



Whitemarsh
Information Systems Corporation

Understanding-Based Data Interoperability Engineered via Army Data Management Program

Whitemarsh Information Systems Corporation

2008 Althea Lane

Bowie, Maryland 20716

Tele: 301-249-1142

Email: Whitemarsh@wiscorp.com

Web: www.wiscorp.com

Topics

- Net Centric Data Goals
- Data Management and Net Centric Goals
- Key Components of a Data Management Architecture
- Key Metadata Models Critical for Data Management
- Reverse Engineering to Achieve Interoperability
- Forward Engineering to Manufacture Interoperable Systems
- Way Ahead to Success (More for Less, Faster & Lower Risk)



DoD Net Centric Data Goals	
Goal	Description
Make data visible	Users and applications can discover the existence of data assets through catalogs, registries, and other search services.
Make data accessible	Users and applications post data to a “shared space.” Data assets, are made available to any user or application under security, etc.
Institutionalize data management	Data approaches are incorporated into Department processes and practices.
Enable data to be understood	Users and applications can comprehend the data, both structurally and semantically.
Enable data to be trusted	Users and applications can determine and assess the authority of the source because the pedigree, security level, and access control.
Support data interoperability	Many-to-many exchanges of data occur between systems, through metadata enhanced.
Be responsive to user needs	Perspectives of users are incorporated to ensure satisfaction.



Understanding Based Data Interoperability	
DoD NC Data Goal	Data Management Support
Make data visible	Identifies data assets within natural contexts of mission, organization and function. Standardizes taxonomies, ontologies and classification schemes to view data semantics including where and how implemented
Make data accessible	Includes Discovery Metadata definition in every data asset. Standardized names, definitions and structures
Institutionalize data management	Standardizes strategies for data definition. Multiple layers to ensure define once use many times. Use of ISO standards 11179 for data elements and SQL for data models.
Enable data to be understood	Standardizes vocabularies, commonly inherited semantics, commonly used data model templates, automatic names and definitions based on well defined words. Automatic abbreviations where necessary.



Understanding Based Data Interoperability	
DoD NC Data Goal	Data Management Support
Enable data to be trusted	Contains consistent semantics, standard reference tables, completely mapped data models across multiple levels of abstraction.
Support data interoperability	Standardizes data structures, well engineered data transactions, automatic XML wrapping of data, supported by accessible data definitions and contexts.
Be responsive to user needs	Supports reuse of already defined data assets metadata, central knowledge of all data assets and distributed access to same.



Four Key Components of Army's AR 25.1

1. IESS (Information Exchange Standards Specifications)

- Each is represented by a Logical data model,
- Contains Business rules for the exchange of information
- Determined by the Community of Interest
- Consistent with the intent for data interoperability
- IESSs help define the requirements for the exchange of structured data
- **Source** for XML Schemas and XML data exchanges



2. ADS (Authoritative Data Sources)

- Reference Data: focused on standardized codes, e.g., country codes, supply codes
- Extended to operational data to ensure that a community is getting the same data
- Level of confidence that the data is accurate and consistent
- Requires Enterprise Identifiers (EIDs) to manage configuration and distribution



3. EID's (Enterprise Identifiers).

- Enable the easy combination of, or more basically, the identification and/or reference of disparate data items
- Enables linkage across arbitrary data sources.
- Facilitates transformation of stove-pipe and independent system database into a single, virtual, distributed database
- “Glue” for linking and mapping data sources.
- Allows arbitrary dependancies to be established to create higher level information structures for decision makers.
- Supports structured and non-structured data.
- Enhances XML solution (So you know “what’s what” and “where’s what”)



4. XML Based Data Exchange.

- XML Schemas for data exchange data models
- SQL DDL Schemas for source data models
- SQL DDL Schemas for data exchange models



XML is NOT the Silver Bullet

- Suppose you have a MIS System.
- Given there are 100 windows, and each window has 3 sub-screens.
- Each sub-screen's data specification has to be represented by a XML Schema.
- Thus, for this ONE system you would need $(100 * 3)$ 300 XML schemas
- The Army's 20% of DoD. It has 125,000 systems. For DoD, that's 185.5 Million
- If each XML schema has 15 XML elements, then the needs to be 2.81 Billion XML element definitions.
- That's 10 times LARGER effort that the DDDS effort that DISA *stopped*.
- If each XML schema takes 8 staff hours to define, that's 713K Staff Years.



