you check the precision of your results, make modifications to the strategy as needed and get immediate delivery of the final results. Displaying records online does not end the search or interfere with further modification and SELECTing terms. However, you do pay for your retrieval time while you are viewing records online.

To display records online, use the TYPE command followed by three pieces of information about the records you want to see: the set number, the display format, and the range of records from the set. The full format of the TYPE command (abbreviated T) with these three parameters is shown below.

```
             Format  ┐        ┌  Range of records

  Set number ─┐      │        │
              ↓      ↓        ↓
         type s3/FULL/1-10
```

The TYPE command above requests that records from set S3 be displayed in Format FULL and requests the first 10 records from the set. Records are typically arranged in reverse chronological order so that the first records in a set are the ones most recently added to the database.

The item numbers can indicate a single record, a range of records, or a non-sequential group of records. You can enter the word ALL to view all the records in the set. Use this option carefully, though, as you may accidentally download more records than you intended, incurring unanticipated costs.

```
?type s5/FREE/1-4

?t s5/FULL/1-4,6,9

?t s5/SHORT/all
```

Formats may also be specified by number (e.g., type s1/6/1). Most of the examples in this book will show a format number.

Note: Most display formats incur a per-record charge and you are charged each time you display the record, so remember to save the output the first time.

Dialog provides default values for the set number, format, and item range if you do not include all the parameters in the command. The default values for the TYPE command are summarized below.

TYPE Command Default Values

If you omit the ...	Dialog defaults to the ...
set number	set last TYPEd or the last set created
format	format last TYPEd or Format 2
range of records	next record

You can use slashes to "fill in" parameters in the TYPE command. For example, after entering the TYPE command above, TYPE S3/5/1-10, you could enter TYPE//11-20 to see the next 10 records in the same set in the same format.

You can also display a record directly by using the Dialog accession number followed by a slash (/) and a format number. For example, TYPE 1127496/9 will display record 1127496 in Format 9. This is useful if you have already viewed the record in a short format in a previous search session and wish to view it again in a more complete format without having to re-run the entire search. The Dialog accession number always displays in the upper left corner of the record.

► **Using the DISPLAY Command Instead of TYPE**

TIP The TYPE command presents records in a continuous display. If you want the display of records to pause automatically after your computer screen is full, use the DISPLAY command in place of TYPE: DISPLAY S3/5/1-10. The output will pause after each record or every 24 lines; press the ENTER key to see the next record or 24 lines.

Formats for Displaying Records

Every database on Dialog offers a variety of pre-defined numbered formats for you to use in displaying search results. Each format displays different parts of a record. A summary of pre-defined formats for each database is shown on the database Bluesheet. While online, you can enter the command HELP FORMAT n, where *n* is the database file number.

In addition to the pre-defined formats available, you can define your own formats, specifying precisely which parts of the records you want to see. For example, TYPE S3/AU,TI/1-5 will display the Author and Title fields of records 1 through 5 of set S3.

Pre-defined formats vary from database to database depending on the type of database (bibliographic, directory, numeric, fulltext), record content (abstract, indexing, tables), and other special features. The numbered formats available in bibliographic databases generally follow the pattern summarized below. Be sure to check the Bluesheets for database-specific formats.

Typical Pre-Defined Formats in Bibliographic Databases

Format Number	Format Name	Record Elements That Display	Comments
2		Full record except abstract	Includes bibliographic record plus indexing terms (descriptors, identifiers, classification codes, etc.).
3	MEDIUM	Bibliographic citation	Includes title, source, date, and word count for fulltext documents
5	LONG	Full record except the text field	
6	FREE	Title and Dialog accession number	Useful for reviewing results before requesting a more complete format.
8	SHORT	Title and indexing terms	Useful for evaluating the search results and for determining additional terms to use to modify the strategy.
9	FULL	Full record, including the complete text of the document when available	

The price of search output depends on the format used. See the specific *Bluesheet* for each file and/or the *Price List* for current database costs, both available on the Dialog Web site. To determine the charge per format, enter the command HELP RATES n, where *n* is the file number. In many files there is no charge for a brief format, such as 6 or 8, but there usually is a charge for the bibliographic citation format, such as Format 3, as well as for the more extensive formats. Also, there may be no cost savings in requesting Format 2 instead of Format 5; typically all formats containing the document source information are charged at the same rate. NOTE: The rates shown online when logged on with a Graduate Education Program password reflect the GEP discounts.

► **Interrupting Search Output**

TIP   If you enter a TYPE command and then decide that you want to stop the records from displaying, use the Break function. The key or key combination for the Break function varies depending on the equipment and communications software you are using; for instance, the "*Esc*" key on the computer keyboard will terminate the display in a web-based interface. You can also use Break to stop the processing of a SELECT command if you decide that you no longer want to see its results.

Key Word In Context (KWIC) Format

When you search a fulltext database, you often want to browse through the records to check their relevance. You can display the section of text where your search terms appear using the KWIC (abbreviated K) format in a TYPE command:

```
[File 471] New York Times Fulltext 1980-2006/Aug 29
(c) 2006 The New York Times. All rights reserved.

? s smokeless(W)tobacco (s) (child? or preteen?)

         521    SMOKELESS
       20224    TOBACCO
      378805    CHILD?
         612    PRETEEN?
S1          18  S SMOKELESS(W)TOBACCO (S) (CHILD? OR PRETEEN?)

? t 1/k/1-2
4/K/3 (Item 1 from file: 471)
New York Times Fulltext
(c) 2006 The New York Times. All rights reserved.

...been setting up booths at rodeo grounds where he warns children of the dangers of smokeless
tobacco. Mr. Hallisey says he is a former user of smokeless tobacco, a habit he says he picked
up at age 15 at the rodeo. He exudes...

4/K/7 (Item 2 from file: 471)
New York Times Fulltext
(c) 2006 The New York Times. All rights reserved.

...regulations in Massachusetts that are aimed at sharply curtailing the advertising of
cigarettes, cigars and smokeless tobacco to discourage children from using those products. For
instance, the regulations, which never took effect, sought to ban...

...at least five feet above the floor so they would not be easily noticed by children.

    The regulations, which were challenged by the tobacco industry, represented an effort "to
'childproof' the...
```

You can use the KWIC format in combination with pre-defined format numbers or field display codes:

```
?t s1/3,k/1

1/3,K/1
New York Times Fulltext
(c) 2006 The New York Times. All rights reserved.

04448431 NYT Sequence Number: 934402040611 (USE FORMAT 7 FOR FULLTEXT)
Taking Aim at the Professional Rodeo Circuit's Drug of Choice
TIMOTHY EGAN
New York Times, Late Edition - Final ED, Col 01 , p 14
Friday June 11 2004
Document Type: Newspaper Language: English
Record Type: Fulltext  Section Heading: SECTA
Word Count: 1536

...been setting up booths at rodeo grounds where he warns children of the dangers of smokeless
tobacco. Mr. Hallisey says he is a former user of smokeless tobacco, a habit he says he picked
up at age 15 at the rodeo. He exudes...
```

Note: In many files, KWIC is not a free format. Adding "K" to an otherwise free format, such as Format 6 or 8, may result in a nominal per-record charge. To find out if a file has a KWIC charge, type HELP RATES *n* (where *n* is the file number).

▶

Increasing the Size of the KWIC Window

TIP
The default KWIC text window displays 15 words before and 15 words after the search term. You can increase the size of the KWIC window up to 50 words. Enter SET KWIC 50 before displaying records. You can also decrease the size of the KWIC window—for example, SET KWIC 10.

LOGOFF Command: Disconnecting From Dialog

To end your search activities for a given session, enter the LOGOFF command. Dialog also recognizes BYE, DISC, LOG, LOGOUT, OFF, QUIT, or STOP. When you enter LOGOFF, Dialog displays estimated costs for the search session, then disconnects you from the Dialog computer.

An example of a LOGOFF message is shown below.

```
? logoff

     COST
     07aug06 11:37:47 User311518 Session D293.2
         $1.80    0.515 DialUnits File1
            $0.80  1 Type(s) in Format  5
            $0.00  4 Type(s) in Format  8
         $0.80  5 Types
    $2.60  Estimated cost File1
    $0.53  INTERNET
    $3.13  Estimated cost this search
    $3.15  Estimated total session cost 0.703 DialUnits
```

The system response to LOGOFF includes the date, time (always U.S. Eastern Standard), User Number, searching costs, network costs, TYPE and PRINT command charges, total estimated costs for that file, and a total cost for the entire session. Searching costs are either given in DialUnits, which represent the amount of system resources your search used, or in Connect Time, which is the amount of time spent in a file.

The regular LOGOFF command disconnects you *and* erases all of your sets. If you want to disconnect temporarily and then reconnect to your search in progress, you can enter the LOGOFF HOLD command instead of LOGOFF. LOGOFF HOLD disconnects you from Dialog and displays a cost estimate, but it also retains your search sets for 30 minutes. You will not incur any charges between the LOGOFF HOLD and when you log back on. You can use LOGOFF HOLD during a search when you need time to rethink the strategy, for example. When you reconnect within the 30-minute window, you will be logged back into the file(s) you last searched. Enter a DISPLAY SETS command to see the list of current sets. Do *not* enter a BEGIN command or the sets will be erased.

☐ **Displaying Online Cost Information**

TIP You can enter the COST command at any time during a search to see the estimated cumulative cost of the search session up to that moment. The cost message is identical to the cost information that displays when you enter LOGOFF or LOGOFF HOLD.

Online cost estimates for Graduate Education Program passwords reflect the reduced rates that Dialog gives to educational institutions for teaching purposes.

Summary

Basic Dialog Commands and Operators

Command or Operator	What It Does
BEGIN or B	Specifies the database to be searched.
SELECT or S	Retrieves records for the search terms and creates a single set.
SELECT STEPS or SS	Retrieves records for the search terms and creates a set for each term.
(W)	Used with SELECT to search words that are adjacent and in the order specified.
(N)	Used with SELECT to search words that are adjacent but in any order.
? the truncation symbol	Used with SELECT to search all words beginning with a word root.
OR, AND, NOT	Creates logical relationships among search terms.
TYPE or T	Displays records online in a continuous stream.
DISPLAY or D	Displays records online, one screen at a time.
DISPLAY SETS or DS	Lists all current sets (since last BEGIN command).
SET CONNECT ON/OFF	Changes pricing from Dial Units to Connect time charges
COST	Shows a cost estimate during a search session.
LOGOFF	Disconnects from Dialog and shows a cost estimate.
LOGOFF HOLD	Disconnects from Dialog, shows a cost estimate, and retains your sets for 30 minutes.

PLANNING THE SEARCH STRATEGY

Objectives

In this section, you learn how to:

- Analyze your search topic and plan an appropriate strategy

- Decide if your topic demands a broad or a focused approach

- Understand how different search styles may apply to your search objectives

- Use a search planning worksheet to organize your search terms and logic

- Evaluate the results of a strategy by measuring its recall and precision

Features

- Step-by-step instructions for planning the initial search strategy

- Searching with concept blocks

- Using 'pearl growing' in a strategy

- Searching by 'successive fractions'

- Searching for a known reference

- Entering multiple commands in a single line

Background

In this section, you begin to develop searching skills and assemble tools for an information professional's "searcher's tool kit." This section introduces the most important first step in online searching: planning the search strategy. While most of the cost of a search is determined by the number and format of records you display, planning the strategy before you go online can reduce how much you spend to get the results you want and can certainly save you time, a cost in itself.

The success of an online search plan depends on three factors:
- How carefully you formulate and execute the search terms and search logic
- Which databases you choose for the search
- How much information is available

To help you focus on the processes involved in good search planning, this workbook presents them separately—in reality, they are closely linked. A later section describes in detail how to choose the best databases for a search topic. This section concentrates on how to organize the search concepts, choose a searching style, and create a search strategy that is appropriate to the topic.

Recall vs. Precision

The success of any search strategy may be measured in terms of its **recall** and **precision**. The number of records retrieved is a measure of search recall. High recall means many records are retrieved; low recall means few records are retrieved. Precision measures the relevancy of the records retrieved, based on your perception of whether the records meet your needs. High precision means most of the records retrieved are relevant; low precision means many records may not be relevant. There tends to be an inverse relationship between precision and recall, but the relationship is not absolute. A search can have high recall and high precision, particularly when searching a topic that is broad, popular, and well-indexed. Note too that high precision will depend on the choice of databases and your skill in articulating the specific concepts being researched.

Both precision and recall may be quantified using rough percentage calculations. Recall is the percentage of relevant records found and is based on the number of relevant records available in the file and the number of relevant records retrieved by the strategy:

$$\frac{\text{Number of relevant records retrieved}}{\text{Number of relevant records in file}} \times 100 = \% \text{ Recall}$$

Obviously, it is extremely difficult to determine the total number of relevant records in a file; knowing when your recall is approaching 100% is a skill that comes with practice and familiarity with the databases being searched.

Precision measures a search strategy's effectiveness at retrieving only those records that are relevant to the topic. It is calculated by dividing the total number of records retrieved into the number of relevant records retrieved and multiplying by 100:

$$\frac{\text{Number of relevant records retrieved}}{\text{Total number of records retrieved}} \times 100 = \% \text{ Precision}$$

Very few searches result in 100% precision. Even a very targeted search almost inevitably retrieves some "false drops"— records that have the search terms you specified but still are not relevant to your research. Attempts to achieve 100% precision almost always result in a significant drop in recall. A well-constructed search often results in a precision rate of somewhere between 50% and 80%.

Several key commands and features help you to adjust the precision and recall of a search strategy. In the next section, four basic searching styles are illustrated. Word proximity searching is described later and search examples are provided that show how different proximity operators can be used to modify a search when you want either greater precision or higher recall.

Basic Searching Styles

Before going online, the first step in a successful search is planning the strategy you will use when you connect to Dialog. This means jotting down notes to yourself about the keywords and search terms you will use. To create the best search strategy, use a Search Planning Worksheet and follow these four steps:

1. **Decide exactly what your topic is and what results you expect to achieve.**
    Start by writing down a one-sentence description of the topic. One good way to get started is to ask yourself, "What would be the title of the perfect article?"

