Whitemarsh
Information Systems Corporation

Data Interoperability Strategy Seminar (Management)

Whitemarsh Information Systems Corporation
2008 Althea Lane
Bowie, Maryland 20716
Tele: 301-249-1142
Email: Whitemarsh@wiscorp.com
Web: www.wiscorp.com

Table of Contents

1.0 Whitemarsh

- Founded in 1981
- Website: www.wiscorp.com
- Key Concentration Area: Data Management
- Delivery Mechanisms
  - ♦ Knowledge Ware (books and papers)
  - ♦ Software
  - ♦ Consulting, Seminars and Workshops

Whitemarsh Clients:

- Bank of America, California
- Delaware Public Safety Database
- DOD Strategic Defense Initiative Office
- DuPont Corporation
- Federal Government of Canada
- Freddie Mac
- Grumman Data Corporation and BDM International
- Hershey Chocolate Company
- Mars Corporation
- International Committee for Information Technology Standards
- Ohio State Supreme Court
- Prince George's County Board of Education
- Social Security Administration
- State of California
- The MITRE Corporation
- U.S. Army
- U.S. Navy
- U.S. Department of Commerce
- U.S. Department of Veterans Affairs
- U.S. Office of Trademarks
- U.S. Office of Personnel Management
- U.S. DoD, Defense Logistics Agency
- U.S. DoD Office of Assistant Inspector General for Audits

Note: Descriptions of Whitemarsh client engagements are at www.wiscorp.com

Goals of the Data Interoperability Strategy Seminar: **Confirm that:**

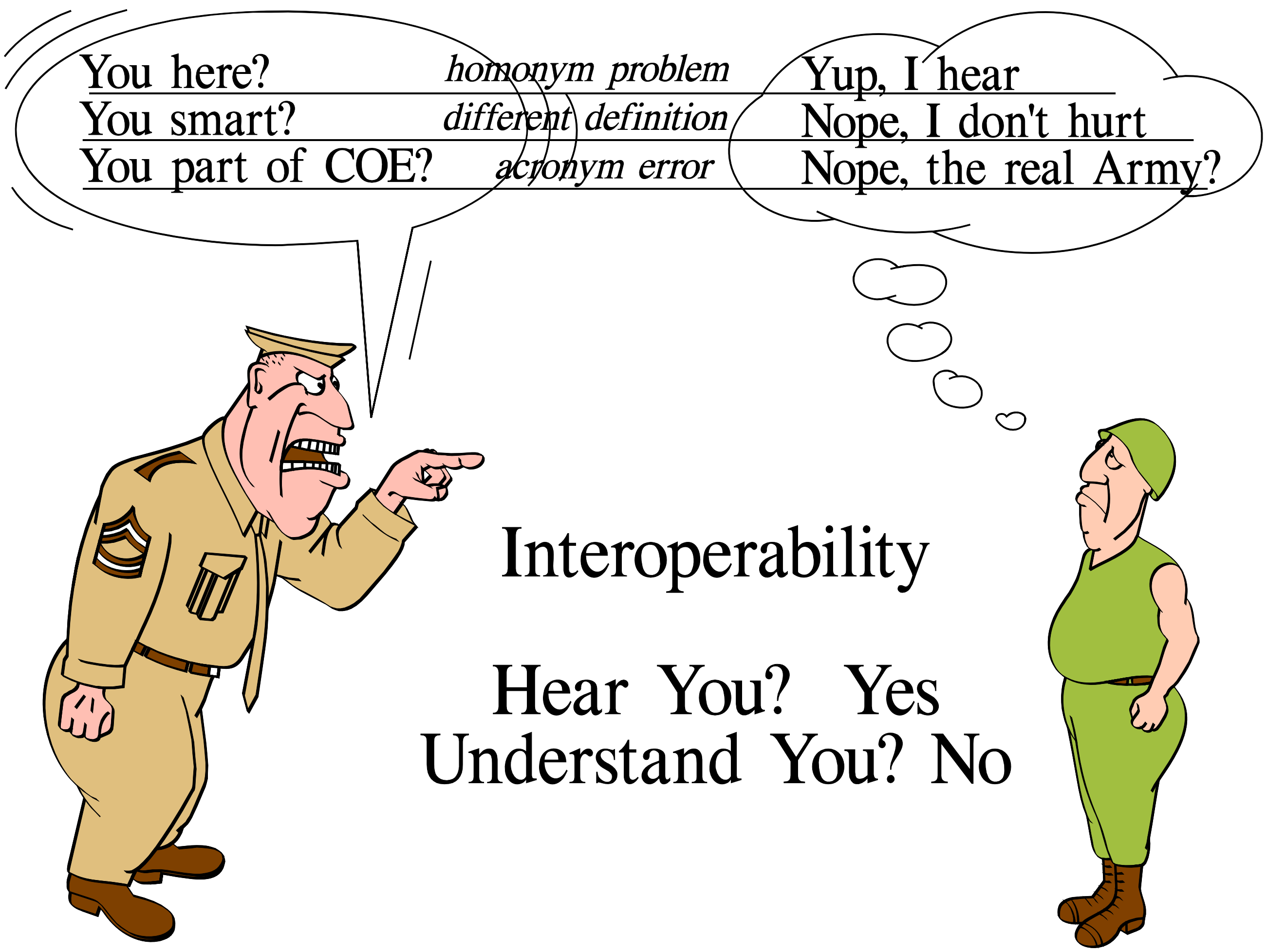
- There can be a viable strategy for achieving understanding-based data interoperability.
- Such a strategy is based on fundamental data management principles that you already know and/or use everyday.
- There is no Silver Bullet that will save us.
- All this is just highly-proceduralized common sense.

Remember, this 2-hour Strategy Seminar not a how-to workshop. Whitemarsh has an 8-hour Strategy seminar, and also a how-to workshop that takes a whole week.

Before we really begin, What is Data?

- Data is "what" remains after ***Policy*** is executed. ***Thus, Data is Executed Policy***.
- Processes are the ***Procedures*** through which the policy execution is accomplished
- "Data" Definitions are thus Policy Definitions
- Enterprise Architecture, Data Administration, Data Standardization, and Database Design is a "Enterprise Policy Specification" activity
- All data (i.e., policy) specifications are metadata.
- All process (i.e., procedure) specifications are metadata
- A Metadata Database (i.e., Metabase) is a database for all Policy and Procedure Specifications.
- This seminar is all about a strategy for Data Interoperability. One based on metadata, which are the specifications of an Enterprise's Persistent Policies and Procedures.

2.0 What is Data Interoperability



Interoperability Requirements

Connectivity Based		Understanding Based
Parse-able Common Formats Persistence Reliability Robustness Timeliness	**Versus**	Composition and Derivation Agreement Data Name and Name-Part Mapping Data Type Mapping Definition Agreement Granularity Agreement Precision Agreement Scale Agreement Value Domain Mapping
Basis–Exchange: fixed strings, "delimited" or Tag-based (e.g., XML)		**Basis–Definition**: Top-down standardization, Community of Interest consensus, External Standards Adoption
Connectivity Interoperability	**is not the same as**	Understanding Interoperability
I hear you		I understand you

If you don't have a Metadata Infrastructure you cannot have data interoperability. Why? Because there will be no "I understand you infrastructure."

The Key Take Away is:

- Connectivity Interoperability is **not** the same as Understanding Interoperability else interoperability would have been solved with the "Telegraph."
- You can have Connectivity Interoperability without Understanding Interoperability.
- You must have Connectivity Interoperability along with Understanding Interoperability to then have Total Interoperability.
- XML **alone** is not a solution.
- XML without a Metadata management infrastructure is just "I hear you."

Data Interoperability, then... The ability to exchange data between two or more agents without loss of precision or semantics, within acceptable limits for latency (timeliness), complexity, and cost.

3.0 Data Interoperability Errors

Type I Error: An error which occurs when a true hypothesis is rejected.

Hypothesis: Social Security Number Of Employee = Parent Social Security Number of Dependent

Rejection: Employee SSN is 9 integers (e.g, 123456789); Parent Social Security Number of Dependent (11 characters) is 3 integers, a hyphen, 2 integers, a hyphen, 4 integers. (e.g., 123-45-6789)

Other Examples:

1. **0 and 1 for Gender**, and **1 and 2 for Gender**. (Value domain mismatch)

2. Mike Gorman vs Michael M. Gorman (Same person, different names errors)

3. Generalized table for valuables (Valuables: ValuableType, Name, Value) vs RealEstate Table, or Jewelry Table, etc. (Data structure mismatches)

Type II Error: An error which occurs when a false hypothesis is accepted.

