enterprise integrity

By DAVID McGOVERAN

Valuing Data, Part 4

Last month we detoured from the "Valuing Data" series in order to acknowledge the passing of Dr. E.F. Codd, inventor of the relational model, father of the relational and SQL DBMS, and godfather of all technologies that depend on relational and SQL databases. This month we'll continue the series, the purpose of which is to provide guidelines for the development of a cost/benefit model of the financial value of data. This model is based on what I call data's cumulative utility — its actual and anticipated utility over the useful life of the data — and treats data as an operational asset that may have the characteristics of either a capital asset or a consumable asset, depending on its context.

Our approach to identifying that context has been to describe the utility life cycle of data in terms of states. As noted previously, some of these states are mutually exclusive while others permit data to be in a mixed state consisting of two or more fundamental states. Furthermore, although these states are somewhat idealized in order to make the problem of data valuation more tractable, I think you will find that they will allow you to develop a reasonably accurate model of your data's value and therefore to make good decisions about acquisition, retention, and the like. Month before last I defined the three of six such states: acquisition, inventory, and operational. The remaining three states, forecasting, historical, and divestiture, are discussed this month.

- **Forecasting** — Data is sometimes used to forecast trends and aid with both near-term and long-term strategic decisions. The relationships between the analysis and the decisions to be made may be pre-determined as, for example, in inventory management. Alternatively, analytic results may be purely advisory and exploratory as is often the case in, for example, market segmentation research. During the forecasting state, data contributes only indirectly to production. Lack of such data may either degrade or enhance production, but is unlikely to stop it. In some cases, such data is used to achieve a goal of unknown, but presumed positive, value such as improved customer satisfaction. Similarly, forecasting data may be used for gaining a better understanding of the business and therefore improve management. If one takes a reductionist approach, the contribution of such forecasting uses to cumulative data value is intangible (i.e., not quantitative) only because we lack either an understanding of the process involved or adequate means to measure the impact. Data in the forecasting state accumulates considerable costs due to the special processes and skills required. Costs also include computing equipment and software, personnel, facilities, and so on, which may be dedicated to forecasting and planning. While the major costs are relatively easy to compute, the value contribution is more difficult to ascertain and we will return to this topic.
- **Historical** — Eventually, data is stored both as a record of what has taken place and in anticipation of a possible future requirement. Associated with data in the historical state is a probability that it will be used (or reused) for some purpose and the value derived from that purpose. That probability usually decreases exponentially over time, although it may actually increase over time (often followed by a sharp decrease after an "opportunity date") for certain special uses of historical data. Once the costs and value associated with each anticipated possible usage have been determined or estimated, these should be weighted by the corresponding amount of data (usable for each purpose). Some historical data is saved by regulatory requirement, resulting in a cost avoidance issue. Avoided costs contribute to value, and should be weighted by the probability that the cost would have been incurred (had the data not been maintained) when adding to the value.
- **Divestiture** — Businesses divest data, both intentionally and unintentionally. Data destruction incurs costs, including personnel and record keeping. Storage materials such as old tapes may require special handling for disposal. In addition to incurring costs, a business may be able to sell some of its data, thus creating either a primary or secondary product with associated profit and loss potential.

Next month we will examine the effects of degradation (and therefore data quality), depreciation, and appreciation. Until then, ponder this: Too many types of uncorrelated data creates an "information catastrophe," making it impossible to converge on a formal analytical model and therefore repeatable management decisions. Lesson? Control your data junkie impulse. It's a matter of enterprise integrity. BIJ

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