

Using ORM in an Ontology Based Approach for a Common Mapping Across Heterogeneous Applications

Baba Piprani

SICOM Canada, Ottawa

Use Case: Transport Canada Aviation Object Dictionary

International Workshop on Object Role Modelling, ORM 2007
Vilamoura, Algarve, Portugal November 2007



“ORM 2007”



Baba Piprani

Consultant – SICOM Canada

- Over 30 years of implementation experience and involvement in NIAM/ORM, Metadata/Database Standards and Data Management
- Canadian Delegate in ISO Database Language SQL (SC32 / WG3), and Metadata Working Groups (SC32 / WG2)
- Past Chair of ISO Conceptual Schema WG and Govt. of Canada Treasury Board IT Standards (TBITS SQL)
- Highly successful (Always successful !) Data Warehouse and metadata based implementations
- Consultant to: Government of Canada Transport Canada; Office of Superintendent of Financial Institutions on Metadata Repository; Corporate Metadata Strategy and Ontology based Aviation Object Dictionary; China National Institute of Standardization
- Author of ‘Using ORM-based Models as a Foundation for a Data Quality Firewall in an Advanced Generation Data Warehouse’ (ORM Workshop, Montpellier, Nov 2006; and Journal of Semantics of Data JODS XI)



Abstract:

- Application Integration efforts face:
 - Lack of common data semantics
 - Conflicting or ambiguous meanings
 - Massive duplication of master data
- Use Case at Transport Canada demonstrates use of
 - Ontology based solution using ORM
 - Using ISO standards IS 19763, IS 11179 and IS 9075
- Enables dynamic integration ON DEMAND!!!
- Common semantics bridged via ontological concepts



Agenda

1. Current Integration Problems
2. Ontology based Approach using ORM
3. Use Case: Aviation Object Dictionary
4. Knowledge-base facilitated dynamic Integration of Aviation Object Dictionary (AOD)
5. Summary



'Helsinki Principle' from ISO TR9007

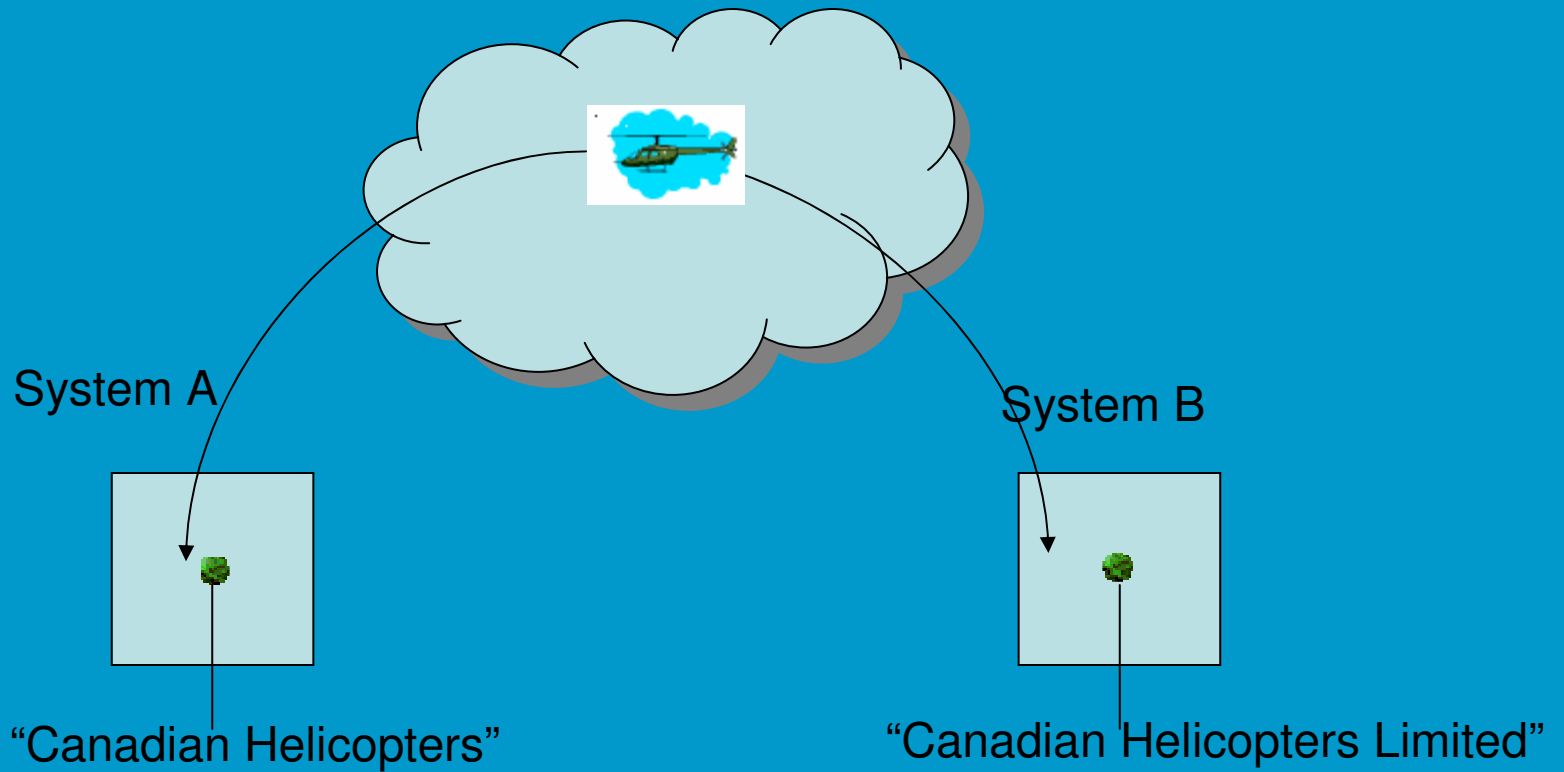
Note: These utterances are to be interpreted (recursively) as international English utterances:

“Any meaningful exchange of utterances depends upon the prior existence of an agreed set of semantic and syntactic rules. The recipients of the utterances must use only these rules to interpret the received utterances, if it is to mean the same as that which was meant by the utterer.”



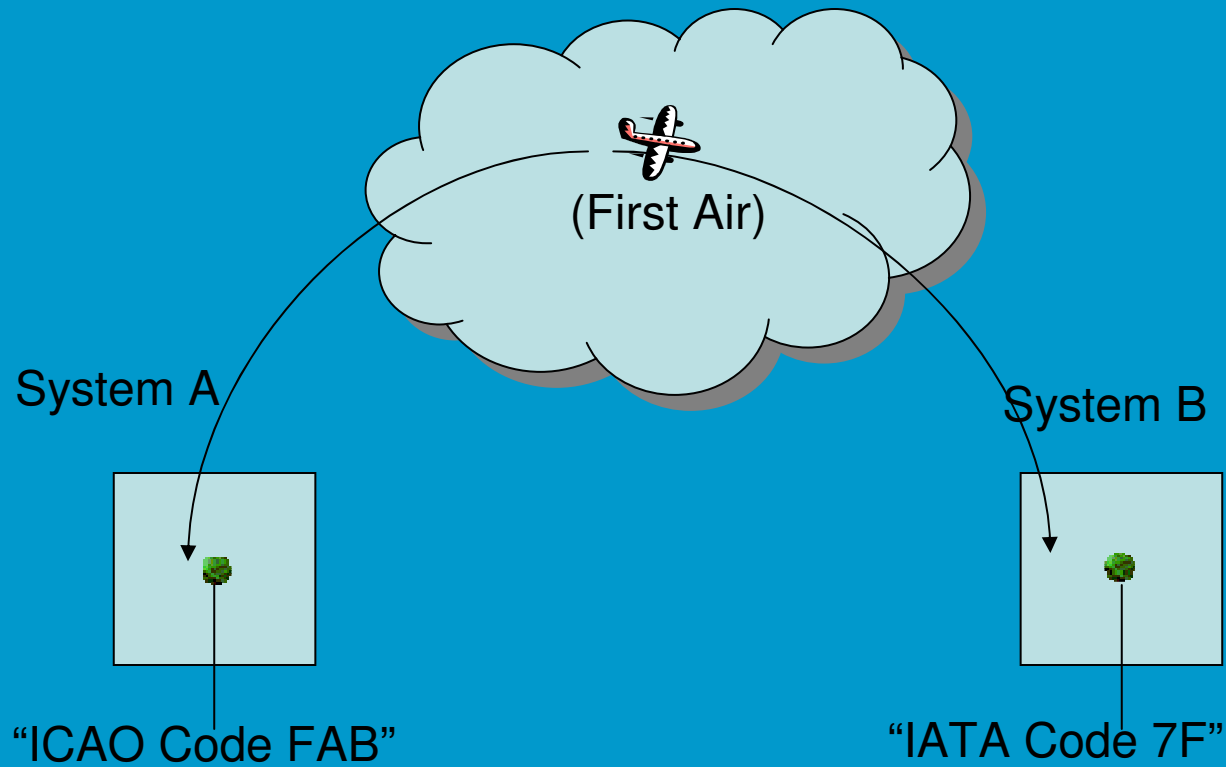
Common Problems in Integration...