Hypothesis: Sales East = Sales for March of NE Division + Sales of March of SE Division

Acceptance: Sales March of NE Division is Net After Expenses Sales
Sales of March of SE Division is Total Monthly Sales

Other Examples:

- 1 for Gender = Female, and 1 for Gender = Male. (Value meaning mismatch)
- Michael Gorman vs Michael Gorman (Same name, different persons error. 1 born 1941, other born 1965 (Sears!))
- Generalized table for valuables (Valuables: ValuableType, Name, Value) vs RealEstate Table, or Jewelry Table, etc. (Data structure mismatches). In the first structure, all valuables were Totaled. In the second, only those that fit tables were known and Totaled.

Complexity and Latency

- Time and effort to sort through similar data that is different and vice versa.
- Time and effort to synchronize granularity, units, precision and reference data codes and meaning.
- Time to find and gather all “right” data for mission

4.0 Data Interoperability Infrastructure

- Frameworks and Metadata Products

- Essential Information Technology Components

  - ♦ Enterprise Identifiers
  - ♦ Authoritative Data Sources
  - ♦ Information Exchange Standard Specifications
  - ♦ XML Data Transport

- Metadata Repository

  - ♦ Characteristics
  - ♦ Key Meta Models
  - ♦ Implementation Alternatives

4.1 A Framework for the Knowledge Worker

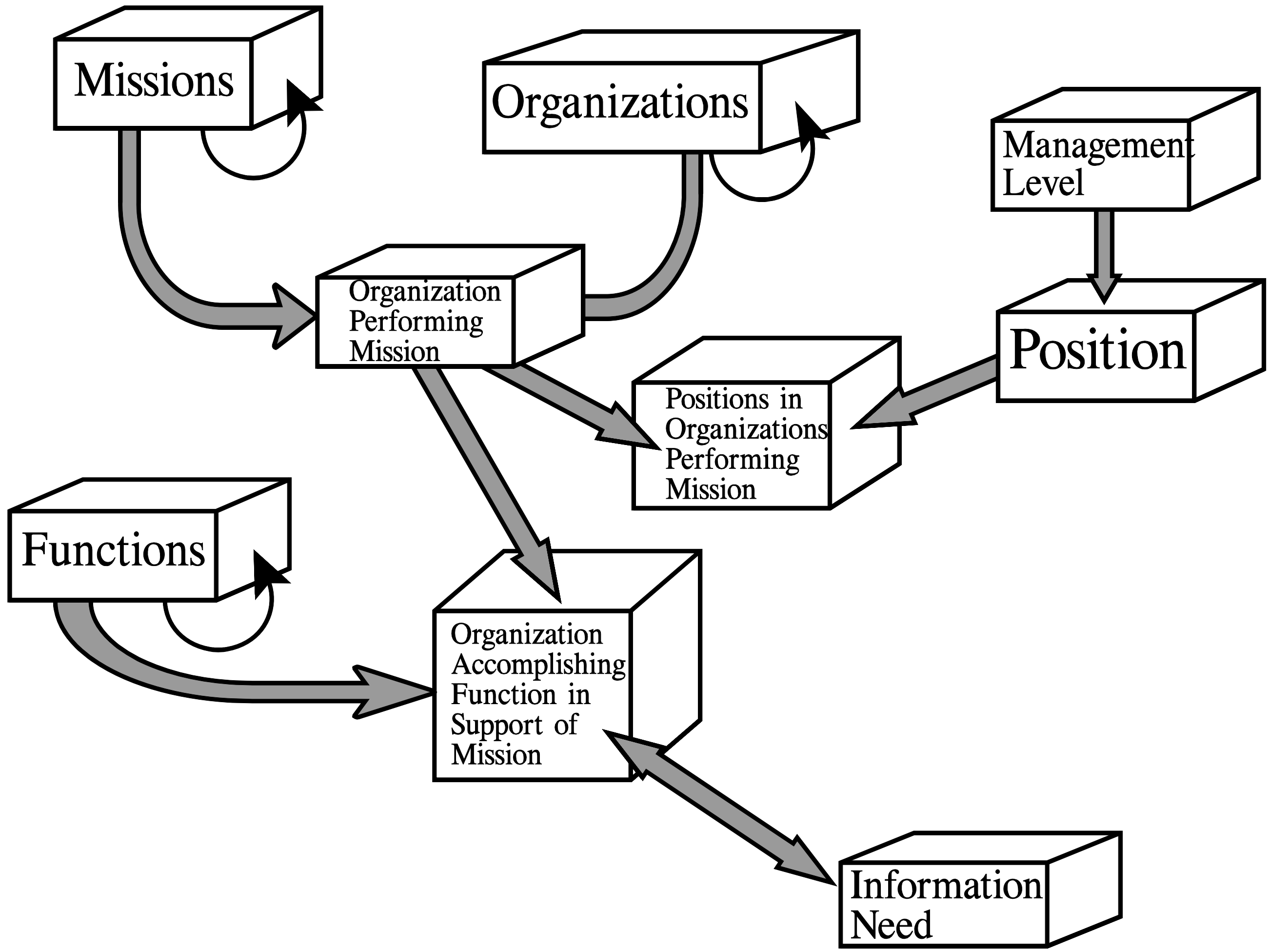
		Man-Machine Interface				
		Machine		Interface	Man	
Deliverables	**Mission**	**Database Object**	**Business Information System**	**Business Event**	**Business Function**	**Organization**
Scope	Business missions	Major business resources	Business information Systems	Interface events	Major business scenarios	Organizations
Business	Mission hierarchies	Resource Life Cycles	Information sequencing and hierarchies	Event sequencing and hierarchies	Business scenario sequencing and hierarchies	Organization charts, jobs and descriptions
System	Policy hierarchies	Specified data models and Identified Database objects	Information system designs	Invocation protocols, input and output data, and messages	Best practices, quality measures and accomplishment assessments	Job roles, responsibilities, and activity schedules
Technology	Policy execution enforcement	Implemented data models and Detailed Database Objects	Information systems application designs	Presentation layer information system instigators	Activity sequences to accomplish business scenarios	Procedure manuals, task lists, quality measures and assessments
Deployment	Installed business policy and procedures	Operational data models	Implemented information systems	Client & server windows and/or batch execution mechanisms	Office policies and procedures to accomplish activities	Daily schedules, shift and personnel assignments
Operations	Operating business	Application Interface data model	Operating information systems	Start, stop, and messages	Detailed procedure based instructions	Daily activity executions, and assessments

Knowledge Worker Framework Relationship to IT System Errors (GAO Studies)

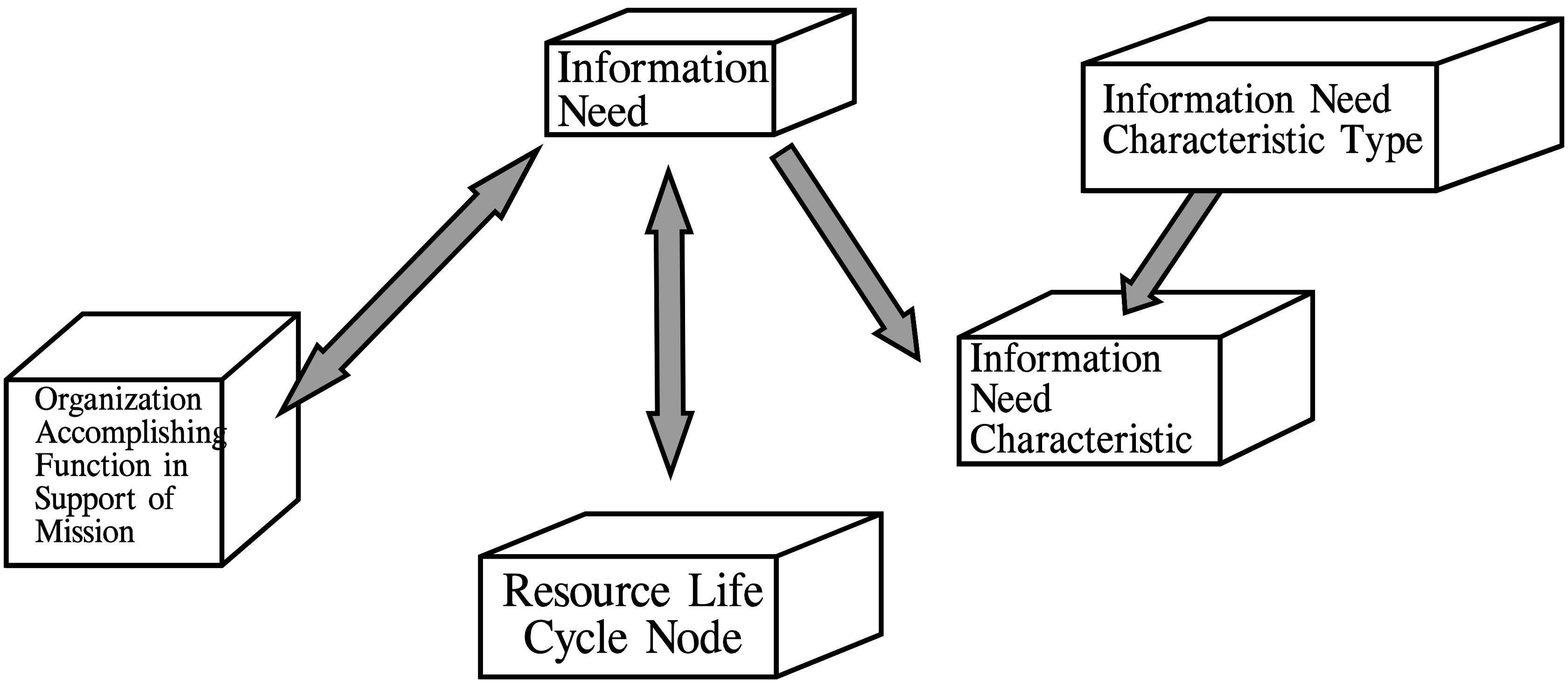
Knowledge Worker Framework						
		Machine		Interface	Man	
	Mission	Database Object	Business Information System	Business Event	Business Function	Organ-ization
Scope	5	2	3	1	3	4
Business	5	3	2	1	6	6
System	3	**2**	**2**	**1**	12	8
Technology	1	**0**	**0**	**0**	8	6
Deployment	0	**0**	**0**	**0**	5	5
Operations	0	**0**	**0**	**0**	3	3
Col. Totals	**14**	**7**	**7**	**3**	**37**	**32**