Key Components of Army's Net Centric Data Management Program (ANCDM) – DA PAM 25.1.1

- DPP (**D**ata **P**erformance **P**lanning) identifies, plans, and manages projects associated with data assets throughout the entire lifecycle of the data asset.
- DPP projects produce products for future and improved data integration and reuse.
- DPP enhances project scoping, responsiveness to business change, and management of systems development sequencing and prioritization.
- Data asset products may be explicitly identified as DODAF views or may exist implicitly within DODAF views.
- All data asset products must be integrated and interrelated within and across COIs.
- All data asset products in the DODAF are metadata and must be Core Architecture Data Model (CADM) conformant, developed through CADM conformant tools.



Key Components of DA PAM 25.1.1 (Cont.)

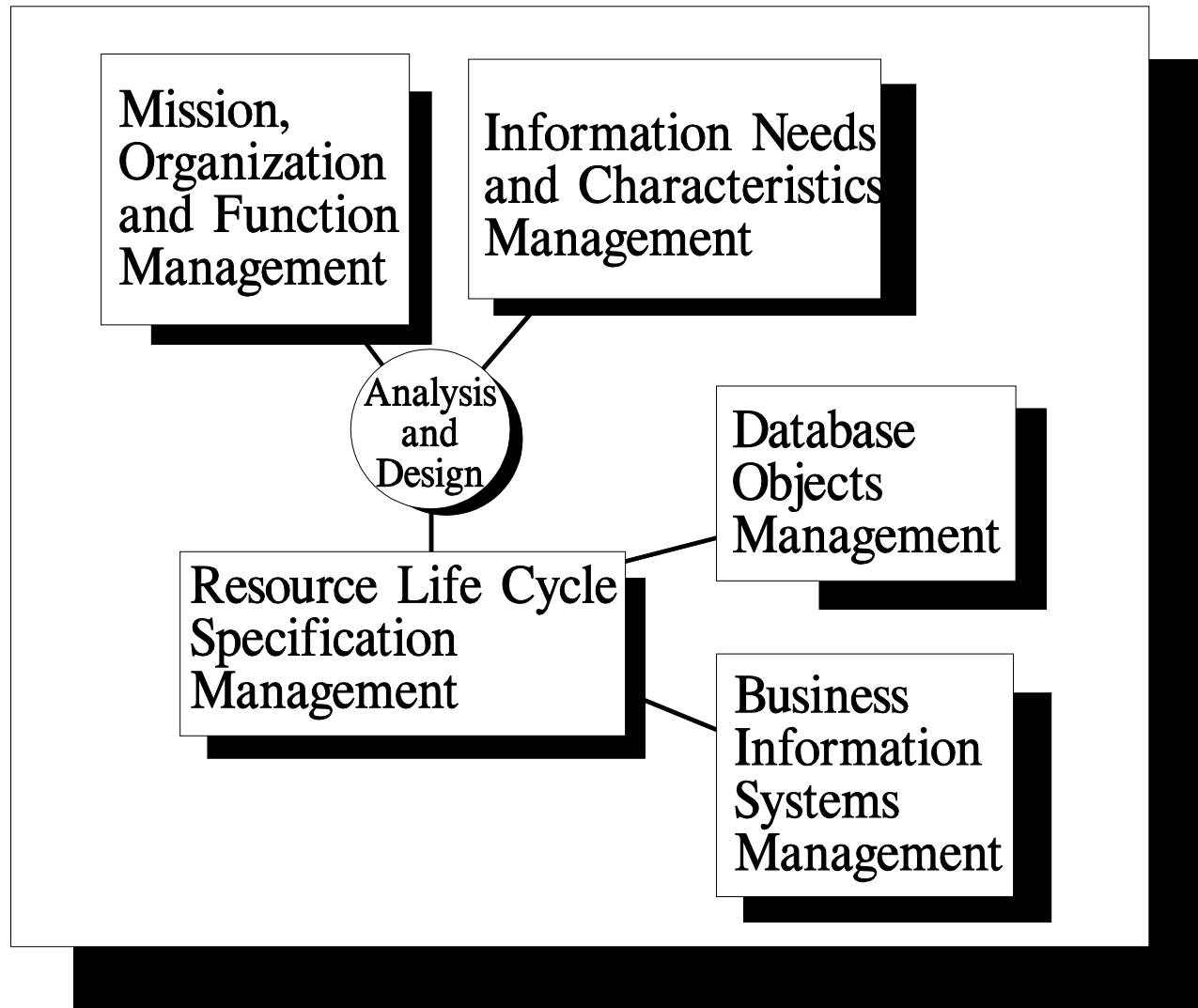
- In order for information to be fully understandable and interoperable...
 - ◆ Semantics and syntax must be well specified;
 - ◆ Data elements must be identified at the enterprise level;
 - ◆ Authoritative data sources must be well defined and managed;
 - ◆ Exchange mechanisms must support current and future demands.

