Lesson 6: Databases and Web Search Engines

Objectives

By the end of this lesson, you will be able to:

✯ 1.11.1: Define and identify Boolean operators and their relationship to search engines.

✯ 1.11.2: Conduct searches using operators common to search engines (e.g., quotation marks, minus sign).

✯ 1.11.3: Identify the importance of keywords and conduct keyword searches.

✯ 1.11.4: Distinguish among search sites, information portals, and pages automatically provided when an error occurs.

✯ 1.11.5: Evaluate research for soft and hard-copy resources, and verify factuality.

✯ 1.11.6: Identify meta search engines, as well as search engine ranking techniques.

✯ 1.11.7: Explain how search engine sites use spiders to enhance search results.

✯ 1.11.8: Cite references to information obtained from Internet-based searches using accepted citation standards (e.g., MLA).

✯ 1.11.9: Search additional Internet resources to fulfill job-role duties.

✯ 1.11.10: Distinguish between authoritative peer-reviewed and open peer-reviewed sources.

✯ 1.11.11: Provide examples of both primary and secondary resources, as well as indexes.

✯ 1.11.12: Explain key differences between popular and scholarly journals.

✯ 1.16.1: Distinguish between relational and non-relational databases.

✯ 1.16.2: Define common relational database concepts, including table, row, record, column, field, data value, join.

✯ 1.16.3: Identify relationships between tables, including one-to-one, one-to-many, many-to-many.

✯ 1.16.4: Explain the purpose of SQL in querying databases.

✯ 1.16.5: Identify common database vendors.

Pre-Assessment Questions

1. How is a many-to-many database table relationship established?

   _____

   _____

   _____

   _____

2. What distinguishes an information portal from a search engine?

   _____

   _____

   _____

   _____

3. Using the Boolean OR operator with keywords in a search means that the search results:

   a. may include a particular keyword.
   b. must include a particular keyword.
   c. may include at least one of the keywords.
   d. must include at least one of the keywords.

Introduction to Data Searching Tools

With so much data on the Internet, finding the exact information you need can be difficult and frustrating. Many tools are available for navigating the Internet, however. All of these tools involve the use of databases. Individuals and search engine companies use databases in sophisticated ways in order to gather and analyze information. To help you learn more about databases, this lesson will focus on the various Web search engines and the search techniques you can employ to find the information you are seeking.

A Web search engine is a large database containing information about Web pages that have been registered with it. To understand search engines, you must understand basic database concepts and the manner in which data is organized and accessed.

CIW Online Resources – Movie Clips

Visit CIW Online at http://education.Certification-Partners.com/CIW to watch a movie clip about this topic.

Lesson 6: Databases and Web Search Engines

Overview of Databases

This section will provide an overview of database components, relational databases, table keys, table relationships, database queries and common database vendors.

OBJECTIVE
1.16.2: Relational
database concepts

database
A collection of data
that can be sorted
and searched using
search algorithms.

table
A collection of data
about a limited
topic, organized
into rows and
columns in a
database.

field
A category of
information in a
database table.

record
A collection of
information in a
database table
consisting of one or
more related fields
about a specific
entity, such as a
person, product or
event.

Database components

A **database** is a collection of data that can be sorted and searched using search algorithms. Databases are useful for storing, manipulating, managing and querying large amounts of data.

Database components are referred to as database objects and are each stored separately within the database. In a database structure, data values are stored in **tables**. A table can contain data such as personnel information, product inventories, customer lists, sales figures and so forth. A database can consist of multiple tables, depending on the complexity of the data being stored.

A **field** is a category of information in a table. Each column in a table represents a field. A **record** is a collection of information consisting of one or more related fields about a specific entity, such as a person, product or event. Each row in a table represents a record.

Figure 6-1 shows a sample database table containing technical specifications for three aircraft.

Fields

Aircraft	Length	Height	Wingspan	Weight	Cruising Speed	Range	Engine	Capacity
PA 700	175 ft.	45 ft.	140 ft.	150,000 lbs	540 mph	5,500 miles	LCAE-450-A	210
PA 800	205 ft.	50 ft.	170 ft.	225,000 lbs	550 mph	7,000 miles	KE-7070-G	300
PA 900	240 ft.	60 ft.	205 ft.	375,000 lbs	560 mph	8,500 miles	BT90-850-D	420

Record: 14 ◄ [1] ►I ►I ►* of 3

Records

Figure 6-1: Sample database table

CIW Online Resources – Online Exercise

Visit CIW Online at *http://education.Certification-Partners.com/CIW* to complete an interactive exercise that will reinforce what you have learned about this topic.

Exercise 6-1: Database components

relational database
A database that contains multiple tables related through common fields.

common field
A field contained in two or more database tables that forms a connection between the tables.

relationship
A connection between two or more database tables that is based on a field that the tables have in common.

Relational databases

A **relational database** contains multiple tables of information that are related through **common fields**. By contrast, a non-relational database consists of multiple tables of information that are not related through common fields. When multiple tables have a field in common, they can be joined to form a **relationship** to access the table data as needed.

To understand the relational database concept, consider a simple database designed for an aircraft manufacturer with three new aircraft that will enter production soon. The manufacturer wants to keep track of aircraft orders, as well as technical information about each aircraft. Figure 6-2 is an example of a relational database.

Figure 6-2: Sample relational database

The Customers table contains information only about the aircraft ordered by each airline; it does not contain information about the aircraft. The Aircraft Technical Specifications table contains information only about the aircraft. The only field the two tables have in common is the Aircraft field; the tables are related through this common field.

Relating the two tables can save time and eliminate data duplication. For example, storing customer and technical specification data in separate tables and creating a relationship between the tables through the common field Aircraft eliminates the need to enter the aircraft information every time you want information about the planes a customer ordered.

Figure 6-3 shows the aircraft ordered by each customer, the quantity ordered and the order dates.

Figure 6-3: Related tables

The design of a relational database is usually determined by four factors:

- The purpose of the database

- The number of tables needed to store information without duplicating data

- The fields needed in each table

- The fields common to more than one table

Whether a database consists of a single table or multiple related tables depends on the data requirements of the database users.

Database table keys

primary key
A field containing a value that uniquely identifies each record in a database table.

foreign key
A field in a related database table that refers to the primary key in the primary table.

A table relationship is established by matching data in key fields between two database tables. Usually, the fields have the same name in both tables; however, the names can be different. In any given relationship, the table in which the common field is the **primary key** is called the primary table. The other table in the relationship is called the related table. The common field in the related table is called the **foreign key**. A database table can have only one primary key.

The previous figure depicting a sample relational database shows the common field, Aircraft. The Aircraft field is the primary key in the Aircraft Technical Specifications table because the data contained in the field is unique; that is, no duplicate data can be contained in the field. The Aircraft field is the foreign key in the Customers table.

OBJECTIVE
1.16.3: Database table relationships

Relating database tables

Three types of relationships can be established between database tables:

- One-to-one

- One-to-many

- Many-to-many

One-to-one relationship

A **one-to-one relationship** is a database table relationship in which each record in Table A can have only one matching record in Table B, and vice versa. The relationship is established only if the common field is the primary key in both tables.

One-to-one table relationships are unusual because most data related in this way would be stored in the same table. However, a one-to-one relationship might be used to divide a large table, to isolate part of a table for security reasons, or to store information that applies to a subset of the main table. For example, several sales representatives in your organization might be part of a training program, and that information could be stored separately.

One-to-many relationship

A **one-to-many relationship** is a database table relationship in which a record in Table A can have multiple matching records in Table B, but a record in Table B has only one matching record in Table A. The relationship is established only if the common field is the primary key in Table A and the foreign key in Table B.

A one-to-many relationship is the most common type of table relationship. For example, a one-to-many relationship can exist between aircraft customers and the aircraft's technical specifications. Each aircraft can have only one set of specifications, but multiple customers can order the same type of aircraft.

Figure 6-4 shows two tables with a one-to-many relationship, joined by the Aircraft field. A record in the Aircraft Technical Specifications table can have several matching records in the Customers table, but a Customers record can have only one matching record in the Aircraft Technical Specifications table.

Figure 6-4: One-to-many database table relationship

Many-to-many relationship

A **many-to-many relationship** is a database table relationship in which one record in Table A can relate to many matching records in Table B, and vice versa. A many-to-many relationship between two tables is established by creating two one-to-many relationships with a third table, called a **junction table**. For example, a many-to-many relationship can exist between airlines and aircraft technical specifications. An airline can order several types of aircraft, and the technical specifications for each aircraft can be combined with the data for each airline.

Figure 6-5 shows a many-to-many relationship between the Aircraft Technical Specifications table and the Airlines table. The Customers table is the junction table that establishes the two required one-to-many relationships. In this example, a one-to-many relationship also exists between the Aircraft Technical Specifications table and the

---

one-to-one relationship
In databases, a relationship in which each record in Table A can have only one matching record in Table B, and vice versa.

one-to-many relationship
In databases, a relationship in which a record in Table A can have multiple matching records in Table B, but a record in Table B has only one matching record in Table A.

NOTE:
Notice the join line that links the common fields between the two tables in the figure. The 1 and the infinity symbol (∞) indicate the one-to-many relationship between the two tables. Also note that the primary key in the primary table appears in bold type.

many-to-many relationship
In databases, a relationship in which one record in Table A can relate to many matching records in Table B, and vice versa.

junction table
A database table containing foreign-key fields that refer to the primary-key fields from the primary tables in a many-to-many relationship.

Engines table, which provides information about the engines used by each type of aircraft.
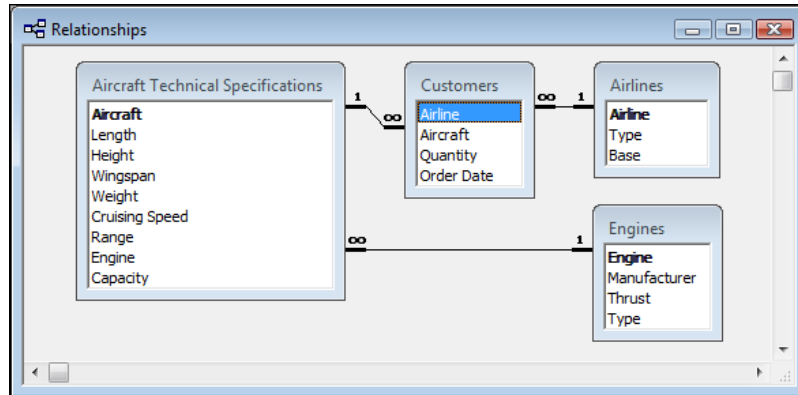
Figure 6-5: Many-to-many database table relationship

Querying databases using Structured Query Language (SQL)

Users input queries to selectively view and analyze data from a database. A **query** isolates and displays only the information you want to examine by specifying the fields and records you want to see. For example, you can use a query to display only the cruising speed, range and capacity of the aircraft that TransCon ordered.

