Enterprise BI Strategy - DoD Business Operations

April 21, 2010
Purpose of Presentation

- Definition of BI
- Business Operations Transformation Roadmap
- The Problem
- Enterprise BI Strategy
- Why Is This Hard? Current State of BI
- Future State
- The Roadmap
- Real World Example - EIW
Scope & Terminology

Business Intelligence (BI): The processes, technologies, and tools needed to turn data into information, information into knowledge, and knowledge into plans that drive profitable business action. (TDWI, 2002)

- BI is a foundational element of enterprise (or business, corporate, etc.) performance management (EPM)
Missions of the DoD

Warfighter Mission Area  Business Mission Area  Intel Mission Area

Enterprise Information Environment Mission Area

DoD BMA CTO & Chief Architect in the Office of the Deputy Chief Management Officer (DCMO)

Defense Business Systems Acquisition Executive and Enterprise Integration, in the DoD Business Transformation Agency (BTA)
Strategy and Roadmap for DoD Business Operations Transformation

Past
(BMA Federation Strategy version 2.4a)

Present
(BOE Execution Roadmap)

Future
(BMA Architecture Strategy version 3.0)

Past

BEA 3.0

Version 2.4a

Roadmap:
Architecture
Governance
Socialization
Services
Infrastructure

Vision & Strategy
Planning & Roadmap
Infrastructure
Governance

Present

DCMO/CIO Policies

Version 2.4a

Initial BOE Experience

Future

DoD Strategic Mgmt. Plan (SMP)

Semantic Information

Data Integration
Business Intelligence
Rules/Workflow
Security

Data Sharing and BI Enablement

Enterprise Stds.

Common Vocabulary (Ontologies)

Rules/Workflow

Security

BEA 8.x

DBSAE SOA Imp. Strategy

Execution

DBSMC/IRBs
DCMO/DCIO; EGB; BECCM

Vision & Strategy
Planning & Roadmap
Infrastructure
Governance
The picture above represents initial data set for systems and interfaces ….
How do I make sense of all this data?

The amount of data doubles every five years…

and…

DoD Business Operations requires relevant business information from operational systems in order to make insightful decisions about strategy and tactics.
We need a Business Intelligence Strategy so Systems and Services communicate

...and Humans can ask the right question, anytime and get an appropriate answer
Purpose: **Information Dominance** thru Secure Info Sharing

To Achieve this Purpose our Milestones are:

- Data is made visible and accessible as services
- Data is intelligible as information for analysis, reporting and presentation
- Information is accessible so users can analyze, report and present it

Vision: Every operator access the information they need when required, wherever it resides, securely.

To Enable the Vision We Must Address:

1. Standardization
2. Resourcing
3. Governance
4. Data origin, provenance and security envelope
Advantage of Semantic Business Intelligence Extractions – Base Case

- Three-year payback, based on IT costs only and excluding the costs of payroll, a breakeven investment would be $840 million.

1. The total number of BMA listed projects is 2,537.
2. Estimated Number of Inter-Application Retrievals of Intelligence - 5000
3. Using a conservative assumption that each application has at least two linkages to other applications, this suggests that there will be at least 5,000 data extraction exchanges among applications
4. 52 x 5,000 intelligence extractions per year.
5. BMA expenditures for operations are $13.9 billion, excluding the costs of payroll for military and civilian personnel.
6. The average daily IT cost of output from BMA applications is $13.9 billion / 2,537 projects / 52 is approximately $106K/week.
7. Assume 10% is BI related, therefore the annual cost of intelligence extraction will be at least $1060 * 52* 5000 = ~$280 million/year.
Establishing Enterprise BI Capability – Milestone Conceptual Roadmap

#1. Establish data services from transactional data sources
Jan 10

#2. Make data understandable for decision making purposes in Data Mart

#3. Decision making from BI Information Delivery

#4. Subsume PoRs data resources into Enterprise Capability, Sunset redundant data storage

---

Use Case transactions
Core Business Mission (CBM) transactions

Services for all CBMAs are available for Data mart

HRM instance
FM
CBM3
CBM4
CBM5

Jan 2011
Jul 2011
Jan 2012
Jul 2012
Jan 2013

Data to Reporting and Analytics

EIW, PM
Others as needed

DBSAE Enterprise Capability Stood Up

PI*
BEIS*
IMIE*
EIW*

Sep 10

Identify & Communicate through Enterprise Transition Planning
Others as identified

* Teams may continue, but support for #1 and #2 is integrated
Why is this hard?
“We don’t speak the same language!”