Entity Naming



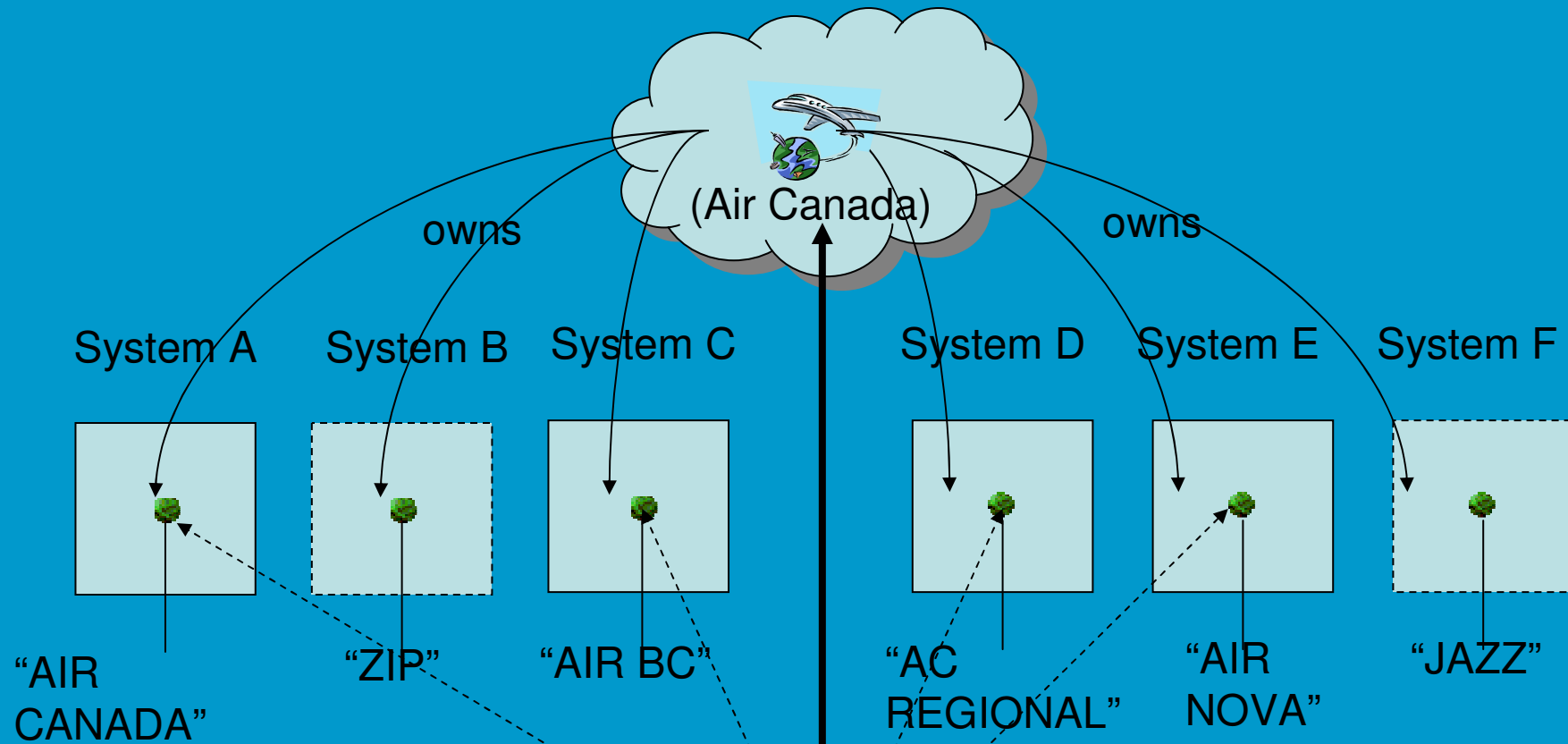
Common Problems in Integration...

