The Business Researcher's Challenge

Where do you turn when you need to find company, industry, new-product, market, or business-trend information? These days business researchers begin on the Internet. There are few traditional business resources that have not migrated in one way or another to the Net, and there have been countless sources born on the Web as well.

But as the savvy researcher knows all too well, in addition to the trusted names and sources that are now Internet-accessible, an enormous number of sources of unknown origin and veracity exist there as well. They're all mixed together in the same vat, so to speak, from which the researcher must filter and scoop. This has given business researchers a new challenge: how to ensure that those sources taken from the open Web are reliable and credible.

On the one hand, the Web has provided the serious searcher with a wealth of worldwide information never before available, speeded up access, and often lowered the cost of—or even made free— some previously expensive sources. For instance, before the Internet, researchers did not have easy—let alone *free*—access to company press releases, daily newspapers from around the globe, and mountains of government reports and statistics.

Today, all of this and more is available by pushing a few keys while sitting at your desk. Moreover, it's typically formatted with all of the original graphics. The Internet has also aided the business researcher by creating remarkably helpful communities of like-minded professionals, on newsgroups and mailing lists, that share resources, answer

questions, identify new sources, and perform other collegial activities. And the Net has made it possible to locate the most obscure data, specialized experts, and those with views and activities far out of the mainstream who previously were difficult to find.

But we've also experienced the flip side of the Web. Free information isn't always quality information, and an abundance of it can cause information overload. Further, the same open characteristics of the Web that provide us with the ability to participate in a worldwide conversation and let us connect with so many interesting people has also made it possible for anyone to be a self-appointed journalist, publisher, expert, or even broadcaster. All of this without necessarily taking any training programs, displaying any qualifications, or having anyone else check their information for accuracy.

None of this is news to today's experienced searcher, though you deal with this reality on a daily basis. You do a Google search for company information and you're directed not to, say, D&B or Standard & Poor's, but to an online directory you've never heard of before. Who is this publisher and where did it get its information? Or you find a local business journal with a Web version of an article that ran in its print version, but is what you found *really* the same as what was in the print publication? And where is the date of this article? Another search, say on trends in a new and growing market that you're researching, sends you to the pages of a popular "Blogger," someone who has created his or her own personal news and opinion page. What are this person's qualifications? Should you believe what he or she is telling you?

These are the problems confronting today's business researcher who ventures onto the Web.

Some Things Haven't Changed

Naturally, the need to evaluate the reliability of an information source goes beyond the Web. Good researchers have always confirmed the origin and veracity of all types of business and nonbusiness information. During the 1980s and early 1990s, most business research was conducted via fee-based online services such as Dialog or LexisNexis (which of course are still heavily used by information

professionals, though today they are typically accessed through a Web browser), as well as print sources. When consulting these online sources, there were (and still are) important questions to be asked. Is the online version the equivalent of the original print version? If not, what was omitted? What sorts of quality-control techniques were utilized when entering the data to avoid keying errors?

During this era there was necessary attention paid to data integrity, which in retrospect seems simpler and less burdensome than in today's Internet age. Now there are virtually no quality controls or filters for information to pass through before it is disseminated to millions of people around the globe.

Even before the online information age, there were basic questions that good researchers knew to ask. What is the reputation and mission of this publisher or source? Does the author have an ax to grind? Is there a bias or some ideology that can call his or her credibility into question? How was the data gathered? If the information was derived from a survey or poll, what were the methodologies used?

In the classic text *The Modern Researcher* (Wadsworth, sixth edition, 2003), which was first published in 1992, authors Jacques Barzun and Henry J. Graff lay out what they view as the most fundamental "virtues" that the serious researcher should learn to cultivate: accuracy, love of order, logic, honesty, and self-awareness (pp. 43–48).

The authors explain that the task of ascertaining truthfulness is a daily occurrence for us: "Everybody daily faces the question: Is this true or false? … The conclusions rest on a combination of knowledge, skepticism, faith, common sense, and intelligent guessing" (p. 96).

