

IMSC Inc.

Information and Business Fusion @ Work

Essential Role of Data Architecture in Business **Architecture & SOA**

William J. Mancuso
President & CTO
Information Management Solutions Consultants *Inc.*
&
SOA Lead BTA, DoD

Information Management Solutions
Consultants Inc. 703-635-0930

William J. Mancuso
William.Mancuso@IMSC.us
Lead SOA Architect for the BTA in the DoD
President IMSC Inc. 703-635-0930

Mr. Mancuso is the President of Information Management Solutions Consultants Inc. and a Internationally recognized expert in the field of Enterprise Architecture, Data Strategy, SOA and EIM. He is currently one of the Lead SOA architects for the BTA in DoD. He spent five years as the Chief Enterprise Architect for the Office of Deputy under Secretary of Defense. Mr Mancuso architected DoD Integrated Data Environment (IDE) and the RFID architecture for the DoD's Supply Chain. He spent 20 years plus in industry with a focus on integrating industry and technology to transform the way business processes are implemented and executed. Mr. Mancuso brings a strong commercial perspective to the federal government.

Topic

We will be discussing data architecture as it relates to business and not just IT with a focus on SOA. We will also be discussing why Architecture is important for a communicating between business and IT. Part of the discussion will be around Ontology and semantics as it relates to business. In the process we will look at SOA maturity models and how and where a Data Architecture and Business architecture come together to achieve SOA enterprise.

The "Enterprise"

United States Military Bases

www.nobases.org

Key to map symbols

- INSTALLATION TYPE
- Active US Base
 - Former US Base
 - ⊙ Forward operating location
 - Spybase
 - Echelon spybase
- COUNTRY STATUS
- Country hosting US base and troops
 - Country hosting US troops
 - Bases under negotiation
 - Country with no evident US military presence

Active US Personnel
United States and Territories*

Army	Navy	Marine	Air Force	Total
388,312	310,429	130,061	299,827	1,128,639

*No. bases: Active: 162; Former: 162

Active US Personnel
Europe

Army	Navy	Marine	Air Force	Total
67,768	11,667	918	33,700	114,053

Active US Personnel
Former Soviet Union

Army	Navy	Marine	Air Force	Total
24	4	84	17	149

Active US Personnel
East Asia and Pacific

Army	Navy	Marine	Air Force	Total
22,318	17,364	16,114	21,479	77,275

Active US Personnel
Sub-Saharan Africa

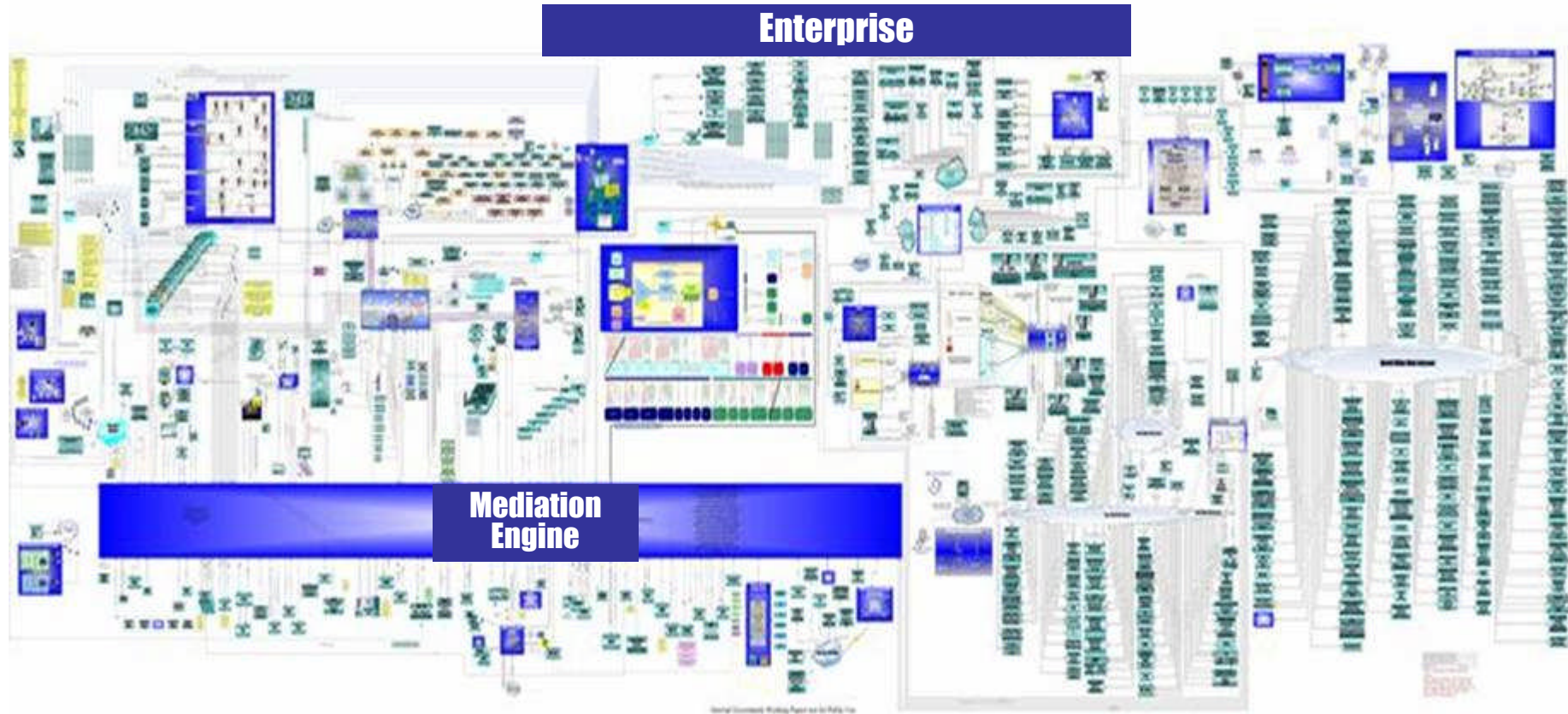
Army	Navy	Marine	Air Force	Total
63	5	193	21	282

Active US Personnel
Western Hemisphere

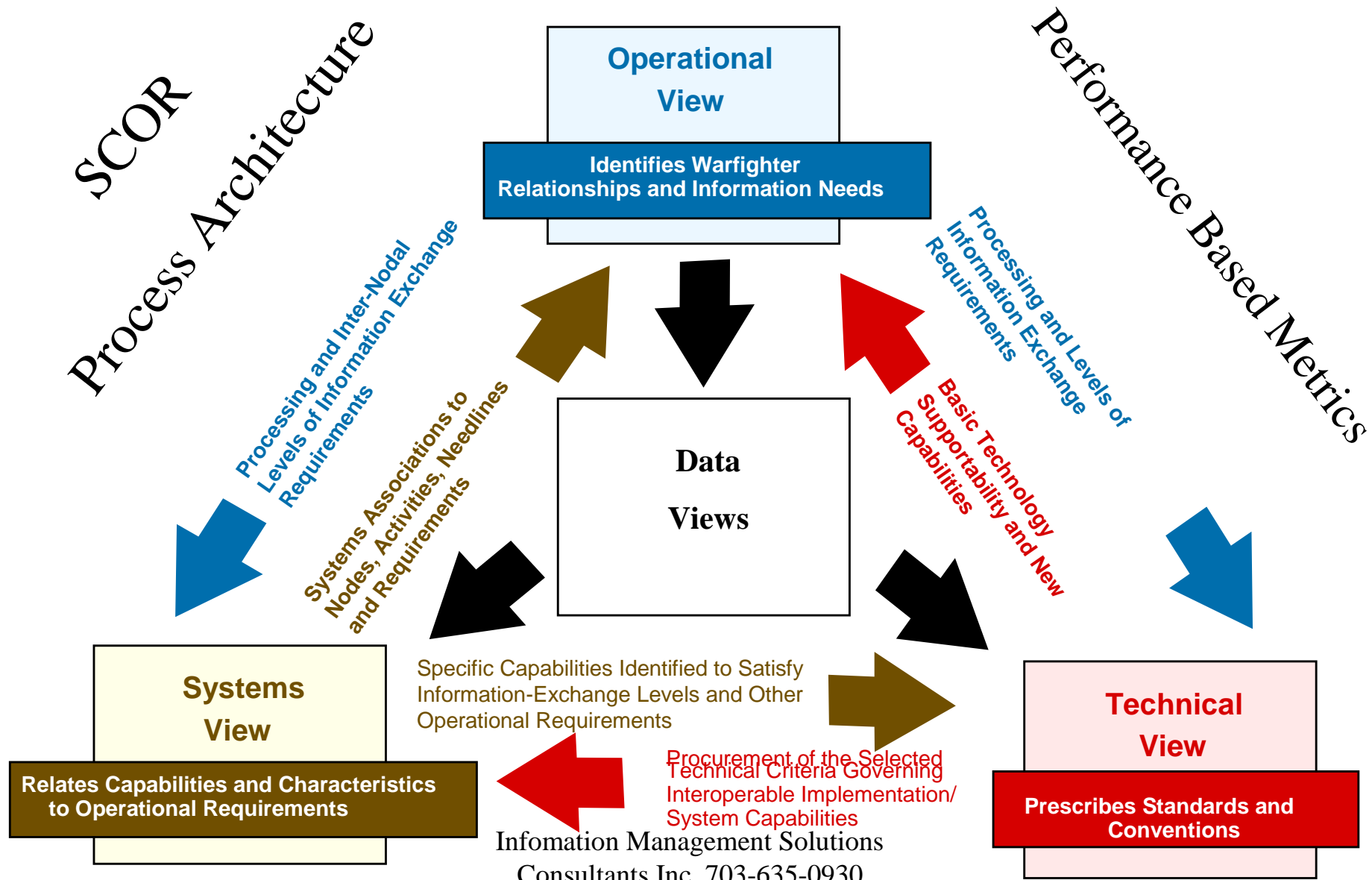
Army	Navy	Marine	Air Force	Total
383	644	433	363	1823

