

Chapter 1

Defining Data Loss

In This Chapter

- ▶ Understanding how the world of data has changed
 - ▶ Identifying IT risk as a big part of your business
 - ▶ Getting *The Big Picture*
 - ▶ Discovering the biggest threat to the twenty-first century
 - ▶ Taking a holistic approach to data loss
-

Why worry about losing data — there's always more where it came from, right? Well, that's part of the problem. Like other kinds of stuff, data accumulates. By 2010, we'll have a vast pile of that electronic stuff, more than if we had a Ph.D. in Having Lots of Stuff: the most extreme an estimate up to *988 billion gigabytes* of information. Where will we put it all?

In addition, how we *handle* all that stuff is speeding up. How are we to keep up? A recent educational estimate from a North American teacher illustrates this change: For students starting studies in technology at the university level in September, over half of what they learn in their first year will be out of date by the time they get to their third year — and *all* of it will be out of date by the time they start work. The Department of Trade and Industry in the U.S. estimates that in 2010, the top ten technical jobs that will be in demand won't even have existed in 2004.

Meanwhile, computer circuitry has crept into nearly everything you use — and many of those things now gather information *from* you and *about* you. If your kids know how to program every electrical device in the house, and you gave up on that about 15 years ago, it's an indicator of what's happening. If you've been shopping for cell phones recently ("mobile phones" if you're in Europe), you've probably seen a sales pitch that points out all the "features" of the phone — especially the ways it can send, receive, and juggle information in various forms. If you're tempted to interrupt with a practical question ("But can I make and receive *telephone calls* with this thing?"), consider asking one more: "How does this thing keep my data safe?" Don't be surprised if you get a blank look.

In effect, we're sending our children to school to prepare them for jobs that don't yet exist, using technologies that haven't yet been invented, to solve problems that we don't yet know are problems. All we know for sure is that this future will be hip-deep in copious quantities of data. Which must be stored somewhere — and (more important) *protected*. Because along with all these cool new capabilities come new ways to poach, pilfer, and nab our sensitive personal and corporate data. Too often, we're too busy trying to get a handle on the technology to be aware of the threat.

How the World of Data Has Changed

Today we're almost entirely reliant on Information Technology — IT — departments at work. Organizations across the world depend on enterprise-wide applications, used by everyone from employees and customers to suppliers and partners. Numerous applications support key business processes such as e-commerce and business intelligence, but they also create a mountain of information that must be successfully harnessed, securely stored, and continually accessible to the right people at the right time. Otherwise it wanders into the wrong hands — of which there are a lot more these days.

Economic growth — a barrage of gadgets

With huge economic growth in emerging countries — India, China, Russia, Brazil, and bits of Eastern Europe (watch Hungary and Romania grow) — the growth of digital information is accelerating to humongous proportions — the primary drivers will be rich media, user-generated content, and in excess of 1.6 billion Internet users — and a boatload of photo phones. An estimated minimum of 50-million-plus laptops and an estimated 2 billion picture phones will be shipped *every year* for the foreseeable future.



One problem with all those portable devices is that they're easy to lose. For example, did you know that in 2004 — in London alone 1— 20,000 cellular (mobile) phones, 11,000 PDAs, and 10,000 laptops were left behind in taxis? It turns out that you're 12 times more likely to lose your phone than your laptop. So, if a typical organization can expect to lose up to 5 percent of its laptops per year, that means about 2,350,000 laptops — and 51,700,000 phones — will go missing this year. Along with *everything that's on them*.

And what's stored on these devices? Information — much of it (unfortunately) *very useful* to people who shouldn't have it. Most folks have multiple copies of their information, stored here and there on various devices. They don't keep track of it, and in fact they don't know where most of it is. It's easy pickings for *cyber-criminals* (bad guys who use computers and the Internet

to commit their crimes). So the more data you have, the bigger the security risk — and these days all that data requires vast amounts of storage. Unfortunately, IT is losing control of the storage — and control of the data. When desktop external storage drives reach terabyte (thousand-gigabyte) capacity and are as easy to use as a little USB storage device, the walls of the data center might as well come tumbling down. Data isn't always stored in the data center. It wanders away to external drives, USB keys, iPods, and cell phones. All the data in your phone's contact list, for example, is *personal identifiable information* — sometimes referred to (irritatingly) as PII. As such it should be protected; in fact, there are laws about protecting information, and most people are either unaware of them or just ignore them.