There are three general types of queries:

- **Menu query** — You are offered a list of options in a menu from which to choose.

- **Query by example** — You specify the fields and data values to be used in a query.

- **Query language** — You use a specialized language called **Structured Query Language (SQL)** to retrieve and manipulate information in a database.

SQL (pronounced "sequel") was developed by IBM in 1974 and was originally called Structured English Query Language (SEQUEL). Oracle Corporation released the first commercial SQL application in 1979.

SQL has become the standard interactive and programming language for accessing information from and updating information in a database, particularly large relational databases. SQL command statements are used for interactively querying information from a database and accumulating data for reports.

Common database vendors

A **database management system (DBMS)** is a program used to store, access and manipulate database information. IT professionals commonly use a DBMS to work with corporate database information. Originally, corporate databases were stored on large mainframe computers. More recently, client/server databases have become more popular and are able to work with smaller networks. Common database vendors and their DBMS products (in parentheses) include the following:

- IBM (DB2)

- Computer Associates (Advantage)

- Oracle (Oracle Database)

- Microsoft (SQL Server, Microsoft Access)

- Sybase (Advantage Database Server [ADS], SQL Anywhere, Sybase IQ)

- MySQL (MySQL)

Databases and Internet servers

An Internet server can be connected to a database to provide information via the Web. In this case, the Web server provides HTML pages that act as the front end, which is another term for the interface that end users see. The database server then acts as the back end, which includes the database and the technologies that enable the front end to communicate with the database. In fact, most companies rely upon such connectivity. Following are examples of ways that an Internet server can use a database:

- **Company sales data** — A company can store sales and marketing data in a centralized database. Employees can then conduct searches of this data when formulating a marketing plan. This process is also called "mining" data. An example of this type of database is Goldmine (*www.goldmine.com*). Such databases can be accessed via a proprietary client, such as one provided by Goldmine, or via a Web server.

- **Banking** — Banks store user account information in databases. Users can then access the databases through the banks' Web servers. Deutsche Bank (*www.db.com*), Citibank (*www.citibank.com*) and Wells Fargo (*www.wellsfargo.com*) are examples of companies that use Web servers and database technology to provide services to their customers.

- **Web storefronts** — Small, medium and large businesses have Web-enabled databases that allow customers to buy and sell goods. Well-known Web storefronts include Amazon.com (*www.amazon.com*), Travelocity (*www.travelocity.com*) and FTD.com (*www.ftd.com*).

- **Web search engines** — Web search sites, such as Google (*www.google.com*) and Bing (*www.bing.com*), populate their databases with information about resources on the Internet, including the World Wide Web. Search engine companies strive to populate their databases with useful information so that people will find value in their services. In turn, they sell space to advertisers to generate revenue. Web search engines will be discussed in greater detail in the following sections.

Open Database Connectivity (ODBC)

The Open Database Connectivity (ODBC) standard was created by The Open Group (*www.opengroup.org*). ODBC was designed to enable an operating system to access databases from various vendors (for example, Oracle, Microsoft and IBM). Using ODBC, a server (for example, a Goldmine or Web server) can access any type of database. ODBC provides a standard database interface.

An alternative to ODBC is Java Database Connectivity (JDBC), which is similar to ODBC except that JDBC uses the Java language and thus is capable of running on any platform. Another alternative to ODBC is the UnixODBC Project (*www.unixodbc.org*), which aims to implement ODBC on UNIX systems.

**Common Gateway
Interface (CGI)**
A program that
processes data
submitted by the
user. Allows a Web
server to pass
control to a
software
application, based
on user request. The
application receives
and organizes data,
then returns it in a
consistent format.

Anytime you conduct a database search, the database may be using ODBC, as well as a **Common Gateway Interface (CGI)** application to connect the front end (e.g., the Web server) to the back end (e.g., the database). CGI is the method by which data is passed back and forth between the server and the application program.

Normally, when you click a hyperlink to access a Web page, the Web server retrieves the requested page. However, when you submit an online form, the Web server typically forwards the form information to an application program that processes the data. You

can think of CGI as the interface between the Web server and the program that collects or generates the form content, such as a database or other application program. CGI uses various languages and technologies, including Personal Hypertext Preprocessor (PHP), Perl and Microsoft Active Server Pages (ASP).

Introduction to Web Search Engines

OBJECTIVE
1.11.3: Keyword searching

search engine
A powerful software program that searches Internet databases for user-specified information.

keyword
A word that appears on a Web page and is used by search engines to identify relevant URLs. Some words, such as "the" or "and," are too common to be used as keywords.

Search engines are powerful software programs that make it easy to find information on the Internet. A search engine consists of a large database that contains information about Web pages that have been registered with the search engine. Most search engines allow you to browse general categories, search for specific topics or find detailed information based on narrowly defined criteria.

Using keywords, you can find information on the Internet about any subject you want to investigate. A search engine database typically contains information about the site such as the title, the URL, a short description of the site contents and keywords to help the search engine reference the page. When a Web site manager registers a Web page with a search engine, the URL and associated information are added to the search engine database.

Not every Web site is registered with every search engine. Some Web site managers register their sites only with particular search engines, neglecting others. Because of this selectivity and the large number of search engines, you may receive different results using the same keywords on different search engines.

OBJECTIVE
1.11.7: Search engine spiders

Some search engines are highly automated and initiate a robot or "spider" program that searches the Web and indexes Web sites. Search engines that use spider programs are sometimes called spider search engines. These programs gather information about Web sites and newsgroups as determined by certain parameters. In some cases, only the page title and URL are indexed. In others, partial or full-page text is also indexed from the site. In addition, spiders check the validity of the URL entered in the database to ensure accuracy.

Registering a Web Site with a Search Engine

OBJECTIVE
1.11.6: Meta search engines

Many popular search engines require that you register your Web site for it to be accessible to Web users. Each search engine has its own registration process.

Many search engines, such as Excite, Lycos and WebCrawler, require you to register by completing online forms and entering the URL of your site. The search engines then use the robot or spider programs to visit your site and search it for relevant source code embedded in the page, often in the <meta> tag.

NOTE:
Be sure that you understand the relationship between spider programs and <meta> tags. Search engines use spider programs to collect keywords to add them to their search databases. The keywords are identified in a Web page using the <meta> tag.

The <meta> tag

The <meta> tag is an HTML element used in a Web page to embed information for use by search engines. This information is placed near the top of the HTML document so search engines can quickly access it. The more descriptive the information included in the <meta> tag, the more likely the Web site is to rank high in search results. The <meta> tag includes information about a document, such as:

- Keywords to be used by some search engines.

- A description of the site's content for display in some search engines.

- An expiration date for the document.

- The document's author.

meta search engine
A search engine that scans Web pages for <meta> tag information.

Most search engines scan Web pages for <meta> tag information. These search engines are sometimes called **meta search engines**. The <meta> tag identifies data in the following format:

<meta name="keywords" content="*keyword1 keyword2 keyword3*"/>

The search engine spider program inspects pages that use the "keywords" value of the *name* attribute. The values specified by the *content* attribute are then ranked against keywords that users entered in the search engine.

You separate keywords in the <meta> tag with spaces or commas. Do not use keywords redundantly when creating <meta> tags for your site; multiple occurrences of the same word can cause the tag to be ignored.

Another useful attribute for search engines is the "description" value of the *name* attribute:

<meta name="description" content="This site provides aircraft technical data."/>

In the preceding tag, the *content* attribute value will appear in some search engines instead of the first few lines of the Web page. The description should be brief.

Spiders and manual registration

Spiders can also search for word occurrences, titles, heading text and images within Web pages. Unfortunately, each search engine spider gathers different information, so no method guarantees a top listing.

Most search engine sites offer a link, such as Add Your Site or Webmaster Tools, which you can access for more information about registering your site. You can visit each search engine's Web site and locate the registration link to submit your site's URL. Submitting your site to a search engine increases your chances for higher placement in the search results.

Not all search engines index the same number of sites. Some have fewer than a million site references, while others have hundreds of millions of references. When you search for topics using a Web search engine, you are never searching the entire Web, only an index or database of referenced sites.

OBJECTIVE
1.11.4: Search sites vs. information portals

NOTE:
View the Yahoo! portal. Notice that the portal organizes information much like the yellow pages of a telephone directory. Then, visit Google. Notice that Google does not organize data the way Yahoo! does. Google simply allows users to conduct searches.

Search engines vs. information portals

The term search engine is used loosely in the Web community to include both search engines and information portals. A search engine, such as Bing or Google, uses a robot or spider program to browse the Web (by following hyperlinks) and index the content that it finds. Therefore, a search engine will eventually find your page, assuming that a link to your site exists somewhere on the Web. However, it is impossible to determine how long it will take for your site to be discovered. Therefore, it is important to submit your site to the search engines. Submitting your site speeds the process and increases your chances for higher placement in search results, thus allowing you as the Web site manager some control over placement.

An information portal differs from a search engine in that it will only find Web sites based on manual submissions. Thus, if you do not manually submit your site to the information portal, it will never know about or index your site, and never list your site in search results. Some information portals, such as Yahoo!, visit your Web site after you register it to determine the content and its relevance to the submitted topic. After the review, you may receive an e-mail message requesting additional information so a Web site description can be created before the site is placed in the portal.

The advantage of information portals is that search results are more likely to contain high-quality content matches to each query. However, increasing the chances for a higher placement in search results is usually not within the control of the Web site manager.

Types of Web Searches

<div style="float:left">

OBJECTIVE
1.11.3: Keyword
searching

index
A catalog of the
contents of a
database. Each
entry identifies a
unique database
record.

</div>

Search engines consist of large databases that allow you to query **indexes**. A search index is an enormous catalog that a specific search engine site has compiled, consisting of millions of Web page references.

You can perform two types of searches with search engines:

- **Directory searches** — The engine displays a list of categories and subcategories that you can browse. At each level, the items are arranged from general to specific.

- **Keyword searches** — You enter one or more keywords into the engine, and the search engine finds specific topics by locating those keywords or phrases in registered Web pages.

You can access a search index using a search engine, or you can manually navigate through an index's structured directory using an information portal. Table 6-1 describes the types of Web searches that users can perform using a search engine, as well as the related indexing method that the search engine uses to enable the search type.

Table 6-1: Types of Web searches

User Search Type	Search Engine Indexing Method	Description
Directory search	**Static index/ site map**	Enables users to manually search information portals to locate indexed information. Some portals are presented as directory trees of all indexed pages. Many Web sites present site maps, which allow users to view the site contents without visiting each page. Examples include Yahoo! and most major Web sites.
Keyword search	**Keyword index**	Enables users to enter keywords in a search engine to query an index. The entered keywords are then compared to the keywords in the index. If the keywords match, the search engine will provide the indexed information. Keywords can consist of one word, several words or a phrase enclosed in quotation marks (e.g., search string). Examples include Yahoo!, Google, Excite, Bing, Dogpile and Lycos.
	Full-text index	Enables a user to enter any search string into the search engine. This type of search requires the full contents of every page to be indexed for results, rather than only the page title or URL. Google is an example of a full-text index search engine.