Sources and notes

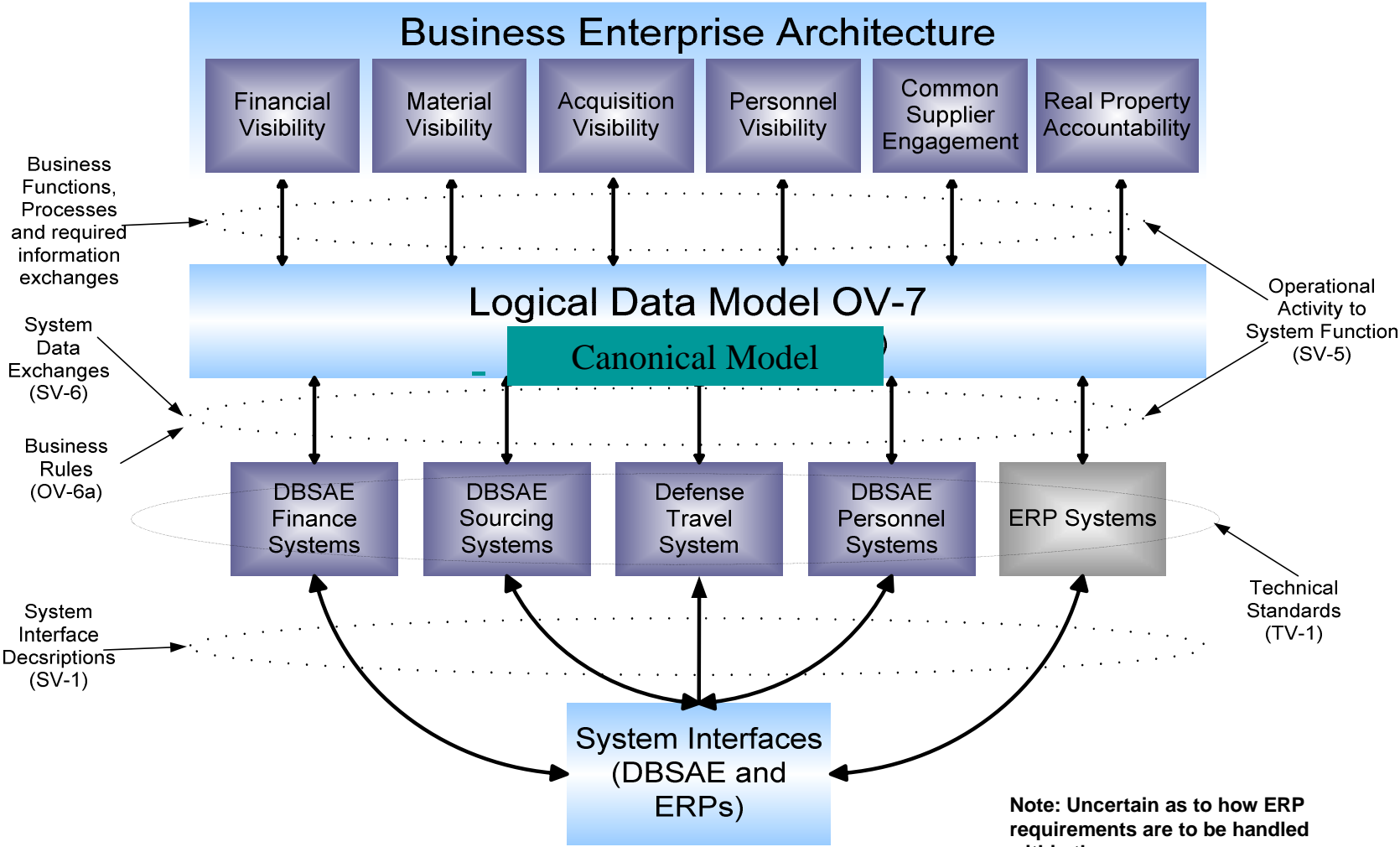
Enterprise Systems Interfaces and Transaction Flows