Note: All numbers expressed as Percent allocations of errors ...12 Gray cells are IT.
All errors from IT system failure are addressed.
This leads us to know what to collect.

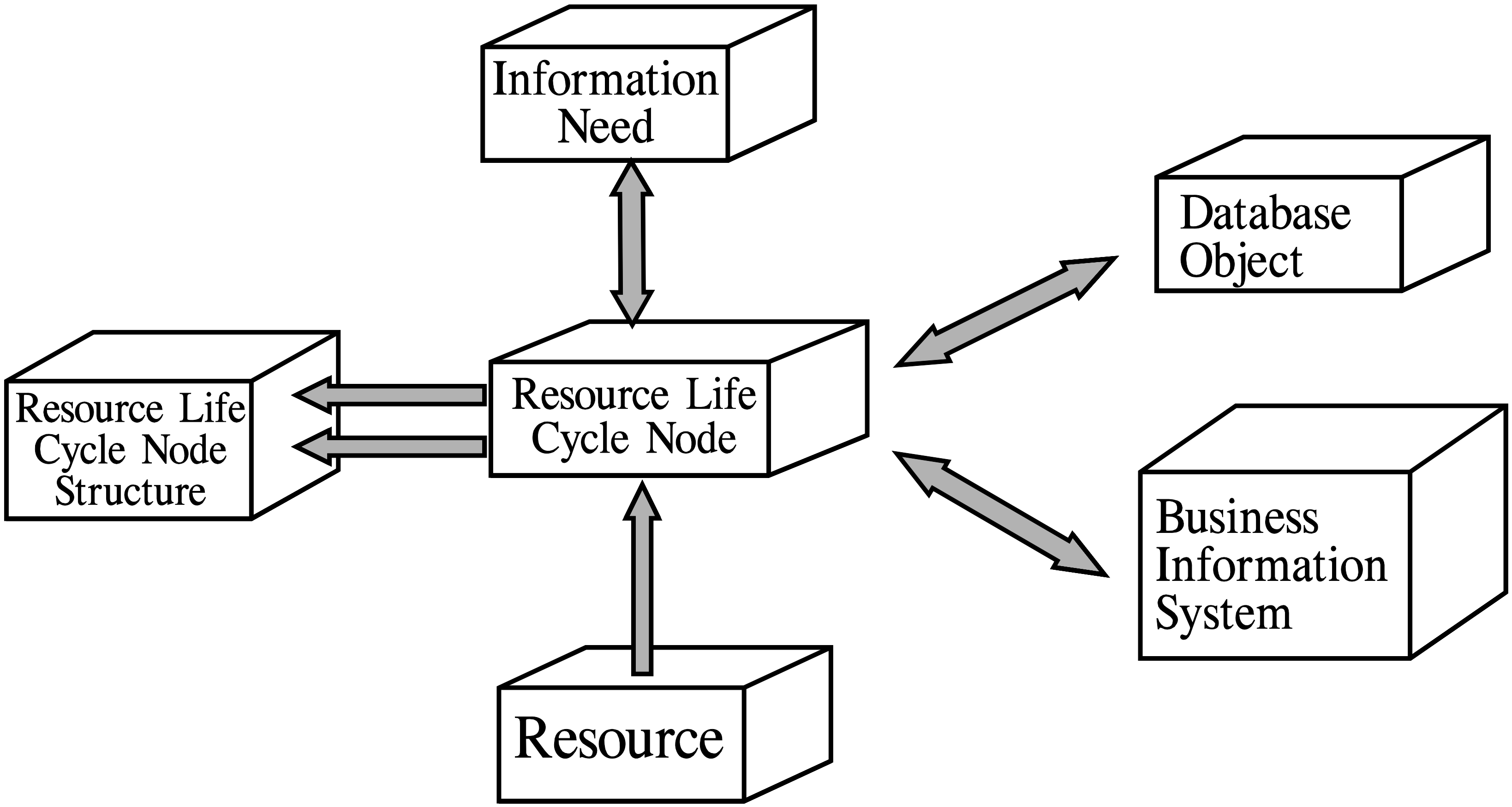
We must therefore collection: Mission, Organization, Function and Position



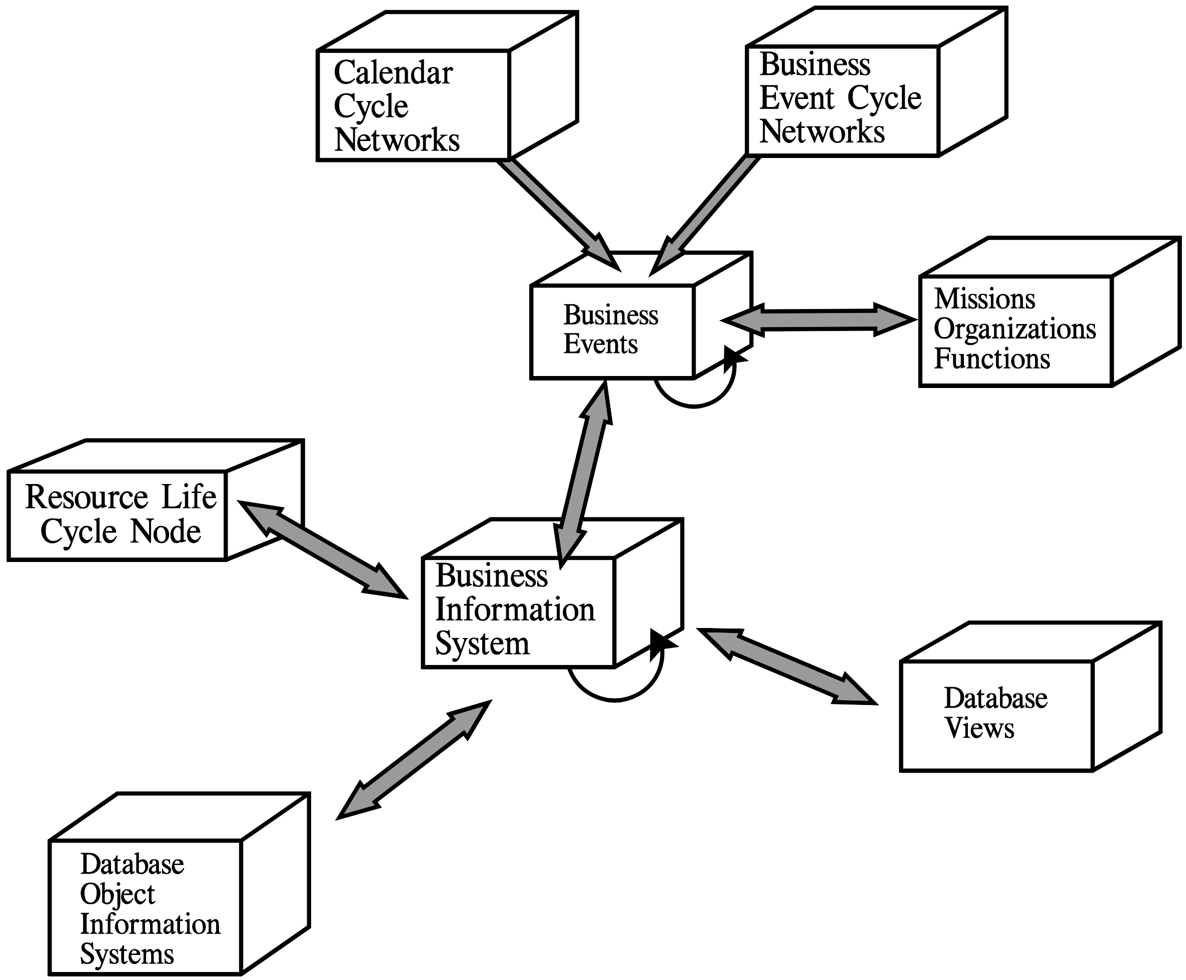
And Information Needs to then support our functions performed by organizations as they accomplish missions.