AltaVista, which provided keyword and full-text index searching, was one of the earliest and most popular Web search engines. Its owner, Yahoo, discontinued the AltaVista search engine in July 2013. Popular search engine alternatives include Bing and Dogpile.

Some search engines are indexed to allow either static directory searching or keyword searching, whereas other search engines offer both. Yahoo! is an example of a search engine site that offers both keyword and directory searching. Figure 6-6 shows the Yahoo! keyword search field.
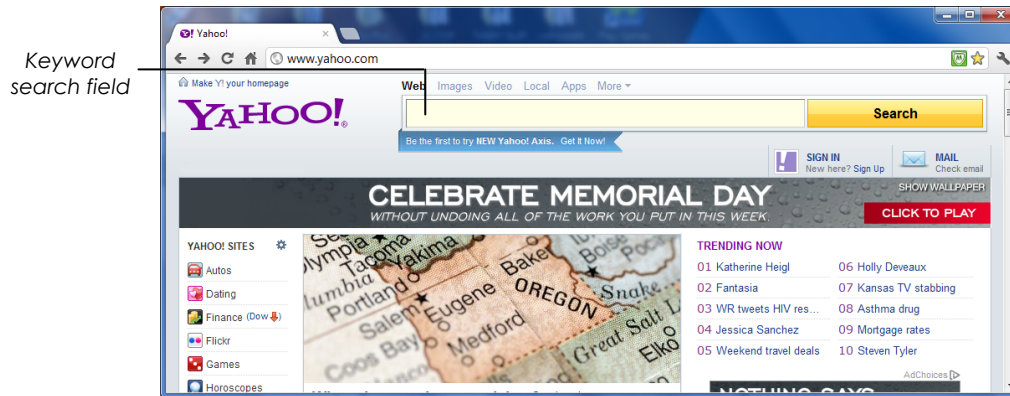
Figure 6-6: Yahoo! keyword search field

Figure 6-7 shows the Yahoo! static directory that appears at the bottom of the Yahoo! Groups page.
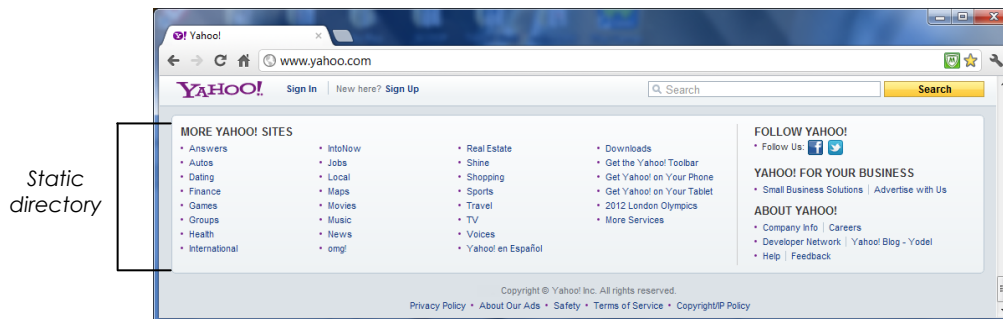
Figure 6-7: Yahoo! static directory

Which type of search is best?

Both directory searches and keyword searches are valuable. The search method you should use depends on the type of information you need. After browsing directories to find general categories, you can use keyword searches to narrow the results. For example, you could use directories to find a Web site about general project management principles, then use keywords to find specific information about the planning phase of a project.

Search engine variations

Search engines vary greatly. Some search only the Web, and others (such as Google) also search for articles in newsgroups. Some engines provide links to sites, and others specify URLs. Some engines may give brief descriptions of the site contents, and others may include paragraph summaries of the sites.

File types searched

Traditionally, the search engines from various sites (such as Google and Yahoo!) focused on indexing HTML pages. However, search engines now have the ability to search for and index the contents of additional file types, including the following:

- **ASCII** — files that are in plaintext format

- **Rich Text Format (RTF)** — files that contain simple formatting (bolding, italics and so forth)

- **Portable Document Format (PDF)** — files that can be transferred across platforms and retain formatting

- **PostScript** — files that can be printed to PostScript printers, or read by various applications such as GhostView (commonly found in Linux systems)

- **Proprietary file formats** — files made by specific vendors' applications, such as Microsoft Word or Excel

---

CIW Online Resources – Course Mastery

Visit CIW Online at *http://education.Certification-Partners.com/CIW* to take the Course Mastery review of this lesson or lesson segment.

IBA Lesson 6 - Part A

Basic Web Searching Techniques

NOTE:
Because so many search engines exist, you may get different results from different engines.

To conduct basic Web searches, you typically enter a keyword or a short phrase within quotation marks into a search field. Basic searches are useful when a simple, distinguishable topic is requested. In the following sections, you will perform basic Web searches with Yahoo! and Google. You will also be introduced to Bing, Excite and WebCrawler.

Yahoo!

NOTE:
According to the Yahoo! site (*http://docs.yahoo.com/info/misc/history.html*), Yahoo! is an acronym for Yet Another Hierarchical Officious Oracle. The founders of Yahoo! state that they chose the name because they liked the basic definition of a yahoo, which is an unsophisticated and rude person.

One of the oldest and most basic search engines is Yahoo! (*www.yahoo.com*). It was created in 1994 by two students at Stanford University as a way to track their personal interests on the Internet. They realized its usefulness, and soon customized their search utility into the hierarchical directory to the Internet that it has become. It was not intended to be a search engine, but rather an index providing several links relating to each topic. Originally, its database of links was very selective, including only carefully chosen and verified links. It is now one of the Internet's major search engines.

The Yahoo! interface provides hyperlinks to help narrow your searches, such as Finance, Games, Groups, HotJobs, Maps, Music, Sports, Travel, Yellow Pages and so forth. It also offers a hypertext listing by alphabetical categories, such as Business & Finance, Computers & Internet, Culture & Community, and Entertainment & Arts. Each of these topics contains subtopics. For example, under the larger topic of Computers & Internet, you will find hypertext links to Communications And Networking, Programming Languages, and Technical Support.

If you select Programming Languages, Yahoo! displays a more advanced search page that offers both a search field and further-categorized hypertext links.

Yahoo! searches titles, categories and text to find results that contain all of your keywords. Yahoo! is not case-sensitive. For example, the keywords "Project Management" would be read the same as "project management."

Yahoo! finds and retrieves three types of information for each search. The searches are based on:

- The alphabetized hypertext categories that match your keywords.

- Web sites that match your keywords.

- Yahoo! categories that list sites that match your keywords.

Yahoo! identifies its categories in bold type. Referenced sites are displayed in plain text. To navigate down the subject hierarchy, select a category, and the next lower level of the hierarchy will display.

In the following lab, you will perform a directory search. Suppose you are a newly hired help desk technician. You have been asked by your IT manager to join an IT association and begin regularly attending its meetings to expand your skill set. What method would you use to find an appropriate organization?

Lab 6-1: Performing a directory search using Yahoo!

In this lab, you will use Yahoo! to perform a directory search.

1. Go to **www.yahoo.com**. The Yahoo! search engine home page will appear.

2. In the left navigation bar, click **Groups** to display the Yahoo! Groups page.

3. Click **Computers & Internet**. You will see the subtopics of the Computers & Internet topic.

4. Click **Internet** to display hyperlinks related to Internet technology.

5. Click **Instant Messaging** to display hyperlinks related to instant messaging applications.

6. Click **Yahoo! Messenger** to display hyperlinks specific to that application.

7. Click one of the hyperlinks in the Group Listings section to display information related to Yahoo! Messenger.

8. In the toolbar, click the **Home** button to return to your browser's home page.

Static directory searches may require more steps than keyword searches. However, when you do not know which keywords to use, the static directory search is an excellent tool.

Google

The Google search engine (*www.google.com*) is the *de facto* standard in search. Like other sites, Google's Web search engine is designed to rank the relevance of a particular site according to the keywords entered by the user. However, Google's system also uses additional techniques:

- It determines a site's relevance based upon the number of links that point to it. Google's developers have created statistics-based applications that, in effect, count each link from one page to the next as a "vote." Essentially, Google determines a site's popularity based on other sites linking to it.

- Google programmers have created sophisticated applications to sift through Internet and Web site content, in addition to searching the keywords entered by users. As a result, Google is also capable of giving weight to particular links. For example, if both the New York Times Web site and a simple user-based site both link to a particular site, the link from the New York Times Web site will receive more weight.

- Advertised sites appear at the top of Google results. Companies pay Google to highlight their sites, and the sites are displayed with a yellow background.

To learn more about Google's search and ranking functions, visit
www.google.com/competition/howgooglesearchworks.html.

Similar to other search engines, Google also allows you to use specific tools to find the
information you need. For example, you can use:

•   Google Maps for finding directions to specific locations.

•   Google Images to search for images.

•   Google Scholar for conducting academic searches.

The Google home page is shown in Figure 6-8.
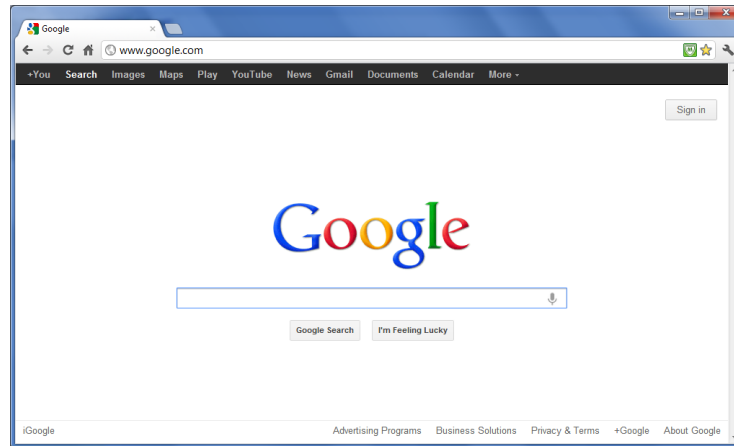
Figure 6-8: Google home page

To perform a keyword search, type the keyword(s) in the search box, then click the Find
button. When specifying keywords, try to use nouns. Verbs and modifiers can be too
variable to be useful, and many search engines ignore them. Articles and prepositions
such as "the," "an" or "with" are also too common to be useful for a search.