Logistics Enterprise Architecture Blueprint










Business Enterprise Architecture



SOA Maturity Models and Business Architecture

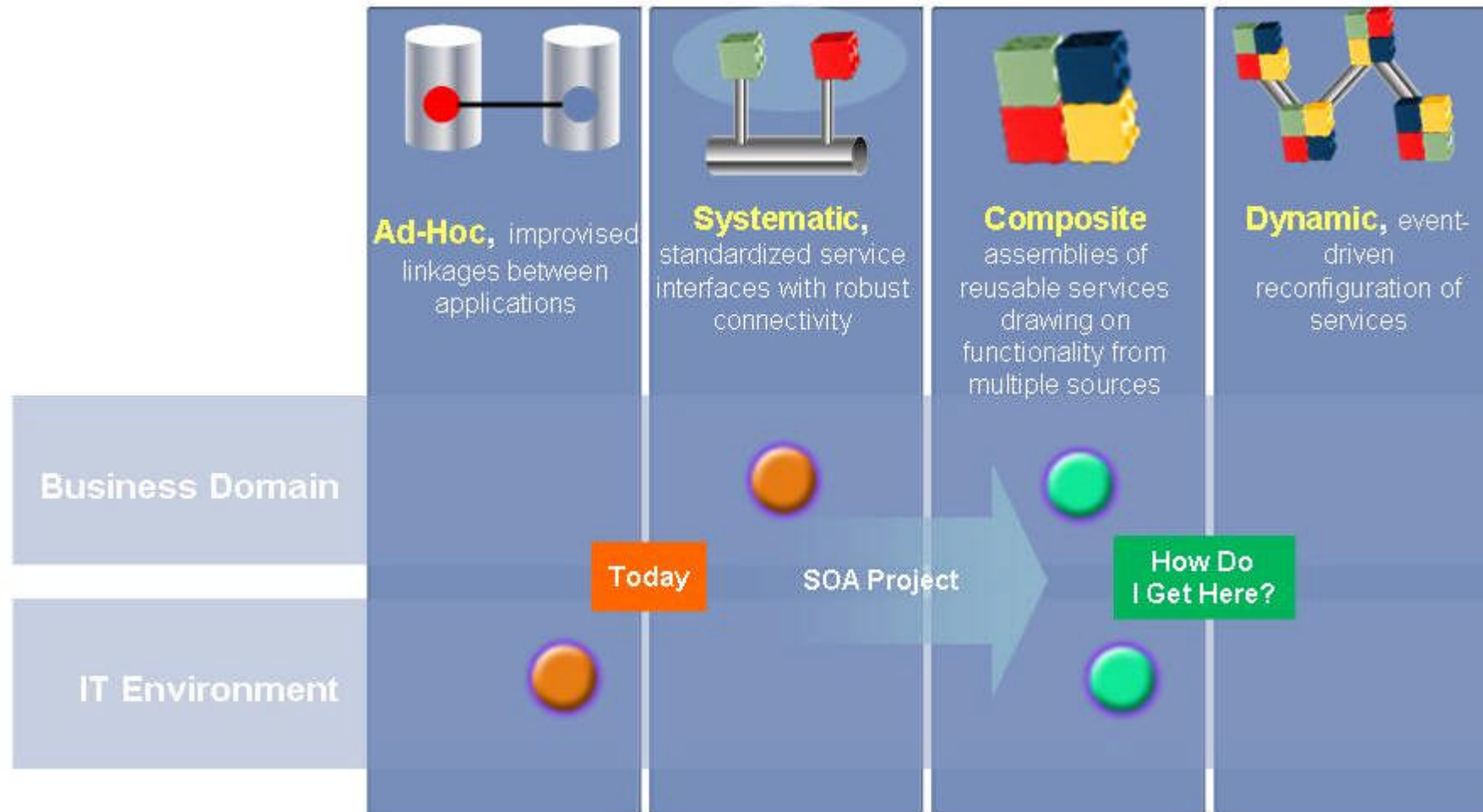
Service Integration Maturity Model

	 Silo	 Integrated	 Componentized	 Services	 Composite Services	 Virtualized Services	 Dynamically Re-Configurable Services
Business	Function Oriented	Function Oriented	Function Oriented	Service Oriented	Service Oriented	Service Oriented	Service Oriented
Organization	Ad hoc IT Governance	Ad hoc IT Governance	Ad hoc IT Governance	Emerging SOA Governance	SOA and IT Governance Alignment	SOA and IT Governance Alignment	SOA and IT Governance Alignment
Methods	Structured Analysis & Design	Object Oriented Modeling	Component Based Development	Service Oriented Modeling	Service Oriented Modeling	Service Oriented Modeling	Grammar Oriented Modeling
Applications	Modules	Objects	Components	Services	Process Integration via Services	Process Integration via Services	Dynamic Application Assembly
Architecture	Monolithic Architecture	Layered Architecture	Component Architecture	Emerging SOA	SOA	Grid Enabled SOA	Dynamically Re-Configurable Architecture
Information	Application Specific	Subject Areas	Canonical Models	Canonical Models	Enterprise Business Data Dictionary	Semantic Data Vocabularies	Semantic Data Vocabularies
Infrastructure	Platform Specific	Platform Specific	Platform Specific	Platform Specific	Platform Specific	Platform Neutral	Dynamic Sense & Respond
	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7

Information Management Solutions
Consultants Inc. 703-635-0930

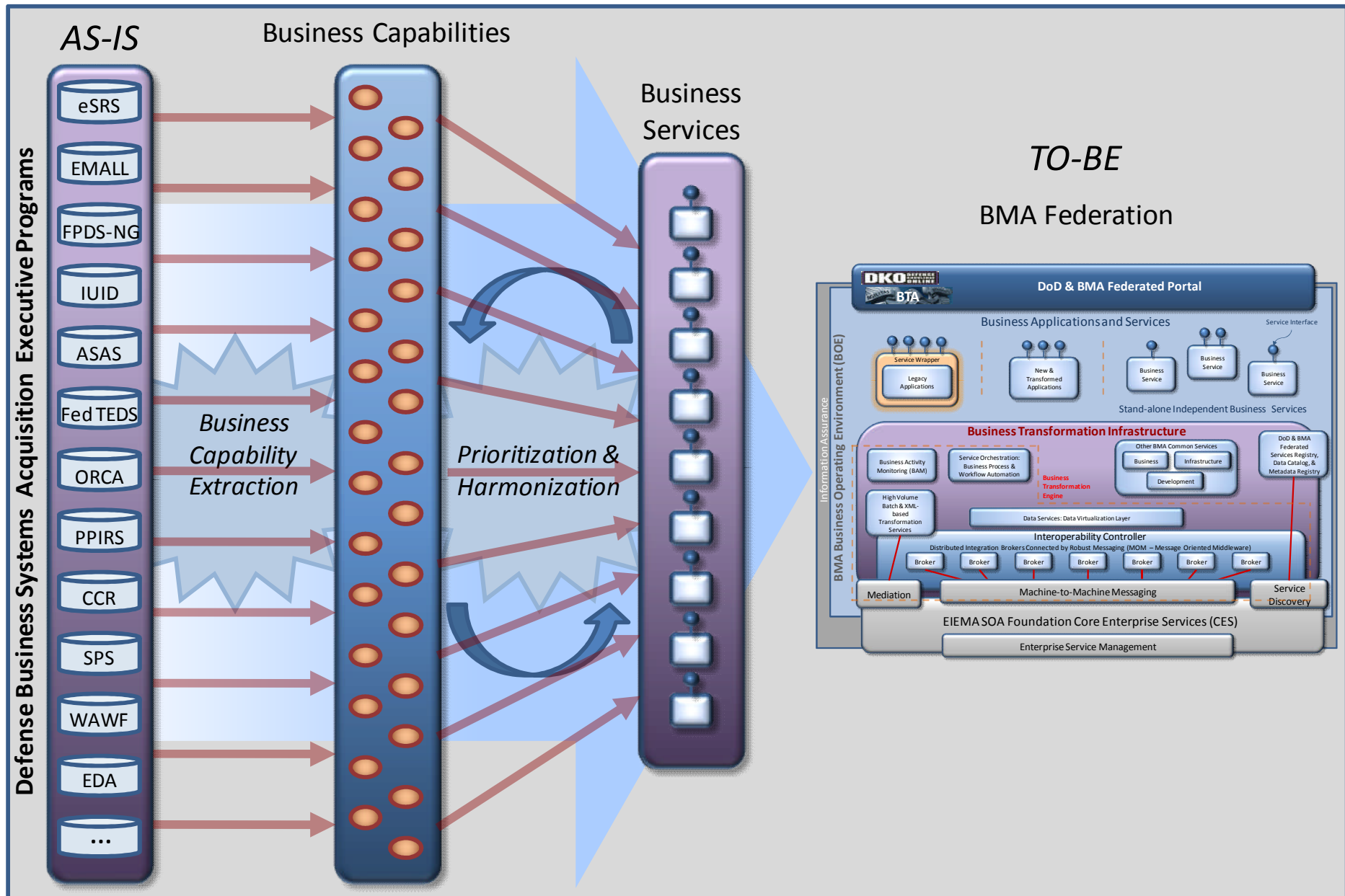
Source: IBM "SIMM" 10

Where We are in BMA SOA and Where We Want to Go

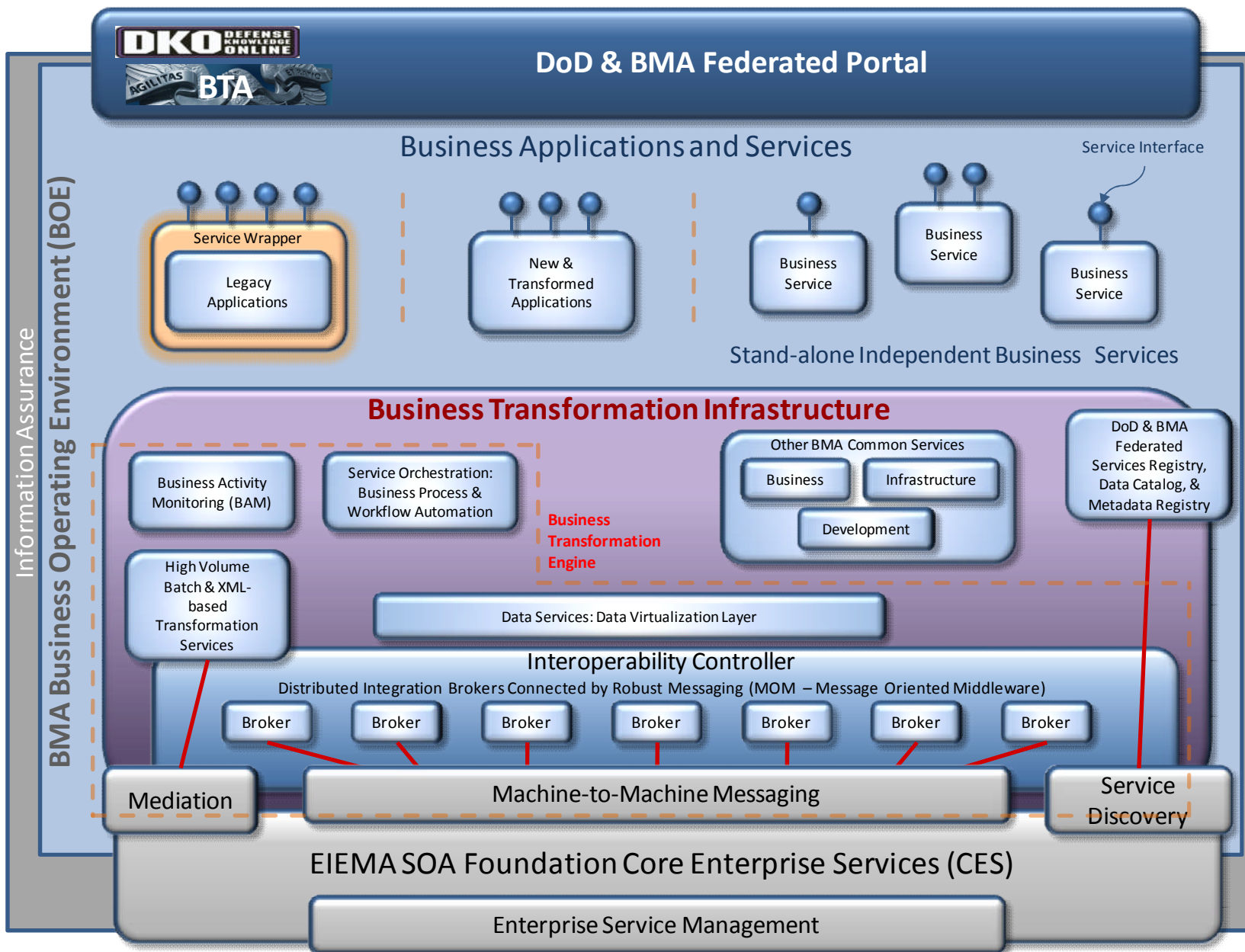


Source: What's New!

