Only the Strong Survive!
The 2002 Corporate Information Factory

This column contributed by Claudia Imhoff and the Intelligent Solutions team.

It's spring, the season that encourages us to approach our tasks with fresh ideas, high hopes for technological breakthroughs and a renewed commitment to building solid environments for our business communities. Each year, we push the boundaries of business intelligence (BI) and data warehousing, discover new and different ways to use the critically important data contained in this environment and appreciate anew the beauty and flexibility of the Corporate Information Factory (CIF) architecture that supports the implementation and maintenance of this environment.

Armed with new information and ever-increasing levels of sophistication obtained through the development of our clients' BI systems, we at Intelligent Solutions continue to improve upon the solid footing of the CIF architecture by adding newly discovered components, expanding the definitions of existing components and renaming components to better describe their functions. Now more than ever, those of you developing BI environments need to use this proven architecture to ensure your success.

This issue of *DM Review* contains the updated CIF architecture poster, developed by Intelligent Solutions and W. H. Inmon, and sponsored by Microsoft. This column is the companion to the poster and covers the latest developments in this established road map for actionable BI and analytical CRM endeavors – recognizing technological advances, new uses for the data, improved processes for construction and maintenance of the components, and, finally the recognition and support for the "actionability" of the BI environment.

Bill Inmon first developed the CIF in the early 1980s. I was privileged to work with him in the late

'80s and, together, we continue to develop and evolve the CIF.

The CIF is a proven and reliable architecture used to create and maintain strategic and tactical decision-support environments. It started life as a data warehouse with associated multidimensional data marts used for newly developed strategic, long-term analyses. Early on, meta data became an important and integral component, gluing the CIF together. Then, the operational data store made its appearance in the mid-'90s to support tactical decision support. Exploration and data mining warehouses were added to complement the OLAP data marts with new and very different analytical capabilities. Operation and administration functions were added to ensure long-term survivability of the environment. The information workshop was added to satisfy the need to bring all the components to bear on a particular workflow or business process. The latest additions are the oper mart and information feedback components.

New Drivers for the CIF

We have identified some new and some continuing drivers for the CIF. We see CRM and ERP implementations and supply chain management as significant initiatives requiring the CIF architecture as the grounding for the technical infrastructure. More sophistication in B2B and B2C transactions lead directly to the need for the CIF to coordinate and ensure integration of the data and information for a seamless enterprise. Business process reengineering (BPR) and XML continue to necessitate a sane and rational approach to their technical infrastructure.

Sensitivity to costs and ROI, in particular, will drive IT implementers to find ways to reduce redundancy of data and processing, increase reuse of components, advance data integration processes and improve the availability

of data and information to their business users. These drivers have long been the real paybacks derived from implementing the CIF architecture. We expect to see these benefits and paybacks contribute substantially to the overall ROI for IT and for their business communities.

Updates to the Existing CIF Components

Let's look at how the existing CIF components have advanced over the past year.

Data Warehouse: Increasingly, corporations recognize that the data warehouse is a mandatory component in their BI environments. The data warehouse is now seen as more than just the source of data for multidimensional or OLAP data marts. It must supply a rich set of data usable for myriad BI analyses such as exploration or mining. The design of the warehouse must be able to provide data for star schemas, cubes, flat files, subsets of relational data, and so on; thus, its design must be as stable, flexible and unbiased as possible.

Data Acquisition: Parallel processing has become an important feature of the data acquisition or ETL (extract, transform and load) tools, especially in environments that add massive sets of new data. Another welcomed occurrence is the reduced price tag of these tools. ETL vendors have come out with "ETL-lite" versions of their software with a much-reduced cost for those smaller companies not needing the robustness of their bigger brothers. We are also seeing the packaging of ETL products with templates for integrating ERP data, new middleware interfaces and better data-quality processing.

Data Delivery: The software that creates the various data marts is now interfacing better with multidimensional databases and includes much more robust meta data. These tools (many are also the data acquisition tools pre-

viously mentioned) also provide newer and better templates for data delivery into analytical applications and have the same reduced price tags as their "lite" versions.

Exploration and Data Mining Warehouses: More and more corporations are realizing the importance of these remarkable components of the CIF. Environments allowing true ad hoc queries, exploration of data, examination of trends, patterns and exceptions, and statistical analyses are becoming obligatory applications.