Intransparency: Stakeholders can’t communicate Needs & Requirements

Waste: Duplicate Developments because common requirements are not identified

Gov’t Mission, Regulatory, Citizen & Corporate Needs

Input

Domain-specific Descriptions

Architects

Unclear Requirements

Monolithic Systems Expensive to Integrate

Arbitrary Specification

Modelers, Designers & Engineers

Isolation: Vertical Practices form “Silos of Excellence”

Isolation: Proprietary and arbitrary service delivery creates effort, expenses

Problem: Architecture & Development
No Governance of Technical Depications

Not This

But This:

Resistor symbol

Capacitor symbol

This agreed upon representation of electrical engineering allows a common understanding…
No Governance of the BI Portfolio

Multiple & dissimilar requests & funds for BI activities

Multiple initiatives at accommodating BI requests

Repetitive efforts in data extractions & diverse locations and conditions of data

Integrated Management Information Environment (IMIE)
- Directorate: EF&I

BI Center of Excellence
- Directorate: EI

Enterprise Information Warehouse (EIW)
- Directorate: EI

Business Enterprise Information Services (BEIS)
- Directorate: DBSAE

DPAP Recovery Act & Procurement Analysis
- Directorate: DBSAE

DCMO

BTA (internal)

DPAP

SPS

EDA

BEA

FPDS-NG

CCR

External 2

GEX XML

SNAP-IT

DAI

ETP

DITPR

SPOT

DTS

DAMIR
BI - The Future
(Federated queries, Mash-ups, Federated queries)

Web Standards-based queries
Integration of data from disparate sources
Real-time access to and integration of information

Consistent, Real-time Answers
Establishing BI Governance

Provide customers with single process & POC to address needs

Centralized body with expertise and process to execute efficiently

Provide PEOs & other data providers a central POC to alleviate repetitive data requests

BI Governance Body
- Governance
- Coordination
- Expertise
- Requirements
- Analysis

DCMO
BTA (internal)
OSD & 4th Estate
MilDep

Originating Taskers

Coordination and Execution

Data Sources
Ontology Spectrum Progression

Interoperable Syntax  Interoperability  Interoperable Semantics

Static Resources  Dynamics  Dynamic Resources

Weak Semantics

Primitives CARP
Relational Model

Taxonomy

Has a narrower meaning then

Thesaurus

Is sub classification of

Conceptual Model

Is subclass

EA w/Ontology AKA DoDAF 2.0 or BEA, UML
Extended ER
RDF/S
XTM
Description Logic
DAML+ OIL, OWL
First Order Logic
Modal Logic

Source: Adapted from MITRE – Dr. Leo Obrst
How Do We Get There

Static Resources

Dynamic Resources

Net-Centricity

Real-time Bindings in Runtime

Ontology and automated reasoning

Semantic Web Services

Strong Ontology Based Mediation Pattern

Expanded Semantic Mediation Pattern

Ontologies

Hub and Spoke Pattern

Loosely coupled, Standards based

Proprietary Standards

Hard Coded Bindings

Text documents and database records

Documents using taxonomy, may be XML based

Canonical Model - XML ontology with controlled mixed vocabularies

Intranet

Web Services

Interoperable Syntax

Interoperability

Interoperable Semantics

Targets

Source: Adapted from MITRE – Dr. Leo Obrst
Members of a federation agree to certain standards to interoperate and relate to each other for the common good.

In a federation participants create their own policies, systems, facilities and delegate some authority to the federal authority.

The United States of America is a federation.

