Do Data Warehouses Challenge Fair Play?
By H. Jefferson Smith

O ne of the most popular concepts in information technology these days is data warehousing, which stores a company's data in a central repository. The information in the database is updated frequently and is made available to the firm's managers and employees for planning, marketing, and decision making.

Data warehouses are designed to support online analytical processing and data mining. These technologies have been described as akin to turning 100 statisticians loose on your data at the same time.

Many kinds of business questions can be answered through these technologies. You can find and track customers, analyze their behavior, segment a customer base, customize products, model past attrition behavior (thus reducing past customer defections), and refine a business strategy by massaging the warehouse data.

For example, one consumer credit company has a data warehouse that contains almost 1,000 attributes per customer. The database is so large that updates take more than 48 hours and rely on 50 different feeder files. But the payback is also large: Analysts are making more than 200 queries per day, and in-depth reports on spending patterns and demographics are available to the company's marketers. The analysts and marketers have also used the warehouse to generate targeted mailings to customers.

Nevertheless, along with the potential benefits of data warehousing come some serious considerations about fair play in the use of customer data. The various issues that arise depend on whether an organization's customers are other businesses or individual consumers.

It's Just Business

Almost every company has relationships with other firms. Some are suppliers that provide the company with products or services, while others distribute or purchase its products and services. In addition, a company has relationships with individual consumers who buy its products—either directly or through a distributor or retailer. Thus, when an enterprise warehouses data about its customers in business-to-business transactions, a corporation should think about what constitutes fair play from the perspective of several different players: the company that is its direct customer, the firms supplying that customer, and the firms or individual consumers buying that customer's products.

In general, the company with the data warehouse should follow a two-edged principle. It is fair to use the customer data to deduce ways in which the relationship with this business customer (or other potential customers of this type) could be strengthened.

It would be fair, for example, to create statistical profiles of current customers based on the warehouse data and to use those profiles to deduce which market segments might be most appropriate for future targeting. It would also be acceptable to conclude which additional products or services would be most appropriate for current customers and to focus special attention on creating and marketing those products or services to those customers.

On the other hand, it is unfair to use the customer data in any of the following three ways: First, it is unfair to do anything that might harm the customer's relationships with any of its suppliers or customers. Suppose, for example, that company B, after careful scrutiny of its warehoused data, realized that most of the purchases customer C made were being resold to one of C's clients, D. Obviously, both B and D could benefit if D bought its products directly from B and bypassed C. B could charge D a price that was higher than what it charged C but lower than what D paid C.

Though this scenario appears economically efficient, B's contacting D to suggest such a deal would constitute gross unfairness to C. B would be using data about its relationship with C to undercut C's position with D.

Second, it is unfair to use customer data in any way that intrudes on the customer's proprietary know-how. Suppose that its data warehouse gave company E knowledge about the specific methods and techniques that one of its customers, F, was using to design and produce its products. It would be unfair for E to reveal this information to others or to use this knowledge to take advantage of F in future negotiations.

Third, it is unfair to use customer data to change an industry structure if that change is detrimental to any of the firm's customers. Suppose, for example, that company G was a supplier to a number of firms in an industry. This industry has a value system of suppliers, manufacturers, distributors, and consumers.

Several companies are involved in the manufacturing process—from raw materials to the final product—with each firm adding some value to the product.

By carefully massaging its data warehouse, G might discover a new scheme for manufacturing and distributing products that would increase the overall efficiency of the system, reduce the cost of production (leading to greater industry profits and more sales for G), and lead to greater sales and profits for some of the manufacturers (G's customers).

On the other hand, the scheme would hurt the sales and profits of other firms that are also G's customers. Although some people might disagree, I would argue that G had an obligation to protect the interests of all its customers and to take no action that would harm any of them. Since G had the data in its warehouse only because of its relationship with its customers, it would be a betrayal to use that data in a manner that would harm any of them.

Making It Personal

When a company's customers are individual consumers instead of other businesses, different rules of fairness apply because concepts of fair information use at this level are often viewed as a human rights issue. Therefore, issues related to consumer privacy—a concept quite distinct from that of corporations' right to proprietary trade knowledge—quickly come into play.

The general rules of fairness in warehousing consumer data should be the same as those that are becoming generally accepted for other applications that involve personal data:
- Consumers should be fully informed of the intended uses of data before the data are collected.
    - Consumers should be allowed to opt out of any uses they find offensive.
- Data collected from consumers for one purpose should not be used for another purpose without the consumer's permission.

The rules suggest that it will be difficult to begin warehousing consumer data unless some up-front work has been done to ensure that the consumers were fully informed of the intended uses ahead of time and were given an opportunity to opt out.

For example, unless consumers are told in advance that transaction data will be used to assess their spending patterns and create psychographic profiles of their activities, such analysis should not be done. Fortunately, the consumer credit company discussed earlier has engaged in just such a notification program.

Assuming that the analysts have access to a set of "clean" consumer data (data gathered under the policies outlined above), they can proceed to mine the data, classifying consumers as appropriate, targeting specific customers for certain offers, and developing plans for soliciting new customers.

However, a word of warning is in order, based on experiences in the database marketing industry: The results of the mining activities should be carefully evaluated to ensure that they produce no socially negative outcomes or, at least, that the outcomes are grounded in business decisions rather than in unintended discrimination. For example, the targeting of specific residents in one urban area for special purchase offers has been called discriminatory because the offers were sent disproportionally to one racial group and excluded members of other groups.

It seems obvious that the use of data warehousing introduces new ethical challenges into both business-to-business and business-to-consumer relationships. However, the lines are not drawn clearly in all areas, and there is still room for judgment calls on many issues. Therefore, in the interest of fair play, corporate and IT executives who want to take advantage of this technology should pay serious attention to all the issues involved.