

# Conference Trip Report

## 2005 DAMA International Symposium & Wilshire Meta-Data Conference

*Renaissance Orlando Resort at SeaWorld, Orlando, Florida • May 22-26, 2005*

**This report compiled and edited by Tony Shaw, Program Chair, Wilshire Conferences**

The 2005 DAMA International Symposium and Wilshire Meta-Data Conference was held before an audience of over 900 attendees and speakers. To receive more information about this conference, and related future events, go to <http://www.wilshireconferences.com>

This report contains a chronological summary of the key discussions and conclusions from almost all of the workshops, tutorials, conference sessions, and special interest groups.

### Reproduction Policy

This document is © Copyright 2005 Wilshire Conferences. It may be copied, linked to, quoted and/or redistributed without fee or royalty provided that all copies and excerpts provide attribution to Wilshire Conferences and the appropriate speaker(s). Any questions regarding reproduction or distribution may be addressed to [info@wilshireconferences.com](mailto:info@wilshireconferences.com).

**WORKSHOPS  
Sunday, May 22  
3:30 pm – 6:45 pm**

**Politics and Effective Principles In Data Management**

*Len Silverston  
President  
Universal Data Models, LLC*

We explored politics, what it is, and 8 principles tackling politics to empower people within data management. We shared different perspectives on what politics means. For example, "poly" means "many" and "tics" means "blood sucking insects"! Another was "doing the right thing with the right people".

Also, we participating in experiential exercises, shared our experiences, and made some commitments regarding the following principles:

- developing a clear, compelling, common purpose
- standing in it
- developing trust,
- re-using versus re-inventing
- relating to our audiences
- balancing long term and short term
- making incremental progress
- appreciating versus being "right"

**Closing the Gaps on Business Rule Projects**

*Barbara von Halle  
Founder  
Knowledge Partners Inc.*

Process excellence, business integrity, and business agility require a Business Rules (BR) Approach. Lessons learned from attendees include:

- This tutorial on a BR Approach maps to solid common sense in a disciplined way.
- There are important considerations prior to automating BRs.
- The Rule Maturity Model provides an incremental low-risk path for getting started with a BR Approach, while keeping a higher vision in sight. One participant learned that they may be aiming too high too soon.
- There is an undeniable need for holistic software for managing all rule types across multiple automation platforms.

## Data Governance & Data Stewardship Programs: Delivering an Effective One-Two Punch

*Robert Seiner*  
*President & Principal/Publisher*  
*KIK Consulting/TDAN.com*

Throughout this year's conference, the terms "Governance" and "Stewardship" echoed from the sessions almost as often as the terms "data" and "metadata". It was obvious to attendees that the disciplines of data governance and data stewardship have become unavoidable terrain on the road to delivering successful data management programs.

Bob Seiner's workshop on data governance and data stewardship concentrated on how to "formalize and execute accountability for the management of enterprise data." Seiner's session focused on his 3-D© approach including identifying and leveraging "de facto" stewards, invoking "discipline" through organizational process, and recording/leveraging steward metadata recorded in a "database".

The first portion of the workshop focused on data governance and the importance of educating senior management and business audiences on practical and pragmatic approaches to addressing the organization's compliance and regulatory initiatives. Mr. Seiner discussed introducing management education as a way to enforce acceptable best practices and data operating policy.

The second part of the workshop focused on identifying appropriate data stewards and roles/responsibilities, engaging stewards during specific phases of the data development life cycle, resolving data conflicts and capturing / utilizing meta-data about the stewards.

## Data Models vs. Process Models, Event Models, Location Models, and Organization Models"

*Steve Farrell*  
*Business Analyst*  
*Advanced Strategies, Inc*

Today, the kind of projects that once delivered "IT solutions" are now expected to deliver full "business solutions" – and rightfully so.

Since most organizations run on information, data modeling is still at the heart of business analysis. But other types of business modeling are needed to provide full business benefit. A government permitting case study showed the following, fairly typical, improvements:

- An **object/data model** provided a common vocabulary for the organization (as well as the framework for the needed data structures).
- A **process model** showed (1) a serial process that had no dependencies. It was converted to a parallel process. (2) that a single process handled all permits; first in, first out; regardless of complexity, The process was changed to begin with an assessment to allow simple permits to be handled in a new fast-track process.
- An **event model** showed why it was so difficult for applicants to find out the "status" of their application.

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- A **location model** showed how the physical facilities had evolved to serve the employees at the expense of the customer. This was remedied.
- An **organization (socio-political) model** made it clear that the applicant had to deal with too many agencies. A new org role (job position) was created to be a single point of contact with the customers.

## **Human Side of Data Modeling**

*Alec Sharp  
Senior Consultant  
Clariteq Systems Consulting Ltd.*

Above all, a data model should provide a basis for communication among decision-makers, content experts, users, analysts, and designers. Unfortunately, the communication often gets lost. Whether the modeler has drifted too quickly into abstraction and generalization, or has taken the “deep dive for detail,” the result is the same – confused, frustrated, or detached subject matter experts, and therefore, inaccurate models. Experience shows that it doesn't have to be this way - simple techniques, consistently and regularly applied, will go a long way to ensuring involvement, buy-in, and communication. These techniques include:

- Skip the "tutorial" on data modeling
- Getting started the right way by choosing between top-down, bottom-up, or sideways-in approaches
- Appealing to all learning styles - visual, auditory, and kinesthetic.
- Guidelines for data model graphics.
- Using workflow modeling and use cases.

## **Business Meta Data Strategies**

*Bonnie O'Neil  
President  
Westridge Consulting*

Business meta data is like a “diamond in the rough”; very valuable but not recognized.

- Good definitions are a great place to start
- Semantics is a critical area
- Since the repository offerings are limited at this time for business meta data, most shops are electing to build their own custom solutions
- Data must be able to be delivered to users in an easy manner
- Measure the results and promote them: toot your own horn!

## **Acquiring, Storing and Using your Meta Data**

*Doug Stacey  
Manager  
Allstate Insurance Co.*

Effective meta data management is much more than just a matter of identification and collection. The whole point is that people in the organization...users, analysts and

developers...actually USE the meta data you're making available to them. The principles in this presentation included:

- Start small
- Sell the benefits
- Utilize the meta data!
- Find an area in which you can make an impact on the business
- Demonstrate the ROI then ask for the initial investment
- Catch a wave!

## **Introduction to Web Services - What does it mean to the Data Architect?**

*James Bean*  
*Lead Data Architect*  
*American Express*

Web services describe a standardized way of integrating Web-based applications using the XML, SOAP, WSDL and UDDI open standards over an Internet protocol backbone (Wikipedia definition).

- XML - eXtensible Markup Language - A self-describing, semantically rich, flexible and extensible metadata language.
- XSD (or XSDL) XML Schema Definition Language - An XML based language used to describe and constrain an XML instance.
- WSDL Web Services Description Language - An XML based language used to describe a web service.
- SOAP - Simple Object Access Protocol - An XML based framework for formalizing and facilitating messages.
- UDDI - Universal Description Discovery and Integration - A directory framework that is used to expose Web Services.
- SOA - Service Oriented Architecture - An application and technology architecture that leverages the characteristics of services, messaging, and collaboration to resolve business requirements.

What do data folks need to think about ?

- The recommended approach: Design the service around the Interface!
- If we have a well-defined, structured, semantically rich, flexible and extensible interface, we also have the best possible contract for collaboration.
- We can then design our services (as well as our client interaction) around this contract.
- We need to "own" the interface, which is fundamentally metadata.
- An XML formatted message is an instance of the interface (e.g. the message is an execution of the terms of the contract for collaboration).
- Although primarily a hierarchy, the message is comprised of structures and data containers.
- The interface incorporates a set of metadata rules and constraints that prescribe the format of the message.

Our role is evolving:

- As data folks, our focus on the interface is similar to our traditional role of acquiring and formalizing data requirements, and then applying structure, form, and standards.

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- However, in the case of a Web Service, we need to extend our thinking beyond the persistence layers and into the interface, and message design.
- Remembering that a Web Service is interoperable and transparent, we also need to use our interface to help insulate our persistence layers from our interface layer.

## **Why and How to Use Dublin Core for Enterprise-Wide Metadata Applications**

*Ron Daniel  
Principal  
Taxonomy Strategies*

*Joseph Busch  
Principal and Founder  
Taxonomy Strategies*

The Dublin Core is the nickname for an ISO standard (ISO 15836) that specifies a small set of 15 elements for describing content. Those elements include Title, Creator, Publisher, Date, Subject, etc.

The tutorial discussed the difference in the meaning of the term "metadata" between the database and library communities, and the role of the Dublin Core elements as common "integration metadata" to help in combining unstructured data with structured data from multiple repositories. The presenters then walked through the elements in the Dublin Core, discussing their definition as well as their practical application (or lack thereof). The session concluded with a discussion of selling a taxonomy and metadata project within a company.

## **Databases Demystified: A Conceptual Framework for Comparing and Contrasting:**

- \* **Operational vs. Analytic Systems**
- \* **Relational vs. Dimensional Modeling**

*Wayne Little  
Sr. Data Architect  
Cutting Edge Computer Consulting*

The workshop explored both the types of systems used in reporting architectures including Operational Systems, ODS, Data Warehouses and Data Marts, as well as the differing modeling techniques used to design them.

Relational design with B-Tree indexes were shown to be better suited to operational reporting due to factors including current-valued changing data, small transaction size, need-it-now and predictable and consistent reporting needs.

Dimensional designs with bitmap indexes were shown to be better suited to analytical reporting due to factors including understandability, static but multiple years of time-variant data, large transaction sizes, relaxed performance needs, and the ad hoc nature of analytical queries.

Participants left better equipped to answer "What report goes where and why?"

**NIGHT SCHOOL  
Sunday, May 22  
7:00 pm – 8:00 pm**

**Driving improved data warehouse performance with data usage governance**

*Ted Serres  
Data Architect  
Nike*

**Implementation Experiences**

- Communicate the governance in terms the business understands, explain the trade offs, give real examples of piranhas and goldfish
- Expect the initial rounds to be a little bumpy until teams become familiar with the process
- Several iterations of breaking down report requirements may be required before clean placement of a report into a layer.
- Consider a business transition team to support business process impacts from changes in report design and placement

**Data Usage Governance Considerations**

- Leverage the data usage governance to protect the operational layer
  - Current data, tactical, low row count, intraday refresh needed
- The analytic layer probably requires the least amount of governance when compared to the operational layer and integration layer
- Some technical strategies to compliment data usage governance are
  - Align the technical topology to the reference architecture
  - Model the operational layer objects narrow – just what is needed
  - Partition for performance in the operational layer
  - Monitor query times and row counts in the operational layer

**Applying Terminology Principles To Metadata Names**

*Judith Newton  
Principal  
Ashton Computing and Management Services, LLC*

Naming of data entities (data elements, value domains, attributes, etc.,) in a rational and organized way is an integral part of the metadata management of an organization. Terminology research represents the application of semantics theory to classifying things in the real world.