We must remain vigilant when encountering even the traditional information sources, which can lead us astray just as easily as a Web site (perhaps *more* easily since our guard may be lowered). In October 2002, for instance, a news story made the rounds on several major news outlets, including the *Washington Post*, the BBC, and several other supposedly trusted sources, that a new World Health Organization (WHO) study was predicting that blonds would soon be a thing of the past, due to the recessive gene associated with that hair color. One minor flaw: WHO never issued the story. In fact, with

a little bit of reflection the blunder should have been obvious, since it doesn't make much sense that an organization concerned with health matters would issue a report about the future of blond-haired people. Media organizations that ran this story could have avoided embarrassment by placing a simple phone call to the organization to confirm it. But apparently none did, resulting in a slew of retractions. (The origin of this error, though not totally clear, was allegedly a 2000 article in a German magazine that erroneously cited WHO.)

Today, though, when using unknown Web sources, determining source reliability becomes even more important. There are special issues that need to be examined and new questions that need to be asked, all of which will be discussed in this book.

The bottom line is that the Web has made the job of the business researcher both easier *and* more difficult. It is easier in that the Web has delivered a great diversity of sources and made them quickly available, at no or low cost. The job is more difficult because of both the sheer amount of data we now have to wade through and the difficulty in determining which sources are trustworthy.

Enter the Quality Checklists

So how have business researchers approached the issue of information quality on the Web, and what steps have they taken to deal with it? Librarians and information professionals have weighed in with their own approaches and views. Some just bemoan the problem. Having been burned once or twice, they eschew searching the open Web and advise sticking with tried and true fee-based databases, supplemented with some favorite bookmarks of trusted sites. Others deal with the issue by sticking with government (*.gov*) or education (*.edu*) sites, and remaining skeptical of the dot-coms.

Some have gone a couple of steps beyond by trying to come up with specific guidelines that can be used by all searchers to assess the quality and reliability of what was uncovered on the open Web. These guidelines typically take the form of a criteria checklist that can be used to evaluate a particular Web site.

Personally, I find these checklists quite useful, though only to a point. On the positive side, they help surface the most likely problem areas and remind researchers of areas where attention needs to be paid when looking at sources. They can also be a nice crutch to lean on when venturing out into unknown Web territories, particularly for beginners or students. On the downside, checklists vary in what they ask you to consider. Some leave out areas that others include. Their emphasis on pinpointing a site's potential flaws can narrow your focus and keep you from seeing the bigger issues surrounding how information is created and disseminated.

The other drawback is that checklists do not really offer practical or long-lasting solutions to truly learn what to trust on the Web. It's hard to imagine actually reading down these lists and measuring and rating individual Web sources as you're surfing the Net. It's just too time-consuming and cumbersome. Furthermore, the longer-term solution (thinking critically and analytically about information) is built up over time, sometimes over years, and cannot be substituted with a checklist. So ultimately, while checklists contain useful reminders of what to look for when examining Web sites, I don't see them as an elegant, workable method of evaluating information on the Web.

That said, I'm not in any way against these checklists, as they can be used as a tool to increase awareness and improve one's searching. (In fact, in Appendix A you'll find a checklist I wrote myself, plus some others I recommend.) Because they help make sense of the Web's information glut, they are also excellent teaching aids for students or novices in Web-based research.

Furthermore, I'm not certain that many checklist creators ever intended them to be used in a literal manner. Just keep in mind that the checklist alone is not going to "solve" the problem of evaluating information on the Web.

The New Internet

The Internet has changed so much for researchers, mostly for the better. It's also had some fascinating effects. One that's particularly noteworthy is how it has changed what we think of when discussing

an information item's "relevancy." Before Web search engines, when you searched a traditional fee-based database, any returned results could and would be deemed as "relevant" to your inquiry, since the search engine made a match between your keywords and the words in the article or other information item. Technically, those items were the relevant ones, even though they weren't always exactly what you wanted.

On the Web today, it seems that you can enter almost any series of random words (as an "AND" search) and nearly always retrieve results. The fact is, when searching such a huge set of information— literally billions of Web pages—probability theory compels that your chances of turning up results—results being defined as pages that include your keywords—are pretty high. But whether those results are truly *meaningful* is another matter altogether.