- The COI structure consists of three layers.
 - ◆ Army data harmonization and integration (ADHI)
 - ◆ Institutional
 - ◆ Expedient COIs

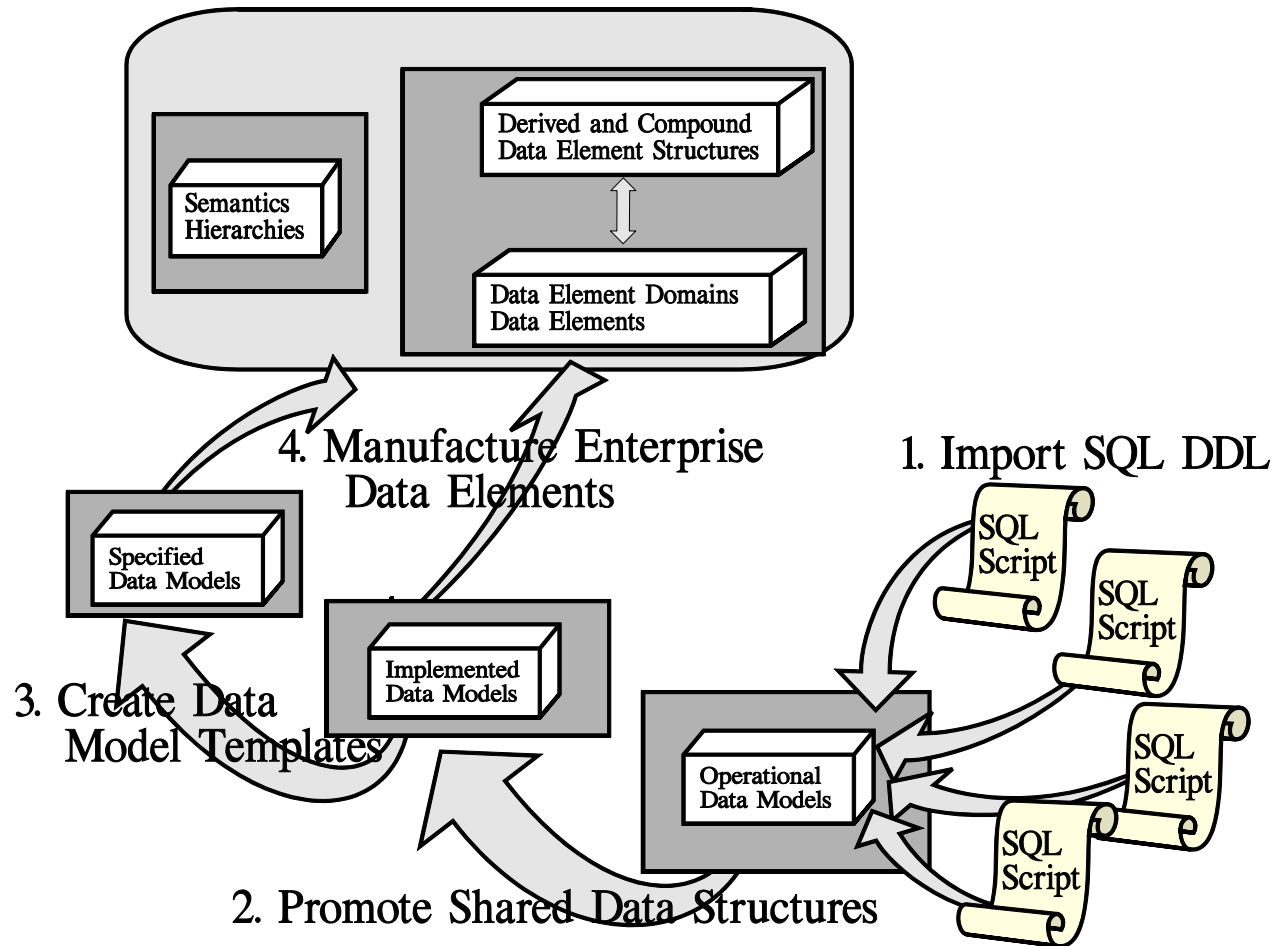
- The ANCDMP consists of three distinct layers:
 - ◆ ANCDMP layer (top-down guidance and bottom-up data asset development)
 - ◆ Data Performance Planning layer (i.e., COI Program Management Plan for the development, scheduling, resource management. and evolution/migration strategy)
 - ◆ Project execution layer (accomplishes the ANCDMP projects)



