". . . we must have interoperability at the data level."

-- U.S. Department of Navy
http://www.chips.navy.mil/archives/00_jul/data_interoperability.html

The Data Is Key!

Jerry Smith
DISA Interoperability Directorate
presentation to
CALS/EC International Symposium
January 2001
Tokyo

- **Criticality of Interoperability**

- **Difficult Challenges**

- **Technology Advances**

- **Past Approaches**

- **Recent Initiatives**

- **Lessons To Be Learned**

- **Summary**
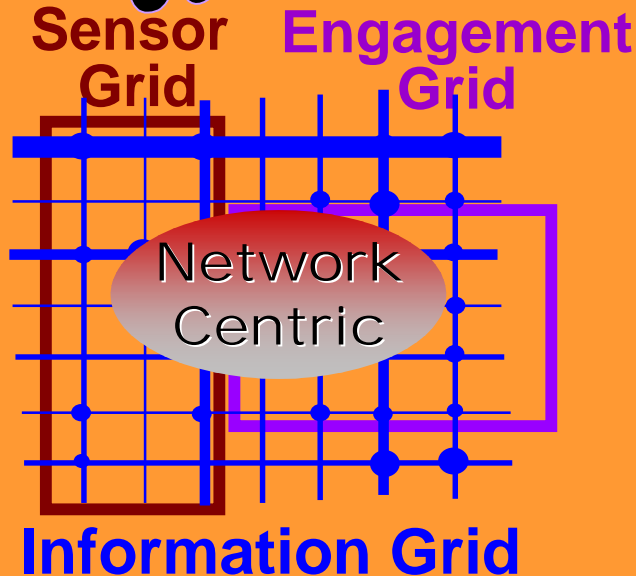
RIGHT DATA ….

RIGHT PERSON ….

RIGHT TIME!

Information Sharing for the Virtual Enterprise

Requires

- Seamless Architecture & Systems Integration

- Responsive Information Collection, Processing & Dissemination
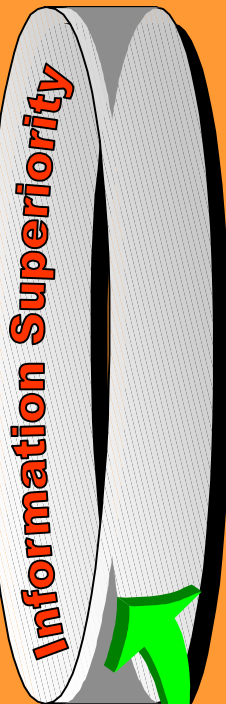
- Offensive & Defensive Information Warfare

Others...

Information Superiority

Enabling Concept

Concept

Sensor Grid

Engagement Grid

Network Centric

Information Grid

End States

Common operational picture

Collaborative planning systems

Networked sensors
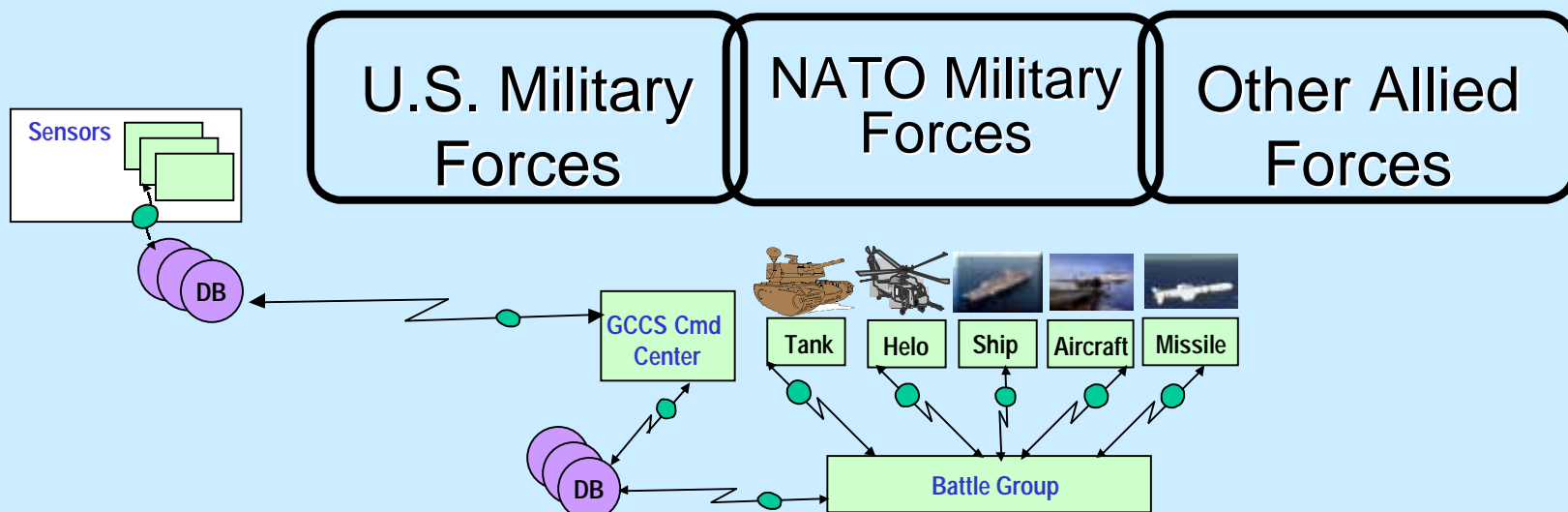
Enhanced C2 systems

Precise battlespace knowledge

Information Superiority

Global Information Grid (GIG)

Intel

Sensor

Intel Reachback

| Comm Svr | Data Svr | App Svr | Client | Client |

ISSE

C2 Reachback

| Comm Svr | Data Svr | App Svr | Client | Client |

JTF Deployed

| App Svr | Client | Client |

SECRET/REL

LOG Reachback

| Data Svr | Data Svr | App Svr |

C2G

SBU/REL

NIPRNET (SBU)

PERSONNEL Reachback

| App Svr | Data Svr | App Svr |

FW

Internet

Shooter

Joint Operations Also Require Interoperability for Multi-National Forces

| U.S. Military Forces | NATO Military Forces | Other Allied Forces |

Sensors

DB

GCCS Cmd Center

DB

Tank · Helo · Ship · Aircraft · Missile

Battle Group

- **Multi-national Joint interoperability depends upon the ability to share data. Realistic assessment includes an evaluation of the effectiveness of IT standards used**

The Defense Department's top adviser to the Joint Chiefs of Staff [LTG Kellogg, J6] on information technology issues says the federal law that entitles the military services to equip their own forces should be revised to ensure that the services buy systems capable of sharing data.

- "When you look at the situation from the level of information [data in context], rather than from the level of data, there are other problems," e.g.:
  - **Difficult to correlate, or fuse, data from the current systems,** meaning that different sensors often provide overlapping coverage at varying levels of accuracy.
  - Duplicate and sometimes conflicting data can **cause confusion and misinterpretation**.
  - Different systems performing only limited data management, mean that **information can be inconsistent, incorrect, or simply lost**.

Source: Carmen Corsetti quoted in "Roving Sands," at http://www.mitre.org/pubs/showcase/roving_sands.html

Data Exchanges Remain Primitive

- ## Operations:
  Air tasking orders transported via "floppy disk" shore to ship in Gulf War. Faulty target identification resulting in destruction of civilian facilities

- ## Logistics:
  Conexes containing a meals rather than other supporting material

- ## Medical:
  Systems unable to track location and condition of injured personnel during evacuation operations

- ## Acquisition:
  Unable to share product data throughout the product life cycle

WHY STEP?