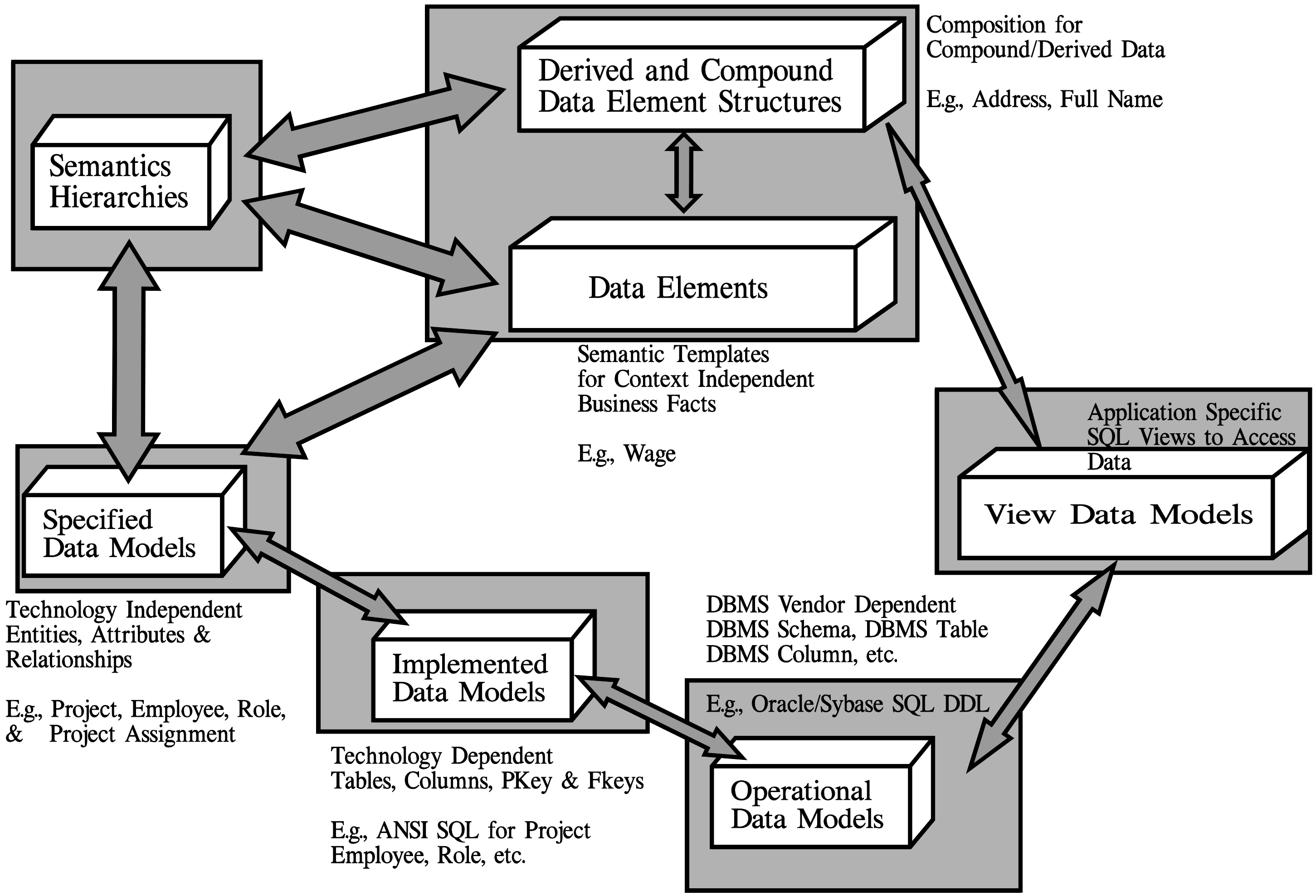
And know we'll know how our information needs are "produced." that is, what databases and information systems within which Resource Life Cycle Nodes satisfy which Information Needs...



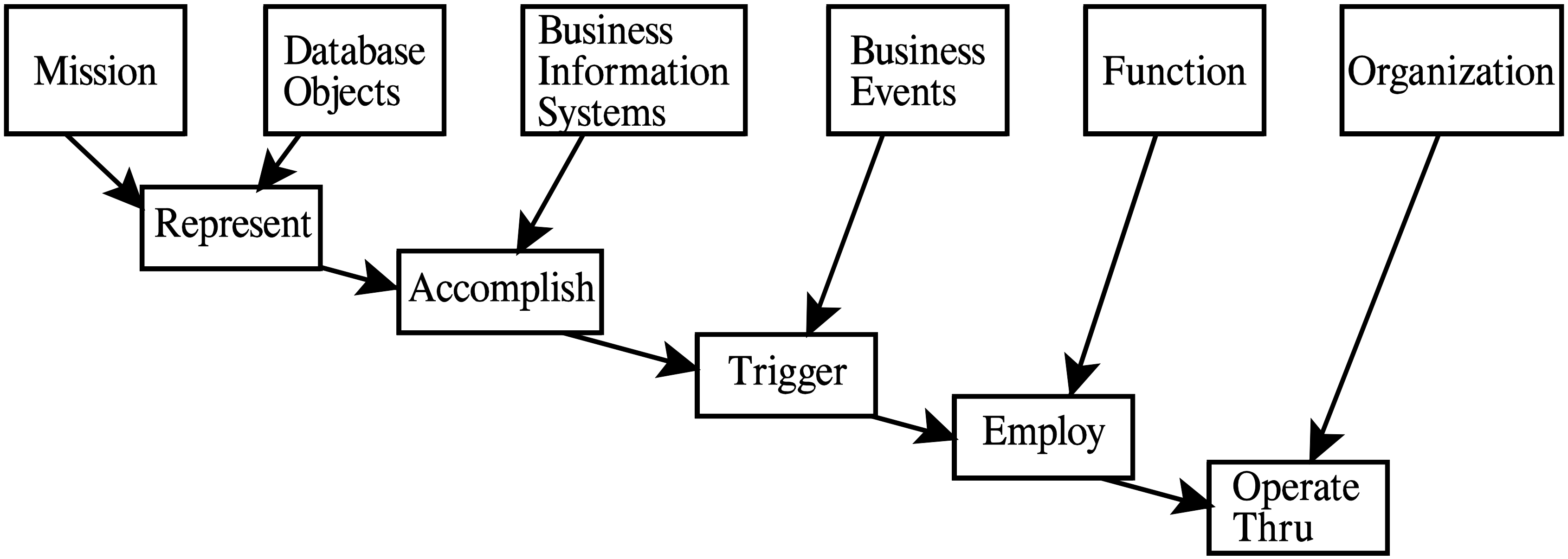
And the Resource Life Cycle Nodes are "IT" supported by : Business Information Systems and Business Events that exist within calendars and cycles...



And finally, the structure and organizations of all the necessary data models



These Meta Models Tell a Story...These six models are non-redundant, independent through many to many relationships, and integrated across the KWF Framework, and down through its rows of unfolding specification, implementation, deployment, evolution and operation.



All these models can be:

- Build bottom up
- Top down
- Independently and then interrelated

Suggestion...

Build Strategy	Model
Top-down	Mission, Organization, Function, Information Needs, Resource Life Cycle, Data Element Semantics (Concepts, Conceptual Value Domains, Data Element Concepts)
Bottom Up	Operational, Implemented, Specified, Data Elements, Value Domains
Maintenance	Put on critical path of every project.

4.2 Information Technology Components

- The IESS (Information Exchange Standard Specifications) gives us our data models for sharing data.
- The ADS (Authoritative Data Sources) gives us our source for truth, accuracy and precision.
- The EID (Enterprise Identifiers) gives us the ability to track critical assets throughout their life span in all involved activities and services.
- XML enables us to exchange data in formats that are not proprietary to either sender or receiver.

