

The PDFs of *eAI Journal's* Data Integration department are sponsored by:



Data Junction sets the standard for data integration with award-winning technology, customer support and technical services. Right now, more than 40,000 customers invested in Data Junction technology to solve their most pressing integration concerns, making our software the most widely deployed in the world. What integration challenges are you facing? Discover how Data Junction's solutions fit into your integration world. For an on-line demonstration of the power and versatility of Data Junction, contact us at 800.580.4411 or info@datajunction.com

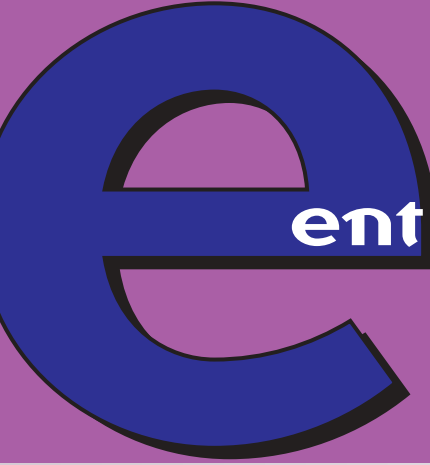


Data Junction Corporation
5555 North Lamar Blvd.
Ste. J-125
Austin, TX 78751
tel: 512-452-6105 ext. 256
fax: 512-467-1331
www.datajunction.com



This article is a PDF version of the one that appeared in a recent issue of *eAI Journal*, the leading resource for e-business, application integration, and Web services.

Integrating the Interconnected World



enterprise integrity



By DAVID MCGOVERAN

Data Integration, Part IX

Although last month presumably wrapped up the data integration series, consider this month an encore. Several conversations and e-mail messages convinced me that a discussion of metadata as related to Enterprise Application Integration (EAI) might benefit readers. Specifically, what priorities does one give to specific metadata and why?

The subject of metadata and methodologies (or any less formal attack on the problem), is enormous. It's hard to cover some 26 years of experience in a dedicated seminar, much less in a column such as this. So I'll simply provide some guidelines and tips in the hope that you'll avoid the mistakes and benefit from the successes of those who have been there before. Remember that these recommendations promote an incremental approach to enterprise integration.

I divide metadata into four categories:

- Type model
- Syntactic
- Semantic
- Historic.

The type model provides the "exoskeleton" in which other metadata is captured and typically evolves as the semantic relationships among data types are understood. Unless you foolishly freeze the model, it can't help but evolve as the business and your understanding of it grows.

Having syntactic metadata will let you test and refine the *functionality* of your EAI solution, since it's this type of metadata that's essential to get an integration server operational. But, as noted throughout the data integration series, just moving data from application to application isn't enough to achieve value.

Having semantic metadata will let you test and refine the *correctness* of your EAI solution. Both syntactic and semantic metadata are essential to successful EAI project deployment.


Finally, having historic metadata will let you maintain your EAI solution since it interconnects past, present, and future understanding of data operations. Unless we understand how data has been used (and abused) in the past, we can neither reconcile the past with the present nor plan the future.

A good metadata repository will capture both the syntax (issues of form) and semantics (issues of meaning) about each "island" of data usage. EAI requires that you have the syntax of both a data source and a data target. A message format,

XML schema, EDI format, data store structure, or API may dictate that data source syntax. The semantics (the subject of most columns in the preceding series) are usually much more difficult to obtain. You *must* focus on capturing *constraints* to get this right. There are many ways to categorize constraints. My preference is the one developed through relational database theory, which readily ties the subject to both abstract data types (a.k.a. domains and object classes) on the one hand and transactions (and so business processes) on the other.

As noted in previous columns, metadata capture can occur incrementally in the context of an EAI effort. If you want to obtain a measurable return on investment from an integration project, it's best to first obtain an understanding of the business processes that are strategic from management's perspective. Integrate the involved business functions and applications first. Segment these into sub-processes that can deliver business value. As you interconnect and deploy each prioritized sub-process incrementally, you must understand only each relevant data source and target. Deliver syntactic metadata first (it's relatively easy and enables the integration team) while semantic metadata is being researched and captured as constraints. Define any necessary semantic transformations before EAI deployment.

Getting an EAI project up and running is only part of the story. Ultimately, if you pursue a data quality objective (as every organization should), you'll be capturing metadata that's directly relevant to implementing EAI and important to maintaining its viability. You'll need to be able to audit data usage and to trace the data's source, and change history (including syntactic versions and semantic refinements) when trying to fix the inevitable problems that arise. You'll want to perform an impact analysis as new applications or other changes are proposed. Historic metadata meets this need.

You can't capture and organize all the metadata pertaining to an entire enterprise before you begin your first EAI project. But you can capture relevant metadata in an extensible repository, driving each project incrementally. That's a great way to make *enterprise integrity* pay. 

David McGovern 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: mcgovern@alternativetech.com; Website: www.alternativetech.com.