Global Harmonization Through Collaboration

Information Models and their Relationship to Semantics

Presented By: *Deborah Cowell*
Dr. Candice Buchanan

Date: *August 29, 2012*

EUROCONTROL

Federal Aviation Administration

AIR TRANSPORTATION INFORMATION EXCHANGE CONFERENCE - (FEATURING AIXM, WXXM AND FIXM)

August 28, 2012 - August 31, 2012
NOAA Auditorium and Science Center
Silver Spring, Maryland

Why Semantics?

- ***What is the problem we are trying to solve?***
- ***What is the requirement?***
- ***What have we built so far?***
- ***How is this different from a traditional approach?***
- ***What are we doing now?***
- ***Where do we go from here?***
- ***How will we support the business community?***

What is the problem we are trying to solve?

We are not meeting the business communities expectations…

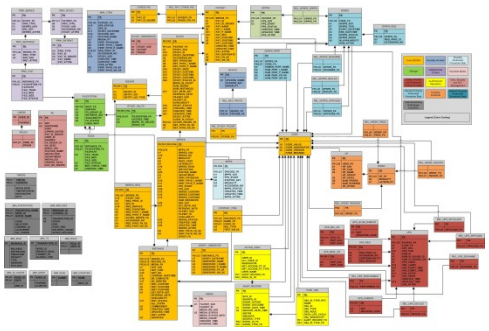
- What we are delivering is **too expensive and too long to deliver**.
- Solutions are **brittle/not adaptable**, we cannot keep up with the pace of the needs of the business.
- We are **drowning them in data** and they are **starving for information**.

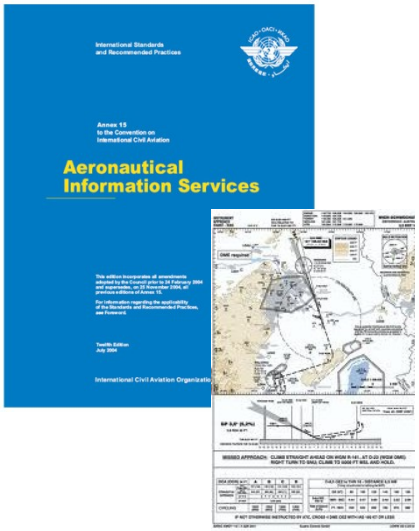
Data Organization

Structured Data (30%)

Semi-Structured Data (25%)

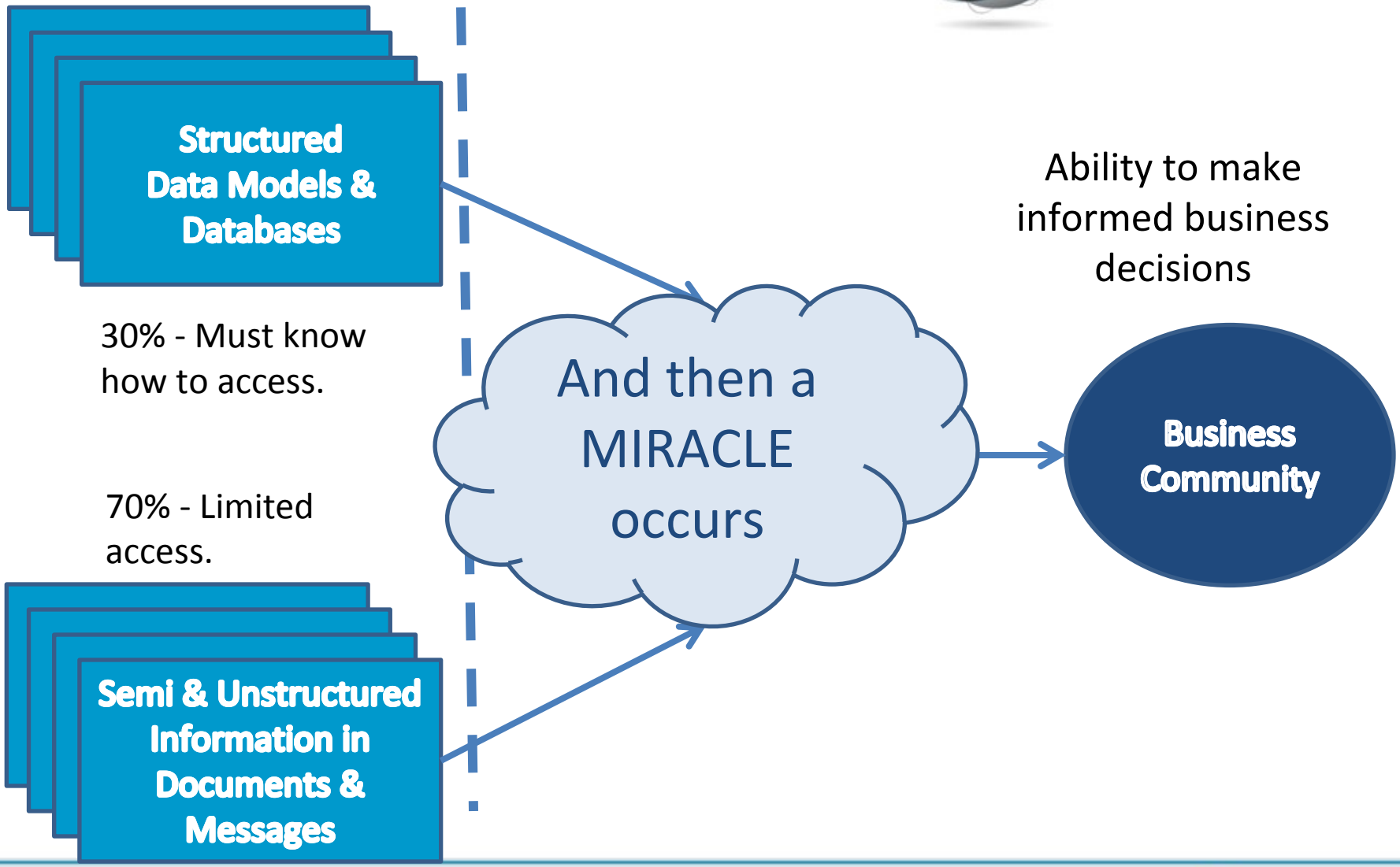
Unstructured Data (45%)