Never underestimate the importance of getting the right data at the right time.

PLCS

A joint industry, government inititative to develop international information standards for through-life product support

An international project designed to produce working data models and draft standards in three years

PLCS utilizes ISO 10303 STEP - the STandard for the Exchange of Product model data

Solving Problems

What version/configuration is it?
The maintenance information is outdated, inaccurate and unavailable when it is needed the most…
I need to reduce my inventory and spares costs
How do I get accurate in-service feedback?

Goals

Significantly improve product availability through improved support capability
Improve the cost, quality and accessibility of Product Life Cycle Support information
Accelerate technical development of the ISO standards
Encourage early implementation by commercial vendors

Through Life Business Model

Operate

Support

Design

Produce

| Effectivity |
| State |
| Configuration |
| Organisation |
| Product ID |

Consistent Linking & Referencing →

Support Engineering

Resource Management

Core Concepts & Referencing
Anomaly
Task
Requirements
Identification
Configuration
State
Applicability
Documentation

Configuration Management

Maintenance And Feedback

Management Information

Target Areas

Configuration Management — Manage change throughout the product life cycle, with the provision of tracking of serial numbers where applicable

Support Engineering — Provide and sustain the support infrastructure

Resource Management — Buy, store, pack, move, issue and dispose of physical products

Maintenance And Feedback — Maintain, test, diagnose, calibrate, repair, and modify physical product, including schedules, resources and feedback
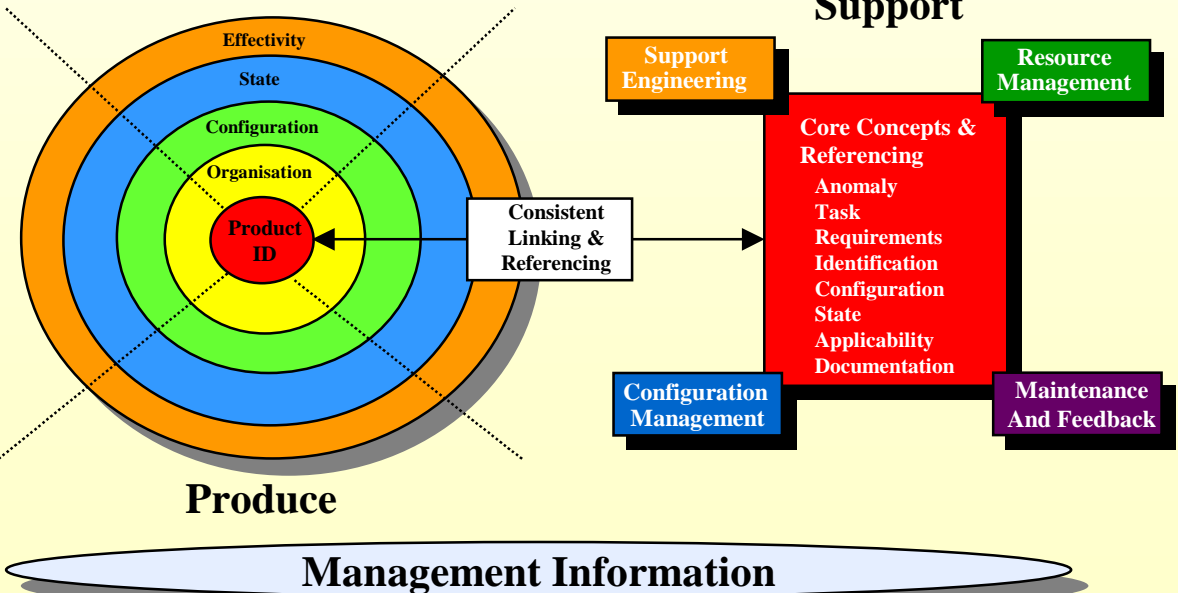
Product Data Example
PROBLEM: product data sources are not integrated vertically or horizontally

Engineering
Drawings
JEDMICS

-- 4 Tech
Order

CAD
Databases

Vendor
Catalogs

How to Share?

Personal
Filespecs

Test
Specs

T&E
Reports

Manufacturing
Process
Descriptions
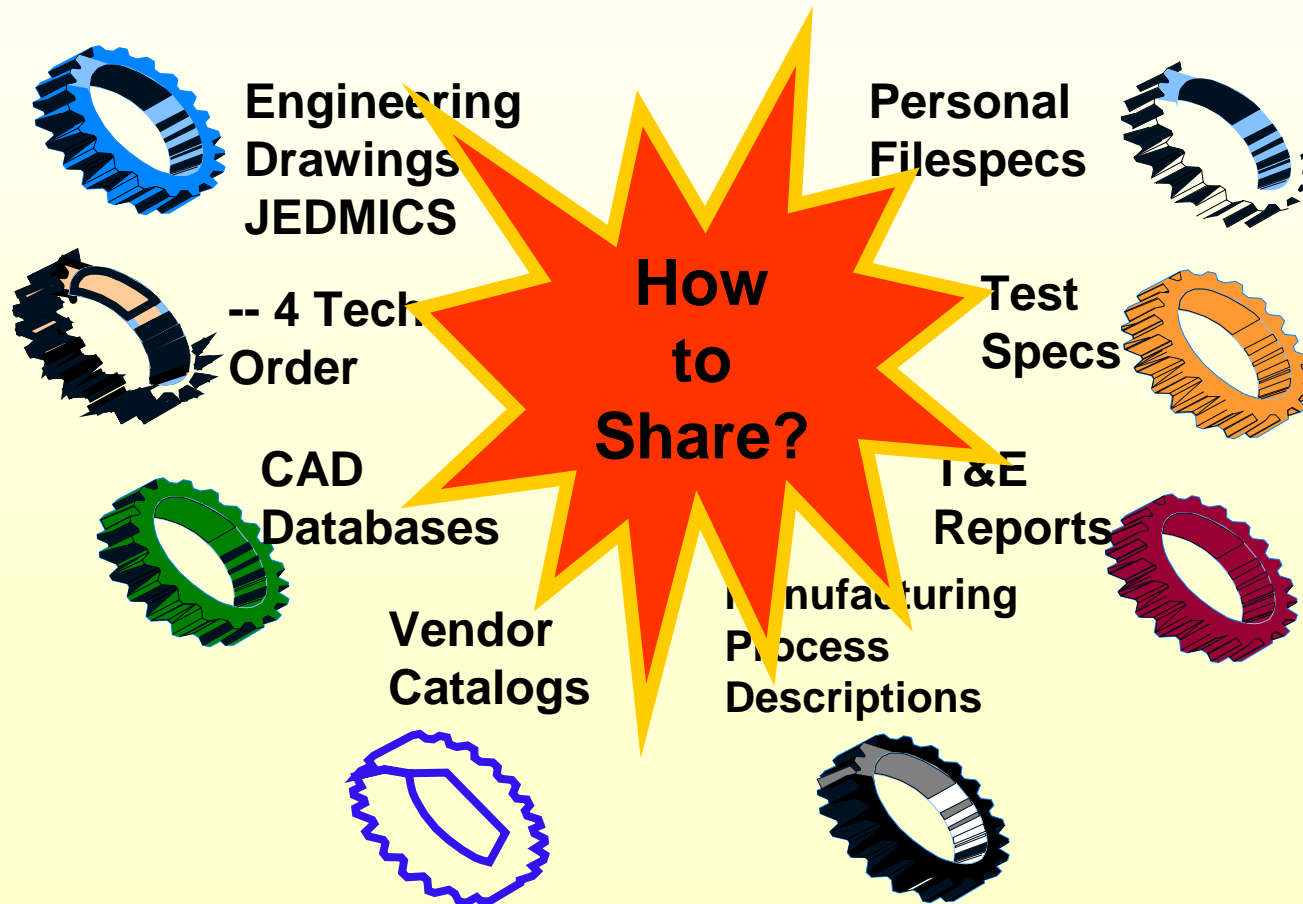
Information Explosion
DATA TSUNAMI!

