Department of Defense
Enterprise Information Web (EIW)

Jonathan Underly
Program Manager
The Enterprise Information Web (EIW) is pioneering the adoption of Semantic Technology and approaches that can be the way forward for enterprise business intelligence and solution architectures in the DoD.
A Vision for DoD Solution Architectures

User executes End-to-End Business Process (E2E BP)

Query BEA directly:
- Enterprise analytics
- Compliance
- IRB/portfolio management

E2E BP executes via BEA directly

OMG Primitives Conformance class 2.0
BP models uniformly described

E2E BP executes via BEA directly

Acq Domain Vocabulary

Real Prop Domain Vocabulary

Log Domain Vocabulary

Fin Domain Vocabulary

HR Domain Vocabulary

Svc Member

OUID

DoD Authoritative Data Source

W3C Open Standards Legend:
- Data described in RDF
- Relationship described in OWL

EIW is Currently in this domain
Problem: Personnel visibility (PV), accurate and timely pay
Alternative: Build an enterprise ERP for HR functionality across DoD

<table>
<thead>
<tr>
<th>Measure</th>
<th>Outcome</th>
</tr>
</thead>
</table>
The HR Enterprise Information Web (EIW) is a mechanism for reaching into Authoritative Data Sources (ADS) to satisfy enterprise information needs. It accomplishes three things:

1. Reports near real-time, authoritative information on-demand
2. Supports enterprise information standards (Open; HRM ES)
3. Supports IT flexibility/agility
EIW Roadmap: Phased Approach

**Ontology Development**
- HR Ontology Development
- HRM Enterprise Standards Modeling & Enterprise Analytical Needs Discovery
- ADS/BPA Modeling: CII
- ADS/BPA Modeling: Separations

**Technology Development**
- RDF Modeling
- RDF Triplestore
- RDF Store Extension
- SPARQL end-points
- Sparqlizer
- Local Federation
- Distributed Federation
- Scaleable Federation
- Secure Federation

Legend: ADS/BPA = Authoritative Data Source/Business Process Area

8 June 2011
Jonathan Underly - EIW Project Manager
PoD Progression

Data Source:
CII (USA)

PoD1:
Models inform the location and extraction of data

Composite App Srvr
Ab Initio App Srvr

Triple Store

SQL
RDF Load
Triple Store Loader
(Semantic Mapping)

SPARQL
Source
Ontology
Web/App Server

PoD2:
• Model Driven Analytics
• Model Driven ETL
• Triple Store

Data Source:
MCTFS (USMC)

PoD3:
Store and query multiple test data sets in a triple store

PoD4:
• Access data from relational store at run time using semantic query engine
• SPARQLizer alpha release with D2RQ mappings

PoD5:
• Access data from relational store at run time;
  - SPARQLizer 1.0 release, support SQL equivalent commands
  - Added Ad-hoc & Faceted Search capability
  - More demographic reports

PoD6:
• Semantic Federation of local data
• Aggregated Svc Member
• Demographics
• Open Social gadgets

Extensibility and Reuse
• Enterprise Perform. Metrics Added
• USMC MCTFS Mappings reused
• NO changes to schemas or other queries
EIW Ontology Architecture

Machine readable policy with automated lineage to the supporting source data.
Enabling Strategic Management

4.0 Preserve and Enhance the All-Volunteer Force

2.0 Support Contingency Business Operations

4.2.10 Percentage of the Dept. AD who meet objectives for time deployed vs time at home

Dwell Time

Army Dwell Time E2E

USMC Dwell Time E2E
BACKUP
Big Picture
Graph 1

HR Dataset

Ontology – Based Information Integration & Analytics

What Pay Grade is Col. Blatt?
How much Dwell Time does Col. Blatt have?

Graph 2

Deployment History

person

hasName

Col. E.J. Blatt

hasDwellTimeStatus

Dwell Time

24 months

hasValue
Who has a Pay Grade of “O6” and has at least 24 months of Dwell Time?
EIW Defined
<table>
<thead>
<tr>
<th>COA</th>
<th>Description</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status quo</td>
<td>Manual aggregation and gathering of information in disparate systems</td>
<td>Process known</td>
<td>Labor intensive (e.g.: daily JPERSTAT report consumes 70 person-hrs); uncertain data lineage</td>
</tr>
<tr>
<td>DIMHRS</td>
<td>Single military personnel and pay system</td>
<td>Efficient; accurate</td>
<td>DIMHRS not fielded; Political change curve substantial</td>
</tr>
<tr>
<td>Traditional Warehouse</td>
<td>Set up a traditional network of data stores to pull and store personnel and pay related information</td>
<td>Known model and technology stack</td>
<td>Duplicates data; costly to develop &amp; to maintain; very costly to modify</td>
</tr>
<tr>
<td>Semantic Approach</td>
<td>Semantically describe personnel and pay information assets, pull, aggregate and display (vice store)</td>
<td>Federated data = data lineage; powerful analytics; virtual data (no duplication); easier to modify and maintain; highly extensible</td>
<td>Maturing technology; Technology change curve exists</td>
</tr>
</tbody>
</table>
EIW Alignment
## DoD Strategy Alignment to DoD Information Priorities