Terminology and classification principles applied to metadata management improve the understandability and usefulness of information. Document availability:

- ISO/IEC 11179-3:2003 (Second Edition) is now available to download
- ISO/IEC TR 20943-1, Procedures for achieving MDR content consistency
- ISO/IEC TR 20943-3, Procedures for achieving MDR content consistency

All of the above are available free at: [www.jtc1.org](http://www.jtc1.org) (select "Freely Available JTC1 Documents")

**TUTORIALS**  
**Monday, May 23**  
**8:30 am – 4:45 pm**

**Enterprise Architecture Principles and Values -  
Straight from the Source!**

*John Zachman*  
*Founder*  
*Zachman International*

In the Industrial Age, it was the products (airplanes, buildings, automobiles, computers, etc.) that were increasing in complexity and changing. We had to learn how to describe them (architecture) in order to create them and maintain them (that is, change them) over time. In the Information Age, it is the Enterprise that is increasing in complexity and changing. By the same token, we will have to learn how to describe them (architecture) in order to create them to accommodate dramatic increases in complexity and keep them relevant as they change over time. The Framework for Enterprise Architecture defines the set of descriptive representations that constitutes Enterprise Architecture and establishes a semantic structure to understand the physics of Enterprise engineering.

For those who understood the value proposition, Enterprise Architecture has always been important. Yet it is only relatively recently that the concepts and benefits of Enterprise Architecture have started to be embraced by a significant number of organizations at a senior level. John explained how and why Enterprise Architecture provides value, and the four reasons why you "do" Architecture including alignment, integration, change management and reduced time to market. Without Architecture, there is no way you can do any of these things.

**Advanced Topics in Data Modeling**

*Tom Haughey*  
*President*  
*Infomodel, LLC*

Tom discussed a number of advanced modeling issues, including:

What is a dimensional model?

- A dimensional model is a model in which the data is structurally classified as fact or dimension.
- General characteristics:
  - Query oriented
  - Structured around data usage not business rules
  - Organized roughly into base facts and dimensions of those facts
  - Based on identification of key grains of data
  - Consisting usually of pre-joined data
  - Looks to reduce the number and depth of joins
- Levels of dimensionality can differ:

- From very atomic to highly aggregated
- From few dimensions to many
- Facts usually contain three or more dimensions

## **Developing an Enterprise Data Strategy**

*John Ladley*  
*President*  
*Navigant/KI Solutions*

- Value = Usage, there is no intrinsic value in information
- IT efficiencies are only a start. Business enabling must also be accounted in the value proposition of IT
- Understanding your maturity path and the value intensity of the metrics will allow for a pragmatic evolution
- A metrics driven information culture must focus on SUSTAINING business value as well as delivering specific deliverables

## **Data Integration Methodology for Data Warehousing**

*Larissa Moss*  
*President*  
*Method Focus, Inc.*

Traditional waterfall methodologies do not support the iterative development of an evolving enterprise-wide data warehouse (DW) environment that contains multiple integrated databases and applications. A data integration methodology is needed, which incorporates the following characteristics:

- Projects are justified not only based on return on investment (ROI) but also return on asset (ROA) because data is regarded as an important reusable business asset.
- Quality (especially data quality) is more important than time to delivery. However, because delivery times must still be kept short (90-180 days), applications are developed using software releases.
- Cross-organizational tasks are embedded throughout the development steps to ensure that DW databases and applications are collectively architected.
- Enterprise information management touch points exist in almost every development step to help the project teams achieve enterprise-wide integration.
- Project teams are self-organizing SWAT teams. Each team includes at least one business representative who is fully dedicated to the project.

## **Designing Quality in to Data Definition and Data Models**

*Larry English*

*President*

*INFORMATION IMPACT International, Inc.*

This tutorial addressed the critical quality characteristics for data definition and information models, including quality of:

- Information standards
- Business term, entity type and attribute names
- Consistency of attribute names across different media (screens, reports, table column names)
- Definition completeness, correctness and clarity
- Business rule completeness, correctness and clarity

Information model quality includes quality of:

- Stability
- Flexibility
- Reuse

The students participated in Root-Cause Analysis workshops to identify root causes of poor quality definition and poor quality data models. The students participated in Process Improvement workshops to identify improvements that can prevent these types of defects.

## **Managed Meta Data Environment Full Life-Cycle Roadmap**

*David Marco*

*President*

*EWSolutions*

Organizations are implementing managed meta data environments (MME) to provide them an enterprise meta data management solution.

- Plan Your Enterprise Around the “To Be Business Processes” not your “As Is” (80-20 Rule)
- Use Commercial Processes and “COTS” Products
- Data Strategy is the Key to Data Interoperability Across the Enterprise
- No Meta Data Management, No Data Strategy
- Compliance is an “On-Going Process”
- Your Business Processes Should Not Be Unique

## **Enterprise Metadata: An Information Odyssey**

*Todd Stephens*  
*Director of the Metadata Services Group*  
*BellSouth*

### Summary of Best Practices

- Perform a Business Plan and an Annual Plan, even if the activity is not required.
- Measure Everything, Promote the good stuff
- Speed is Essential, Utilize it, Embrace it, Live it
- Continually sell, market, and brand the world of metadata and the adjoining disciplines.
- Pick your team, leadership, products, and timing wisely.
- Success is about people and relationships; Metadata is simply a tool and/or a means.
- Commoditize metadata, the product, the services and yourself. Before someone else does it for you.
- Welcome the progression that will occur over the years. From the raw material of metadata to the enterprise implementation of the experience.
- Never stop learning and expanding the body of knowledge. Never stop demanding that from yourself, your team, leaders or the vendor community.
- Enjoy the Journey.

## **Semantics: The New Era for Information Management**

*Dave McComb*  
*President*  
*Semantic Arts, Inc.*

Semantic Technology is about to explode in the marketplace: independent projections suggest a \$60 billion industry by the year 2010. The good news for the data management community is that people with Data Modeling and Analytic skills are in the best position to capitalize on this trend.

Semantics is the study of meaning. Semantic Technologies are those that take advantage of a more formal, and computer interpretable, expression of the meaning of the information stored in our systems. With Semantic Technology we will be tagging meta data as well as instance data with ontologically agreed terms, in a way that our systems can use to improve search, integration and more.

## **Consulting Skills for Data Professionals**

*Graeme Simson*  
*Senior Fellow*  
*University of Melbourne*

Consulting skills are critical because:

- data managers work largely through persuasion rather than direct authority

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- a substantial proportion of data management problems are 'political' rather than technical
- Most data management people (and most IT people) have not developed their consulting skills to the same extent as their technical skills.

Some key ideas:

- Manage expectations – of the economic, user and technical clients
- Avoid working alone – and have mentor backup available
- Make the client take some responsibility for deliverables
- Encourage strong relationships between own manager and client managers (at highest possible level)
- Re-write the project plan a week or two after project commencement (agree this step in advance!)
- Start writing the report early – “no (unpleasant) surprises in writing”
- Organize follow-up when planning the project

**NIGHT SCHOOL  
Monday, May 23  
5:00 pm – 6:00 pm**

**Facing the Outsourcing Boogie Man**

*Paul Thompson  
Senior Advisor  
Metagyre, Inc.*

Offshore Outsourcing is providing a unique opportunity for data management in:

- Modeling current state systems
- Elevating the value of logical and design models
- Reducing management and communication costs

Data management professional have always lived in complexity and can now:

- Improve their enterprise's capability to orchestrate complexity
- Encourage business silos to participate in enterprise level modeling
- Obtain a zone of optimal modeling

**Data Warehousing 2.0: Lets Visualize the Future**

*Neil Raden  
President  
Hired Brains Research*

Required BI Platform Qualities

- De-emphasizes technical, favors domain expertise
- Declarative modeling (no programming or scripts)
- Intelligence about data (implies active metadata)
- Collaboration capabilities
- Complete abstraction from physical data

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- Zero-impact from changes in the physical model
- Expressive; build models from scratch, assemble from components, protected by referential integrity
- Standards-based

## **Navigating the Data Quality Analysis Tool Market**

*Celeste Harris*  
*Consultant*  
*Metaview360, Inc.*

### **Factors That Will Influence Your Tool Choices**

- What methodology is included in the tool or supported by the tool
- Meta data features in the tool and integration with other meta data tools
- Business factors that influence requirements
  - Timing
  - Price
- Infrastructure implications
  - Current environment
  - Storage and processor requirements
  - Security implications
- End users of the tool – data analysts, business analysts, etc.

## **Resolving Taxonomy Challenges and Information Architecture Conflicts**

*Seth Earley*  
*President*  
*Earley & Associates, Inc*

### **Enterprise Taxonomy is:**

- A dictionary of preferred terminology that can be created for any business environment
- A structured hierarchy of relationships of terms and concepts to support the businesses information needs
- A navigational aid for classifying content and finding business information
- A common language for site architecture

### **Process for Deriving a taxonomy**

- Define the scope of the project
- Determine application framework
- Gather user behavior
- Survey content
- Perform knowledge audit
- Review existing term sources
- Characterize audiences
- Organize terms
- Create “straw man” and validate
- Create search and navigation scenarios
- Iterate

## Vocabulary! Vocabulary! Vocabulary! Managing XML Metadata in Medium and Large Enterprises

*William K. Neils*  
*V. President, Consultant*  
*Quantum Resources Group*

Why do we care about Vocabularies?

- The more ambiguity we have, the harder it is to communicate.
- Undocumented Vocabularies are difficult to understand (i.e. “go ask the programmer what the data means”). Today's world?
- Known Vocabularies can be more readily translated – easier to communicate and integrate. Especially if using an agreed to meta language (i.e. XML) to describe it
- Enforce rules for managing Vocabularies (thru IT Standards, Guidelines and Compliance for developers) to reduce data redundancy / ensure better reuse.
- Online Vocabularies / Data Dictionaries can provide the “masses” with an understanding of their data ... what a concept!
- Controlled Vocabularies support automated systems that will help integrate business processes and enterprises. Initially with ‘statically’ defined integration ... evolving to dynamic, adaptive, autonomic integration!

## What's in a Name? Metadata about personal names within databases

*Jack Hermansen*  
*CEO*  
*LAS, Inc.*

People's names are usually treated as a simple character string, but there is a wealth of data within people's names:

- Names often reveal very specific information about a person, such as gender, culture, religious affiliation, and military or social status (e.g., “Professor”).
- This data can now be mined quickly using new, knowledge-based technology, for names from every country in the world.
- Statistical analyses now make it possible to continuously monitor and improve the data quality of personal name files by more accurate parsing of complex or unfamiliar name types into given name and surname field:  
(e.g. Bustamante de la Fuente, Maria del Carmen).
- Sophisticated new technology now supports advanced searches of your data for both business and personal names.
- Organizations that continue to use archaic name-handling technology will not be well-prepared for our increasingly multicultural world.

## **Data Retention and Risk Management**

*John Murphy*  
*Principal*  
*Apex Solutions*

Data management needs to incorporate practices that are in sync with formalized data retention policies. The cost of retention is not only in the DASD occupied, but in the legal exposure through failure to comply with stated policies. Risk is often increased with retaining data that has exceeded its formal retention period and is of no analytical purpose to the organization.

- Develop Policy statements that have specific actions and measures associated with them.
- Reduce your total volume first by eliminating content that has expired policy retention limits.
- Incorporate retention requirements in your metadata under governance.
- When implementing, ensure you have:
  - Rules Engine
  - Audit Services
  - Reporting Services
  - Disposal Services
- Ensure you are incorporating all data sources
  - Relational
  - Unstructured / Content Managed
  - Email, Paper etc.

## **DAMA International's CDMP Certification Program**

*Patricia Cupoli*  
*DAMA International ICCP Liaison*  
*DAMA International*

DAMA International's Certified Data Management Professional (CDMP) program has been in existence for the last year. This program requires three exams: IT Core, one data exam and one exam of choice. Also, data work experience is required plus continuing education for recertification.