The messaging boom throws data everywhere

Data is everywhere, e-mail has become the language of business, okay, so we still use the phone, but it's e-mail that transfers data. E-mail, in fact, has become the *business record* — essential documentation for compliance with the laws that affect your industry. E-mail, which is so easy to send to the wrong person — a data leak right there! Of course, e-mail is now seen as *too slow*. So people are supplementing it with instant messaging and texting on mobile phones. We're all going digital in our contact with other people — compulsively sending data off to who-knows-where, usually unaware of unintended side effects like these:

- ✔ The number of text messages sent daily is bigger than the population of the planet. (Even if your teenager seems to be sending half of those, it's still a huge amount of stuff floating around in the ether.)
- ✔ The average distance you have to be from a colleague before you resort to e-mail is (apparently) a mere 6 feet. (More usage means more data wandering around until someone grabs it.)
- ✔ One out of eight couples who married in 2006 in the U.S. met online. (Imagine the sheer amount of personal data the sweethearts must have exchanged without bothering to encrypt it.)
- ✔ MySpace has over 300 million registered users; if it were a country, it would be the fourth largest in the world, between Indonesia and the U.S. The average MySpace page gets over 30 visits per day. (Are *some* of those visitors sneakily checking social-networking sites for unauthorized data? You bet they are.)
- ✔ In 2007, the U.K. saw 160,000 cases at hospital A&E (Accident and Emergency) rooms that were related to people not looking where they were going while texting. That's the equivalent of 160,000 people walking carelessly into lampposts while staring at their phones.

Okay, that might be understandable if you're distracted by a real, live, good-looking member of the opposite sex. One of your authors (we won't say which one) did that once, and then apologized to the lamp-post. But at least he *wasn't* generating personal data the whole time.

- ✓ The development of the Internet has seen a rapid increase in its use for everyday purposes — and a lot of that use is careless, which makes it a great target. Security is poor (because we haven't had to worry about it before), the data is valuable, and too many users are unaware of its value.

Technology gone wild; data gone missing

Not only do we have to deal with a Brave New World of mutating technology, but a much smaller one too; business is always going on *somewhere*, and people worldwide take advantage of it. Old-fashioned bankers' hours don't slow it down a bit when so many people bank online. From Beijing to Bournemouth, a startling array of banks lines up from every country to vie for market share. This shrinking world also has more people in it — close to 7 billion — most with personal information that can be used to help or harm them. This vital resource often resides in *multiple databases around the globe*. If you have a credit card and a bank account, shop a little online, and pay taxes, then your data is held in an *average* of 700 databases! Do you ever wonder whether it's *safe* there?

Sure, all the new forms of e-commerce benefit businesses. You can do business around the world at a click of a mouse; companies can be faster in getting to market, responding to requests, dealing with customers, and making contact with partners and suppliers. Result: huge opportunities to play big even if your company is small. But if you play big, you can also wind up taking big risks.

Organizations and individuals have something in common here: Leaving their data unprotected data puts them under threat from short- and long-term *data loss* — the straying of sensitive information into unauthorized hands.



Irresponsibly (or criminally) used, lost data can easily cost you revenue, reputation, customer loyalty, share value, brand equity, and market share. For openers.

Today's turbulent, networked world presents many risks to corporate information: security breaches, leaks, and losses, infestations of malware, and deliberate attacks on PCs and computer infrastructures can devastate businesses. The risks range from careless to malevolent.

Although user and operator errors account for more than 32 percent of permanent data loss, cyber-crime and computer-virus attacks are the fastest-growing cause of business disruption.

Organizations put challenging demands on their information — not only integrity, but rapid, secure, and continual availability. Although server and storage hardware continue to improve, those goals are harder to achieve as IT infrastructures grow more complex — and add new points of potential failure and vulnerability.

Watertight data-protection and security strategies are critical for every organization. Nonstop, 24-hour operations require some heavy-duty capabilities:

- ✓ Robust, scalable storage management
- ✓ Secure backup and disaster recovery
- ✓ Protection of data from desktop to data center
- ✓ Effective management of a wide range of environments from a single location

No wonder the cost of management, for both security and storage, is escalating. The IT infrastructure not only has to store more information than ever before, but also to secure, log, and discover where the most sensitive electronic information is hiding — in an environment that often isn't equipped to handle those tasks.

Gone are the days when data lived peaceably in the data center and everyone who had access worked in the same building. These days, the data — and the risks — can be found all over the map (Figure 1-1 shows what a typical organization has to deal with). Every stage of the business, any business, has the potential for data loss, data leaks, and data breaches.

As corporations expand their business operations globally, into different geography and various time zones, continuous availability of e-commerce services is expected. E-commerce is a blessing and a curse to business:

- ✓ It makes communication to customers easy to open up new avenues for sales and marketing
- ✓ It puts pressure on IT to minimize the risks associated with managing the exchange and storage of data.