You can also use more than one keyword in a search. Using several keywords can help
narrow the search results to yield hyperlinks that are more pertinent to your topic. Some
search engines assume when you list multiple keywords that you want the results to
contain either one or the other of the keywords. However, some search engines, such as
Google, assume that you want the results to contain all keywords. Check a search
engine's Help section to determine the logic that is used by default.

To further narrow your search, specify a search string by using quotation marks to
combine words that must appear together in the specified order. Because results are
restricted to exact matches, using quotation marks will treat "information technology" as
a single phrase and you will receive only hyperlinks to Web pages that contain both
words in that order.

In the following lab, you will perform a keyword search. Suppose you are a Web developer
and you have been asked by your Web Manager to research the most recent guidelines
surrounding Web development. You know that the World Wide Web Consortium (W3C)
provides standards and guidelines in the form of W3C Recommendations that may
provide the information you need. How would you quickly find the W3C home page on the
Internet?

NOTE:
In Step 2, notice the line containing the number of sites found, or "hits."

NOTE:
In Step 3, many results will not be related to your searches. The site keywords may not represent the site accurately, or the keywords may be taken out of context. Sometimes the site is related to the search, even if the description is not. You must follow a link to determine whether the site is relevant.

Lab 6-2: Performing a keyword search using Google

In this lab, you will use Google to perform a keyword search.

1. In your browser, go to *www.google.com*. The Google search engine home page will appear.

2. In the Search box, type *world wide*, then click the **Search Google** button. You will see hyperlinks to all Web pages containing the keyword "world" or the keyword "wide" or both. Notice that each hyperlink is accompanied by a description of the Web site. Notice also that Google found billions of sites using these keywords.

3. Click one of the hyperlinks and determine the relevance of this site to your search and your needs.

4. In the browser toolbar, click the **Back** button (or similar) to return to the list of search results.

5. In the Search box, add quotation marks before and after the existing text so that it reads *"world wide"*, then click the **Search Google** button. You will see hyperlinks to all Web pages containing the keywords "world wide" in the order specified. Notice that the number of Web sites found, or hits, is still in the tens of millions but is dramatically lower than in your first search.

6. In the Search box, change the search string to *"world wide web consortium"*, then click the **Search Google** button. Notice that the number of hits has again been reduced and now is in the hundreds of thousands.

7. Click the **World Wide Web Consortium (W3C)** link to display the site.

8. On the toolbar, click the **Back** button (or similar) until you return to the Google home page.

Notice that finding the W3C Web site is only a two-step process using a keyword search (Steps 6 and 7 in this example). Using a directory search to find the W3C Web site would involve many more steps.

Bing

Microsoft's Bing search engine (*www.bing.com*) was formerly known as MSN Search and Windows Live Search. It is the search engine that powers Yahoo! search. Bing provides search suggestions as queries and also includes an Explore pane that displays related searches. To accommodate social networking, Bing includes a Sidebar feature that searches a user's social network (Facebook and Twitter). It uses the social network to find more relevant information for a particular user's query. The Bing home page, shown in Figure 6-9, includes a background image that changes each day.

Figure 6-9: Bing home page

Excite

The Excite search engine (*www.excite.com*) offers a completely different way to search the Web. Like other search engines, it allows keyword searches. However, the Excite database also contains cross-referencing fields for conceptual searches. The Excite database knows that one word or phrase may have the same meaning as another, and it will return those related results as well.

WebCrawler

WebCrawler (*www.webcrawler.com*) was started in 1994 as a private project to offer free searching to Internet users. WebCrawler is a much smaller database than Lycos, and its search results do not supply the same amount of detail, but it is fast, simple and reliable. WebCrawler is a good search engine to use when searching for a broad topic, such as travel, sports, entertainment and so forth.

Site rankings

Many search engines, such as Google, conduct statistical analyses on Web sites in order to return a ranked list of sites. Such search engines have become very popular because they help the user differentiate between popular and unpopular sites.

Manipulating search engine results

Web users have learned the methods that search engines, such as Google, use to create their vote-based systems. Users have created automatic applications that spoof links to Web pages, making the statistical analysis applications return inaccurate rankings. As a result, it is even more important to evaluate the results of your searches. You should not necessarily trust the most popular sites simply because a Web search engine ranks them highly.

Boolean operator
A symbol or word
used in Internet
searches to narrow
search results by
including or
excluding certain
words or phrases
from the search
criteria.

Boolean Operators

Although search engines have sophisticated programming to help you conduct searches, this programming is nearly useless unless you learn how to conduct a relevant search. Anyone can conduct a simple word-based search. However, fewer people know how to use key search tools and techniques. Learning these techniques will help you become a more efficient user of the Internet and the Web. You will save time and find resources that others will not discover.

Just as standard mathematics uses addition, subtraction, division and multiplication, Boolean logic uses the AND, OR and NOT operators. **Boolean operators** enable users to narrow their Internet search results by requiring or excluding keywords from the search criteria.

The AND, OR and NOT operators

When you use the AND operator between two keywords, results must contain both keywords. The more terms you combine using the AND operator, the fewer results you retrieve.

When you use the OR operator between two keywords, results will contain one keyword or the other or both. The more terms you combine with an OR operator, the more results you retrieve.

When you use the NOT operator with a keyword, you are specifying that the results cannot contain the keyword that follows the operator.

NOTE:
Boolean logic is
named after British
mathematician
George Boole
(1815-1864). Boole is
known as the
founder of
mathematical logic.

For example, if you were searching for laser printers, a standard search using the keyword *printers* may yield millions of Web sites. These documents will include printing companies, ink jet printers, dot matrix printers and many other types of printers, whether they are relevant to your search or not.

To limit your search results, you could use a combination of keywords and Boolean operators, such as *printers AND laser NOT ink jet*. By adding *AND laser*, you are requiring your results to contain the words *laser* and *printers*. By further narrowing your results with the NOT operator, you will exclude all pages that contain the words *ink jet*. This method will narrow your search results dramatically. By using additional Boolean operators to exclude other keywords irrelevant to your needs, such as *NOT dot matrix*, you can narrow your search even further.

Additional operators

Many search engines allow the use of the plus sign (+) to signify the AND operator and the minus sign (-) to signify the NOT operator.

You can also use parentheses to combine Boolean phrases. This strategy identifies single concepts and indicates the order in which the relationships should be considered.

NOTE:
Take time to read
the Help section in
each search engine
to determine the
operators and/or
symbols it supports,
and the default
logic it uses.

Most operators are interchangeable among search engines. However, not all search engines recognize the same operators, or they may use commands other than Boolean operators. Table 6-2 lists some common Boolean operators and the actions they perform.

Table 6-2: Boolean operators

Boolean Operators	Action Performed
AND, &	Results must include both words.
OR, |	Results must include at least one of the words.
+ (plus sign does not work in advanced search)	Results must include the word specified after the operator.
NOT, !, - (minus sign does not work in advanced search)	Results must exclude the word specified after the operator.
NEAR, ADJ (adjacent), FAR, BEFORE	Two keywords on a page must be within a certain proximity to each other (near or far).
" "	Keywords must appear in phrases in the order specified.
( ), < >, [ ], {}	Boolean operator phrases must be performed in the order specified.
* (Example: *color** will return *color, colors, colorize, Colorado*, etc.)	The root word can include variations in spelling. This Boolean operator is called a wildcard character.
. (Example: *color.* will return only *color*, not *colors, colorize, Colorado*, etc.)	The root word cannot be expanded.

Tech Note: *Boolean operators in Excite and Google are case-sensitive; in other search engines they are not. Some search engines allow user-defined variations of the Boolean operators. For more information, view the Help files at each site.*

CIW Online Resources – Online Exercise

Visit CIW Online at *http://education.Certification-Partners.com/CIW* to complete an interactive exercise that will reinforce what you have learned about this topic.

Exercise 6-2: Boolean operators

In the following lab, you will use Boolean operators to narrow Web search results. Suppose that the IT manager has asked you to train the new sales interns on how to quickly access information on the Internet by narrowing their search results to a manageable number. Which Boolean operators would you teach them about first?

Lab 6-3: Using the plus sign (+) and minus sign (-) operators in Google

In this lab, you will use the plus sign (+) and minus sign (-) operators in Google.

1. First, you will use the plus sign operator to search for Web pages that contain the search string "IT administrator" and the keyword "network." Display the **Google** home page if necessary.

2. In the Search box, type *"IT administrator"* (with quotation marks), then click the **Search Google** button. You will see hyperlinks to all Web pages containing the keywords "IT" and "administrator" in the order specified. Notice that the number of sites found is in the millions.

3. Change the text in the Search box so it reads *"IT administrator" +database*, then click the **Search Google** button. You will see hyperlinks to all Web pages containing the original search string and "database." Notice that the number of sites is far less than your first search.

Note: In most search engines, the plus sign (+) or minus sign (-)must directly precede the keyword it requires or excludes; do not add a space between them.

4. Next, you will use the minus sign operator to narrow the search. Change the text in the Search box so it reads *"IT administrator" +database -network*, then click the **Search Google** button. You will see hyperlinks to only a few Web pages containing the original search string and "database," but excluding pages that contain the word "network." Notice that the number of sites found has again been reduced.

5. The search has become too narrow and the results are not necessarily helpful. To widen the search, remove the quotes from "IT administrator" and run the search again as **IT administrator +database –network**. The search results will expand and offer more relevant choices, including jobs available.

6. Use the plus sign and minus sign operators to perform your own Web searches. Do you find these operators helpful in narrowing your search results to display more relevant hyperlinks?

Note: The plus sign and minus sign operators do not work in all search engines. Check the search engine's help files to determine the advanced search tools available.

Advanced Web Searching Techniques

A basic search using a keyword or a search string is sufficient when a simple, distinguishable topic is requested. However, for more complex subjects, or to narrow the results of a standard search, you should use advanced search techniques. Some search engines use a separate interface for advanced searches, whereas others allow you to use Boolean operators in the standard search interface.

In this section, you will learn about the advanced search features of Google, Excite and Yahoo! Not all search engines provide a separate interface for advanced searches. Some sites, such as Bing, expect users to use Boolean operators and other search tools in the basic search field.

CIW Online Resources – Movie Clips

Visit CIW Online at http://education.Certification-Partners.com/CIW to watch a movie clip about this topic.