DAMA has been working with the Institute for Certification of Computing Professionals (ICCP) since 1993 in the area of professional data certification. Jointly developed DAMA / ICCP exams include Data Management, Data Warehousing, and Database Administration. The next exam to be developed is Data & Information Quality.

Reasons to certify include professional growth, credentials, self-assessment, professionalism, and challenge. Certification benefits include increased credibility, assessed ability, and increased compensation (bonuses and raises). Attendees were offered the opportunity during the Meta-Data Conference / DAMA International Symposium to take ICCP exams for the CDMP.

**SPECIAL INTEREST GROUPS**

**Tuesday, May 24**

**7:15 am – 8:15 am**

**The New Jersey Enterprise Information Architecture and the Zachman Framework**

*Dan Paolini*

*Director, Data Management Services*

*State of NJ - OIT*

An Enterprise Information Architecture (EIA) consists of the organization's information architecture vision, goals, objectives, policies, standards, tools and technologies, as well as the methodology and the expertise to employ them. The creation of an EIA is a major yet essential commitment to any long-term strategic initiative to support an enterprise architecture. By aligning the EIA with the Zachman Framework, we insure that the EIA is a consistent component of the Enterprise IT Architecture.

How do we get there?

- Recognize that we are our own worst enemy
  - o We are the people that the geeks call geeks – de-geek thyself
- Data isn't Technology, it's Business
- Data Administrators should be the closest IT professionals to the Business
- Don't "model" – engineer

**KEYNOTE PRESENTATION**

**Tuesday, May 24**

**8:45 am – 9:45 am**

**Data Warehouse and Unstructured Data**

*Bill Inmon*

*Founder*

*Inmon Data Systems, Inc.*

Data warehousing concentrates on only structured data - the formal, transaction based data of the corporation. In order to really open the floodgates it is necessary to open data warehousing up to a different kind of data - unstructured data. In unstructured data, emails, email attachments, spreadsheets, documents, and other forms of unstructured data are able to be integrated into the data warehouse.

New tools and techniques are now available to bring structure (and therefore manageability) to unstructured data. Once data is structured, then standard structured transactions have no problem mixing structured and unstructured data. And a whole new world of processing emerges. There is a new and exciting world of business intelligence – unstructured business intelligence – that promises to truly unleash the promises that have been made in years past.

## CONFERENCE SESSIONS

Tuesday, May 24

10:15 am – 11:15 am

### **Enterprise Architecture, Service Oriented Architecture & Knowledge Management at Motorola**

*Joe Danielewicz*  
*Enterprise Data Architect*  
*Motorola*

Motorola Enterprise Architecture is built on the ProACT Framework. The ProACT Enterprise Architecture Framework (EAF) supports four inter-related views of the enterprise architecture:

- Business (or Mission, for non-business enterprises)
- Operational
- Systems
- Technology (outsourced)

Each of these views is typically planned and controlled by different groups of architects. These four views are represented as layers in the framework. Each layer is fed requirements from the layer above and is offered capabilities from the layer below. The challenge for today's architects is to maintain top to bottom alignment between these requirements and IT-enabled capabilities.

### **Data Governance in Action: Cases from the Insurance and Securities Industries**

*Barbara Nichols*  
*President*  
*Metaview360, Inc.*

Data governance is not a technical principle. It has no snazzy interface, runs on no software platform, and isn't based on mathematical principles. Yet the establishment of effective data governance is an indispensable part of data management particularly for those organizations that are in the midst of transformation. Key ingredients to success are:

- Compelling business imperative with visibility
- Key business personnel authorized and involved
- Key barriers and how to deal with them
  - o Change management
  - o Time
  - o Language
- Key elements of a program
  - o Charter, Mission, Policies, Priorities, Implementation Plan
- Key players and their roles
  - o Business-directed data governance
  - o IT as facilitators and implementers

## **Data Quality and Enterprise Application Integration (EAI)**

*Alex Kerezy*  
*Data Architect*  
*Nationwide Insurance*

General Considerations for Data Quality Guidelines:

- Stop the ongoing bad design.
- EAI is NOT a magic bullet. Plan it out and implement it step by step. Review and revise your plan as needed.
- Make certain people with the right skills and knowledge are involved in designing your common semantic or language.
- Create a road map/project plan for your EAI efforts.
- Go through the efforts of defining your business /systems processes and services
- Prioritize your systems to become true service providers. Identify consumers and a migration plan. A system might have to function two different ways for some time.
- Keep the hardware/ middleware separate.
- Create a centralized team to manage the common message or definition format.

## **Strategic and Tactical Issues in Metadata Management**

*Ian Rowlands*  
*Vice President of Product Development*  
*ASG Software Solutions*

*Alan Perkins*  
*Chief Solutions Architect*  
*ASG Software Solutions*

A CIO Executive Board report published in January 2005 called out IT governance as a top concern. The same survey identified “Strategy and Planning” and “Data and Knowledge Management” as two competences of world-class I.T. organizations. This session took the conference theme of “Aligning Data Strategy with Business Strategy” and extended it to metadata based on Perkins’ and Rowlands’ experience that the key differentiators between metadata management successes and failures are good governance and a clear delineation of strategic and tactical issues.

Perkins and Rowlands reviewed key aspects of governance structure, project planning and design, technology and vendor selection and staffing and the impact of key decisions on projects at several major organizations.

## Data Semantics and Data Ontologies in a Data Warehouse or Enterprise Information Integration Project

*Robert Klopp*  
*Managing Director*  
*Skyland Technologies*

Ontological commitment is a process that allows two systems to agree to use a common Ontological Spec. Eg. You say "Earnings"... I say "NetProfit" let's agree to call it "Earnings". You say that Earnings are related to Taxes... I say that Earnings are unrelated to Taxes (before Taxes)... is there a Class in your spec that is Earnings before Taxes? And so on...in other words... the Ontologies provide the basis for finding semantic agreement...to allow information to be meaningfully exchanged.

OWL (Web Ontology Language) is:

- A W3C standard
- A formal language for specifying an ontology
- Based on RDF which is based on XML
- Provides the basis for ontological commitment between two applications

### Can We Share?

#### **Convergence: Library Sciences, Document Management, Digital Asset Management, Knowledge Management, Content Management, Data Management**

*Loretta Mahon Smith*  
*Information Architect*

Library Sciences, Document Management, Digital Asset Management, Knowledge Management, and Data Management are responsible for organizing and safeguarding different information assets for enterprises. Each discipline has separate tools and techniques which aid in the development, management, and maintenance of the items that are within their domain. Information in most companies is fragmented, managed separately as physical media, virtual documents, digital files, institutional knowledge and structured data. This presentation is an attempt to provide the reference that will help us understand each other's practice, and promote a conversation results in a more mature, coherent and useful environment. The presentation includes a single slide for each of the disciplines that include hot-links to the websites she feels would be a good starting point for research. Additionally, she identified the following concepts that should be immediately reusable, and good talking points for conversations professionals from the other disciplines:

- Library Sciences: Cataloging, Provenance
- Document Management: Fulfillment
- Digital Asset Management: Repurposing
- Knowledge Management: Learning Theory, Communities of Practice
- Content Management: Audience Taxonomy
- Data Management: Graphical Modeling of Information, Metadata

#### Management Techniques

Loretta suggests that DAMA professionals build their own personal network of relationships with the other individuals in the enterprise who manage information!

## **Bridging the Gap Between Technology and Business**

*Jeff Tash  
ITscout & CEO  
Flashmap Systems, Inc.*

Enterprise architects are supposed to operate like city planners, mapping out infrastructure, establishing standards, and achieving cost efficiencies. Of course, the cyber equivalents of roads, schools, hospitals, water, sewer & gas pipes, electric, cable & phone lines, etc., are IT infrastructure components such as database managers, operating systems, a myriad of middleware stuff, and scores of jargon-laden product categories. This presentation examines how Technology Architecture can be used to organize, visualize, and communicate IT standards and how this impacts software architects. In this session, Jeff spoke about:

- how the same Technology Architecture approach used to communicate and enforce IT standards can also be used to reduce complexity, improve collaboration, and increase knowledge sharing among software architects
- the role of Technology Architecture and standards for harnessing and controlling the creation and maintenance of taxonomy, ontology, and classification for SOA services
- TOGAF's (The Open Group Architecture Framework) four kinds of "architecture" that are commonly accepted as subsets of an overall Enterprise Architecture: Business Architecture; Data Architecture; Application Architecture; and Technology Architecture
- how standards can be graphically communicated using a 3 layer/4-model Technology Architecture framework consisting of INFRASTRUCTURE at the bottom, APPLICATIONS (Build or Buy) in the middle, and DATA at the top
- one (of many) beneficial byproducts of standardization is consolidation which results in increased efficiencies, so everyone looks like cost-cutting heroes to the CFO and CEO

### **CONFERENCE SESSIONS**

**Tuesday, May 24**

**11:30 am – 12:30 am**

## **Improve Your Bottom Line – Increase Your Modeling Return on Investment**

*Marcie Barkin Goodwin  
President & CEO  
Axis Software Designs, Inc.*

The absence of a Model Management infrastructure increases costs by contributing to an inefficient modeling environment which perpetuates:

- Lack of reusability
- Insufficient analysis leading to implementation & production errors
- Procedural confusion leading to a slower project development life cycle & missed deadlines
- Lack of communication causing inconsistent & unacceptable model quality

## **Credibility is Critical. How Intel transitioned to a managed enterprise metadata environment.**

*David Ashton*  
*Enterprise Metadata Architect*  
*Intel Corporation*

*Jacque Gibson*  
*Manager, Metadata Services*  
*Intel Corporation*

- Success must be measured by Quality not Quantity
- Scope needs to focus on the Enterprise
- Standard Deliverables and Repeatable Processes must be baked into the development lifecycle
- Align to a corporate initiative (e.g. Quality)
- Use a phased implementation approach
- Fixed deliverables facilitate metadata management and governance.
- Standardized tool set: We had history of attempting to satisfy varied customer requirements – it does not work
- A “build it and they will come” strategy doesn’t work. You must have buy-in and ownership from the beginning
- Metadata should be available to everyone, but have security around maintenance

## **The Buy Solution Secret : The Critical Role of Data Architecture and Administration in Implementing Application Packages**

*April Reeve*  
*Information Architecture Consultant*  
*Metro Solutions, LLC*

This is a very relevant topic because most organizations are buying software solutions rather than building custom solutions. The work of Data Management on a purchased package implementation is much greater and much more complex than custom development. It is very profitable for everyone to get Data Management involved during package assessment including performing a package model assessment. The goal during a model assessment is not to pronounce the model "good" or "bad" but to understand the applicability to your organization's needs. Prior to purchase can be the time a software vendor will be most willing to provide information about their data model. Other areas where Data Management should be involved in a package implementation:

- Conversion (of course)
- Reference Data integration
- Transaction Data integration
- Metadata integration
- Integration with Data Warehouse, Data Marts Consolidated and application reporting

Also helping with data issues in:

- Security
- Change Control, Performance Management, Backup and Recovery

## **Is XML Still Relevant to Data Professionals?**

*David Plotkin*  
*Data Quality Manager*  
*Wells Fargo Consumer Credit Group*

- Avoid chaos by managing XML metadata across the Enterprise:
  - Create reusable base definition schemas
  - Create and document:
    - Widely used elements (with their attributes)
    - Complex data types
    - Simple data types
  - Map relationships between schemas (which schemas use other schemas)
  - Keep track of which documents are validated by which schemas.
- Track usage of elements/components
- Understand the structure of XML schemas
- Avoid proliferating redundant XML structures
- Avoid changes that “break” systems which depend on XML structures
- Version XML Schemas
- Establish stewardship for elements and schemas
- Manage Schemas from a central location (repository)

## **The Zachman Framework, Model Driven Architecture and Meta data: Tying the Enterprise together for reuse**

*Lowell Fryman*  
*Founder*  
*Enterprise Integration Solutions LLC*

The Zachman Framework has long been an approach for identifying architectural constructs and models for enterprise re-use. More recently, the OMG-Model Driven Architecture (MDA) has been developed for similar objectives. This seminar presented both approaches as being complementary and not conflicting. Both make use of engineering models for enterprise re-use and long-term systems maintainability. The seminar discussed the 6 rows of the Framework, and the difference between primitive and composite models. Most often project implementation deliverables will be composites, i.e., the Requirements document. The MDA was discussed in detail, specifically the 3 primary models; the Computational Independent, Platform Independent, and Platform Specific models. These 3 MDA models, as supported by UML 2.0, map almost exactly to the Framework rows 2, 3, and 4 respectively. Thus, we could visualize the alignment of both approaches. Lastly, we discussed that all models should be recognized as being a form of Meta-data. Thus, it is critical that we establish a strong meta-data lifecycle to manage those models to enable enterprise re-use and long-term systems maintainability.