Entity Identification

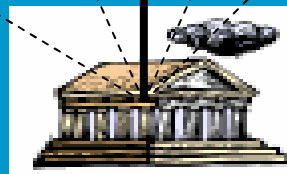


Common Problems in Integration...

Entity Transparency



**TC: Occurrences concerning
all operations of Air Canada
Prior 14/3/2002?**

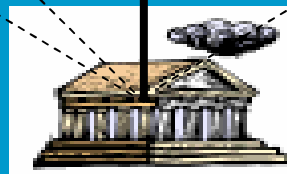
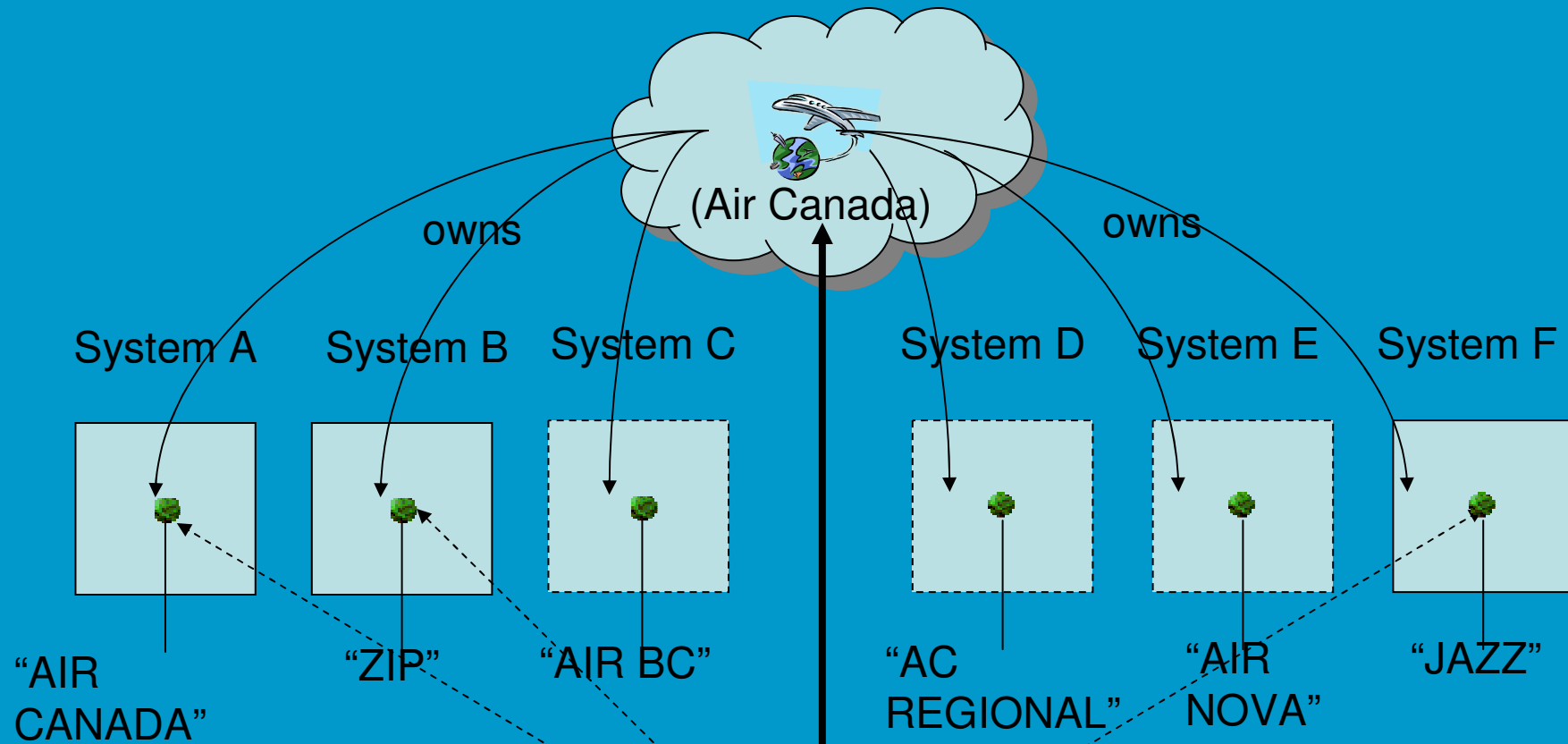


