Who Collects Data and Why?

By Bob Martin

GOVERNMENT AGENCIES

The most politically charged, and therefore the most discussed, databases have in the past been those maintained by the Federal government: e.g., those of the Census Bureau, the Social Security Administration, all branches of the military, the FBI and other security agencies, and myriad lesser-known Federal agencies which inevitably compile extensive databases of use to someone.

Less known to the public, but equally important, are various obscure computer files established by the 50 state governments and thousands of local government entities. Unlike the well known Federal databases, many state and local data files are in the public domain—available for free and for a fee to any interested party. The compilation and sale of such files is a growing business for database companies.

Important state and local data archives—often computerized and readily accessible—include:

- Public registries of births, deaths, marriages and divorces, real estate property sales and assessments, and so on.
- State department of motor vehicles listings of automobile ownership and accidents.
- Local tax records (such as property tax payments), lists of residents and voter registration lists.

BANKS AND INSURANCE FIRMS

Commercial banks, insurance companies of all kinds, and health care insurers in particular all assemble massive amounts of data on the finances, health status, and other personal business of nearly all Americans. While this data is not public information as such, it is very often available for a fee to outside parties.

Even the most confidential kinds of information are often available with little difficulty to those who know how to gain access: for example, certain health insurance firms will provide employers with detailed monthly reports on the health care expenditures and diagnoses of their employees (e.g., John Jones saw a psychiatrist four times for depression last month).

Bob Martin is the administrator for a large health care organization.
Other commonly used criteria include what are called the “re-cency, frequency and amount” of each list member’s purchases during a certain period. For example, you could purchase a mailing list whose members had each bought at least $50 worth of books on at least three different occasions during the past six months.

Various databases firms specialize in compiling and continuously updating lists based on certain public sources. For example, one firm tracks every house sold in the nation, using local property tax and courthouse records as raw material. The firm compiles not only the buyer’s and seller’s identity but the sales price, location, and other data. This data is valuable to many other firms—for example, to carpet and furniture retailers for whom recent homebuyers are prime customers.

**HOW IS THIS DATA USED?**

As indicated above, all of this data is compiled, sold, and used for many commercial purposes. The most obvious of these purposes is direct marketing—especially the sale of retail consumer goods by mail and by telephone.

As computer memories become ever-larger and the speed and complexity of computer capabilities grow, the cross-tabulation and synthesis and the many types of database mentioned above is becoming remarkably sophisticated.

One example of such sophistication is the Claritas Corporation, a database and marketing firm which has analyzed the U.S. population on the basis of zip code and sub-area classifications. Each small geographical zone—down to the neighborhood level—is given one of about 40 tag names, such as “Shotguns and Pick-up trucks,” “Brains and Money,” and so on. These categories are based not only on income, but other variables such as occupation, family size, property values in the area, spending habits, and so on. Claritas continuously updates and sells this data not only to direct marketers but to many types of businesses deciding where to initiate new ventures.

Besides direct marketing, other uses of data include the following:

- Deciding where to locate new stores, malls, distribution hubs, office buildings, residential developments and other enterprises.