## **MetLife's Experience with Semantics (and the ACORD Messaging Standard)**

*Mary Lee Miller*  
*Enterprise Architect*  
*Metropolitan Life Insurance*

Where MetLife sees Semantic Information Management taking us:

- Providing Semantic Convergence
  - Identifying semantically equivalent data elements regardless of name/format
  - Information Model becomes the Lingua Franca of the enterprise
- Supporting One Business Language
  - All LOBs can discuss Party data and understand each other
- Supporting Data Tracking and Transparency
  - Tracking source/target of transformations and chaining them together to describe overall information flow
- Support the integration of Data Warehouses
  - by connecting to the Information Model to support business reporting, etc.
  - Allowing the business process to respond in real time to trends in the Warehouse.

The benefits of SIM:

- ROI of SIM
  - Productivity of Data Mapping
  - Higher quality of mappings produces higher quality IT products/processes
  - Exposing Redundancy
- Business can respond more quickly to change
- Quicker adoption of standards like ACORD
  - Supports changing versions
  - Supports the addition and migration of Extensions in the Information Model

## **Telling IT Like it is: Assessing and Aligning the Data Management Portfolio**

*Doug Laney*  
*Founder & Chief Research Officer*  
*Evalubase Research, Inc.*

With hundreds or thousands of IT solutions in their portfolio, enterprises are tasked unwittingly with innumerable IT decisions throughout the year: When to upgrade each product? When to replace each product? Which vendors and consultants to consider? How to best implement each product? In this session, Mr. Laney will share current findings and insights from Evalubase Research's continuous analysis of IT vendor/solution ratings and reviews based exclusively on feedback from actual IT professionals, not analysts. His recommendations are:

- Develop and fine-tune your BS Meter to help interpret vendorspeak.
- Be heads-up about gettin' down with various research resources.
- Most IT portfolio management solutions can keep you in compliance, but fail to track complacency.
- IT portfolio benchmarks should be internal and external, periodic, and as non-disruptive and inexpensive as possible.

## **CONFERENCE SESSIONS**

**Tuesday, May 24**

**1:45 pm – 2:45 pm**

### **Data Quality: Research and Practice**

*Richard Wang*

*Director, MIT Information Quality Program*

*Massachusetts Institute of Technology*

Typical data quality issues:

- People no longer ask “What is data quality?”
  - They now say “I know I have problems, but I don’t know...”
- Real-time cleansing and matching
  - E-business is real-time, so it should be the processes
- Meta data management
  - Standards, definitions, and application meta data sharing
  - Application-specific data and context
- DQ dashboards
  - Top quantitative data measurements per function
  - Rolled up into weighted metrics for entire organization
  - Can be converted to \$\$\$ showing cost savings
- New Government Regulations
  - HIPPA, USA Patriot Act (OFAC Compliance), Sarbanes Oxley,

### **The Road to Building an Enterprise Data Model**

*Candace McCabe*

*Technical Lead- IS Model Team*

*JB Hunt Transport, Inc.*

Incorporating the EDM Into the Development Life Cycle

- Validate new projects against the EDM
  - Does the project change the EDM?
  - Does the project validate the EDM?
- Constantly update and refine the model
  - Pull existing entities into project-specific models
  - Refine the entities and relationships as needed
  - Update the EDM with the new project requirements for long-term re-use
  - Maintain both the logical and physical

The EDM and 3rd Party Packages

- The buy decision
  - Use the EDM to validate and identify gaps when evaluating packages
  - Logical mismatches can be used to estimate additional costs of customization
- Maintaining 3rd party packages
  - Incorporate the 3rd party package concepts into the EDM
  - Map similar concepts and document all logical mismatches

- All proposed changes follow EDM standards
- Ensure that changes are updated in the EDM using model management process

## **Use the process! How Entity Life cycles can help define the data.**

*Dagna Gaythorpe  
Independent consultant*

Dagna started by admitting that she is an introvert (as described by the Myers-Briggs classification), and has a lot of strategies for avoiding talking to people. But there are times when she can't avoid it any longer, and has to go and find out about the data that people use and need; this session explained how she does it.

If you ask a business person to tell you about their data, they will often panic and give you a set of reports and screen dumps. These may be useful, but they miss a lot of vital information, so Dagna recommended starting by asking "what is it that you do with this thing", and getting the life story of "thing", identifying the stages it goes through, and what changes at each stage, before asking "so, what do you need to know at this stage". The answer to this last question is the first time that you talk about data, and the answers can be normalised into a model of the data the business actually uses and needs; the "stages" can be recorded in documents or diagrams (in whatever notation suits the business people), and then used as scenarios to test the data models (the logical, physical and database design), define the application or process, test the application, and scope the user guide (on the rare occasions when such a thing is produced!) Above all, the whole thing must be fed back to the business people, so that they can check the result, and see how their piece fits in. Dagna then moved on to talk about questions (because we all have a set of useful questions along with our set of patterns), and covered the useful ones ("what can go wrong"), and dangerous ("what would make your life easier"), and what you should get out of each question.

## **Metadata Management 2005 - an interactive discussion of the latest metadata usage trends, corporate issues, and vendor software solutions**

*Stu Carty  
Founder & Principal  
Gavilan Research Associates*

This presentation provided attendees with a discussion of "best practices and approaches" in developing or reassessing their metadata strategies and successfully implementing and selecting metadata solutions and software. The presentation discussed:

- The top issues that Global 5000 companies are facing in metadata management
- Criteria to effectively evaluate a metadata management solution
- Key implementation strategies to gain immediate use, acceptance and return on investment for your metadata solution
- How to increase the chance of success in evaluating, selecting, and implementing the best metadata solution
- Survey results on the latest metadata vendor tools - software products, capabilities, comparisons, best-fit, features, strengths, and weaknesses

## **Make Your Metadata Repository the Center of Your Company's Data Regulatory Governance Program.**

*David Schlesinger*  
*Data Policy RBAC Architect*  
*Intel Corporation*

*Peggy Schlesinger*  
*Senior Data Analyst*  
*Intel Corporation*

- Data is impacted by a growing number of regulations including Privacy, SEC Insider, Sarbanes-Oxley, HIPAA, §Graham-Leach-Bliley, U.S. Patriot Act, and many others. It should all be handled as a single class called "Regulated Data."
- Data must be categorized centrally by governmental regulatory categories as related to time-span of regulation, need to know, constraints required, and type of regulation. Regulations combined with security policies may result in data having multiple constraints.
- Data Retention schedules required by law must also be captured in metadata definitions.
- The attributes should contain the regulatory categorization as part of defining the metadata and should be housed in the metadata repository. The metadata repository is the correct place to keep this data definition.
- A major aspect of complying with governmental regulations is managing the entitlement profile, which is made up of the inventory of data content in an application and database view, the security classification, and regulatory requirements defined in the metadata of the entities and attributes, and the business rules around each user's job-role.
- The Business Decision of who is entitled to see and handle information is called an "Entitlement Decision" and is quite separate from the security process of access control. This business decision approves a subject's authorization.
- The metadata repository should have a key role, if not the complete role, in managing this Entitlement Decision, based upon the metadata description of the data content under consideration.
- A centrally managed regulatory constraint taxonomy needs to be created so each business unit uses the same terms to define regulatory requirements of their data. As the data flows across the enterprise, other business units must accommodate this as they pull their data definitions from the Metadata Repository to create their local systems.

## **Beyond Metadata: The Low-Down on Ontologies in Words You Can Understand**

*Neil Raden*  
*President*  
*Hired Brains Research*

Ontologies have migrated from an interesting idea to a brand new way to do your job. There is no question that writing a good ontology will require more domain-specific knowledge than you probably have, but that's a good thing for you because it strengthens the link between your work and the business at large. Starting out as a mysterious term and a poorly understood concept,

ontologies have moved into the mainstream with organizations like IBM, CA and Sun developing standards for ontologies.

**PANEL: Guess Who's Next!**

*Michael Brackett (Moderator)*  
*Consulting Data Architect*  
*Data Resource Design & Remodeling*

*John Zachman, Zachman International*  
*Larissa Moss, Method Focus*  
*Dawn Michels, Andersen Corporation*  
*John Ladley, Navigant/KI Solutions*

There are four critical resources in any public or private sector organization - finances, real property, the human resource, and the data resource. Over the years the first three of these critical resources have become more professionally and formally managed. The data resource is the only critical resource that has not become professionally and formally managed. This situation is potentially disastrous because it is the data resource that provides the real understanding about the business. If the first three critical resources are mismanaged there are civil and criminal penalties imposed. The time will come when there will be severe penalties for the mismanagement of the data resource.

**CONFERENCE SESSIONS**

**Tuesday, May 24**

**3:15 pm – 4:15 pm**

**Information Modeling - we've made it work**

*Martin Richley*  
*Head of Army Information Coherence Group*  
*Ministry of Defence*

Information Modelling with CBML enables:

- Business information to be stated
- Information to be reconciled
- Enterprise models to be developed incrementally
- Information exchange needs to be stated

The presentation described:

- Benefits to the business, e.g. common understanding of coherent information
- Benefits to the modeler, e.g. business commitment to precise, accurate information
- Benefits to the enterprise, e.g. business and modelers communicating effectively
- How these symbiotic benefits promote more effective exploitation of business information

## Implementing Business Driven Information Management Practices: Organizing Information Stewardship

*Ron Klein*  
*Sr. Information Management Specialist*  
*BMO Financial Group*

*Linda Mirtle*  
*Sr. Information Management Consultant*  
*BMO Financial Group*

In this information-rich case study there were many take-aways for the audience. Just one was the explanation of the BMO repository development.