By 2000 AD
- Information doubles 1 time every 3 years
- Data will double > 20 times in only 6 years

Thus, in 6 years
Data 125,000 > Information!

e.g., "Documents" now include digitized

Text   Photographs
Graphics   Audio
Color   Animation
Video

Next 6 Years

DATA

INFORMATION

SOURCE: George Gilder, "Life After Television"

Jerry Smith Tokyo January 2002

Interoperability Problems

- **Portability and dependability is required across heterogeneous environments.**
- **In spite of all the technological advances, data exchanges remain primitive.**
- **Significance of the data itself is evidenced by operational problems and lack of interoperability.**
- **There has been too little focus on data meaning within context!**

- CALS Origins
- CALS Vision: Implement a Logistics Virtual Enterprise
- Name Evolution:
  - ***C**omputer **A**ided **L**ogistics **S**upport*
  - ***C**ontinuous **A**cquisition and **L**ogistics **S**upport*
  - ***CALS/EC***
  - *"**C**ommerce **A**t **L**ight **S**peed!"*
- CALS/EC in U.S. DoD Today

Some organizational changes but still pursuing the original CALS goals and objectives

CALS Vision

CALS Leads the Transition to Integrated Digital Operations Utilizing Value Added Technical and Business Information to Increase Productivity

CALS Objectives

- Implement Integrated Data Environments
- Integrate Configuration & Data Management Practices
- Foster Global Competitiveness
- Develop ISO Standards as Enablers
- Implement a Virtual Logistics Enterprise
- Reduce O&S Costs
- Support Paper Free Contracting

Shared Information: Foundation for 21st Century Productivity

❖ Integrated Data Environments:
  – Collaboration
  – Enterprise Connectivity

❖ Productivity Gains Realized through:
  – Shared Information Access
  – Improved Processes
  – Common Infrastructure

Much Progress … BUT

- **Computing**
- **Communications**
- **Systems, Software & Data Engineering**
- **Standards**

Exchanging the Meaning of Data Within Context is Still a Major Challenge!

Data in Context

- **<u>CONTEXT</u>** really matters . . . not just communications links

- Critical portion among combat systems and data links that involve tracking and delivery of ordnance. . .

- Sharing accurate and timely data is a technical objective, programmatic focus, and most importantly it is an operational requirement (*warfighter position*).

Meaningful Data Is Key

- *"Data plays a huge role in the interoperability equation."*
- *"Further investigation of data integrity, data reference table synchronization and data standardization is needed to resolve interoperability issues."*
- *"Data management and data standardization should be key areas for near term investigation."*

Source: Worldwide DoD CIO Conference, circa 2000
http://www.c3i.osd.mil/doc/dodcio-2000conf/New%20Pages/RTFinal_Interoperability.htm

Data Interoperability Challenge

<u>Expect</u> Heterogeneity!

- No Single Standard *can* be defined!
- Various Defense communities <u>will</u> adopt multiple "standards:"
  - Government (message, database, symbology), Commercial, International, de facto/legacy etc.
- Implementation <u>will</u> vary across systems *even* within communities

DoD 5000.2-R *"It is DoD policy to develop software systems based on... use of <u>standard data.</u> Additional guidance is contained in DoDD 8320.1"*

DoDD 8000.1 "It is DoD policy that ... standard DoD data definitions shall be used for all ISs, to include the interfaces between weapon systems and the ISs"

Joint Technical Architecture "The mandated standards for DoD data definitions are DoD 8320.1-M-1and the DDDS"

"...applies to data elements and values that are unique to the operation of equipment and software that are an integral part of a weapons system and related test equipment"

Data Administration: "The responsibility for the definition, organization, supervision and protection of data within an enterprise."

DoDD 8320.1 DoD Data Administration

DoD 8320.1- M - 1 Data Standardization Procedures

"Levy burden and cost of conversion to nonstandard data, regardless of the origin of the requirement for the information on the Component using nonstandard data"

Data Element Life Cycle

Developmental — Candidate — Modified / Approved / Disapproved — Archived
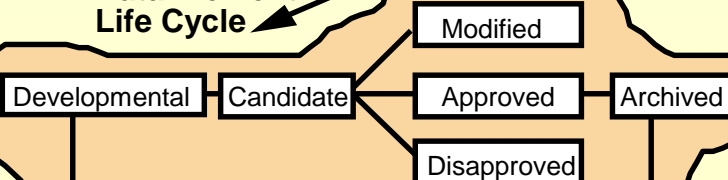
Defense Data Dictionary System (DDDS)

Responsibilities:
- **Functional Areas**
  - USD (A)
  - USD (C3I)
  - . . .
- **Components**
  - Army
  - Navy
  - Air Force
  - Defense Agencies

PAST INITIATIVES
Standard Data Elements

C3	1,678
Intelligence	317
Information Mgmt	737
Logistics	2,235
External Standards	1,335
DDR&E	2,589
Personnel & Readiness	1018
Health Affairs	769
Finance	449
Procurement	652
Installations	230
Environmental Security	684
Other	628
Total	13,321

