Overview of Enterprise Data Architecture – What's In YOUR Data Architecture?

Anne Marie Smith, Ph.D.
Principal Consultant, Director of Education
AMSmith@ewsolutions.com

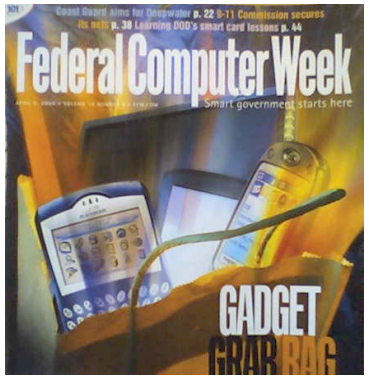
EWSolutions

EWSolutions is a Chicago-headquartered strategic partner and full life-cycle systems integrator providing both **award winning** strategic consulting and **full-service implementation services**. This combination affords our clients a full range of services for any size enterprise information management, managed meta data environment, and/or data warehouse/business intelligence initiative. Our notable client projects have been featured in the Chicago Tribune, Federal Computer Weekly, Crain's Chicago Business, and won the 2004 Intelligent Enterprise's RealWare award, 2007 Excellence in Information Integrity Award nomination and DM Review's 2005 World Class Solutions award.

Information Integrity Coalition

2007 Excellence in Information Integrity Award Nomination

Chicago Tribune



DM Review 2005

World Class Solutions Award Data Management

intelligent enterprise REAL Awards TRANSFORM

2004 WINNER

Best Business Intelligence Application Information Integration Client: Department of Defense

For more information on our Strategic Consulting Services, Implementation Services, or World-Class Training, call toll free at 866.EWS.1100, 866.397.1100, main number 630.920.0005 or email us at Info@EWSolutions.com

EWSolutions Partial Client List

Arizona Supreme Court
Bank of Montreal
BankUnited
Basic American Foods
Becton, Dickinson and Company
Blue Cross Blue Shield companies
Branch Banking and Trust (BBandT)
British Petroleum (BP)
California DMV
College Board
Corning Cable Systems
Countrywide Financial
Defense Logistics Agency (DLA)
Delta Dental
Department of Defense (DoD)
Driehaus Capital Management
Eli Lilly and Company
Federal Aviation Administration
Federal Bureau of Investigation (FBI)
Fidelity Information Services
Ford Motor Company
GlaxoSmithKline
Harris Bank
The Hartford
Harvard Pilgrim HealthCare
Health Care Services Corporation
Hewitt Associates
HP (Hewlett-Packard)
Information Resources Inc.
International Paper
Janus Mutual Funds
Johnson Controls
Key Bank
LiquidNet
Loyola Medical Center
Manulife Financial
Mayo Clinic
Microsoft
National City Bank
Nationwide
Neighborhood Health Plan
NORC
Physicians Mutual Insurance
Pillsbury
Quintiles
Sallie Mae
Schneider National
Secretary of Defense/Logistics
South Orange County Community College
SunTrust Bank
Target Corporation
The Regence Group
Thomson Multimedia (RCA)
United Health Group
United States Air Force
United States Navy
United States Transportation Command
USAA
Wells Fargo
Wisconsin Department of Transportation
Zurich Cantonal Bank

GSA Schedule
Contract GS-35F-0453M

SeaPort-e

For more information on our Strategic Consulting Services, Implementation Services, or World-Class Training, call toll free at 866.EWS.1100, 866.397.1100, main number 630.920.0005 or email us at Info@EWSolutions.com

Instructor Background

Anne Marie Smith – AMSmith@ewsolutions.com

- Internationally recognized expert and speaker in the fields of enterprise information management, data modeling, meta data management and data warehousing, one of the top industry experts in data governance and information management strategy and planning
- Over 20 years experience in delivering solutions in enterprise information management for numerous companies / clients across varied industries
- Published over 50 articles that have been featured in some of the industry's most prestigious magazines and newsletters
- Contributing author to the DAMA Data Management Body of Knowledge (DMBOK)
- Has taught at several institutions of higher learning including LaSalle University
- Holds certifications in PMP, CDPM and has earned a PhD in Management Information Systems

Agenda

- What Is Architecture?
- The Zachman Framework
- Enterprise Architecture
- Enterprise Data Model
- Information Value Chain Analysis
- Data Delivery Architecture
- Managing Your Data Architecture

Enterprise Data Architecture

- "Enterprise data architecture is
  - An *integrated set* of specification *artifacts*
  - That define strategic data requirements,
  - guide integration of data assets
  - And align data investments with business strategy." (***DAMA-DMBOK***)
- Master Plan / Blueprints
  - For enterprise integration and alignment

- Your Organization's Treasure Map
  - A guide to unlocking business value
- Data architecture itself is a precious knowledge asset
  - "Ability to fully leverage your information assets? *Priceless!*"

Enterprise Data Architecture

- An integrated set of specification artifacts
  - Data Architecture is part of a complete enterprise architecture
  - What would be in a comprehensive enterprise architecture?
- Managing architecture is a significant investment
  - Building, maintaining, enriching and using architecture
  - No organization maintains a comprehensive collection - choices
- Decisions must be made
  - Which artifacts to build and maintain?
  - What level of detail to define and maintain?
  - What business risks need to be managed?
  - What will it cost and whose time will it require?
  - What artifacts do we need the most ***today***?
  - How to build our collection iteratively over time?

"What's In YOUR (data architecture) wallet?"

What Is Architecture?

- "The design of any complex object or system"
  - Enables management of complexity
  - Inherent organization of natural things – biology, geology, mathematics
  - Design of human-made things – buildings, music, literature, machines, *organizations, processes, software, databases, semantics*
  - Macro Level – city planning, the universe
  - Micro Level – Machine parts, computer chips, atoms
  - Abstraction of the system – not the system itself
  - The more complexity, the greater the need / value

- "An organized arrangement of component elements…
  - …to optimize function, performance, feasibility, cost and/or aesthetics"
  - Helps attain a goal
  - Addresses requirements and constraints
  - Requires both analysis and design

- Also a skill – an art – a discipline – a field of study – a profession

Architectural Frameworks

- Ways to Think About and Understand Architecture
  - "Architecture for Architecture"

- Including:
  - The Zachman Framework For Enterprise Architecture
  - TOGAF – The Open Group Architecture Framework
  - RM-ODP - Reference Model of Open Distributed Processing (ISO/IEC 10746)
  - ANSI/IEEE 1471-2000 "Recommended Practice for Architecture Description of Software-Intensive Systems"
  - PRISM Architecture Framework (1986 -- Hammer, Champy and Davenport)
  - CAP Gemini and other consulting firms
  - Government and Defense Frameworks
    - FEA – US Federal Enterprise Architecture – from the Office of Management and Budget
    - DODAF - US Department of Defense Architecture Framework
    - MODAF -- The UK Ministry of Defence Architecture Framework
    - AGATE -- The France DGA Architecture Framework
    - GEA – Government Enterprise Architecture – Queensland, Australia provincial government

The Zachman Framework For Enterprise Architecture

- The most widely known and adopted architectural framework
  - "A logical structure for identifying and organizing the descriptive representations (models) useful in the management of enterprises and the development of their systems (automated and manual)." (John Zachman)
  - A generic classification scheme for designing any complex system, not just enterprises and information systems
  - White paper published in IBM Systems Journal, 1986 – still available!
  - Studied the fields of architecture and construction (buildings) and aerospace engineering (airplanes)