- Began our journey in late 2001 and are within a couple of years of it being fully integrated in our day-to-day activities
- Led by Information Resource Management function, but strongly integrated with our Data Warehouse Group
- Hired a Champion who had experience with an established Metadata Function
- Heavily integrated with Business Initiatives on Information Quality
- Focused on bridging business and technical metadata

## Homeland Security Information Sharing through Metadata Management

*Michael Daconta*  
*Metadata Program Manager*  
*Department of Homeland Security*

Mr. Michael Daconta, the DHS Metadata Program Manager, briefed the audience on the strategy and current status of the department's data management and information sharing efforts. The presentation contained four key areas: the overarching vision of data-centric and metadata-driven applications; the Metadata Center Of Excellence and the services it will offer to achieve that vision; the Department's leadership of the Federal Enterprise Architecture (FEA) Data Reference Model (DRM) Working group; and lastly the joint initiative with the Department of Justice to establish the National Information Exchange Model (NIEM). Mr. Daconta explained how the data vision requires a new way of thinking about data where applications revolve around shared data to achieve relevance: "the right information to the right person at the right time". To describe the Metadata Center of Excellence (MCOE), he first covered the basics of metadata and then how his program was structured to deliver metadata management services like a metadata registry/repository, data standards and data governance. He explored details of the strategy in regard to the selected Metadata Registry/Repository and the operations of the DHS Metadata Working Group. In addition to the MCOE, he briefed the status of the FEA DRM and NIEM. The FEA DRM is a government wide model for how agencies will structure their data, exchange their data and categorize their data. Mr. Daconta walked through the primary use case: inter-agency information sharing that is guiding the work. Lastly, he discussed the National Information Exchange Model, a joint DOJ/DHS initiative to establish a set of Reusable XML Schemas for DOJ/DHS information exchanges. He discussed the memorandum of agreement between the organizations, sample exchanges and the current status of the initiative. More information on NIEM can be found at <http://www.niem.gov>.

## **Meta Data Models of the Managed Meta Data Environment**

*Michael F. Jennings*  
*Business Intelligence Strategist/Architect*  
*Hewitt Associates*

Managed Meta Data Environment - The managed meta data environment represents the architectural components, people and processes that are required to properly and systematically gather, retain and disseminate meta data throughout the enterprise.

Universal Meta Data Model (UMM) - The Universal Meta Data Model represents the data and meta data components that are required to support information and knowledge dissemination throughout the entire enterprise

The UMM is:

- Starting point for development of logical MME meta model
- NOT shrink wrapped solution that is ready for implementation
- Litmus test for existing, planned, or purchased MME projects

Model Diagrams:

- The Complete View of the UMM:  
[http://www.dmreview.com/editorial/dmreview/200404/200404\\_060\\_1.gif](http://www.dmreview.com/editorial/dmreview/200404/200404_060_1.gif)
- Enterprise Systems Component Model:  
[http://www.dmreview.com/editorial/dmreview/200406/200406\\_008\\_1.gif](http://www.dmreview.com/editorial/dmreview/200406/200406_008_1.gif)
- IT Portfolio Management Component Model:  
[http://www.dmreview.com/editorial/dmreview/200407/200407\\_jennings\\_online\\_1.pdf](http://www.dmreview.com/editorial/dmreview/200407/200407_jennings_online_1.pdf)
- Business Rules, Business Meta Data and Data Stewardship Component Model:  
[http://www.dmreview.com/editorial/dmreview/200408/200408\\_060\\_1.gif](http://www.dmreview.com/editorial/dmreview/200408/200408_060_1.gif)
- Business Rules, Business Meta Data and Data Stewardship Component:  
[http://www.dmreview.com/editorial/dmreview/200410/200410\\_jennings\\_1.pdf](http://www.dmreview.com/editorial/dmreview/200410/200410_jennings_1.pdf)

## **Taxonomies, Meta-Analysis, and Ontologies**

*Malcolm Chisholm*  
*President*  
*Askget.com Inc*

Taxonomies are:

- A set of categories that spans a domain populated by apparently similar “things”.
- The categories are usually arranged in a hierarchy.
- The “things” being classified get put in the lowest level of the hierarchy.
- The higher levels of the hierarchy are classifications of the lower level categories.

Meta Analysis is:

- a statistical practice of combining the results of a number of studies...Meta-analysis is a collection of systematic techniques for resolving apparent contradictions in research findings. Meta-analysts translate results from different studies to a common metric and statistically explore relations between study characteristics and findings. (Wikipedia)

### Ontology

- In computer science, an ontology is the attempt to formulate an exhaustive and rigorous conceptual schema within a given domain, typically a hierarchical data structure containing all the relevant entities and their relationships and rules (theorems, regulations) within that domain. (Wikipedia)

### Comparison:

- Taxonomy: Permits additional value to be derived from data in a given database. The data model usually has to be extended.
- Meta-Analysis: Permits the questions that can be asked about one set of data to be asked about many combined sets of data. The context is the same. Value is obtained from the combination of data.
- Ontology: A different context, or perspective, on a given set of data. The new context includes a different set of requirements for the data. Value is obtained by having the same data meet the new requirements.

Remember: We have been preaching the value of data as a corporate resource. Ontologies are beginning to ask for that promise to be delivered on.

## **Sustaining Information Management**

*John Ladley*  
*President*  
*KI Solutions*

- Information assets are the expressed valuation of an organization's ability to execute business plans via application of structured and unstructured content. This valuation consists of cost of production, market valuation of the business reflective of managed information, and a statement of the subjective value of corporate information
- Without the mindset of an asset, sustaining IM is difficult
- Value = Usage, there is no intrinsic value in information
- Understanding your maturity path enables many pragmatic responses to roadblocks
- The new information culture must focus on SUSTAINING business value as well as delivering specific deliverables

**SPECIAL INTEREST GROUPS**

**Wednesday, May 25**

**7:15 am – 8:15 am**

**Lessons Learned from Sarbanes-Oxley -a Data Perspective**

*Gwen Thomas*

*Editor*

*SOX-online.com, a publication of the Data Governance Institute*

Action plan for data professionals:

- Become familiar with Risk Management language
  - o risk grids & approaches (accept – avoid – reduce – transfer)
  - o COSO risk management framework
  - o COBIT risk management framework
- Become familiar with language of controls
- Become familiar with Governance concepts

Opportunities for data professionals:

- Chance to demonstrate business savvy
- Chance to be seen as part of solution, not problem
- Chance to move into outsource-safe role

**Unleashing Legacy Data**

*Peter Everitt*

*CTO*

*Modulant*

Any data can be considered legacy data if knowledge about the data (meaning, context, dependencies, constraints, or business rules) is ambiguous - whether the data is years, months, or merely days old. To solve this problem, enterprises need a systematic approach to aggregating knowledge about data, validating it (identifying voids, inconsistencies and redundancies), visualizing it and, ultimately, making data more seamlessly reusable both internally and with third parties.

- Legacy data is similar to material inventory, it is costly to an organization until it can be used to support the business purpose.
- Understanding the business information within the context of the business is critical migrating information as organizations reduce IT support costs through system migration, consolidation and/or elimination
- Semantic Knowledge Management can provide repeatable and predictable results for an organization's migration requirement

## CONFERENCE SESSIONS

Wednesday, May 25

8:30 am – 9:30 am

### Mapping UML Class Diagrams to Relational Databases (Classes to Tables will Never Work!)

*Paul Dorsey*  
*President*  
*Dulcian, Inc.*

When polled, about 80% of the participants were working in shops that were Oracle-based. Almost everyone was working in some kind of RDBMS environment so there was a high level of concern about how to merge OO and relational thinking. Key ideas from the talk were:

- 1) Mapping a class diagram to tables and columns does not really work. There is too much richness in class diagrams and there are too many limitations in a relational database.
- 2) All of these limitations can be solved if the class diagram simultaneously generates both a database and some sort of OO-friendly other object. This other object can be an INSTEAD OF trigger view in the database, an Enterprise JavaBean, a TopLink element, or even an Oracle ADF BC entity object.
- 3) Paul also presented an algorithm demonstrating what tables and views can be used to generate from a class diagram the "right" way. However, he acknowledged that this was open to argument. The important idea is that generating both the database and a middle object was the right way to think about the problem.
- 4) No mainstream products support this idea very well. Paul's own company's product (BRIM) uses this idea but it is not big enough to be considered a mainstream product. If users want to use this idea in their own environments they have to either
  - a. Do it all by hand
  - b. Convince the vendors to write generators
  - c. Write their own generators
- 5) The hard part is thinking through what you want to generate. The actual generators themselves to not require a great deal of work.

### Approaches for Implementing a Multi-Faceted Architecture

*Dawn Michels*  
*Information Architect*  
*Andersen Corporation*

There are four highly recognized facets to Enterprise Architecture:

- Business architecture uncovers and understands the business requirements needed in projects, corporate initiatives and future strategies.
- Information or data architecture concentrates on the corporate data asset and designs needed to support the business requirements.
- Application architecture focuses on the tools and solutions that the business requirements and design specifications demand

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- Infrastructure or Technical Architecture provides the underlying technical mechanism, capturing the total cost of ownership and flexibility for future demand to make it all work. Each of the disciplines has specific emphasis, tools, artifacts and metrics. Some of the ways a company can successfully implement an Enterprise Architecture Strategy are to:

- staff an EA team with strategic thinkers,
- attain sponsorship from the CIO or management as near to them as possible,
- align the goals and strategies for the EA to the strategies important to the business,
- and to establish real goals and metrics to work through.

There is no single "right way" to implement an EA, but rather it should be tailored to match a particular corporation's needs.

### Estimating The Monetary Value Of Data Modeling

*Gary Flye*

*Manager, Enterprise Data Administration  
Wachovia Corporation*

Data Modeling is far more valuable than most people realize:

- The ability of data modeling to accurately and completely capture business data requirements is widely recognized, however...
- The most significant monetary value of data modeling is to be found in its ability to reduce the cost of system maintenance, specifically by reducing the number of errors and subsequent re-work

How much money can be saved?

- There is no definitive answer, however, even if data modeling only reduced system maintenance by 1% for 5 projects at Wachovia, the NPV savings would be over \$1 million. A 5% reduction would yield a NPV savings of \$11 million, and a 10% reduction would yield a NPV savings of \$24 million.

This is a message to take back to your organization. Data modeling is a sound investment!

### **Data, metadata, and ITIL**

*Charles Betz*

*Technical Consultant  
Accenture/Best Buy*

The concept of IT Service Management (ITSM) and its flagship standard, the Information Technology Infrastructure Library (ITIL), have been sweeping global corporations in the past few years, providing a common framework for managing IT in the enterprise. Metadata, in its broadest sense, can be thought of as "data about data AND the systems that process it." However, when the metadata practice moves beyond the basic data dictionary into application systems and hardware, it starts to converge with the pivotal ITIL concept of "Configuration Management Database," or CMDB.

- Metadata is a problematic term. Hard to distinguish "meta" data from other internal IT data.
- Metadata professionals need to start paying attention to the processes by which metadata is originated and maintained.

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- IT tools market and concepts are fluid. Be wary of research firms; they are subdividing the problem domain in non-useful ways.
- Data professionals should cast a critical eye on the data structures being implemented under ITSM auspices
- Metadata repositories and CMDBs will inevitably converge. Be ready for “ERP for IT” suites!