A key question left is:

Where and how do we store all our metadata (i.e., specifications and interrelationships) about all these things?

4.3 Metadata Repository

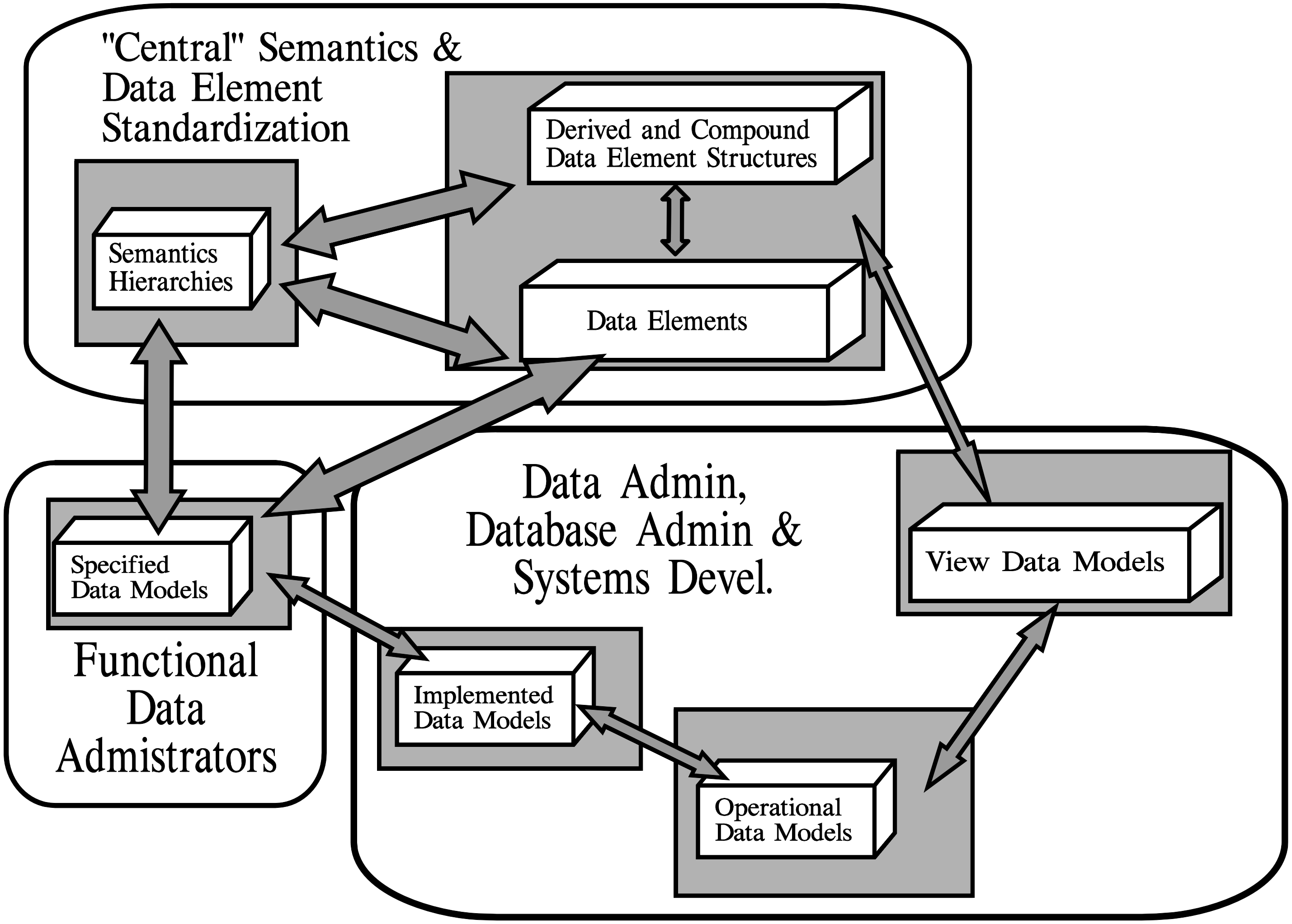
A metadata repository is just a database where all the data is “metadata.”

- The data architecture is just a schema of tables and columns of metadata types
- The application architecture is just systems of “CASE” (**C**omputer **A**ided **S**ystem **E**ngineering) like processes
- The technology architecture is just your garden variety SQL engine, presentation layers, diagram makers, and report writers.
- Buying a metadata repository is faster, but then you’re buying somebody else’s design and your critical metadata may be trapped in “secret structures.”
- Building a metadata repository is slower but then you have what you want and need, and the metadata is likely in open-access structures.
- Buying a template based Metadata Repository and then evolving to what you need is preferred and the metadata can be in open-access structures.

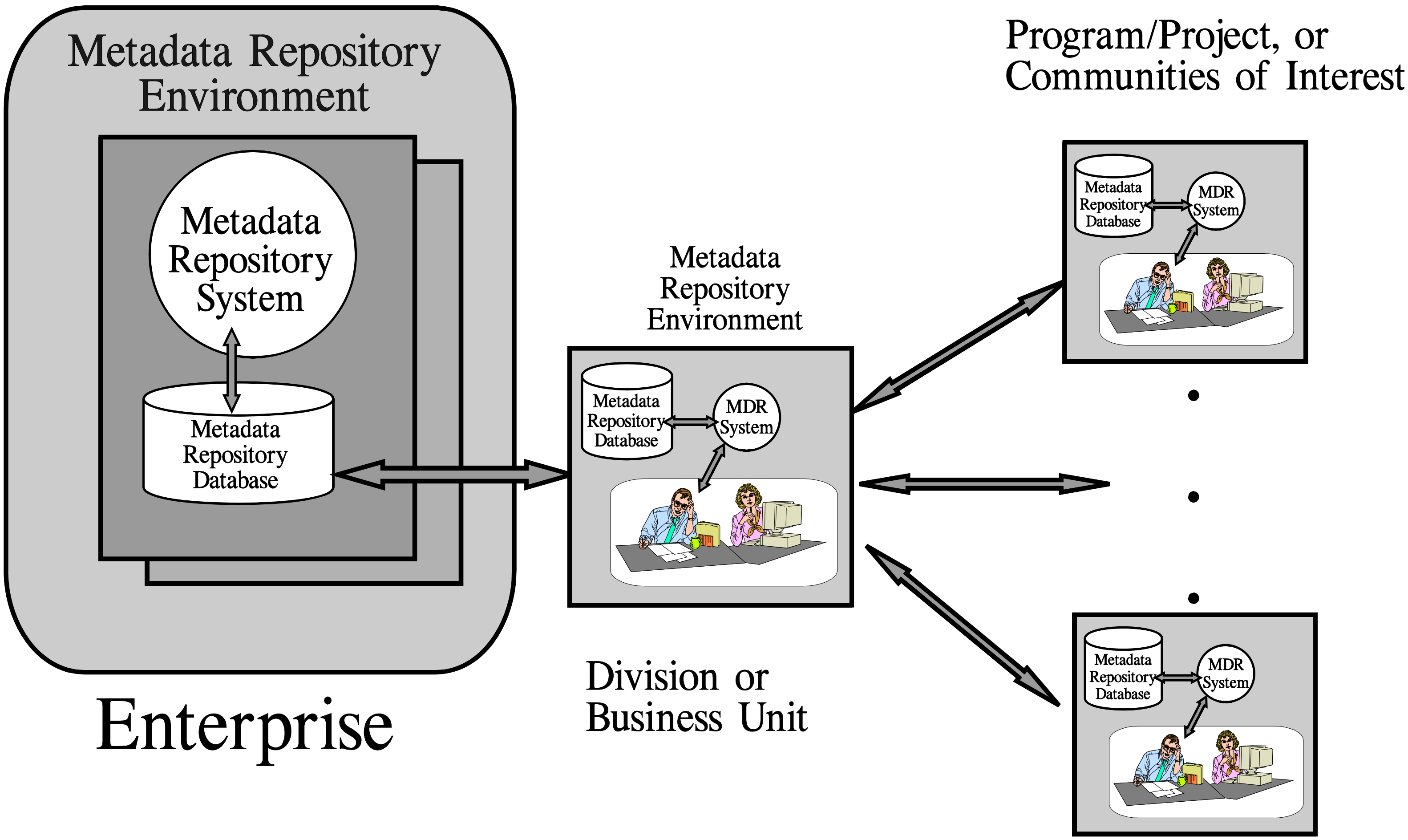
What is the Scope of the Metadata Repository

Knowledge Worker Framework						
	Mission	**Database Object**	**Business Infor-mation System**	**Business Event**	**Function**	**Organ-ization**
Scope	**To be successful at Knowledge Worker Management, all rows and columns must be integrated through:** ● An overall metadata model ● Methodology or work breakdown structure to move intra and inter-cell ● Metrics to predict the effort for small to large projects ● Deliverable specifications supported by objective quality control measures					
Business						
System						
Technology						
Deployment						
Operations						

Functional Distribution of “Data Model” Metadata



Distribution of Metadata



5.0 Data Interoperability Governance

- Data Interoperability Program Infrastructure
  - Program Committee
  - Finance Board
  - Procedures Board
  - Standards Policy Board
  - Standards Development Board
  - Study Groups

- Community of Interest Infrastructure
  - Community of Interest Committee
  - Operational Technical Committee
  - Systems Engineering Technical Committee
  - Data Modeling Technical Committee
  - Test and Evaluation Technical Committee
  - Configuration Control Technical Committee

Good Governance and Dilbert are Inversely Proportional

6.0 Data Interoperability Scenarios

- **Building and Employing Enterprise Architecture Models**:
  Scope and Business Rows, all columns. About 41% of all GAO IT errors.