2. **Determine what individual concepts are within the topic's description.**
    Not all words in the topic's description represent concepts that are important to your final results. For example, 'the relationship between,' 'research on' or 'effects of' are not concepts that need to be part of your search strategy. Most topics have two to four concepts.

    Similarly, the content of the database you choose may imply certain concepts so that you do not have to include terms for them in the strategy. For example, there is little reason to include EDUCAT? in a search in the ERIC database because the database is dedicated to educational topics.

3. **For each of the concepts, think of alternative words or synonyms.**
    As you work through steps 2 and 3, ask yourself: what words *must* a record contain in order to be relevant? Also, make certain that your search words are a good fit for the database you search. Databases are not unlike printed publications in that each has a target audience and editorial guidelines. Content is selected and included based on these criteria. The words you might choose for a search about a medical condition, for example, would be highly technical in a medical research file, but in a database covering popular health publications, the words should be drawn from a layperson's terminology. Searching on 'kidney cancer' will work in a general magazine file, but you might need to use 'renal neoplasms' in a database for medical professionals.

4. **Perform the search, one concept at a time.**
    Review the results carefully and think about how you can modify your search to fine-tune precision and recall. Remember, there is no one perfect search strategy…stay flexible as you go and be prepared to change your strategy based on the results of the search.

Searching with Concept Blocks

Almost all searches begin with a search plan based on **concept blocks**, listing synonyms or alternate terms for each concept and combining those synonyms with the OR connector.
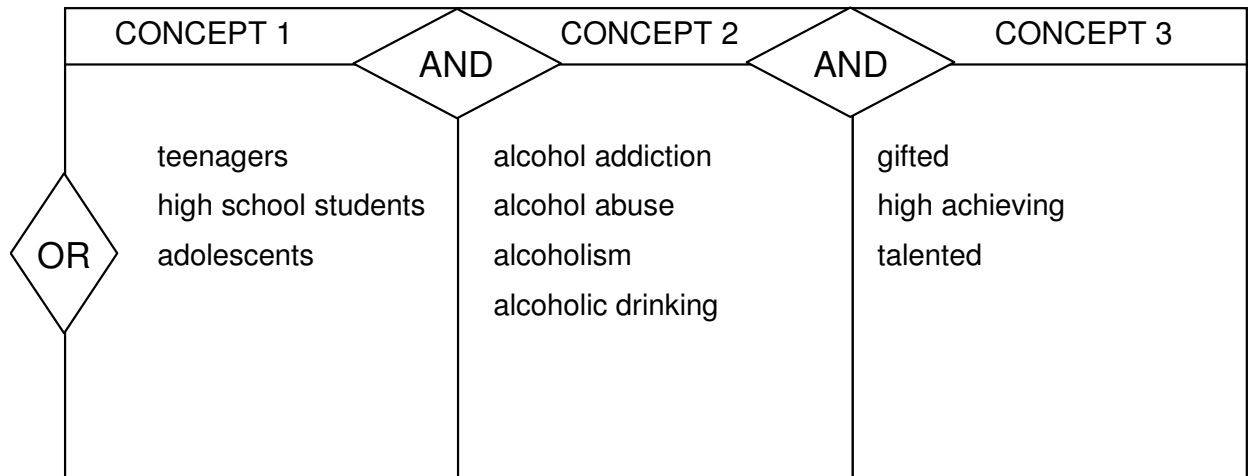
A Search Planning Worksheet is helpful for both listing synonyms and combining concepts. Below, a worksheet has been filled out for a sample search on the topic of alcoholism among gifted teenagers. A blank Search Planning Worksheet is available for download at the GEP Web site.

Search Planning Worksheet

Topic: Alcohol abuse among gifted teenagers.
File(s): ERIC (File 1)

Important Ideas

CONCEPT 1	CONCEPT 2	CONCEPT 3
teenagers	alcohol addiction	gifted
high school students	alcohol abuse	high achieving
adolescents	alcoholism	talented
	alcoholic drinking	

(Concept 1 terms joined by **OR**; Concept 1 **AND** Concept 2 **AND** Concept 3)

Once you have a good list of terms, planning the commands to use is straightforward. The search plan above would use the commands:

> **?select teenage? or high(w)school(w)student? or adolescen?**
> **?s alcoholism or alcoholic? or alcohol(w)(addict? or abuse?)or drinking**
> **?s gifted or high(w)achiev? or talented**
> **?s s1 and s2 and s3**

'Pearl Growing' as a Search Strategy

One of the easiest ways to improve on a search and retrieve additional relevant records is to examine the records resulting from a preliminary strategy. This searching style is called ***pearl growing***—it might also be called "how to take what you have and get more like it."

If you are having trouble thinking of synonyms or alternative terms for a concept, try doing a limited search with just the key concepts you already have. Even if you locate only a few records from this initial strategy, TYPE them in a format that includes their indexing terms and evaluate which new terms might be used in a revised strategy. In bibliographic or fulltext files, Format 8 is usually your best choice. In other databases, consult the Bluesheet and use the display codes for fields containing indexing terms. For example, the field display code PC is often a good choice in business files that use NAICS (North American Industry Classification System) or product codes.

In the sample search that follows, assume that you want research studies on the effect of birth control pills on bone density. Since you are searching a database designed for medical searchers, the terms used may not be "birth control pills" and "bone density." Assume that you do not already know the descriptors used in EMBASE for these concepts. The preliminary strategy begins with a basic search on the two concepts:

```
?begin 72

[File 72] EMBASE 1993-2006/Aug 03
(c) 2006 Elsevier Science B.V. All rights reserved.

      Set  Items  Description
      ---  -----  -----------
?select birth(w)control(w)pill?
         68725  BIRTH
        836799  CONTROL
          5806  PILL?
      S1     95  BIRTH(W)CONTROL(W)PILL?

?s bone(w)density
        234183  BONE
        168221  DENSITY
      S2  19170  BONE(W)DENSITY

?s s1 and s2
            95  S1
         19170  S2
      S3      5  S1 AND S2
```

Although you retrieved only five records, TYPE them in a format that includes the indexing fields.

```
?type s3/8/1-5

3/8/1
EMBASE
(c) 2006 Elsevier Science B.V. All rights reserved.
11044093    EMBASE No: 2000383232
Female athlete triad: Diagnosis, therapy and prevention of the syndrome of disordered eating,
amenorrhea and osteoporosis

FEMALE ATHLETE TRIAD: DIAGNOSE, THERAPIE UND PRAVENTION VON GESTORTEM ESSVERHALTEN, AMENORRHOE
UND OSTEOPOROSE
Publication Date: 2000
DRUG DESCRIPTORS:
estrogen derivative--drug therapy--dt; oral contraceptive agent--drug therapy--dt; gestagen--
drug therapy--dt; calcium--drug therapy--dt; vitamin D--drug therapy--dt
MEDICAL DESCRIPTORS:
* eating disorder--diagnosis--di; *amenorrhea--diagnosis--di; *osteoporosis --diagnosis--di;
*osteoporosis--drug therapy--dt; *osteoporosis--prevention--pc; sport; sports medicine;
syndrome; anamnesis; exercise; estrogen deficiency --drug therapy--dt; symptomatology; dual
energy X ray absorptiometry; disease association; bone mass; training; athlete; oral
contraception; physical performance; human; female; article
CAS Registry Number: 7440-70-2 (calcium)

3/8/2
EMBASE
(c) 2006 Elsevier Science B.V. All rights reserved.
11010426    EMBASE No: 2001061166
Flexural wave propagation velocity and bone mineral density in females with and without tibial
bone stress injuries
Publication Date: 2001
MEDICAL DESCRIPTORS:
* tibia fracture; *stress fracture; bone density; geometry; osteometry; statistical analysis;
human; female; clinical article; controlled study; adult; article

3/8/3
EMBASE
(c) 2006 Elsevier Science B.V. All rights reserved.
07507080    EMBASE No: 1998394567
Lumbar bone mineral density in adolescent female runners
Publication Date: 1998
DRUG DESCRIPTORS:
* bone mineral; *hormone--endogenous compound--ec
MEDICAL DESCRIPTORS:
* bone density; *lumbar spine; *running; amenorrhea; menstrual cycle; training; body fat;
dietary intake; calcium intake; dual energy X ray absorptiometry; hormone blood level; human;
female; human experiment; controlled study; adolescent; adult; article
```

Notice in the displayed records that the descriptors include *oral contraceptive agent* and *oral contraception*. You can increase the recall percentage of set S1 by SELECTing these two terms, using truncation appropriately:

```
?s s1 or oral(w)contracept?

            95  S1
        317003  ORAL
         24271  CONTRACEPT?
         13427  ORAL(W)CONTRACEPT?
    S4   13467  S1 OR ORAL()CONTRACEPT?
```

Likewise, displayed records included the descriptors *osteoporosis* and *bone mass*, which can be added to set S2 to increase recall.

```
?s s2 or bone(w)mass or osteoporosis

         19170  S2
        234183  BONE
        244125  MASS
          8485  BONE(W)MASS
         30491  OSTEOPOROSIS
    S5   41932  S2 OR BONE(W)MASS OR OSTEOPOROSIS
```

Finally, SELECT these new, expanded sets S4 and S5. The result is a set with much higher recall.

```
?s s4 and s5

         13467  S4
         41932  S5
    S6     714  S4 AND S5
```

Searching by Successive Fractions

A search that uses a **successive fractions** approach begins with a broad strategy that produces high recall. Then, restrictions are applied as needed to reduce and focus the final results. Successive fractions searching is sometimes called the 'lawn mower' approach, because you take broad swipes at the most recent set until you have results with the level of precision you want. The restrictions in a successive fractions search may be additional terms or concepts that are ANDed into the strategy, or they may be field suffixes or LIMITs.

A successive fractions approach often works well when you want very precise results. The following example assumes that you are looking for some articles on how libraries use RSS.

```
[File 148] Gale Group Trade & Industry DB 1976-2006/Aug 30
(c)2006 The Gale Group. All rights reserved.

? s rss
S1       9750   S RSS

? s librar?
S2     499744   S LIBRAR?

? s s1 and s2
         9750   S1
       499744   S2
S3        808   S S1 AND S2
```

Casting a fairly wide net (for greater recall) yielded more than 800 records that mention the words RSS and the truncated word LIBRAR? Since you are looking for articles specifically about RSS, you can use the field suffix /TI to specify that the word RSS must appear in the title (or headline). (Field suffixes are explained in detail in the next section.)

```
? s s1/ti
S4         586    S S1/TI

? s s2 and s4
        499744   S2
           586   S4
S5          72   S S2 AND S4

? t s5/8/1-3

5/8/1
Gale Group Trade & Industry DB
(c)2006 The Gale Group. All rights reserved.

0020847793      Supplier Number: 149873260
Yale law library implements catalog-driven RSS feeds.

August, 2006
Company Names: Innovative Interfaces Inc.--Licensing agreements
Industry Codes/Names: BUSN Business; LIB Library and information science
Descriptors: University and college libraries--Technology application; University and college
libraries--Service introduction; University and college libraries--Licensing agreements;
Computer software industry-- Licensing agreements; Applications software--Usage; Online
databases-- Service introduction
Geographic Codes: 1USA United States
Product/Industry Names: 8232000 (Academic Libraries); 7372400 (Applications Software)
Event Codes/Names: 366 Services introduction;389 Alliances, partnerships
Product/Industry Names: 8231 Libraries; 7372 Prepackaged software
NAICS Codes: 51412 Libraries and Archives; 51121 Software Publishers
Trade Names: Innovative Interfaces Feed Builder (Application development software)-- Usage
File Segment: TI File 148

5/8/2
Gale Group Trade & Industry DB
(c)2006 The Gale Group. All rights reserved.
0019642962      Supplier Number: 146957681
Browsing books via RSS feed.(front desk)

May 15, 2006
Industry Codes/Names: LIB Library and information science
File Segment: TI File 148

5/8/3
0019327061      SUPPLIER  NUMBER:  141707577      (USE FORMAT 7 OR 9 FOR FULL TEXT)
The nuts and bolts of delivering new technical reports via database-generated RSS feeds.

Feb, 2006
WORD COUNT: 2874     LINE COUNT: 00223

Company Names: HP Labs--Information management
Industry Codes/Names: COMP Computers; LIB Library and information science
Descriptors: RSS (Content syndication)--Usage; Computers--Services; Computers--Information
management
Geographic Codes/Names: 1USA  United States
Event Codes/Names: 260 General services
File Segment:  TI File 148
```

If you wished to further increase precision, you could narrow down the concept of LIBRAR? by requiring that each retrieved record mention special or academic libraries.

Searching for a Known Item

When you need to verify a reference or to find a complete citation for an article about which you have just a few key facts, you are searching for a *known item*. The objective of a known item search is to locate a record for the specific article in which you are interested. A known item search requires a strategy that produces high precision. The next illustration uses a successive fraction strategy to locate a known item. For the sample search that follows, assume that you are trying to find an article someone recalls reading that describes how pigs in Portland, Oregon are being used to sniff out illegal drugs. This topic has four concepts: Portland, pigs, sniffing, and drugs. You can narrow the search to Portland by selecting the file containing the full text of articles from the Portland *Oregonian*.

```
?b 704

File 704:(Portland)The Oregonian  1989-2006/Aug 03
(c) 2006 The Oregonian. All rights reserved.

     Set  Items  Description
     ---  -----  -----------
?s pig? ?

     S1   5619  PIG? ?
```

Since this database contains the complete text of newspaper articles, make the *pig* set more precise by restricting this word to the Title field.

```
?s pig? ?/ti

     S2    357  PIG? ?/TI
```

Notice how the set has been reduced from more than 5,600 records to 357. Now reduce it further by ANDing it with the concept of *sniffing*. If the set is still too big, reduce it further by including the concept of *drugs*.

```
?s s2 and (sniff? or smell?)
          357  S2
         3168  SNIFF?
        12213  SMELL?
     S3    20  S2 AND (SNIFF? OR SMELL?)

?s s3 and drug?
           20  S3
        68428  DRUG?
     S4     3  S3 AND DRUG?
```

Now that there are only three records in the most recent set, TYPE them in a brief format, and then TYPE the record you want in the fulltext Format 9.

```
?t s4/6/all

4/6/1
08251297
HARLEY THE DRUG SNIFFING PIG
FRIDAY, September 9, 1995
Word Count: 19

4/6/2
07326107
POLICE PIG JUST SAYS `OINK' TO DRUGS
MONDAY, November 22, 1993
Word Count: 595

4/6/3
06178077
IF IT WAS A BOMB-SNIFFING PIG, OUR GOOSE WOULD BE COOKED FOR SURE
THURSDAY June 27, 1991
Word Count: 735

?t s4/9/2

 4/9/2
DIALOG(R)File 704:(Portland)The Oregonian
(c) 2006 The Oregonian. All rights reserved.

07326107
POLICE  PIG JUST SAYS `OINK' TO  DRUGS
Oregonian (PO) – MONDAY, November 22, 1993
By: ERIC L. WEE – of the Oregonian Staff
Edition: FOURTH  Section: LOCAL STORIES  Page: A01
Word Count: 595

TEXT:
Summary:  Harley  the  pot-bellied  pig joins the Portland bureau's drug-sniffing animal team.

    This is a story about drugs, the police and a pot-bellied pig named Harley.

   Actually, he's a miniature Vietnamese pot-bellied pig, and he could be the first drug
sniffing porker on the Portland police force.
                                        •
                                        •  lines deleted
                                        •
```

►

TIP

Entering Multiple Commands: 'Stacking'

For many searches, you will know what your next two or three commands will be. At times like this, you may want to 'stack' the commands by entering them on the same line, separated by semicolons (;).