What I mean by this is as follows. Say you had a hunch that companies that outsource have lower morale among their employees and you wanted to go to the Net to see if there was any confirming data to support this. You go to a search engine, enter, "low morale" and "outsourcing" and lo and behold, you turn up scores of Web pages and discussion forums, which to your surprise—seem to confirm your theory! You find, that, yes there are people discussing and exploring how outsourcing is lowering morale. Similarly, say you are wondering if there is currently a housing bubble in the real estate market, and you enter "housing bubble" into a search engine and, yes, find *lots* of discussion on the subject of an upcoming housing bubble.

But you need to keep in mind that because you are searching such a vast amount of information, the odds become pretty good that you're going to find someone, somewhere, who also happens to believe or is exploring what you are looking into. And finding these kinds of "supporting" views can easily feed into something that all psychologists know about—a phenomenon called "self-confirmation bias"! Simply put, this means we tend to give credence to information and views that support our own pre-existing beliefs and opinions.

Furthermore, the fact that you are searching a set of billions of Web pages means that the odds increase that you'll turn up pages just because they happen to include the keywords you've entered. But many pages will just include your keywords somewhere on the page,

unrelated to each other. In this case, you'll get back lots of results, but they won't necessarily be relevant to your query.

So if it's on the Web you seek, you'll more than likely find—just be careful that what you find is relevant, meaningful, and, if you're looking to verify or confirm a theory or point of view, something more than just a case of self-confirmation bias!

Another way in which the Internet has completely changed the scene has been in empowering the researcher to go beyond merely scooping up facts and secondary information and to become a more proactive, directed editor or fact-checker. Indeed, smart researchers today understand that the information-quality problem on the Web isn't necessarily in the believability or credibility of the individual who is posting it. Rather, it's the lack of standard editorial processes that traditional publications go through, such as topic screening, peer review, copyediting, fact-checking, and proofreading. For this reason, it behooves you, the researcher, to assume those roles yourself—that is, to examine what you find with skepticism.

Interestingly, while the Internet has eliminated much of this function, it has also given birth to some tools and sites that make it possible to reintroduce these critical functions, at least partially. In some cases, with additional verifications you're able to go well beyond what the typical researcher has been able to do. For example, in 2002 President George W. Bush said something controversial, and the media took him to task for it. One particular blogger took umbrage at this, but rather than simply note his disagreement, he offered visitors to his site the ability to link to an archived video clip of the event. This way they could judge for themselves whether the President's comments were inappropriate. Now that's the kind of media check that, until not so long ago, hadn't been readily available to information users!

Another way in which the Web has changed information flow is in how quickly both good and bad data is spread. When a dramatic news story breaks or a rumor begins to emerge, the Web, primarily in the form of e-mail forwarding, rapidly and seemingly uncontrollably accelerates the data. In the case of bad information, the result can be

that, in a very short time, it's possible for a great many people to believe something that isn't true.

In an informative and amusing April 2000 interview on National Public Radio's *All Things Considered* ("A Cautionary Tale About Facts on the Internet"), author and expert searcher Reva Basch was interviewed about the problem of commonly believed, but untrue information that's passed on via the Web. In response to NPR host David Kestenbaum's query about this phenomenon, Basch replied jokingly that this situation "does call the definition of reality into question. If enough people arrive at a consensus on a fact, does the nature of reality change?" (See Appendix B for the link to the audio of this broadcast.)

Defining reality is a bit beyond the scope of this book, but I will observe that, fortunately, business research does not normally rely on instant access to breaking news or rumors. Typically, it involves more in-depth, deliberative research over a longer period of time. This allows ample opportunity to question the source, analyze it, and get additional confirmations, and then apply the various safeguards we'll discuss throughout this book.

Why This Book?

Why do we need a book to help us evaluate business information on the Web? The problems in establishing the veracity of Web information in general are clear, but what's special about business information and why does it need its own treatment?

For one thing, there's usually much at stake in locating accurate data on the Web and in making sure you're not deceived by erroneous or even deliberate misinformation. Business research is normally performed prior to making some significant business decision. It could be research on whether to introduce a new product, partner with another company, plan to enter a new market, create a competitive strategy, understand changing buyer habits, or keep track of emerging opportunities and threats within your own market. And while a good business decision requires more than accurate and reliable information, it's impossible to make a good decision that's based on bad information. Plus, your own personal credibility is on

the line whenever you provide information to another party, as you are implicitly vouching for its accuracy simply by choosing to pass it along.