BMA Roadmap to SOA, Top-down approach



Systems Federation Via the Business Transformation Engine Within the BOE





SOA Platform Challenge

Web Services Interoperability: Profiles



Establishing guidance for Web Services, expressed in profiles on the use and applications of standards—specifies and constrains their use, helps to assure interoperability of autonomous systems and applications across the network.

Web Services Management

- Web Services Management Framework (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)
- Web Services Management and Policy (WS-M) (WS-I)

Security Specifications

- WS-Security (WS-I)
- WS-Security: Username Token Profile (WS-I)
- WS-Security: SOAP Message Security (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)
- WS-Security: Policy (WS-I)

Resource Specifications

- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)
- WS-Resource Framework (WS-I)

Metadata Specifications

- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)
- WS-Metadata Exchange (WS-I)

Transaction Specifications

- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)
- WS-Transaction (WS-I)

Business Process Definition, Execution and Management

- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)
- Business Process Definition Language (BPEL) (WS-I)

XML Specifications

- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)
- XML Schema (WS-I)

SOAP

- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)
- SOAP (WS-I)

Web Foundation Standards

- HyperText Markup Language (HTML) (WS-I)
- HyperText Transfer Protocol (HTTP) (WS-I)
- Web Services for Remote Portals (WSRP) (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)
- Department of Defense Metadata Registry (WS-I)

Internet

- Internet Protocol version 6 (IPv6) (WS-I)
- Cryptographic Message Syntax (CMS) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)
- Secure Multipurpose Internet Mail Extensions (S-MIME) (WS-I)

Messaging Specifications

- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)
- WS-Addressing (WS-I)

Message Format / Content

- Business Object Document (BOD) (WS-I)
- Web Document Archiving & Wrapping (WDAW) (WS-I)
- WS-Data-Correlation (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)
- Universal Business Language (UBL) (WS-I)

Strategic and Tactical Data Modeling

What is a Vocabulary?

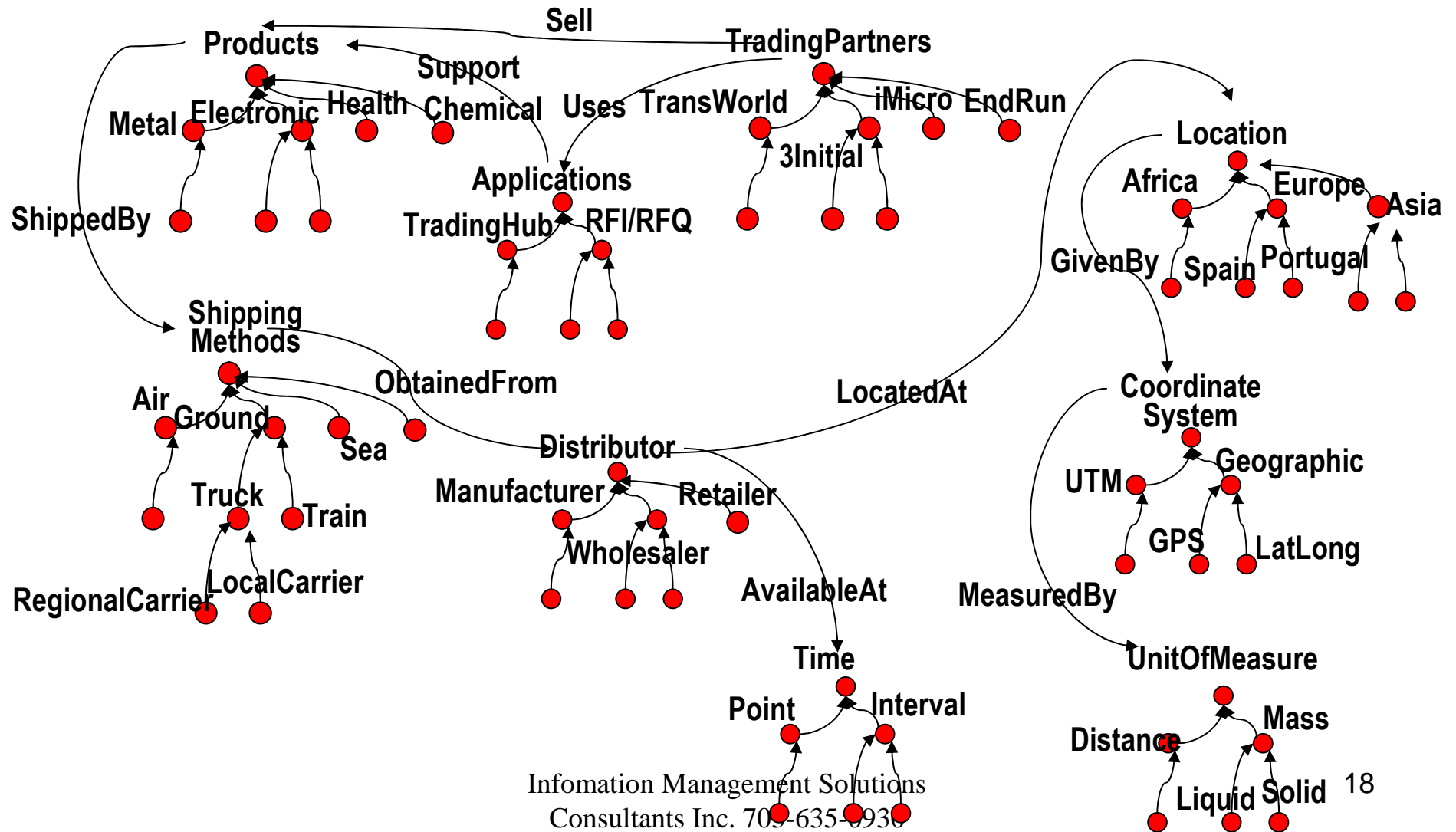
- **All terms used, developed, or understood by, a particular person or group for a particular domain or purpose.**
- **Domain lexicons, glossaries, subject headings, metadata element sets, topic maps, taxonomies, thesauri, and ontologies are all types of vocabularies.**
- **Vocabularies are often specific to domains and communities of interest**