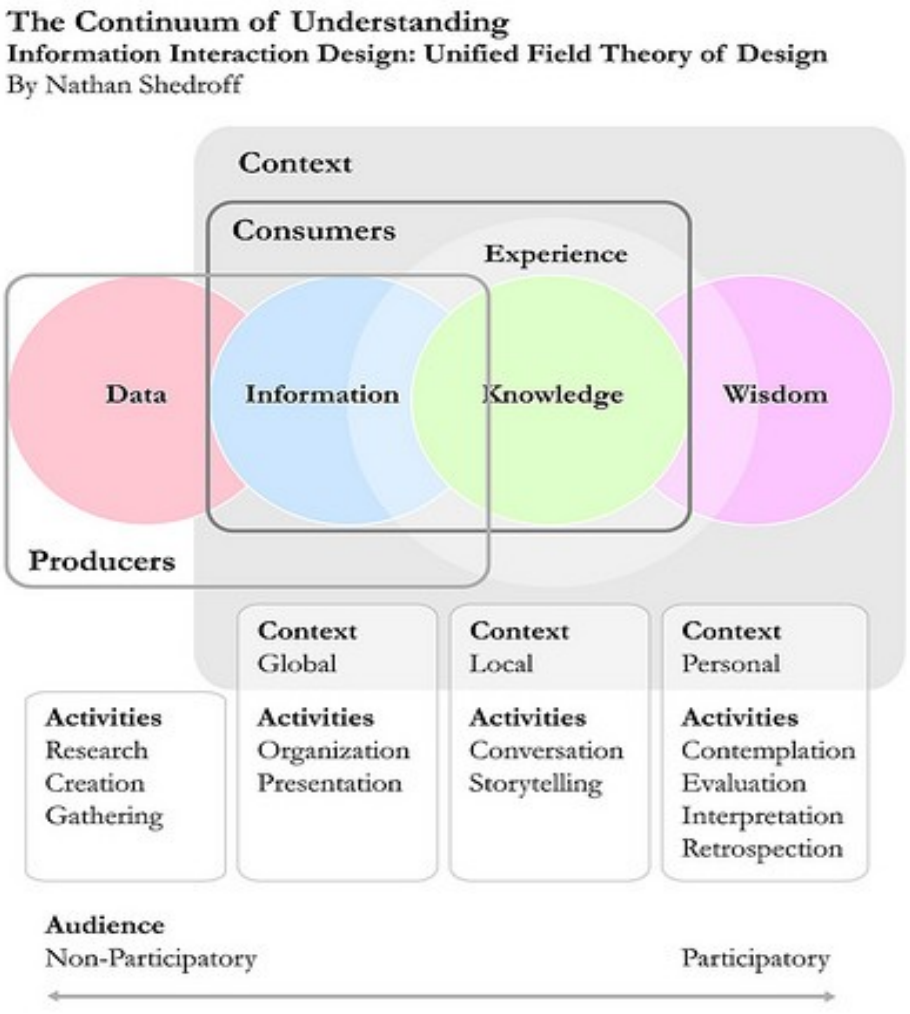


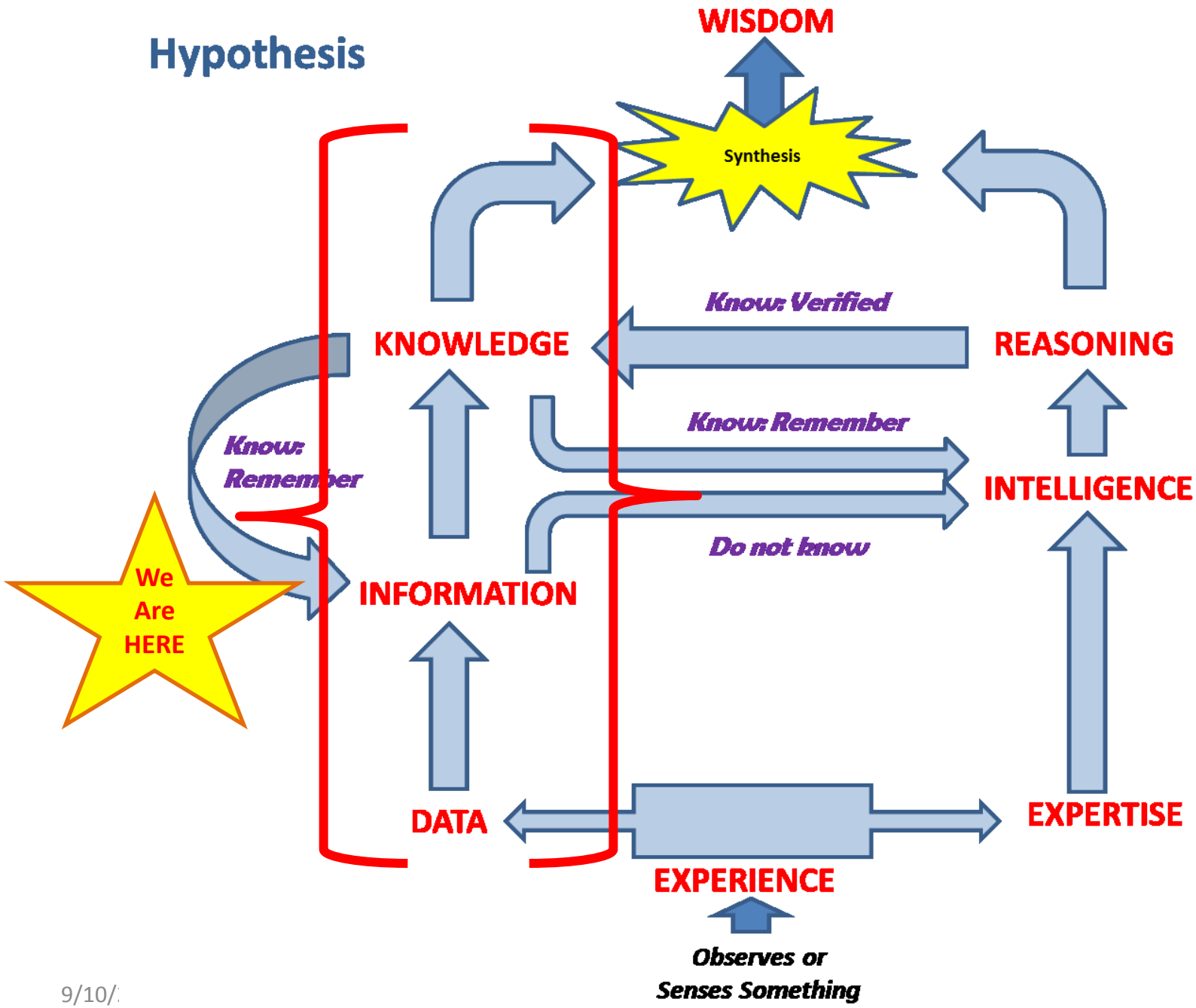
Source of percentages: http://www.allianza.com.au/pdf/Discovery_paper_Brief.pdf