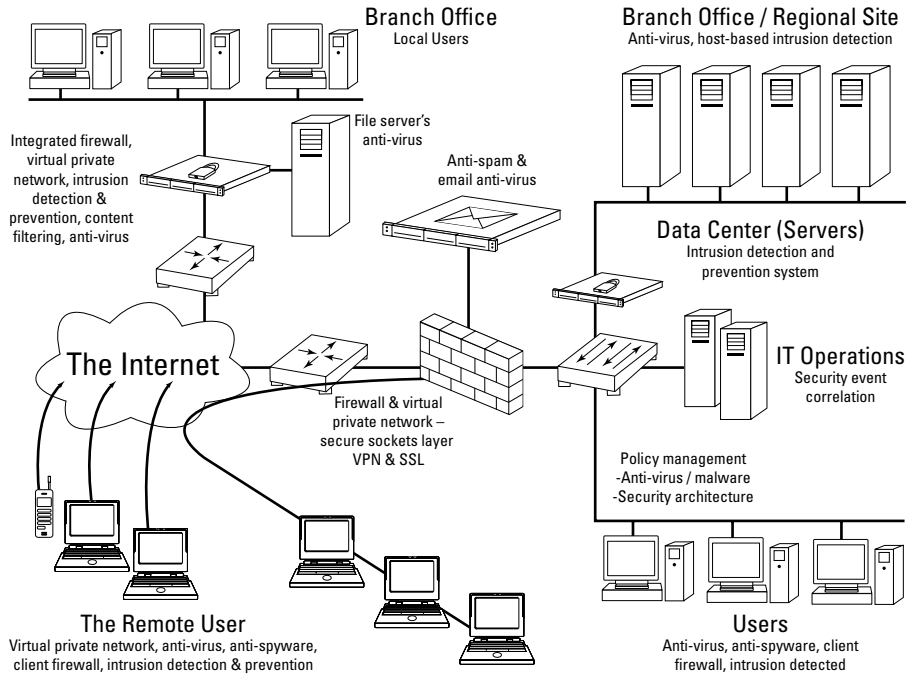


Figure 1-1:
Data-loss
proliferation
in a maze
of security
needs.

A few quick numbers illustrate why this is a growing problem. IDC estimates that in 2008 alone, human beings created more data than in the previous 5,000 years combined. That gives the digital universe an approximate size of over 250 exabytes (billion gigabytes). And IDC projects a hefty growth in the amount of information generated between 2006 and 2010 — to the tune of 600 percent per year. In addition, in 2010 around 70 percent of the digital universe will be *generated by individuals*. Guess who'll be looking after it . . . you guessed it: In 2010, *business organizations will have responsibility* for the security, privacy, reliability, storage, and legal compliance of at least 85 percent of that information.

Where will we put it all?

Managing a universe, even a digital one, will be a huge challenge for organizations. Consider: On average, the annual demand rate for digital storage was only 35 to 40 percent from 2006-2008, and the average level of disk allocation for storage (on Unix/Linux systems, anyway) was only 30 to 45 percent.

Kiss those days goodbye.

The amount of information created, captured, or replicated in 2007 exceeded available storage for the first time. Suppose that currently the amount of new technical information is doubling every year. If all things remain equal (and they won't), those growth numbers tell us that by the end of 2010, the amount of information will be doubling every 72 hours. Which is amazing or ridiculous, depending on your perspective.

If, in 2006, we created 161 exabytes of digital information — three million times the information in all the books ever written (and currently we are publishing around 3,000 books worldwide — daily), here's what that would look like if we literally piled it up: a stack of books from outside your back door to the sun and back — six times. That's 93×12 or 1,116 million miles of books. In one year. That's a cosmically awful amount of stuff we've got to store and protect — all of it with major implications for individuals, businesses, and society in general.



- ✓ The great mass of information will put a considerable strain on the IT infrastructures that organizations have in place today. (Got it in one, Sherlock!)
- ✓ This huge growth will change how organizations and IT professionals do their jobs, and how we consumers use information. (Of course it will! We must be geniuses!)

Who are we kidding? The numbers just scream at you: over 2.7 billion Google searches performed every month (by the way, to whom did we refer these searches B.G. [Before Google]?) Every year we cumulatively wait 32 billion hours for Internet pages to load. We simply can't survive without this stuff. Even if IT bursts at the seams, we've got to have our information. But if IT succumbs, we *won't* have our information. So . . .

Where will it all end?



Somebody's got to get smart about this — we must take steps, as an industry and as individuals, to make sure that we create infrastructures that make information secure, reliable, scalable, and highly available. The name of that game is *information management*, and it won't get easier. Organizations will have to use more sophisticated techniques to manage, store, secure, and protect their information if they expect to survive. To handle the increased amounts of information that we'll have to protect, store, and manage in the future, we have to start now — by getting control of the information we already have.