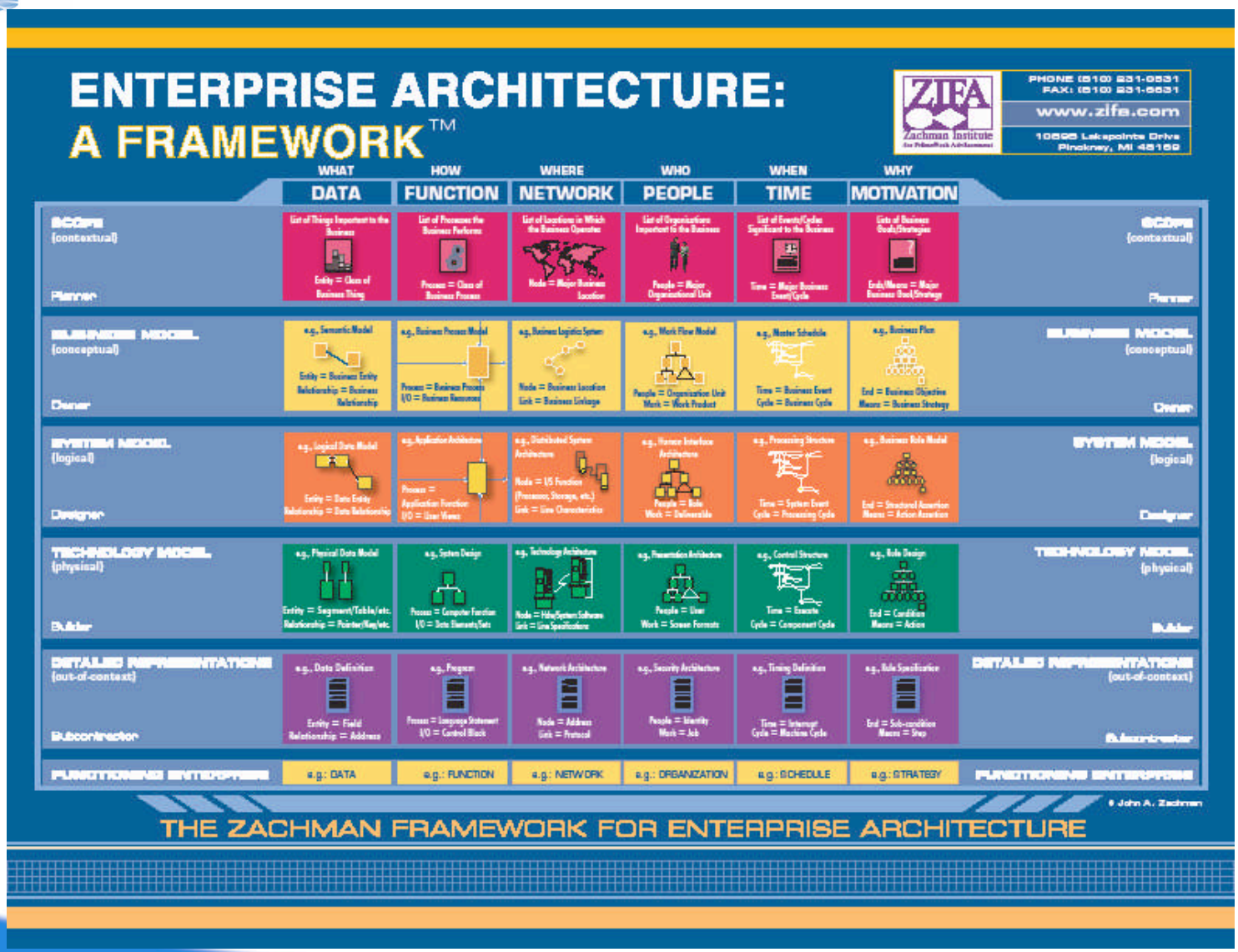
- Two dimensions of systems architecture – a 6 by 6 matrix
  - Different stakeholders required different levels of abstraction (rows)
    - The planner view – lists of system elements defining scope
    - The owner view – a semantic model showing the relationships between the elements
    - The designer view – a logical view detailing requirements and unconstrained design
    - The contractor view – a physical view optimizing the design for specific use and constraints
    - The implementer view – an out-of-context view of how components are assembled and operate
    - The actual implementation
  - Different perspectives answered different questions (columns):
    - What – the "data" column -- materials used to build the system
    - How – the "function" column -- processes performed
    - Where – the "network" column – topography and technology
    - Who – the "people" column – roles and organizations
    - When – the "time" column – events, cycles and schedules
    - Why – the "motivation" column – goals, strategies, rules
  - Each cell represents a unique type of model

The Zachman Framework

THE ZACHMAN FRAMEWORK FOR ENTERPRISE ARCHITECTURE TM

	DATA *What*	FUNCTION *How*	NETWORK *Where*	PEOPLE *Who*	TIME *When*	MOTIVATION *Why*	
SCOPE (CONTEXTUAL) *Planner*	Things Important To the Business Entity = Business Thing	Processes Performed Function = Business Process	Business Locations Node = Location Site	Important Organizations People = Major Organization	Events Significant To the Business Time = Major Business Event	Business Goals and Strategy Ends/Means = Business Goal	**SCOPE (CONTEXTUAL)** *Planner*
ENTERPRISE MODEL (CONCEPTUAL) *Owner*	Semantic Model Ent = Business Entity Rel = Business Relationship	Business Process Model Proc = Business Process I/O = Business Resources	Business Logistics System Node = Business Location Link = Business Linkage	**Work Flow Model** People = Organization Unit Work = Work Product	Master Schedule Time = Business Event Cycle = Business Cycle	Business Plan End = Business Objective Means = Business Strategy	**ENTERPRISE MODEL (CONCEPTUAL)** *Owner*
SYSTEM MODEL (LOGICAL) *Designer*	Logical Data Model Ent = Data Entity Rel = Data Relationship	Application Architecture Proc = Application Function I/O = User Views	Distributed System Architecture Node = IS Function Link = Line Characteristics	Human Interface Architecture People = Role Work = Deliverable	Processing Schedule Time = System Event Cycle = Processing Cycle	Business Rule Model End = Structural Assertion Means = Action Assertion	**SYSTEM MODEL (LOGICAL)** *Designer*
TECHNOLOGY MODEL (PHYSICAL) *Builder*	Physical Data Model Ent = Segment/Table Rel = Pointer/Key	System Design Proc = Computer Function I/O = Data Elements/Sets	Technology Architecture Node = Hardware/Software Link = Line Specifications	Presentation Architecture People = User Work = Screen Format	Control Structure Time = Execute Cycle = Component Cycle	Rule Design End = Condition Means = Action	**TECHNOLOGY MODEL (PHYSICAL)** *Builder*
DETAILED REPRESENTATIONS (OUT – OF- CONTEXT) *Sub - Contractor*	Data Definition Ent = Field Rel = Address	Program Proc = Language Statement I/O = Control Block	Network Architecture Node = Addresses Link = Protocols	Security Architecture People = Identity Work = Job	Timing Definition Time = Interrupt Cycle = Machine Cycle	Rule Design End = Sub-Condition Means = Step	**DETAILED REPRESENTATIONS (OUT - OF - CONTEXT)** *Sub - Contractor*
FUNCTIONING ENTERPRISE	**Data**	**Function**	**Network**	**Organization**	**Schedule**	**Strategy**	**FUNCTIONING ENTERPRISE**
	DATA ***What***	**FUNCTION** ***How***	**NETWORK** ***Where***	**PEOPLE** ***Who***	**TIME** ***When***	**MOTIVATION** ***Why***	

Another View



The Zachman Framework

- Is it correct? Debatable – no framework is perfect
- Is it useful? Yes!
  - Simple (only two dimensions) and easy to understand.
  - Non-technical language helps people think and communicate
  - Addresses the enterprise but also divisions
  - Helps frame issues and teach topics and solve problems.
  - Enables focus on details while seeing the “big picture” context.
  - Independent of specific tools or methodologies
- Some cells have proven more useful than others
  - Some people adopt the Zachman framework loosely, others apply it strictly

Enterprise Architecture in Practice

- An ***integrated collection*** of (business and IT) models and documents reflecting enterprise integration and standardization requirements and high-level design.
  - Usually defines both an “as is” and “target” state
  - May also include “reference” and “transition” states
  - All versions must be kept current to be relevant and useful

- A tool for planning, IT governance and portfolio management, that helps:
  - Align information systems with business strategy.
  - Align organization and operating model with business strategy.
  - Guide integration of data, processes, technologies and efforts.
  - Enable effective coordination of resources.
  - Improve communication and understanding across the organization.
  - Reduce the cost of managing the IT infrastructure.
  - Guide business process improvement.
  - Enable leadership to respond effectively to changing market opportunities, industry challenges and technological advances. Enterprise architecture helps evaluate business risk, manage change and improve business effectiveness, agility and accountability.