A/O 9 March 1999

Logistics
2,235
17.8%

Environmental
Security
684
5.1%

Intelligence
317
2.4%

Information
Management
737
5.7%

C3
1,678
13%

Other
628
4.7%

External
Standards
1,335
10.3%

Installations
230
1.7%

DDR&E
2,589
18.7%

Health Affairs
769
5.3%

Finance
449
2.6%

Procurement
652
4.8%

Personnel &
Readiness
1018
7.3%
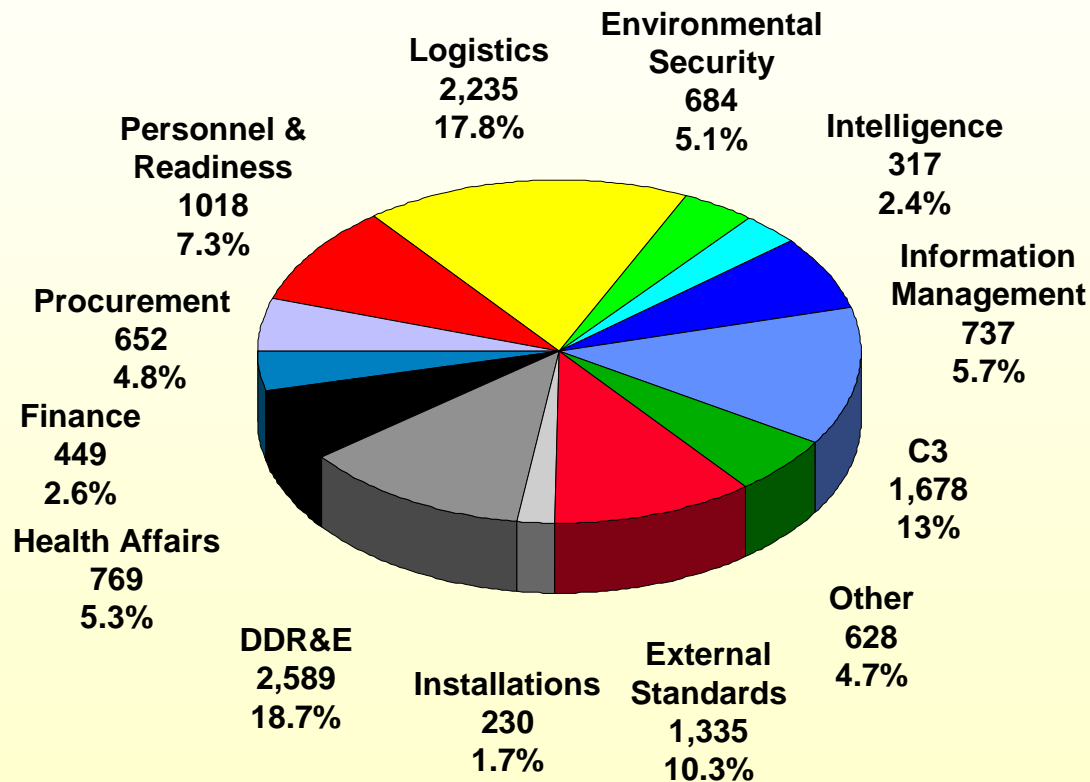
Implemented in AISs: 6,937

What's Broken?

- ## Multiple standardization efforts

  - Data Standards
  - Symbology Standards
  - Terminology Management Standards

  - Message Standards
  - De Facto Standards
  - Commercial Standards

  Result: Multiple terms, definitions and structures evolving in different CM cycles (and wasted resources)

- ## Weak Metadata support

  – Poor visibility and distribution of "in situ" data assets

  – Hard to link existing GOTS & COTS data resources

  – No consistent way to show which systems use what data

- ## Slow change mechanisms

  Too much *PROCESS* -- not enough *PRODUCT!*

Some Current DoD Approaches to Data Issues

- Data Engineering (COE SHADE – shared data elements) & Initiatives
- Market-driven Data Management
- XML Repository
- PDML Project
- Semantic Mediation
- DAML (DARPA Agent Markup Language)
- PLCS Project

Shared Data Engineering (SHADE)

- **Data services infrastructure for the DII COE that promotes information**
  - Sharing
  - Interoperability
  - Software reuse

  . . . in a secure, reliable, global environment

- **The infrastructure is implemented as a set of**
  - Shared schema
  - Services
  - Tools
  - Operating procedures

  . . .supporting COE-based mission applications.

Addresses data issues necessary to achieve system interoperability and data sharing goals within DOD
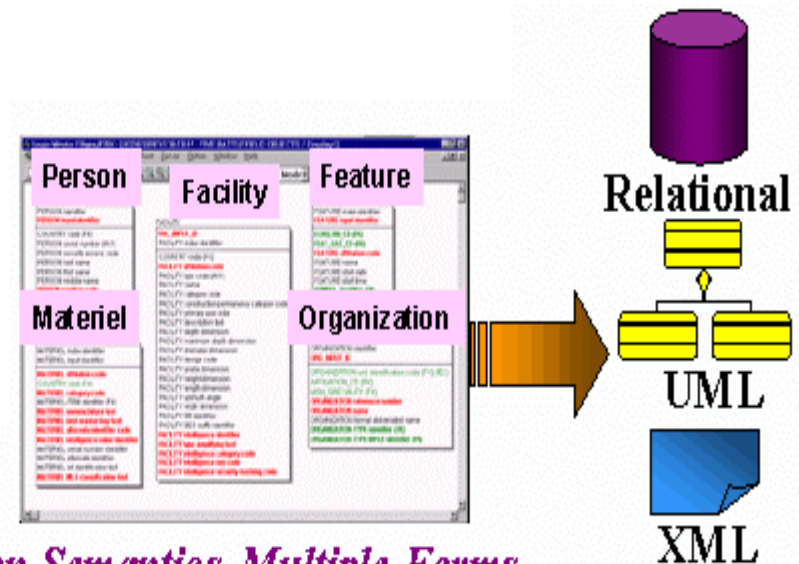
Common Semantics, Multiple Forms

Market-Driven Data Management Objective

- **Strategy for Data Resource "Accreditation" that Allows Network Components to be Independently Upgraded without Requiring Complete "System" re-baselining and re-accreditation**
  - Add/Remove Data Sources or Services
  - Add/Remove Communications Paths
  - Update Infrastructure Components
  - Add/Remove Applications

> **Define "Publish and Subscribe" Architecture for Data Resources**

New Data Source

New Warfighter Applications

New Comms Path

Upgraded Infrastructure (COTS & GOTS)
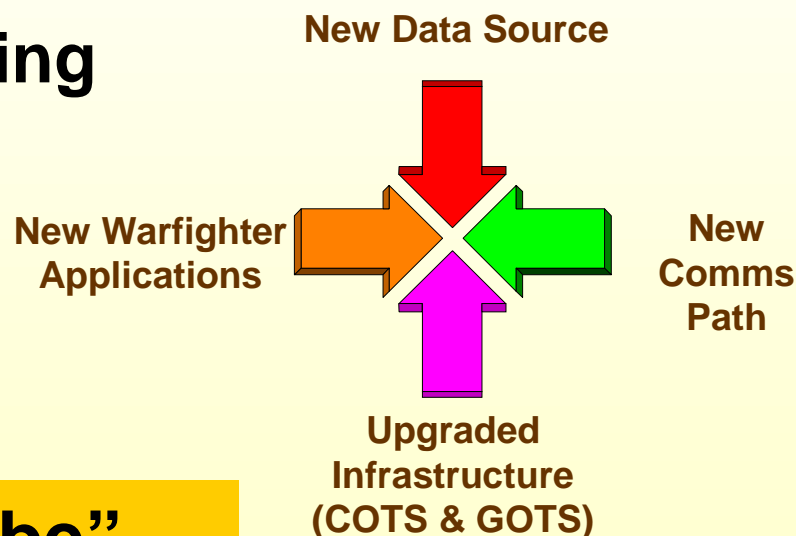
Market-Driven Data Management
New Info Systems Paradigm

➢ Publish and Subscribe Challenges

- How, where and what do data resources publish?
- How do users find resources and subscribe?
- How is data product or service delivery achieved?

• Context: Global Information Grid (GIG)

- Massively Networked Environment
- Many complex interconnections
- Numerous, frequently changing data resources
- Dynamic network architectures (e.g., crisis-specific)

Flexible and Responsive Management is Crucial!

Market-Driven Data Management GIG Electronic Marketplace

Consumers shall . . . easily discover, retrieve, and manage . . . information based upon its characteristics as advertised by producers . . .   Accordingly:

• Info producers shall advertise information availability and accessibility using DoD standard meta-data, data schema, and producer profiling mechanisms.

• Info awareness, access and delivery shall be facilitated . . . common mechanisms such as producer profiles and source registries.