But it isn't easy finding reliable, trustworthy business information on the Web. Although we've come to appreciate Google's uncanny ability to retrieve relevant results, it's easy to be swamped with too many pages of dubious quality if you don't know how to create a focused, precise search and take advantage of all its powerful features.

As impressive as Google is, it still cannot screen out pages where data is too old, biased, or created from an untrustworthy source. Google cannot tell you which directory has the latest sales figures of the companies you want to track. It can't tell whether a poll was conducted scientifically or not. It cannot tell you if a company is fudging its profits, or if an analyst is giving you biased recommendations, or if an organization's agenda is different than what it appears to be. These are still matters best left to human analysis.

In order to really end up with the best business information on the Web, you need to know when and how to go beyond search engines, when to use other information-gathering tools, and how to pull together collections of pre-screened sites.

This book is designed to tell you what you need to know about all these matters—how to create the best business-information search, use the right research tools, and scrutinize Web sites for credibility—so that you can rely on the business information you find on the Web.

Seeing the Big Picture

For better or worse, doing any type of research has become synonymous with going to the Web. How did we get to this point, and what are some of the larger implications for the serious researcher?

Remember Marshall McLuhan's admonishment: First we shape our tools, and then they shape us. The Internet is an incredibly powerful tool, and it has an equally powerful influence on us. One effect is in how we view the research process, and what it means to be a good researcher. As outlined in *The Modern Researcher*, these virtues include accuracy, love of order, logic, honesty, self-awareness, and imagination.

The Web certainly challenges researchers to maintain these virtues. According to Barzun and Graff, a key characteristic in maintaining accuracy is "the habit of unremitting attention" to what you find. Can you maintain that level of scrutiny when surfing through Web pages and scanning for data while under time pressure? As for imagination, the ability to "imagine the kind of source [you] would like to have before [you] can have it" is imperative. But if research now means simply going to Google and typing in keywords, how good are you in imagining your ideal source?

Similarly, any good investigative researcher will tell you that a thorough investigation means lots of nitty-gritty, detailed work, and often takes many months. It means you'll have to comb through primary documents, find people who know what you need, and interview them. In the old-fashioned journalist vernacular, you need to use up a lot of shoe leather. On the Web, as professors today lament, it's almost too easy to "perform research" simply by scanning the first few pages returned by a search engine.

In *The Road Less Traveled*, M. Scott Peck suggests that if there's anything in our nature that can be termed "original sin," it would be laziness. This is the inertia we all feel that prevents us from doing the necessary work and making the effort that's required to do what we know needs to be done. And unfortunately, doing research on the Web makes it easy for us to indulge our propensity for laziness, to the detriment of our finished product.

Like any media or communication tool, the Web embodies inherent characteristics that determine and influence the way we obtain and use information. But this is a subtle process. Like the old saying about a fish's inability to discover water, we too are becoming immersed in the medium, which may make it impossible to distinguish the new information environment clearly. But because the Web is a new medium, we still have the opportunity to observe its impact. The following are some of the special hazards we'll encounter as we move to a mostly Web-based research world:

1. **Missing Context**—A search engine does not care about context, it merely retrieves for you the page that contains

your keywords. But understanding the context of what you're reading is critical to understanding it. You'll need to do the work necessary to find this larger context.

2. **Emphasis on speed**—If there is a single primary "message" to the Web, it may be speed. Everything is instantly available. Researching can get news that's only a few minutes old. But while speed is important, it's not necessarily a friend of good research. In fact, emphasizing speed increases the chance of making mistakes and merely allows for data access, rather than knowledge building—a process that takes time.

3. **Emphasis on style and graphics**—Web designers and online journalists alike know that it's difficult to read long blocks of text on the Web, and that a page must look appealing to draw people to it. But an emphasis on design can come at the expense of text length, and the written word is still the best way to convey complex ideas. Relying only on shorter "chunking" of text and eschewing longer analyses means there's a greater potential of getting a superficial read of the subject.