Enterprise Architecture in Practice

- **Data architecture**
  - Subject areas, business entities, business relationships, data attributes, business definitions, taxonomies, entity lifecycle states, valid reference values, data quality rules, data security classifications, data flow
- **Process architecture**
  - Functions, activities, tasks, steps, workflow, products, events, cycles, procedural rules
- **Business architecture**
  - Goals and objectives, strategies and initiatives, roles and job positions, organization structures, locations, operating principles
- **Application architecture**
  - Business system portfolio, software components (SOA), program structure and flow, portals and user interfaces, implementation projects
- **Technology architecture**
  - Hardware and software platforms, standards, protocols, network topology
- **Information value chain analysis**
  - Mapping the relationships between data, process, business, applications and technology

Enterprise Data Architecture

Data Architecture

-- Enterprise Data Model
- **Subject Areas**
- **Entity Hierarchies**
- **Conceptual Views**
- **Logical Views**
- **Business Glossary**
- **Taxonomies**
- **Entity Life Cycles / States**
- **Reference Data Values**
- **Data Quality Rules**

-- Related Data Architecture
- **Database Architecture**
- **Data Integration Architecture**
- **DW/BI Architecture**
- **Reporting Architecture**
- **Content Mgmt. Architecture**
- **Meta Data Architecture**

Process Architecture

-- Functional Decompositions
-- Process Flows / Workflows
-- Information Products
-- Events and Business Cycles
-- Procedural Rules

Business Architecture

-- Operating Principles
-- Goals and Strategies
-- Organization Structures
-- Roles and Job Positions
-- Locations

Information Value Chain Analysis

Matrix Relationships between Data, Processes, Organizations, Roles, Locations, Goals, Applications, Projects and Technology Platforms

Application Architecture

-- Business System Portfolio
-- Portals and User Interfaces
-- Program Structure and Flow
-- Software Component Architecture
-- Implementation Project Portfolio

Technology Architecture

-- Network Topology
-- Technology Platforms
-- Standards and Protocols
-- Software Tool Portfolio

Enterprise Data Architecture

- The master blueprints for *semantic* and *physical* integration of enterprise information assets.
  - Shared data requirements expressed in business terms
  - Guides implementation tailoring choices for “The Perfect Fit”
- Enterprise Data Model
  - Subject areas, business entities, relationships, super and sub-types
  - Business definitions, data stewardship assignments
  - Essential data attributes
  - Entity lifecycle states, valid reference values, data quality requirements
- Information Value Chain Analysis
  - Alignment with process, technology and strategy
- Data Delivery Architecture
  - Data Integration Architecture
    - Macro-level data flow: “The Corporate Information Factory”
    - Reference Data and MDM Hubs, ODS, Data Warehouses and Data Marts
    - SOA Data access services
  - Database technology architecture
  - Information content and delivery architecture – portals, taxonomies, …
  - Meta data architecture – integration, control, delivery, meta model

Enterprise Data Model

- An enterprise data model (EDM) is an *integrated subject-oriented* data model defining the *essential* data produced and consumed across an entire organization.
  - *Essential* means the data critical to the effective operation and decision-making of the organization. Few (if any) enterprise data models define all the data within an enterprise. Decisions must be made (and revisited) about the scope of enterprise data modeling efforts. “Essential” does not mean “common” or “shared.” Essential data requirements may or may not be common to multiple applications and projects. Some data defined in the enterprise data model may be shared by multiple systems, but other data may be critically important yet created and used within a single system. Over time, the enterprise data model should define all data of importance to the enterprise.
  - *Integrated* means that all of the entities, attributes and rules in the model are defined once, without redundancy. The concepts in the model fit together as the CEO sees the enterprise, not reflecting separate and limited functional or departmental views. There is only one version of the Customer entity, one Order entity, etc. Every data element also has a single name and definition. The data model may also identify common synonyms and important distinctions between different sub-types of the same common business entity.
  - *Subject-oriented* means the model is divided into commonly recognized subject areas that span across multiple business processes and application systems. Subject areas are focused around the most essential business entities.

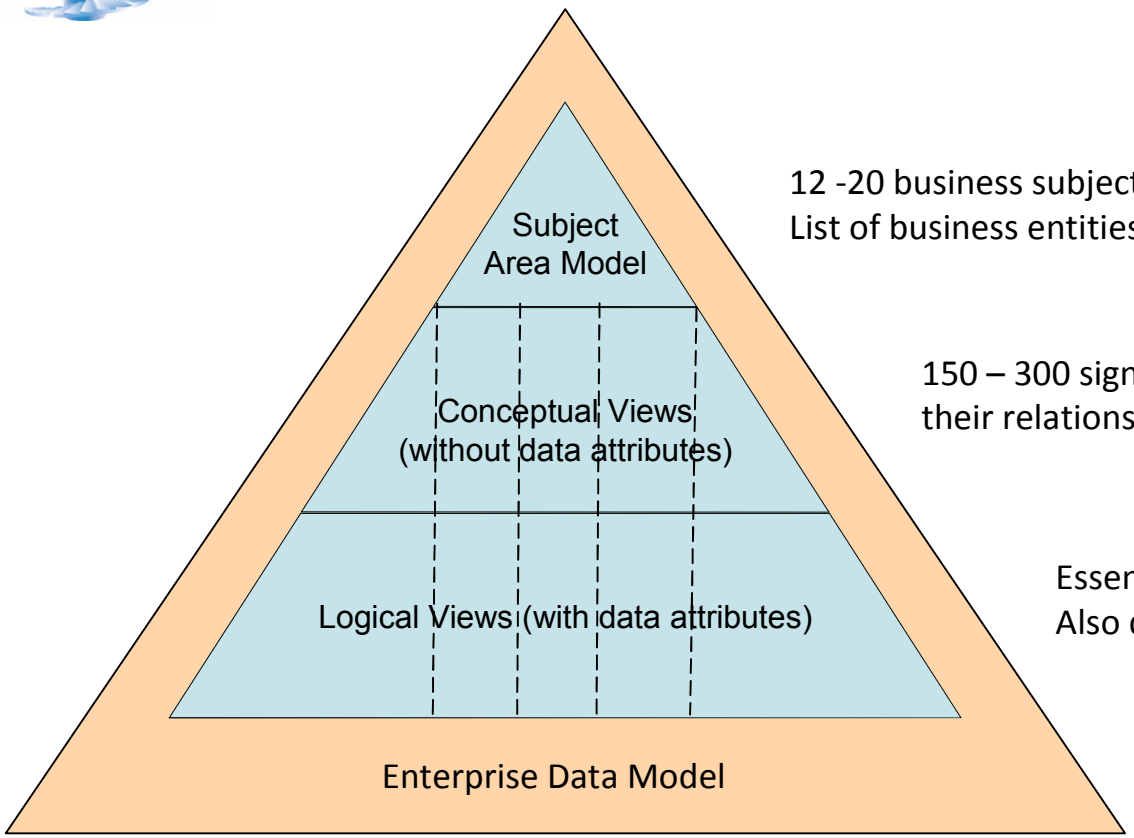
DAMA-DMBOK, 2008 – used with permission

Enterprise Data Model

- Not Just Data, But Also Words
  - Defines a common enterprise vocabulary
  - "Enterprise semantic model?" "Enterprise ontology?"

- Not Just Alignment and Integration
  - Improves data quality
    - Less redundancy, less variability, more reuse
  - Deeper business understanding
  - Wiser data interpretation
  - Enables data governance and stewardship

Enterprise Data Model Layers



12 -20 business subjects. One diagram.
List of business entities in each subject area.

150 – 300 significant business entities and their relationships. Developed iteratively.

