



# Data Integration Center of Excellence

by jonathangeiger

Last month, I described some of the major factors to address when establishing a center of excellence. These dealt with ensuring that the center has the authority to influence actions, that it is properly placed within the organization, that the roles and responsibilities are well defined and that it has the proper technological support. The data integration center of excellence (DICOE) promotes collaboration and best practices for bringing data in from disparate sources. The target may be a data warehouse, an operational data store or an integrated data store for master data management (MDM). In all cases, the underlying objective is to provide a data store with an integrated, enterprise view of the data.

There are two fundamental structures for a DICOE. In one structure, the DICOE is a line organization responsible for data integration, and in the other, it is a support organization assisting others who perform data integration.

## Line Organization

A DICOE is sometimes the group that actually performs the data integration activities. In this environment, the group has three major functions. The dominant and most obvious one is the actual cleansing, transforming and integrating of the data. A related function is the data capture. I separate these two functions because the capture is not always performed by the DICOE. For example, if the operational system environment is outsourced, the outsourcer may extract the data and provide it. Both of these functions place responsibility for moving the data into the integrated data store within the realm of the DICOE. A third function is internal to the DICOE. It has a role similar to the support organization described as follows, except that its target audience is the DICOE itself.

When the DICOE is a line organization, its staff is called upon to perform work embedded within data warehouse and MDM projects. This can be accomplished by placing some project responsibilities within the realm of the DICOE or by having a matrix organization whereby DICOE staff members are assigned to the project. The latter maintains primary responsibility for the project results with the overall project manager, while the former establishes a mini-project within a larger project framework.

## Support Organization

Another approach is to have individual project teams assume the responsibility for the data cleansing, transformation and integration activities. In such an environment, these activities are planned the same as other project activities, and members of the project staff are assigned to these responsibilities. The DICOE supports this organizational structure by providing expertise, templates and best practice information related to data cleansing, transformation and integration.

The group, for example, may be responsible for selecting the extract, transform and load (ETL) tool, for making decisions concerning upgrades, for maintaining the relationship with the tool vendor and for providing internal training and consulting on the tool. In addition, it may provide templates for data profiling and transformation mapping and may even provide a quality assurance role by reviewing related project team deliverables. Another DICOE responsibility would be the maintenance of information on data that has been integrated so that new projects can fully leverage work already performed.

## Choosing the Organizational Structure

There is no single structure that will work best in all companies. Possibly the greatest single factor that would impact the selection is the firm's current practices and culture. Questions to ask include: Does the culture support a centralized group that dictates practices? Does it promote a collaborative approach? Does it promote individualized approaches? Does the firm effectively apply matrix management? Answers to these questions combined with information about previous data integration attempts, the linkage that can be established between data integration and corporate goals, and the degree of executive backing will dictate the approach to be pursued, at least initially.

A company that has successfully applied matrix management, established a linkage between data integration and enterprise goals, and possesses a strong, visible sponsor for establishing the enterprise perspective can succeed with either structure. One that has strong backing and a poor history of achieving the enterprise perspective may be well served by placing the DICOE as a line organization so that it serves as an agent for change. A company that is highly distributed in its management structure will likely require a support structure for the DICOE, and as successes are achieved and documented, it may be able to move to a line-organization structure.

A DICOE is critical to an organization for achieving an enterprise view of data to support data warehousing and MDM initiatives. The group can be structured so that it performs all (or most) of the data integration work or it can be established as a support organization that provides expertise and templates. The best approach for you depends on your circumstances, and you may be well served by an evolutionary approach that begins with one type of structure and migrates to the other.

In my next column, I will describe the center of excellence for getting information out. Unlike the DICOE, getting information out is often a highly distributed function, and the role of the center of excellence must be organized accordingly. 

*Jonathan G. Geiger is executive vice president at Intelligent Solutions, Inc. He may be reached at [jgeiger@intelsols.com](mailto:jgeiger@intelsols.com).*