"ORM 2007"



Common Problems in Integration...

Entity Transparency



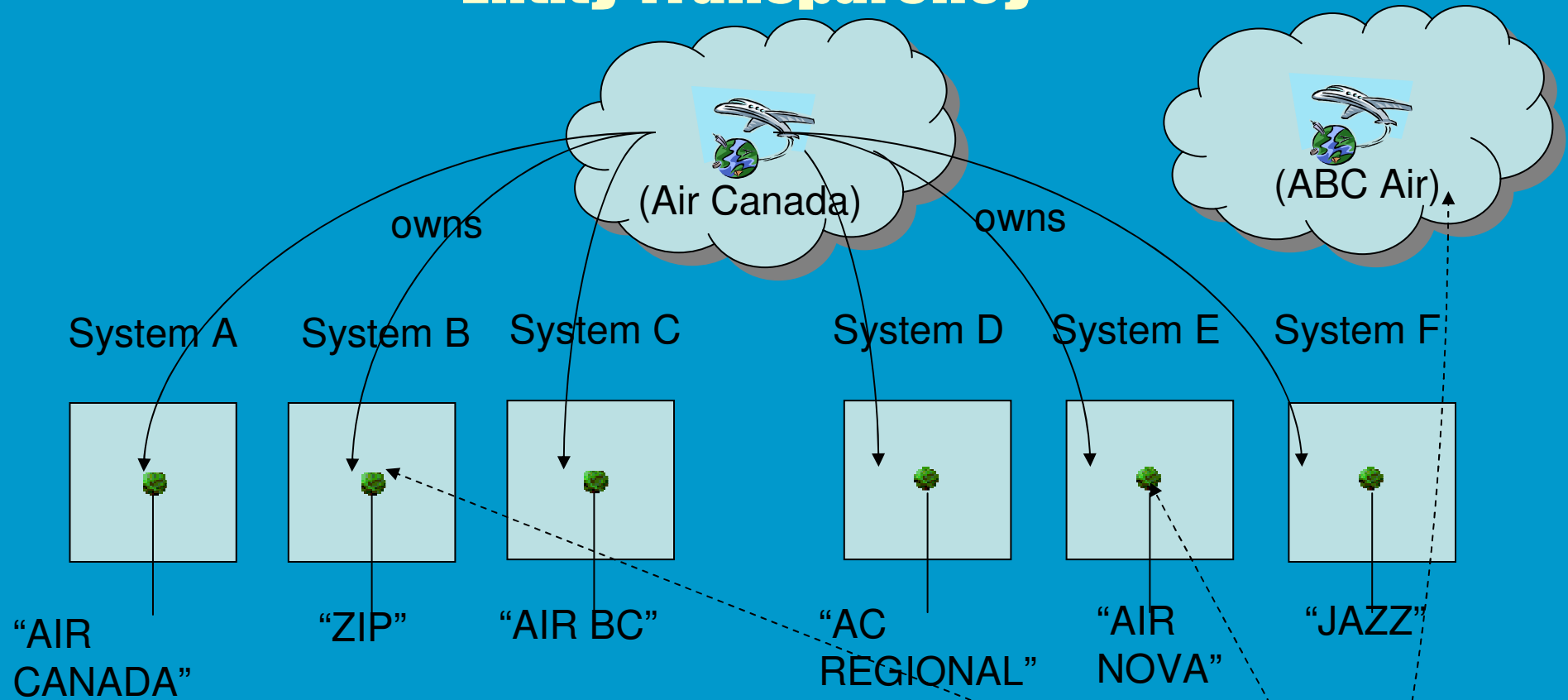
"ORM 2007"

TC: Occurrences concerning all operations of Air Canada After 14/3/2002?