Lesson 6: Conducting an Advanced Search

Advanced searching in Google

The advanced search features of Google enable users to narrow search results from a standard search to specific criteria. Figure 6-10 shows the Google Advanced Search page.
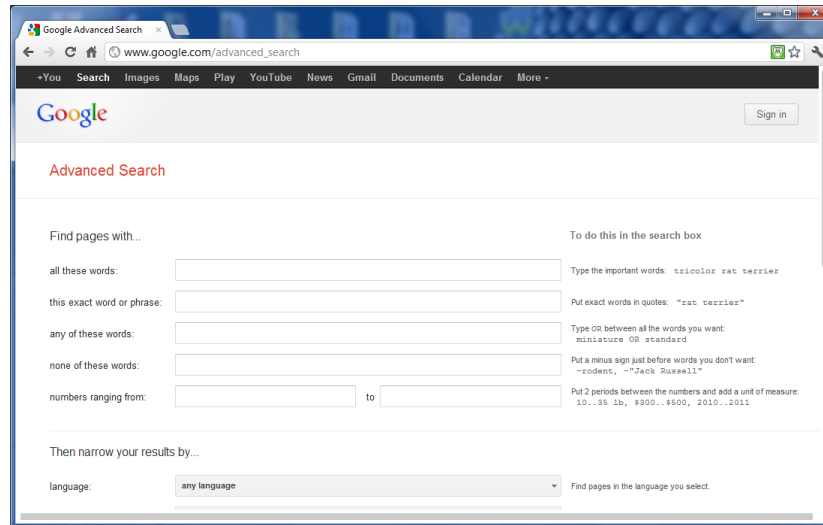

Figure 6-10: Google Advanced Search page

In the following lab, you will use Google's Advanced Web Search feature to perform an advanced search. Suppose the sales interns at your company are familiar with basic keyword search techniques. What other search techniques can you show them to make their Internet searches as efficient as possible?

Lab 6-4: Performing an advanced search in Google

OBJECTIVE
1.11.2: Common search engine operators

In this lab, you will build a query to perform an advanced search in Google.

1. Display the **Google** home page.

2. Go to *www.google.com/advanced_search*. The Google Advanced Search page will appear.

3. Click anywhere in the **This Exact Word or Phrase** box, type *project management*, then press **ENTER** or click the **Advanced Search** button. Notice that hundreds of millions of sites are found.

4. Click the **Back** button (or similar) to return to the Advanced Search page. Click anywhere in the **All These Words** text box, type *phases "life cycle"*, then press **ENTER**. The search results will include only those Web pages containing the original search string plus pages that contain the words "phases" and "life cycle." Notice that the number of sites found is significantly reduced. Click the **Back** button (or similar) to return to the Advanced Search page.

5. Delete the text in both text boxes.

6. In the This Exact Word or Phrase text box, type *database administration*, then press **ENTER**. Notice that millions of sites about database administration are found. Click the **Back** button (or similar) to return to the Advanced Search page.

7. Click anywhere in the **None Of These Words** box, type *software Microsoft*, then press **ENTER**. Most sites that focus on Microsoft database administration software are removed from the results. Notice that the number of sites found is dramatically reduced. Click the **Back** button (or similar) to return to the Advanced Search page.

8. In the None Of These Words box, delete **Microsoft** but keep **software**. Press **ENTER**. Notice that the number of sites found has increased from the previous search. Click the **Back** button (or similar) to return to the Advanced Search page.

9. Capitalize the first letters of *database* and *administration* so that they read ***Database Administration***, then press **ENTER**. Did this change affect the number of sites found?

In this lab, you conducted an advanced search in Google.

Advanced searching in Excite

The Excite search engine (*www.excite.com*) obtains its search results from Google, Yahoo! and Bing. Excite allows conceptual searches that will return results related to the keyword(s) you enter, as well as sites that exactly match the keyword(s).

For example, if you entered the keywords *small business tax forms*, Excite would find sites that exactly match the keywords. Expanding on exact matches, the database also knows that a relationship exists between tax forms and government, and would also return sites that discuss the concept of taxes and government in addition to taxes and small businesses.

Excite supports the Boolean operators AND, OR and NOT, as well as parentheses. If you use these operators, Excite will automatically disable its conceptual search mechanism. Therefore, the results will display only documents containing your exact keywords.

Advanced searching in Yahoo!

Yahoo! has a specific interface for advanced searches (*http://search.yahoo.com/web/advanced*) that provides word filtering, URL and site name searching, and language querying. To access the Yahoo! advanced search, you can also click the Advanced Search link under the More menu on the Yahoo! home page (*www.yahoo.com*).

Using Web Searches to Perform Job Tasks

OBJECTIVE
1.11.9: Internet searches for job tasks

As you have seen, you can conduct Web searches to find information about any subject imaginable. You can use any search engine to find any type of data you need from the Internet. This vast resource can help you perform tasks at your job, regardless of what your job may be.

For example, many jobs require employees to perform research periodically. Search engines can help researchers find the information they seek or the correct resource to provide that information. If you are performing research on the job, you need only specify keywords about the topic being researched, then analyze the resulting references and choose the best sources of information.

Search engines can help you find other online services that can assist you in your job. You can accomplish many tasks quickly with instant access to maps, travel services, product comparisons, couriers, supply ordering and delivery, Web hosting services, yellow pages, news, weather reports, people searches, and much more.

Whatever tasks your job requires you to perform, using the Internet can often save time and money by providing fast, convenient access to a vast amount of information and services.

In the following lab, you will use the Internet to perform a job task. Suppose you are a novice network engineer and your company is setting up a branch office in an area outside the city. The IT director has asked you to research various kinds of wireless transmission media with which to connect the network equipment. The current equipment is old and the IT director wants to upgrade the equipment in the new office. She has also asked you to research wireless networking standards to ensure that data will be secure in the new office. She is concerned that wireless networking may compromise confidential business information sent across the network. How can you use the Internet to perform this job task?

Lab 6-5: Using the Internet to perform job tasks using Bing

In this lab, you will use Bing to access Internet resources to perform the job task assigned to you in the previous scenario.

1. Go to ***www.bing.com***. The Bing search engine home page will appear (Figure 6-11).
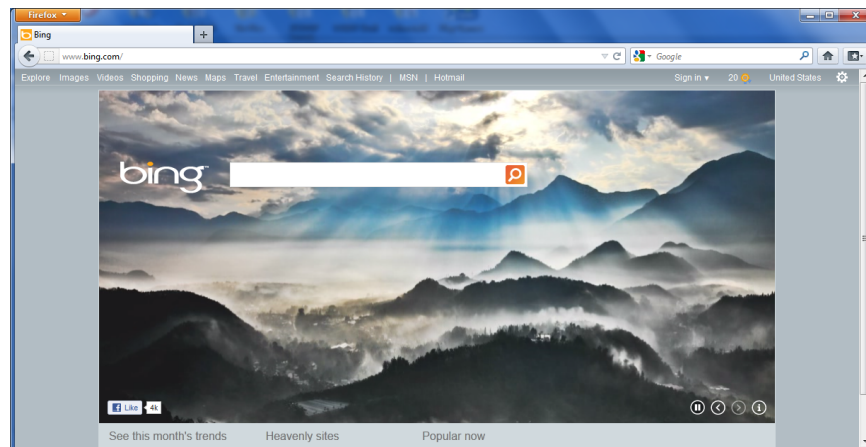
Figure 6-11: Bing home page

2. In the Search box, type *"**wireless network**"* (with the quotation marks), then press **ENTER**. Notice that millions of links are found containing the keywords you specified. However, several links on the first page appear to be very helpful.

3. Click several of the links on the first page to determine whether they provide information about wireless transmission media.

4. In the Search box, type *"**wireless network**" AND "**data security**"*, then click the **Search** button. Notice that the number of sites found is dramatically reduced.

5. Click several of the links on the first page to determine whether they provide information about secure wireless networks.

6. In the Search box, replace *network* with ***standards*** so that the text reads *"**wireless standards**" AND "**data security**"*, then click the **Search** button. Notice that even fewer links are found during this search.

7. Click several of the links on the first page to determine whether the information you find is more helpful than any of the previous sites you visited. Notice that subtle changes in keywords can yield different results.

8. Are there better keywords and search tools you can use to find the information? Change the keywords and use different search tools to provide the best information.

In this lab, you used the Internet to conduct searches for job tasks. When you are conducting Internet searches, you should try various keyword combinations in order to retrieve information as closely related as possible to the job task you are trying to complete.

Unexpected Web Search Results

Search engines sometimes yield unexpected results such as an error page, an alternative search engine or an advertisement instead of the topic for which you were searching.

Error pages most commonly occur when you enter erroneous search criteria, click a dead link or try to access a busy server. When an error page appears due to a busy server, you can click the browser's Back button to return to the previous screen, wait a few seconds, then click the link again. Provided that the link is not a dead link and you entered the search criteria correctly, the site you are trying to access should display.

Unexpected Web pages and advertisements most commonly display when you click a link for a site that was added to a search engine's database by its spider program. Spider programs can inadvertently add unrelated sites to a search engine's database if links to that site exist on the Internet.

Portals vs. keyword searching

OBJECTIVE
1.11.4: Search sites vs. information portals

You may remember that Web sites are added to information portal databases manually. Because you commonly search for data in information portals by clicking subtopic links to reach a specific site, it is much less likely that you will encounter unexpected results. However, if you conduct keyword searches in search engines or information portals, unexpected results can occur if unrelated Web sites are encountered.

In the following lab, you will access a Web site that yields helpful information about various methods you can use to search the Web efficiently for the information you need. Suppose the sales interns at your company try to access a certain Web site to sign up for sales training, and an error page displays instead. As the help desk technician, what would you recommend they do to obtain the information they need?

| **Lab 6-6: Reviewing Web search tips and tricks** |

In this lab, you will use Google to search for Web sites that provide valuable techniques, tips and tricks to help you efficiently search the Web for the information you need.

1. Display the **Google** search engine home page (***www.google.com***).

2. In the Search box, type ***Google Help Basics***, then press **ENTER**. You will see links to Web pages containing the specified keywords.

3. Click the **How To Search On Google – Search Help – Google Help** link to display a Web page containing information about how to conduct efficient Web searches using the Google search engine.

4. After you are finished reviewing the Help page, return to the **Google** home page.

5. In the Search box, type "***web search tips" AND tricks***, then press **ENTER**. You will see links to Web pages containing the specified keywords.

6. Scroll down the page and click several links to learn more about Web searching tips and tricks.

7. As a class, discuss the following questions:

    • Which Boolean combinations are the most useful to help you narrow your search results?

    • Which search engines and search engine types are the most useful to conduct Web searches?

    • What are some of the similarities and differences among the help sites listed in the search results?

In this lab, you learned about techniques, tips and tricks to efficiently conduct Web searches for the information you need.

Web Search Strategies

Before beginning any search, think about the information you are seeking. If you want to browse the Web just to see the types of information available, start with an index on a site such as Yahoo! If you are looking for specific information, use a major keyword search engine such as Bing, Google or Yahoo! If you want as much information as you can find about a topic, try your search on several search engines.

Check the default settings for each search engine you use. Some automatically insert a Boolean AND between keywords, whereas others automatically insert a Boolean OR.

Consider the following tips when you enter keywords in search engines:

• Use keywords that are specific, and try to use nouns rather than verbs.

• Put the most important keywords first in your keyword list.