• Authoritative info repositories shall be established, and organizations shall be identified and authorized to create, compile, distribute, and dispose of data and metadata in these repositories . . .

DoD CIO G&PM No. 7-8170-082400 Global Information Grid (GIG) Information Management (IM)

Market-Driven Data Management Build-Time Marketplace Rules

Data Component Registration

- Consult Emporium before creating new components and reuse existing data where practical
- Indicate planned use of components by formally subscribing to them
- Register additional components or recommended mods

Communities of Interest (COIs) Formation

- Created "as required" *when* someone will agree to manage
- Requirements for new COIs staffed with:
  - Existing COI Managers
  - Senior Service/Agency engineers
  - Flag Level Review Board

Market-Driven Data Management Run-Time Data Market Players

- **Data Producers**
  - "Advertise" their data products and services <u>and</u> convey access information

- **Data Consumers**
  - Use cataloged metadata to drive precision search and retrieval tools

- **Operational Data Managers**
  - Adjust network content in response to user demand as reflected in Run-Time Market transactions

- **Defense Acquisition Sponsors**
  - Use market metrics for acquisition oversight (e.g., reflecting *true* Program value thru specific data service usage data etc.)

- A Namespace is a technical mechanism that allows various, overlapping XML collections to be tagged with distinguishing labels.

- For DoD XML administrative purposes, Namespaces constitute a collection of data constructs that share a common context within a Community of Interest (COI).

  - These collections are managed by team leaders called "Namespace Managers" supported by "Namespace Working Groups."

  - The Configuration Review and Control Board (CRCB), established by DISA for the ASD/C3I, will charter Namespaces and designate managers for them.

Data Exchange Problem and XML

- XML by itself cannot resolve data engineering and application interoperability problems.

  - **The data problems that exist will continue to exist - and could become worse!**

- A serious problem lies in the semantics of the schemas.

  - **schemas are developed independently and are not semantically consistent with one another, therefore data cannot be consistently exchanged or interpreted "between schemas".**

  - **XML *encourages* the ad hoc development DTD's, thereby exacerbating the problem.**

Intelligent Use of XML

XML Tower of Babel

- XML is a great new technology breakthrough -- but not a silver bullet -- it will not solve all existing data problems.
- The *only* problem that XML contributes a solution to is the physical syntax of exchanged data, *i.e., marked-up ASCII text.*
  - **it contributes in that it is a globally standard approach for the exchange format of data between systems.**

- The DoD data problem is still there - it is a:
  - **human problem - interpersonal communication**
  - **schema problem - systems integration**
- Exchanging data successfully is not the same thing as application interoperability.
- Successful employment of XML-like technologies requires Business Process Re-Engineering as a Foundation

XML Repository Governance
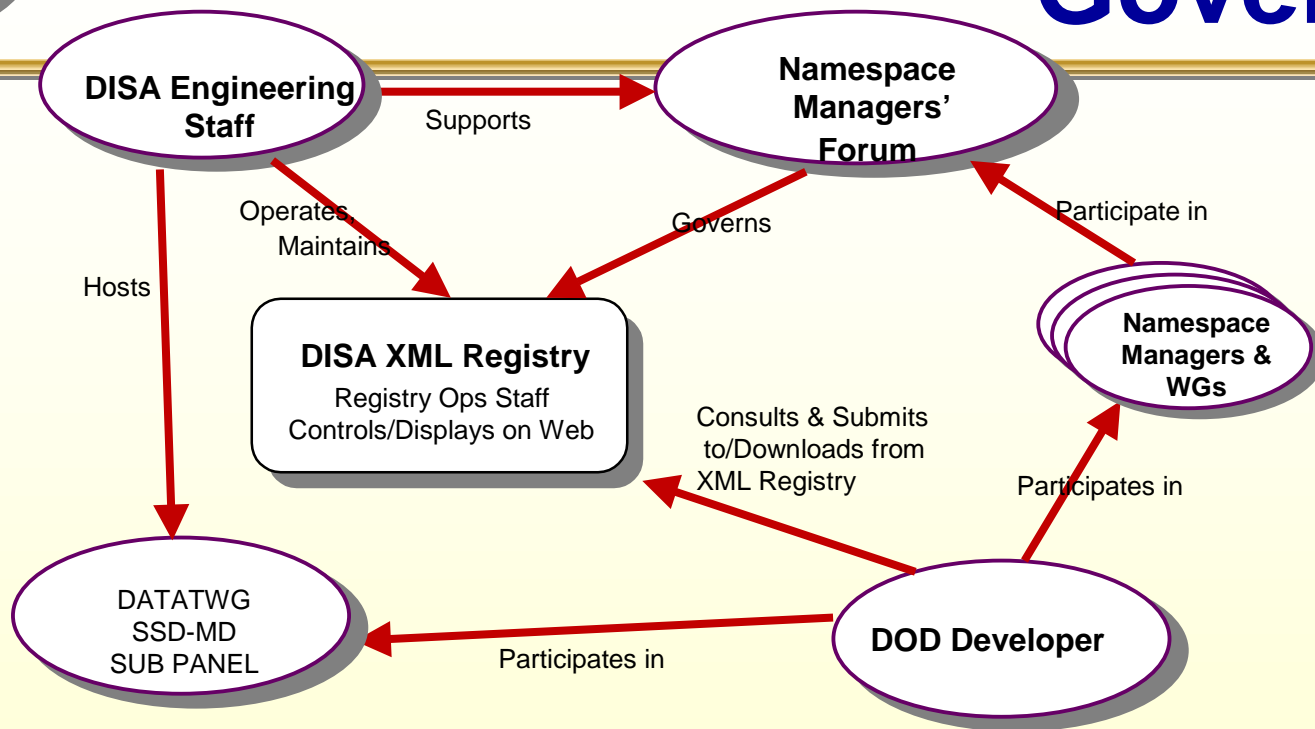
- **Management arrangements to accomplish registration and to perform the Clearinghouse function.**
- **Organizations and processes provide developers with a straightforward means to**
  - **Comply with the registration requirement**
  - **Acquire detailed XML technical information**
  - **Have a voice in formulating DoD XML directions**

John Zachman on Managing Change in the "**Knowledge-based**" **Information Age**

ENTERPRISE ARCHITECTURE - A FRAMEWORK

- The key to accommodating change in the knowledge-based, Information Age enterprise lies in
  - The "engineering" discipline for building and managing the enterprise models
  - The cultural discipline to employ the models [in an architecture framework] in the operation of the enterprise.
- Build models, store models, manage (enforce) models and change models [and use models in an architected knowledge framework]... the only rational Enterprise response to change in the "Information Age" [to gain competitive advantage.]

