Leveraging Data Models to Support Development of Conceptual Enterprise Models

Enterprise Information Architecture products are often largely based on database designs. These are useful as bottom-up grounding but need conceptual analysis to provide an integrated view of Enterprise data such as demanded by data interoperability to support flexible business processes.

Enterprise Architecture Use of Traditional Information & Data Models - Enterprise architecture frameworks are typically built on more traditional architecture methodology and formalism such as found in IT applications and database design. Some enterprise architecture efforts represent aggregations of products from these “lower level” approaches to leverage existing model and analysis. At its simplest this is a “federated” approach to enterprise data modeling and integration, based on DB schemas view integration. However, the scope & depth needed for enterprise models requires some extensions to these methods, especially in the area of conceptual analysis & modeling. Conceptual business analysis has long been identified as an important tool for the integration of separate databases (Ceri & Pelagatti, 1986) and data engineering as a whole. However, it does not often receive proper support in DB design and even less in the necessary integrated sense of enterprise modeling to support data integration and interoperability. The need for something beyond a record-based ER model can be seen in Date’s (1986) summary on the semantics of databases:

"most database systems, relational and otherwise, really have only a very limited understanding of what the data in the database means. They typically 'understand' certain simple atomic data values and certain many-to-one relationships, but very little else. Any more sophisticated interpretation is left to the user."

ANSI 3 Schema Approach - An initial idea of a conceptual level modeling was described as the front end of the ANSI three-schema architecture approach to engineering databases (Tsichritzis and Klug, 1977). In this approach, each modeling level is independent of the others and, in particular, can represent information in different ways a suited to the model’s purpose. In the ANSI methodology database development starts with Level 1, the Conceptual Model, (also called the Conceptual Schema) created to formalize data requirements by means of business analysis. We might call it a conceptual database model (CDM) to distinguish it from the broader enterprise conceptual model (ECM). A CDM documents the semantic relationships in the smaller scope of a database. In a conceptual view the relation between data elements has little or no consideration for structure that logical and physical models will provide. When formed by a valid analysis, an ANSI conceptual schema guides and constrains the design specification of a database and its operation based on key business entities and relationships which are independent of any technical implementation. It is, however, by no means a comprehensive or even systematic view of enterprise concepts. A series of independently designed CDMs might capture a major portion of the enterprise business area concepts, but particular concepts are often in conflict. This arises because particular database functionality is scoped to some subset of business function(s). Level 2, the Logical Model (aka the External Schema), converts the general entity-relations of the

1 In database design a “schema” defines how data is organized and constrained. A model may do this, but can be thought of as a broader approach.
conceptual model to a DB model relation. As an intermediate product these are typically application level models structured as record-based tables. These use a relations model (entities & key-attribute based relations) to optimize data for the range of “applications” specified in requirements. They are thus not business-level models. If we use a city planning analogy, logical data models are building-specific blueprints used to guide building construction. These plans largely ignore framing concepts and constraints of a “city” plan which is not so much concerned with individual building-level details as with how buildings integrate into the larger city architecture. The lesson is that analysis preceding database design is typically too focused to provide a broad business conceptualization. Completing the “logical” design feeds final DB design work at Level 3 – a physical schema which represents the implementation of a DB using a specific RDBMS platform. The difference between these 3 levels can be viewed as a hierarchy of constraints that apply at each level. Physical model represents those constraints that are highly specific to aid performance. Logical models provide constraints such as referential integrity or “unique rows” that are critical to databases design, while conceptual models may ignore all of these to capture the semantics and constraints from a purely business point of view – central to enterprise models. As part of enterprise modeling “Logical Data Models” sometimes are used in an effort to structure enterprise-level data types and related business process rules. This is mostly not appropriate using the formalisms of a logical data model, but it can be done in a conceptual model and this is what is needed for integration at the enterprise level. When done properly a conceptual model provides an interpretation that should be readily understandable by business analysts who are not trained in data models. It is this type of model that is of most interest to business analysts and thus it has a central role in enterprise models