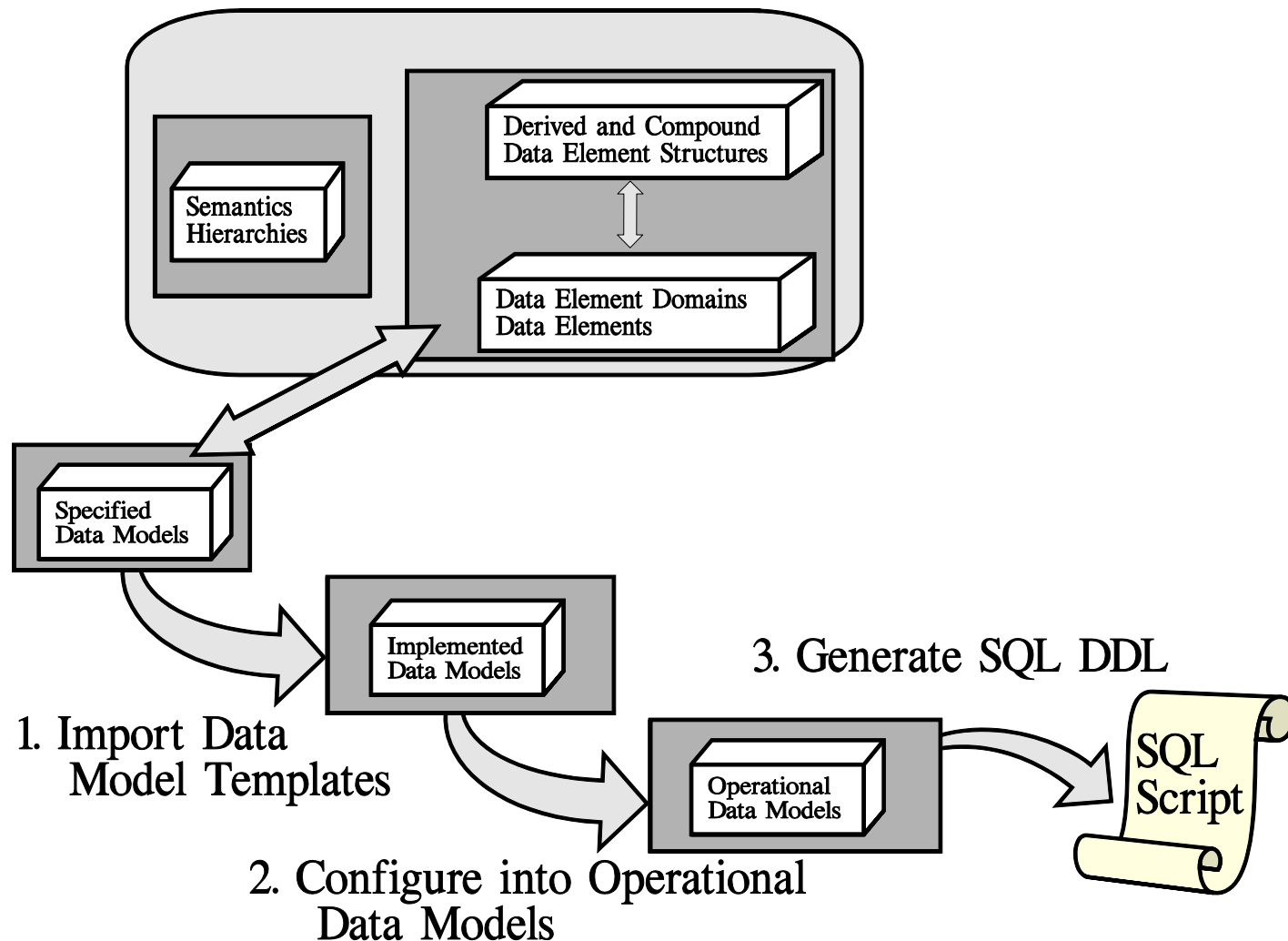
Key Metadata Models Critical for Data Management



