- "It is estimated that the amount of medical knowledge doubles every two years."—*JAMA* 275 (1996): 1637–1639.
- "Knowledge is doubling every 3–5 years."—*Daily Motivator,* Mar. 27, 1996.
- "At the current rate, the entire body of scientific knowledge will double every two years."—*American Druggist,* June 1996.
- "The total of all printed information doubles every five years."—*Professional Reading Guide for Educational Administrators,* Feb./Mar. 1997.
- "Human knowledge is doubling every 30 or so years."—4th Hong Kong (AsiaPacific) Medical Informatics Conference, Oct. 1997.
- "The reservoir of world knowledge is doubling every two years."—Kuala Lumpur *New Straits Times,* Dec. 2, 1997.
- "Human knowledge is doubling every 10 years."—London *Times,* Mar. 16, 1998.
- "The store of human knowledge is doubling every five years."—*Journal of Commerce,* Mar. 26, 1998.
- "Recorded knowledge is doubling every 15–20 years."—Los Alamos National Laboratory, *Research Library Newsletter,* Jan. 1999.
- "Scientific and technological knowledge doubles every three to five years."—Tokyo *Daily Yomiuri,* Jan. 25, 1999.
- "In 25 years, knowledge will double every three months."—*San Jose Mercury News,* Sept. 15, 1999.
- "All printed knowledge doubles every five years."—Kuala Lumpur *New Straits Times,* July 20, 2000.
- "The total of all printed knowledge doubles every four or five years." —Jack Trout, *Differentiate or Die* (Wiley, 2000).
- "The amount of available knowledge is doubling every 18 months." —Bayer Group Webzine, Sept. 12, 2001.
- "Advances in technology and technique now double the total information astronomers gather annually."—*Washington Post,* Oct. 30, 2001.

I have long had serious reservations about statements such as these. Let me explain why.

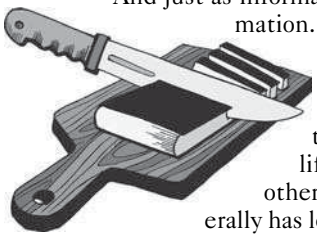
Data, information, knowledge

My first concern is simply the lack of precision about what exactly is growing. While most of the statements speak of "knowledge," I discovered the majority of them while searching for "information." They were taken from speeches and articles about "The Age of Information" or "the information explosion," yet they cite how rapidly knowledge is growing. Information and knowledge, however, are not the same thing. (Nor is either the same as data, another term people sometimes confuse with them.) We need to understand clearly the distinctions among these terms.

Pieces of *data* are things lying about. They can be natural phenomena—the temperature of the air, the size of a seed, the weight of a lion, the composition of the soil, and so forth; or they can be man-made—prices of goods, sports scores, dimensions of buildings, and so forth. There seems to be no end to data.

Information is more than data. (As Clifford Stoll in his book *Silicon Snake Oil* [Doubleday, 1995] has observed, "Data isn't information any more than fifty tons of cement is a skyscraper.") It has been defined in many ways, but two of my favorite definitions are "facts without context and therefore without priority" (Wendell Berry, *Orion Online,* Oct. 30, 2001), and "the meaning that

someone assigns to data" (Peter J. Denning, *The Invisible Future* [McGraw-Hill, 2002]). The critical aspect is that for information to exist, a person must gather and ponder data and arrive at some decision that usually results in an action being taken or at least a choice being made. This action or choice may be immediate, or the information may be filed away for future use.



And just as information is built from data, so is *knowledge* built from information. Knowledge is information that has been internalized, that has become "mine" and is different from "yours." It is what results when I gather and ponder information, evaluate various conflicting ideas and experiences, then apply it to my circumstances to make it part of my life. It is also something that rarely results from anything other than sustained effort or direct experience. Thus, it generally has longer staying power. It is more permanent, less open to the winds of change.

Knowledge is also difficult to communicate. When you try to share with me what you know, it comes to me not as knowledge but as information or perhaps even as data, if there is distance between us either in time or in the number of intermediaries who have translated, reworded, or summarized the ideas along the way. I must recreate my own knowledge with the information I have received. Thus in the quotes above, the term "printed information" is acceptable, but "printed knowledge," "recorded knowledge," and even "available knowledge" are more difficult for me to accept.

One way to illustrate the differences in these concepts is to picture yourself driving in your car toward a railroad crossing on Main Street in your town. The tracks are data: they are just there. The times of departures and arrivals are also data.

Information is formed when you see the tracks, read the timetable, and recognize it as the correct one for your location. You then combine this with data such as the day of the week, the time of day, and the fact that you plan to be in town several times this week.

Knowledge is the realization not only that the train arrives at 12:05 on Tuesday afternoon but also that the day is Tuesday, the time is 12:04, and you are approaching the crossing. All the data and information you have gathered and considered now combine to become valuable to you. (Wisdom, of course, is to stop your car and wait for the train to pass.)

All this is probably moot, however. The people who made the doubling statements likely never thought about my concern. For them the terms are almost certainly interchangeable.

I have no doubt that the amount of data is growing at an astounding pace, as we discover more about our natural world and create mountains of artificial data in our man-made world. If these statements spoke of "data" instead of "knowledge" or "information," I would have a much smaller quarrel with them. The fact that they do not discriminate between the terms is a serious weakness they all share.