A problem in search of a miracle…



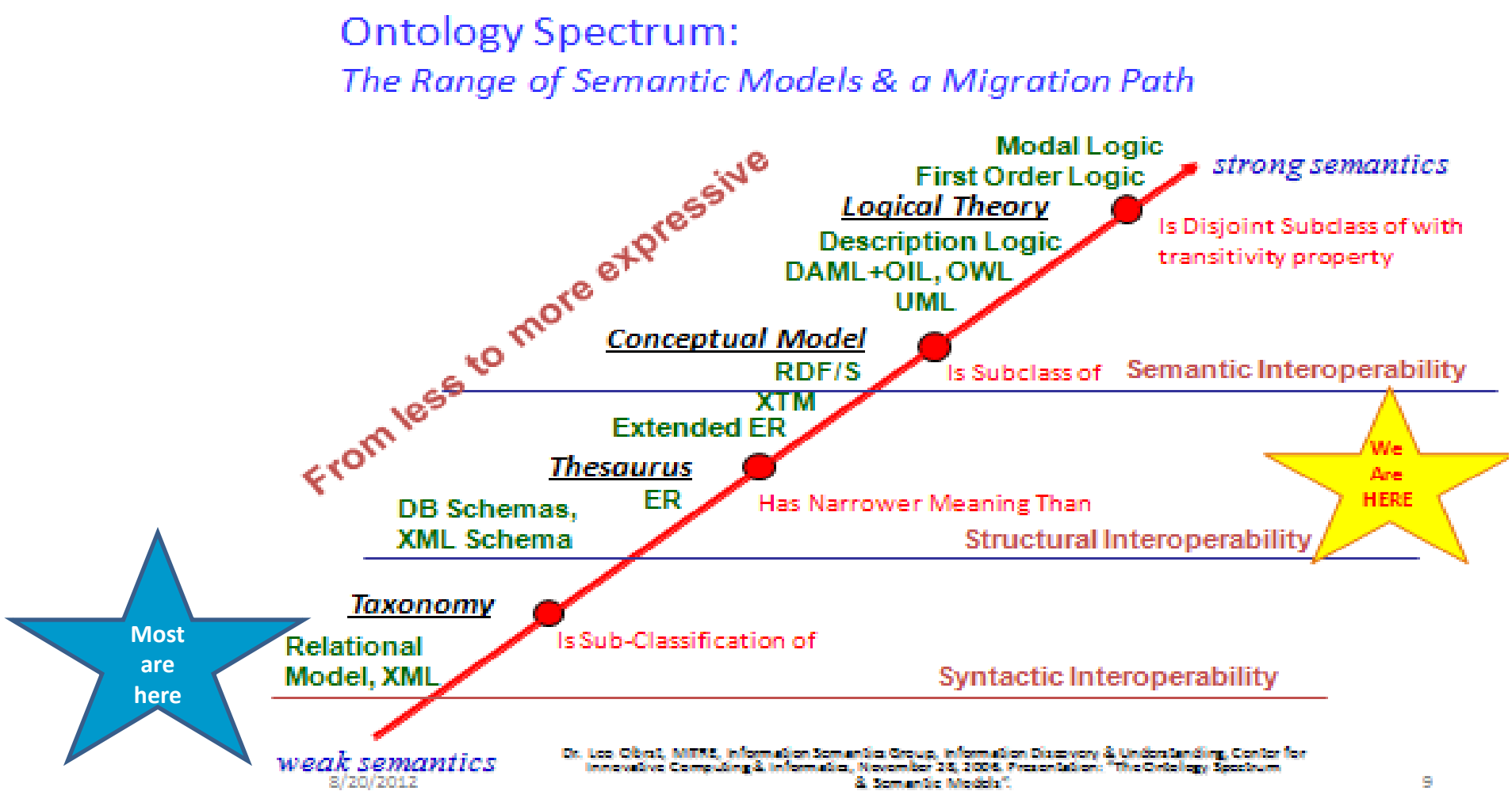
Understanding our perspectives




Hypothesis
WISDOM
Synthesis
Know: Verified
KNOWLEDGE
REASONING
Know: Remember
Know: Remember
INTELLIGENCE
Do not know
We Are HERE
INFORMATION
DATA
EXPERTISE
EXPERIENCE
Observes or Senses Something