4. **Lack of historical perspective**—The Web is a new medium and, other than some specialized archival collections that have been digitized, most pages are no older than the early 1990s. That's fine for most business research purposes, as you rarely need old business information. But on those occasions where you do require historical, archival-type information, it's important to go off the Web and check specialized hard-copy or microfilmed collections.

5. **Search Engine/Keyword Mindset**—Being a good researcher goes beyond creating keywords or even searching online databases. *The Modern Researcher*'s Barzun expressed concern that online searching "presupposes the searcher's knowing ahead of time what he should find out."

Further, he notes that "a database is good for the person who wants to know the length of the Brooklyn Bridge; it is not for the one who wants to understand the respective effects of Norman, Arab, and Spanish influence in Southern Italy" (personal correspondence, October 1, 1995). Or, to put Barzun's point in a business-related context, an online search won't help you understand how the economic forces of the 1990s are going to influence the success of your new product introduction in 2005. In other words, building an in-depth knowledge and understanding of any complex matter takes more than a search.

6. **Bots vs. People**—Scientists and researchers continue to create smarter, more capable automated tools to collect and display information. Some of these are impressive indeed and have the capability to search through millions of articles. These tools cull those articles that meet the profile of what you want to track, summarize them, and then display them in an intuitive, graphical manner. U.S. intelligence agencies even use these programs to identify linkages among people and organizations that are suspected of terrorism. These tools and bots are generally touted for their ability to turn information into knowledge. But only humans have the ability to provide true analysis and insight, and it will remain this way for the foreseeable future. Only people can answer your specific questions by calling on their wealth of experience and knowledge.

The Internet Emerges

It's interesting to go back 10 years and read what was written in the monthly newsletter I founded and edit, *The Information Advisor*. The article that appears in the following sidebar was published in December 1992. George H.W. Bush had just lost the presidential election to Bill Clinton, and the European Union was created earlier that year in Maastricht, Netherlands. An excerpt of the article, "Catching Business

Information on the Internet," provides an interesting glimpse into the technology as we knew it then.

Catching Business Information on the Internet

What do you need to know about "the Internet"? During the last 6 months or so, interest and discussion about this international electronic network has exploded. But much of the writing has been overly technical and dense. Furthermore, few—if any—of these analyses have been relevant for the business researcher. This article will explain, in jargon-free language, what the Internet is, how it can be of use to business researchers, how to sign on, and where to find the latest and most useful sources of information.

What Is the Internet?

The Internet is an enormous conglomeration of electronic computer networks. Currently, it links about 4 million people on 1 million computers in over 100 countries.

Here's a capsule history: The Internet was begun in 1969 by the Department of Defense. It grew in size and scope over the years, and eventually became a publicly run network for fostering study and communication in education, research, and academia. Up until the last year or so, [its] use was restricted to those persons associated with an academic or research institution already on the network. Today, while commercial use of the Internet is still restricted, anyone can sign on and join the network simply by hooking up with one of a number of telecommunication vendors (listed later in this article). The Internet is funded today by the National Science Foundation, and has been supported and nurtured by Vice President-elect Albert Gore.

What do people do on the Internet? Well, as with the well-known and popular bulletin board services like CompuServe, Prodigy, The Well, and others, a primary use is electronic mail and user-to-user communications. Internet users also can subscribe to electronic journals, join special interest groups, and access various information databases.

Business Information Uses

But what can the Internet offer specifically to the business researcher? There are probably four major applications: locating experts and answers, sharing information and problems, accessing documents and books, and searching online databases.

Let's look at each of these areas separately, beginning with the locating of experts.

We spoke with Sharyn Ladner, business librarian at the University of Miami. Ladner is the author of How Special Librarians Really Use the Internet and was speaker at the December 7–8 Meckler Conference on Document Delivery and Libraries/Internet and Libraries. Ladner told us that "the best information tool is a Rolodex. And the Internet is an electronic Rolodex." She enthusiastically describes how researchers have "thrown out hard-to-answer questions" onto the Internet, and in a matter of hours, have received a handful … of replies. "How many phone calls would you have had to make to get an answer?" Ladner asks.