DOD is a federation.
DOD Federation Problem

- DOD is made up of many domains within domains
  - Army, Navy AF, Marines, OSD, JFCOM, …
    - Logistics, HR, Finance, Command and Control, intelligence …

- Each domain fields its own applications and creates its own information to execute its mission
  - It is often not possible to federate and integrate applications within domains

- In many cases it is necessary to share data within domains and across domains

- DOD will never meet its information sharing needs until it first solves the federation problem
1. Information Systems

2. Application information wrapped in web services

3. Service bus connects web services to DOD networks

4. No way to know how the information is related

5. No integrated data for the user, No federation
Federating communications networks does not mean that “information” is federated.

- Many people think that NCES is an information federation architecture, it isn’t
- It is a network federation architecture

TCP/IP, HTTP, XML let machines know how to read packets of bytes, but they do not tell machines what the information contained in the packets means

- It is easy to tell a machine what a TCP/IP packet means so we just build the knowledge into the hardware or firmware

Semantics must be added to the packets so that machines know how the information contained in the packet is related and what each term means.

- It is complex to tell a machine what a term means, so we build ontologies
- A “tank” in one domain is related to liquid, and in another domain “tank” is related to vehicle

If a machine can understand the similarity or difference of meaning of terms in domains, then we have a federated information architecture
Federation Solution

- **Enterprise Information Web**
  - Any information from any system can be shared with any other system on the Global Information Grid

- **Steps**
  - Describe all of the artifacts in each domain using standards (RDF, OWL)
    - We currently do this description work, but we do not use standards – Excel, Word, Powerpoint, Visio
    - The formal description of a domain is called a Domain vocabulary
  - Use these standards based descriptions to say how domains are related
    - this is the big missing piece of the current approach
  - Use these standards to say how all of the data in each domain is related to the domain vocabulary
  - Query the Domain vocabularies for any information

- The result is an Enterprise Information Web that meets the goals of information sharing as laid out in numerous DOD policy statements
1. Information Systems

2. Expose as RDF web services or SPARQL endpoints

3. GIG contains self-described data

4. GIG is a big federated knowledgebase of any information

5. Any authorized user or system can query the GIG for any information
Federation Progress

- DOD has led the charge and made progress on solving the Information Systems federation problem
    - This standard has allowed data networks to be federated
  - V.6, SMTP
    - Facilitates the federation of telecommunications networks and email

- In the past ten years standards to federate information have arisen
  - DARPA creates the Defense Agent Markup Language program in 1998 to facilitate information federation
  - W3C takes the work funded by DARPA and creates the Resource Description Framework (RDF) and Ontology Web Language (OWL) specifications
  - These specs are an integrated part of the W3C stack – HTTP, HTML, XML, XSD, namespaces, URI, and URL

- Taken together they form the standards on which an Enterprise Information Web can be formed
Leverage Existing Investment

- Enabling the GIG as an EIW leverages all of the existing infrastructure
  - Same networks, same security, same applications, same organizations
  - COIs already formed and new COIs will have to do their work in standards based tools instead of proprietary tools

- DOD is doing a lot of this description work now, it simply requires some redirection
  - Must use standards like any other federation

- The result of this relatively minor change and expense will be an astounding advance in information management capability
On Jan 16th 2009, Deputy Secretary of Defense issued a memorandum with a directive regarding enterprise level personnel and pay information requirements:

- **To meet the Department’s requirements for enterprise-level information visibility to support the needs of OSD and the Combatant Commands**, the enterprise will continue to manage the delivery of these information capabilities under the leadership of the Office of the Deputy Chief Management Officer (DCMO). This will include **establishing an enterprise level information warehouse** and the necessary functional and technical requirements to enable the delivery of this capability in close coordination with the delivery of the personnel and pay transaction systems by the individual military departments.
DoD currently lacks the capability to quickly and accurately account for personnel, manage troop strength, and war plan based on enterprise level authoritative, real-time personnel information.

Impact: challenges the objective of having the right personnel in the right place at the right time to fight battles and win wars.

Sample COCOM personnel data needs:

- Requirement 28.0: Provide timely and accurate information on the location, status, and identity of units, personnel, equipment, and supplies with emphasis on Personnel including patients.