• Use the plus sign (+) before each keyword to ensure that it is included in the result.

• Whenever possible, combine keywords into phrases by using quotation marks to indicate exact wording, and limit your search results by using the Boolean operator AND.

• Use all uppercase letters when typing Boolean operators in search engines (for example, AND). Some engines will accept both lowercase and uppercase operators, but all will accept uppercase. Conversely, you should use all lowercase letters when typing keywords so that all variations of the keywords (whether they appear in lowercase or uppercase in the Web pages) will be found.

• Use parentheses to combine Boolean phrases in order to identify single concepts and to indicate the order in which the relationships should be considered.

Web search relevancy

Search engines display results in a similar manner, generally a list of hyperlinks to Web pages containing the information that most closely matches the keyword(s) you entered. Each database uses its own scoring system. Some engines, such as Google, are useful for deciding whether the information is current; others, such as Excite, offer ways to sort results. The more searches you conduct, the better you will become at reading and quickly navigating through the results.

A ranking system, or relevancy, is used to help you determine how closely a site matches the topic for which you are searching. The more frequently your specified keywords are found in a particular document, the higher the relevancy score that document receives. The more powerful search engines use both the words you enter and their synonyms to perform a search, which yield more relevant Web pages.

Usually, a scoring system ranks sites from 0.000 to 1.000 (1.000 being the highest probability of a match), but other variations are used. If you find that one site has the highest score of all the results displayed, you should still look at other results. Because these scoring systems used by search engine databases are based on programming instructions, savvy Web developers sometimes construct Web pages with many keyword references. Thus, you might not always find the best information on the first site you visit.

Additional search tips and tricks

You have learned how to browse the Web using various search techniques to find the information you want. In addition to the techniques you have used so far, there are other tips and tricks you can use to quickly access exactly the information you want.

In the following lab, you will use additional Web search techniques that can assist you in Web research. Suppose the CEO has asked you if you know of any searching techniques she can use to quickly access certain types of information that she requires regularly. You have just learned about some additional search techniques in Google and decide that perhaps the CEO may find these techniques useful.

Lab 6-7: Using additional search tricks

In this lab, you will use Google to learn about some additional search tricks and fun techniques that can enhance your browsing experience.

1. Redisplay the **Google home page**, if necessary.

2. Rather than typing simple keywords to search for a topic, you can use verbs, questions and probing phrases to help you find information that is more complex than simple fact listings, and to find it faster. In the Search box, type ***define program management***, then press **ENTER**. You will see links to Web pages containing definitions of program management.

3. In the Search box, type ***what time is it in Tokyo***, then press **ENTER**. You will see links to Web pages that provide the current time in Tokyo, Japan.

Tech Note: Most search engines ignore common words such as "is," "it" and "what." Today, however, search engines can also readily tell when a question is being asked, and thus they will use sophisticated, custom-created applications to catch these questions and return the most relevant sites in the search engine results page.

4. In the Search box, type ***100 US dollars in euros***, then press **ENTER**. You will see links to Web pages that provide currency conversions so you can determine how many euros equal US $100.

5. At the top of the Google home page, click the **Images** link to display the Google Images Search page, shown in Figure 6-12.
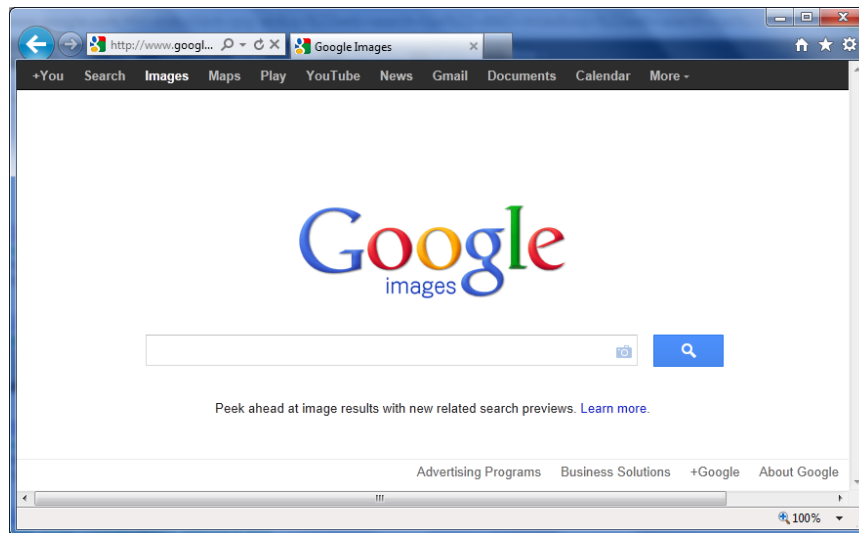
Figure 6-12: Google Images Search page

6. In the Search box, type ***Taj Mahal***, then press **ENTER**. You will see pages of thumbnail clips of the Taj Mahal.

7. Experiment with Google to search for videos, news and maps.

8. Close your browser.

As you use search engines more often, experience will teach you that you can find a much wider range of information types on the Web than you may have thought possible. Going beyond simple keywords to use verbs, questions and probing phrases can help you to find information that is more complex than simple fact listings, and to find it faster.

Evaluating Resources

It is vital that you understand how to evaluate information you retrieve from resources on the Internet. Evaluating a resource involves making judgments about the resource. Judgments can include determining the quality of the research that was likely to be involved in creating the Internet article, as well as determining the type of bias presented, if any.

Too often, however, people are told to avoid a particular resource because it is considered suspect for one reason or another. Students are often told to avoid retrieving information from Wikipedia (*www.wikipedia.org*), for example, because information in it is considered by some to be suspect, or incorrect. Problems with open information sources such as Wikipedia can occur because anyone can add to or remove the information presented. Also, contributors are most often not authorities in any particular field. Contributions to Wikipedia are loosely restricted, if they are restricted at all.

However, it is also incorrect to assume that Wikipedia contains no accurate or valuable information. Wikipedia has taken steps to ensure the reliability of content that is posted about subjects. Article postings that are questionable are flagged and noted so that the reader will understand that they may need to question the validity of the content. Wikipedia also requires contributors to create accounts prior to editing documents, which helps track down sources of content deemed biased or incorrect.

By contrast, authoritative encyclopedias such as Encyclopedia Britannica (*www.britannica.com*) are often suggested and trusted because contributors to it are considered true experts in their fields, and contributions are strictly controlled.

In general, it is important to understand that no source, even one that is considered authoritative and is "peer-reviewed," is consistently accurate. Therefore, it is vital for you to develop critical reasoning and evaluation skills to judge information on the Internet.

Types of resources

OBJECTIVE
1.11.10:
Authoritative vs. open peer-reviewed sources

Traditionally, a peer-reviewed resource is one that accepts input only from recognized scholars and authorities. The "peers" who review these entries are likewise scholars and authorities in their fields. With the advent of sites such as Wikipedia, "peers" have now become defined as individuals with similar interests, or simply as people who want to contribute articles.

Reviewed and non-reviewed resources

Some resource types describe the verified level of expertise that its contributors possess. The source of information can directly affect the quality and factuality of the information. Following are three types of resources:

- **Authoritative peer-reviewed** — These sites and resources only accept input from scholars and recognized authorities. Traditional examples include Encyclopedia Britannica and Merriam-Webster (*www.merriam-webster.com*). If you do not feel confident about your ability to evaluate resources, you may want to use this type of resource exclusively.

- **Open peer-reviewed** — Content is open to review by anyone; "peers" do not necessarily possess any particular level of expertise or qualification in the topic. The most common example is Wikipedia. Many others exist.

- **Individual** — The Internet is full of sites created by individuals whose abilities to conduct research and whose levels of objectivity vary widely. Such sites typically do not undergo any sort of review, and they represent research of varying quality.

OBJECTIVE
1.11.11: Primary and secondary resources, indexes

Primary and secondary resources

Another way to describe resource types is by the proximity to original data that the resource possesses. You will generally be viewing three categories of resources on the Internet:

- **Primary** — source material that has not been edited or compiled in a significant way by a researcher. Primary source material can include correspondence (letters/e-mail), photographs, physical objects and relics, debate transcripts, surveys, speeches, some podcasts and videos, autobiographies, and any records and transcriptions of an event.

- **Secondary** — research based on primary resources, preferably by a credible individual or group. Secondary resources are the most common type of resource you will use. Resources includes textbooks, magazines (print or online), newspapers, biographies, most wikis, commentaries on speeches, reference books, histories and books written by an author about a particular topic.

- **Indexes** — encyclopedias and reference material that list additional titles. These additional titles can be either primary or secondary sources.

Popular and scholarly resources

It is also important to distinguish between popular and scholarly resources. Although both types of resources will often discuss the same topics, their approaches are generally quite different.

- **Popular resource**— generally bases information on secondary resources. Examples include online and print resources such as Newsweek (*www.thedailybeast.com/content/newsweek.html*) and The Register (*www.theregister.co.uk*).

- **Scholarly resource** —bases its research on primary resources. Examples include The Computer Journal (*http://comjnl.oxfordjournals.org*), Briefings in Bioinformatics (*http://bib.oxfordjournals.org*), and the British Journal of Aesthetics (*http://bjaesthetics.oxfordjournals.org*).

Directional and informational resources

Resources are also often described in terms of whether they provide information themselves or simply provide a listing of other resources on a topic. Following are two types of research sources:

- **Directional** — a resource that contains references to one or more additional resources. Indexes and abstracts contain directional sources. Examples of directional resources include Google Scholar (*http://scholar.google.com*) and ERIC (*www.eric.ed.gov*).

- **Informational** — a resource that provides either essential or detailed data about a particular topic. Examples of informational resources include most primary and secondary resources. Informational resources can also contain directional resources. For example, at the bottom of almost any encyclopedia or Wikipedia article will be a list of references.

Developing evaluation skills

After you have found the information you need on the Web, you must still evaluate its accuracy and value. The fact that you found it on the Internet does not make it true or reliable. You can expect to find all sorts of resources on the Web: personal pages, reference sources, white papers, hoaxes, advertisements, scientific papers and personal opinions. Web pages are not overseen or controlled by a central authority. You must determine whether the pages are objective or biased, accurate or inaccurate. Take the following steps to decide whether information is reliable or factual:

- **Discern between fact and opinion** — As you first conduct research, identify information that both sides of an issue can usually all agree on, including dates, events and chronology.

- **Identify bias** — Almost all resources you read will have a bias of some kind. Bias is not the same as prejudice or hate speech. Some of the most informative and useful resources you will find will have a bias. The best sources, however, take care to ensure that this bias does not cause the author to distort or misrepresent facts. As you read, ask yourself the following questions:

  "What is the author's tone?" In other words, how does the author address his or her audience? Is the tone angry? Sarcastic? Condescending? Tone is a fairly accurate indicator of an author's bias.