Related to the opportunity to locate experts is the ability to share information and problems with peers. The Internet currently has a forum for business librarians, called "BUSLIB-L." Its mission is to deal with "all issues related to the collection, storage, and dissemination of business information within a library setting." While still populated mainly by academic rather than purely commercial researchers, Ladner expects that this group will grow quickly. (Other special-interest lists on the Internet cover government documents, chemicals, law, library reference, medicine and health, science, and many other topics.)

Another use for the Internet is directly accessing documents and books from other network participants. Users can search hundreds of libraries' catalogs online, and information specialists can use this function to facilitate traditional interlibrary loans.

Online Searching

Finally, Internet users can conduct database searches of certain popular online hosts, as well as smaller, more specialized databases. Among the big names currently accessible are Dialog and Mead Data Central (Nexis/Lexis). Why would you choose to search these hosts on the Internet? There are pros and cons to searching these files here. Let's look at searching Dialog first.

A major reason why you might want to search Dialog through the Internet is cost. Currently, accessing Dialog through one of the major telecommunication vendors like Tymnet or Sprintnet costs an average of $12 per hour in telecommunication fees. However, it costs only $3 per hour on the Internet (plus regular Internet access fees). The other advantage to searching Dialog on the Internet is more abstract; it relates to the synergy you could achieve by having an electronic communications experience that consisted of a mixture of posting queries, downloading other users' data resources, and searching databases all during a single session. Dialog's Internet connection is managed by ANSI, [which is] located in Michigan.

On the negative side, searching Dialog on the Internet is rather user-unfriendly. You cannot perform certain functions such as capturing log files or utilizing accounting functions [that are] otherwise available.

Searching Mead's files on its own Meadnet system costs $13 per hour in telecommunications costs. Mead has created a version of its communication software to search the Internet. However, the firm tells us that the product is made available to law school and business school users only.

Dialog and Mead are not the only online systems accessible through the Internet. There are also individually accessible database files, which include

economic and census reports, and science and technology files. However, no complete list of databases exists.

Access Options

Once you've determined that you want to give the Internet a try, how do you sign on? There are many different ways to connect. … As mentioned earlier, if you belong to an organization that already is connected to the system, all you need is a password. Otherwise, you'll need to contact a commercial or specialized provider.

Probably the best-known names are CompuServe and Delphi. Once you decide to sign up, be prepared to deal with a sometimes difficult interface and confusing procedures. …

So, other than evidence that Al Gore was at least a nurturing father figure to the Internet, what can we glean from this 1992 article?

As expected, much has changed since then. Certain Internet service providers mentioned are either no longer in existence or have been absorbed by other companies. And the technical aspects of connecting and navigating the Internet have clearly improved.

But the Internet is still doing what it was doing best even at those nascent, pre-World Wide Web days: It allows us to find other *people*, who remain the best, most reliable sources of expertise. They might be trained information experts like journalists or librarians, or just have knowledge, interest, and a willingness to share what they know. The scholar, teacher, businessperson, author, and passionate amateur expert can all be located in a matter of seconds through the power of the Internet.

You, then, as an intelligent business researcher, will use the Internet as an extraordinarily powerful data-gathering tool and an excellent method for locating knowledgeable people. You must

employ your skepticism and skills to ferret out the good and bad data. You can then take the next step and ask your insightful questions, critically assess the information you receive, and work to build your own knowledge and further your own understanding.

Scope of This Book

It's important to make clear what this book covers and what it does not. The most important point to keep in mind is that I am limiting the discussion of business information to what can be found on the *free* Web. Specifically, I am referring to the portion that is accessible by anyone, at no charge. (This includes sites where you may need to register, but not pay a fee.) Also included here are those sites and pages that constitute the "Invisible Web"—these contain information that resides on the Web, but for a variety of reasons are hidden from search engines. These "invisible" pages are covered here, as long as they are free.

What I don't cover are those fee-based traditional databases that have migrated to the Web: large online vendors like Dialog, LexisNexis, and Factiva; the individual databases such as ProQuest, Gale, Moody's, and D&B; and smaller, more specialized data providers such as the Economist Intelligence Unit. Also not covered are fee-based business databases that were born on the Web: SkyMinder and OneSource, for instance.