**Meta-data Based Impact and Lineage Analysis Across Heterogeneous Meta-Data Sources**

*John Friedrich  
Director of Professional Services  
Meta Integration Technologies, Inc.*

Many organizations have several CASE and modeling tools, data warehouses development tools, ETL engines, BI reporting and analysis tools, EAI environments, and EIP builders. Each contains, uses, manages and is driven by meta-data. In such complex environments, questions that the business user feels should be simple, such as “why is the field in my report inaccurate”, or “how is it calculated”, become excruciatingly difficult in practice, as one must trace the data lineage across BI, to ETL, to Mart, to DW, to ETL, to CASE, etc., tools. The reverse question, i.e., what is the impact of a system or data change upon all of these related environments, is equally difficult. The difficulty initially stems from the inability to collect meta-data from these disparate environments in a common representation with relationships, or data mappings, from tool to tool captured and their lifecycle managed.

**Turning Reference Data Sideways**

*Ted Kowalski  
Data Architect  
Shell Oil Products*

The speaker offered a process and design that puts all of your reference data in a repository that is both simple and yet powerful (about ten tables). Company, Products and Customer references can span across multiple countries, cultures, and measures. A generic or common value can be related to various location-specific implementations (Do all countries hold the same meaning for “station wagon”?) This design, where the structure IS the metadata, opens you even to identifying your information enablers and disablers. This can be the beginning of integrating all of your reference data from legacy or SAP systems into a centralized repository and thus reacting to change more quickly. The benefits of this approach are:

- Simplicity
- Integrity, Centrally Focused
- Flexibility
- Low cost
- Re-use, add easily, not buried in program
- Great opportunities for consulting
- Attributes / Constraints Predefined & Grouped

## **CONFERENCE SESSIONS**

**Wednesday, May 25**

**10:00 am – 11:00 am**

### **Gaining Data Element Insights**

*David Beulke*

*Principal*

*Pragmatic Solutions, Inc.*

The practices and techniques of Business Process Analysis can help you quickly learn all about your applications:

- Discovering your data elements
- Naming and definition conventions
- Finding system integration points
- Business Process Analysis Steps
- Job and process interaction

The benefits of element analysis are:

- Do more with existing staff
- Improve service levels
- Improve compliance & risk management
- Support for portfolio rationalization
- Focus existing resources on servicing and delivering to the critical business priorities
- Leverage core IT assets and rationalize your software portfolio

### **Information Management at Mayo Clinic**

*Tim Brennan*

*Data Administrator*

*Mayo Clinic*

The Mayo IM Program Framework comprises:

- Institutional Data Repositories
- Institutional Data Architecture
- Standard Data Infrastructure
- Coordinated Data Stewardship
- Operational Support
- Program Management

Key take home messages are:

- Answer “why” with muscle
- Define a framework that works for your company
- Always be prepared
- Don’t underestimate the power of communication & education
- Manage relationships

## **Ready to Use XML Metadata Frameworks – An overview**

*Peter Aiken*  
*Founding Director*  
*Data Blueprint*

*M. David Allen*  
*Director of Technology*  
*Data Blueprint*

This talk provided an overview of XML Frameworks and their value from a data management perspective, including industry specific frameworks such as RosettaNET (Electronics), Acord (Insurance), Envera (Chemical) as well as more generalized frameworks such as ebXML and bizTalk. The conclusions are that for well-specified domains, these frameworks provide immediate utility, complimenting existing data management efforts. If implemented correctly, these frameworks also provide a sound basis for the development of information architecture components including basic data modeling templates, standard data items, domain information, and inter and intra organization metadata.

## **Governance and Synchronization of Shared Metadata: a User Perspective**

*Jeff Dirks*  
*President*  
*SchemaLogic*

*Conrad Gottfredson*  
*Enterprise Content Architect*  
*LDS Church*

**How Do You Get to Where You Need to Be with Enterprise Content Asset Management?**  
(Guiding Principles)

- ECAM does not require the stopping or slowing down of any single initiative. It does require foresight and decisions that will allow their eventual integration.
- Acquiring this foresight requires cultivating a common vision across the enterprise.
- The cultivation of this common vision most naturally occurs during the establishment of a common metadata architecture.
- This common architecture must attend to metadata that addresses content management and usage as well as access requirements.
- Content management and usage requirements are best identified through a content analysis process.
- The content analysis process will not only yield information needed to define the metadata architecture and help grow a common vision of ECAM, it will also help identify functional requirements that span current initiatives.
- These cross-functional requirements can help guide decisions that might otherwise restrict full-featured integration of all these initiatives.

## **Beyond Data About Data**

*Kevin McDearis*  
*Vice President of Data and Delivery*  
*Checkfree*

*James Baker*  
*Data Architecture Manager*  
*Checkfree*

### **Process & Data Principles**

- Processes support Core Values
- Process and all work activity occurs at the furthest possible point upstream
- Organizational boundaries do not limit process planning
- Processes & data must be agile, adaptable, and extensible
- Process inputs and outputs and data sources/outputs are clearly defined & usable upon delivery
- Incorporate continuous learning & improvement elements into all processes and work activity
- Use proven methods and tools to ensure customer satisfaction
- Data is a corporate asset with defined attributes; information derived from understanding data provides a strategic competitive advantage
- Process results are reproduce-able, repeatable & not dependent on individuals

## **Patterns and Practice: Data, Object, and Interface Design**

*Deborah Coleman*  
*Assistant Professor*  
*Rochester Institute of Technology*

*Bill Murray, Software Analysis Manager, Paychex Corporation*  
*Carla Lorek, Manager, Communications, Xerox North America*

Data and process requirements are at the heart of data analysis, object-oriented analysis and interface design methods. Since the dependency of requirements is laden upon customers, it is important to identify areas of synergism that are possible between these methods and the methodologists that employ them. This presentation introduced synergistic opportunities and patterns between data, object-oriented, and interface design methods.

Object/Relational Differences include:

Similarities between Methods

- Both diagrams Depict data
- Class and Entity Attributes
- Depict structure between data
- Cardinality and Optionality

Cultural Differences

- Object Modelers believe relational databases are not necessary
- Relational Modelers want the object models to reflect the relation schema

Technical Differences

- Prohibit a one-to-one mapping of constructs
- Inheritance, polymorphism, encapsulation, specialization, and directional associations are not supported in relational models
- Different set of data types
- Different method to uniquely identify items
- Object Modeling is based on software engineering principles
- Relational Modeling is based on mathematical principles

**CONFERENCE SESSIONS**

**Wednesday, May 25**

**11:15 am – 12:15 am**

**Four Basic Steps to Data Security -- Shielding Your Enterprise in a Highly Regulated Environment**

*Sherman Lee*  
*Data Architect*  
*A.G. Edwards*

*Becky Kirkpatrick*  
*Associate Vice President*  
*A.G. Edwards*

Key enablers for a robust and comprehensive data security strategy are:

- Embracing a systematic approach to data security
- Achieving not just “business involvement” but “business ownership” of the requirements for securing corporate data

Simplify a very complex business problem of securing data consistently, by breaking it down to four basic steps:

1. Classifying the risk of each corporate data item
2. Specifying the access to data by roles
3. Implementing consistent security measures
4. Monitoring and auditing to ensure compliance

Some valuable lessons:

- Having a granular perspective of data security requirements enables the business to be responsive to business, environment and regulatory changes
- Involve the business in the commitment to address security at a much more granular level than the traditional siloed application and database security solutions.
- Determine the appropriate scope based on return on investment for your firm
- Realize immediate benefits to undertaking such a granular data security initiative:
  - o Requirements for desensitizing data
  - o Standard data dictionary for your enterprise
  - o Business ownership of data

## **Implementing a Metadata Repository at a Data Driven Company**

*Ray McGlew*

*Enterprise Information Architect Consultant*

*Software Methods and Faculty Practitioner, University of Phoenix*

### **Key Lessons Learned:**

- Communicate
  - Team, Management, Peers, Future Users
- Promote throughout organization
  - Get in front of staff meetings
  - Keep gathering requirements-
- Dedicated team
  - Tendency to “borrow” staff
  - Hard to manage goals
- Manage expectations
  - Communicate roll-out strategy
  - Lower resources = less progress
  - Quickly reflect new challenges
  - Quickly reflect new organizational goals
- Project Management
  - Monitor budget
  - Identify scope changes
  - Manage impacts
  - Produce useful status reports

## **An Enterprise Approach to Metadata Management (or How We Learned NOT to Manage Metadata)**

*Dan Paolini*

*Director, Data Management Services*

*State of NJ - OIT*

A sound Metadata architecture incorporates five general characteristics. It should be...

- Integrated
- Scalable
- Robust
- Customizable
- Open

In addition to the general characteristics of good architecture, all good data repositories share a set of key elements that are essential for success, regardless of the architectural approach used to build the repository. In short, all good repositories:

- Are based on clear, well-defined management direction
- Use the same front end as the data warehouse
- Use the same entity and attribute naming standards throughout
- Incorporate multiple sources of Metadata
- Include automated and reusable processes

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

- Use a standardized integration process
- Use a flexible meta model
- Manage multiple versions of Metadata
- Incorporate update facilities
- Use a component-based multi-tier architecture
- Incorporate a security management scheme
- Incorporate cross-tool Metadata dependency and lineage

**The Semantic Web: What is it and Why Should You Care?**

*Dave McComb  
President  
Semantic Arts*

The Semantic Web (SW) is not a different web, but an extension of the current one. It is essentially the addition of “meaningful” tags to content on the web, such that software can help us find things. The power of the SW comes from a set of standards, each of which provides just a few capabilities. Semantic Web = shared ontologies + tagged content + inference engines + service composition. Key to the SW is RDF - Resource Description Framework - comprising Subject/Predicate/Object “Triples” and “Triple Stores” which allows the user to make assertions and merge identities.

**PANEL: The Changing Role of the Business Analyst/System Designer**

*Gil Laware  
Assoc. Professor  
Purdue University*

*Andrew Walters, Associate Professor, Purdue University  
Jack Hilty, Managing Principal, SentientPoint, Inc.*

What new demands will be placed on the role of the Business/Systems/Data Analyst in the Service Oriented Architecture (SOA) World? Data Management professionals can transition to the new roles required by SOA by acquiring and enhancing their skills. In an SOA environment, the following will take on greater importance:

1. Acquiring business knowledge as well as technical knowledge.
2. Defining services and incorporating them into service oriented processes.
3. Translating or interpreting business models into systems models using Unified Modeling Language (UML).
4. Mastering new skills (facilitation, organization, communication, technical, etc.)

Using their existing pre-SOA skills and newly mastered ones appropriate for an SOA environment, Data Management professionals can provide business solutions; effectively simplifying, designing, and communicating business models and optimizing current and evolving technologies.

**CONFERENCE SESSIONS**  
**Wednesday, May 25**  
**1:15 pm – 2:15 pm**

**Developing Enterprise Architecture Strategies**

*Jane Carbone*  
*Partner*  
*infomajic LLC*

In this presentation, we looked at where enterprise architecture (EA) strategies fit in the EA process. Sometimes models and standards need further explanation of intent. For example, we may depict a messaging service in our EA models and define standard middleware to be used for implementation. But if the intention of using messaging is not clear, we may want to develop a strategy. The strategy may constrain what standard is selected.

A strategy is a small set of succinct statements or very brief position papers that each address a specific, significant change implied in the target state architecture, e.g., messaging. Strategies are “optional” - necessary only where intent needs to be clarified or standards constrained. An individual strategy ought to be a short story versus a novel – it is not a technical specification.

We identify where a strategy is needed by:

- Analyzing target state architecture models for implied strategies
- Assessing if there are current organization-specific considerations that require additional explanation, e.g., Security, Network-centric, Outsourcing, eBusiness
- Reviewing all the “core” components of target architecture—data, functions, technology and people—and assessing where strategies are necessary to constrain decision-making.