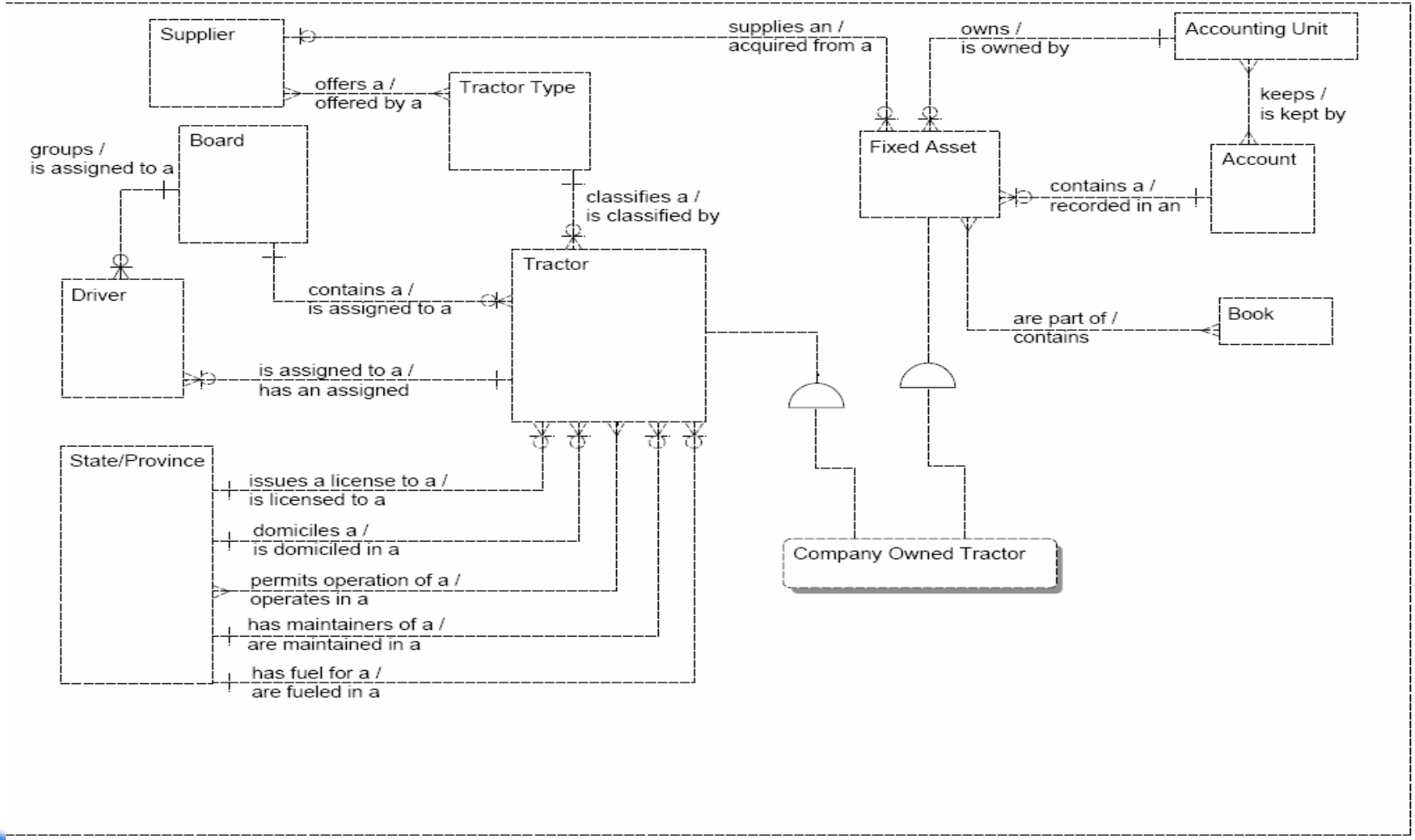
Essential business attributes added to entities.
Also developed iteratively over time.

		Application Logical Data Models					
		Application Physical Data Models					

The Subject Area Model

- Zachman Framework Column 1 Row 1 (Scope View) Model – a list!
- Organizes the Enterprise Data Model
  - A very significant enterprise taxonomy! Get it "right" from the start!
- An essential structure for data governance and stewardship
  - Entities frequently appear in multiple subject areas, but should be assigned one primary subject area for governance
  - Business data stewards are assigned accountability for entities or entire subject areas
  - Data stewardship teams organized by subject area for modeling, data quality requirements definition and reference data (code table) management
- How many subject areas are needed?
- Will you need multiple levels (macro and mini subject areas)?
- Depict all subject areas graphically on one diagram (group related subject areas)
- Each subject area should have a business definition
- Some subject areas include all the entities supporting a business function ("Finance"), while other include a super-type and its family of sub-types ("Party")
- Subject areas are named after their "core" entity / identify that entity

Sample Entity Relationship Diagram -- Conceptual View



What Is Essential?

- What are the “essential” subject areas?
- What are the “essential” business entities?
- What are the “essential” data attributes?
  - Familiar to stakeholders
  - Can’t operate or make decisions without them
  - Not application or implementation specific
  - Not derived

Data Modeling Styles

- Information Engineering (IE) syntax
  - "crow's feet"
- IDEF0
  - "eye-deaf-zero"
  - "dark or white circles, solid or dotted lines"
- Object Role Modeling (ORM)
  - Rich syntax for complex relationships/rules
  - Smaller subject areas to keep readable
  - Use it (with Visio) to zoom in on special problems with data and process
- UML
  - Conceptual entities and classes are synonymous
  - Normalized data entities and business objects are different
  - Physical database tables and program objects are different
  - Unifying object methodologies is one thing, but…
  - Not well accepted among most data modelers
  - Popular for industry reference models

Data Modeling Standards

- Diagrams, reports and other artifacts
- Diagramming syntax
- Naming conventions
  - Word order and form
  - Key words
  - No abbreviations for business/logical names
  - Abbreviated physical names
  - Standard abbreviations
- Attribute domains (logical data types)
- Definitions, examples and other extended properties
- Completeness
- Consistency
- Participation, review and approval

Defining the Conceptual Views

- Identify and name business entities
- Identify super-type / sub-types
  - "X is a kind of Y"
  - Isolate most class hierarchies in separate subject areas
- Draft, review and refine entity business definitions
- Identify and specify examples, synonyms, acronyms
  - Document distinctions from closely related terms
- Assign primary subject area
- Assign data stewardship accountability
- Define potential business identifiers
  - Define initial draft primary key attribute
- Identify and specify business relationships
  - Relationship names, cardinality

Defining the Logical Views

- Still usage (application) neutral – NOT the application logical model
- Identify essential data attributes
  - Unique identifiers, others
  - Brainstorm by data type (names, codes, dates, measures, …)
  - Analyze existing data models, databases, screens and reports
  - Do NOT accept the status quo
  - Exclude anything application or implementation specific
- Assign to logical data types (domains)
  - Inherit standard length for the domain
- Name according to standards
- Draft, review and refine attribute definitions
- Determine null-ability – should some value be mandatory?
- Identify best (if any) default value

Enterprise Logical View Issues

- Identify primary and foreign keys
- Natural or surrogate keys?
- Third normal form?
  - Retain many-to-many relationships?
  - Only “business” entities / no “data” entities?
  - Depends – ok if there are no essential attributes dependent on both parent keys

"Above and Beyond" Data Modeling - 1

- Entity lifecycle state-transition diagrams
  - (Status values and trigger events)
- Valid reference values
  - (codes, names and meanings)
- Business glossary
  - (including general terms, more than entities)
- Enterprise content management taxonomies
  - Standard topic hierarchies

“Above and Beyond” Data Modeling - 2

- Data attribute quality rules and requirements
  - Integrity rules
  - Format requirements
  - Data cleansing rules and procedures
  - Match / merge rules
  - Accuracy / precision requirements
  - Timeliness / “freshness” requirements
  - Consistency requirements
  - Security / privacy protection requirements
  - Security classification (“restricted”, ”confidential”, ”internal only”, ”public”)
  - Retention and archival requirements
  - Regulatory compliance requirements
  - Audit requirements

Identify the most critical dimensions of data quality

Not Just Primitive Models

- Zachman Framework identifies primitive model artifacts
  - Data-to-data relationships only
- Other relationships are also critical to the enterprise – "composite" models
  - Data-to-process
  - Data-to-organization
  - Data-to-role
  - Data-to-application