- Requirement 124.0: Identify the availability of personnel within the active forces, or within the reserves and other sources when analyzing mobilization alternatives, to satisfy filler, replacement and force expansion personnel requirements.
The HR EIW is a mechanism for reaching into service applications to satisfy enterprise HR information needs. It accomplishes three things:

1. It reports real-time, authoritative HR information on-demand.
2. It creates HR enterprise information standards.
3. It supports IT flexibility.
HR EIW At a Glance

Multiple Sources

- DMDC Data Stores
- Services Data Stores
- External Data

DMDC Data Stores
HR Data Store
HR Data Store
HR Data Store

Services Data Stores
HR Source
HR Source
HR Source
HR Source

External Data

Single view

Common Vocabulary
Combine and Analyze

Combine multiple data and service sources into single view
Activities and Benefits

WHAT WE’RE DOING

Building a DOD HR ontology (or Common Vocabulary) using W3C standards called RDF/OWL resulting in:
- A conceptual model that is queryable in a run-time environment;
- Unambiguously described information in the DoD HR Domain
- A common vocabulary for information integration (federation)

MAJOR BENEFITS

- A virtual web of DoD authoritative source HR systems
- Answers for any HR information requirements with existing HR systems or ability to show the gaps
- A plug-and-play federated environment so new systems or analytical needs can come online and go offline without disrupting the overall environment
- Near real-time, on-demand, dynamic and authoritative data
Activities and Benefits

WHAT WE’RE DOING

Building a DOD HR ontology (or Common Vocabulary) using W3C standards called RDF/OWL resulting in:
- A conceptual model that is queryable in a run-time environment;
- Unambiguously described information in the DoD HR Domain
- A common vocabulary for information integration (federation)

MAJOR BENEFITS

- A virtual web of DoD authoritative source HR systems
- Answers for any HR information requirements with existing HR systems or ability to show the gaps
- A plug-and-play federated environment so new systems or analytical needs can come online and go offline without disrupting the overall environment
- Near real-time, on-demand, dynamic and authoritative data
**Process Ontology**

Translation of BPMs to OWL. Activity/Message descriptions include relationships to domain vocabulary terms.

**Domain Vocabulary**

Description of concepts in Domain. For HR, will initially be based on CHRIS’s Mapping of domain vocabulary terms to the physical data elements they represent in the Authoritative data source systems; currently mapping MCTFS, DIMHRS

**Authoritative Data Sources**

**Analytic Requirements**

Models analytic requirements for the Domain Ontology and how they relate to the concepts in the Domain Vocabulary, including SPARQL queries (CHRIS and COCOM 129s)

**References**

Shows from which authoritative documents concepts have been extracted; eg: DoDIs, etc
Domain Ontology Defined

- The Domain Ontology is a conceptual description of the domain covered by the relevant business processes
  - The “domain” is defined by the business processes and rules, the information sources, and any reports which are required
- Instances in this ontology are the same instances which are currently stored in information sources (i.e. databases)
CHRIS Modeling

- Each CHRIS is mapped to a property in the ontology
- A Property has a Domain and Range
  - Domain – Defines types of things that can have a value for the property
  - Ranges – Defines types of things that can be the value for the property
- If a CHRIS has Permitted Values defined, they are represented as an enumerated list of range classes
  - Represents that there is a specific view of the range of the property
## HR EIW Progress: CHRIS Metrics

<table>
<thead>
<tr>
<th></th>
<th>CHRIS Modeled</th>
<th>% Total CHRIS Modeled</th>
<th>CHRIS Aligned to DIMHRS</th>
<th># CHRIS Signed Off</th>
</tr>
</thead>
<tbody>
<tr>
<td>P&amp;R Supplied CHRIS</td>
<td>328</td>
<td>100%</td>
<td>247</td>
<td>121</td>
</tr>
</tbody>
</table>

**Legend**

- P&R Supplied CHRIS: # of CHRIS to complete (this number will fluctuate throughout project)
- CHRIS Modeled: progress metric; # CHRIS that have been modeled
- CHRIS Aligned to DIMHRS: progress metric; # of CHRIS processed through DIMHRS model alignment activity
- CHRIS signed Off: completion metric: CHRIS statements that have been “signed off” by P&R as accurately aligned and modeled
HR EIW Technology Architecture Roadmap