Common Problems in Integration...

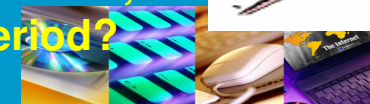
Entity Transparency



TC: Occurrences concerning particular aircraft and who owns it, in what period?



"ORM 2007"



Transport Canada (TC) Corporate Horizon



Manufacturers



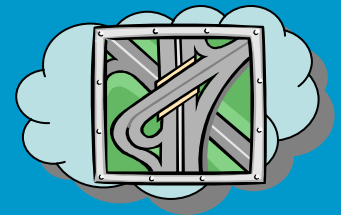
Regulatory
Agencies



Operators

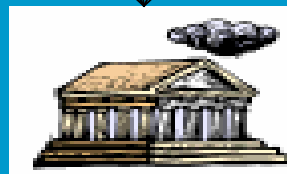


Legislation



Infrastructure

Data relationships need to be captured, applied and managed for TC mandate



"ORM 2007"

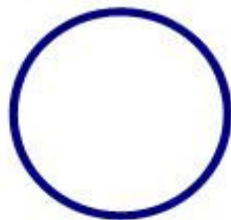


Situation

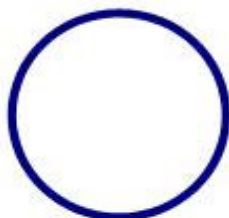
- Not able to share data
- Data searches require convoluted processes
- Official ownership of Aviation Objects are not defined
- Problem was much larger than it appears at the surface
- Serious lack of data quality with high accompanying costs



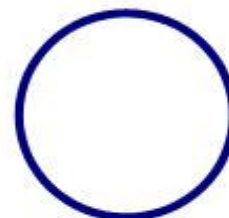
Strategies and Integration's
Aviation Dialect