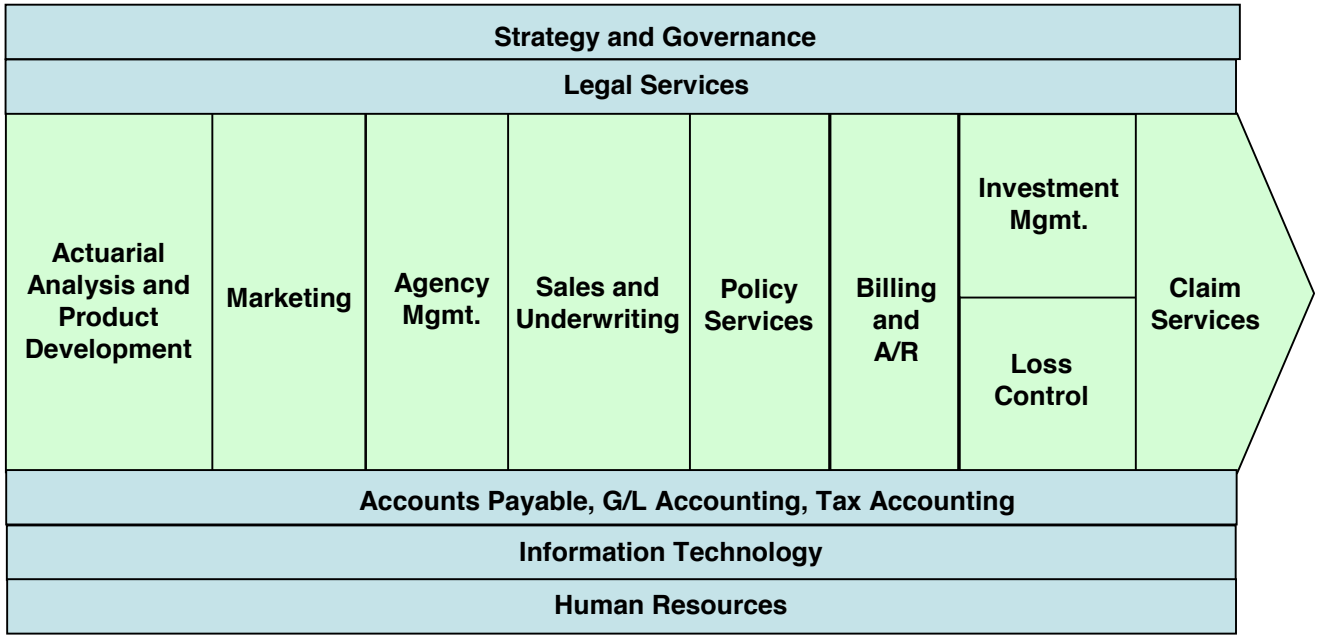
Information Value Chain Analysis

- Matrix mapping of relationships between two sets of elements
  - Most data relationship matrices are “CRUD” matrices
  - “Create, read, update, delete”
- Developed in early 1980’s for IBM business systems planning (BSP)
  - Incorporated into james martin’s information engineering methodology -- information systems planning (ISP) phase
- Still a critical component of enterprise architecture!!!!

Data Classes															
Process	Policy / Regulation / Law	Mission and Function	Agency Agreements	Planning	Work in Progress	Products	Financial Status	Procurement and Contracts	Man Power	Training	Personnel	Security and Safety	Reference Information	Fixed Assets and Expenses	ADP Information
Implement Policy Guidance from OCE	U	C	C	C	U	U	U	U	U	U	U				
Provide Strategic Direction	U	C	C	U	U	U									
Formulative Army RDT&E Programs	C	U	U	C	U	U	U	U	U	U	U	U	U	U	
Manage Research and Development Programs	U	U	U	U	C	C	U	U	U	U	U	U	U	U	U
Manage Reimbursable and Small Problem Program	U	U	U	U	C	C	U	U	U	U	U	U	U	U	U
Develop Improved Planning, Design & Construction Pr	U	C	C	U											
Develop Improved O&M Procedures	U	C	C	U											
Manage Financial Resources		U	U	U	U	C	U	U	U						
Manage Procurement & Contracts		U	U	U	C	U	U	U	U						
Manage Manpower	U	U	U	U	U	U	U	C	C	C	U	U	U		
Manage Organizational Effectiveness Programs	U	U	U	U	U	C	C	U	U						
Manage Personnel Training	U	U	U	U	U	U	C	U	U	U	U				
Provide Staff Review & Approval		U	U	U	U	U	U	U	U	U	U	C	U	U	U
Investigate and Solve Personnel Problems	U	U	U	U	U	C	U								
Manage Safety and Security Activities	U	U	U	C	U										
Manage Public Affairs Program	U	U	U	U	U	U	U	U	C	U	U				
Manage Laboratory Support Functions	U	U	U	U	U	U	U	U	U	C	C	U			
Manage Automation and Information	U	U	U	U	U	U	U	U	U	C					

Information Value Chain Analysis

- Technique re-named after Michael Porter's Business Value Chain concept
  - Directly contributing functions sequenced left to right
  - Indirect functions support from below or above



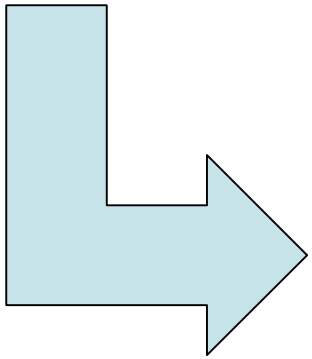
Information Value Chain Analysis

- Element Sequenced Using the Business Value Chain
  - Familiar, intuitive sequence
  - Functions AND subject areas?
  - X axis (left to right) AND Y axis (top-to-bottom)?
- Data / Process CRUD Matrices: Different Levels of Detail
  - Subject areas / business functions
  - Business entities / functions or processes
  - Data attributes / processes and their information products
- Other Potentially Useful CRUD Matrices
  - Data / organization CRUD matrix – who?
  - Data / role crud matrix –who?
  - Data / location crud matrix – where?
  - Data / application system crud matrix – where?

Affinity Analysis

Order Processing System CRUD Diagram 1

	Customer	Customer Order	Customer Account	Customer Invoice	Vendor Invoice	Product
Receive Customer Order	R	C	CR			
Process Customer Order	CRU		RU			R
Maintain Customer Order	U		U		RU	
Terminate Customer Order	U		U		RU	
Fill Customer Order	RU		RU			RU
Ship Customer Order			U		C	
Validate Vendor Invoice					R	
Pay Vendor Invoice					RU	
Invoice Customer	RU		RU	C		
Maintain Inventory						CRUD



Affinity analysis sorts one or both axes to group related processes and subject areas/entities together

Order Processing System CRUD Diagram 2

	Customer	Customer Order	Customer Account	Customer Invoice	Vendor Invoice	Product
Maintain Customer Order	U		U		RU	
Terminate Customer Order	U		U		RU	
Receive Customer Order	R	C	CR			
Process Customer Order	CRU		RU			R
Fill Customer Order	RU		RU			RU
Invoice Customer	RU		RU	C		
Ship Customer Order			U		C	
Validate Vendor Invoice					R	
Pay Vendor Invoice					RU	
Maintain Inventory						CRUD

Why Value Chain Analysis?

- Alignment with process models
- Validating the data model
- Understanding data sources
- Analyzing data quality issues
- Change impact analysis
- Defining information products

"Above and Beyond" Data/Process Analysis

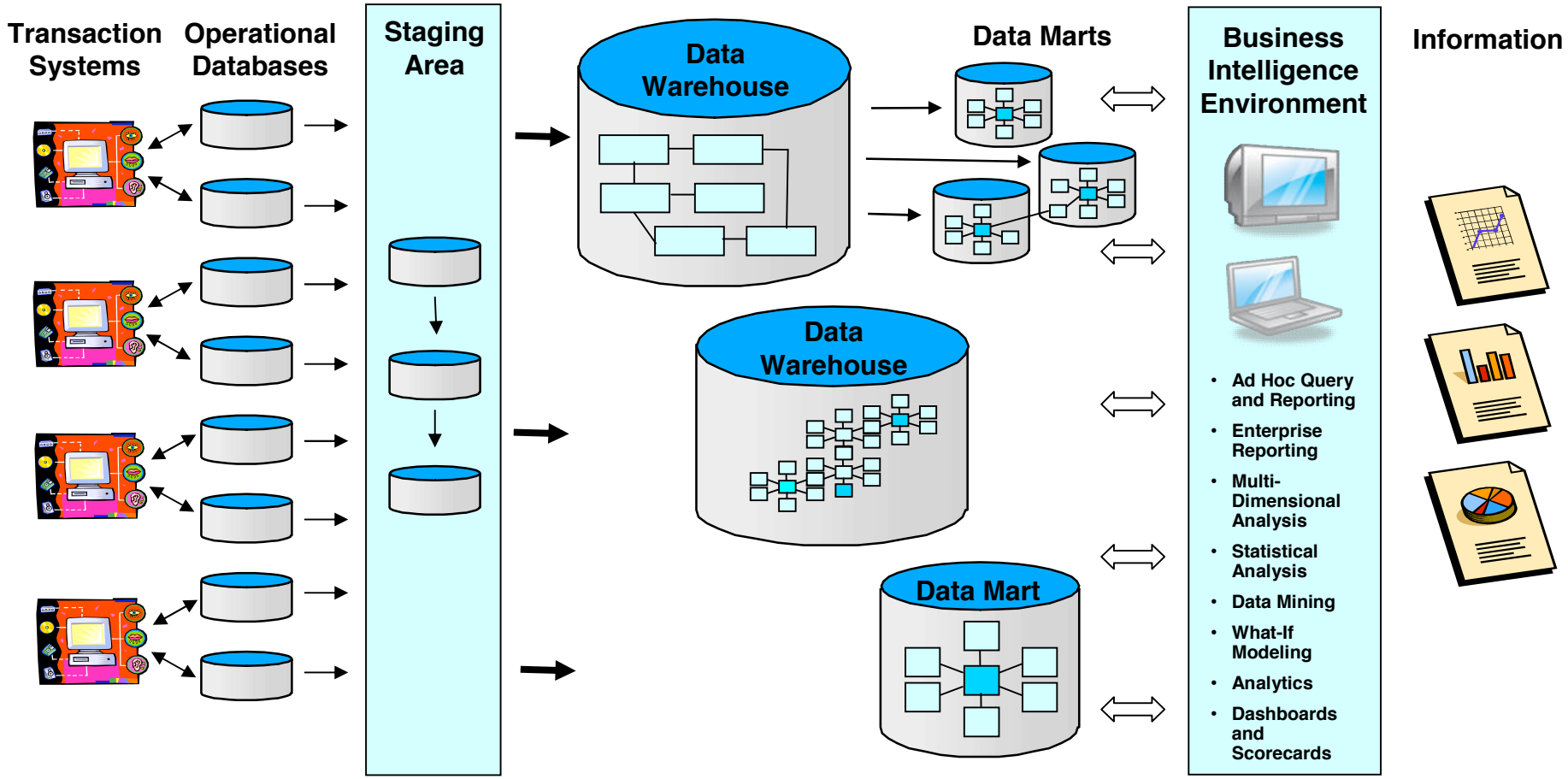
- Data/Process Flow Modeling
  - "Information Supply Chain Analysis"
  - Defining information products (process outputs) and their sources
  - Collaborative business process modeling – data analysts, business process analysts and subject matter experts – developing process flow diagrams
  - Working ***backwards*** from need to source
  - Identifying the data attributes needed in information products
  - Identifying essential data attributes from information products

