What is EIM?

- **Enterprise Information Management (EIM)** is a critical business initiative at the FAA, endorsed by executive leadership.
- It is a **business discipline** that applies management best practices and governance techniques to effectively and efficiently deliver information as a service.
- While EIM is **not a singular technology, system or an IT project**, by tapping into existing and new enterprise resources in the agency, it will make available needed enterprise data and information capabilities.
Information Centric

By Transforming FAA into an Information-Centric Enterprise we can:

• **Support and enhance LOB/SO capabilities** to manage and deliver data and information assets

• **Discover, access, and utilize** the potential of FAA data and information

• Create a partnership of **information and governance best practices**

• Discover and share **operational best practices** across the Agency

• Utilize **enterprise-wide shared enabling resources** (technology platforms & services)
Semantic Components

To make FAA Information-Centric, we need central principles:

- Identifying **Cross-cutting information domains**
- Enabling **extensive tagging and metadata enrichment**
- Navigating across **concepts and information domains**
- **Converging information around the user**
An Example

SAFETY ANALYSIS TODAY

I spend time on process.

Resolve the data manually

CCMIS  RSTS  Incident Reporting

Analytic A  Analytic B  Analytic C

Analyze manually in different systems

DATE
?
Answer in Hours or Days

OARS EIM TOMORROW

I spend time making decisions.

OARS Portal:
I ask for the information and analytics I need, in the format I want to do my analysis

Shared Services handle data pulls, integration, and analytical loading

Information  Data  Analytics

Answer in Minutes
A Holistic Approach

• New initiatives involving process or technology are often “silos”, very system-centric, and struggle as a result

• To be successful, this had to be an “integrated” effort, including:
  – Business Information Driven
  – Forward looking Governance
  – User Needs Discovery
  – Technology Demonstration
How To Implement

• Governance
  – FAA Steering Committee - prioritization
  – Information and Data Advisory Board – enterprise policy
  – COI’s – information domain layer
  – COP’s – data subject layer, data infrastructure

• Needs Discovery
  – Process for understand business need

• Technology Demonstration
  – Data Management
  – Advanced Analytics
  – **Semantic Enrichment and Search** (Specific Example: Dynamic Regulatory System (DRS))
DRS Architecture

Search
- REST API
- Library Services REST API
- Start page
- DRS UX
- Librarian UX

Analytics
- Elasticsearch
- MapReduce
- R
- Stanza/GP
- MongoDB
- Thrids
- Noods

Infrastructure
- FAA Doc Repo
- DRS Boundary
- SOR Doc Retrieval
- Data
- Messaging

Data
- SORs
- MongoDB
- ACs
- ADs
- E-CFIs
- TBD

Extract Transform and Load Process
- Ingest Processing
  - Concept Extraction
  - Transform
  - Normalize
  - Load

Implementation status:
- Implemented
- Not implemented
DRS Use Design

1. **End User**
   - Request specific document
   - Search query

2. **Web UI**
   - Extract concepts
   - Search results
   - Doc by ID
   - Document

3. **REST API (JSON)**
   - Find related concepts and documents
   - Get concept results, snippets, and facets

4. **Elasticsearch**
   - Get full document

5. **Graph DB**

6. **Data Store (MongoDB)**

---

Air Transportation Information Exchange Conference - Global Information Management
DRS Ontologies

• Ontologies Integrated
  – Aircraft-List.rdf
  – CFR_91.rdf
  – CFR_129.rdf
  – DRS-Aircraft.rdf
  – DRS-AirOperator.rdf
  – DRS-Atmosphere-Airspace.rdf
  – DRS-Document.rdf
  – DRS-GeneralFlightRules.rdf
  – DRS-MaintenanceRepair.rdf
  – DRS-OperationsForeignCarrier.rdf
  – DRS-PilotCertificate.rdf
  – Manufacturer-List.rdf
DRS Technologies

- Technologies Used:
  - mongDB
  - Titan
  - REST APIs
  - TTL/SKOS Ontologies
  - TopBraid, Protégé
  - Elasticsearch
  - Regular Expression
  - Stanford NLP
  - Elasticsearch
  - Cytoscape
  - JAX
  - Gremlin
Search Principles

- Three Methods for Driving Relevance, Recall, and Precision
  - Semantic Traverse (Gremlin/Titan)
  - Keyword tuning (use of sets in Elasticsearch)
  - Scoring (use of basic relevance scoring functions in Elastic Search)
Semantic Traverses

• Specific to Titan, use the ontology to retrieve and / or prioritize relevant semantic concepts
  – Focus on traversing the ontology
    • Add vertices that are 1 degree from search concepts to broaden recall
    • Identification of common concepts and using this to specify an intersection narrows precision
    • Traverse up or down the ontology branch to broaden or narrow the search
• Use the concepts to provide an intelligent interactive user interface to “specify” the search
Direct User Benefit

• **Provide Relevant, Trusted Information in an Actionable Format to Enable Agile Decision-Making**
  – Semantic Search means more **relevant, information-driven results**
    • See across repositories
    • See documents you weren’t seeing before
    • Improve relevancy
  – Graph interface means better understanding which leads to rule rationalization (IG report) and **better decisions**
    • Document and rule cross-linking to allow users “complete” view of domain

• **Guide Organizational Culture to Embrace an Information-Centric Enterprise**
  – **Improves effectiveness** of personnel in day-to-day work
  – Moving from **document to knowledge discovery**