After we have identified where strategies are needed, we construct them. An individual EA strategy should include:

- Purpose or intent of change - How it supports the business target state
- What is required – what is implied
- High-level costs and benefits
- How it supports/where it fits in the target architecture.

**Data Privacy - Legal Risks and Ethical Considerations**

*Alan Wernick*  
*Attorney*  
*Querrey & Harrow, Ltd.*

Privacy laws and regulations are increasingly affecting how companies (and government agencies) obtain, use, and distribute data. Failure to comply may result in legal liabilities including regulatory oversight of your business, monetary damages, loss of data, and criminal penalties. This presentation provided an overview of legal and ethical considerations of data management including privacy, infringement, and data negligence liabilities.

Consumers think the most important elements to be verified in order to increase confidence:

- that security procedures are adequate (90%);
- that the company not release customer personal data without permission or unless required by law (89%);
- that access to personal data within the company be limited (84%);
- that the company only collect customer information that its privacy policies indicate (84%); and
- that information use or sharing follows stated privacy policies (81%).

## **Partnering with Application Development: Integrating Data Development with Iterative Design Methodologies**

*Katherine Sivier  
Senior Data Analyst  
Caremark, Inc.*

Information Technology has turned to methodologies such as Agile and RUP as a way to quickly implement new functionality to software applications. These methodologies support the need for applications to be flexible. However, they work at cross-purposes to the development of a data design to support the business's requirements. Operational systems require data to be a stable resource. This presents a challenge for the data analyst: how to create a cohesive data design within an iterative design environment. Lessons of experience with agile development:

- Every project will be different and needs to be handled as unique
- Application developers' focus on a narrow task is in conflict with flexible data
- Refactoring data is not a simple task
- An accurate physical data model is still a critical tool

## **Combining Business Metadata Delivery with Knowledge Management**

*Jim McQuade  
Data Administrator  
Giant Eagle, Inc.*

Knowledge management systems are ones that harness technique and technology to collaborate and facilitate knowledge transfer. We've been able to compress dissemination of proven practices from weeks to days. Using knowledge management to disseminate metadata is a relatively new application, but it meets all of our key criteria for what knowledge management is good for. Some people have said that metadata is the glue that keeps all the different systems of an enterprise working together, and "working together" is the key idea in collaboration. So, in the way that we hope to combine knowledge management with metadata delivery, we hope we can make detailed information about critical data concepts readily available throughout our enterprise, or, at very least, to our target audience. One of the big ideas is that knowledge management does not work without people who are accountable for making it work, at all stages of the knowledge creation chain. Knowledge management can also facilitate Data Capability Maturity both through actually writing stuff down, knowing where to go to find it, and also in managing your policies and practices. Do not spend a lot of money on hardware, software, packaged systems, or any other of the traditional capital expenses

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

associated with information systems. Without the enabling conditions found in your people systems, knowledge management won't work. Organizational culture is the key enabling factor, and most cultures have a long way to go. A Learning Culture, one that values true team learning and ego-less behavior, is critical for the exchange and sharing of information with one goal and one goal only: the growth of the enterprise.

## **Using Benford's Law for Business Rule and Metadata Discovery**

*David Loshin  
President  
Knowledge Integrity, Inc.*

Benford's law is an interesting phenomenon noted in the 1920's by GE physicist Frank Benford, associated with a logarithmic size law associated with number distributions. For data sets that conform to a certain set of criteria, there is a predictable, yet counterintuitive distribution of initial digits of numbers. Other interesting logarithmic laws and principles can be applied to your data profiling/metadata discovery program.

Metadata may be enhanced through progressive integration of natural size laws:

- More complex mathematical functions
- Newer approaches to identifying business rule compliance
- Integrating natural behavior laws with Business Intelligence techniques

## **PANEL: The Next Big Thing in Data Management**

*Tony Shaw  
Principal  
Wilshire Conferences*

*John Ladley, Navigant Consulting  
Karen Lopez, InfoAdvisors  
Mike Gorman, Whitemarsh Information Systems  
Jeff Pollock, Cerebra*

This lively session invited the diverse views of a group of experts in the industry, as well as the audience, on the next major trends in data management. The areas the panelists raised were:

- Information valuation - increased measurement of the value of information and data, if not on the balance sheet, then at least as an internal management metric. (John Ladley)
- Ontologies - seen as the next enabler for better data integration, metadata management, business intelligence and search, among other applications. (Jeff Pollock)
- Professional Certification - the commitment by DAMA to bringing about recognized professional credentials was seen as increasing the recognition of value which data management brings to the enterprise, and increased opportunity to the individuals within the profession. (Karen Lopez)
- There is no silver bullet except for good management practice, including extensive use of re-use and templates (and generating code from the templates), and learning from the lessons of previous practitioners. (Mike Gorman)

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

Also discussed at length was the impact of RFID which will impact many in the audience in the near future. Key issues are the huge quantity of sub-transactional data which will be generated (raising questions of where to store it, how to mine it and what exactly to do with it?) as well as a number of socio-regulatory issues related to privacy and identity.

## **Normalization - Achilles Heel of Data Modeling**

*Gordon Everest  
Professor Emeritus, and Adjunct  
University of Minnesota*

The bottom line to normalization is:

- In a record-based (ER) data model, there are both inter-record (explicit) and intra-record (implicit) relationships.
- So, start with individual domains for all entities and attributes, and explicitly represent all relationships.
- Build up your model from elementary facts
- Thus, defining all functional dependencies
- Enabling the system to produce a normalized data structure
- If you don't a priori cluster attributes into records, you won't ever need to normalize!

### **KEYNOTE PRESENTATION**

**Wednesday, May 25**

**2:45 pm – 3:45 pm**

## **Ethics After Enron: Why Taking the High Road is Good for Data Management**

*Bruce Weinstein  
CEO  
Ethics at Work, Inc.*

Bruce gave everyone a nice break from the usual conference "all data all the time" themes with his entertaining insights and solutions to a number of ethical situations, based on his framework of 5 essential "Life Principles." (also the title of his new book). These principles are:

- Do No Harm
- Make Things Better
- Respect Others
- Be Fair
- Be Loving

## SPECIAL INTEREST GROUPS

Wednesday, May 25

5:30 pm – 6:30 pm

### An Introduction to Eclipse and the Eclipse Modeling Framework

*Robert Kitzberger*

*IBM DB2 Information Integration-Strategy*

*IBM*

Eclipse is:

- A Technology Offering
  - o Universal platform for building and integrating tools through common UI and common metadata
  - o Open, extensible architecture based on standards
  - o Licensed for royalty free world-wide distribution
- An Open Source Project
  - o Delivers and supports the technology
- A Community
  - o Contributes to the project and technology
  - o Brings to market commercial offerings
- Started with \$40M software/R&D from IBM as initial Eclipse technology in Nov 2001.

### Fundamentals of the Capability Maturity Model

*Anne Marie Smith*

*Lead Instructional Consultant*

*LearnQuest*

CMM is based on the concepts of Watts Humphrey and Philip Crosby. It outlines a series of process improvement levels that an organization can achieve by following sets of best practices. It is not a methodology, rather a framework to be applied to manage the process of developing and implementing software. Each level of the CMM (2-5) has a set of Key Process Areas, whose application in an organization can point to that organization's level of software development maturity. KPA's are issues to address to reach each maturity level and related activities for achieving goals. There are no KPA's for Level 1, since this is the foundation level.

Recommendations for Continual Process Improvement

- CMM is an excellent process improvement model for Software Development
- Commit to applying a consistent process improvement model to all IS processes
- Learn about the CMM through orientation sessions
- Perform an initial CMM assessment and develop plans for applying assessment results
- Adopt a redesign approach to all IS processes, based on results from your CMM assessment
- Train technical managers in quality management principles
- Apply best practices to all Information Systems efforts, using the CMM

## **Enterprise Architecture in ACTION! Supercharging your Local DAMA Chapter**

*Derek Strauss  
President & CEO  
Gavroshe USA Inc.*

We were able to kick off multiple development initiatives:

- Event Registration
- Member Registration
- Member Tracking

We have developed:

- An EA Process and the ability to prioritize our systems efforts
- A common Blueprint for multiple development groups
- A data architecture to ensure our membership database is properly architected and managed
- A schema for storing all the design and implementation artifacts using the Zachman Framework

We now have CONTINUITY – we have documented our enterprise’s fundamental architecture – and the EA is enduring beyond the team that developed it. This initiative created excitement and gave the Board and the Membership an opportunity to rally around a common enterprise design, leading to better consistency and improved operations – the SWOC DAMA enterprise has been supercharged!!!

## **Electronic Health Record Strategy: Evaluating Metadata with a Web-based Registry**

*LuAnn Whittenburg  
Certified Nurse Practitioner, Health Informaticist  
Computer Technology Associates, Inc.*

*Marco Johnson  
Deputy Director, Operations, Information Management  
Department of Defense (Health Affairs)*

- Some new versions of standards do not identify a delta between the versions, so a metadata difference must be determined
- Due to the many points where errors can be made, the Quality Assurance responsibility puts a tremendous load on the team analyzing the standard, determining the differences between the versions, reformatting the metadata into the standard format, and loading the information
- Changes to standards may be needed to correct errors. Changes may be in the structure of the Standards
- Changes may have a cascading affect that may change meaning
- Changes need to be reviewed to determine if data elements still meet the Consolidated Health Informatics needs
- [www.usihk.org](http://www.usihk.org) online NOW!

## CONFERENCE SESSIONS

Thursday, May 2

8:30 am – 9:30 am

### **Data Modeling Layers - Communicating to everyone in their native language**

*Michael Nicewarner*

*Data Analyst*

*John Deere Credit*

There are many audiences for your data models...business users, various IT professionals (including requirements analysts, business analysts, DBAs, and management)...and all of them have a stake in the contents of your models. But selecting the correct model type and level of detail is critical to effective communication. Mike discussed the various types of models and summarized his conclusions briefly as follows:

- There are a number of styles and notations
- Establish corporate standards
- Pick appropriate style/form
- Need tools that can “tie it all together”
- Communicate

### **The Practitioner’s View of Business Rules: The Latest in Practice from the Front lines**

*Ronald G. Ross*

*Principal*

*Business Rule Solutions, LLC*

This session discussed successful techniques for applying business rule techniques in business analysis, business vocabulary development, data modeling, and requirements capture.

The business rule approach offers innovative, yet practical answers for engaging business people in successful data-oriented initiatives. New light is being shed on core concerns of data professionals, particularly data definition and analysis. The bottom line is higher-quality and more adaptive solutions for the business.

### **Just When We Thought It Was Safe: New Technologies and the Metadata Future**

*Evan Levy*

*Partner*

*Baseline Consulting*

Key technologies in the future of metadata are:

Extract Transformation & Loading

- Development software supporting the integration and transformation of data.

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

**Enterprise Application Integration**

- A code package providing application data access with standard interfaces.  
Dramatically simplifies application-to-application data access.

**Enterprise Information Integration**

- An integration technology that combines data from multiple systems “on the fly” without storing it.

**Data and Meta Data of a Different Kind**

*Gail Ritchie*

*Project Leader, Meta Data Repository Project*

*Bank of Canada*

*Suzanne Young*

*Assistant Director, IT Services*

*Bank of Canada*

The session presented the statistical metadata repository recently implemented at the Bank of Canada to help staff locate the statistical resources available to them. The repository describes both structured and unstructured statistical resources. The presenter reviewed:

- the scope of the project
- the business and technical environments
- the application

Major points raised include the business and IT relationship at the Bank of Canada and the importance of learning and sharing across disciplines.