?**b 72; s birth(w)control(w)pill? ; s bone(w)density**

```
File  72:EMBASE 1993-2006/Aug 03
(c) 2001 Elsevier Science B.V. All rights reserved.

      Set  Items  Description
      ---  -----  -----------
           68725  BIRTH
          836799  CONTROL
            5806  PILL?
      S1      95  BIRTH(W)CONTROL(W)PILL?

          234183  BONE
          168221  DENSITY
      S2   19170  BONE(W)DENSITY
```

?**s s1 and s2 ; t s3/8/all**

```
             95  S1
          19170  S2
      S3       5 S1 AND S2

3/8/1
EMBASE
(c) 2006 Elsevier Science B.V. All rights reserved.
11010426     EMBASE No: 2001061166
Flexural wave propagation velocity and bone mineral density in females
with and without tibial bone stress injuries

Publication Date: 2001
MEDICAL DESCRIPTORS:
* tibia fracture; *stress fracture; bone density; geometry; osteometry;
statistical analysis; human; female; clinical article; controlled study;
adult; article

3/8/2
EMBASE
(c) 2006 Elsevier Science B.V. All rights reserved.
11044093     EMBASE No: 2000383232
Female athlete triad: Diagnosis, therapy and prevention of the syndrome of
disordered eating, amenorrhea and osteoporosis

FEMALE ATHLETE TRIAD: DIAGNOSE, THERAPIE UND PRAVENTION VON GESTORTEM
ESSVERHALTEN, AMENORRHOE UND OSTEOPOROSE
Publication Date: 2000

DRUG DESCRIPTORS:
estrogen derivative--drug therapy--dt; oral contraceptive agent--drug
therapy--dt; gestagen--drug therapy--dt; calcium--drug therapy--dt;
vitamin D--drug therapy--dt

MEDICAL DESCRIPTORS:
* eating disorder--diagnosis--di; *amenorrhea--diagnosis--di;
*osteoporosis --diagnosis--di; *osteoporosis--drug therapy--dt;
*osteoporosis--prevention --pc
sport; sports medicine; syndrome; anamnesis; exercise; estrogen deficiency
--drug therapy--dt; symptomatology; dual energy X ray absorptiometry;
disease association; bone mass; training; athlete; oral contraception;
physical performance; human; female; article
CAS Registry Number: 7440-70-2 (calcium)
```

… lines deleted

HOW DATABASES ARE CONSTRUCTED

Objectives

In this section, you learn:

- Basic terminology for the structure and design of databases

- The general types of databases available on search systems

- What a typical database record contains and how it is organized

- How a database is constructed

- The kinds of indexing applied to a database to make it searchable

- How to use Bluesheets to understand a database's structure and search options

Features

- Using field suffixes to focus a search in a database's Basic Index

- Using prefixes to search in the Additional Indexes of a database

- Searching effectively in both word-indexed and phrase-indexed fields

- Searching a range of numeric terms

- Using the EXPAND and PAGE commands in the Additional Indexes

- Author, journal, and company names (AU=, JN=, CO=)

- Corporate source name (CS=)

- Bluesheets

Background

This section introduces some of the key terminology needed to discuss the structures of databases and the processes involved in making a database searchable.

A *database* is defined as a collection of machine-readable information accessible through a computer. Some databases are also published in printed form, but many are available only online. The term *file* is used interchangeably with the term database, although functionally a single database can be divided into multiple files. For example, the complete MEDLINE□ database is contained in File 155, and is also broken down to a smaller time period in File 154 (1990-present). On Dialog, each database has a corresponding *file number*.

A *record* is a discrete unit of information in a database. In a bibliographic file, a typical record may include a citation, an abstract, and indexing terms. In a fulltext database, a typical record may be an entire journal article. In a directory database, a typical record may consist of a company name, address, and financial information.

A *field* is a distinct part or section of a record. In a bibliographic record, the fields typically include title, author, journal name, publication date, and language. For a non-bibliographic record, the fields vary depending on the type of data included. For example, a business directory may include company name, Standard Industrial Classification (SIC) code or North American Industry Classification System (NAICS) code, city name, ZIP code, and sales figures. As with their print counterparts, the records in a given database may lack uniformity in both form (for example, an author's name may be entered as *Smith, Abigail W.*; or *Smith, A.W.*; or *Smith, A.*) and content (for example, with or without abstracts).

An *information provider* is an organization or company that compiles the records in a database and is responsible for its content. Indexers employed by the information provider evaluate the publications or source data, organize it, and assign indexing terms, such as descriptors or subject classification codes. For example, the American Society of Health-System Pharmacists is the information provider for the Drug Information Fulltext (File 229) database.

Types of Databases

Databases on Dialog can be broadly classified into five types:

- **Bibliographic**
  Each record in a bibliographic database contains citations or references to, but not the full text of, documents such as journal articles, news stories, books, patents, conference papers, technical reports, dissertations, or other similar publications. Many bibliographic records include abstracts or summaries of the documents.

- **Fulltext**
  Each record in a fulltext database contains the text of a newspaper article, magazine article, encyclopedia article, dictionary entry, code of law, or other such document as provided by the information publisher.

- **Directory**
  Each record in a directory database contains factual information about organizations, companies, products, people, or materials.

- **Numeric**
  Each record in a numeric database contains data in a tabular or statistically manipulated form, often with some added text.

- **Hybrid**
  Some databases contain a mix of record types. For example, a database may have fulltext records for some publications and citations and abstracts for other source documents. A company intelligence database may have financial data and news abstracts within the same record. Some databases have images as well as text (notably patents and trademarks).

Sample Database Records

With such a wide variety of databases, one of the main challenges for a search system is to provide consistency for the searcher. It is crucial for the novice searcher to develop an appreciation for the structure and content of a database in order to maintain a perspective on what could be a bewildering array of databases and data. By understanding the structure of a typical record, you can perform effective searches and make good choices when you select databases for a search.

Citations to the same publication may appear in two or more databases in very different formats. In the records below, note the difference in the form and content of the citations for the same document retrieved from two different databases. ERIC (File 1) is an example of a *bibliographic* database; Gale Group Computer Database™ (File 275) is a *hybrid* database, containing mostly fulltext records and some bibliographic records.

Record from ERIC (File 1)

```
[File 1] ERIC 1966-2006/June
(c) format only 2006 The Dialog Corporation. All rights reserved.

01172352 ERIC Number: EJ674996    Clearinghouse Number: IR548050
How Much Security Does Your Library Need?
Banerjee, Kyle
Computers in Libraries v23 n5 p12-14,54-56 May 2003
2003(20030000)
ISSN: 1041-7915
Language: English
Document Type: 80 (Journal articles); 141 (Reports--Descriptive)
RECORD TYPE: ABSTRACT
Journal Announcement: CIJFEB2004

Explains how to keep library systems healthy and functioning by taking sensible security
measures. Examines why hackers would target library systems and how library systems are
compromised. Describes tools that can help, including: firewalls; antivirus software; alarms;
network analysis tools; and encryption. Identifies several strategies for making systems more
secure and discusses what to do if the system is attacked. (AEF)

Descriptors: Computer Networks; *Computer Security; Libraries; Library Administration; *Library
Equipment; *Library Planning; Library Policy
Identifiers: Computer Hackers; *Data Security; *Library Security
```

Record from Gale Group^{SM} Computer Database™ (File 275)

```
[File 275] Gale Group Computer DB(TM) 1983-2006/Jul 21
(c) 2006 The Gale Group. All rights reserved.

02704495     Supplier Number: 100606918 (This Is The FULL TEXT)
How much security does your library need?
Banerjee, Kyle
Computers in Libraries,23,5, 12(6)
May, 2003
ISSN: 1041-7915
Language: English     Record Type: Fulltext
Word Count: 3277    Line Count: 00266

Text:

 Just like eating well, exercising, and getting health, following a few often protects your more
effectively than trying to prevent every conceivable threat with overjealous security measures.

 How much should you worry about security? The answer depends on what your library does and who
your users are. Protecting systems is a lot like protecting your own health--it's mostly a
matter of common sense. Just as people who eat good food and exercise moderately are usually
healthier than those who pop pills and order aggressive medical care for every minor ailment,
```

```
observing a few sensible practices often protects your services more effectively than trying to
prevent every conceivable threat with overzealous security measures.
        Many people think that the point of security is to thwart attacks, but the real purpose of
security is to make hardware, services, programs, and data available to those who need them,
when they need them. This is a very difficult goal to achieve because at least to some extent,
security involves constructing barriers that prevent people from doing things.

. lines deleted
.
.
        Security policies should be based on your organization's goals and not the other way
around. Excessive paranoia will make your services hard to use and may undermine security. If
services don't work properly, staff and users will be frustrated, and people will take actions
that may severely undermine measures designed to protect the network. For example, if staff is
forced to use passwords that are impossible to remember and that expire frequently, you may find
people writing them down and taping them to monitors. If services don't work, you may find staff
setting up rogue servers or taking other measures to circumvent policies. Staff and users should
be allies in promoting security policies. Remember that the purpose of security is to ensure
healthy, reliable systems--not to prevent attacks.

COPYRIGHT 2003 Information Today, Inc.

Geographic Codes/Names: 1USA United States
Event Codes/Names: 260 General services
Product/Industry Names: 8230000 (Libraries); 9912610 (Information Systems)
SIC Codes: 8231 Libraries
NAICS Codes: 51412 Libraries and Archives
File Segment: TI File 148
```

Designing a Database

When a database is added to Dialog, decisions are made on the following file design components:

- which fields in the records will be searchable
- what type of indexing (parsing rules) will be applied to each searchable field
- how special characters will be treated
- what LIMIT capabilities will be built into the file
- what pre-defined (numbered) display formats will be used

Each searchable field is indexed or **parsed**, which means that the words or phrases are extracted and entered into an alphabetical list so that they can be searched separately. A field may be parsed into:

- single words only (e.g., the subject COMPUTER ORIENTED PROGRAMS would be searched as COMPUTER(W)ORIENTED(W)PROGRAMS)

- multiple-word phrases only (searched as ET=Exact trademark or SL=A Bolder Spirit Always Stands Out)

- both single words and multiple-word phrases

- word fragments (such as segments of chemical compounds – the word PROTOPORPHYRIN would be retrieved by the search term PROTOPORPHYRIN and by PROTO(W)PORPHYRIN)

To Dialog, a "single word" is any string of alphanumeric characters surrounded by blank spaces or punctuation. When a field is word-indexed, each individual word is indexed separately, with the exception of nine non-searchable words called **stop words**:

AN	FOR	THE
AND	FROM	TO
BY	OF	WITH

Stop words are not searchable as separate words, but they may appear in the index when embedded in a phrase. For example, the stop word "to" is found in the index when it is part of the descriptor phrase *Back To Basics*.

The complete record is stored under its **accession number** in a **linear file** on the computer. A record in the linear file is analogous to a page in a book. When a record is displayed, it is pulled from the linear file.

Database Indexes

The system software analyzes each discrete record in a database according to the parsing rules for the file and indexes all significant words and phrases. These online indexes are also called **inverted files** and are analogous to the printed index of a book, but are much more exhaustive.

The two indexes prepared for Dialog databases are:

*   the **Basic Index**, which includes all the words from the *subject-related fields*, such as title, abstract, text, descriptors, and identifiers. In a company directory file, the Basic Index usually includes the company name and subject or product codes.

*   the **Additional Indexes**, which include *all other searchable fields*, such as author, date, journal name, classification code, and many others specific to each database. Company directory databases typically include the company name in both the Additional Indexes and Basic Index.

A sample record and a summary of the fields included in the Basic Index and Additional Indexes are provided in the database Bluesheet.

Constructing a Database, Step by Step

This section illustrates the step-by-step process for constructing a database by taking two records from the ERIC database and processing them to build a Basic Index and the appropriate Additional Indexes.