EIA Vendor Architectures

Process Automation

Enterprise Applications

- Custom Appl.
- Legacy Sys.
- Packaged Appl.
- Databases

Connector

Publish & Subscribe

Analyzer

Communicator
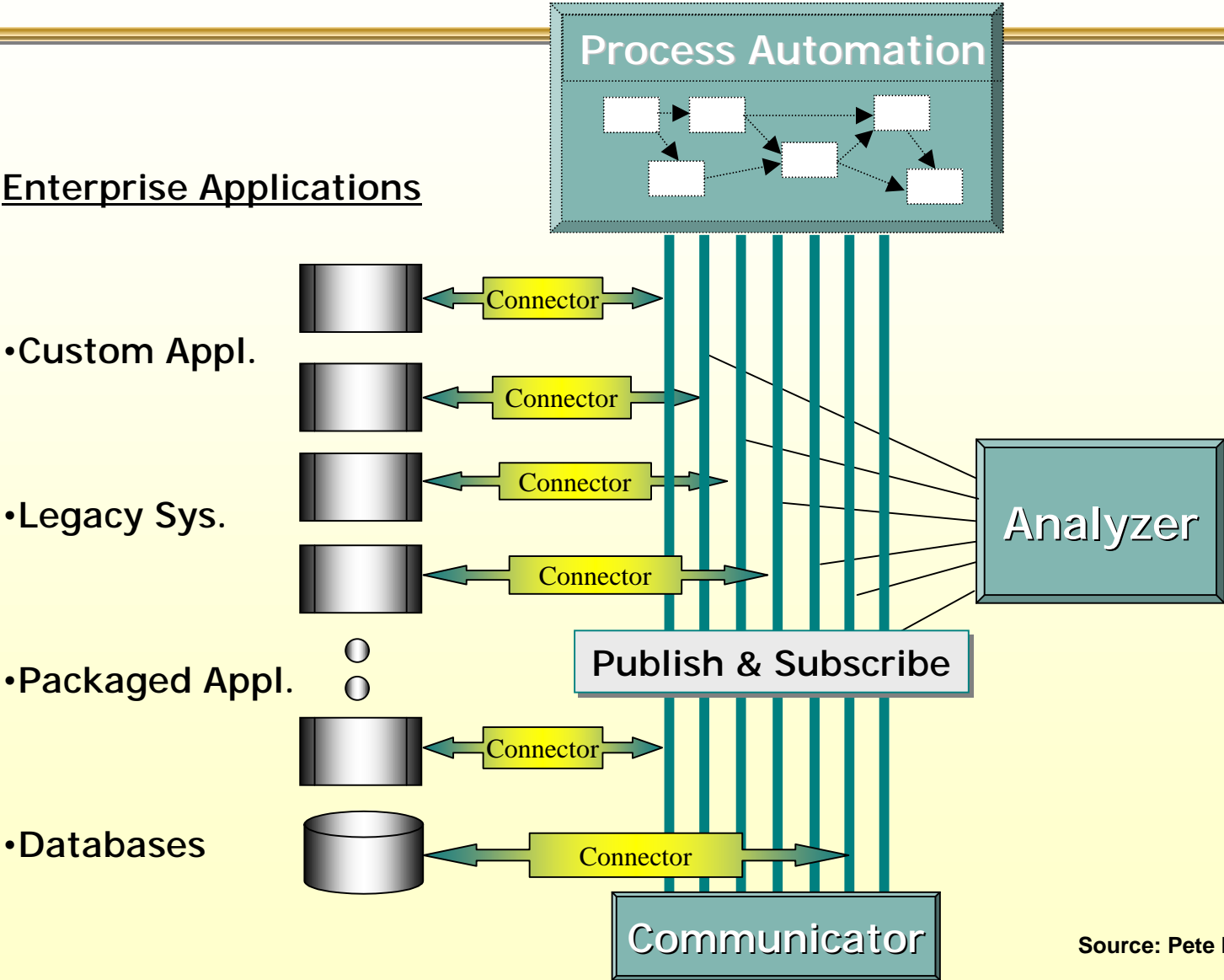
Source: Pete Everitt

1. Infrastructure
2. Formal data capture from existing records
3. Interoperability of data
4. Data authentication
5. Security

Information Integration

Database Level

Event Level

DataMart/DataWarehouse

?????

ERP Vendors

EAI Vendors

Lacks Validation of Legacy and Context Management from External Information Sources

Unified View

Cross-System Consistency
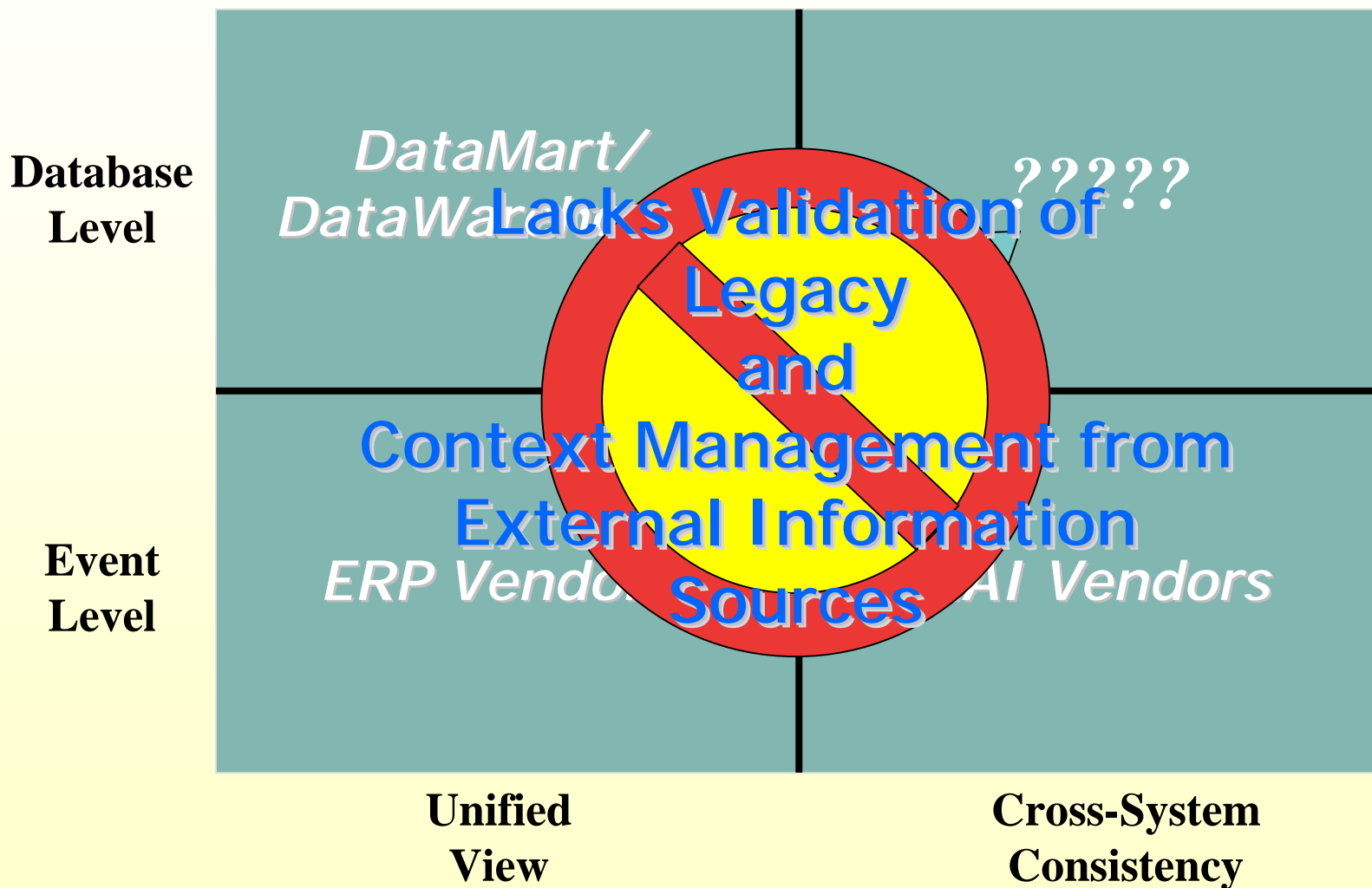
Adapted from: Michael Stonebreaker - EAI Journal