Information and Communication: Risky Business

Information is our most valuable asset, and yet is coming under continual, increasingly sophisticated attack from cyber-criminals who target it for financial gain. The situation has been exacerbated recently with in wide-spread investment in new, more efficient communication technologies. Communication is essential to business; no argument here. But if you define *communication* as “the sharing of information,” you get several developments immediately:

- ✔ The more easily information is available, the more it tends to be shared.
- ✔ The more widely information is shared, the more ways it can be abused.
- ✔ If your business depends on information (what business doesn't?) but can't control its communication technology, you're in trouble.

Web 2.0 and the dark side of progress

Presently, the rapid emergence of constantly changing forms of communication is the norm. One of the most visible of these developments is *Web 2.0* — a set of economic, social, and technology trends that collectively form the basis for the next generation of the Internet. The goal: a more mature, distinctive medium, characterized by user participation, openness, and networked capabilities. One consequence: The scope of what cyber-criminals can do has opened up and left no boundaries. *Ack. We're all doomed!*

Okay, panic aside (for now, anyway), here are just two examples of why Web 2.0 is an open challenge (so to speak) to effective data security:

- ✔ **Unsecured, multiple-user technologies abound.** Examples are *wikis* (collaborative information projects) and *blogs* (online diaries), along with *services* like Flickr (sharing pictures) and YouTube (sharing videos) are prime examples of how the Web has evolved to bring about increased community participation. What these services really do is bring about freedom of speech to the masses? Unfortunately, though the masses include *the good*, they must inevitably also include *the bad* and *the ugly*.
- ✔ **Web 2.0 technologies rely heavily upon Web services.** Web services are designed to support interoperability between hosts over a network. But in the rush to develop Web services, the underlying Web applications

that use them aren't receiving as much security auditing as traditional client-based applications and services. Furthermore, the policies and procedures for using these new services haven't kept up with the technology and the working practices that go along with it. As a result, threats to confidential information are on the rise.

But even before Web 2.0 is fully implemented, the IT risks that go along with it are entrenched. The next few sections take a closer look at these risks.

The business of cyber-crime

With more people going online all the time, the latest security-threat reports from the IT industry show a worrisome shift in attackers' behavior, motivation, and execution over the past five years. Malicious hacking isn't just an obnoxious prank anymore. Today's security-threat environment is characterized by an increase in data theft and data leakage, and in the creation of malicious code that targets specific organizations for information that the attacker can use *for financial gain*. Attackers are becoming more "professional" — even commercial — in the development, distribution, and use of malicious code and services. Figure 1-2 shows how the same processes used to develop commercial products are now used by cyber-criminal gangs to bring new "products" efficiently to market.

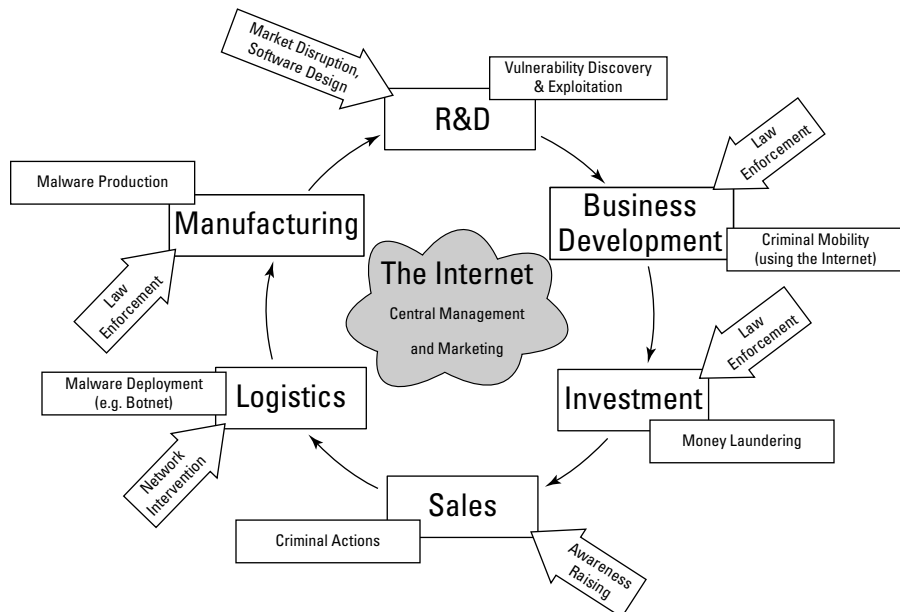


Figure 1-2:
Industrialization of e-crime.

There is an *underground economy* where servers are used by criminals to sell stolen information, usually for later use in identity theft. This data can include government-issued identity numbers, Social Security Numbers, national insurance numbers, credit cards, bank cards, personal identification numbers (PINs), user accounts, and e-mail address lists. And the bad guys are selling it all, at bargain-basement prices, to other bad guys.