- **Creating and Evolving Information Systems Plans**:
  System Row, all columns. About 8% of GAO IT errors.

- **Architecture and Engineering of Data Models**:
  Technology Row, database object column. Less than 2% of Errors

- **Performing Reverse Engineering of Legacy Systems and Databases**:
  Operations, Deployment, Technology and System Rows. Database Objects Column.
  Less than 2% of Errors

- **Forward Engineering Manufacture of New Systems and Databases**:
  Operations, Deployment, Technology and System Rows. Database Objects Column.
  Less than 2% of Errors

- **Employment Errors**:
  Systems through Operations Rows. Operations and Functions Columns.
  About 49% of Errors.

Distribution of Workshops Across Knowledge Worker Framework

<table>
<tr><th colspan="7">Knowledge Worker Framework</th></tr>
<tr><th></th><th>Mission</th><th>Database Object</th><th>Business Infor-mation System</th><th>Business Event</th><th>Function</th><th>Organ-ization</th></tr>
<tr><td>Scope</td><td colspan="6" rowspan="2">Enterprise Architecture Workshops</td></tr>
<tr><td>Business</td></tr>
<tr><td>System</td><td colspan="6">Information Systems Planning Workshops</td></tr>
<tr><td>Technology</td><td rowspan="3">Post Imple-menta-tion</td><td rowspan="3">Reverse and Forward Engineering Workshops</td><td colspan="2" rowspan="3">Accomplished by IT after systems and database are defined</td><td colspan="2" rowspan="3">Post Implementation Efforts to change enterprise to take advantage</td></tr>
<tr><td>Deployment</td></tr>
<tr><td>Operations</td></tr>
</table>

6.1 Building and Employing Enterprise Architecture Models

Workshops necessary to build the Enterprise Architecture Model:

- Missions, Organizations, Functions and Positions
- Database Domains
- Database Objects
- Information Needs Analysis

6.1.1 Missions, Organizations, and Functions

Missions Workshop	
Objective	Identify the set of missions that frame the essence of the enterprise. The collection of missions should be the business, no fewer and no extra.
Why	Missions set the context and stage for everything. Without missions, you're "Alice in Wonderland"
Time	Should only take two or three months for enterprise. Build from existing materials. This should result in about 100 mission leafs.
Cost	3 full time senior business and IT staff. $60K

Note: 100 missions contains one top mission node, 7 sub-missions on each of next two levels, and two sub-missions at the fourth level. That's a total of 99, or for the purposes of the estimation, 100.

It is practical to consider analyzing, designing, and building at the granularity of the second level. The third level is essentially applications.

Seven parallel efforts could thus be conducted for an enterprise-wide effort, or seven successive efforts if there was to be a serial effort. Estimates through these sections is at an enterprise-level which is neither recommended nor practical. Further cost refinements occur at the end of Section 6.

Organizations Workshop

Organizations Workshop	
Objective	Identify and describe the organizations of the enterprise charged with executing the missions.
Why	Organizations are the bureaucracy for accomplishing enterprise policy. It's key to know how they organized and which is responsible for different missions. Conflicts? Redundancy? Missing?
Time	Should only take two or three weeks for enterprise. Build from existing materials.
Cost	1 full time senior business. $20K

Functions Workshop

Functions Workshop	
Objective	Identify and describe the functions (human processes) of the enterprise charged with executing the missions from within the context of organizations.
Why	Functions are the human processes. Critical to know what they are, which organizations are doing them, and whether they conflict, are redundant,
Time	Assume four functions per mission leaf. That’s 400 functions. Assume 25% duplication. At 2 hours per formulation, that’s about four staff months, over two calendar months.
Cost	2 full time senior business and IT staff. $80K

6.1.2 Database Domains

Database Domains Workshop	
Objective	To identify and describe the data intensive areas necessary to fulfill the missions of the enterprise. Results in both leaf database diagrams and also enterprise database diagram.
Why	Database domains are initial descriptions of the data required for each mission. These ultimately result in database designs that support the missions. Ensures that there's complete coverage of all the mission areas of the enterprise.
Time	One day per mission leaf. One day for the leaf ER model. Several weeks to merge, discover database objects and create final ER model diagram.
Cost	If there are 100 mission leaves, then about 225 staff days total. That's about 10 staff months, or $200K.

6.1.3 Database Objects

Database Objects Workshop	
Objective	Identify the non-redundant set of database objects from across a set of "nouns" within the database domain diagrams.
Why	Database objects are really important because they keep you way above all the weeds. Each database object is really an important policy specification in the enterprise. Database object transformations are your major corporate processes.
Time	High level specification of each database object should take at most two days each.
Cost	If Enterprise has 50 database objects, then that's 5 staff months. Cost is $100K.

6.1.4 Information Needs Analysis

Information Needs Analysis Workshop	
Objective	Identify every key information need within mission area and fully describe each.
Why	Information is needed by individuals as they perform their functions through an organization in accomplishing missions. This becomes a good cross check on the sufficiency of a set of database models. If comprehensive and valid, ancillary and/or private databases and information systems can be eliminated. Eliminated too are then conflicting data, reports, analyses, etc.
Time	About 1 hour per Information Need
Cost	Assuming 4 information needs per function, then that's 8 staff months, or $160K.

6.2 Creating and Evolving Information Systems Plans

Workshops necessary to build the Information Systems Plan:

- Resource Life Cycle Analysis
- Business Information Systems
- Create Information Systems Plans

6.2.1 Resource Life Cycle Analysis

Resource Life Cycles Workshop	
Objective	To identify, describe, and determine the life cycles for the critical resources of the enterprise.
Why	Resources are the essential "enterprise possessions" (abstract or real) that are manipulated by the enterprise that cause income, expenses,
Time	About one week per resource. And about 2 weeks for all the enablement vectors.
Cost	Assume 20 resources in the enterprise. That's 5 staff months, or $100K