- **Current 2-year Schedule**
  - Implement or Build Standards-Based Semantic Information Management Platform
    - ✓ Modeling PoD
    - **RDF Warehouse**
      - RDF Services / SPARQLizer
      - Federated SPARQL Engine
  - Implement or Build SPARQL-driven Business Intelligence (BI) Platform
  - Risk Mitigation Strategy includes Implementation of Relational Warehouse (ETL)

- **Current Plan for Years 3 & 4**
  - Select, Scale, & Deploy Operational Technology
Objectives Achieved:
- Web Service
- DKO CAC Authentication
- Data Virtualization
- ETL Process
- DMDC MOU
- P&R HR Ontology Models
- DIMHRS Reuse
RDF Warehouse Architecture (POD2)

User Agent (Web Browser)  HTTPS

Host Network

Web / Application Server

- SPARQL Data Access
- Wiki Content
- Modeling

Model Driven Analytics

Triple Store

RDF Triple Store

Schema

Triples

RDF Load

SQL

DIMHRS – HRTS2
Scrambled DB

Relational Data Source

Model Driven ETL

Real Time Interface

Batch Process

API

NIPRNet / Internet
Backup
# Ontology Graph Key

<table>
<thead>
<tr>
<th>OWL Construct</th>
<th>Relationship to CHRIS Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td></td>
<td>Text inside a bubble represents a class. In this example, &quot;Person&quot; is the class being represented.</td>
</tr>
<tr>
<td>Object Property</td>
<td></td>
<td>An arrow connecting two classes represents an object property. These arrows are always purple. In this example, the property &quot;has Hair Color&quot; is being represented by the purple arrow.</td>
</tr>
<tr>
<td>Datatype Property</td>
<td></td>
<td>An arrow connecting a class to some datatype element represents a datatype property. These arrows are always orange. In this example, the property &quot;has Social Security Number&quot; is being represented by the orange arrow.</td>
</tr>
<tr>
<td>Domain</td>
<td></td>
<td>A class with an arrow pointing away from it represents the domain of a property. The domain is the class that can have a specific property. In this example, the class &quot;Person&quot; has the property &quot;has Hair Color.&quot;</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>A class with an arrow pointing toward it represents the range of a property. The range is the class that can be referred to by a specific property; it constrains the values of this property. In this example, the property &quot;has Hair Color&quot; refers to the class &quot;Hair Color.&quot; So, the values for the property &quot;has Hair Color&quot; must be part of the class &quot;Hair Color.&quot;</td>
</tr>
<tr>
<td>Subclass</td>
<td></td>
<td>A bubble inside another bubble represents a subclass. In this example, the class &quot;Member&quot; is a subclass of the class &quot;Person.&quot;</td>
</tr>
<tr>
<td>Subproperty</td>
<td></td>
<td>An arrow originating from a longer arrow represents a subproperty. In this example, the property &quot;has Address&quot; is a superproperty, and it has subproperties &quot;has Legal Residence Address,&quot; &quot;has Mailing Address,&quot; and &quot;has Residence Address.&quot;</td>
</tr>
<tr>
<td>A Class that has Allowed Values of a CHRIS</td>
<td></td>
<td>Green text inside a bubble represents a class that contains Allowed Values of a CHRIS. In this example, the class &quot;Blood Type&quot; contains allowed values of the CHRIS &quot;Blood Type&quot; (i.e., A, AB, B, or O).</td>
</tr>
<tr>
<td>A property that represents a CHRIS</td>
<td></td>
<td>A property that appear in pink font represents a CHRIS. In this example, the property &quot;has Blood Type&quot; represents the CHRIS &quot;Blood Type.&quot;</td>
</tr>
</tbody>
</table>
Business Transformation and Interoperability Goals for BI:

Objective: De-conflict redundant capabilities and informational silos

1. Policy to Establish functional interoperability through informational interoperability
   - Identify redundant capabilities using common vocabulary
   - Create understandable business processes using standardized representation (Primitives)
   - Create consistent and reusable vocabularies using CARP