Target: your information

Security vendors report a rise in threats to confidential information. Of the top 50 samples of malicious code, two thirds threaten confidential data in some way. So attackers are continually refining their attacks, or enhancing their number or quality, to get what they're after: personal information, which means money. If you were a cyber-criminal, who would you target? If you zero in on individuals, you might get lucky and get one person's information. If you set your sights on a business — where you can potentially get *millions* of people's information, or even governments — where the haul can be information belonging to tens or hundreds of millions of individuals — which target is more tempting? One guess.

Spam continues to rise as a percentage of e-mail traffic, extending a long-observed trend. But it's more than just a tacky nuisance these days. Increasingly, spam is part of coordinated attacks that also use malicious code and online fraud, including data theft. A prominent example is *phishing*, a type of *social engineering* (essentially lying in order to steal) that uses a plausible pretext to lure e-mail recipients into sending valuable personal information to cyber-criminals. The information might be PIN numbers for bank accounts (guess who's making the next withdrawal), your mother's maiden name, or your date of birth.



It only takes a name, address, and date of birth to fake a passport application, open a bank account, or obtain a driver's license in your name — opening a new floodgate of identity fraud.

As enterprises increasingly adapt to the changing threat environment by implementing better security practices and creating in-depth strategies for defense, attackers respond by changing their techniques — sometimes reviving old approaches, sometimes inventing new ones:

- ✔ **More application-targeted malicious code:** Increasingly, these attacks are aimed at client-side applications, such as Web browsers, e-mail clients, word processors, and spreadsheets — any of which can open untrustworthy content downloaded by a network client.

- ✔ **More social engineering:** This is an older, non-technical means of compromising security; it shifts the attack activity away from computer networks and operating systems and toward the end-user as the weak link in the security chain.
- ✔ **Smishing and/or SMS (text) phishing:** In this new variant of phishing, the phisher uses SMS (Short Message Service — that is, texting) messages to tell victims they're being charged for services they didn't actually sign up for. They're asked to go to a Web site to correct the situation — a process that requires the victim to enter credentials that are useful to the bad guy.
- ✔ **Vishing and/or voice phishing:** This approach uses traditional e-mail phishing to ask the victim to call a phone number owned by the attacker who can then fake an interactive voice-response tree — including hold music — that extracts information while lulling the victim into a false sense of security. Cyber-criminals love voice-over-IP (making telephone calls over the Internet, also called VoIP) because it makes the attacks so economical — the calls are free or cost a few cents.

More connections, more risk

The more people work online, the more opportunity exists — for doing business *and* for committing cyber-crimes. Data leakage and identity theft have grown to epidemic proportions worldwide over the last two years. They affect everybody, and they're hard to detect until it's probably too late. Such fraud may account for as much as 25 percent of all credit-card fraud losses each year. For the criminals, identity theft is a relatively low-risk, high-reward endeavor. Issuers of credit cards often don't prosecute thieves who are apprehended; they figure it isn't cost-efficient. They can afford to write off a certain amount of fraud as a cost of doing business.

Most victims, whether individual or corporate, don't even know how the perpetrators got their identities or other sensitive information — *or how they managed to lose the data in the first place.* (Hint: There's a leak somewhere.) Companies that have lost data often have difficulty answering some basic inquiries:

- ✔ Describe in detail the categories of information compromised from a lost company laptop (for example, name, address, phone number, date of birth, driver's license number, or other personal information).
- ✔ Describe all steps that your company has taken to track down and retrieve the personally identifying information.

- ✔ Identify all steps taken to contact and warn consumers that their information may have been compromised.
- ✔ Provide an outline of the plan that will prevent the recurrence of such a data breach — and your timeline for implementing it.

The extent to which they can't provide these answers is a clue to how much control they've lost over their data.

How IT Risk Affects Business Risk

Without electronic information, business would cease to function — which is why *data loss is the biggest risk that businesses face in the twenty-first century*. Reducing that risk means meeting a daunting challenge: protecting electronic information. The risk is more intense now, because of two technological developments:

- ✔ More advanced and pervasive communication devices (as described in the preceding section).
- ✔ A massive reduction in the size of portable storage.

Both of these have business advantages — but they also make it easier to get away with more!

As these technologies continue to develop, IT organizations are faced with the requirement that critical information must be readily available for exchange to, from, and about customers, partners, and employees. Security measures have not kept pace; no wonder data leakage is rapidly becoming a major concern for businesses and consumers alike. The sad story of a data leak has become a familiar news item — complete with its embarrassing loss of customer information, potential monetary loss, and (worse) loss of faith in organizations and their ability to protect critical information.