- Strategic Business Intelligence Requirements Analysis
  - Starting with the information needs of executives
  - Identifying key performance indicators and other critical measures
  - Identifying most frequently needed dimensions

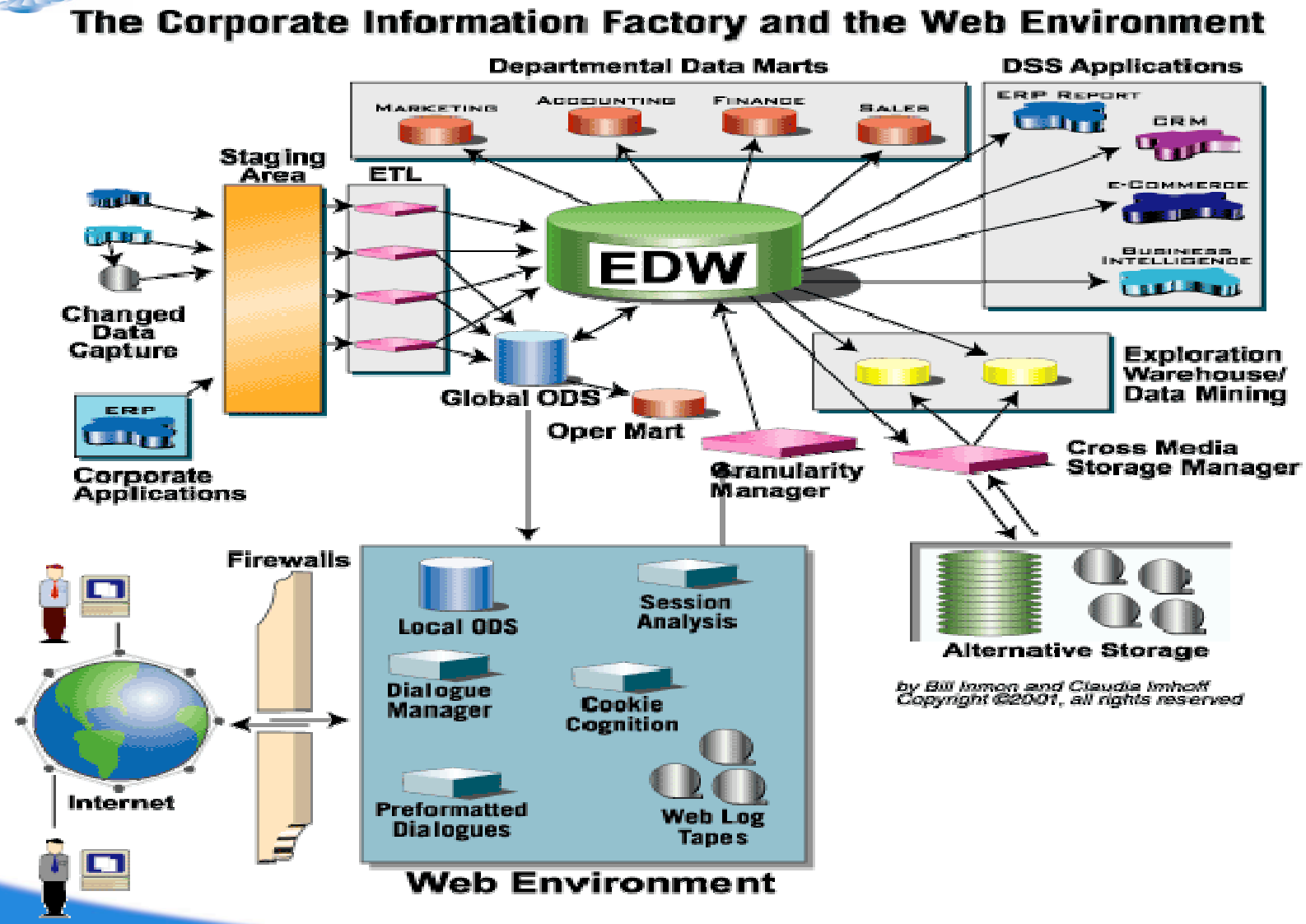
Data Delivery Architecture

- Data Integration Architecture
  - How data flows across databases and applications (OLTP, MDM, DW, BI)
  - "The Corporate Information Factory"
  - Master data management hubs
  - Operational data stores
  - Data warehouses and data marts
  - Data replication and transformation
  - Subscribe and publish
  - Batch vs. near-real time (asynch MQ bus, …)
  - XML, Web Services and SOA
  - "Replacing feeds with reads"