Step 1

The actual records obtained from the information provider in machine-readable form are stored sequentially by accession number in a linear file. Dialog assigns sequential accession numbers if they are not provided by the database supplier. The two records below are extracted from the linear file of ERIC.

```
EJ289520
  Education as Growth: Life-Long Learning.
  Brodbelt, Samuel
  Clearing House, v57 n2 p72-75, Oct 1983
  Language: English
  Document Type: 120 (Opinion papers)
  Journal Announcement: CIJMAR84
  Develops a sociological and philosophical perspective on lifelong education  as  a goal that
is integral to an advanced, technologically directed, democratic society. (FL)
  Descriptors: Adult Education; *Educational Philosophy; Educational Trends; *Futures (of
Society); *Lifelong Learning; Technology


EJ289519
  Servants of Socrates in the Land of Oz.
  Blackburn, Harold
  Clearing House, v57 n2 p69-71, Oct 1983
  Language: English
  Document Type: 120 (Opinion papers)
  Journal Announcement: CIJMAR84
  Discusses the role of the microcomputer in the classroom. (FL)
  Descriptors:  *Educational Improvement; *Educational Quality; *Educational Technology;
*Futures  (of  Society);  Learning  Theories;  *Microcomputers; Technological Advancement
```

Step 2

Dialog analyzes each record for significant words and phrases for entry into the indexes. Each record is divided into fields, each field is labeled, and the position of each word within a field is noted. For example, the first word in the Title field may be labeled **TI1**, the second word, **TI2**, etc. The position of

any stop word is noted, but the word is not indexed. If a field is phrase-indexed only, individual words are not labeled separately, e.g., **AU** represents the full Author field entry.

```
289520      <Accession Number>

Education as Growth Life Long Learning
    TI1  TI2   TI3   TI4  TI5    TI6

Brodbelt, Samuel
 AU

Clearing House, v57 n2 p72-75 Oct 1983
     JN                        PY

Language:  English
           LA

Document Type:  Opinion Paper
                   DT

 Develops  a  sociological and philosophical perspective  on
   AB1   AB2      AB3            AB5              AB6      AB7

lifelong   education   as   a   goal  that  is  integral  to
  AB8       AB9     AB10 AB11 AB12 AB13 AB14   AB15

an advanced, technologically directed, democratic society.
     AB18          AB19          AB20       AB21     AB22

  Descriptors:    Adult   Education;  *Educational  Philosophy;   Educational
                 DE1      DE2            DE3          DE4            DE5

Trends;   *Futures   (of  Society);  *Lifelong Learning; Technology
 DE6       DE7             DE9        DE10      DE11       DE12


289519      <Accession Number>

Servants of Socrates in the land of Oz
  TI1        TI3    TI4     TI6    TI8

Blackburn, Harold
  AU

Clearing House, V57 n2 p69-71 Oct 1983
   JN                        PY

Language:   English
            LA

Document Type:   Opinion Paper
                    DT

Discusses  the  role  of  the  microcomputer  in  the  classroom.
  AB1         AB3             AB6     AB7          AB9

Descriptors:    Educational  Improvement;  Educational  Quality;
                   DE1          DE2           DE3          DE4

Educational  Technology;  Futures  (of  Society);  Learning
  DE5           DE6         DE7        DE9           DE10

Theories; Microcomputers; Technological Advancement.
 DE11      DE12              DE13         DE14
```

Step 3

Words and phrases from fields containing subject-related terms and their positions in the field are entered into the Basic Index, excluding stop words. Words from the Descriptor and Identifier fields are entered as individual words and also as phrases. Each occurrence of a term is stored in the index by accession number, field, and position, providing an "address" referring from the indexed term back to the record in the linear file.

Term	*Address*
a	289520 AB2
	289520 AB11

Term	*Address*
adult	289520 DE1
adult education	289520 DE1,DE2
advanced	289520 AB18
advancement	289519 DE14
an	289520 AB17
as	289520 TI2
	289520 AB10
classroom	289519 AB9
democratic	289520 AB21
develops	289520 AB1
directed	289520 AB20
discusses	289519 AB1
education	289520 TI1
	289520 AB9
	289520 DE2
educational	289520 DE3
	289520 DE5
	289519 DE1
	289519 DE3
	289519 DE5
educational improvement	289519 DE1,DE2
educational philosophy	289520 DE3,DE4
educational quality	289519 DE3,DE4
educational technology	289519 DE5,DE6
educational trends	289520 DE5,DE6
futures	289520 DE7
	289519 DE7
futures (of society)	289520 DE7,DE9
	289519 DE7,DE9
goal	289520 AB12
growth	289520 TI3
improvement	289519 DE2
in	289519 TI4
	289519 AB7
integral	289520 AB15
is	289520 AB14
land	289519 TI6
learning	289520 TI6
	289520 DE11
	289519 DE10
learning theories	289519 DE10,DE11
life	289520 TI4
lifelong	289520 AB8
	289520 DE10
lifelong learning	289520 DE10,DE11
long	289520 TI5
microcomputer	289519 AB6
microcomputers	289519 DE12
on	289520 AB7
oz	289519 TI8
perspective	289520 AB6
philosophical	289520 AB5

Term	Address
philosophy	289520 DE4
quality	289519 DE4
role	289519 AB3
servants	289519 TI1
society	289520 AB22
	289520 DE9
	289519 DE9
sociological	289520 AB3
socrates	289519 TI3
technological	289519 DE13
technological advancement	289519 DE13,DE14
technologically	289520 AB19
technology	289520 DE12
	289519 DE6
that	289520 AB13
theories	289519 DE11
trends	289520 DE6

Step 4

Additional Indexes are created for the terms in all the remaining searchable fields. These indexes are searched with prefixes. For example, the prefix AU= is assigned to the Author field.

Term	Address
au=blackburn, harold	289519
au=brodbelt, samuel	289520
dt=opinion paper	289520
	289519
jn=clearing house	289520
	289519
la=english	289520
	289519
py=1983	289520
	289519

Applying Knowledge of the Database Structure to a Search

Once you understand the basics about how databases are constructed, you can take full advantage of this knowledge to perform effective and efficient searches.

Searching the Basic Index

All of the words resulting from the parsing of the major subject fields are entered into the Basic Index. By default, a SELECT command on any term automatically searches the entire Basic Index:

```
?select oz

     S1      82  OZ
```

You can restrict a term within the Basic Index to a specific field using a *field suffix*, and the Basic Index fields are sometimes called the *suffix fields*. To restrict retrieval to a field, enter a slash (/) followed by the two-letter suffix, such as TI for Title, after the term:

```
?select oz/ti
      S2      41  OZ/TI
```

All of the field suffixes available in a database are defined on the Bluesheet. You can use suffixes to specify the field location of terms in order to make your strategy more precise.

When you combine terms using AND, your search becomes even more specific:

```
?s s1 and socrates/ti
            82  S1
            77  SOCRATES/TI
      S3     1  S1 AND SOCRATES/TI
```

You can retrieve multiple terms in the same title in a single SELECT by either of two methods. To limit two terms to a specific field, you can use parentheses to indicate the limitation:

```
?s oz/ti and socrates/ti

            41  OZ/TI
            77  SOCRATES/TI
      S4     1  OZ/TI AND SOCRATES/TI

?s (oz and socrates)/ti

            41  OZ/TI
            77  SOCRATES/TI
      S5     1  (OZ AND SOCRATES)/TI
```

You can apply more than one suffix to a term, separating the suffixes by commas. This locates the term in at least one of the fields. To look for the term SOCRATES in either the Title or Descriptor field, type:

```
?s socrates/ti,de
      S6      77  SOCRATES/TI,DE
```

You can apply multiple suffixes to groups of terms using parentheses:

```
?s (computer? or microcomputer?)/ti,de

         67592  COMPUTER?/TI,DE
         10186  MICROCOMPUTER?/TI,DE
      S7 69073  COMPUTER? OR MICROCOMPUTER?)/TI,DE
```

You can also apply a field suffix to a set number to focus the results:

```
?s computer? or microcomputer?

         83505  COMPUTER?
         11677  MICROCOMPUTER?
      S8 84730  COMPUTER? OR MICROCOMPUTER?

?s s8/ti,de

      S9 69073  S8/TI,DE
```

Each record in set S7 and set S9 has either COMPUTER? or MICROCOMPUTER? in either the Title or Descriptor field.

However, if a suffix is applied to a set number representing a set created using AND or NOT logic, the suffix will apply to at least one but not necessarily all the terms in the set. Compare the results of set S11 with set S12 shown below.

```
?s moscow and union?

        428686  MOSCOW
        751476  UNION?
   S10   12368  MOSCOW AND UNION?

?s s10/ti

   S11     1250  S10/TI

?s (moscow and union?)/ti

        38382  MOSCOW/TI
        32527  UNION?/TI
   S12      41  (MOSCOW AND UNION?)/TI
```

Each record in set S11 contains:

- both the term MOSCOW AND the term UNION?
- the term MOSCOW in the title field OR the term UNION? in the Title field

Each record in set S12 contains:

- the term MOSCOW in the Title field
- AND the term UNION? in the Title field

To be certain that *all* terms in the set are qualified, SELECT the terms a second time (as shown in set S12), enclosing them in parentheses, and then enter the suffixes. If only OR logic has been used to create the set, you can safely qualify (use a suffix) on the set number.

Searching a Database's Controlled Vocabulary

In many files, the information provider assigns subject-related terms taken from a thesaurus, or controlled vocabulary, to each record. These terms, called **descriptors**, may be single words or multiple-word phrases. Multiple-word descriptors—also known as "bound descriptors" or "descriptor phrases"—receive double indexing: the phrase is indexed as a whole and each individual word (except stop words) is indexed. For example, the descriptor *Educational Improvement* is indexed under:

> EDUCATIONAL
> IMPROVEMENT
> EDUCATIONAL IMPROVEMENT

Some database producers also assign index terms called **identifiers**. Identifiers are assigned to a record by an indexer but are generally not from a controlled vocabulary. Identifiers frequently are proper names, geographic locations, or terms that have not yet been added to the thesaurus.

In the ERIC (File 1) database, terms from the Descriptor and Identifier fields are entered in the Basic Index as both words and phrases. Words from the Title and Abstract fields are entered only as individual words. If you SELECT a complete multiple-word phrase, including spaces and any punctuation, you only retrieve records with that exact phrase in the Descriptor or Identifier field. Notice the difference in retrieval for these two SELECT commands:

```
?s teaching skills

     S1    4464   TEACHING SKILLS

?s teaching(w)skills

       231235   TEACHING
       162498   SKILLS  (COMPLEX MENTAL AND/OR PHYSICAL BEHAVIORS THA...)
     S2    5443   TEACHING(W)SKILLS
```

If you select a word restricted by the suffix /DE, the system retrieves records in which that word appears as part of any descriptor. The SELECT command below retrieves records assigned descriptors, such as TEACHING CONDITIONS, TEACHING EXPERIENCE, STUDENT TEACHING, and REFLECTIVE TEACHING.

```
?s teaching/de

     S3  137016   TEACHING/DE
```

To retrieve a single-word descriptor, but *not* any descriptor phrases that contain the word, enter the suffix /DF, for "full descriptor." For example, to retrieve only the single-word descriptor TEACHING:

```
?s teaching/df

     S4    1031   TEACHING/DF
```

Similarly, you can use the suffix /IF to retrieve a single-word identifier:

```
?s mexico/if

     S5     1415   MEXICO/IF
```

The /DF and /IF suffixes are available in most databases that have descriptors and identifiers. Consult the Bluesheet to find out exactly what suffixes are available.

Searching the Additional Indexes

After the Basic Index is created, all remaining searchable fields are put into the Additional Indexes. Typical fields in the Additional Indexes for a bibliographic database include author, journal name, document type, language, and publication year. Non-bibliographic databases often have numerous Additional Indexes. Take a look at the Bluesheet for DISCLOSURE Database (File 101), which has more than 100 searchable Additional Index fields. Occasionally, some fields are indexed both in the Basic Index and in an Additional Index—the Company Name field is typically made searchable in both indexes.

Each of the Additional Indexes uses a **prefix** to designate a particular field and the Additional Indexes are sometimes called the **prefixed fields**.

Typical Additional Index Fields and Their Prefixes

Prefix	Field Name	Prefix	Field Name
AU=	Author Name	LA=	Language
BN=	ISBN (International Standard Book Number)	PC=	Product Code or Standard Industrial Classification (SIC) Code
CO=	Company Name—also searchable	PD=	Publication Date

	in the Basic Index using the /CO suffix		
CD=	CODEN	PY=	Publication Year
CS=	Corporate Source	SN=	ISSN (International Standard Serial Number)
CY=	City	ST=	State
DT=	Document Type	TS=	Ticker Symbol
JN=	Journal Name	WD=	Word Count

Within a prefixed field, there may be variations in the form of entry for the data—for example, the names of the journals may be abbreviated or spelled out in full. Always check the Bluesheet for a list of prefixed fields and examples of typical entries in the fields. For many prefixed fields—including author, company name, and journal name—you will get the best results if you first use the EXPAND command to verify how data is entered in the field. The EXPAND command is illustrated later in this section.

Whereas using suffixes is an *option* when searching the Basic Index, searching the Additional Indexes *requires* a prefix in front of your search terms. Retrieval is automatically restricted to the field indicated by the prefix. The following pages show examples of searching in the Additional Indexes.

►

Word-Indexed vs. Phrase-Indexed Searching

TIP

Terms in the Basic Index are word-indexed; that is, each word (with the exception of stop words) is indexed individually. Suffix codes are used to restrict retrieval to specific fields within the Basic Index. Since the suffix fields are word-indexed, proximity operators and truncation can be used in the SELECT statement. The Additional Indexes include every field that is indexed using prefix codes, and these fields are generally phrase-indexed. When searching phrase-indexed fields right-hand truncation can be used. However, internal truncation and proximity operators should not be used since the complete phrase is indexed together and if internal truncation or proximity operators are used then 0 (zero) postings will be retrieved.

```
?s co=general electric
```

```
S1      0  CO=GENERAL ELECTRIC
```

Nothing is retrieved because GENERAL ELECTRIC doesn't appear as a phrase with only these two words in the indexes. It is either preceded or followed by other words in the phrase. Words following GENERAL ELECTRIC can be retrieved using right-hand truncation in the SELECT statement:

```
?s co=general electric?
```

```
S2   20816  CO=GENERAL ELECTRIC?
```

```
?s general()electric/co
```

```
114942   GENERAL/CO
 71765   ELECTRIC/CO

S3   21358   GENERAL()ELECTRIC/CO
```

Many more results are retrieved using the word-indexed suffix because this strategy will pick up GENERAL ELECTRIC anytime it appears together in the company field, regardless of whether the company name includes words preceding or following GENERAL ELECTRIC. Examples of records retrieved using the last statement—the word-indexed field—show results that could not be retrieved using prefix codes because the GENERAL ELECTRIC company name is preceded by CANADIAN or PORTLAND and would not be picked up in phrase searching.

```
S4      585  S3 NOT S2
```

```
?t 4/k/1-2

4/K/1

DIALOG®File 148: Gale Group Trade & Industry DB
 ©2006 The Gale Group. All rights reserved.

COMPANY NAMES: Canadian General Electric Company Ltd. …

4/K/2
DIALOG®File 148: Gale Group Trade & Industry DB
 ©2006 The Gale Group. All rights reserved.

COMPANY NAMES: … Portland General Electric Co …
```

Searching Document Type (DT=)

Databases with records for different types of documents often have a Document Type field that you can use to restrict retrieval. Typical document types include: journal article, conference paper, conference proceeding, dissertation, research report, and book.