**Enterprise-Level Conceptual Models** - With this as background it is possible to better understand the different needs of enterprise architecture compared to a database-focused architecture. An EA framework needs all three levels of the ANSI approach since all are needed to build IT systems. However, the model products developed as part of an EA need to be more general than those of the purely DB kind and cannot just be formed by aggregating data using a bottom up process from extant DBs. EA products should provide support for many implementation technologies, e.g. XML-based messaging for sharing data not just DBs. The modern data concept within an enterprise analysis needs to be broader than that found in a DB model because it includes both non-structured data as well as structured and because data needs to support multiple business functions. Crossing business functions means that the “logical” view of one DB cannot be expected to apply to another. Thus, in an enterprise approach the ANSI 3 schema approach for database design needs to be generalized to incorporate both ideas of heterogeneous information and heterogeneous logical models describing the data. This is illustrated as “upper-level concepts” shown in Figure 1. At foundational concepts like this upper level very broad concepts are defined such as “Process”. Within the frame established top-down by the upper level, mid-level concepts can be specified. This level of an ECM includes the main business concepts such as “Business Transaction”. These provide a starting point for top-down development. Mid-level concepts
in turn serve as a bridge to the lower levels and provide conceptual classes for them. As shown in the Figure, DB schema contribute to the lower level model which is made up of business areas. A conceptual DB schema may be aggregated using conceptual analysis working “bottom up” to contribute to the mid-level model. Interpretation of an individual DB “blueprint model” at the bottom level, such as with “federated” DB schemas, takes several conceptual analysis steps. Work often starts with a pre-integration analysis in which input schemas are re-arranged to make them more “similar” (syntactically and semantically). This allows conceptual identification of correspondence of related items in the DB schemas and precise description of inter-schemas relationships. The final integration step actually unifies corresponding items into an integrated schema which includes mappings.

From a broad enterprise-level an Entity Relationship model often defaults to too low-level a design-oriented a logical model, similar to the relational model of data. This makes sense since most “information modeling” targets creation of a database rather than an integrated model of business data. One exception is data warehouse efforts. A data warehouse design has to integrate data from multiple, dissimilar sources into a common database, or to a target data warehouse. Thus data warehouse analysis handles some integrated data requirements for a portion (several functional areas) of a business and provides some mid-level model concepts.

Summary - Useful enterprise models are built by a combination of top-down and bottom up processes. Existing DB schemas provide bottom-level candidate entities and relations for business areas, which need to be integrated by common concepts. To do this they need to include conceptual levels that are both broader and more abstract than lower level IT models.

In a true enterprise view conceptual analysis determines high-level/global classes of data which may cross functional business lines and establishes an abstract infrastructure of the data architecture. Such mid-level and upper-level models provide semantic relationships that frame definition of data at lower, implementation levels.

References
About the Author: Dr. Berg-Cross (Ph.D.) is a Cognitive Psychologist with over 20 years of experience in enterprise analysis and modeling. Over the last 3 years he has worked on enterprise architectures and as part of EM&I’s BMMP team is currently coordinating efforts to integrate architecture using coordinated ontology and business vocabularies for Service-Oriented Architectures as part of the DoD Business Modernization effort. He provides guidance and strategy for the DoD Business Management Modernization Program’s on Architecture & data management including issues on how to leverage business models to support data sharing and system interoperability thru development of a standard business vocabulary from architecture products. He has published Cognitive Science research including knowledge representation, intelligent and adaptive systems, and data warehouse development. Prior to joining EM&I he worked at DoD TRICARE Management Authority IM and eBusiness divisions and NIH Institutes using web technology to represent, visualize and disseminate information and coordinating the operational and enterprise architecture development. He supported DARPA knowledge engineering, semantic modeling and message understanding projects, the ANSI X3H4 repository technology standard and the HL-7 Reference Information Model.

About EM&I: Engineering, Management & Integration (EM&I) Incorporated is a client-focused, management-consulting firm bridging the gap between business and technology. Our teams of highly skilled specialists provide clients engaged in technology intensive business situations with solutions geared to reduce risks and ensure successful results. In an ever-changing business environment, we support our client agencies in four key areas: Strategy, Architecture, Business Solutions, & Governance. Please visit [http://www.em-i.com](http://www.em-i.com) for more information.

Contact EM&I: If you are interested in hearing more about EM&I’s services to the government and private sector or would like to discuss potential teaming arrangements, please contact Dr. Malcolm Slovin at 703.742.0585.

© 2005 Engineering, Management & Integration, Inc. – All Rights Reserved.