

---

**Section 3**

---



## **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<sup>®</sup> 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<sup>SM</sup> 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 **T11**, the second word, **T12**, etc. The position of

### Section 3

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	<u>TI1</u> <u>TI2</u> <u>TI3</u> <u>TI4</u> <u>TI5</u> <u>TI6</u>
Brodbelt, Samuel	<u>AU</u>
Clearing House, v57 n2 p72-75 Oct 1983	<u>JN</u> <u>PY</u>
Language: English	<u>LA</u>
Document Type: Opinion Paper	<u>DT</u>
Develops a sociological and philosophical perspective on	<u>AB1</u> <u>AB2</u> <u>AB3</u> <u>AB5</u> <u>AB6</u> <u>AB7</u>
lifelong education as a goal that is integral to	<u>AB8</u> <u>AB9</u> <u>AB10</u> <u>AB11</u> <u>AB12</u> <u>AB13</u> <u>AB14</u> <u>AB15</u>
an advanced, technologically directed, democratic society.	<u>AB18</u> <u>AB19</u> <u>AB20</u> <u>AB21</u> <u>AB22</u>
Descriptors: Adult Education; *Educational Philosophy; Educational	<u>DE1</u> <u>DE2</u> <u>DE3</u> <u>DE4</u> <u>DE5</u>
Trends; *Futures (of Society); *Lifelong Learning; Technology	<u>DE6</u> <u>DE7</u> <u>DE9</u> <u>DE10</u> <u>DE11</u> <u>DE12</u>
289519	<Accession Number>
Servants of Socrates in the land of Oz	<u>TI1</u> <u>TI3</u> <u>TI4</u> <u>TI6</u> <u>TI8</u>
Blackburn, Harold	<u>AU</u>
Clearing House, V57 n2 p69-71 Oct 1983	<u>JN</u> <u>PY</u>
Language: English	<u>LA</u>
Document Type: Opinion Paper	<u>DT</u>
Discusses the role of the microcomputer in the classroom.	<u>AB1</u> <u>AB3</u> <u>AB6</u> <u>AB7</u> <u>AB9</u>
Descriptors: Educational Improvement; Educational Quality;	<u>DE1</u> <u>DE2</u> <u>DE3</u> <u>DE4</u>
Educational Technology; Futures (of Society); Learning	<u>DE5</u> <u>DE6</u> <u>DE7</u> <u>DE9</u> <u>DE10</u>
Theories; Microcomputers; Technological Advancement.	<u>DE11</u> <u>DE12</u> <u>DE13</u> <u>DE14</u>

### 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.

<b>Term</b>	<b>Address</b>
a	289520 AB2 289520 AB11

<i>Term</i>	<i>Address</i>
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

<b>Term</b>	<b>Address</b>
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.

<b>Term</b>	<b>Address</b>
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.

### Section 3

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.

<b>?s moscow and union?</b>		
	428686	MOSCOW
	751476	UNION?
S10	12368	MOSCOW AND UNION?
<b>?s s10/ti</b>		
S11	1250	S10/TI
<b>?s (moscow and union?)/ti</b>		
	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<sup>□</sup> 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 □

<b>Prefix</b>	<b>Field Name</b>	<b>Prefix</b>	<b>Field Name</b>
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 must 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:

## Section 3

?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=BEHAVIORIAL 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=BEHAVIORIAL SCIENCE
S1 22 JN='BEHAV SCI' OR JN='BEHAVIORAL SCIENCE' OR JN='BEHAVIORIAL 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

<i>Basic Index</i>		
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

<b><i>Additional Indexes</i></b>		
<b>Search Suffix</b>	<b>Field Name</b>	<b>Indexing</b>
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.

### Section 3

#### ?help field 1

FIELD1 :ERIC

#### SEARCH OPTIONS

##### BASIC INDEX

SEARCH DISPLAY	PREFIX CODE	FIELD NAME/ 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	PREFIX CODE	FIELD NAME/ 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.