You can combine subject terms in a search with document type restrictions:

```
? S TEACHING SKILLS AND DT=DISSERTATIONS


        4466  TEACHING SKILLS
        8867  DT=DISSERTATIONS
   S4     42  TEACHING SKILLS AND DT=DISSERTATIONS

? S TEACHING SKILLS AND DT=(DISSERTATIONS OR JOURNAL ARTICLES)

        4466  TEACHING SKILLS
        8867  DT=DISSERTATIONS
      521333  DT=JOURNAL ARTICLES
   S6   1318  TEACHING SKILLS AND DT=(DISSERTATIONS OR JOURNAL ARTICLES)
```

Although entries in the Document Type field tend to be more standardized than entries for names—author, journal, or company names—it is often a good practice to EXPAND on the DT= prefix to learn exactly which document types are indexed in the file. See instructions on the EXPAND command later in this section.

Searching Numeric Indexes

Additional Indexes that contain numeric information can be searched with a Dialog feature known as *range searching*. Range searching retrieves a sequential list of entries in a given index.

Searching a numeric range can be done in several ways. To enter a specific range, use a colon (:) between the lower and higher numbers. To retrieve a range of numbers or all values greater than or less than a certain number, enter one of the following:

:	inclusive range (lowest and highest number)
>	greater than
<	less than
>= (*or* **=>**)	greater than or equal to
<= (*or* **=<**)	less than or equal to

For example, to search the range of publication years (PY=) from 2005 to 2006, enter the command:

```
?s py=2005:2006

     S1  117335  PY=2005:2006
```

To search all companies with sales (SA=) of $300 million or more, enter the command:

```
?s sa=>300m

     S1   10362  SA=>300M
```

You can also truncate an entry in an Additional Index to retrieve all entries beginning with the stem specified. For example, to retrieve all ZIP codes (ZP=) starting with 200, enter the command:

```
?s zp=200?

     S2   30902  ZP=200?
```

► **Searching the Additional Indexes**

TIP
Remember, you <u>must</u> include the prefix when searching the Additional Index fields.
Compare the following results of a search for articles by Naomi S. Travers:

```
?s travers(w)naomi
            727  TRAVERS
           2272  NAOMI
    S1        0  TRAVERS(W)NAOMI
```
    Wrong Logic

```
?s au=travers, naomi s.
    S2       16  AU=TRAVERS, NAOMI S.
```
    Correct Logic

Viewing Database Indexes: EXPAND Command

The EXPAND command can be used in a variety of ways to verify how terms are entered in the index and to improve your search strategy. These techniques are illustrated in the following section.

To view the contents of a database index, enter EXPAND (or E) followed by the search term of interest. EXPAND is completely literal—always enter the term to be EXPANDed without any proximity or truncation operators. To view terms in the Basic Index, simply type the search term following the word EXPAND.

```
? expand intranet

Ref   Items  Index-term
E1        1  INTRANATAL
E2       16  INTRANATIONAL
E3      265  INTRANET
E4      165  INTRANETS
E5        1  INTRANSIGEANCE
E6       14  INTRANSIGENCE
E7        2  INTRANSIGENCY
E8       16  INTRANSIGENT
E9       89  INTRANSITIVE
E10       4  INTRANSITIVE CHOICE BEHAVIOR
E11       4  INTRANSITIVE VERBS
E12       1  INTRANSITIVELY
E13       5  INTRANSITIVES
          Enter P or PAGE for more
```

EXPANDing in the Basic Index

You can use the EXPAND command in the Basic Index. The EXPAND list shows all terms from the Basic Index fields, including descriptors, identifiers, and title, abstract, and text words.

```
? expand microcomputer

Ref   Items  Index-term
E1        1  MICROCOMPUTE
E2        1  MICROCOMPUED
E3     5018 *MICROCOMPUTER
E4        1  MICROCOMPUTER APPLICATIONS
E5        1  MICROCOMPUTER BASED LABORATORIES
E6        1  MICROCOMPUTER BASED TRAINING
E7        1  MICROCOMPUTER CONTROLLED SYSTEMS
E8      200  MICROCOMPUTER DATA
E9        1  MICROCOMPUTER ENCAPSULATION
E10       1  MICROCOMPUTER INVENTORY
E11       1  MICROCOMPUTER PROGRAMMING
E12       1  MICROCOMPUTER SOFTWARE AND INFO FOR TEACHERS

        Enter P or PAGE for more
?p

Ref   Items  Index-term
E13       1  MICROCOMPUTER SPECTRUM ANALYSIS MODELS
E14       1  MICROCOMPUTER STRUCTURE
E15       1  MICROCOMPUTER STUDENT RATIO
E16       1  MICROCOMPUTER TERMINALS
E17       1  MICROCOMPUTER TO MAINFRAM CONNECTIVITY
E18       1  MICROCOMPUTER TRAINING
E19      22  MICROCOMPUTERIZED
E20       1  MICROCOMPUTERNETZE
E21    8965  MICROCOMPUTERS
E22       1  MICROCOMPUTERS MATHEMATICAL MODELS
E23       3  MICROCOMPUTERS UNITED STATES
E24       7  MICROCOMPUTERS

        Enter P or PAGE for more
```

Single words in an EXPAND list in the Basic Index may be from any of the word-indexed fields. Bound phrases that display (e.g., those in lines E4 through E18) are from one of the phrase-indexed fields, typically the Descriptor and Identifier fields.

In some databases, you can EXPAND in the Basic Index to view the online thesaurus. This special feature is described in *Using Controlled Vocabulary*, available on the GEP Web site.

EXPANDing in the Additional Indexes

EXPAND is especially useful for viewing the Additional Indexes. Always check the Bluesheet to determine the prefixes for the fields you want to use. EXPAND is strongly recommended for fields containing names that can have many variants, such as Company Name, Author, and Journal Name. To view the contents of an Additional Index, use the appropriate prefix. To see a list of the Company Name index terms for Whole Foods (and to see the number of records indexed to each term):

```
? expand co=whole foods

Ref   Items  Index-term
E1        2   CO=WHOLE EARTH NETWORKS LLC
E2        1   CO=WHOLE FOOD CO.
E3        1  *CO=WHOLE FOODS
E4        6   CO=WHOLE FOODS INC.
E5        3   CO=WHOLE FOODS MARKET
E6        3   CO=WHOLE FOODS MARKET CALIFORNIA INC.
E7     1109   CO=WHOLE FOODS MARKET INC.
E8        1   CO=WHOLE FOODS MARKET SOUTHWEST L.P.
E9        2   CO=WHOLE FOODS MARKETING
E10       1   CO=WHOLE HEALTH DISCOUNT CENTERS
E11       2   CO=WHOLE HEALTH INSURANCE NETWORK INC.
E12       1   CO=WHOLE HEALTH MANAGEMENT

        Enter P or PAGE for more
```

Notice that these lists of index terms include reference numbers or "E numbers" (E1, E2, E3 ...) under the column headed **Ref**. The term you entered after the EXPAND command appears on line E3. Beneath

the next column headed **Items**, the number of records for each **Index-term** is shown. Since there was one entry for the exact term searched (CO=WHOLE FOODS), the term listed on E3 shows 1 record.

An EXPAND list on DialogClassic always shows twelve lines from the index, using E numbers E1 to E12. To see the next twelve lines, enter the PAGE command (abbreviated P). As you page through the list, the E numbers will go up to E50, then the numbering begins again at E1. Be sure to SELECT relevant E numbers from the list before entering a new EXPAND command or paging through past E50. PAGE is illustrated in the Author section that follows. An EXPAND list on DialogWeb shows 50 lines from the index.

SELECTing Terms after EXPANDing

To display the records associated with an E number (or a group of E numbers) you must first SELECT them and create a set. Use OR logic and/or the colon range operator (:) to group together all the variant terms. Since the index terms for Whole Foods on lines E3 through E8 are all of interest, enter:

```
?  select e3:e8

S1       1123    CO='WHOLE FOODS':CO='WHOLE FOODS MARKET SOUTHWEST L.P.'
```

Set S1 contains all 1123 records that have some form of the company name Whole Foods in the Company Name (CO=) field.

Searching on Author Names (AU=)

EXPANDing on an author name is essential when searching for all articles by a given author, or when uncertain about consistency and use of initials, hyphens, Jr., Ed., or other variations in form of entry. A given author often has multiple entries in the Author field, as shown below.

```
?e au=pialorsi, f

Ref   Items  Index-term
E1       1  AU=PIAGET, J.
E2      13  AU=PIAGET, JEAN
E3       0 *AU=PIALORSI, F
E4       2  AU=PIALORSI, FRANK             ←
E5       1  AU=PIALORSI, FRANK PAUL        ←
E6       1  AU=PIALORSI, FRANK, ED.        ←
E7       1  AU=PIALOUX, BERNARD
E8       1  AU=PIAMONTE, JOHN S.
E9       3  AU=PIAN, CANTA
E10      1  AU=PIANAROSA, ALBERTINA
E11      1  AU=PIANKO, NORMAN
E12      3  AU=PIANKO, SHARON

        Enter P or PAGE for more
```

In the EXPAND display above, there are three entries (E4, E5, and E6) for the name Frank Pialorsi. The second entry includes a middle name, and the third entry is for him as an editor. Notice also that entries E1 and E8 include initials. These variations are typical of the Author field.

While truncation can be used to retrieve variations of an author name, the EXPAND command is recommended because you know exactly which variants are being retrieved when you select E numbers. Before EXPANDing an author name, check the Bluesheet to find out if the database uses a comma and a space or only a space after the surname. Then enter the surname, the comma if needed, and the first initial.

```
File 7:Social SciSearch(R) 1972-2006/Aug W3
       (c) 2006 The Thomson Corp

?e au=pialorsi f

Ref   Items  Index-term
E1        1  AU=PIALLAT B
E2        1  AU=PIALLEPORT T
E3        2 *AU=PIALORSI F      ←
E4        1  AU=PIALOUX
E5        1  AU=PIALOUX B
E6        2  AU=PIALOUX G
E7        6  AU=PIALOUX M
E8        4  AU=PIALOUX P
E9        1  AU=PIAMARTA F
E10       1  AU=PIAMENTA R
E11       3  AU=PIAMJARIYAKUL U
E12       2  AU=PIAMONTE DPT

          Enter P or PAGE for more
```

In the next example, you want to locate articles authored or edited by Aaron J. Miller. Enter the appropriate EXPAND command:

```
[File 1] ERIC 1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

? expand au=miller, a

Ref   Items  Index-term
E1        1  AU=MILLER-WOOD, D. J.
E2        1  AU=MILLER,
E3        0  AU=MILLER, A
E4        2  AU=MILLER, A.                    •
E5        2  AU=MILLER, A. CAROLYN
E6        2  AU=MILLER, A. CHRISTINE
E7        1  AU=MILLER, A. G.
E8        3  AU=MILLER, A. H.
E9        1  AU=MILLER, A. J.                 •
E10       1  AU=MILLER, A. KATE
E11       1  AU=MILLER, A. PATRICIA
E12       2  AU=MILLER, A. R.

          Enter P or PAGE for more
```

Notice the relevant entries on lines E4 and E9. Since it's possible that additional variants of the name may appear after line E12, use the PAGE (or P) command at this point.

```
? page

E13       1  AU=MILLER, A. T.
E14       1  AU=MILLER, A. THERESE
E15       1  AU=MILLER, A.J.
E16      17  AU=MILLER, AARON J.          ←
E17       1  AU=MILLER, AARON J., ED      ←
E18       3  AU=MILLER, AARON J., ED.     ←
E19       2  AU=MILLER, ABRAHAM
E20       1  AU=MILLER, ABRAHAM H.
E21       9  AU=MILLER, ADAM
E22       1  AU=MILLER, ADAM DAVID
E23       1  AU=MILLER, ADAM W.
E24       3  AU=MILLER, ADAM W., JR.

          Enter P or PAGE for more
```

The next "page" presents the next twelve lines from the Author index, with reference numbers from E13 to E24. To retrieve all records in which Aaron J. Miller is indexed as an author, SELECT lines E4, E8 and E15 through E18:

```
?S E4 OR E9 OR E15:E18

        2  AU=MILLER, A.
        1  AU=MILLER, A. J.
       22  AU=MILLER, A.J.:AU=MILLER, AARON J., ED.
S1     25  AU='MILLER, A.' OR AU='MILLER, A. J.' OR AU='MILLER, A.J.':AU='MILLER, AARON J., ED.'

?t s1/3/5

 DIALOG(R) File 1:ERIC
(c) format only 2006 Dialog. All rts. reserv.

00607080 ERIC NO.: ED263324 CLEARINGHOUSE NO.: CE042777
Preparing Ohio's Youth through Occupational Work Adjustment and Occupational Work Experience
Programs: Prospects for the Future.
   Miller, Aaron J.; Bragg, Debra D.;
CORP. SOURCE: Ohio State Univ., Columbus. Coll. of Education. (RUF67460)
   112pp.
   1985 (19850000)
SPONSORING AGENCY: Ohio State Dept. of Education, Columbus. Div. of Vocational Education.
(RUF67055)
```

Searching on Company Names (CO=)

EXPANDing in the Company Name field is helpful when you are uncertain of the complete name, or when searching for all forms of the name.

```
[File 502] Teikoku Databank:Japanese Companies 2006/Jun
(c) 2006 Teikoku Databank. All rights reserved.

?e co=volkswagen

Ref   Items   Index-term
E1       2   CO=VOLKS CO LTD
E2       1   CO=VOLKS INC
E3       0   CO=VOLKSWAGEN
E4       1   CO=VOLKSWAGEN AUDI PARTS CENTER
E5       1   CO=VOLKSWAGEN FINANCE JAPAN KK
E6       1   CO=VOLKSWAGEN GROUP JAPAN K K
E7       1   CO=VOLKSWAGEN TOKYO KK
E8       1   CO=VOLLEY BOARD CO LTD
E9       1   CO=VOLLEY CO LTD
E10      1   CO=VOLLMER JAPAN CORP
E11      1   CO=VOLT CO LTD
E12      1   CO=VOLTA 92 CO LTD


        Enter P or PAGE for more
```