Source: Pete Everitt

Information Integration

Database Level

Event Level

DataMart/ DataWarehouse

ECI

ERP Vendors

EAI Vendors

Validated Info

Unified View

Cross-System Consistency

Adapted from: Michael Stonebreaker - EAI Journal

PDML Methodology

Source Data

- **Mil-Std 25**
- **JEDMICS**
- **TechOrder-4**
- **PDM Sys**
- **PDM Enablers**
- **PDM Schema**
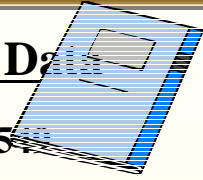
PDI Integration Schema in EXPRESS
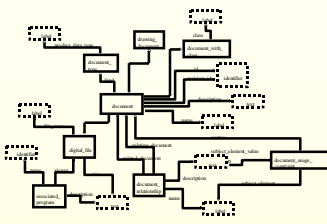
Mapping of Source Data to Advanced Integration Model

PDML Transaction Sets in EXPRESS

PDML Transaction Sets in XML

Product Structure 2549

Etc.

Analysis Tool → Auto Tools
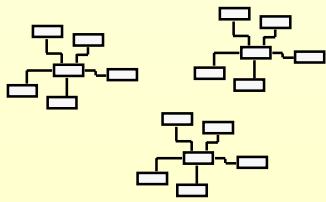
Semantic Vs Data Mediation

Data Mediation

Source System → Data Table → Target System
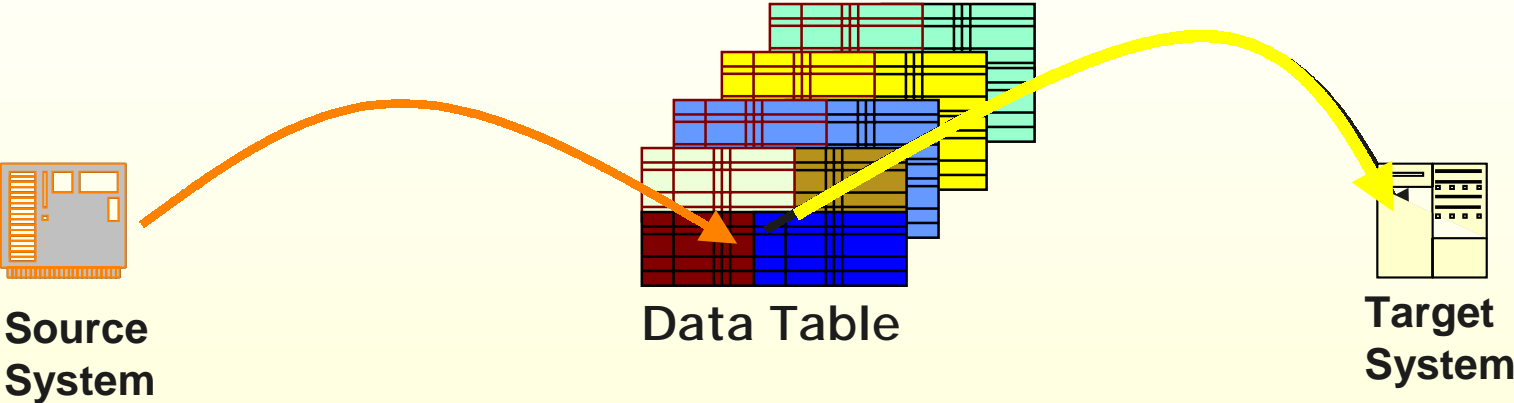
Semantic Mediation

Source System → Semantic Interpretation → Interoperability Engine → Context Transformation → Target System

Neutral Integration
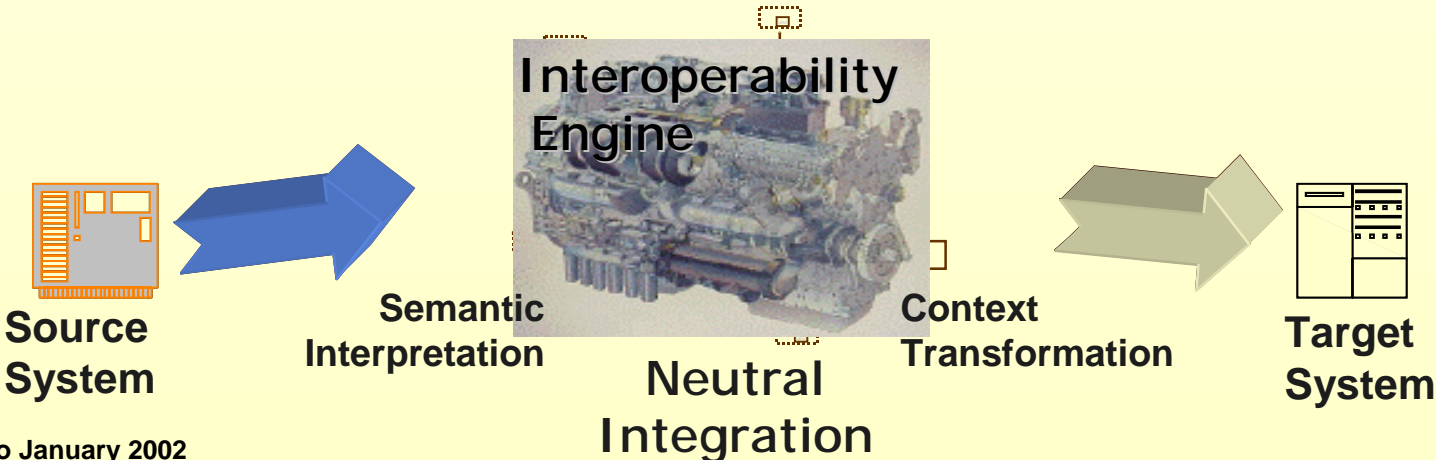
Semantic Mediation Provides...

...a discipline for semantic knowledge capture and integration from legacy sources

➢Compliant with XML Schema as recommended by W3C
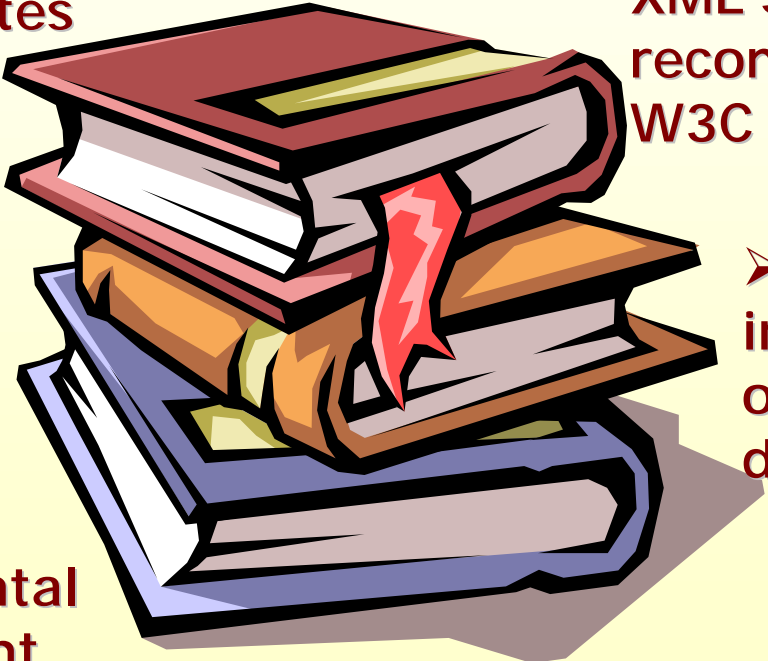
➢Captures and relates new knowledge to existing knowledge

➢Enables rapid start-up

➢Readily interoperates with other XML modeled data

➢Supports incremental process improvement

➢Supports human and machine interpretation

DAML
(DARPA Agent Markup Language)

The goal of the DAML effort is to develop a language and tools to facilitate the concept of the semantic web.