Fortunately, the loss of sensitive information — whether by inadvertent or malicious means — can be controlled. Although information leakage is difficult to plug completely without impeding business processes, it has to be done to reduce the risk of malicious data breaches.

IT risk — buckets of it

All organizations run according to risk. Traditionally this has been limited to financial and operational risks; the operational side of the house didn't

normally consider IT as a major component of risk. But the world has changed since then. Businesses can't run without IT — and IT is under attack, so the risk needs to be broken out and examined carefully. Figure 1-3 shows how IT risk is a component part of the overall risk to the business.

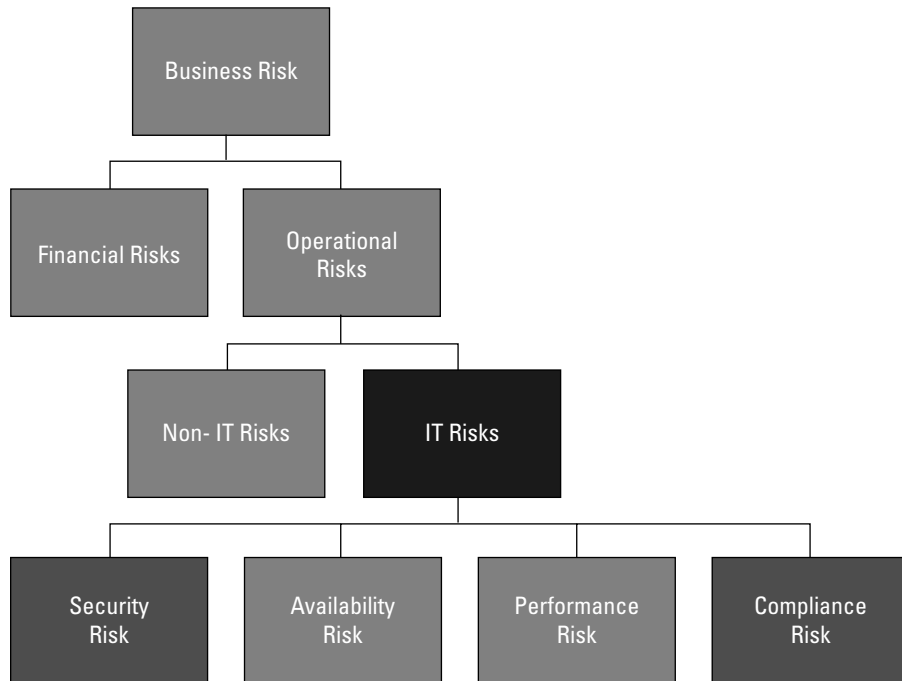


Figure 1-3: IT risk is a big part of business risk.

Note that within IT risk, there are four main *buckets*:

- ✓ **Security:** What are the security risks to the systems — hackers and the like — besides data loss?
- ✓ **Availability:** Will the systems be around when you need them?
- ✓ **Performance:** Will the systems work as quickly and efficiently as you require?
- ✓ **Compliance:** Increasingly, rules and regulations set by industries and governments shape how businesses actually *do* business. Getting it wrong can put you out of business. And guess what — data loss is one big way to get it wrong.



Of the four types of IT risk, two are directly related to data loss.

All organizations could benefit from a greater appreciation of the mounting risks to IT — a firmer, more practical understanding of the risks we're likely to be subjected to. Three areas of consideration stand out:

- ✓ How risk manifests itself to information technology and users (including employees, partners, and suppliers).
- ✓ How organizations should assess and address their level of exposure to risk.
- ✓ How an environment full of communication technologies, sensitive intellectual property, and personal customer information aggravates risk.

The issue is much more complicated than initially thought.

Until recently, organizations always assumed that they must somehow to be protected from *outside* threats. *After all*, they reasoned, *our employees are all good, right?* That's as may be — but these days, the whole concept of an organizational boundary is old hat — in effect, no longer valid — in even the most limited of organizations. This makes any kind of risk assessment tricky. Although IT risk management is becoming increasingly important to all organizations, creating a full-fledged, ongoing program takes time. But it isn't a bad idea to kick off this process yourself — and we're here to help:



- ✓ To be successful, you need senior management on your side so the effort gets decent support.
- ✓ IT department heads must talk with business people and vice versa.
Excruciating? Yes, we can appreciate that — but both the functional managers and the IT administrators must be able to review business operations, workflow, and the technology that affects data loss. And not just once. You'll have to keep this dialog going . . .
- ✓ Periodically you have to carry out thorough security reviews to analyze changes to manage new and unseen threats and vulnerabilities created by changes in business processes, and to determine the effectiveness of existing controls.