<table>
<thead>
<tr>
<th>DoD Strategy</th>
<th>Alignment</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DoD Information Enterprise Strategic Plan</td>
<td>DoD and mission partners will obtain an information advantage when timely, secure and trusted information is available to all decision makers</td>
<td>EIW has demonstrated the capability to federate locally distributed data; currently working on dispersed distributed data sets</td>
</tr>
<tr>
<td>DoD Strategic Management Plan</td>
<td>Improve IT Acquisition Performance</td>
<td>Agile method to rapid capability deployment; Surfacing authoritative data in the H2R E2E BP; Exposed data to support SMP performance measure “Meet End-Strength Goals”</td>
</tr>
<tr>
<td>DoD Net-Centric Data Strategy</td>
<td>Visible, Accessible, Understandable, Trusted, Interoperable, Responsive</td>
<td>EIW enables all these goals save Trusted, currently planned for PoD 8 &amp; 9</td>
</tr>
</tbody>
</table>
EIW - Benefits
Interoperable Systems
Fact-based Portfolio Analysis
Data Lineage Traceability
Cross-Service Analytics
Compliance Assessment
Arbitrary Extensibility
Common Language

With this
Come These

Data Concept System Std
Airman Svc Mem Pers Sys
Soldier Svc Mem Pers Sys
Lawyer Svc Mem Pers Sys

Capability System Y/N
Leave PersPay1 Y
Leave PersPay2 Y
Leave PersPay3 N

8 June 2011
Jonathan Underly - EIW Project Manager
EIW Approach
EIW Capabilities

Current

- Federated dashboards
  - Operational view
  - Demographic reports
  - Faceted search capability
  - Ad hoc query capability
  - PDA pilot capabilities
- Gap analysis
- Standards compliance
- Policy alignment
- ADS traceability
- Gadget repository
- Process alignment
- Performance measures

Future

- Automated policy compliance
- Interoperability analysis
- Portfolio management/strategic initiative management
- BPR analysis
- Cross domain analytics
- Live data sources
- Scalability
EIW PoD6 Demo
Interoperability Framework

DoD Analytics
- Human Resources
- Weapon System Lifecycle
- Material Supply and Service
- Real Property and Installation
- Financial
- Portfolio Management
- Interoperability Analysis
- Policy Compliance
- Performance Analysis
- BPR

Semantic Technology
- SPARQLizer
- Federator
- Rules Engine
- Ontology Architecture
- Processes
- Domain
- Systems
- Mappings
- Standards
- Policies
- Projects
- Schemas
- Requirements

Semantic Models (by BPA)

Authoritative Data
- Authoritative Data Sources
  - Legacy Systems
  - Services
  - Data Warehouses

Operational Governance

Warfighter Requirements

Security & Hosting

Center of Excellence

Executes on

8 June 2011
Jonathan Underly - EIW Project Manager
The Root of Information Visibility (IV): Information Federation

IV consists of many domains within domains

Each Service is a domain. Each Service fields its own applications and creates its own information to execute its mission

It is prohibitively expensive and fragile to federate and integrate relationally based applications within & across domains

DoD will not solve ISV problem until we solve the INFORMATION FEDERATION problem

8 June 2011
Jonathan Underly - EIW Project Manager
Semantic Technology Layer Cake

- User Interface & Applications
- Trust
- Proof
- Unifying Logic
- SPARQL
- Ontology: OWL
- RDF-S
- Rule: RIF
- RDF
- XML
- URI/IRI
- Crypto
EIW Alignment with DoD Information Priorities

• DoD Information Enterprise Strategic Plan:
  – **Information** is an asset—a source of power and a force multiplier. DoD and mission partners will obtain an information advantage when timely, secure and trusted information is available to all decision makers.
  – The Department must also gain wider acceptance of rapid acquisition methods for fielding new IT capabilities that demonstrably improve operational performance.