  Also, constantly ask yourself the following question: "How did the author reach his or her conclusions?"

- **Consider your sources** — Always consider the source of the information. Remember that you can often learn a lot about a site from the URL. Does your information come from a commercial organization? An institution of higher learning? A non-profit organization? Information you find on a .gov or .edu site will probably be accurate because these institutions do not publish sites with the goal of making money. You may need to verify information you find on a .net, .org or .com site because these types of sites generally present material favorable to their own agendas, services or goods.

- **Identify the contributor** — Identifying the contributor includes determining the bias of the author, as well as determining whether any organizational ideology somehow governs or influences the information you are reading.

- **Consult authoritative, peer-reviewed resources** — Determine the depth of research. It is always useful to start with an overview of essential information. Eventually, however, you will have to delve more deeply into information in order to be truly informed about a subject.

- **Compare information you obtain** — Consider two or three separate resources about the same topic. As you read multiple articles, confirm that information from one source is not simply repeated in another. With Internet-based resources, it is very simple to copy and paste information from one site to another.

- **Compare information you obtain with information find in print** — Remember that hard-copy resources — such as professional journals, scholarly writings and reference books — contain accurate and current information that you can use to corroborate the information you find during Web searches. If hard-copy reference materials refute your Web search findings, you may want to search further or use the reference materials as your basis of information.

- **Withhold judgment** — Avoid making snap judgments about the validity of a resource simply because you agree or disagree with the resource. Make sure that you take the time to compare information among articles. Try to make educated guesses about the topic, and then research the topic further to determine whether your guess is correct or not.

- **Identify essential concepts** — Many free resources provide accurate information, but very little context. Therefore, you may find an overwhelming amount of data about a topic that may be unrelated, as opposed to helpful, contextualized information.

- **Consult with trusted individuals —** Work with your parents, guardians, instructors, mentors and other reliable sources to determine the validity of information.

The above steps will help you become an authoritative source yourself. You will be able to more easily discern bias in the resources you read, identify underlying motives, and eliminate poor information.

Comparing resources

As you compare resources, identify common language usage. It is possible — even likely — for multiple resources on the Internet to have one common source. In other words, even though you can quote multiple sites, each of these sites may simply be repeating the same original source. If the use of language is similar, including similar word usage, verb usage and sentence construction, then you may not truly be consulting multiple resources.

Questionable resources

If you encounter resources that overtly or implicitly contain hate speech or otherwise preach a hatred toward a particular ideology or ethnicity, you should immediately stop using that resource. Take care that you are not tricked into participating in forums that advocate such speech or activities. Consult with trusted advisors to answer any questions you may have.

In the following lab, you will review types of resources and evaluate information you find on the Internet. Suppose you work in your company's marketing department and you have been assigned to write a white paper for a line of your company's technology products. You need to research some similar products offered by competitors, as well as include reviews and feedback that your company's products have received. It is important that the information you provide in your white paper is accurate. You can use the Internet to perform your research, provided that you carefully evaluate the information you find there for factuality.

Lab 6-8: Evaluating Internet resources

In this lab, you will review types of resources and evaluate information you find on various resources.

1. In your browser, go to ***www.youtube.com***. The YouTube home page will appear (Figure 6-13).
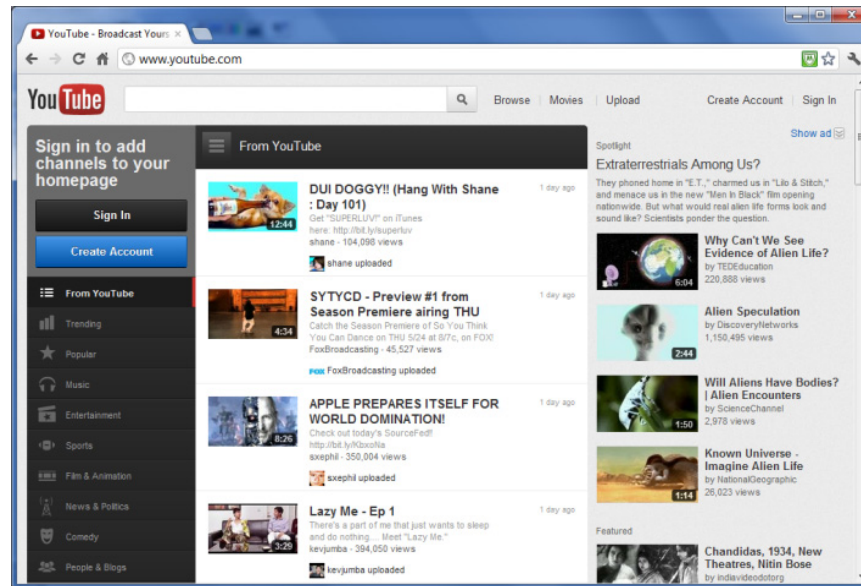
Figure 6-13: YouTube home page

2. Explore the YouTube home page and conduct a search for videos of your choice.

3. Now go to ***www.bbc.co.uk/news/***. The BBC News home page will appear (Figure 6-14).

Figure 6-14 BBC News home page

4. Compare the BBC News page with the YouTube page you explored in Step 2. Which of the two sites is most likely to provide primary sources, and which is most likely to provide secondary sources?

5. In the spaces below, write down the names of three scholarly or popular journals (online or print):

_____

_____

_____

6. Go to the Wikipedia (***www.wikipedia.org***) or the Encyclopedia Britannica (***www.britannica.com***) site and search for information about the journals you entered in Step 5. Review the links and articles you find that are related to your search. Can you find directional sources in any of the articles? What conclusions can you draw about each author's perspective based on the article's content? Can you detect any biases?

7. Use the Wikipedia or Encyclopedia Britannica site to conduct research about **Open Source** and **Microsoft**. Based on your research, answer the following questions:

- Do the articles you find about both topics seem to be accurate?

  _____

- Which resource seems to provide more complete information about the topics?

  _____

  _____

- Consider that resources such as Wikipedia and Encyclopedia Britannica are preliminary sources and are more directional than true informational references. What in these resources can you use to obtain more credible resources?

  _____

  _____

  _____

Organizing Internet Research

The Internet can be an excellent research tool because it enables you to find information about any subject imaginable. However, unless you have some way to organize this information, your research can turn into an endless cycle of surfing, copying, pasting, bookmarking and deleting. Internet research tools such as Zotero (*www.zotero.org*) enable you to organize the information you find and store it on your local hard drive.

Zotero is a free, open-source Firefox add-on that you can use to collect, manage and cite reference material you find on the Internet. As you conduct research, Zotero can detect when you are viewing a book or article, and then save the full reference information to a local file. If the source is a Web page, Zotero can also store a copy of the article to your local file. Zotero is shown in Figure 6-15. It occupies the bottom half of your screen by default.
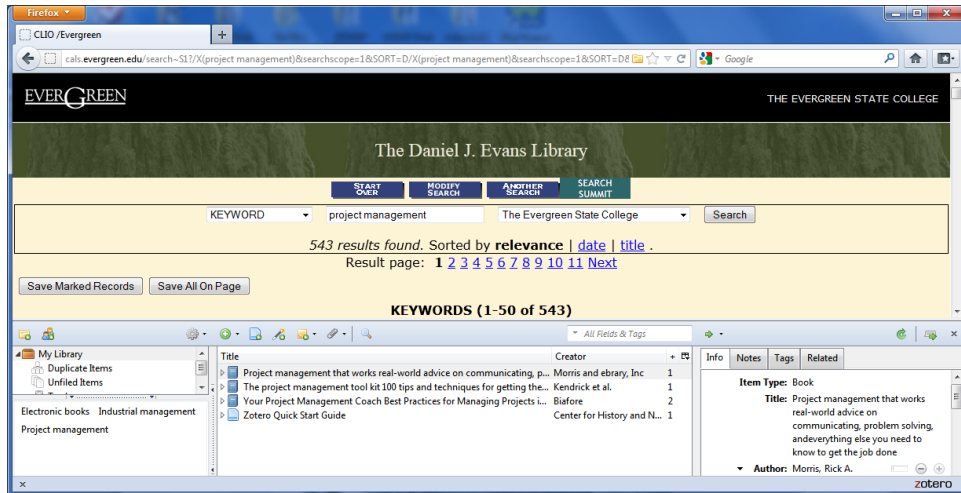
Figure 6-15: Zotero

The left column of the Zotero window contains your library (My Library) and any collections you add. The middle column shows the items contained in the collection selected in the left column. The right column shows information about the item selected in the middle column.

In the following lab, you will conduct Internet research using the Firefox add-on Zotero. Suppose you are a high school senior and your instructor has assigned a term paper in which you must conduct research about project management fundamentals, especially as they relate to the IT industry. You can use Zotero to collect and organize the information you find on the Internet relating to your topic. As you write your term paper, your resources will be easily available, so you can cite authorities and construct your bibliography correctly and appropriately.

Lab 6-9: Organizing Internet research using Zotero

In this lab, you will use the Zotero add-on in Firefox to consolidate information that you find as you conduct an Internet search.

1. First, you will download and install the Firefox add-on Zotero. In the Firefox browser, go to *www.zotero.org* to display the Zotero home page.

2. Click the **Download Now** button. Scroll down to the Downloads section and click the **Zotero 3.0.7 for Firefox** link (your version may vary). An alert window may appear stating that Firefox has blocked the site from asking you to install the software. Click the **Allow** button.

3. In the Software Installation dialog box, click **Install Now**, then click **Restart Now** to complete the installation.

4. When Firefox restarts, the Zotero Download page will appear. Click the **Documentation** link, then click the **Quick Start Guide** to learn more about Zotero. If time permits, watch the Video Tour. Notice the Zotero button on the right side of the status bar. You click the Zotero button to toggle Zotero on and off.

5. Click the **Zotero** button on the right side of the status bar to open the Zotero window in the bottom half of your screen.

6. In the Firefox Location Bar, go to ***www.evergreen.edu/library/*** to display the home page for The Evergreen State College library.

7. On the library's home page, click in the Search field near the top of the page, type ***project management AND IT***, then press **ENTER**.

8. Scroll down in the results list and view the links to books that match your keyword search criteria. Click a link that interests you to display information about that book.

9. Notice the book icon that appears at the right end of the Firefox Location Bar. You click the book icon if you want to add the current book to your Zotero collection. Click the **Book** icon now. Notice that the book title and author appear in the middle pane of the Zotero window, and information about the book appears in the right pane.

10. In the Evergreen window, click the **Back To Results List** link, click another book link, then click the **Book** icon in the Location Bar to add the book to your collection.