Canonical Data Model

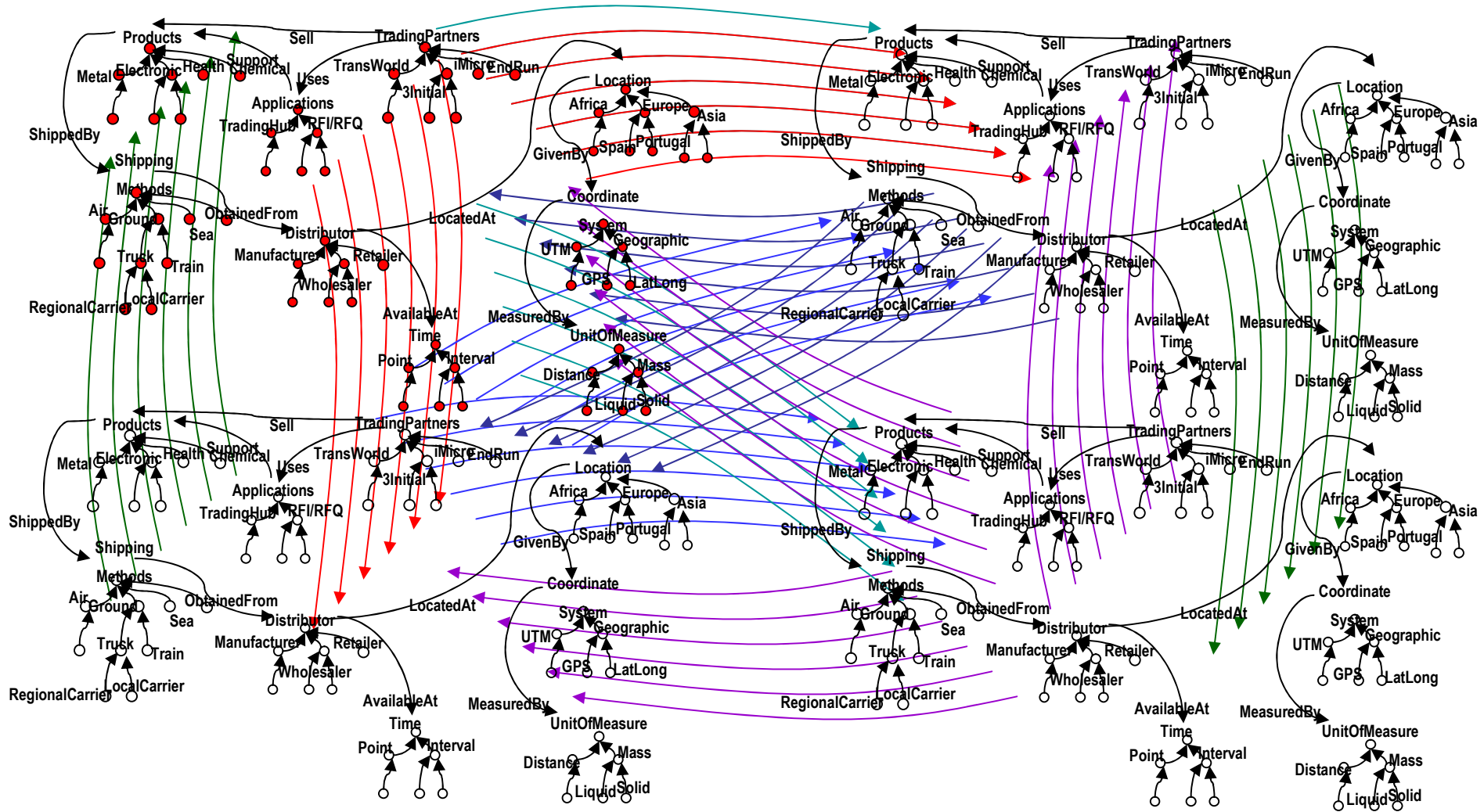
“A Canonical Data Model defines message formats that are independent from any specific application so that all applications can communicate with each other in this common format. If the internal format of an application changes, only the *message translator* between the affected application and the common *message channel* has to change, while all other applications and *message translators* remain unaffected.” – ***Enterprise Integration Patterns, Gregor Hohpe, Bobby Woolf***

“A data model that represents the inherent structure of data without regard to either individual use or hardware or software implementation.” – ***Vertaasis Inc.***

Generic Electronic Commerce Example:



Now Assume Each Company Has Separate Enterprise Semantics, Multiply by the Number of Companies, & Have Them Interoperate and Preserve Semantics – Look Familiar?

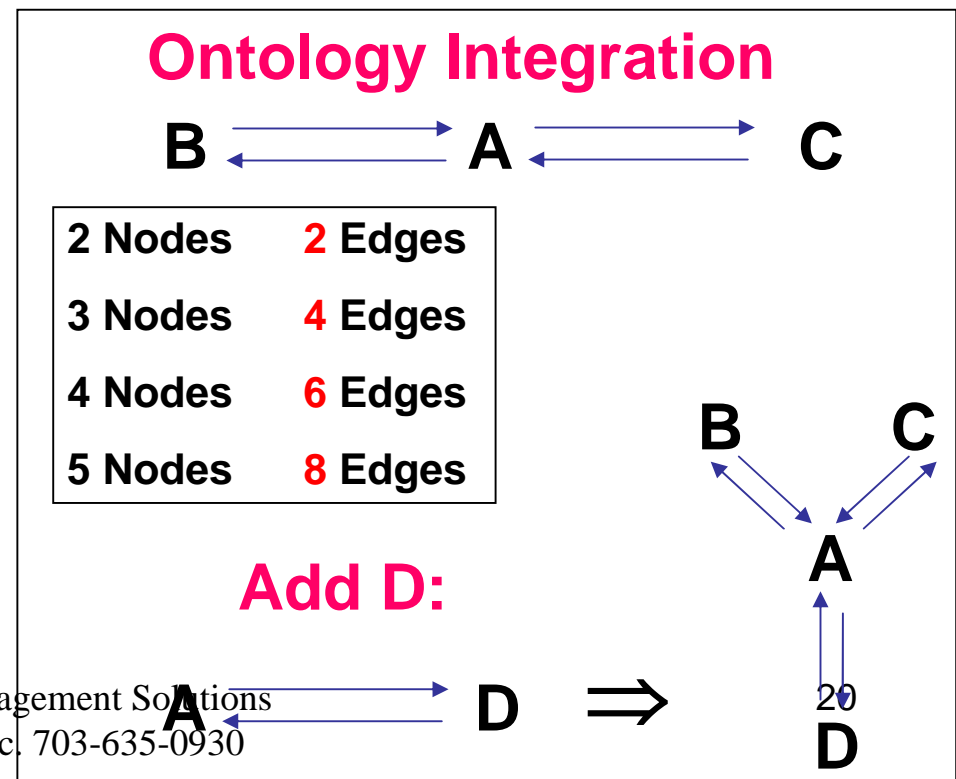
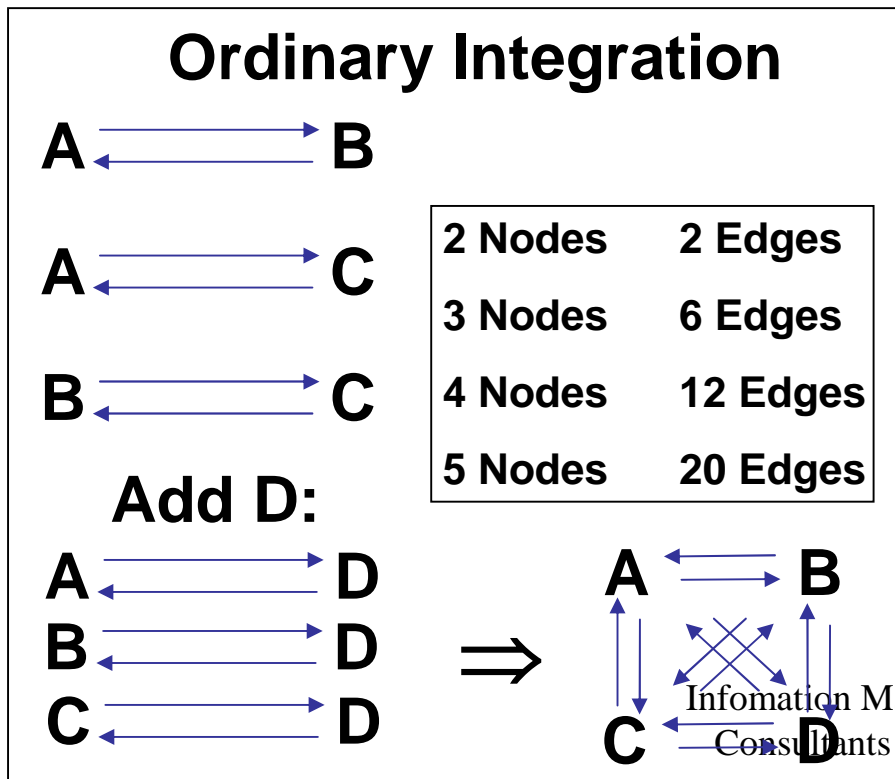


Try doing this without Ontologies! You can, but it's a Nightmare, and it COSTS: Now & Later!

Semantic Issues: Complexity

- An ontology allows for near linear semantic integration actually $2n-2$ rather than $n^2 - n$ integration
 - Each application/database maps to the "lingua franca" of the ontology, rather than to each other

Dr. Leo Obrst



Canonical Data Model Life-cycle

Bootstrapping

The screenshot displays the Rational Data Architect interface for a project named "Data - WAWFtoEDALexicalMapping.msl". The main workspace is divided into three vertical panes: "Source", "Mappings", and "Target".