6.2.2 Business Information Systems

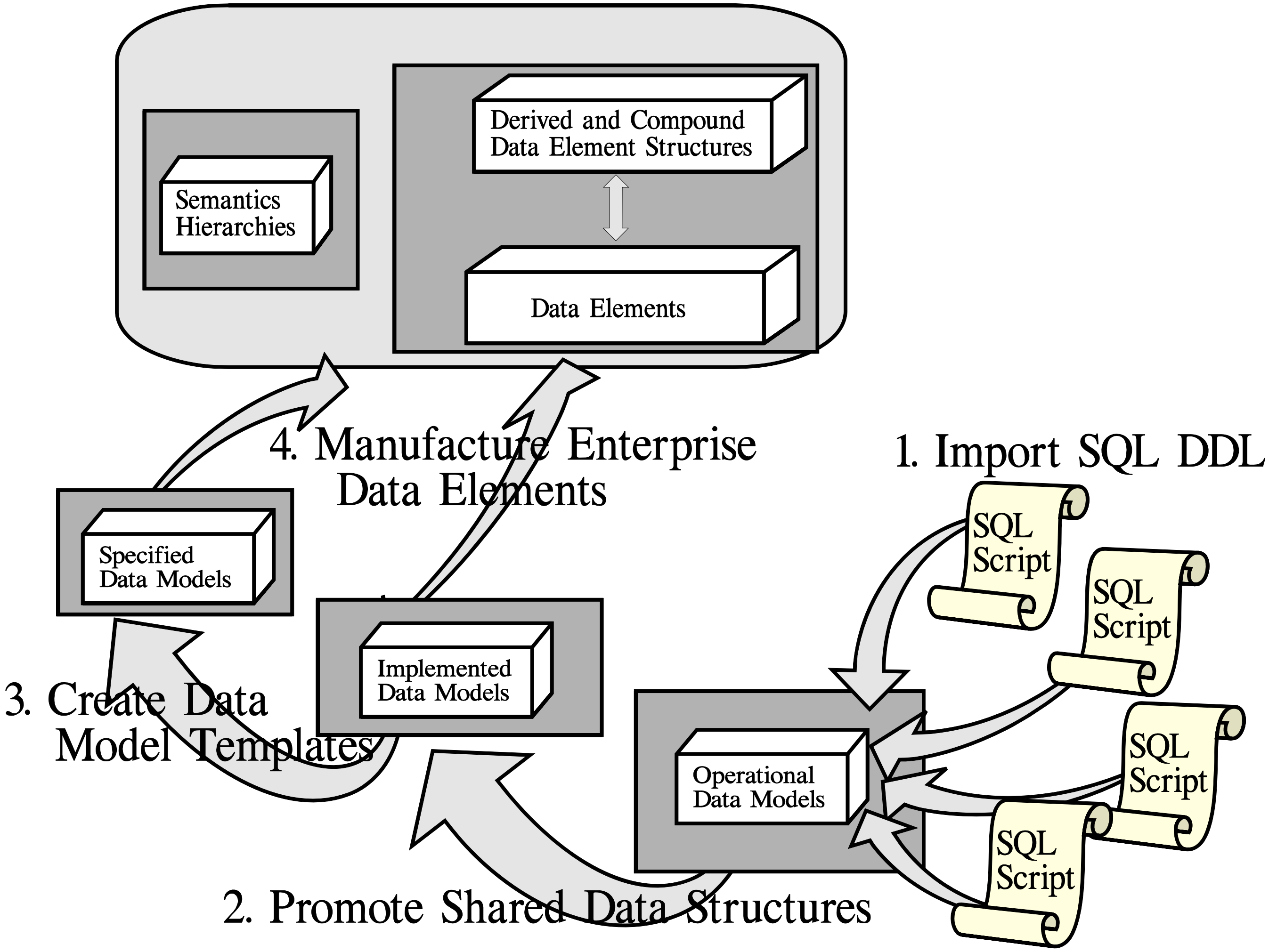
Business Information Systems Workshop	
Objective	Identify, describe, and create high level specify business information systems. Specify view data models as interfaces to operational data models.
Why	Business information systems need to be identified and described because a key objective is their creation, modification, or elimination (antiquated or combined).
Time	No more than about half-day per Business Information System
Cost	Assume 100 business information systems (one per mission leaf). That's 2.5 staff months, or $50K

6.2.3 Information Systems Plans

Information Systems Plan Workshop	
Objective	Create an Information Systems Plan from DBOs, BIS, and RLCA
Why	The information systems plan provides the sequence and interrelationships among the systems that are to be build, modified, combined or discontinued. It further provides a work breakdown structure, and estimates of staff resources.
Time	Given a full set of RLCs, BIS, and DBOs, then it is about 2 days for each information systems plan at the RLC node level.
Cost	Given 20 resources and 7 nodes each, that's 140 individual information system plans combined into one overall information system plan. That's a cost of $260K.

Traditional Approach computational note.
Under the least expensive traditional approach to Information Systems Planning the minimum staff years is 15. Hence, $3.6 million for a single, paper-based Information Systems Plan. Research performed at the MITRE Corporation in the middle 1990s.

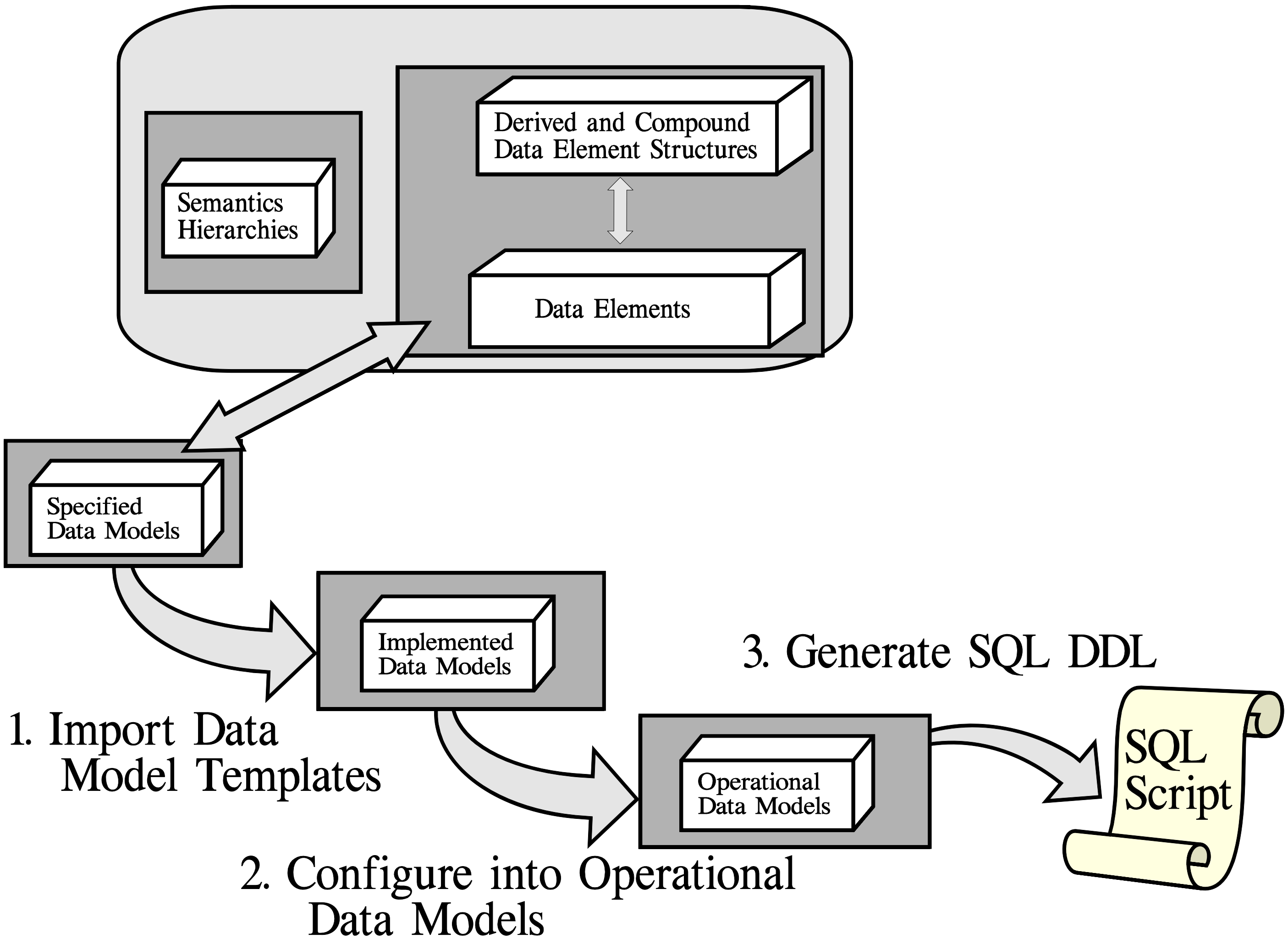
6.3 Reverse Engineering



Reverse Engineering Workshop	
Objective	To build a complete set of metadata at the Operational, Implemented, Specified, and Data Element layers.
Why	Building metadata from existing, legacy environment is the safest and most comprehensive approach. From the legacy layers, building the upper layers that are then consensus based semantics enables the construction of new databases and also building of the data interoperability transformation of the "As-Is" environment.
Time	For 100 schemas of 100 tables each and 15 columns per table, total staff days is 239. (See computation note).
Cost	11 staff months, or $220K

Step	Computation Element Notes	Hours
1	1 hour to import the legacy SQL schema. Assume 100 schemas and 100 tables for each schema.	100
2	1 hour to promote the schema to an implemented data model. Promote 1 out of 7 because of functional overlap	14
3	2 days to map/promote a new legacy schema to an existing IDM model. Mapping 93 ODM schemas to 7 IDM schemas. 186 days = 1116 hours (6 hours per day of productive work)	1116
4	1 hour to promote IDM schema to specified data model schema. Promote one IDM schema	1
5	2 days to map/promote a new IDM model to an existing SDM model. Map 6 IDM schemas. 12 days = 72 hours	72
6	10 minutes to promote a column/attribute to data element. Assume 30:1 column/attribute to data element produces ((7 schemas * 100 tables * 15 columns) /30) * 10 minutes. Creates 350 data elements	60
7	1 week to craft upper layers of data element model.	40
8	5 minutes to map a data element to a data element concept. 350 * 5 minutes	30
	Total Hours	1433
	Staff Days	239
	Cost at 22 staff days per month and $20,000 per month of staff cost	$220K