Searching on Journal Names (JN=)

EXPANDing on a Journal Name (usually JN=) is particularly useful when the entry may be abbreviated, spelled out, or even misspelled. Three different entries for the journal *Behavioral Science* appear in the example below:

```
[File 1] ERIC 1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

? e jn=behav

Ref   Items   Index-term
E1       1   JN=BCLA REPORTER
E2       4   JN=BEA JOURNAL OF METROPOLITAN NEW YORK
E3       0   JN=BEHAV
E4       6   JN=BEHAV SCI
E5       1   JN=BEHAV. RES. & THERAPY
E6       1   JN=BEHAVIOR GENETICS
E7      16   JN=BEHAVIOR IN OUR SCHOOLS
E8     170   JN=BEHAVIOR MODIFICATION
E9       1   JN=BEHAVIOR MODIFICATION MONOGRAPHS
E10      1   JN=BEHAVIOR RESEARCH METHODS, INSTRUMENTS, & COMP
E11      1   JN=BEHAVIOR RESEARCH METHODS, INSTRUMENTS, AND CO
E12     30   JN=BEHAVIOR THERAPY

        Enter P or PAGE for more
```

```
? p

E13     138  JN=BEHAVIORAL & SOCIAL SCIENCES LIBRARIAN
E14      25  JN=BEHAVIORAL AND SOCIAL SCIENCE TEACHER
E15     673  JN=BEHAVIORAL DISORDERS
E16       1  JN=BEHAVIORAL RESEARCH METHODS AND INSTRUCTION
E17      15  JN=BEHAVIORAL SCIENCE
E18       4  JN=BEHAVIOUR & INFORMATION TECHNOLOGY
E19      10  JN=BEHAVIOUR AND INFORMATION TECHNOLOGY
E20       7  JN=BEHAVIOUR PROBLEMS BULLETIN
E21       1  JN=BEHAVORIAL SCIENCE
E22       1  JN=BELL JOURNAL OF ECONOMICS AND MANAGEMENT SCIEN
E23       1  JN=BENCHMARKS
E24       1  JN=BENEFITS2

? s e4 or e17 or e21

          6  JN=BEHAV SCI
         15  JN=BEHAVIORAL SCIENCE
          1  JN=BEHAVORIAL SCIENCE
S1       22  JN='BEHAV SCI' OR JN='BEHAVIORAL SCIENCE' OR JN='BEHAVORIAL SCIENCE'
```

Because you often encounter abbreviations when EXPANDing on journal names and company names, enter only the first part of the name or just enough to put you into the right part of the index to see possible entries for your term. Sometimes it may be necessary to enter more than one EXPAND to view all variants.

Searching on Corporate Source (CS=)

The Corporate Source (CS=) field is word-indexed in most databases so that proximity operators must be used. In the example below, the items indexed under CS=LAWRENCE (shown on line E3) include corporate source names, such as *Lawrence Berkeley Laboratory*, *Lawrence Livermore Laboratory*, *St. Lawrence University*, and *Lawrence Hall of Science*.

```
[File 1] ERIC 1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

?e cs=lawrence

Ref   Items  Index-term
E1        3  CS=LAWNDALE
E2        1  CS=LAWNDALE SCHOOL DISTRICT, CA.
E3      664 *CS=LAWRENCE
E4        1  CS=LAWRENCE COUNTY ADULT EDUCATION, MOULTON, AL.
E5        2  CS=LAWRENCE INST. OF TECH., SOUTHFIELD, MI.
E6        1  CS=LAWRENCY PUBLIC SCHOOLS, CEDARHURST, NY.
E7        4  CS=LAWRENCE UNIFIED SCHOOL DISTRICT 497, KS.
E8        3  CS=LAWRENCE UNIV., APPLETON, WI.
E9        1  CS=LAWRENCE UNIV., APPLETON, WI. DEPT. OF SOCIOLO
E10       2  CS=LAWRENCEVILLE
E11       3  CS=LAWS
E12       2  CS=LAWYER

            Enter P or PAGE for more
```

To retrieve the corporate source *Lawrence Hall of Science*, use the SELECT command below. Notice how parentheses are inserted after the prefix and at the end of the command—this makes the CS= prefix apply to all of the terms.

```
?s cs=(lawrence()hall(1w)science)

          664  CS=LAWRENCE
          203  CS=HALL
         6941  CS=SCIENCE
    S1    110  CS=(LAWRENCE()HALL(1W)SCIENCE)
```

► **EXPANDing on Terms That Contain Punctuation**

TIP
When you use the EXPAND command to view an index, you are looking up your term as though using a dictionary. Whatever you enter after EXPAND is interpreted literally and simply inserted into the index where it fits alphabetically. If you need to EXPAND a term that contains punctuation, enter the term exactly.

```
?e jn=women's history

>>>Warning: unmatched quote found

Ref   Items  Index-term
E1      17   JN=WOMEN & POLITICS
E2       2   JN=WOMEN'S ANNUAL
E3       0  *JN=WOMEN'S HISTORY
E4      37   JN=WOMEN'S HISTORY REVIEW
E5      68   JN=WOMEN'S STUDIES
E6      40   JN=WOMEN'S STUDIES INT. FORUM
E7       2   JN=WOMEN'S STUDIES INT. Q
E8     160   JN=WOMEN'S STUDIES INTERNATIONAL FORUM
E9       5   JN=WOMEN'S STUDIES QUARTERLY
E10     76   JN=WORD & IMAGE
E11      4   JN=WORD AND IMAGE
E12      7   JN=WORKING PAPERS FOR A NEW SOC

           Enter P or PAGE for more
```

Dialog Bluesheets

To find out about a specific database's content structure, refer to its Bluesheet. Each Bluesheet shows a typical record and describes the content of the Basic and Additional Indexes in the Search Options section:

SEARCH OPTIONS

Basic Index		
Search Suffix	**Field Name**	**Indexing**
/AB	Abstract	word
/DE	Descriptor	word & phrase
/ID	Identifier	word & phrase
/LP	Lead Paragraph	word
/TI	Title	word
/TX	Text	word

Additional Indexes		
Search Suffix	**Field Name**	**Indexing**
AU=	Author Name	phrase
CS=	Corporate Source	word
JN=	Journal Name	phrase
LA=	Language	phrase
PD=	Publication Date	phrase
PY=	Publication Year	phrase
WD=	Word Count	numeric

*** Composite Bluesheet showing common search options.

This same information is also available online using HELP commands. For example, for a list of searchable fields in ERIC (File 1), enter HELP FIELD 1.

```
?help field 1

FIELD1 :ERIC


                        SEARCH OPTIONS


   BASIC INDEX
SEARCH DISPLAY  FIELD NAME/
PREFIX CODE     INDEXING/EXAMPLES


None    None    All Basic Index Fields
                Word
                S INNER(W)CITY(W)NEIGHBORHOOD?
/AB     AB      Abstract
                Word
                S NEIGHBORHOOD(W)AFFILIATION?/AB
/DE     DE      Descriptor -1
                Word & Phrase
                S MANIPULATIVE(W)MATERIALS/DE
                S MATHEMATICS ACTIVITIES/DE
/ID     ID      Identifier -2
                Word & Phrase
                S AFFILIATIVE(W)BEHAVIOR/ID
                S AFFILIATIVE BEHAVIOR/ID
/NT     NT      Note -3
                Word
                S CHILD(W)DEVELOPMENT/NT
/TI     TI      Title
                Word
                S  GEOMETRIC(W)THINKING/TI


ADDITIONAL INDEXES
SEARCH DISPLAY  FIELD NAME/
PREFIX CODE     INDEXING/EXAMPLES


AA=     AA      ERIC Document Number
                Phrase
                S AA=ED430060
AU=     AU      Author
                Phrase
                S AU=VAN HIELE, PIERRE  M.
AV=     AV      Availability
                Word
                S AV=FAX
None    AZ      DIALOG Accession Number
BN=     BN      International Standard Book Number (ISBN)
                Phrase
                S BN=0-7246-1346-3
CH=     CH      Clearinghouse Number
                Phrase
                S CH=SE561085


                                •
                                • Additional Index Fields omitted
                                •
```

Sample Bluesheets

Files representing three different types of database are:

- **ERIC (File 1)**
  A bibliographic database containing references and abstracts for journal articles, conference papers, and reports in the field of education and related disciplines.

- **Gale Group Computer Database™ (File 275)**
  A hybrid database containing many fulltext records as well as bibliographic records with abstracts.

- **Standard & Poor's Register - Corporate (File 527)**
  A directory database where each record is a description of a company, including key facts, such as address, phone, names of top officers and directors, and financial data.

To review the Bluesheets for these and other files, please visit the Dialog Web site.

Section 4

CHOOSING THE RIGHT DATABASE & SEARCHING MULTIPLE DATABASES

Objectives

In this section, you learn how to:

- Decide before going online if a database is a good candidate for a search
- Display database-specific information online to help you plan a search
- Obtain Dialog user documentation
- Use DIALINDEX® (File 411) to help you choose appropriate databases for a topic
- Perform effective searches in multiple databases simultaneously
- Remove duplicate records in a multi-file search

Features

- Overview of Dialog user documentation
- Dialog Bluesheets database
- DIALINDEX® (File 411)
- Dialog OneSearch®
- REMOVE DUPLICATES command

Background

This section provides an overview of the printed documentation and online user aids that can help you decide which databases are best for a given topic. Many of these user aids are also available on Dialog's Web site at *http://library.dialog.com.* Later in this section you learn about online tools that help you choose databases.

To get started, use this short checklist of questions about the topic and the search objective:

- What is the general subject area? (medicine, business, science, etc.)
- What is the focus of the question? (technical, consumer-oriented, latest news, etc.)
- How much information is needed? (comprehensive search or "a few good articles")
- What general type of database is needed? (bibliographic, fulltext, directory, etc.)

The answers to these questions will help to narrow your choice from the more than 600 databases listed in the *Dialog Database Catalog*.

Evaluating a Database: 4 Cs to Consider

With hundreds of databases from which to choose, the challenge of finding your information "needle" begins with choosing the best "haystack" to search. *One of the easiest ways to waste money online is to conduct the right search in the wrong database.* In this section, you learn how to select databases that are most likely to produce the results you seek.

Databases vary in several important ways. Keep in mind these "4 Cs" to evaluate the databases you are considering for a given search objective:

- **Content**
    - o **Subject** content (law, medicine, mathematics, engineering, etc.)

    - o Sources: **where the data comes from** (newspapers, patents, company interviews, trade publications)

    - o Record content: exactly **what data is in a record** (fields, text, abstracts, numeric data) and **how can the data be searched** (e.g., do you need a directory file that indexes NAICS codes?)

    - o Intended audience: what **technical level** is presented (medical research for health professionals vs. medical information for consumers)

- **Coverage**
    - o **Size** of the database (number of records, years covered)

    - o **Scope** of coverage (national or international, number of journals indexed, types of companies included, etc.)

- **Currency**
    - o How often is the file **updated** (daily, weekly, monthly)

    - o Timeliness: what is the **lag time** (newswires, published or unpublished data)

- **Cost**
    - o How much can you expect to pay for what you need: DialUnit or Connect Time; per-record charges (online Types or offline Prints)

Dialog User Documentation

Print Tools

Dialog has extensive documentation on database content and search features. The *Dialog Database Catalog* contains brief descriptions of all databases, including subject categories, dates of coverage, and name of the information provider. A subject index is also included.

The *Dialog Pocket Guide* and *Successful Searching with Dialog Command Language* are the major reference manuals on all system features and commands. Both are available on the Dialog Web site.

Bluesheets provide essential documentation for day-to-day searching. Each database-specific Bluesheet shows a sample record with all fields labeled and lists details about search options and output formats. Bluesheets are available online in File 415 and through the Dialog Web site. Note that the most up-to-date Bluesheet for any databases is the PDF version found at: *http://library.dialog.com/bluesheets*. The hyperlink to the PDF is found at the top of the HTML Bluesheet for any given database.

In addition to the documentation described above, you can get Dialog user aids online through Dialog Homebase and through the Dialog Web site.

Consulting the Bluesheet

After you identify a few candidate files, examine their Bluesheets to see what a typical record in the file looks like. Check on search and display features that you know you will need. The Bluesheet sample record is often your best guide to deciding if a given database is worth pursuing for your topic.

Use the Bluesheet to find out if a file has the search options and coverage you need. For example,

- If you need to create a list of companies that fit certain criteria, make certain that there are fields for what you need. For example: Can the file be searched by sales figures? Are phone numbers included in the records?

- If you want to have complete text for relevant articles—not just abstracts or summaries—check to make certain that the records include full text.

- If you need to locate very recent publications, check the Dialog File Data section on the Bluesheet to verify the timeliness of updates to the database.

- If you are going to search on an author's name, look at the author name (AU=) example on the Bluesheet to see how names are formatted in the file: Are commas used after the surname? Are first names spelled out or abbreviated to initials only?

Dialog Bluesheets™ (File 415) Database

The complete text of all Dialog Bluesheets is available in File 415. Each record in the file is equivalent to one Bluesheet. The records in the file always reflect the most recent information available. You can retrieve a single Bluesheet by SELECTing the file name or number in the Title field. For example, SELECT FILE(W)275/TI retrieves the Bluesheet for File 275. To display the complete Bluesheet, use Format 9. You can use File 415 to locate databases that have a particular searchable field or subject area.

Database Selection Tools

DIALINDEX®

Even after looking in the *Dialog Database Catalog* and reviewing appropriate Bluesheets, you can't always be certain that you've chosen the best databases until you actually conduct the search using your own keywords and search terms. An interim step for many situations is a quick search in DIALINDEX (File 411), a preview file where you can scan multiple files at a low, flat rate on DialogClassic and find out if your search terms appear. This step can save you both time and money because you use DIALINDEX *before* BEGINing in a potentially expensive database.