- Source Pane:** Lists various database fields from the source system, including:
 - PAY_REJECTED_IN [CHAR(1) Nullable]
 - PAY_REJECTION_DT [DATE Nullable]
 - PAY_WORKED_USER_ID [VARCHAR2(30) Nullable]
 - PAYMENT_WORKED_IN [CHAR(1) Nullable]
 - PAY_WORKED_DT [DATE Nullable]
 - GROSS_WT [NUMBER(6, 0) Nullable]
 - VOID_DT [DATE Nullable]
 - WEIGHT_MEASURE_CD [VARCHAR2(5) Nullable]
 - INVOICE_RCV_DT [DATE Nullable]
 - VOID_IN [CHAR(1) Nullable]
 - STATUS_DS [VARCHAR2(25) FK]
 - ARP_CHECK_IN [CHAR(1) Nullable]
 - COC_CHECK_IN [CHAR(1) Nullable]
 - ESTIMATED_DT_IN [CHAR(1) Nullable]
 - DOC_HOLD_IN [CHAR(1) Nullable]
 - LPO_CERTIFIED_IN [CHAR(1) Nullable]
 - LPO_USER_ID [VARCHAR2(30) Nullable]
 - LPO_REJECTED_IN [CHAR(1) Nullable]
 - LPO_CERTIFIED_DT [DATE Nullable]
 - LPO_REJECTED_DT [DATE Nullable]
 - LPO_HOLD_USER_ID [VARCHAR2(30) Nullable]
 - LPO_HOLD_IN [CHAR(1) Nullable]
 - LPO_HOLD_DT [DATE Nullable]
 - GBL_REF_NR [VARCHAR2(30) Nullable]
 - TCN_NR [VARCHAR2(30) Nullable]
 - SHIPMENT_NO [VARCHAR2(25) Nullable]
 - SHIPMENT_DT [DATE Nullable]
 - CREATED_BY_CD [VARCHAR2(2) Nullable FK]
 - PV_POINT_OF_CONTACT_NM [VARCHAR2(80) Nullable]
 - INPUT_TYPE_CD [CHAR(4) Nullable FK]
 - REJECTED_BACK_TO_CD [CHAR(1) Nullable FK]
 - COMBO_IN [CHAR(1) Nullable]
 - FORM_TYPE_CD [VARCHAR2(5) Nullable]
 - LAST_USER_ACTIVE_ID [VARCHAR2(30) Nullable]
 - SMALL_BUSINESS_IN [CHAR(1) Nullable]
 - FMS_IN [CHAR(1) Nullable]
 - VENDOR_INVOICE_NO [VARCHAR2(30) Nullable]
 - TASK_ORDER_NO [VARCHAR2(30) Nullable]
 - SOURCE_INSPECT_IN [CHAR(1) Nullable]
 - PBP_LAST_COM_DT [DATE Nullable]
 - PBP_ENCUMBRANCE_DT [DATE Nullable]
 - PBP_LAST_FIN_STMT_DT [DATE Nullable]
 - DIRECT_BILL_CD [VARCHAR2(2) Nullable]
 - GOVT_APPROVED_AM [NUMBER(17, 2) Nullable]
 - INV_RR_DISCOUNTS_FLAG [CHAR(1) Nullable]
 - STATUS_DT [DATE]
 - LAST_UPDATE_DT_BKUP [DATE Nullable]
 - RELATED_SEQ_NO [NUMBER Nullable]
 - DOC_CERTIFIED_MFT NUMBER(15, 2) Nullable
- Mappings Pane:** A central area showing a dense network of yellow lines that connect fields from the Source pane to fields in the Target pane, representing the data mapping logic.
- Target Pane:** Lists fields from the target system, including:
 - SHIPMENT_NUMBER [VARCHAR2(8) Nullable]
 - MODIFICATION_NUMBER [VARCHAR2(6) Nullable]
 - BATCH_NUMBER [VARCHAR2(6) Nullable]
 - BRANCH_OF_SERVICE [VARCHAR2(10) Nullable]
 - CAGE_CODE [VARCHAR2(5) Nullable]
 - INVOICE_AMOUNT [VARCHAR2(16) Nullable]
 - PAYMENT_AMOUNT [VARCHAR2(13) Nullable]
 - ADPE_NUMBER [VARCHAR2(6) Nullable]
 - ISSUE_DODAAC [VARCHAR2(6) Nullable]
 - ADMIN_DODAAC [VARCHAR2(6) Nullable]
 - LOCATION [VARCHAR2(256) Nullable]
 - DOC_SIZE [VARCHAR2(6) Nullable]
 - MESSAGE [VARCHAR2(80) Nullable]
 - TIMESTAMP [DATE Nullable]
 - DFAS_DD 1716_LOAD
 - FUNCTION [CHAR(1) Nullable]
 - FORM_NAME [VARCHAR2(8) Nullable]
 - INDEX_DATE [DATE Nullable]
 - DIVISION [VARCHAR2(10) Nullable]
 - CONTRACT [VARCHAR2(19) Nullable]
 - DELIVERY_ORDER [VARCHAR2(19) Nullable]
 - SHIPMENT_NUMBER [VARCHAR2(8) Nullable]
 - MODIFICATION_NUMBER [VARCHAR2(6) Nullable]
 - BATCH_NUMBER [VARCHAR2(6) Nullable]
 - BRANCH_OF_SERVICE [VARCHAR2(10) Nullable]
 - CAGE_CODE [VARCHAR2(5) Nullable]
 - INVOICE_AMOUNT [VARCHAR2(16) Nullable]
 - PAYMENT_AMOUNT [VARCHAR2(13) Nullable]
 - ADPE_NUMBER [VARCHAR2(6) Nullable]
 - ISSUE_DODAAC [VARCHAR2(6) Nullable]
 - ADMIN_DODAAC [VARCHAR2(6) Nullable]
 - LOCATION [VARCHAR2(256) Nullable]
 - DOC_SIZE [VARCHAR2(6) Nullable]
 - DFAS_DD 577
 - RECORD_KEY [RAW(16)]
 - INDEX_DATE [DATE Nullable]
 - SYSTEM_NAME [VARCHAR2(8) Nullable]
 - LAST_NAME [VARCHAR2(30) Nullable]
 - FIRST_NAME [VARCHAR2(25) Nullable]
 - MIDDLE_INITIAL [VARCHAR2(1) Nullable]
 - CITY [VARCHAR2(25) Nullable]
 - STATE [VARCHAR2(3) Nullable]
 - ORGANIZATION [VARCHAR2(50) Nullable]
 - ORIG_PROTOCOL [VARCHAR2(5) Nullable]
 - ORIG_SERVER [VARCHAR2(50) Nullable]
 - ORIG_PORT [VARCHAR2(6) Nullable]
 - LOCATION [VARCHAR2(256)]
 - SERIALIZED_LOCATION [VARCHAR2(60) Nullable]

The bottom of the image shows the Windows taskbar with the system clock at 11:15 PM on Thursday, 5/3/2007, and a taskbar containing several open applications like "Re: Status - IBM Lotu...", "Data - Rational Data ...", "Yahoo! Mail - rockgop...", "CNN.com - Breaking ...", and "AT&T Network Client ...".

Bootstrapping Ontologies

- **Step 1: Start at the bottom**
 - Build vocabularies from existing physical systems
 - Finds and uses the terminology that's important
- **Step 2: Collaborate**
 - The community can document, review, discuss and change
 - Human-readable documentation and formal vocabulary definition
- **Step 3: Share and Use**
 - People access the vocabularies through web browsers to view the natural language documentation and navigate formal relationships
 - Machines can download vocabularies and use

Step 1: Start at the Bottom

- **Databases and message structures don't keep semantics**
 - What information do we have?
 - What does it mean to the enterprise?
 - How does all this information align?
- **Find and use the semantics!**
 - Combine the terms used with knowledge bases to discover and assign semantics to information
 - Store the terms, definitions and semantics in vocabularies
 - Use semantics to align (match) information