The train is leaving the station…

The path to semantics



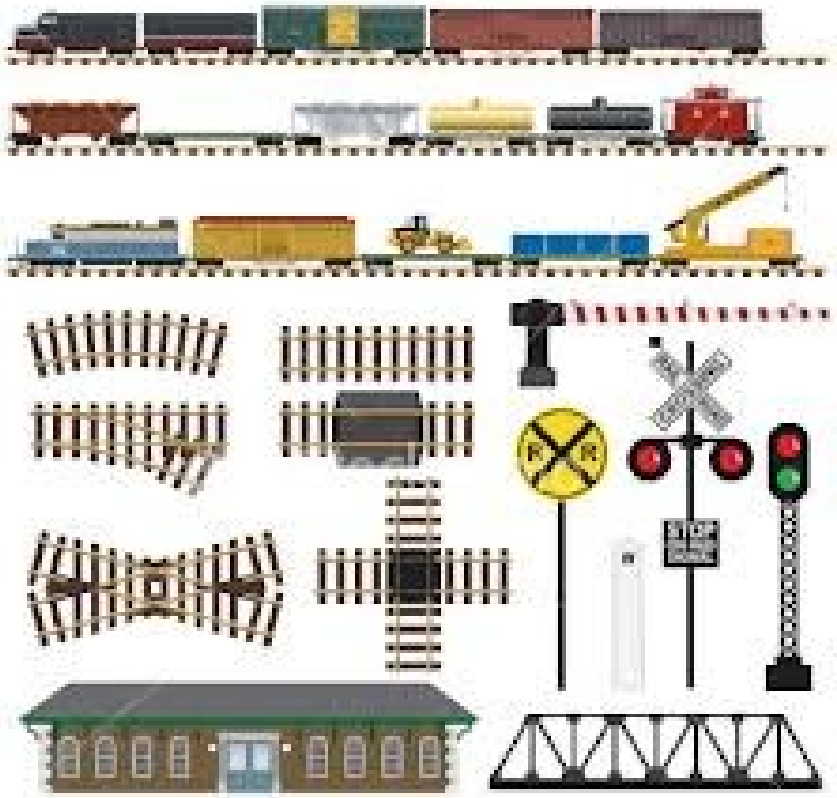
What is the requirement?

Requirements?

- **Aviation System Block Upgrades (ASBU):**
  - Requires the preparation of the Information to cover "both the semantic and syntactic aspects of data composing information and the information management functions."
- **Document Management (JPAMS):**
  - Requires the discovery of document components and the ability to tag content.
- **Service Discovery & Delivery (OGC – OWS-9)**

Working Document for the Aviation System Block Upgrades – 16 Nov 2011

What have we built so far?



What are the primary components?

1) **Terms:**
   - **Lexicon of Terms**
   - **Thesaurus**
2) **Concepts:**
   - **Conceptual Model**
3) **Relationships:**
   - **Taxonomy (Hierarchical)**
   - **Ontology (Hierarchical with Horizontal Relationships)**
   - **Logical Theory/ Business Rules**

What have we done so far?

1) Terms:

- ✓ **Lexicon of Terms** (using FAA, Eurocontrol, ICAO docs)
- ✓ **Thesaurus** (using FAA, Eurocontrol, ICAO and standard acronyms, abbreviations, synonyms, codes, etc.)

2) Concepts:

- ✓ **Conceptual Model** (classes and categories in spreadsheets)

3) Relationships:

- ▪ **Taxonomy (Hierarchical)**
- ✓ **Ontology (Hierarchical with Horizontal Relationships) –** (currently populating in TopBraid Composer Maestro tool)
- ▪ **Logical Theory/ Business Rules**

All for the Air Transportation Domain

Different Perspectives

All need to be defined

FAA's AI Domain

Information Categories (Architecture)

- AERONAUTICAL INFORMATION
- TOURIST ATTRACTION
- OPERATIONAL THEATRE CONFIG.
- NAS CONFIGURATION INFORMATION
- MARITIME SYSTEM CONFIGURATION
- HISTORICAL STRUCTURE
- MILITARY TARGET
- AVIATION OBSTACLE
- MARINE NAVIGATION AID

Data Subjects

- LIGHTHOUSE
  - ONSHORE LIGHTHOUSE
  - OFFSHORE LIGHTHOUSE
    - PIER BREAKWATER
    - TEXAS TOWERS

DATA:
44701 State Highway 498
Piney Point
MD
(301) 994-1471
Onshore Lighthouse

We are stoking the fire using ontologies

How is this different from a traditional approach?