Scanning a Large Group of Databases: DIALINDEX (File 411)

For most topics, your search preparation and database selection planning is done before logging on. There are many topics, however, that do not fit easily into the scope of a single database or small group of databases. For these searches, you can go online and scan a large group of files quickly and inexpensively using DIALINDEX to determine which files are most appropriate for your search. DIALINDEX is a single database which includes the indexing from most of the other databases on Dialog.

It's important to understand what DIALINDEX does and what it does *not* do. Put simply, you BEGIN in DIALINDEX, choose a group of databases to scan, and then enter a *single* SELECT command to find out how many records contain your search terms. Because DIALINDEX is an index only, you will *not* be able to display records while in DIALINDEX. You are also *not* able to enter search terms in multiple SELECT commands or to combine set numbers. In fact, no sets are formed while in DIALINDEX.

When you finish a search in DIALINDEX, you typically BEGIN in the databases that demonstrated the best retrieval, perform the search strategy (refining as needed), and display the records. Using DIALINDEX, you can avoid costly searching in databases that ultimately prove to contain very few or no records on your topic.

Searching DIALINDEX on DialogClassic

To use DIALINDEX on DialogClassic, follow these steps:

1. BEGIN 411.

2. Enter a SET FILES command followed by the databases you want to scan. SET FILES may be abbreviated as SF.

    You can enter any combination of file numbers or database category acronyms. DIALINDEX categories identify subject groupings of databases so that you can easily scan a related group of files. Categories and their acronyms are listed at the back of the *Dialog Database Catalog,* on the File 411 Bluesheet, and on the Dialog Web site at *http://library.dialog.com/bluesheets*.

    There are two types of categories in DIALINDEX: SuperCategories and regular categories. SuperCategories encompass very large groups of subject-related files. For example, the SuperCategory acronym ALLSCIENCE groups together ~150 databases covering scientific subjects. Regular categories are smaller and more specific, typically including 10 to 15 databases. Examples include ARCHITEC for architectural files, BIOTECH for biotechnology, and NUTRIT for nutrition and food-related topics. The fulltext newspaper files are combined in categories by state and by region— PAPERSCA for California newspapers, PAPERSSE for newspapers in the U.S. Southeast, PAPERSUK for newspapers in the U.K., and so on.

    You can create your own grouping of files to be scanned by entering any combination of file numbers and category acronyms, separated by commas. You can even use the NOT operator to exclude files from a category. Here are a few sample SET FILES commands:

```
?SET  FILES  ALLSCIENCE
?SF  BIOTECH,PATENTS  NOT  636,144
?SF  2,295,8,32
?SF  TRADEMKS,751
```

3. Enter your search terms in a single SELECT command.

    Remember that because no sets are formed, the complete strategy must be entered in *one* SELECT statement. You may want to use only the most important search terms at this point and not try to include all the synonyms you have in mind; you will usually perform a more refined strategy when you BEGIN the files you identify as the best ones. Remember, this is simply a way of seeing what files look the most promising for your search. Be sure to use parentheses as needed so that your search logic is correct; if you want terms ORed together, enclose them in parentheses. Examples:

    ```
    ?SELECT  (SMOK?  OR  TOBACCO)  AND  HEART(W)DISEASE?
    ?SELECT  CHOLESTEROL  AND  (DOG  OR  DOGS)
    ```

    Keep in mind that fulltext files will have higher record counts if you SELECT terms in the full Basic Index of the files selected. Avoid controlled vocabulary terms that may not apply to all files in the group.

    You can EXPAND while in DIALINDEX, but no E numbers will display.

4. Enter the RANK FILES command to see the files with the most records on your topic listed first.

5. BEGIN in the databases that have the best coverage of your topic.

6. Perform the search and display records.

Two DIALINDEX sample searches follow. The first search uses a regular DIALINDEX category acronym and the second one uses a SuperCategory.

DIALINDEX Sample Search—1

```
File 411:DIALINDEX(R)

DIALINDEX(R)
   (c) 2006 Dialog

*** DIALINDEX search results display in an abbreviated ***
*** format unless you enter the SET DETAIL ON command. ***

? SET FILES FOODSCI

You have 15 files in your file list.
   (To see banners, use SHOW FILES command)

? S (GARLIC OR ONION?) AND (ODORLESS OR ODOURLESS OR DEODORI?)

Your SELECT statement is:
   S (GARLIC OR ONION?) AND (ODORLESS OR ODOURLESS OR DEODORI?)

          Items   File
          -----   ----
             22     5: Biosis Previews(R)_1969-2006/Aug W3
              6    10: AGRICOLA_70-2006/Aug
             15    50: CAB Abstracts_1972-2006/Jul
             45    51: Food Sci.&Tech.Abs_1969-2006/Aug W4
             42    53: FOODLINE(R): Science_1972-2006/Aug 29
             11    79: Foods Adlibra(TM)_1974-2002/Apr
             29    94: JICST-EPlus_1985-2006/May W3
              8   144: Pascal_1973-2006/Aug W1
              4   203: AGRIS_1974-2006/Apr
            112   399: CA SEARCH(R)_1967-2006/UD=14510

   10 files have one or more items; file list includes 15 files.

? RANK FILES

Your last SELECT statement was:
   S (GARLIC OR ONION?) AND (ODORLESS OR ODOURLESS OR DEODORI?)

Ref        Items   File
---        -----   ----
N1           112   399: CA SEARCH(R)_1967-2006/UD=14510
N2            45    51: Food Sci.&Tech.Abs_1969-2006/Aug W4
N3            42    53: FOODLINE(R): Science_1972-2006/Aug 29
N4            29    94: JICST-EPlus_1985-2006/May W3
N5            22     5: Biosis Previews(R)_1969-2006/Aug W3
N6            15    50: CAB Abstracts_1972-2006/Jul
N7            11    79: Foods Adlibra(TM)_1974-2002/Apr
N8             8   144: Pascal_1973-2006/Aug W1
N9             6    10: AGRICOLA_70-2006/Aug
N10            4   203: AGRIS_1974-2006/Apr

   10 files have one or more items; file list includes 15 files.
```

DIALINDEX Sample Search—2

```
File 411:DIALINDEX(R)

DIALINDEX(R)
   (c) 2005 The Dialog Corporation

*** DIALINDEX search results display in an abbreviated ***
*** format unless you enter the SET DETAIL ON command. ***

? SF ALLSCIENCE
   You have 233 files in your file list.
   (To see banners, use SHOW FILES command)

? S BUCKYBALL? ?

Your SELECT statement is:
   S BUCKYBALL? ?

          Items   File
          -----   ----
            125     2: INSPEC_1969-2005/Jul W3
              4     5: Biosis Previews(R)_1969-2005/Jul W4
             21     6: NTIS_1964-2005/Jul W3
             52     8: Ei Compendex(R)_1970-2005/Jul W3
             38     9: Business & Industry(R)_Jul/1994-2005/Jul 26
              2    10: AGRICOLA_70-2005/Jul
             79    15: ABI/Inform(R)_1971-2005/Jul 27
            263    16: Gale Group PROMT(R)_1990-2005/Jul 26
             78    18: Gale Group F&S Index(R)_1988-2005/Jul 26
            135    20: Dialog Global Reporter_1997-2005/Jul 27
              3    31: World Surface Coatings Abs_1976-2005/Jul
            234    34: SciSearch(R) Cited Ref Sci_1990-2005/Jul W3
             16    35: Dissertation Abs Online_1861-2005/Jun
             10    36: MetalBase_1965-20050725
              3    42: Pharmaceuticl News Idx_1974-2005/Jul W3
            223    47: Gale Group Magazine DB(TM)_1959-2005/Jul 27
              1    49: PAIS Int._1976-2005/Jun
              1    50: CAB Abstracts_1972-2005/Jun
              1    58: GeoArchive_1974-2005/May
             33    62: SPIN(R)_1975-2005/May W3
             14    65: Inside Conferences_1993-2005/Jul W4
                      .
                      . files omitted
                      .
            225   636: Gale Group Newsletter DB(TM)_1987-2005/Jul 26
              1   637: Journal of Commerce_1986-2005/Jul 27
             21   649: Gale Group Newswire ASAP(TM)_2005/Jul 15
            373   654: US Pat.Full._1976-2005/Jul 14
             24   660: Federal News Service_1991-2002/Jul 02
              1   674: Computer News Fulltext_1989-2005/Jul W3
              1   761: Datamonitor Market Res._1992-2005/Jul
            126   764: BCC Market Research_1989-2005/Jul
              1   766: (R)Kalorama Info Market Res._1993-2000/Aug
              1   768: EIU Market Research_2005/Apr 21
              9   813: PR Newswire_1987-1999/Apr 30

   94 files have one or more items; file list includes 233 files.

? RANK FILES

Your last SELECT statement was:
   S BUCKYBALL? ?

Ref        Items   File
---        -----   ----
N1           373   654: US Pat.Full._1976-2005/Jul 14
N2           355   440: Current Contents Search(R)_1990-2005/Jul 27
N3           263    16: Gale Group PROMT(R)_1990-2005/Jul 26
N4           234    34: SciSearch(R) Cited Ref Sci_1990-2005/Jul W3
N5           225   636: Gale Group Newsletter DB(TM)_1987-2005/Jul 26
N6           223    47: Gale Group Magazine DB(TM)_1959-2005/Jul 27
N7           216   148: Gale Group Trade & Industry DB_1976-2005/Jul 27
N8           189   349: PCT FULLTEXT_1979-2005/UB=20050721,UT=20050714
N9           158    98: General Sci Abs/Full-Text_1984-2004/Dec
N10          135    20: Dialog Global Reporter_1997-2005/Jul 27

   94 files have one or more items; file list includes 233 files.


       - Enter P or PAGE for more -
? P
```

```
Your last SELECT statement was:
   S BUCKYBALL? ?

Ref        Items   File
---        -----   ----
N11          126   764: BCC Market Research_1989-2005/Jul
N12          125     2: INSPEC_1969-2005/Jul W3
N13           95   399: CA SEARCH(R)_1967-2005/UD=14305
N14           85    99: Wilson Appl. Sci & Tech Abs_1983-2005/Jun
N15           79    15: ABI/Inform(R)_1971-2005/Jul 27
N16           78    18: Gale Group F&S Index(R)_1988-2005/Jul 26
N17           71   103: Energy SciTec_1974-2005/Jul B1
N18           59   144: Pascal_1973-2005/Jul W3
N19           57   211: Gale Group Newsearch(TM)_2005/Jul 27
N20           52     8: Ei Compendex(R)_1970-2005/Jul W3

   94 files have one or more items; file list includes 233 files.

        – Enter P or PAGE for more –
```

► Shortcuts for DIALINDEX Searches

TIP   *SELECT Command*

You can specify a new group of databases with the SET FILES command, and then process the same SELECT command in the new files by simply entering a single S (for SELECT). This makes it easy to run the same search in additional files without re-entering the search terms. The S always applies to the search statement most recently entered.

Saving the Strategy

Enter SAVE TEMP to save the strategy, and then BEGIN in the files you want. Enter the EXS command to execute the strategy.

Using the Strategy in Files with 'Hits'

If you want to search all of the databases that produced hits in DIALINDEX, enter SAVE TEMP, and then BEGIN HITS;EXS.

Using the N Numbers After a RANK FILES command

You can use the N numbers that display after a RANK FILES command to specify the databases you want in the next BEGIN command. For example, BEGIN N1-N2 would be equivalent to BEGIN 654,440.

Dialog Finder Files

When your topic requires that you search on a specific company name, product name, or journal name, you can take advantage of the Dialog Finder files. Using a Finder file, you can quickly determine which database(s) contain records on the company, journal, or product you are seeking. The type of information included in each database is described in the Finder record. Three Finder files are currently available:

File 413	Dialog Product Code Finder™
File 414	Dialog Journal Name Finder™
File 416	Dialog Company Name Finder™

To learn more about the Finder files, see *Using Controlled Vocabulary* on the GEP Web site.

Searching Multiple Databases

One of the most important features of Dialog is the access it provides to many different databases and the capacity to use the same search strategy for several files. Searching in multiple databases, however, can entail problems that are not encountered or are not obvious when working with a single file. Because the source data come from many different information providers, files vary in content, record format, and searchable fields.

In planning a search to be run in multiple databases, keep in mind the following:

- Codes, limits, prefixes, or suffixes that work in one file may not work in another file because the data elements or their equivalents are not present or differ. Be prepared for such differences by reviewing the Bluesheets and including alternative codes or forms of entry.

- Use proximity operators, particularly with descriptors, and consider qualifying to descriptor (/DE) and title (/TI). Bound descriptors are not likely to be the same across databases.

- Use limit suffixes to modify the search after it is run on the basic concepts.

- If you SELECT from an EXPAND display while constructing your strategy, be sure to use the OR and range (colon) operators so that the E numbers are saved as actual search terms (see Section 3).

Using the Dialog OneSearch feature, you can search multiple files simultaneously. Using the SearchSave feature, you can save a search strategy and run it in another file or use it again at a later date in the same file to retrieve the latest records. Dialog customers can get automatic updates on a search topic using the Dialog Alert service described in *Customizing Dialog* on the GEP Web site.

Searching Multiple Files with Dialog OneSearch

Every time you enter a BEGIN command that accesses more than one database, you are using the Dialog OneSearch capability. You can enter any combination of file numbers and DIALINDEX/OneSearch categories, up to 60 files. The DIALINDEX SuperCategories may not be used. The categories are listed in the *Dialog Database Catalog*, on the File 411 Bluesheet, and on the Dialog Web site.