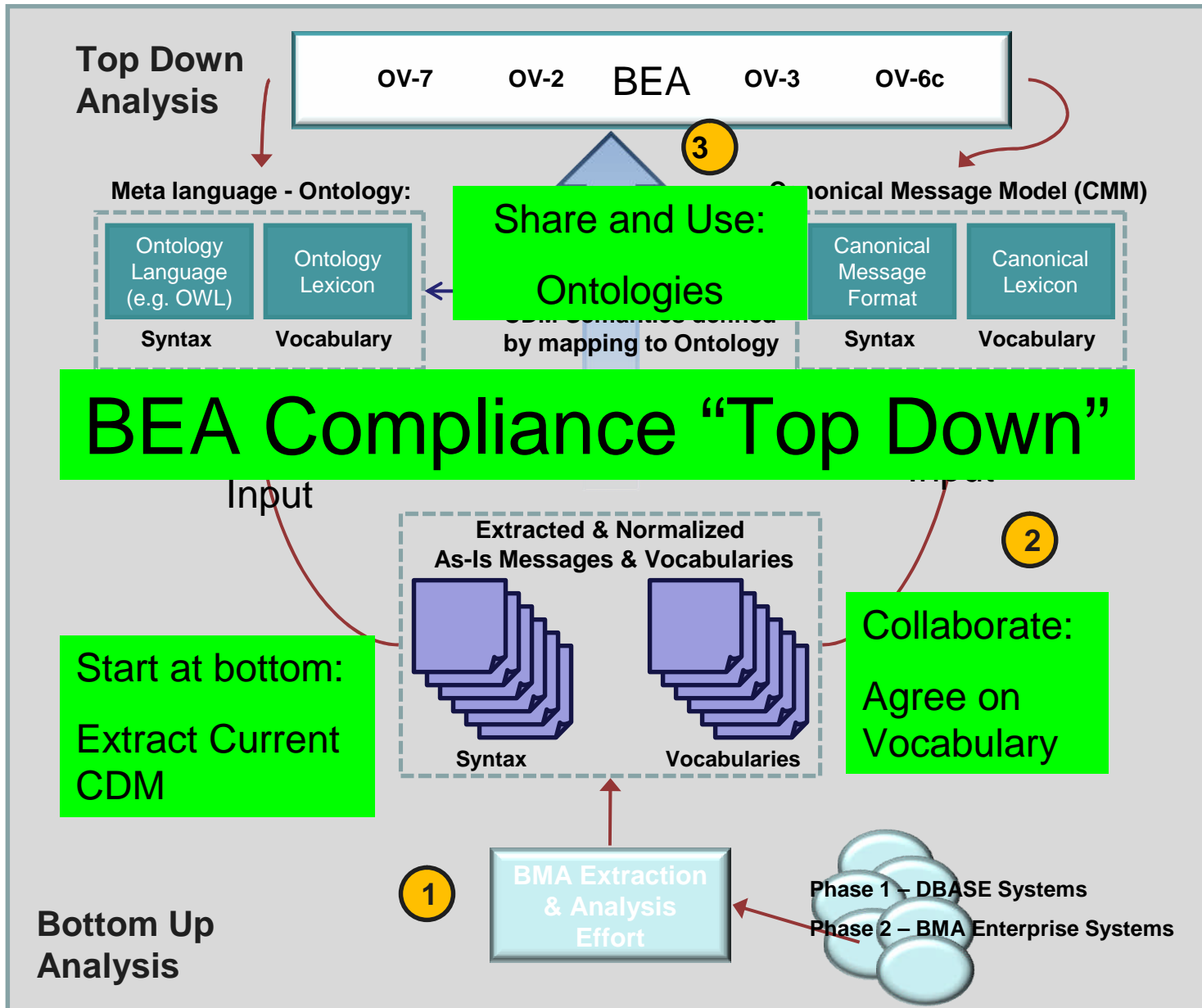
Step 2: Collaborate

- **Creating vocabularies is naturally collaborative**
 - **Tasks: harvest, identify, define, disambiguate, document, standardize, edit, visualize, review, audit and publish**
 - **Involve the right people**
 - **Reuse other vocabularies: benefit from the experts**
- **Community-oriented**
 - **A community consists of members that share experience, expertise and interest in a particular domain**
 - **Communities manage memberships, content, and access privileges**
- **Semantic Wiki**
 - **A wiki is a website that can be edited by users through their browsers**
 - **Captures the efforts of many over time**
 - **Adds semantic richness to wiki markup language**

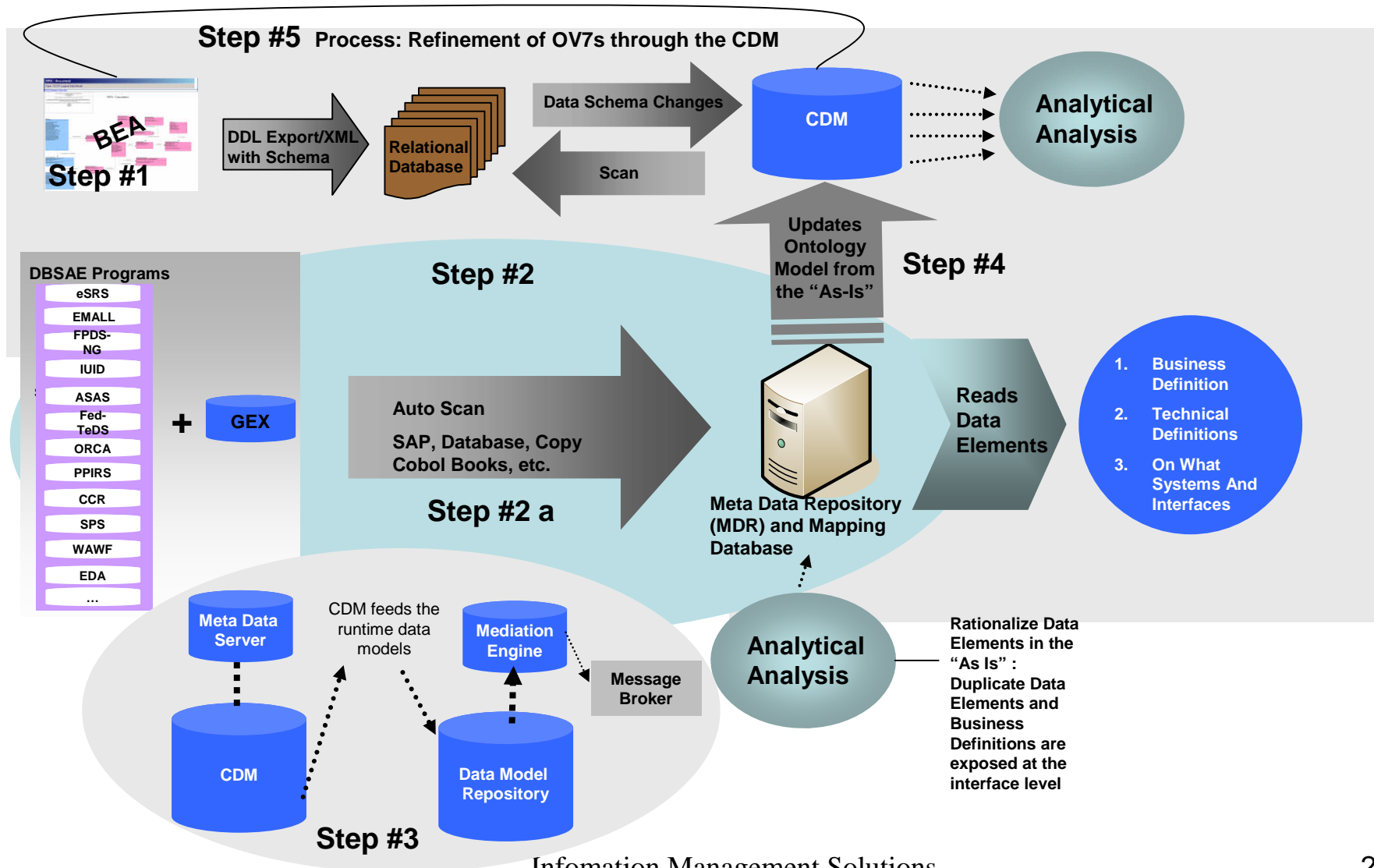
Step 3: Share and Use

- **Machines use ontologies**
 - The vocabularies are represented with formalism that are rich and precise enough for software
 - Vocabularies can be downloaded as OWL ontologies
- **People use natural language**
 - (Most) People don't understand XML, OWL, RDF, or even HTML
 - People understand text, images, tables, charts, links
 - Follow existing web paradigms that people are comfortable with (browsers, links, pages, addresses, search, discussions, etc.)
- **Keep the two parts together**
 - People have to understand the vocabulary to maintain and use it
 - If parts are kept separate, more difficult to diverge
 - It's simply easier this way! (Manually aligning documentation with models is too much work)