- **Create technologies that will enable software agents to dynamically identify and understand information sources**
- **Provide interoperability between the agents in a semantic manner**

What STEP is Not!

- An Enterprise Data Model
  - Would not be complete
  - Does not cover all necessary data

- A Standard for Databases
  - Does provide an interface (SDAI), and many definitions and associated data
  - BUT it is not possible nor appropriate to standardize internal schemas

- A Graphic Standard
  - Does include data to support graphics

Source: Eurostep, 1999

LESSONS TO BE LEARNED

- **CALS/EC**
- **Technology**
- **Standards**
- **DoD Data Program**
- **STEP**

- **Effective change comes from the individuals who must implement the change. It cannot be imposed from outside them.**

- **Change must be User and Requirements Driven.**

- **It is a long, slow process – be prepared.**

- **Form an 'Executive Team' for senior management buy-in, support and direct participation.**
  - **They identify Process Owners who will assume responsibility from end to end.**
  - **They formulate meaningful and ambitious goals, identify processes, decide difficult staffing and organizational issues, develop performance measures, guide implementations.**
  - **They maintain effective relationships, enlist involvement of all, and solicit feedback .**

Source: Dr Herve' LeBoeuf, et al

- **Need commercial buy-in – Tailored uniques only exacerbate legacy problems.**

- **Acquisition and Procurement Implementation**
  - **Acquisition must support all logistics business processes**
    - **Management**
    - **Supply**
    - **Training**
    - **Transportation**

- **Mandates from high don't work without associated funding and bottom-up buy-in and support.**

Source: Dr Herve' LeBoeuf, et al

Lessons to be Learned

- Modernization Depends on Legacy Data Availability
- Leverage from Already Existing Capabilities and Adapt from There
- Always Involve the User First
- Life Cycle Documentation/ Management is <u>Very</u> <u>Important</u>

- **Existing functional and DoD Component structures drive stove-pipe system solutions, impeding collaboration and a DoD-wide focus.**
- **Effective use of ERP/COTS applications must drive process change … else "paved cow paths".**
- **Beware of hidden Proprietary land mines**
- **Must overcome IT solution 'traditions', DoD unique needs, and "*not invented here*" syndrome.**
- **Change is contentious and painful.**
- **Must champion changes at highest levels of Component, Function and DoD enterprise.**
- **Collaboration by all increases each individual chance of success**
- **Take enterprise data out of the hands of the vendors and tool makers and give it back to the enterprise**

Source: Zach Goldstien, et al

XML is *GREAT*! <u>BUT</u> . . .

- **It is not magic.**
- **By itself it will not resolve all the data problems.**
- **Need to manage or else the tower of Babel may fall!**
- **Exchanging data successfully is not the same thing as application interoperability!**

XML Tower of Babel

SOLUTION: Reengineering Business Processes for proper use of models and registered XML components (schemas, DTDs, TAGs) are required to achieve "intelligent use of XML."

Intelligent Use of XML

Requires . . .

- **Proper use and implementation**
- **"Balanced approach" (not autocratic)**
- **Avoid fragmentation**
- **Consistent application**
- **Coordination**
- **Service and agency Buy-in and support**
- **Need to model our data and business requirements, then build XML components**
- **Must have agreement between "sender" and "receiver"**
- **Collaborative development**
- **Vocabulary re-use**
- **Education for intelligent use**

What Management Style will best work?
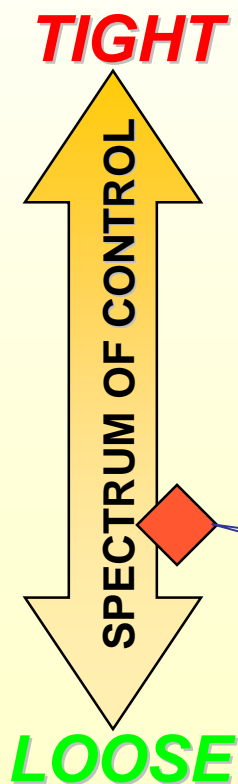
Top-down, "**Command**"

versus

Market-Driven

TIGHT

SPECTRUM OF CONTROL

LOOSE

Recommended Approach:
Market Driven with *Some* Controls

- **De jure vs Consortia**
- **Organizations don't actually compete:** each has a role, scope, and purpose
  - **Consortia best rapid for technology development**
  - **Formal de jure process best for consensus-building**
  - **but not vice versa!**
- **PLCS & SC4**
  - **Best of Both Worlds eg.**

Information Resources

- **"One issue pushing defense spending on information technology is that warfighters now see IT as another weapons system."**
  - **Data is the "ammo" of IT -- So how can IT applications function without its ammo?**

- **Information Is A Resource**
  - **Manage as a Critical Asset**

Source: Michael Kush, EDS Corp, quoted in http://www.washingtontechnology.com/news/14_13/federal/818-1.html

- **Document the form of the data in a manner independent of the transfer syntax**

- **Support multiple implementation approaches**

- **Support automation of implementation**

- **Include rules which the data must satisfy**

- **Support standard selections to meet specific requirements**

Source: Stefan Lindahl, Eurostep

- **Collaborative** and **concurrent** design and development leads to increased number of changes affecting a greater number of product components, people and systems

- **Changes need to be controlled across many systems, processes and geographic barriers**

- **The change process must be harmonized across applications andf egineering domains**

Source: Stefan Lindahl, Eurostep

Data, Meaning, & Communication

- Purpose of data
  - convey information to humans or software
  - data that are not intended for delivery to humans is ultimately intended to facilitate the delivery of information to humans

- Subtleties of human communication faculties not applied to data
  - Linguistics, philosophy, sociology all ignored
  - "Meaning" is what perceived data signifies to a stakeholder

- **Need to recognize these problems**

SUMMARY

- **True Interoperability is CRITICAL**
  - **Solving Data Issues Key to Success**
- **Many Significant Advances**
  - **But Still have Data Problems**
- **Need More Attention to Data Issues**

> *How to Preserve and Exchange the Meaning of Data Within Context ?*

Information – not Military Might –
Will Dominate Battlefields of 21st Century

- **Historically, the force that occupied the high ground had the greatest advantage. 'High Ground' now consists of *information* from satellites and aerial surveillance systems. -- Secretary Cohen**

KM is "obsoleting what you know before others obsolete it and profit by creating the challenges and opportunities others haven't even thought about" -- Dr. Yogesh Malhotra, *Inc. Technology*