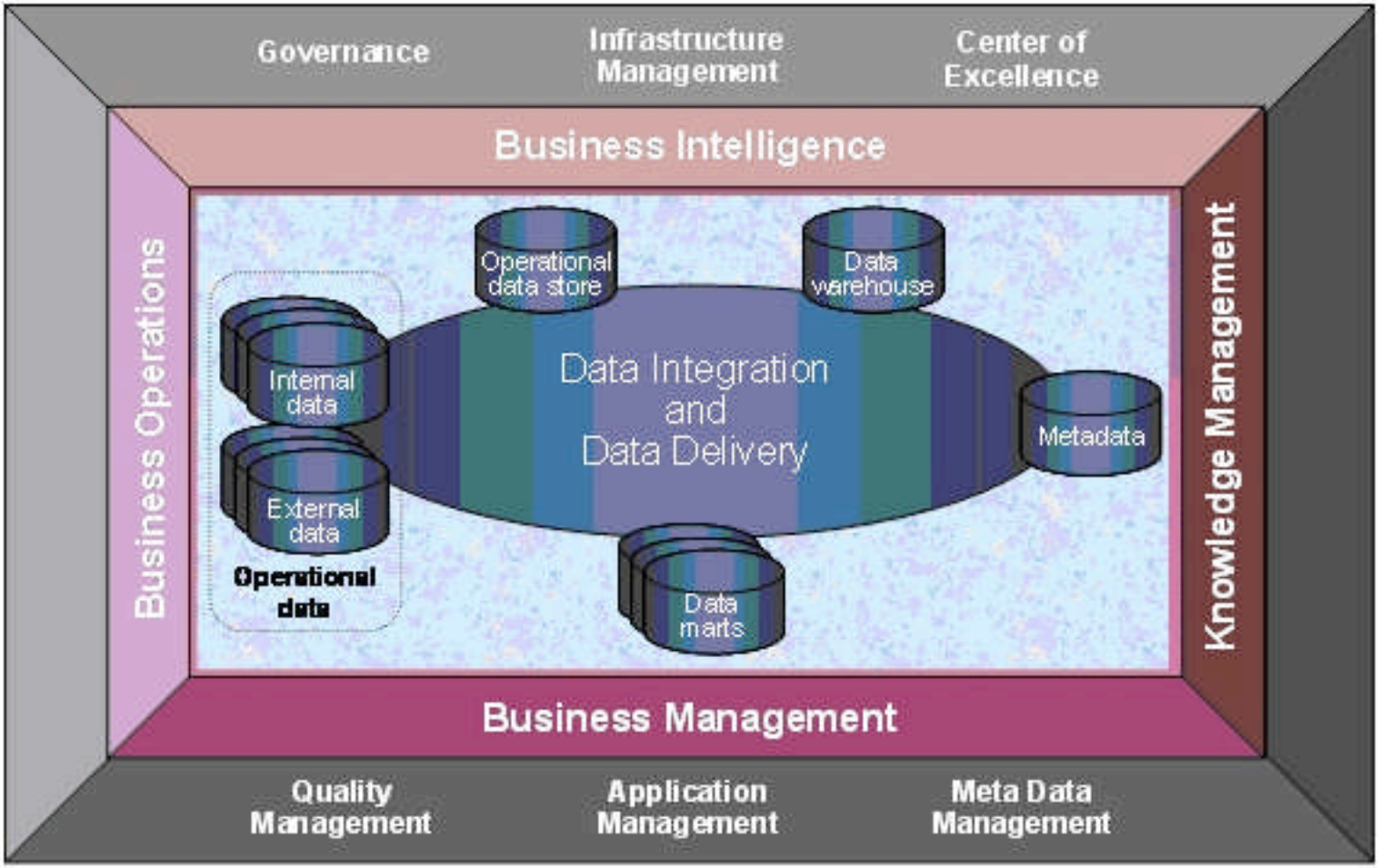
DW / BI Architecture



The Corporate Information Factory



The Extended Corporate Information Factory



Courtesy of Intelligent Solutions, Inc. and BI Research

Database Architecture

- Which DBMS tools?
- Which integration tools?
- When/where to use what technology?
- When to distribute? When to federate?

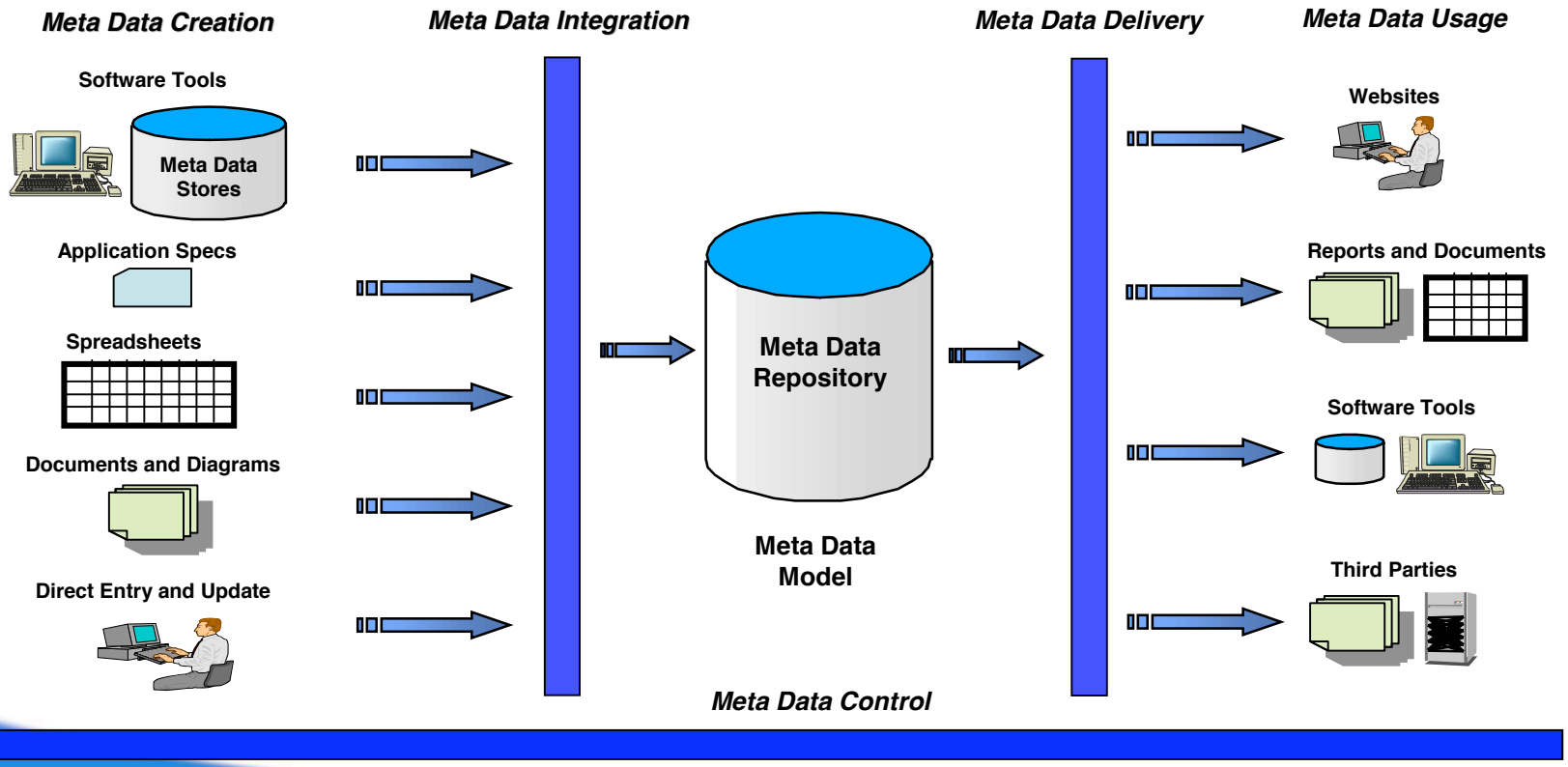
Unstructured Information Architecture

- Enterprise content management
- Enterprise taxonomies
- Enterprise portal strategy
- Document mgmt and imaging systems
- Storage management / archival and retrieval
- Report format, storage and distribution

What's essential for data alignment and data integration?

Meta Data Architecture

- **The Managed Meta Data Environment (MME) Architecture**
  - Dedicated hardware, software, staff and processes
  - For ***integration, control*** *and* ***delivery*** of meta data
  - Providing easier access to integrated meta data
  - Centralized, hierarchical, distributed or federated?



What's Worthwhile Doing?

- "Someday you will want to have everything defined to excruciating level of detail" – John Zachman
- What are the significant business risks today?
- What models and artifacts manage those risks?
- "Whatever is worthwhile doing is worth doing well" – Your Mom
- And then, "What's next?" – President Bartlett, The West Wing

Getting Started

- Set the context – define the ultimate objectives
- Secure commitment, initial funding and participation
- Define the big picture first – subject area layer
  - Subject area names and definitions
  - 1st draft list of business entities within each SA
  - 1st draft outline of entity super/sub-type hierarchy
  - 1st draft definitions of entity definitions
  - High level matrices: SA/function, SA/org, SA/application
  - With a small, respected group of business data stewards
  - Do most work in one single 3 day offsite facilitated workshop
  - Prepare first – study and summarize existing data models
  - Review, revise and approve by data governance council
- Prioritize the subject areas
- Build an iterative development plan

For Each Subject Area Iteration:

1. Conceptual View
    1. Entity names and definitions
    2. Super-types and sub-types (separate subject areas)
    3. Business relationships
    4. Examples, synonyms, acronyms
    5. Entity lifecycle (state-transition diagrams)
    6. Data steward assignments
2. Info Value Chain Analysis
    1. Entity/function
    2. Entity/org
    3. Entity/application
1. Logical View – Data Attributes
    1. Attribute names and domains
    2. Definitions, acronyms, synonyms, examples
    3. Unique identifiers
    4. Entity normalization (some – for foreign keys)
2. Reference Data Values
    - Codes, names, meanings
3. Data Quality Requirements
4. Process Flow Modeling / Information Product Analysis
    - "Info Supply Chain Analysis"

Data Architecture Management

Building, maintaining, enriching the models requires:

- Iterative and incremental development
  - A long-term plan
  - Short-term and sustained funding
- Data governance
  - Sponsorship, funding and scope
  - Oversight
  - Review and approval
- A shared partnership
  - Actively involved and accountable business data stewards
  - Other participating subject matter experts
  - Data architects and other data analysts
  - Alignment with other enterprise architects
- Collaboration models
  - Facilitated brainstorming and synthesis
  - Strawman draft, review and revision
- Data model quality control
  - Standards and reviews
  - Model configuration management

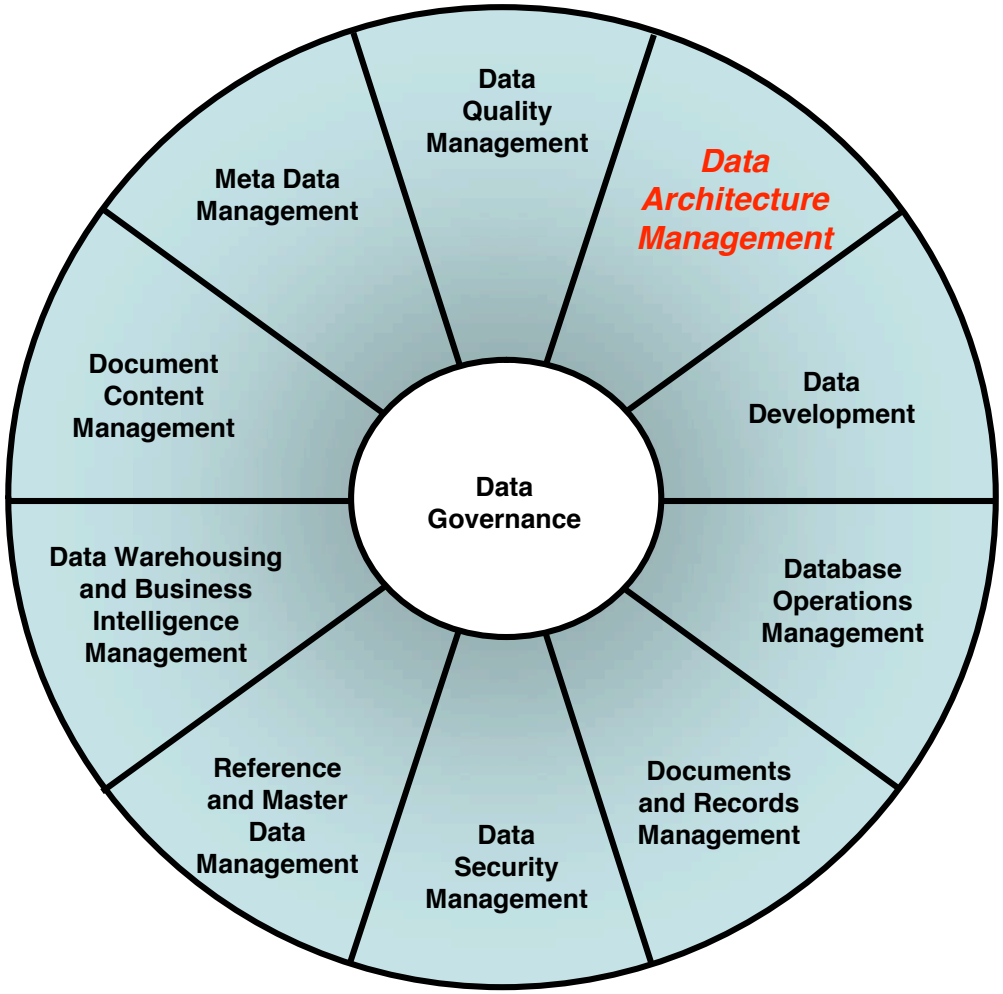
Data Architecture Stewardship

Business data stewards have many responsibilities in developing and maintaining the information architecture

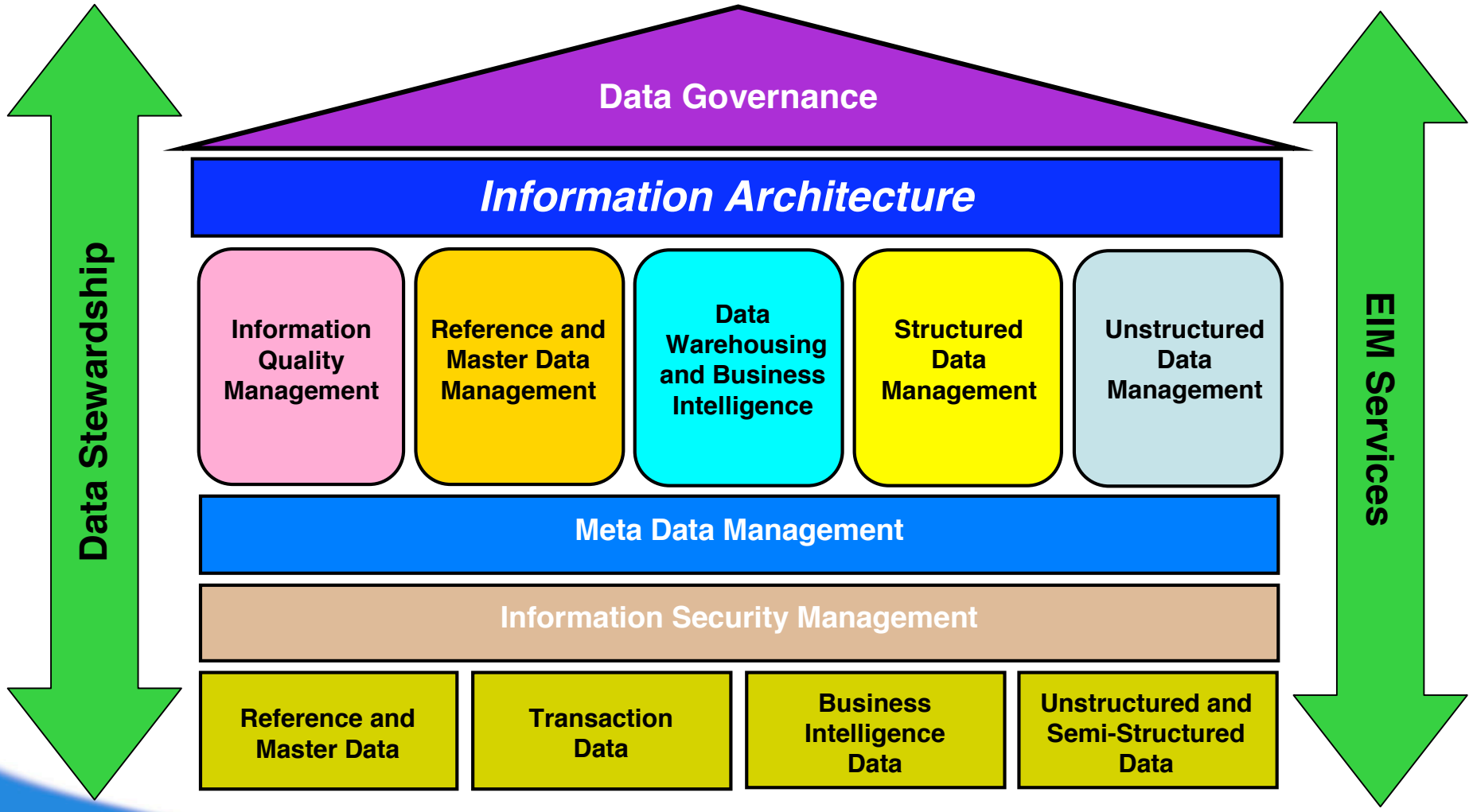
- Propose, draft, refine and approve the names and definitions for subject areas, business entities and data attributes.
- Analyze, identify and approve the entity class hierarchy relationships (super-types and sub-types)
- Analyze, identify and approve entity relationship names and business rules
- Define data quality requirements and expectations
- Review / approve data architecture components

The DAMA-DMBOK Framework

Version 3



EWSolutions EIM Framework



Conclusion

- Data architecture is part of an enterprise architecture
- EA can drive data architecture or reverse – both are ultimately essential to a fully functional enterprise
- Data architecture requires skills in several areas, and is a discipline for experienced data professionals, includes technical knowledge

Questions

EWSolutions, Inc.
15 Spinning Wheel Road,
Suite 330
Hinsdale, IL 60521
Office 630.920.0005
Fax 630.920.0008

Anne Marie Smith, Ph.D.
Director of Education, Principal Consultant
Office 856-468-6194
AMSmith@EWSolutions.com

http://www.EWSolutions.com