6.4 Forward Engineering



Forward Engineering Workshop	
Objective	To generate a new database design from metadata created during the reverse engineering process. Or to create a data interchange mapping between two or more legacy schema applications.
Why	New database designs need to be based on standardized semantics so that there is a slow but sure change over from legacy, non-standard, redundant and unintegrated data to standard, non-redundant, and integrated data. This will greatly ease the creation of data warehouses, the integration of ERP packages, use of general purpose query systems, report writers such as Crystal Reports, and analysis tools such as Cognos.
Time	For 50 schemas of 100 tables each and 15 columns per table, total staff days is 492. (See computation note).
Cost	11 staff months, or $438K

Computation Element	Hours
50 hours to create a new Implemented Data Model schema (100 tables) from metadata within Specified Data models. Estimate of 30 minute per table. Assume 50 schemas.	2,500
1 hour each to create an Operational Data model from the Implemented Data model. 50 schemas	50
8 hour each to adjust data types across all schemas	400
Total Hours	2950
Staff Days	492
Whitemarsh Cost at 22 staff days per month and $20,000 per month of staff cost	$438K
Traditional Cost ((160,000/2,950) * $438K)	$23.756 M

Traditional Approach Computational Note.

Under the traditional approach, a table takes about 2 hours per to identify and describe. Each column takes about 2 hours to identify, describe and define. Thus, for 50 schemas of 100 tables with 15 columns per table, that's 5,000 tables, and 75,000 columns. Total 160,000 staff hours.

6.5 SQL Generation and Application Generation

SQL DDL Generation Workshop	
Objective	Generate a SQL Schema for a particular Operational Data Model. This includes also the creation of SQL views as needed by applications.
Why	Every application that requires access to a database has to accomplish through the database's schema. Additionally, the application requires the creation of SQL views to then shield the application from changes that would occur in the schema.
Time	For 50 schemas and 2500 SQL views, total staff days is 425. (See computation note).
Cost	11 staff months, or $378K

Computation Element Note	**Hours**
1 hour to create the SQL DDL for the target DBMS. Assume 50 schemas	50
1 hour to create a SQL view. Assume 50 view per schema, and 50 schemas	2500
Total Hours	2550
Staff Days	425
Cost at 22 staff days per month and $20,000 per month of staff cost	$378K

Cost Comparison Summary		
Workshop	**Whitemarsh Approach**	**Traditional Approach**
Mission	$60K	Not accomplished
Organization	$20K	Not accomplished
Function	$80K	Not accomplished
Database Domains	$200K	Not accomplished
Database Objects	$100K	Not accomplished
Information Needs Analysis	$160K	Not accomplished
Resource Life Cycle	$100K	Not accomplished
Business Information Systems	$50K	Not accomplished
Information Systems Planning	$260K	$3.6 Million
Reverse Engineering (100 schemas)	$220K	Not accomplished
Forward Engineering (50 schemas)	$438K	$23.756 million
Schema Generation	$378K	$378K
Total Cost Across Enterprise	$2,066 million	$27.734 million

Traditional approach: takes longer, costs more, accomplishes less, continues stove-pipes.

6.6 Data Interoperability Scenarios Summary

Scenario	Description or Use
Enterprise Architecture Models 41% of all GAO IT errors.	Provides the overall context within which all data interoperability is created. Critical. Enables Information Systems Planning, Data Architectures, and all productive reverse and forward engineering
Information Systems Plans 8% of GAO IT errors.	Provides the ability to know what to do, when to do it, and a transformation map from As-Is to To-Be. Provides business basis for project sequence.
Data Architecture Reference Model <2% of GAO IT Errors.	Provides "technical" structure for capturing data model metadata. Provides architecture for having a define-once, use many-times approach. Provides basis for semantics standardization.
Reverse Engineering <2% of GAO IT Errors.	Enables new models and structures to be founded on a strong, legacy-based foundation. Causes the bottom-up build of the metadata critical to forward engineering.
Forward Engineering <2% of GAO IT Errors.	Causes new models to be built with a fraction of the resources and in a fraction of the time. All new models are standards-based and support data interoperability from the very start.

7.0 Key Measures of Return on Investment

Measure	ROI	Way to Measure	Cost Savings
Cost to develop Data Elements	Hard	Either use your existing metrics against new collections, or use industry metrics against your new collections	See Savings #1
Cost to develop Data Models	Hard	Either use your existing metrics against new collections, or use industry metrics against your new collections	See Savings #2
Cost to develop software applications	Hard	Measure the time and resources given modified processes and tools to create software systems	See Savings #3
Improve Answer Quality	Soft	Attempt to measure a greater consensus or a reduced quantity of discord in making decisions	
Improved Customer Support	Soft	Reduced complaints by customers, that end up being true that their accounts are not in order.	

Savings # 1: Consolidation, Non-Redundant Metadata ROI

Activity	Quantity	Cost via technique employed for definition
Starting quantity of columns/fields	19,000	$4.75 million
Elimination of closely named columns and fields reduced the quantity to	3,000	$1.06 million
Elimination of same concept but very differently named columns and fields reduced the quantity to	560	$200,000

Savings # 2: Data Model Development ROI

Quantity of Tables From estimate of a prototypical database	**Average Columns per Table**	**Total Col-umns**	**Process Driven Approach (2 hours per)**	**Data Driven Approach (2 hours per table, and 1/30th 11179 Data Elements**	
400	15	6,000	5.75 staff years	0.7 Staff years	**Cost Difference at $100 per hour**
		Cost	$1,200,000	$120,000	$1,180,000 in favor of data driven.

Savings # 3: System Development ROI

Quantity of Tables	Function Points Per Table	Cost Per Function Point	Type of Software	Total Cost	Cost Difference (in favor of data driven)
400 (If Data Driven)	80	$400	Information	$12,800,000	$43,008,000
1744 (If Process Driven (400 * 4.36)				$55,808,000	
400 (If Data Driven)	80	$1,000	Military	$32,000,000	$107,520,000
1744 (If Process Driven (400 * 4.36)				$139,520,000	

Estimates exclude hardware, computing infrastructure, travel, testing, documentation, evolution and maintenance.

8.0 Data Interoperability Summary

- Essential to achieve Data Management goals
- Approach and infrastructure has been proven and improved over 25 years.
- Metadata are the "books" of IT just as ledgers are the "books" of finance
- Quality metadata management is the "SOX" of IT (truth, integrity, and quality)
- The ROI from quality meta management is
  - Increased productivity: faster specifications, re-usable artifacts
  - Increased quality: consistency, integrity, authority, configuration management
  - Decreased cost: define once, use many times. Manufacturing databases and systems
  - Decreased risk: confidence based consistency. Proven through use and reus

If we have the resources to do IT Databases and Systems wrong, many times,

Don't We have the resources to do IT Databases and Systems Right, just one time?

Whitemarsh Website Links

The point of the link should be self-evident in the link's text.

http://www.wiscorp.com/
http://www.wiscorp.com/why_whitemarsh.html
http://www.wiscorp.com/whats_new.html
http://www.wiscorp.com/membership_purchase_process.html
http://www.wiscorp.com/data_interoperability_workshop.html
http://www.wiscorp.com/SQLStandards.html
http://www.wiscorp.com/tdan.html
http://www.wiscorp.com/products.html
http://www.wiscorp.com/new_updated_products.html
http://www.wiscorp.com/EnterpriseDatabase.htm
http://www.wiscorp.com/DatabaseProjects.htm
http://www.wiscorp.com/MetabaseProducts.htm
http://www.wiscorp.com/DatabaseDesignInformation.html
http://www.wiscorp.com/metabase_demo.html
http://www.wiscorp.com/featured_papers.html

Data Interoperability Community of Interest Handbook

(http://www.wiscorp.com/printed_books.html)

This book was modeled after three very successful communities of interest from IT. This is a very practical, day to day, engineering, construction and development guide for the creation of Data Interoperability Standards.

This 320 page book provides:

1) Rationale for Shared Data Environments across the Enterprise;

2) Blueprint for Collaborative Data Sharing;

3) Why Communities of Interest are Ideal Organizational Structures;

4) How Communities of Interest are Best Engineered;

5) Step-by-Step Strategies to Build Common Data Architectures;

6) Methodology & Plan to Create Focused Organizations;

7) Strategies for Meetings, Decision Making, and Voting; and

8) Much, Much More!