enterprise integrity

By DAVID McGOVERAN

Valuing Data, Part 3

Developing a model of data valuation would be futile if we were to attempt to list all the many uses of data in detail. We can reduce the complexity of this task by categorizing the data according to certain usage characteristics. In doing so, we'll develop an understanding of the utility and value of each category. When using the model for a particular business, it's important to consider each of these categories. Some categories are mutually exclusive, meaning that a particular data element can accrue value through, at most, one of these categories at a time. Others will be compatible, meaning a particular data element can accrue value from multiple categories more or less simultaneously. This month, we begin to identify a data categorization that will permit us to develop a general model of data valuation in subsequent columns.

The use categories of data can be better understood if we think of data as having a kind of life cycle. At some point, a business will acquire a particular datum. Eventually, the business will no longer have access to the data or will be unlikely to use it. In between, data may be used in various ways. We'll consider six states, each corresponding to a category of data utility. As we'll see, the states of this life cycle aren't strictly sequential, nor are they necessarily mutually exclusive. This month, we'll define the first three states: acquisition, inventory, and operational:

> The use categories of data can be better understood if we think of data as having a kind of life cycle.

- **Acquisition** — Data can be acquired by purchase, data entry, electronic feed, online capture, and so on. Online applications often incorporate data acquisition during production operations, so a distinct acquisition state may not exist. For the most part, data in the acquisition state contributes only to costs. Each method of acquisition has a wide variety of associated costs. The costs of equipment, software, consumables, facilities, and personnel required to implement, operate, and maintain data acquisition should be included. Furthermore, data cleansing is part of the data acquisition state even if it's deferred. We'll consider the possibility of late data cleansing (i.e., subsequent to use) in terms of a data quality factor.
- **Inventory** — Often, data is stored for some time subsequent to acquisition or between specific uses. By definition, data in the inventory state isn't being used. As such, it contributes only to costs and, like all assets held in inventory, the longer it remains in this state, the greater the costs. Again, the costs of equipment, software, consumables, facilities, and personnel required to implement, operate, and maintain data storage should be included. Backup and recovery costs for maintaining data inventory should be included. It's easy to overlook some of the many forms of data inventory. For example, we must be certain to include data stores on desktops and all the data stored for application integration, including message queues.
- **Operational** — Sometimes, data is used directly in the production of goods and services. If a data element (or group of data elements) is required for production to continue, then it's in the operational state. Online reporting and decision support data are in this state if it's required for production. For example, some historical data is required to make scheduling decisions. Of course, policy determines data requirements of production activities as much as any logical requirements. As with capital equipment, some of the value of the goods and services being produced must be allocated to this operational data. Some uses of operational data influence production decisions, while other uses are collateral and, at best, influence the decisions of other entities. For example, a shipping address is critical operational data, but often doesn't influence any production decision. Later, we'll return to additional analysis of the contributions of operational data to our data valuation model.

Data cost and value accrue from both direct and indirect associations with production decisions and from more administrative, infrastructure, and strategic business activities. Next month, we'll define the remaining three possible states of data: historical, forecasting, and divestiture. Meanwhile, give some real thought to the costs (and value) of the data that assures your enterprise integrity. _eAI_

David McGoveran is president of Alternative Technologies, Inc. He has more than 20 years' experience with mission-critical applications and has authored numerous technical articles on application integration. e-Mail: mcgoveran@alternativetech.com; Website: www.alternativetech.com.