Interoperability through Canonical Data Model

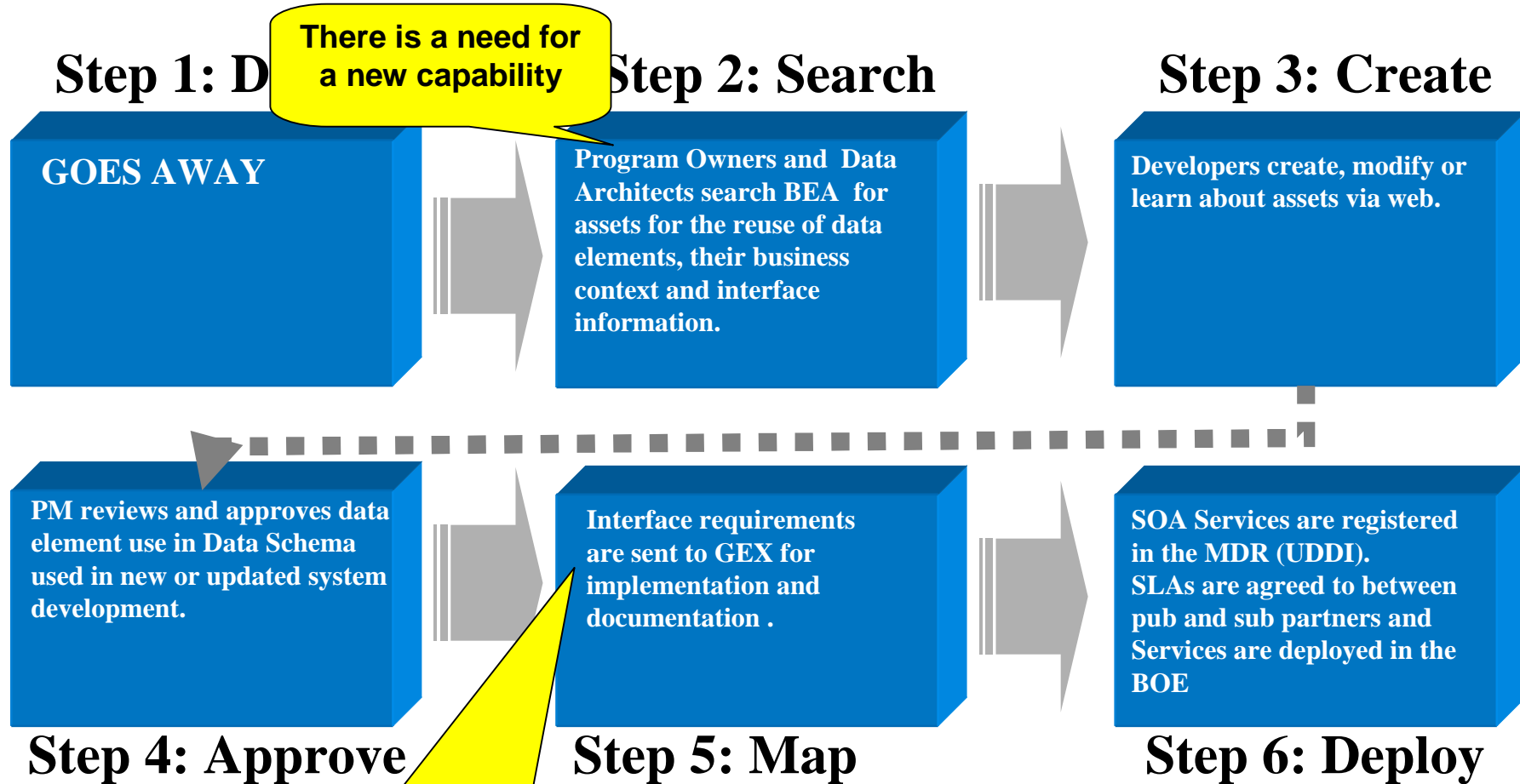


“To-Be” Architecture Meets “As-Is”



Canonical Data Element – Use Process

Search > Create > Approve > Map > Deploy



Real Life Examples

Mapping COI to the OV-7

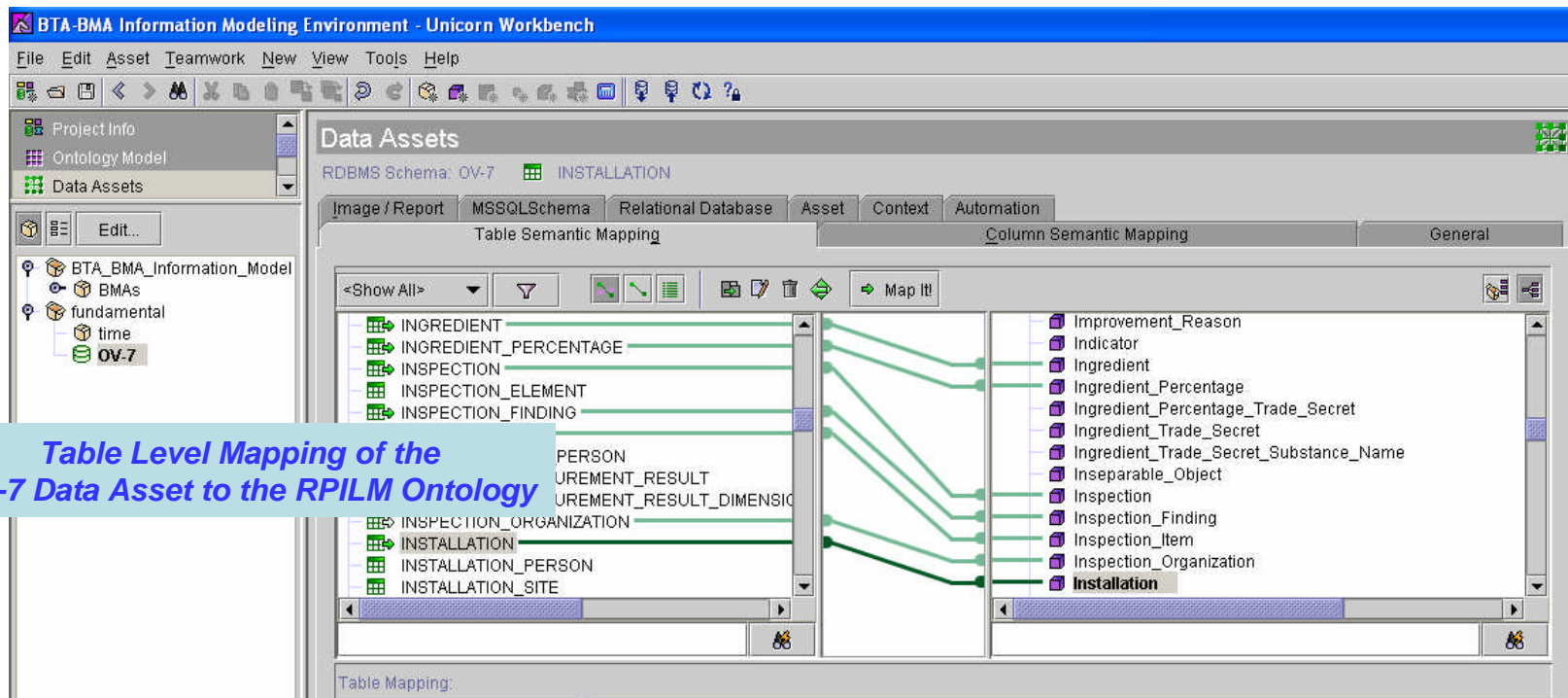
The screenshot displays the Unicorn Workbench interface for mapping data assets. The main window is titled "Data Assets" and shows the "INSTALLATION" table from the "OV-7" RDBMS Schema. The "Table Semantic Mapping" tab is active, showing a list of columns on the left and their corresponding ontology properties on the right. The columns are: INSTALLATION_CODE (char), INSTALLATION_PRIMARY_ACTIVITY_TYPE_CODE (char), INSTALLATION_OPERATING_STATUS_CODE (char), INSTALLATION_VIRTUAL_INDICATOR (char), ORGANIZATION_ROLE_CATEGORY_CODE (char), ORGANIZATION_UNIQUE_IDENTIFIER (varchar), INSTALLATION_DESCRIPTION_TEXT (char), INSTALLATION_NAME (char), and ORGANIZATION_ROLE_START_DATE (smalldatetime). The ontology properties are: installationCloseDate ([1..1]Instant), installationCode ([1..1]String), installationCreationDate ([1..1]Instant), installationName ([1..1]String), installationOperationalStatusCode ([1..1]Operational_Status), code ([1..1]String), description ([1..1]String), status ([1..1]String), installationPrimaryActivityTypeCode ([1..1]Primary_Activity), activity ([1..1]String), code ([1..1]String), description ([1..1]String), installationReportingComponentCode ([1..1]Component), code ([1..1]String), description ([1..1]String), installationVirtualIndicator ([1..1]Indicator), code ([1..1]String), and description ([1..1]String). Green lines indicate the mapping between the columns and the ontology properties. A blue callout box at the bottom left of the mapping area contains the text: "Column Level Mapping of the OV-7 Data Asset to the RPILM Ontology".

Column Level Mapping of the OV-7 Data Asset to the RPILM Ontology

Using Ontology for Compliance

Using a COTS, map the OV-7 Data Asset to the Real Property and Installations Lifecycle Management Model and determine compliance with the OV-7 Data Asset and completeness of the Model.

- The OV-7 Data Asset was mapped to the Ontology at both the Table and Column level to the corresponding Ontology Model Class and Attribute.



“Breaking up is hard to do”



Information Management Solutions
Consultants Inc. 703-635-0930

For Further Information:

**Contact William.Mancuso@IMSC.us
703-635-0930**