11. Repeat Step 10 as many times as you want to add book references to your Zotero collection. When you are finished, go to your browser's home page. Notice that the Zotero window remains on your screen. Your screen should resemble Figure 6-16.
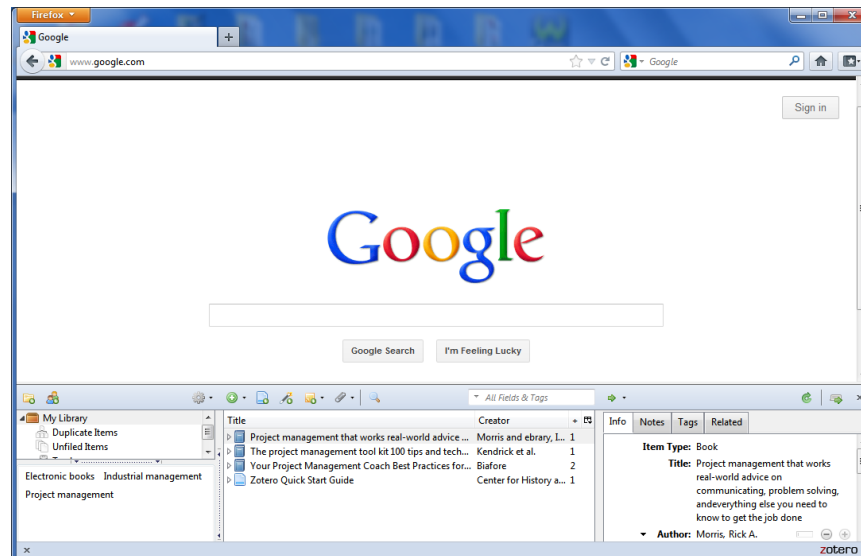

Figure 6-16: Zotero window with collection

12. Ensure that one of the books you added to your collection is selected in the middle pane, then view the **Info** tab in the right pane. Notice that information about the book appears, such as title, author, publisher, publish date, page count, ISBN and so forth.

13. Click the **Notes** tab. You can use this tab to add any notes you like regarding the selected item.

14. Click the **Tags** tab. You can use this tab to add relevant keywords that you would like to associate with the selected item. Tagging items makes it easy for you to categorize them later. You may notice that Zotero has already added some tags based on the subject headings and keywords associated with the items you added to your

collection. For the example shown in the previous figure, the keywords "Electronic Books" and "Project Management" appear on the Tags tab.

15. Click the **Related** tab. You can use this tab to associate other items to the selected item that are not covered by the other tabs.

16. Click the **Paper Clip** icon on the Zotero toolbar to add attachments. You can use this button to attach files, such as word-processing documents, PDFs and images.

17. Click the **Zotero** button in the status bar to close the Zotero window, then close **Firefox**.

Researching new software

As IT professionals, you will probably be directed (or want) to conduct Internet research about advances in the IT industry, particularly new software as it becomes available. As you saw in the previous lab, you can use add-ons such as Zotero to help organize the vast amount of available data and help you limit it to information in which you are interested.

Following are some excellent online sources for researching new software:

• **Lifehacker** (*http://lifehacker.com*) — a blog about Windows, OS X and Linux software that includes time-saving productivity tips and tricks (known as life hacks).

• **The Register** (*www.theregister.co.uk*) — a technology news and opinion Web site from the United Kingdom that employs heavy doses of sarcasm and satire in its articles.

• **PC Magazine** (*www.pcmag.com*) — an online computer publication that reviews and provides previews of hardware and software advances in the IT industry.

• **Free Code** (*http://freecode.com/*) — a Web site that reviews and provides previews and articles on the latest software updates and releases. The majority of the reviews and articles deal with open-source software.

• **Sourceforge** (*http://sourceforge.net/*) — a hosting repository for software developers to store programming code in an effort to manage and control the development of open-source software applications.

• **Slashdot** (*http://slashdot.org/*) — a technology news Web site that provides articles about current affairs. User-submitted articles are subject to an editorial process before being posted to the site.

Citing Copyrighted Online Resources

<div style="border:1px solid #000; background:#cceeff; padding:4px; width:140px;">
OBJECTIVE
1.11.8: Reference citation standards
</div>

It is important to properly cite information that you obtain from an Internet search to avoid plagiarizing someone else's work. Plagiarizing information you obtain from the Internet is the same as plagiarizing someone else's printed work and could constitute copyright infringement. When using information from any medium, you must cite the source to give appropriate credit to the creator.

Style manuals, such as the *MLA Handbook for Writers of Research Papers* and *The Chicago Manual of Style*, are good sources for determining the way to properly cite Internet-based information. The citation for an Internet source should be the same as that for a print source.

According to the *MLA Handbook for Writers of Research Papers*, a reference to Internet-based information should be cited as follows:

- Lane, Patrick. "BYOD and Exchange ActiveSync." <u>CompTIA IT Pro Community</u>. May 2012. < http://itpro.comptia.org/networking/b/weblog/archive/2012/05/23/byod-and-exchange-activesync.aspx>.

In this example, the author is Patrick Lane; the title of the article is "BYOD and Exchange ActiveSync"; the source of the article is the CompTIA IT Pro Community Web site; and the issue date was May 2012. Notice that the author's last name is displayed first, and that name is used to sequence this entry in an alphabetical list of citations.

The Chicago Manual of Style uses the following format for citing the same Internet-based information:

- Patrick Lane, "BYOD and Exchange ActiveSync," *CompTIA IT Pro Community*, May 2012, http://itpro.comptia.org/networking/b/weblog/archive/2012/05/23/byod-and-exchange-activesync.aspx (accessed May 24, 2012).

In this example, the components of the entry are the same as in the first example. However, the author's last name is not displayed first, and the date on which the Web site was accessed is displayed parenthetically after the Web site reference. The punctuation scheme also differs.

Either citation method is acceptable, although you should choose one standard and use it consistently. Most organizations designate standard style manuals or conventions for company publications.

Copyrighted material

Remember that any published work or expression is protected by copyright. You cannot copy and use someone else's work for your own site or publication without express written permission. Always check copyright notices on any site you want to borrow from. In many cases, the copyright owner may only require proper credit and citation for you to use the material, but sometimes you must pay a fee. The copyright owner sets the terms. Always observe fair use of others' copyrighted material. Copyright infringement is a punishable offense.

CIW Online Resources – Course Mastery

Visit CIW Online at *http://education.Certification-Partners.com/CIW* to take the Course Mastery review of this lesson or lesson segment.

IBA Lesson 6 - Part B

To Be Valid or Not to Be Valid?

Ellen is an anthropologist who has been working at Olduvai Gorge in Tanzania for the past six months. She has been conducting research on the Internet. She finds a Web site with an article that proposes a revolutionary theory regarding human migration patterns from east-central Africa that may have occurred about 150,000 years ago. Ellen wants to determine whether the information is valid before she presents it to her fellow anthropologists for consideration.

Ellen performs the following tasks:

- She obtains information about the organization on whose Web site the article was presented. Based on the domain name, Ellen determines that the organization is an educational institution.

- She researches other organizations recommended by the Web site containing the article by following some of the links. She finds that the sites are hosted by trustworthy sources known in her field.

- She checks the credentials of the author and determines that he is a known professional in anthropology who has published reputable works.

- She accesses current scholarly writings and professional journals, and finds that this new and exciting theory has the support of other anthropologists.

Ellen feels comfortable presenting the information she found to her fellow anthropologists because she knows it is valid.

*   *   *

As a class, discuss this scenario and answer the following questions:

- What other ways can Ellen evaluate the information presented in the article to determine its legitimacy?

- Because various other resources have cited the theory, does Ellen still need to cite her source when sharing the information? Which source(s) should she cite?

- Consider each of the tasks Ellen performed in the scenario. For each task, what might she have instead found that could have indicated the information was not valid?

Lesson Summary

Application project

Each search engine has advantages and disadvantages for different types of Internet searching. Now that you have used several search engines, which do you think will work best for your needs?

Suppose your company wants to open a branch office in Athens, Greece, or Lisbon, Portugal. Locate commercial real estate in Athens and Lisbon for your new office using an advanced search with Boolean operators and wildcard characters as needed. Use Google, Excite and Bing for your advanced searches (remember that your results will vary). Attempt to narrow each search as much as possible, using only one search entry. Which search engine provided the most useful information? Attempt to locate the rates per square meter of available office space in both cities. What results would you report back to your company?

Skills review

In this lesson, you learned about database components, relational databases, database table keys and relationships, database queries using SQL, and common database vendors.

You also learned about search engines and Web searches, registering a Web site with a search engine, and the types of Web searches you can perform. You learned common basic and advanced search techniques, and the Boolean operators that can help you narrow the scope of your search results.

You also learned how to conduct Web searches to perform job tasks, you learned the reasons you may sometimes encounter unexpected search results, and you learned Web search strategies, tricks and techniques to make the browsing experience more effective and enjoyable.

Finally, you learned evaluating information and resources you find on the Internet, tools you can use to organize your Web research, and ways to properly cite information found on the Web when you use it in your own work.

Now that you have completed this lesson, you should be able to:

✓ 1.11.1: Define and identify Boolean operators and their relationship to search engines.

✓ 1.11.2: Conduct searches using operators common to search engines (e.g., quotation marks, minus sign).

✓ 1.11.3: Identify the importance of keywords and conduct keyword searches.

✓ 1.11.4: Distinguish among search sites, information portals, and pages automatically provided when an error occurs.

✓ 1.11.5: Evaluate research for soft and hard-copy resources, and verify factuality.

✓ 1.11.6: Identify meta search engines, as well as search engine ranking techniques.

✓ 1.11.7: Explain how search engine sites use spiders to enhance search results.

✓ 1.11.8: Cite references to information obtained from Internet-based searches using accepted citation standards (e.g., MLA).

✓ 1.11.9: Search additional Internet resources to fulfill job-role duties.

✓ 1.11.10: Distinguish between authoritative peer-reviewed and open peer-reviewed sources.

✓ 1.11.11: Provide examples of both primary and secondary resources, as well as indexes.

✓ 1.11.12: Explain key differences between popular and scholarly journals.

✓ 1.16.1: Distinguish between relational and non-relational databases.

✓ 1.16.2: Define common relational database concepts, including table, row, record, column, field, data value, join.

✓ 1.16.3: Identify relationships between tables, including one-to-one, one-to-many, many-to-many.

CIW Practice Exams

Visit CIW Online at *http://education.Certification-Partners.com/CIW* to take the Practice Exams assessment covering the objectives in this lesson.

IBA Objective 1.11 Review

IBA Objective 1.16 Review

Note that some objectives may be only partially covered in this lesson.

Lesson 6 Review

1. Briefly describe a relational database.

   _____

   _____

2. What is a search engine?

   _____

   _____

3. What is the function of a <meta> tag?

   _____

   _____

4. List three types of search engine indexing methods.

   _____

5. List at least three standard Boolean operators.

   _____

6. What is the purpose of a search engine ranking system?

   _____

   _____

   _____