Enterprise Data Management: Companies are starting to understand the real beauty of partitioning and archiving the data from the warehouse. In addition, they are realizing the need to conform dimensions across the OLAP data marts.

Meta Data: Recently, the biggest development in this component is recognition of the need to centralize the capture of meta data created and maintained by ETL tools. The ETL programs contain much more in-depth information about the data and its environment than the meta data we usually think of (source of data, target of data, transformation rules, etc.). We are also noting the expanding usage of meta data beyond BI or data warehousing, thus recognizing the enterprise nature of this important component.

Decision Support and Transaction Interfaces: We are now able to push BI to the masses through non-traditional interfaces such as PDAs and cell phones. The software supporting these interfaces (access tools) now contains rich visualization capabilities, sophisticated algorithms and analytical competencies, and the ability to trigger alerts or alarms when a specific threshold or set limit has been reached.

Operational Data Store: This component has perhaps had the best year. The ODS's very existence has been unquestionably validated by the critical role it plays in customer relationship management. The ODS is no longer an optional piece of the CIF. Rather, it is now mandatory in supplying the current, integrated customer data supporting true CRM. In addition, its role has expanded beyond just the integration of up-to-the-minute operational data into one where it is now creating data that exists nowhere else in the enterprise.

Operation and Administration Functions: These functions supporting the maintenance and sustenance of the CIF have been given a boost through the usage of new, mobile administrative utilities available on PDAs (e.g., Pocket DBA).

Operational Systems: Enterprise application interfaces now allow access to multiple operational sources. XML is coming to the fore as the standard language to be used in this data exchange. Also, administrative tools are becoming much more sophisticated and user-friendly, more robust in functionality and meta-data savvy.

New Names

OLAP Data Marts: We renamed the data mart component to OLAP data mart to reflect the multidimensional nature of these popular marts. The big news here is the creation of customizable analytical applications (see the Intelligent Solutions column in the April 2001 issue of *DM Review*) for OLAP data marts. Many vendors are creating templates or starter sets of data models, queries, reports and KPIs with flexible points in their implementations, allowing customization to your company's specific needs.

Information Workshop: This component was renamed from enterprise portals to reflect its workflow nature rather than indicate a specific technology used in support of the function. The ultimate goal of the information workshop is to mask the technological complexity of this environment from the business community and incorporate the various components into the appropriate business process at the appropriate points.

Newest Additions to the CIF

Probably the most important characteristic of the CIF is its ability to accommodate and adapt to the ever-changing technological and business world. Two new additions to the components of the CIF evolved through usage of the CIF and demonstrate the ability of the CIF to incorporate evolv-ing best practices. They are:

Information Feedback: This is the set of processes that transmits the intelligence gained through usage of the CIF to appropriate data stores. This is the mechanism by which we push BI "to the masses." Examples abound of storing the results of BI analyses in operational systems or, more appropriately perhaps, the operational data store. One such example is to store the *results* of a customer life-time value (LTV) analysis. The numerical values generated from such an analysis are stored in the ODS and are viewed by all customer-facing employees *before* they touch the customer. Behavior toward each customer is altered based on the knowledge of the customer's LTV score.

Oper Marts: These are small subsets of the operational data store data used for tactical analysis. They are usually stored in a multidimensional manner (star schema or hypercube) and updated by transactions occurring in the ODS (no history stored here!). They are created in a temporary manner and dismantled when no longer needed. They must be sourced from the ODS (see the September 2001 Intelligent Solutions column in *DM Review*).

The CIF has been the steady beacon for many corporations and IT personnel, consultants and vendors in building strategic and tactical decision-making environments – guiding the implementations of the various components for many years. And no wonder! This architecture has repeatedly been dependable and yet flexible enough to encompass new uses of analytical capabilities, fascinating and futuristic technologies, and entrepreneurial implementations of BI.

It is with a sense of pride and satisfaction that we roll out an updated version of this highly successful architecture. We hope you find the current additions and updates useful in your own BI work and look forward to sharing with you other innovations and discoveries in future implementations of the CIF. **DM**

Claudia Imhoff is the president of Intelligent Solutions, Inc. Imhoff is a popular speaker and internationally recognized expert on the Corporate Information Factory, business intelligence and CRM. She has co-authored four books and over 40 articles on these topics. She may be reached at cimhoff@intelsols.com.

Reprinted from DM Review • March 2002

www.dmreview.com