• DoD Strategic Management Plan:
  – Improve IT Acquisition Performance: agile method to rapid capability deployment
  – Surfacing authoritative data in the H2R E2E BP
  – Exposed data to support SMP performance measure “Meet End-Strength Goals”
The DoD Net-Centric Data Strategy, issued by the DoD CIO May 9, 2003, is the plan to make the Department’s information resources:

- **Visible**: Is an information resource discoverable by most users?
- **Accessible**: Is it available on the network, and are tools readily available to use it?
- **Understandable**: Can it be intelligibly used? Are the semantics well documented?
- **Trusted**: Are the source, security level and access controls of the data available to users?
- **Interoperable**: Can it be combined or compared with other information? Can it be mediated?
- **Responsive**: Is the data what users need? Are robust user feedback mechanisms in place to improve it?
Evolution of Information Management Technologies

Defense Advanced Research Projects Agency
- Relational Database Technology
- TCP/IP
- OWL/RDF

- DARPA creates the Defense Agent Markup Language program in 2000 to facilitate information federation - DAML.org
- W3C takes the work funded by DARPA and others to create the Resource Description Framework (RDF) and Ontology Web Language (OWL) specifications
- These standards are not just for modeling or serialization; they comprise a complete information management technology architecture
- There are no other standards that can be used to accomplish the goal of information federation
PoD 1 Architecture

Objectives Achieved:
- Web Service
- DKO CAC Authentication
- Data Virtualization
- ETL Process
- DMDC MOU
- P&R HR Ontology Models
- DIMHRS Reuse

Firewall

DMZ

HTTPS Port 443 for web traffic

SOAP/XML ↔ XML File

Web Service Call (bind)

DKO

Port 443 open for web traffic to DMDC

HTML

Voluntary Retirement

Requests Received

% of Approved Requests

Discharges by Types

Port 443 for web traffic to DMDC
PoD 2 Architecture

User Agent (Web Browser) -> HTTPS

Host Network:
- **Knoodl**
  - SPARQL Data Access
  - Wiki Content
  - Modeling

Web / Application Server:
- Protocols:
  - HTTPS

**Mulgara**
- RDF Triple Store
- Schema
- Triples

**Oracle**
- DIMHRS – HRTS2 Scrambled DB
- Relational Data Source

**ETL (Semantic Mapping)**
- RDF/Relational Mapping

**Real Time Interface**
- Batch Process API

**API**
- Model Driven Analytics
- Triple Store
- Model Driven ETL
PoD 3 Architecture

User Agent (Web Browser) -> NIPRNet / Internet (HTTPS)

Google Maps

Derby DBMS
- Dashboard Cache, Geographic & Language Information (Public Domain)

Dashboards (Fusion)

DWR Service

Knoodi
- SPARQL Endpoint
- Presentation & Business Logic
- Wiki Content
- Modeling

Mulgara Triple Store
- Ontology
- Triples

Oracle 11g DBMS
- Test Data:
  - USMC (ODSE)
  - Army (DIMHRS) (No PII)

Oracle Triple Store
- Ontology
- Triples

USMC Network

Data Scramble
- Manual File Transfer (No System Interface)
- Scrambled USMC Records (No PII Allowed)

USMC Records (Relational)

USMC Records (Relational)

Army Records (Relational)

Server (m1.large)

Web / App Server (m1.large)

API

SPARQL Endpoint - Installed and tested, not yet tested with DWR Service

SPARQL Endpoint

SPARQL Endpoint

SPARQL Endpoint

SPARQL Endpoint

DWR Service

Test Data:
- USMC (ODSE)
- Army (DIMHRS) (No PII)

API

Triples

Ontology

Dashboards (Fusion)

Real Time Interface

Batch Process

API
PoD5 Architecture

Information Flow for EIW POD Dashboard
Information Flow for EIW Common Vocabulary Views
Web Browser Application

User Agent (Web Browser)

Google Maps

SPARQL Endpoint:
A standard web interface for accessing data described in RDF via the SPARQL Query Language.

Ontology Repository:
A OWL/RDF knowledgebase where descriptions of things and relationships between things can be managed. For example, the BEA can be fully described, managed, and maintained in an Ontology Repository.

SPARQLizer:
Reusable software component (service) that provides SPARQL access to any relational ADS for which a Mapping File has been developed.

SPARQLizer:
Description of how to access data via a Common (Domain) Vocabulary from an authoritative data source (ADS).

Mapping File:
Description of how to access data via a Common (Domain) Vocabulary from an authoritative data source (ADS).
EIW IA Plan

Information Assurance

IATT

SLA Sign-off

Encryption

ERB and CCB

Status and Accreditation IATT

ATO

TBD

March, 2011

4/20/11

4/22/11

4/29/11

5/20/11

6/11 – 12/11

December, 2011

8 June 2011

Jonathan Underly - EIW Project Manager
EIW IA Current Status

Initiation
- Full Time IA resources on board (100%)
- ADCF and DMDC Site Visits (100%)

IATT
- Review current status (100%)
- Doc. Physical Architecture (0%)
- SLA sign-off (0%)
- Doc. IA Control & Procedures (0%)

ERB and CCB
- All Controls updated in EMASS (0%)
- Status review with DAA (0%)
- Accreditation Decision IATT

ATO
- COOP, CM, Testing (0%)
- Dec 2011

(As of 4/6/11)

- 0% Task not started
- 25% Task started
- 50% Task in full execution
- 75% Task finishing
- 100% Task finished

- No issues
- Task behind schedule
- Anticipated impact to parent task or overall project schedule