Newspapers Fulltext (U.S. Pennsylvania) PAPERSPA

633	Philadelphia Inquirer
718	Pittsburgh Post-Gazette
731	Philadelphia Daily News
738	(Allentown) The Morning Call

To search in all of the fulltext Pennsylvania newspaper files, enter:

```
?b paperspa

[File 633] Phil.Inquirer 1983-2006/Aug 02
(c) 2006 Philadelphia Newspapers Inc. All rights reserved.
[File 718] Pittsburgh Post-Gazette Jun 1990-2006/Aug 03
(c) 2006 PG Publishing. All rights reserved.
[File 731] Philad.Dly.News 1983- 2006/Aug 02
(c) 2006 Philadelphia Newspapers Inc. All rights reserved.
[File 738] (Allentown) The Morning Call 1990-2006/Aug 02
(c) 2006 Morning Call. All rights reserved.

     Set  Items  Description
     ---  -----  -----------
?
```

While in OneSearch, your search looks exactly like a search run in a single file. When you SELECT a term, you see the total number of records retrieved for all of the files.

```
? b 7,1

File 7] Social SciSearch(R) 1972-2006/Jul W5
(c) 2006 The Thomson Corp. All rights reserved.
[File 1] ERIC 1966-2006/June
(c) format only 2006 Dialog. All rights reserved.

? s (autistic or autism) and environment?

        5797   AUTISTIC
       10454   AUTISM
      227362   ENVIRONMENT?
S1        717   S (AUTISTIC OR AUTISM) AND ENVIRONMENT?
```

Similarly, when you EXPAND a term, you receive an E-numbered list of terms showing a merged index for all the files. Be sure to look carefully to note differences in how phrase-indexed fields are handled and how abbreviations differ. For example, journal titles may be identical in the two files, but author name abbreviations may be different. Be sure to use the PAGE command or scroll down the EXPAND display to see *all* possible terms.

```
?e jn=journal of autism

Ref   Items  Index-term
E1        7  JN=JOURNAL OF AUSTISM AND CHILDHOOD SCHIZOPHRENIA        ← note misspelling of title
E2       11  JN=JOURNAL OF AUSTRALIAN RESEARCH IN EARLY CHILDH
E3        0  JN=JOURNAL OF AUTISM
E4      512  JN=JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA
E5     2807  JN=JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS
E6        2  JN=JOURNAL OF AUTOIMMUNITY
E7       15  JN=JOURNAL OF AUTOMATIC CHEMISTRY
E8        1  JN=JOURNAL OF AUTONOMIC PHARMACOLOGY
E9        2  JN=JOURNAL OF AVIAN BIOLOGY
E10       1  JN=JOURNAL OF AVIAN MEDICINE AND SURGERY
E11      22  JN=JOURNAL OF BACK AND MUSCULOSKELETAL REHABILITA
E12      18  JN=JOURNAL OF BACTERIOLOGY

        Enter P or PAGE for more

?e au=karnes f

Ref   Items  Index-term
E1        2  AU=KARNES EL
E2        2  AU=KARNES EW
E3        2 *AU=KARNES F
E4       75  AU=KARNES FA
E5        1  AU=KARNES HL
E6        3  AU=KARNES J
E7        1  AU=KARNES JL
E8        2  AU=KARNES LJ
E9        2  AU=KARNES LR
E10      13  AU=KARNES MB
E11       1  AU=KARNES MR
E12       1  AU=KARNES ST

        Enter P or PAGE for more
?p

Ref   Items  Index-term
E13       1  AU=KARNES T
E14       1  AU=KARNES TI
E15      14  AU=KARNES TL
E16       4  AU=KARNES WE
E17       1  AU=KARNES, DIANNE E.
E18       1  AU=KARNES, ELIZABETH L.
E19       5  AU=KARNES, FRANCES
E20      94  AU=KARNES, FRANCES A.
E21       1  AU=KARNES, FRANCES A., ED,
E22       3  AU=KARNES, JAMES B.
E23       1  AU=KARNES, JUDITH A.
E24       1  AU=KARNES, LUCIA R.

        Enter P or PAGE for more

?s e3:e4 or e19:e21
```

You can use TYPE, DISPLAY, and PRINT commands while in OneSearch to display search results. A notation is included in the header of each record to show which file it came from. Records display in this order: all records from the first file in the BEGIN command display in reverse accession number order, then all records from the second file display in reverse accession number order, etc.

```
?t s1/6/1-10

  1/6/1 (Item 1 from file: 7)

04394631   Genuine Article#: 063QR   Number of References: 73
Title: The CHARGE study: An epidemiologic investigation of genetic and environmental factors
contributing to autism  (ABSTRACT AVAILABLE)
2006

1/6/2 (Item 2 from file: 7)

04393975   Genuine Article#: 064DG   Number of References: 70
Title: Bidirectional effects of expressed emotion and behavior problems and symptoms in
adolescents and adults with autism  (ABSTRACT AVAILABLE)
2006
                                   .
                                   .  lines deleted
                                   .
1/6/8 (Item 8 from file: 7)

04372068   Genuine Article#: 046CE   Number of References: 30
Title: Genetic heterogeneity between the three components of the autism spectrum: A twin study
 (ABSTRACT AVAILABLE)
2006

1/6/9 (Item 9 from file: 7)

04371252   Genuine Article#: 047BO   Number of References: 29
Title: Does pervasive developmental disorder protect children and adolescents against drug and
alcohol use?
(ABSTRACT AVAILABLE)
2006

1/6/10 (Item 10 from file: 7)

04369639   Genuine Article#: 045KZ   Number of References: 32
Title: Parental perceptions and use of complementary and alternative medicine practices for
children with autistic spectrum disorders in private practice  (ABSTRACT AVAILABLE)
2006
```

If you want to order the records so that records from all files are sorted together, you can do so. However, you can sort only on PD or PY when using OneSearch.

Searching multiple files simultaneously is a practical way to save online search time and there are several options you can use to make it even more cost effective.

Listing Record Counts Separately

You can request that the record counts for your OneSearch sets be listed separately using the DISPLAY SETS FROM EACH command. This may be useful if you need to determine how many records are retrieved from each file. Notice the difference in the display in the two different DISPLAY SETS commands:

```
?ds

Set    Items   Description
S1       717   S (AUTISTIC OR AUTISM) AND ENVIRONMENT?

?ds from each

Set   File   Items   Description
        1     362
        7     355
S1            717   S (AUTISTIC OR AUTISM) AND ENVIRONMENT?
```

If you prefer, you can have this detailed form of record counts maintained during a search session so that you see individual file record counts each time you SELECT, EXPAND, or DISPLAY SETS. Enter the SET DETAIL ON command, which shows the record counts for each file throughout your search process:

```
?set detail on

DETAIL set on

?s (autism or autistic) and environment?

File 1:ERIC 1966-2006/June
         1646   AUTISTIC
         3373   AUTISM
       118105   ENVIRONMENT?
          362   (AUTISTIC OR AUTISM) AND ENVIRONMENT?

File 7:Social SciSearch(R) 1972-2006/Jul W5
         4151   AUTISTIC
         7081   AUTISM
       109257   ENVIRONMENT?
          355   (AUTISTIC OR AUTISM) AND ENVIRONMENT?

Total:
         5797   AUTISTIC
        10454   AUTISM
       227362   ENVIRONMENT?
S2        717   S (AUTISTIC OR AUTISM) AND ENVIRONMENT??

?e jn=journal of autism

Ref File   Items   Total   Index-term
E1    1              7    JN=JOURNAL OF AUSTISM AND CHILDHOOD SCHIZOPHRENIA
E2    1             11    JN=JOURNAL OF AUSTRALIAN RESEARCH IN EARLY CHILDH
      1      0           JN=JOURNAL OF AUTISM
      7      0           JN=JOURNAL OF AUTISM
E3                   0    JN=JOURNAL OF AUTISM
      1    199           JN=JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA
      7    313           JN=JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA
E4                 512    JN=JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA
      1    983           JN=JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS
      7   1824           JN=JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS
E5                2807    JN=JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS
E6    7              2    JN=JOURNAL OF AUTOIMMUNITY
E7    7             15    JN=JOURNAL OF AUTOMATIC CHEMISTRY
E8    7              1    JN=JOURNAL OF AUTONOMIC PHARMACOLOGY
E9    7              2    JN=JOURNAL OF AVIAN BIOLOGY
E10   7              1    JN=JOURNAL OF AVIAN MEDICINE AND SURGERY
E11   7             22    JN=JOURNAL OF BACK AND MUSCULOSKELETAL REHABILITA
E12   7             18    JN=JOURNAL OF BACTERIOLOGY

         Enter P or PAGE for more
```

Notice that when you use the EXPAND command, a *File* column is added to the list and the *Items* column shows file-specific counts. A *Total* column is also added to show the record count from all files. E numbers are assigned only to the *Total* items. If you want to SELECT an E number and restrict to an individual file, use the FROM option. For example, to SELECT the records included on line E4 from File 7, enter:

```
?s e4 from 7

  File 7:Social SciSearch(R) 1972-2006/Jul W5

            313   JN="JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA" FROM 7

TOTAL: FILES 7
     S3      313   JN="JOURNAL OF AUTISM AND CHILDHOOD SCHIZOPHRENIA" FROM 7
```

SET DETAIL ON remains in effect until you LOGOFF or enter SET DETAIL OFF.

Displaying Records from Individual Databases

If you want to display records from one or more of the files within the group of files you are searching, you can use the FROM option. FROM is entered as part of a TYPE command (or DISPLAY or PRINT) and is always the last parameter entered. For example, to type the first record from each file in set S1 using display codes TI and PY, enter:

```
? t s1/6/1 from each

 1/6/1 (Item 1 from file: 1)

01213205 ERIC Number: ED491806
Young Children with Disabilities in Natural Environments: Methods and Procedures
2006 (20060000)

1/6/363 (Item 1 from file: 7)

04394631    Genuine Article#: 063QR    Number of References: 73
Title: The CHARGE study: An epidemiologic investigation of genetic and environmental factors
contributing to autism  ( ABSTRACT AVAILABLE )
2006
```

You can also use FROM followed by a file number or several file numbers to see a subset of records from a set:

```
?t s1/6/1-3 from 1

01213205 ERIC Number: ED491806
Young Children with Disabilities in Natural Environments: Methods and Procedures
2006 (20060000)

1/6/2 (Item 2 from file: 1)

01212866 ERIC Number: EJ736517
Using Acceptance and Commitment Training in the Support of Parents of Children Diagnosed with
Autism
March 14, 2006 (20060314)

1/6/3 (Item 3 from file: 1)

01212558 ERIC Number: EJ731706
Approach-Avoidance and Happiness Indicators in Natural Environments: A Preliminary Analysis of
the Stimulus Preference Coding System
2005 (20050000)
```

Details on the FROM Option

The FROM option can be used with many different commands while in OneSearch, as summarized below. For a few commands, using FROM is required while in OneSearch.

OneSearch FROM Options with Different Dialog Commands

Command	Example
DISPLAY SETS	DISPLAY SETS FROM 16
	DS FROM EACH
DISPLAY SETS <set numbers>	DISPLAY SETS S2 FROM 16
	DS S2-S3 FROM EACH
EXPAND	EXPAND JN=JOURNAL OF AUTISTIC FROM 7
PRINT	PRINT S2/9/ALL FROM 29
	PR S2/9/ALL ADDR AD023 FROM 16,675
SELECT	S AUTISTIC/TI FROM 7
TYPE	TYPE S4/6/1-3 FROM EACH
	T S4/6/1-10 FROM 2
FROM is required in the following commands and must be followed by a single file number:	
ORDER	ORDER S1/3 SRC1 FROM 7
SAVE ALERT	SAVE ALERT FROM 7
SORT on fields other than PY or PD	SORT S1/ALL/AU FROM 675

Removing Duplicate Records

One of the most powerful features built into Dialog OneSearch is the capability for removing duplicate records in bibliographic and fulltext files. An algorithm examines key portions of records and determines if they are citations to identical publications. Duplication typically occurs across files that index some of the same journals, conference proceedings, or similar documents. The process of removing duplicates can reduce your search costs by reducing per-record download charges.

To remove duplicates and create a new set of the unique records, use the RD (REMOVE DUPLICATES) command followed by a set number. Records are chosen for retention based on the order in which the files were entered in the BEGIN command. (You can change the file order using the SET FILES command—see the command summary in this section.) Unless you want to use the SET FILES command, though, be careful of the order you BEGIN files. If you are searching in both a bibliographic and a fulltext file, put the fulltext file first in your BEGIN command. This way, a REMOVE DUPLICATES command will find a duplicated fulltext record first and a bibliographic citation next and drop the latter rather than the former.

In the sample search below, the search results are reduced from 61 records to 35 unique records when duplicates are removed using the RD command.

```
?b 89,58,292,144

[File 89] GeoRef 1785-2006/Jul B2
(c) 2006 American Geological Institute. All rights reserved.
[File 58] GeoArchive 1974-2006/May
(c) 2006 Geosystems. All rights reserved.
[File 292] GEOBASE(TM) 1980-2006/Jul W4
(c) 2006 Elsevier Science Ltd. All rights reserved.
[File 144] Pascal 1973-2006/Jul W2
(c) 2006 INIST/CNRS. All rights reserved.

? s parkfield/2005:2006

         1528   PARKFIELD
       830582   PY=2005 : PY=2006
S1         97   S PARKFIELD/2005:2006

?  rd s3

S2         83   RD S2  (UNIQUE ITEMS)
```

Several other commands are available to help you with duplicate detection. A command summary is included below.

Duplicate Detection Command Summary

Command and Abbreviation	Example	Function	Applications
IDENTIFY DUPLICATES ID	ID S5	Groups duplicate records together	Use ID to create a sorted set of records in which duplicate citations are grouped together. Use ID to identify duplicate citations, while retaining all of the records in the set. Use ID to create a set that is approximately sorted by title across all the files.
REMOVE DUPLICATES RD	RD S3	Removes duplicate records	Use RD to create a set of unique records in which only one citation from each set of duplicate citations is retained. Records are chosen for retention based on the order in which files were entered in the most recent BEGIN or SET FILES command. Use RD followed by ID to create a set of unique records that are approximately sorted by title.
IDENTIFY DUPLICATES ONLY IDO	IDO S4	Creates a set that contains only the duplicate records	Use IDO to create a set of the records that have been identified as duplicates. Use IDO to check the results of the duplicate detection process.

Note: Sets created by ID and IDO *cannot* be used in a subsequent SELECT command. Sets created by RD *can* be used in a SELECT command.

Two additional commands are available to manage the files being searched:

Command Abbreviation	Example	Applications
SHOW FILES	SHOW FILES	Use SHOW FILES to review the current order of files.
SET FILES <file number(s) >	SET FILES 72,154	Use SET FILES to change the current order of files. If SET FILES is not used, the files are in the order specified in the BEGIN command. Any files not included in the SET FILES command remain in the order of the BEGIN command, coming after the other files.