Departments whose units handle or manage information assets or electronic resources should conduct regular, formal risk assessments. A risk assessment must

- ✓ Determine what information resources exist
- ✓ Identify what information resources require protection
- ✓ Help IT and the business to understand and document potential risks from electronic or physical security failures

- ✔ Identify issues that may cause loss of information confidentiality, integrity, or availability
- ✔ Provide management with appropriate strategies and controls for the management of information assets.

Although getting this juggernaut underway looks like a daunting process, it isn't rocket science. You can start with simple procedures — say, to start reining in the security of end-user laptops and desktops — or researching and listing best practices for protecting restricted data, or perhaps working out what your organization considers *restricted data*. The problem, in a hectic, 24/7 world, is that you have to make time for all this — and if you're in IT, senior management may be struggling to understand why you exist at all if you aren't directly generating income.



Actually you're *protecting* income. Here are some reasons why:

- ✔ Although IT professionals agree with consumers about the severity of data-leakage incidents, they may underestimate their frequency.
- ✔ IT professionals expect IT incidents to occur about once a month; if the preceding point is correct, then these events probably happen *more* than once a month.
- ✔ Work-process issues cause 53 percent of IT incidents — most often because no process is in place to manage the incident.
- ✔ IT risk management is more than a defensive exercise — it identifies trade-offs among risks, costs, and controls for confident, risk-aware pursuit of opportunities. (Hint: Opportunities generate income.)



From a career-enhancement perspective, all this is great news. You have no doubt heard of the CIO (Chief Information Officer), but new roles are being created, such as the CISO (Chief Information Security Officer) and CIRO (Chief Information Risk Officer). These roles are becoming prevalent in large companies; before long, they'll make it into smaller ones. If you're the person who understands the problem *and can fix it*, then it may be time to recommend that your company needs a CISO or CIRO — and you know just the right person for the job: You. Just do your homework first. (But you knew that.)

Electronic records — incoming!

There's an information tsunami on the horizon. CIOs in 2009 are under increased pressure to deliver business growth, but complexity and tight budgets are still the enemy. But if one of your basic assets is at risk, it makes just as much sense to focus on data storage and data security — you've got to

get a handle on data loss, leakage, breaches, all the places data is wandering away into the wrong hands. It's out of control and growing, but hard to put in front of a CIO who's looking to trim costs, migrate to Linux (or to Windows) or not, drag the company into virtualization, develop new services or applications, reconsider managed services versus in-house operations, and the rest of the standard IT brouhaha. And now here comes this "little" data issue — that's about to get a whole lot bigger.

There's already a bunch of sensitive stuff bouncing around the ether. It's estimated that 1 in 50 documents contain confidential and/or sensitive information, given that we do everything by e-mail then: If we send a minimum 50 e-mails per day, an organization of some 20,000 strong will create over 10 million sensitive e-mails a year, just waiting to be stolen. And there are between 35 and 60 billion e-mails sent worldwide — *each day!* That's 700 million that are sensitive *every day* (best-case scenario) — that's 255 billion potential targets per year for a skilled criminal (still best-case scenario). However you look at it, this is a massive problem — and one that needs to be resolved. Meanwhile, as individual users become an ever-larger source of information, here come some more scary statistics



- ✓ Percentage of companies citing employees as the most likely source of hacking: 77 percent.
- ✓ Annual growth rate of e-mail spam message traffic: about 350 percent (estimated 2006).
- ✓ Average number of spam e-mails delivered every 30 days: 3.65 billion (estimated 2006).
- ✓ Average size of an e-mail message in 2007 (estimated): 650 KB
- ✓ Percentage of all e-mail traffic that is unwanted: about 84 percent (estimated 2006).

Even if 70 percent of the digital universe is generated by individuals, most of this stuff will be handled along the way by an enterprise, businesses, public services, governments and associations: could be on a laptop, USB key, CD, phone, PDA, iPod via a network, stored in a data center, or a hosting site, across wireless or IP network, or Internet switch, on some storage or even more likely in a backup system. This means that organizations must take responsibility for security, privacy, reliability, storage and compliance for an estimated 85 percent of all the information. Or to convert to numbers, a mere 840 billion gigabytes of information.

Getting the Whole Picture

So how do you get the *holistic* (big-picture, everything-accounted-for) perspective you need if you're going to bring your data under control? You could get hold of a risk-assessment tool to identify your assets (as well as the risks to those assets), to estimate the likelihood of security failures, and to identify appropriate controls for protecting your assets and resources. The problem with these tools is that they often have an inclination toward the technology that the particular tool vendor is touting. Worse than that, most of the tools are aimed at the world as it was yesterday — back when it wasn't front-page news to lose a laptop or have a CD-ROM vanish in the mail. Too often, the tools miss one of today's unpleasant realities: Losing a laptop can do more damage to the reputation of a company than losing a whole data center.