Different Approaches

Traditional Approaches

- Are **Brittle**, new annotations to existing content/knowledgebase require modification to schemas
- **Does not permit inference*** among existing facts, they must be explicitly stated
- **Rules must be coded and are buried in the technology**
- **New relationships are made in a controlled environment** that is slow, labor-intensive, and costly

Semantic Approaches

- Are **Flexible**, new annotations can be added to existing content/knowledgebase without modification to schemas
- **Permits inference*** among existing facts
- **Rules can be added at any time and exist independent from technology**
- **New relationships can be made extemporaneously**

*Inference is the ability to derive new information from existing information, and requires application of rules or logic constructs.

EUROCONTROL

Federal Aviation Administration

Different Approaches (cont.)

Traditional Approaches

- **Data from different sources can be merged / integrated, but require a new/modified schema**
- **New connections to different systems require extensive pre-coordination and cooperation** among involved parties
- **Enables continuous improvement,** (but takes a long time) and it affects legacy applications

Semantic Approaches

- **Data from different sources can be merged / integrated, without a new schema**
- **New connections to different systems can be made easily**
- **Enables continuous improvement,** (almost immediately) without affecting legacy applications

What are we doing now?

Currently we are…

Developing:

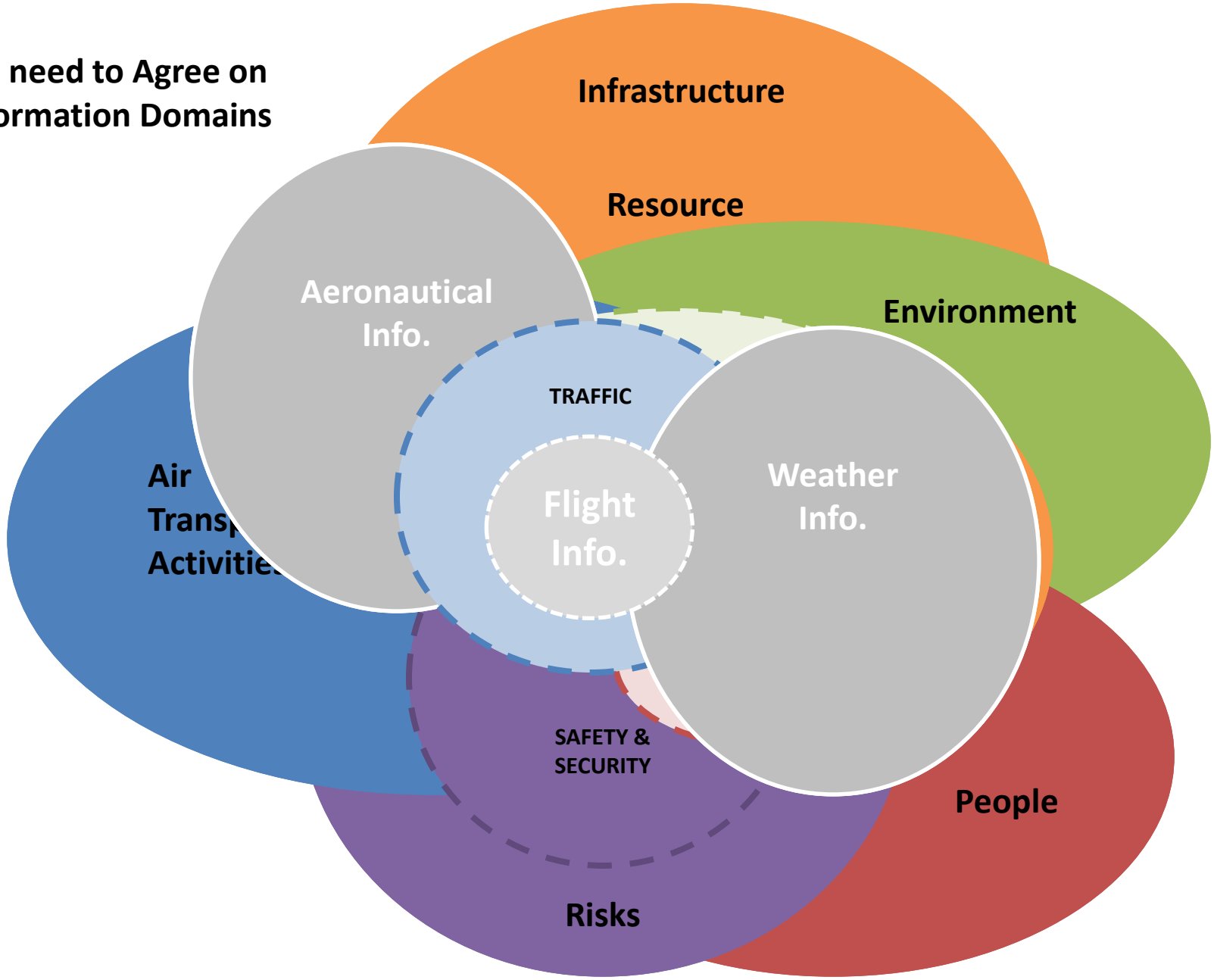
- **Categorization scheme for:**
  - **Tagging content** (manually, then automated)
  - **Finding content** (once it has been tagged)
  - **Retrieving content**
  - **Finding information services** (by human or machine)
- **Vocabulary/Thesaurus Management to:**
  - **Enable term and definition look-up**
  - **Capture and retrieval of information and "meta-data" about a term** (in a central repository)