2. Limit the creation of new data services to only those cases when they cannot be created from existing information exchanges

3. Remove the need for custom interfaces by creating implicit interoperability
   - Make information understandable using information models √
   - Create standardized and reusable methods for accessing data
   - Create physical instantiation of the common vocabulary
“Unify heterogeneous data sources to facilitate measuring DoD Strategic Management Plan performance targets & metrics output reports for action and remediation.”
BI Service Vision Applied to Performance Measurement

- DCMO Business Operations
- Services, Components, Agencies
- Status
- History
- Standards
- OSD
- External Sources
- Authoritative Process Data
-Query Service
- Source Data Values
- Virtual Project Teams
- Service Pulls Data
- Ad Hoc and Standard Displays
- Performance Service Level Agreements

Owners make data available
BI Is A TEAM SPORT – Matrixed Teams

Provide customers with single process & POC to address needs

Centralized body with expertise and process to execute efficiently

Provide PEOs & other data providers a central POC to alleviate repetitive data requests
**Business Intelligence 2 Year Transformation Roadmap Template**

- **Pilot** - Implementation using Traditional BI Technical Approach – limited scope

- **IOC** - Implementation using Traditional BI Technical Approach (non-standard Semantic Mediation Strategy & Technology w/data services)

- **FOC** - Implementation using Semantic Strategy & Technology

Fuzzy Timelines for Transition Decisions

- Start: Sep 09
- Jan 10
- Oct 10
- Oct 11
Governance Required

Tier of Information Exchange

- **FEDERAL TIER**
- **DoD TIER ENTERPRISE SERVICE**
- **SEGMENT TIER ENTERPRISE SERVICE**
- **COMPONENT TIER LEVEL SERVICE**
- **PROGRAM TIER LEVEL GOVERNANCE**

**Design and Maintenance of Policy**
- DoD CIOs
- SEGMENT COIs
- Components

**Application of Policy to Design and Execution of Runtime**
- Service Providers
- Service Consumers

This is where we concentrate
**BI - How it works – EIW Example**

Eg: To satisfy for a mobilization planning exercise (requirement 124.0) a COCOM user demands a specialized troop roster to include Service members from each Service. Problem: The services call Service members something different:

- Soldier
- Airman
- Sailor

Request forwarded

Svc specific answer

Request forwarded

Svc specific answer

Request forwarded

Svc specific answer

Service Member

EIW

Roster
DoD BI Pattern – FOC

Implementation using Semantic Strategy & Technology

Source Data

| HR | USMC Data Store | Navy Data Store | USAF Data Store | DoD 4th Estate Data Store |
| Performance | USMC Data Store | Navy Data Store | USAF Data Store | DoD 4th Estate Data Store |

Data Organization: Virtualization, Caching and Modeling

- Virtualization – Basic
- Service Registry and MDR
- Validation
- Upper Ontology

ETL or Load

- Data Virtualization – Basic
- Data Virtual - Composite Business Data Service

Build

- Service Registry and MDR
- Ontology Model

Query

- Information Warehouse
- Information Warehouse - RDB as Triple Store
- Soldier, RDF
- Soldier, RDF
- Metric, RDF
- Metric, RDF

Presentation & User Access

Single View of Truth

User access to information and BI capabilities

WebPages (Result)

BI Visualization (Analytics)

Information Flow

Interoperability Controller

Mediation Side

Federated DoD/IC Data

Mission Area Data

COI
Best Practices for BI Implementation in Industry – Currently Applied to EIW and PDA

- **Conform to Users**: what information they need, how they like to receive it, and how often they want to interact with it.
- One BI Tool will not satisfy all audiences, but organizations should **standardize tools** for each category (OLAP, Dashboard/Scorecard, Querying, Mining, & Modeling).
- **Take Inventory**: what tools currently exist, users served, cost analysis (BCR), etc.
- Enterprise BI initiatives require **upper management mandate**
- **Formalize process** on what information is needed by which organizations.
- Established **master data management** (common vocabulary)