Knowing and controlling what you have

The toughest part of protecting the data is *finding* it. If you don't know where it is, how can you protect it? Subsequent chapters in this book help you achieve this — and much more. If you want to jump-start the process, then you're probably better off trying to find some kind of discovery technology. SRM (Storage Resource Management), for example, may be a bit old-fashioned but it can still discover the file types you have in storage. A more recent technology, DLP (Data Loss Prevention), analyzes the data it finds — and can identify and protect confidential information on file servers, databases, collaboration sites, e-mail systems, Web servers, and other data repositories — such as (yes) laptops. This kind of technology can discover and create an inventory of confidential data stored on laptops and desktops, as well as help prioritize the high-risk areas of data storage.

When an organization knows what it has and where, it can then monitor (or prevent) downloading or copying as needed — both internally and externally. Data being copied (for example) to those handy keychain-size USB devices, burned to CDs or DVDs, downloaded to local drives, sent via Web mail, instant messaging, or peer-to-peer networks and generic TCP — can all be monitored and controlled.

A one-size solution does not fit all

It isn't good enough just to motor down the technological route in search of instant data-leak prevention solutions. Too often, it's thought that technology will solve all problems — to which we can only say, *Dream on*. Often technology, especially when it's applied badly, makes the situation worse — unless

you've considered all the options. A much wider approach is needed, taking into account such vital data-management activities as these:

- ✓ Creating data-protection policies
- ✓ Classifying your data
- ✓ Organizing data storage into tiers
- ✓ Archiving your data
- ✓ Encrypting your data
- ✓ Digital rights management
- ✓ Discovery of confidential data
- ✓ Applying data policies consistently

Technology by itself can prevent small-scale stuff — say, keep an engineer from copying confidential CAD diagrams to a USB stick, or prevent a call-center representative from inappropriately copying the customer database to a CD-ROM or DVD. Technology can even manage offline machines and remote office systems. And it can give on-screen warnings and notifications to employees who attempt to violate a company's data-leak prevention (DLP) policies. What it *can't* do is manage the growth and development of the cyber-criminal's arsenal, or catch and correct the inconsistent practices of the end-user.



Much of what we do that's called "user error" happens simply because we don't know what we're doing wrong. One more thing technology can't do: Write the policies and procedures in the first place.

A mind map of data loss

The subject of data leaks is huge. You might think it impossible to put on a single page — but we have: Figure 1-4 shows a mind map that provides a bird's-eye view of data loss. In essence, this diagram shows all the major components that make up the data-loss problem. Each area is then subdivided further. It's an example of a holistic view — a Big Picture.

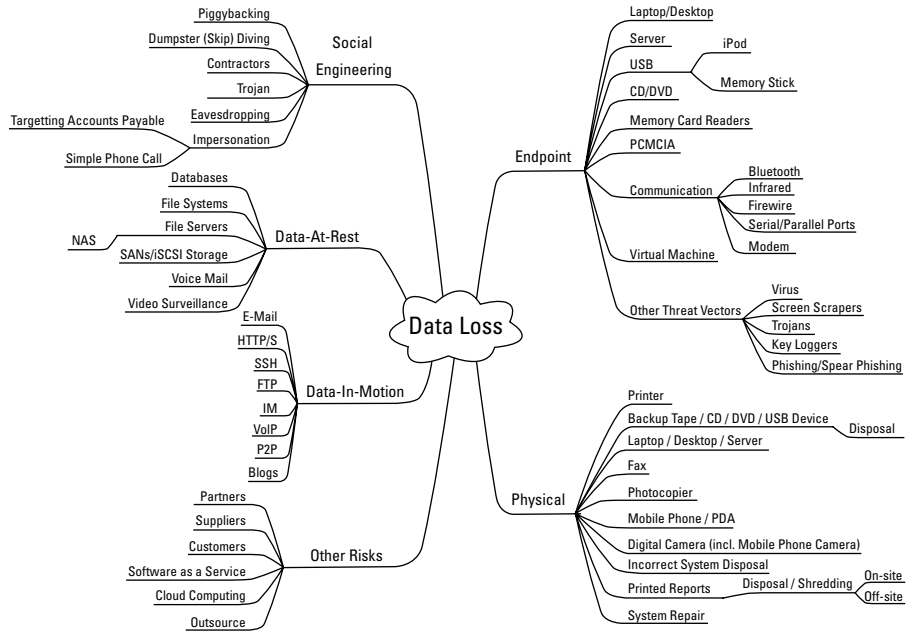


Figure 1-4:
A mind map
of data loss.

A specific data-loss solution can be helpful in multiple areas, so when you start looking at the problem, look at *your* Big Picture — and at how you maximize your investment while minimizing the risk.