Where do we go from here?

What's Next?

- Agree on Information Domains
- Form/Establish Information Domain Communities (if not already established) including governance
- Determine linkage of Information Domains to current data models (OV-7 & AIRM)
- Collaborate on these initiatives/efforts to harmonize global air transportation information!

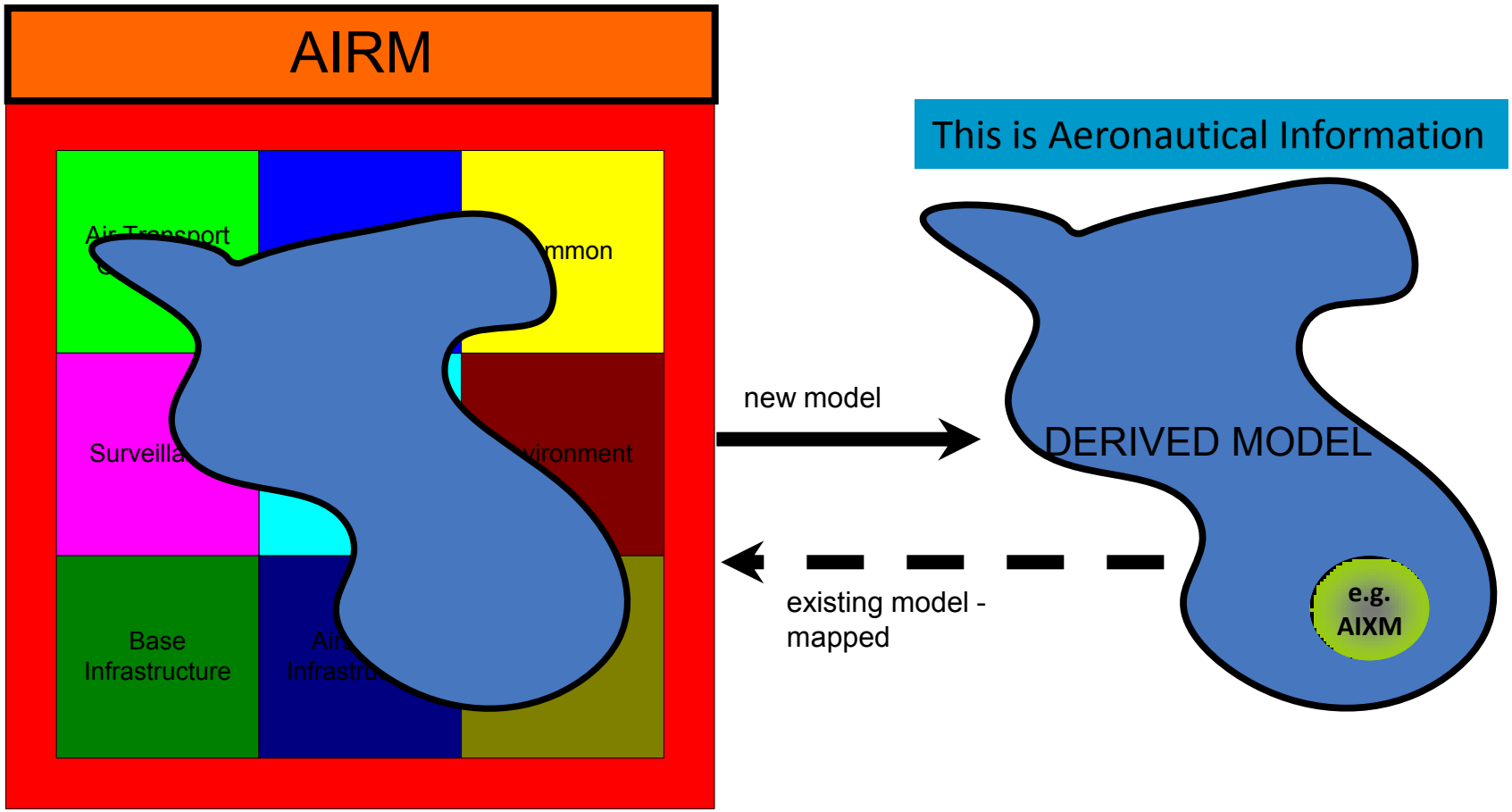
We need to Agree on Information Domains



What about Information Domains?

- **Information Domains need to be managed within Communities.**
- **The Information Domain Process begins to describe how to manage the domains within communities and across domains including oversight and governance.**
- **Additional sections describe differences between Information Management, Information Oversight, and SWIM.**
- **There are roles/responsibilities described for community members, governance body, management council and regulator.**

Shows Nature of Information Domain…



Establish understanding/agreement by:

- Acknowledging by their nature, that they overlap
- Understanding that they are not just discrete data subject areas (classes)
- Identifying an information domain and sub-domain by certain criterion
- Evaluating them by certain quality criterion

What is the Risk?

We will never get this right…

if we don't understand/agree on the **major *business* information domains required to support Air Traffic Management Activities**

How will we support the business community?

How will all of this support the business community?

For NextGen, we have agreed to develop and support:

- Common Operating Picture
- Business and Performance Analytics
- Situational Analysis
- Collaborative Decision Making

Semantics will enable us to:

- **Be clear** about what we mean.
- **Express, find and use information.**
- Build and extend the foundation for **Knowledge Management** (declarative knowledge).
- Perform **inferences** (procedural knowledge/intelligence).
- Specify **business rules** (procedural knowledge/reasoning).