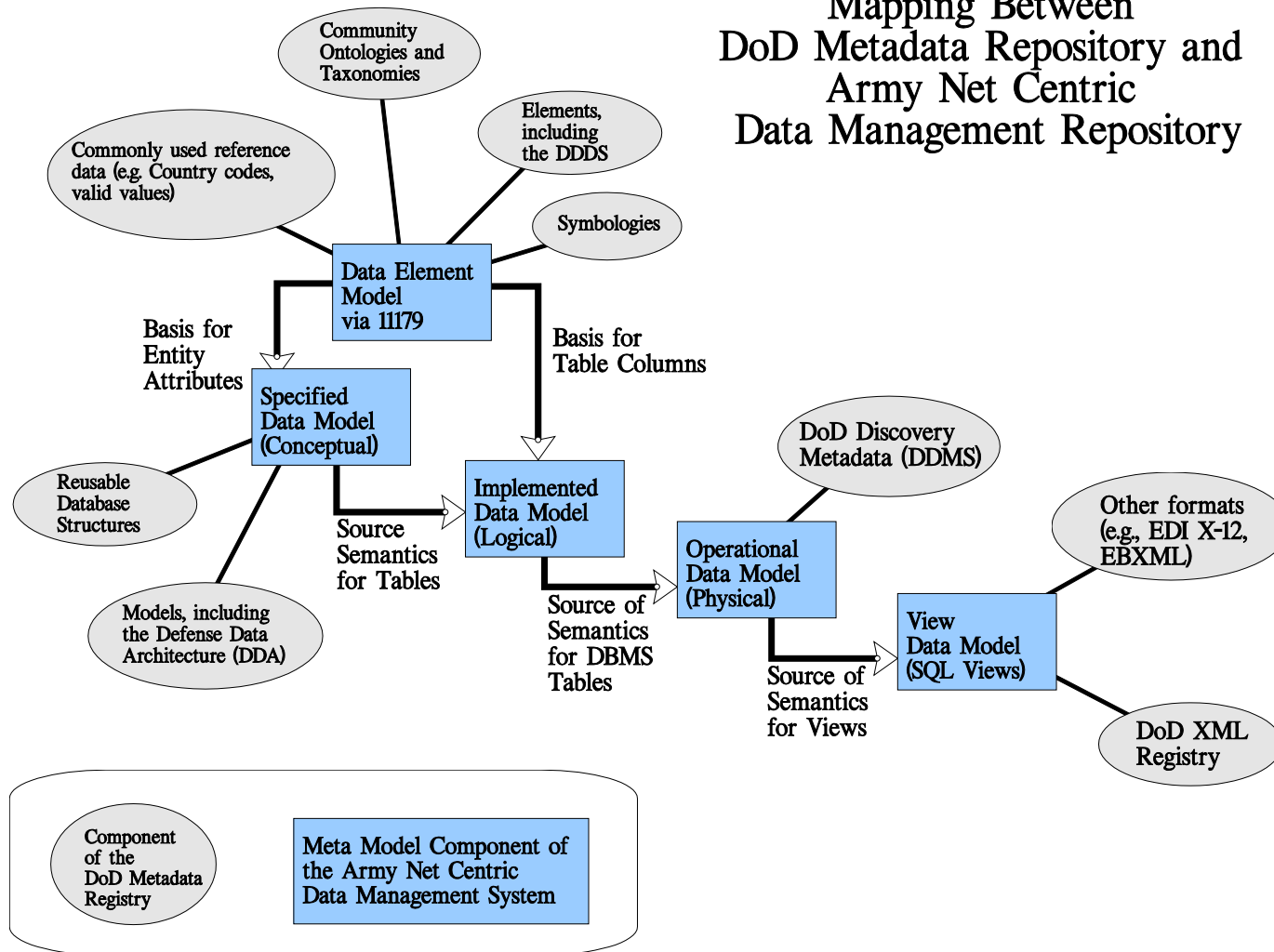
Reverse Engineering to Achieve Interoperability



Forward Engineering to Manufacture Interoperable Systems



Mapping Between DoD Metadata Repository and Army Net Centric Data Management Repository



Key Benefits (Data Model Development)

Quantity of Tables From estimate of a prototypical database	Average Columns per Table	Total Columns	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.



Key Benefits (System Development)

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.



Way Ahead to Success (More for Less, Faster at Lower Risk)

- For information superiority, you must first have a data management reference model
- You must engineer and accomplish data standardization, heeding lessons from the DDDS and DDA.
- You must engineer data asset products into clear and precise specifications
- Store, interrelated, and employ all data asset product specifications from a metadata repository
- Metadata repositories must be federated.
- Create prototypical data architectures for define once, use many times.
- Employ Enterprise Identifiers, Authoritative Data Sources, Information Exchange Standards Specification, and XML to ensure integration, re-use and non-redundancy
- Automate XML schema generation from metadata repository
- Manufacture databases by assembling pre-engineered standardized parts.