**Structured and Unstructured Data Integration: A Case Study**

*Mark Popivchak*

*Senior Systems Architect*

*Booz Allen Hamilton/ASE*

*Kristopher Schroeder*

*Associate*

*Booz Allen Hamilton*

The primary reason for integration is that there may be critical business information buried in either a structured or unstructured silo. Querying one or the other may miss that info. Querying both takes extra effort. Benefits of unstructured and structured data integration:

- Federated Query
- Normalized Access
- Agile Data environment: decouple Data and application layer = Reduced O&M
- Better Analysis, sharing of data across domains

Examples:

- Pharmaceutical: Gene research – where does protein A effect a reaction with enzyme B
- Financial: Acquisitions and Mergers – three corporations with 3 different financial databases merged and they need a consolidated report, on-demand.
- Federal: Logistics Agency: Match databases of parts and equipment with manuals and failure reports

## CONFERENCE SESSIONS

Thursday, May 26

9:50 am – 10:50 am

### Success Story, Realizing Value of Data Management, Enterprise Architecture, and Portfolio Management

*Paul Fairbourn*

*Sr. Advisory Engineer*

*Idaho National Laboratory (INL)*

*Dave Brown*

*Sr. Consulting Technical Specialist*

*Idaho National Laboratory (INL)*

The Idaho National Laboratory (INL) has implemented a Data Management, Enterprise Architecture and Portfolio Management program that has consistently provided value to internal customers by:

- Providing an integrated repository that supports Data Management, Enterprise Architecture and Portfolio customer needs.
- Delivering baseline of the as-built Information Technology configuration and assets
- Providing an enterprise model identifying valid data sources
- Identification of software development standards
- Communication of repository information to stakeholder groups
- Supporting audits, ad hoc requests, and transition activities.

Benefits realized are:

- \$4 million cost savings realized to date
- Provides information for improvement opportunities
- Improved decision making
- Used outside of Information Technology (IT)
- Baseline activities are well established
- Value is no longer questioned
- Supported contractor change and company split
- Integration with IT Portfolio

Conclusions thus far:

- Although we are taking a moment to declare success, there is still much to do.
- The Architecture framework provides a structure to manage a complex environment and to plan for the future
- Minimize data redundancy – drive to official data sources
- Reduce technologies, e.g., DBMSs, development tools, etc.
- Improve delivery time and product quality

## **When to use a Rules Engine (or Why Business Rule Management Matters)**

*Paul Vincent*  
*Product Manager*  
*Fair Isaac Corporation*

Rules engines should be strongly considered by anyone who needs to automate business logic that has any of the following characteristics

- Changes frequently
- Needs to be deployed across different systems
- Needs to be maintained and/or monitored by business managers
- Includes complex logic that is difficult to define in a procedural language

## **What Are You Worth? Why Companies are (Re)Investing in Data Management**

*Bryan Rockoff*  
*Director*  
*Baseline Consulting*

We (particularly those in the data warehouse world) have to get away from the notion that data is static and lives out its life on a big platform. The fact is, data is dynamic and is moving around organizations all the time. So, where should you begin to provide business value?

- Measure "baseline" data quality
- Analyze the moving parts
- What data management processes are still manual?
- What data skills need to be acquired?
- Quantify the cost of bad/missing/inaccurate data
- Create data-specific job descriptions
- Begin data profiling of key source systems
- Devise a Data Qualification Questionnaire for new projects
- Volunteer to perform data audits for key applications
- Assign data stewards to key subject areas

## **Promoting Metadata to full Data Status**

*David Downs*  
*Business Consultant*  
*CTG*

Meta-data is data, so good data management practices should apply. Yet we abandon them when it comes to metadata. If we are frustrated with the success of our metadata management efforts it may be because we are underestimating the "data" nature of it.

Meta Data Management is a critical component for today's enterprise Data Warehouse and Web-deployed applications. Meta Data is the foundation that defines the business structure of these applications and their lineage back to the originating enterprise and organizational

systems. The repository provides the foundation for acquiring, maintaining, and exploiting disparate Meta Data throughout the enterprise from a common management view. It reduces application development and maintenance costs and helps organizations derive more value from existing information assets.

## **Managing Metric Meta Data at Allstate**

*Doug Stacey*  
*Manager*  
*Allstate Insurance Co.*

Why manage Metric Metadata?

- Need for users to better understand their data
- Current process allows for too much 'interpretation' of what a number means
- Know where it came from
- Moving "information" to "knowledge"
- Provide solid benefit to the business
- Reduce rework – save time and money!

Benefits are:

- Metrics defined as a combination of metric components that can be metrics themselves allows for storing and managing both simple and complex metrics.
- Metric components are shared and defined only once for all metrics that are using them (single definition of an item across all of its uses). Promotes reuse and consistency and allows for impact analysis.
- Flexibility, new operators and constants can be added by the users as needed if the existing ones can not describe a new metric.
- Allows common understanding between business users
- No more wasted time at the start of a meeting discussing what numbers are being used, where they came from, and what they really mean.
- Reuse = cost savings
- Raises the stature of Data Management as a true contributor of value to the business!

## **Contextual Data Extraction: Data Elements and Business Value from Free Text**

*David Allen*  
*Director of Technology*  
*Data Blueprint*

*David Rafner*  
*Program Manager*  
*Data Blueprint*

What's the difference between search and extraction? In search, finding something based on criteria is the goal of the game. In extraction, cataloguing some aspect of data is the goal.

**DAMA Symposium + Wilshire Meta-Data Conference  
Orlando, May 22-26, 2005**

Free text search tends to be lousy at complicated queries that cut across lines. (“Show me all documents where Person X is mentioned, but not as an author”) and is better at locating a document, not as good at locating classes of well-specified documents.

Extraction allows bits and pieces to be recombined in any way, because it’s not unstructured any more. It’s structured through extraction. Text extraction provides a bridge between two types of uses – free text and structured databases.

**PANEL: DAMA Model Curriculum Framework for Data Management**

*Deborah Henderson*

*Data Architect, Cap Gemini*

*Chair, DAMA-I Curriculum Committee*

*Anne Marie Smith & Pat Cupoli, DAMA-I Curriculum Committee*

Curricula exist for many computing and technology disciplines but none for information/data management. The lack of information/data management curriculum leads to lack of understanding of the value and role of information in an organization’s successful performance.

- Goal: Gain endorsement (2005) from the following key accreditation bodies for North America:
  - ABET - Engineering
  - AACSB - Business
  - ACM - Computer Science
- The Committee will engage these Groups in a review and endorsement process for the Curriculum Framework
- These endorsements will be critical to adoption of the Curriculum by post secondary institutions

**KEYNOTE PRESENTATION  
Thursday, May 26  
11:10 am – 12:30 pm**

**PANEL: Metadata Best Practices**

*Robert Seiner (moderator)  
President and Principal, KIK Consulting  
Publisher, TDAN.com*

*Todd Stephens, BellSouth  
Art DeSilva, RBC Financial Services  
Doug Stacey, Allstate Insurance  
Nikhil Sharma, Intel Corporation  
Kevin McDearis, CheckFree*

The closing keynote included a panel of industry-leading metadata practitioners whose companies have excelled in the Wilshire Conferences Metadata Best Practice Award competition over the past several years. The panel was enjoyed by several hundreds of attendees that became actively involved in an informal and open forum.

The panelists started by briefly describing several of their company's "Off the Beaten Path" metadata best practices, or best practices that they discovered while getting their metadata programs off the ground and driving them toward successful implementations. Several panelists mentioned compliance procedures, change management/notification, leveraging enterprise models (not EDM), and metadata branding as best practices that became instrumental components of their award winning implementations.

Several microphones circled the floor for more than an hour as thought-provoking questions and answers were shared by panelists and attendees.

The moderator closed the program by posing a question to all regarding the long-term use of the term "metadata". The first response from an attendee was that the term can not be retired without the data world coming to an agreed upon spelling, once and for all. The panelists wrapped up the event by agreeing with the attendees that the term metadata will be an important term for the foreseeable future as the use of the term and the use of metadata will continue to evolve. Everyone agreed that the Wilshire/DAMA annual event will continue to be the leading event that covers the subject matter.

**SEMINARS  
Thursday, May 26  
2:00 pm – 5:00 pm**

**Presenting Data Management to the Enterprise**

*Michael Scofield  
Assistant Professor  
Loma Linda University*

Data management, as a function in many organizations, is fighting for its life. The value which may be obvious to most “data bigots” as Mr. Scofield describes DAMA people, may not be obvious to many executives, application developers, and hardware-oriented I.T. personnel. So for DM to be successful and survive, it must find new ways to bring value to the enterprise, and deliver that value quickly.

Mr. Scofield introduced a five-level model which was used throughout the discussion. It has, from top to bottom, the business, the data describing the business, the applications which capture, store, and display that data for micro-decisions, the operating system which forms an environment for the application, and the hardware infrastructure.

One way to create new value is to create a data asset inventory. Part of such an inventory would be an assessment of the quality of the data asset, and its completeness.

Data integration is a major way in which DM can contribute to ensuring success of data warehousing and application integration projects. Understanding the differences in source data architectures, and aiding in the resolving of those differences in the design of the target database can eliminate one potential point of failure in a project.

Moving an enterprise from a legacy business data architecture to an ERP package also involves significant data architectural issues, often overlooked in superficial consideration of the functional requirements for the package selection. The function may be there, but competitive advantage may be lost in altering the architectural components of the relationship to the customers.

Delivering the value quickly is not always obvious to a culture which tolerates a long development life cycle. Few companies now will tolerate an 18-month project to build a corporate data model, and wait for payback even after that. Quick payback on data knowledge can be accomplished, in part, by creating an environment where the data stewards and meta-data experts have quick read-only access to the data asset throughout the organization, and can create ad hoc reports quickly to meet spontaneous questions from executives and business functions all over the organization. This role does overlap with data warehousing, of course.

In data warehousing, it is essential to make the information created in the DW processes relevant to the culture of the business. This includes employing cultural sequencers to present the data in a consistent way across all the delivery mechanisms (reporting and graphics).

Finally, Mr. Scofield discussed ways of communicating to the rest of the organization about what DM can do, and has done. A variety of media and methods are employed for such

communication. But it requires some entrepreneurial behavior on the part of DM practitioners who often are more reclusive and introverted. It means understanding your message and being ready to deliver it in any form suitable to the situation. It means being creative in getting “out of the cubicle” and engaging the rest of the organization, looking for new ways to add value with the meta-data assets and the skills you have.

## **Data Certification: A Collaboration Method for Effective Information Resource Management**

*Andres Perez*

*Senior Information Management Consultant*

*IRM Consulting, Ltd. Co.*

This workshop presented the Data Certification approach and associated tools and techniques for managing information across the enterprise. Common Data Certification pitfalls are:

- Starting implementation without management accountability
- Starting with procedures and guidance only
- Avoiding difficult or complex but critical impact issues
- Creating (or attempting to create) a large inflexible organization
- Staffing the core team without proper skills
- Assuming that the process will work “fine” (no need for improvement)
- Failing to manage the participants’ time effectively
- Limiting Participation to IT Areas only
- Limiting Participation to Business Areas only
- Focusing on tools, techniques, or mechanics instead of process & business value
- Waiting for (or building) the Repository