

**WORKING GROUP
DYNAMIC INFORMATION INTEGRATION**

BUSINESS RULES IN MAPPING

Wednesday, May 22, 2002

Programs at the Center are possible due to the support of its Sponsors:

Amdocs • Cable & Wireless • Deutsche Telekom • Fujitsu • IBM • SAP • Sun Microsystems

SUMMARY

As a result of a series of programs in autumn of 2001, the Center hosted a Working Group on May 22, 2002, to continue an examination of “Dynamic Information Integration.” An on-going group of professionals in the information and data management arena discussed whether or not the application of business rules to data integration assists in extrapolating high-quality information from data. The discussion took place in the light of:

- The role and importance of explicit models
- Achieving a working feedback loop
- The impact of existing technologies and components
- The state of relevant standards

Participating in the working group were Dejima, Inc.; iSpheres, Inc.; KnowMadic; SAP; Stanford University; Strategies, Concepts & Solutions Lab; TekMetrix Corporation; WhizBang! Labs; and consultants Dennis Andrie, David Brooks, Guy de Bruyn and Nicholas Hoddock.

The scenario presented to the participants was the following:

When integrating systems and data, mapping between taxonomies often becomes necessary. However, in some cases, a static mapping of objects in one taxonomy to objects in another taxonomy does not present all the relevant data for a business process to take place. Rather, business rules often need to be applied during mapping to achieve the greatest accuracy of the information. The following illustrates this situation. A credit check is offered as a Web Service and requires the DUNS number and credit amount as input parameters. An ERP (enterprise resource planning [software]) system with its own way of numbering customers will be calling the credit check service. A mapping between the customer ID of the ERP system and the DUNS number thus becomes necessary. But this simple, static mapping may not be enough to provide accurate information. For example, agreements may exist that a parent company takes over payment responsibilities for certain subsidiaries or different forms of bank guarantees are in place, complicating the necessary mapping.

The group examined this example with the questions:

- Is there a need for a generic solution to combined data mapping with business rules? If so, in what situations is this needed?
- Do systematic approaches to this problem exist?
- How can rules be generated and maintained efficiently?

IT DEPENDS

Participants heard that mapping is not simply mapping. Some initial mappings are static, that is, they occur directly between two pieces of information that are sufficient to describe the business function to be achieved. However, the “it depends” factor weighed in heavily on the conversation. Business rules only describe so many circumstances or combinations of factors. At some point, when things become complicated or intricate enough, the answer will be, “It depends.” It depends on several particulars. “When [the mapping] is more dynamic,” said one participant, “the rules are not so easily expressed. Integration with existing ERPs can get pretty messy.” At that point, these particulars can be described only by the IT person sitting down and changing the coding. Every business-rule change requires a programmer to make the change to the software. “Is each case so specific that we just keep coding?” asked one participant.

The group questioned what cases give rise to there being a need for complex mapping where point-to-point, static data mapping is no longer sufficient for precise information acquisition. Some of those events were identified as the opening up of trade networks, company mergers and the change of ticker symbols. At these points, information integration becomes complex due to the merging of diverse systems and the simple changing of information, such as a ticker change. How do systems recognize one another or merge into one system when these changes occur?

SHOW ME THE DATA

When looking at how data currently resides in corporations, participants noted that only 20 to 30 percent of data is structured in data bases, while the other 70 percent is in documents that require it be picked up through the use of taxonomies. The group identified several challenges including how to model documents, and once that is done, how to integrate these into a given system, how to assign semantics to the modeling and how to keep track of changing profiles. “How do we create a profile-persistent event that kicks in as your profile changes?” asked one participant.

SOFTWARE DEFINING THE BUSINESS RULES

Part of the problem of data management seems to be that those who have defined the business rules for many companies are not the business professionals, but the technology experts. “The key issue,” opined one participant, “is that we have allowed the programmers to put together the business rules.” The group agreed that this phenomenon has occurred as a result of businesses not having defined and known what their own business rules are. Software such as SAP’s enterprise resource planning offerings are often the only framework many companies have for describing business rules. And they use them for such without having first established their own business processes, definitions and procedures. When software is too flexible, different divisions of a single company customize it differently from one another, increasing the communication problem. The group agreed that there is a need for more simple, easily scripted rules than those expressed through C++ and COBAL. They also talked about the struggle to maintain rules and rely on standards as a path to better rules.

Standards are part of the solution, but they do not solve the problem of relevant information. Information is domain specific. One needs the domain expert to describe how the information should all be put together. So, if the information resides within human resources or finance, language will have different connotations in each of these domains. What is needed to address this issue, suggested one participant, are standard company and people descriptors, for example.

LESSONS LEARNED

The topic proved to be complex, with the concept of the context dependency of mapping taking center stage. One participant suggested that perhaps the issues are too application-specific for simple solutions. Another stated strongly that breaking mapping out as a separate module is not the right approach to the problem. “We need to be looking at the higher level problem,” he suggested, “which involves the context present at the time you are trying to do the mapping. That is a solvable problem, but it is not necessarily amenable to a general domain-independent solution.” This opinion met with some resistance as another participant opined that if mapping cannot be treated independent of the domain, it would mean the ultimate doom of achieving general dynamic look-up on the Web. The hope is that for the major part of data, there is a generic solution. And the consensus was, indeed, that 80 percent of the information should be negotiable, following the 80/20 rule. Nevertheless, the context dependency of the mapping emerged as a major revelation to some participants. It also became clear that, in order to automate processes, they need to be well understood. As a result, professionals in this arena may always be behind the curve on what they hope to accomplish and may never achieve 100 percent automation. Another participant expressed that, although it is good to have an explicit model and the tools to help evaluate that model, the model is not the magic bullet. Because of the context of knowledge, people will need to be involved to describe the rules and the boundaries. To do that, rich tools are required. XML is only a beginning. What is needed is a pallet of sophisticated, standard tools, such as sets, that most people use with the same infrastructure and the same framework to work collaboratively. (An ongoing project at Stanford is trying to build this infrastructure. Information is available at <http://www.tap.Stanford.edu>).

*700 Welch Road, Suite 250 Palo Alto, CA 94304 • Tel: 650.462.9800 • Fax: 650.462.9818 •
www.worldinternetcenter.com*