Measuring or guessing?

My second reservation with these statements is their significant disagreement as to exactly how rapidly information/knowledge is growing.

The differences cannot be explained because of the years in which the

words were written, as if the pace of growth were accelerating. The two quotes about "scientific knowledge" illustrate this problem: It is doubling every two years in 1996 but every three to five years in 1999. It is also important to note that the extremes are separated by only two years: We have 30 years' doubling time in 1997 but only 12 months doubling time in 2001.

Similarly in 1995, different statements have this doubling occurring within years, months, and even days. Either we are not all using the same instrument to measure growth, or the tool we are using is inaccurate. It is also possible that we are not using any measuring instrument at all but only making educated guesses, based on limited observations of data that is sketchy at best.

Furthermore, in order to claim that something is growing, we must not only measure it now but must have also measured it sometime in the past using the same (imprecise) tool. Has this been done? Yes, but rarely, since there are few situations where accurate measurements are even possible. One is to count the number of publications in a particular field and calculate how rapidly it increases. For example, we can determine how long it took Chemical Abstracts to index its first millionth entry, its second millionth, and so forth. Because fewer years elapsed between each milestone, it is reasonable to conclude that the number of publications in the field of chemistry is growing. But the data simply does not support the contention that any sort of doubling is occurring very quickly.

If we define "doubling" as the total number of items cited since Chemical Abstracts began keeping track, the data shows that the most recent doubling took about 17 years, from 9,904,000 items in 1983 to 19,754,000 in 2000. If we define "doubling" as the number of citations listed each year, the data shows an even longer period: In 2000 there were 725,195 citations, but we must go back to 1975 to find a figure about half this size. Thus, the data reveals periods of 17 or 25 years, not two or three or five, as these experts claim. (A similar analysis of Biological Abstracts showed that the literature of biology increased by about 80% between 1984 and 1994—a faster pace than that of chemistry but still far from doubling in just a few years.)

Even if we acknowledge that the pace of scientific publication is accelerating, can we conclude that all knowledge is growing at a similar rate? Such an extrapolation, I believe, is unjustifiable.

Documentation

The third factor that erodes the value of these statements is that the authors provide no documentation for their claims. One statement does mention "those who calculate such things," and others (not cited here) occasionally make passing reference to the number of articles or books published in a year, or some similar bit of data that could be verified (or not). But every time I have tried to uncover the source of a growth number, I have run into a dead end.

For example, a scholar was quoted in *American Libraries* as saying that "the amount of electronic information is doubling every 60 minutes" (Aug. 2000, p. 34). I contacted him and asked for his documentation. I learned that what he was really talking about was the number of electronic records generated everyday, including short-lived banking and business transactions, and satellite downloads of weather and military surveillance data that are not really "information." He had gotten his figures from a colleague, who in turn had read them in a 2000

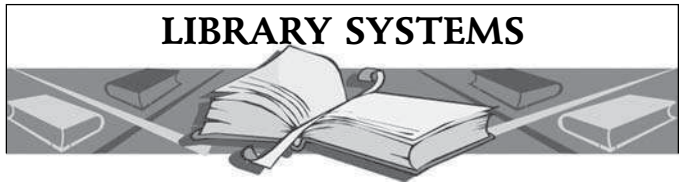
Collaborative Electronic Notebook Systems Association report, the methodology of which was described as being "at best, back of envelope."

What originally appeared to be a rare case of a well-documented study of the growth of information proved to be a mirage. (I should also note that the source where I first found the statement neglected to mention that it was a prediction, that the growth was not happening today, but was expected to occur "within 10 years.")

It is undoubtedly true that data is being discovered and generated at a fantastic pace and that the amount of information we deal with daily is prodigious, but none of the statements about the doubling of knowledge have enough real evidence to persuade me that it is increasing at anywhere near the rate many people are claiming.

This leads me to conclude that the "Knowledge Is Doubling" monster is, to a large extent, more a mirage than a reality. Some sort of creature is lurking under the bed, and it is probably large, but it is not so huge that we should be paralyzed by its mere shadow. Shining a flashlight on it reveals its lack of substance and shrinks it to a manageable size.

SOURCE: Martin H. Raish, "Shining Some Light on the Monster under the Bed: A Closer Look at the 'Doubling of Knowledge,'" in Martin H. Raish, ed., *Musings, Meanderings, and Monsters, Too* (Lanham, Md.: Scarecrow, 2003), pp. 145–156.



Trends in integrated library software

by Marshall Breeding

THE MAINSTREAM LIBRARY AUTOMATION SYSTEMS available in 2004 can all be considered fully developed, at least regarding the core ILS functionality. Although each product has its strengths and weaknesses relative to its competitors, all the products include the major modules and sophisticated functionality within each module.

As the table on page 431 illustrates, some systems on the market benefit from more than 20 years of product development and evolution. Although some systems have evolved significantly in their underlying components, the features available in the system have grown steadily. Even the newest of the current systems have been on the market for more than five years.

Having developed previous generations of automation software, some companies offer history and experience that exceed the development history of their latest product.

The expected feature set of library automation systems has been evolving even longer. The basic methods and procedures associated with library automation were established in the precursors of these systems that have since become extinct.

Few multiuser library automation systems have entered the market in the last five years. The Destiny product released by Follett Software Co., which