The reason I've chosen not to include those sites is that they represent known quantities to researchers. And the fact that they can now be accessed on the Web (rather than on proprietary dial-up services or through larger database vendors) does not raise too many new, significant quality issues. So if you've searched, say, D&B's Million Dollar Directory by dialing up to Dialog's service in the past, but you now search it either on DialogWeb or even on D&B's own services, there aren't too many critical quality issues to consider. (Not that there aren't *any* new issues that have arisen from having Web access to some of these sources. For example, you'll want to confirm that the same quality and currency is available via the Web interface as

before, in respect to matters like length of archive, update frequency, coverage, advanced search capabilities, and so on.)

There are some sites that fall into a gray area when it comes to categorizing them as free or fee-based. For instance, market-research-report aggregators on the Web allow anyone to search their databases, view initial results, and sometimes view even more (such as a summary or table of contents) for no charge—and some don't even require registration. But if you decide to purchase a report, then you'll need to pony up some money. These kinds of sites—free searching/fee-based viewing—*are* included in the scope of this book.

It might be helpful for me to describe the scope of what I am covering under the broad umbrella of what is called "business research." Obviously, "business research" is a very broad term and can be defined as anything that answers any business-related query, so theoretically any kind of data or source could support a business related question.

In practice, though, there are certain types of information sources that are relied on and utilized most often for business research purposes. These include company information sources, market research reports, statistical data, news and media sources, and online discussion groups and Weblogs. I'll discuss each of these sources in detail in Chapters 5 and 6.

One type of source not covered in this book is information geared toward investors, which generally consists of advice and news on choosing and trading stocks, bonds, commodities, or other investment instruments. These really don't fall under the category of professional business research. Rather, they're mainly geared to consumer or institutional investing.

How to Use This Book

Following this introductory chapter, Chapter 2 will examine some strategies to use before using a Web search engine. I'll review the important non-Internet-based business sources as well as Web sites that have already proven themselves as reliable sources of business-related facts. We'll also examine how information-finding

tools (other than search engines) can be more effective for locating reliable business data than standard search engines.

Chapter 3 will in fact focus on search engines—specifically, on using them effectively to pinpoint just the business information you need. I'll tell you how to get the most out of Google: how to make its advanced search features work, filtering out questionable sites from the get-go, analyzing its initial results page to spot the most promising pages, and more. Also included is a summary of what I think are the best of the other search engines and advice on when to use them.

In Chapter 4, I'll tell you what to do when you arrive at a site or source of questionable origin or veracity. Here you'll get some tips from some of the best investigative reporters around on analyzing the truthfulness of a source. You'll also discover the top signs that you may be dealing with a source of dubious quality. You'll learn how to go beyond just reacting to what you see on your screen by using some proactive evaluation strategies to determine whether a source is credible.

The two chapters that follow will help you evaluate specific types of commonly used business information sources. Chapter 5 looks at company information from corporate home pages, market research reports, company directories, and a few other sources. Chapter 6 focuses on statistical data, polls, and surveys. Chapter 7 examines online news and Web discussion groups. For each source I'll discuss the most common quality problems, examine specific Web-related concerns, offer tips on finding out if your source suffers from those problems, and offer advice on how to ensure that you locate quality information.

In Chapter 8, we'll look at some of the overriding, big-picture issues that we need to know about to feel confident that we're gathering quality information on the Web. Here I'll discuss the importance of building your own knowledge on a topic so as to more easily spot bad data, along with the roles of critical thinking, intuition, and trust.

Finally, Chapter 9 is a compilation of filtered and screened business sites, organized by category, that have proven themselves to be reliable and useful sources for business research purposes. An important part

of the process of doing research on the Web is to have a set of trusted, stable Web sites at your disposal. The sources listed in this chapter will help you develop such a collection.

Appendix A offers my Web site evaluation checklist along with recommendations of other checklists you'll find freely available on the Web. Appendix B is a chapter-by-chapter listing of resources recommended in the book.

To make *The Skeptical Business Searcher* easier to read, I have placed the URLs for all recommended Web sites in Appendix B. The companion Web page to this book features hot links to these Web pages; this page will be updated on a semiannual basis.

I'd love to hear from you and hope you'll share some of your own tips with me. I'll include them at the Web page and perhaps in a future edition of *The Skeptical Business Searcher*.

Now, let's get out onto the Web and start finding those reliable business sources you need.