- Identify **patterns and trends** (reasoning and intelligence).

Combine all these to support analytics, common operating picture, situational awareness, and collaborative decision-making

We have reached our destination!

Questions

Contact Information

Air Transportation Information Exchange Conference - (featuring AIXM, WXXM and FIXM)

Deborah Cowell, Deborah.Cowell@faa.gov
202-385-7077

Dr. Candice Buchanan,
Candice.CTR.Buchanan@faa.gov
202-385-7095

Backup Slides

The relationship between XML and RDF[1]:

Not so much

XML	RDF
• **Provides limited semantics and IS ambiguous** – XML is the first step to ensuring computers can communicate • **Like the alphabet, but it is not a language** • **Best to share data and exchange information between different platforms and applications**	• **Has richer semantics and is NOT ambiguous** – It is a standard that can be "understood" AND permit inferences by computers • **Is an ontology language that expresses concepts and relationships** • **Best to express, find, and use information and knowledge obtained from different sources and produced by different platforms and applications**

[1] Yi, L. (2011). A Developer's Guide to the Semantic Web. Springer: Heidelberg.

The relationship between XML and RDF[1]:

Not so much

XML

- **Parsing XML statements depends on the tree structure**
  - **Not scalable** on a global basis
  - **Structures are hard to handle**, especially in large amounts
  - Does not provide what we need for construction of the Semantic Web

RDF

- **Has a very simple data structure – RDF graphs**
  - Is scalable for large datasets
  - Graphs can easily be converted in to Statements or Triples
- **Able to break info and knowledge into smaller pieces**

  SUBJECT PREDICATE OBJECT
  - **Each piece has its own semantics so…**
- **Represents and models info and knowledge so that machines can understand it and use it to do useful things**

[1] Yi, L. (2011). A Developer's Guide to the Semantic Web. Springer: Heidelberg.