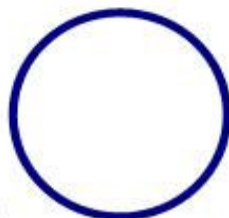
Civil Aviation's
Aviation Dialect



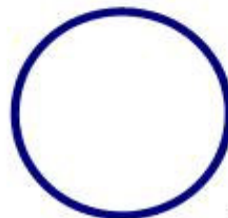
Policy's Aviation
Dialect



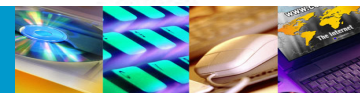
Information
topology



FAA's (US)
Aviation Dialect



TSB's Aviation
Dialect



Strategies and Integration's
Aviation Dialect

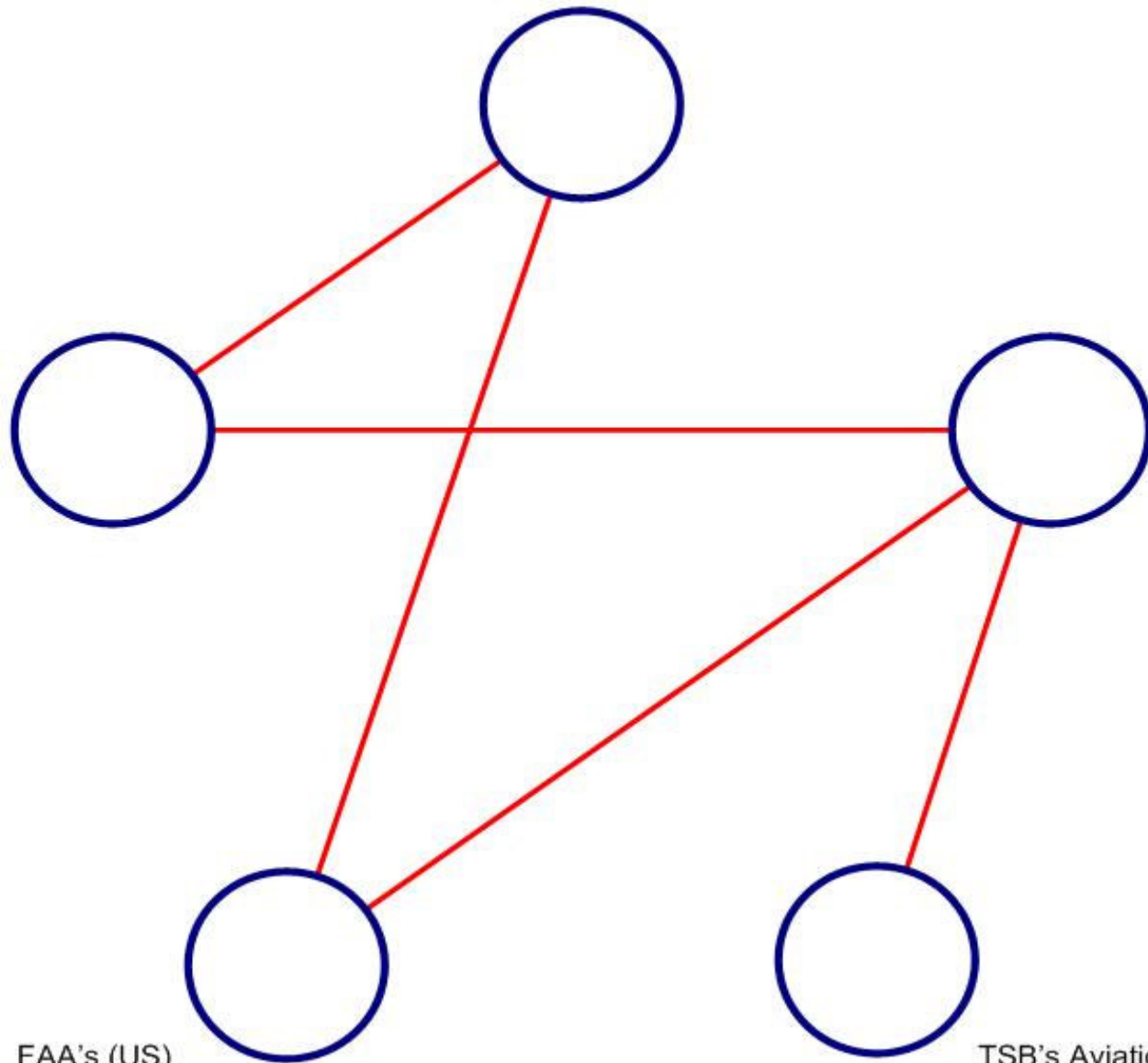
Civil Aviation's
Aviation Dialect

Policy's Aviation
Dialect

Information
topology

FAA's (US)
Aviation Dialect

TSB's Aviation
Dialect



Strategies and Integration's
Aviation Dialect

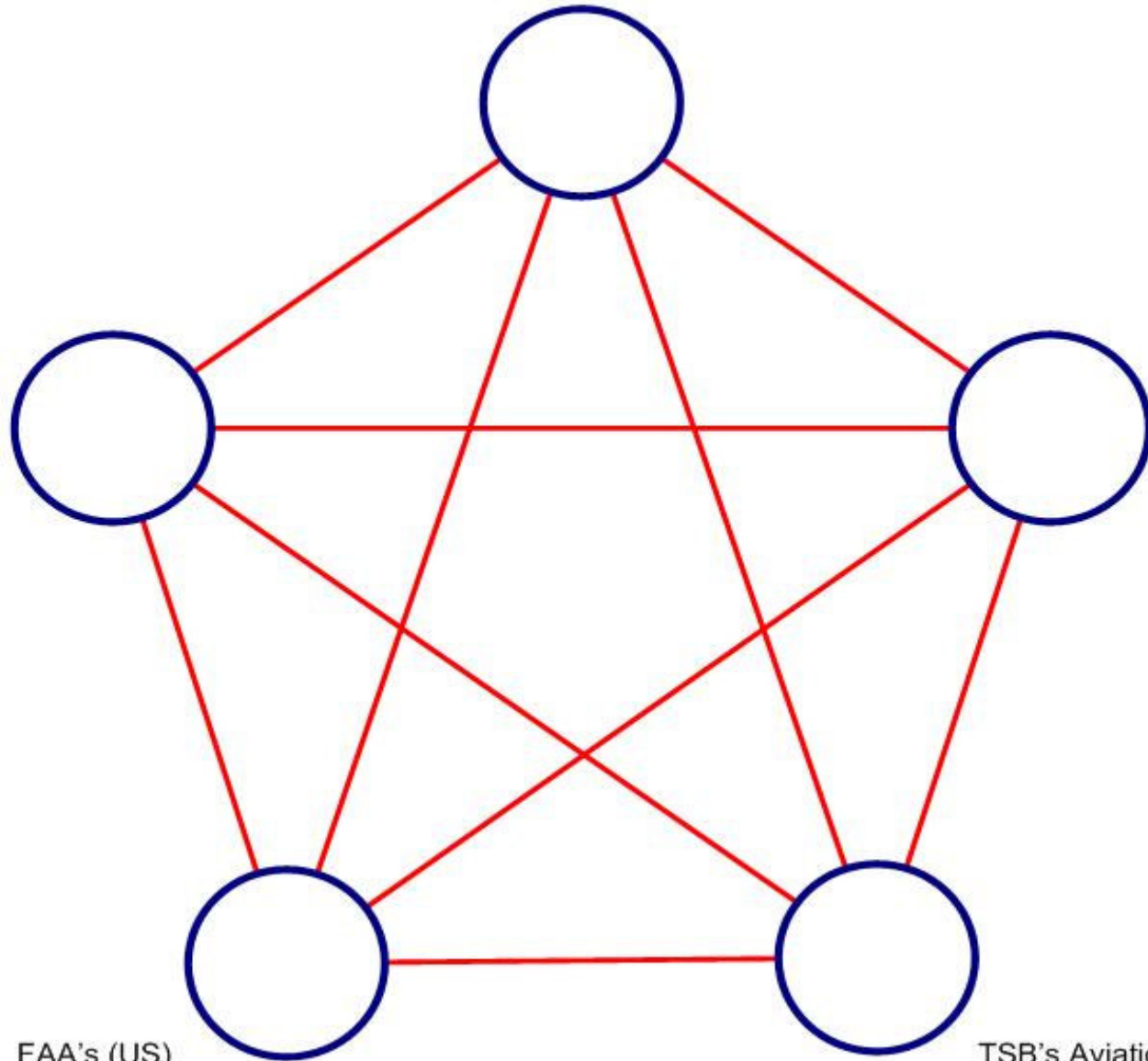
Civil Aviation's
Aviation Dialect

Policy's Aviation
Dialect

Information
topology

FAA's (US)
Aviation Dialect

TSB's Aviation
Dialect



Defined Policies Needed To:

- Access to various air related data is becoming increasingly important due to raised level of security and safety requirements in the air sector (especially since 9/11)
- Requirement to access data from various angles or dimensions is seen as imperative to be able to analyze various crises situations



Enter...Aviation Object Dictionary (AOD)

- Mandated to establish logical relationships between TC Air Information Systems
- Enable logical integration of existing systems through standards
- Establish TC standards for defining and naming of Aviation Objects
- Enable harmonization of various identification schemes and industry standards (ICAO etc) for TC use



Knowledge based concepts:

Goal: Allow unrelated applications to share a common pool of concept-based synchronized data

How to Achieve?

Create and maintain mechanisms to enable an enterprise-wide “single version of record” for the core business entities through concept-based links.



What to Achieve:

- Develop a centralized registry of data elements for standardized objects w.r.t. naming and meaning within Civil Aviation. e.g. objects:
 - Carrier, Aircraft, Airport, Contacts , Operational, Occurrence and Regulatory
- Develop data element mapping standards to facilitate integration of existing systems towards consistent analysis, reporting and data entry



Involved systems...

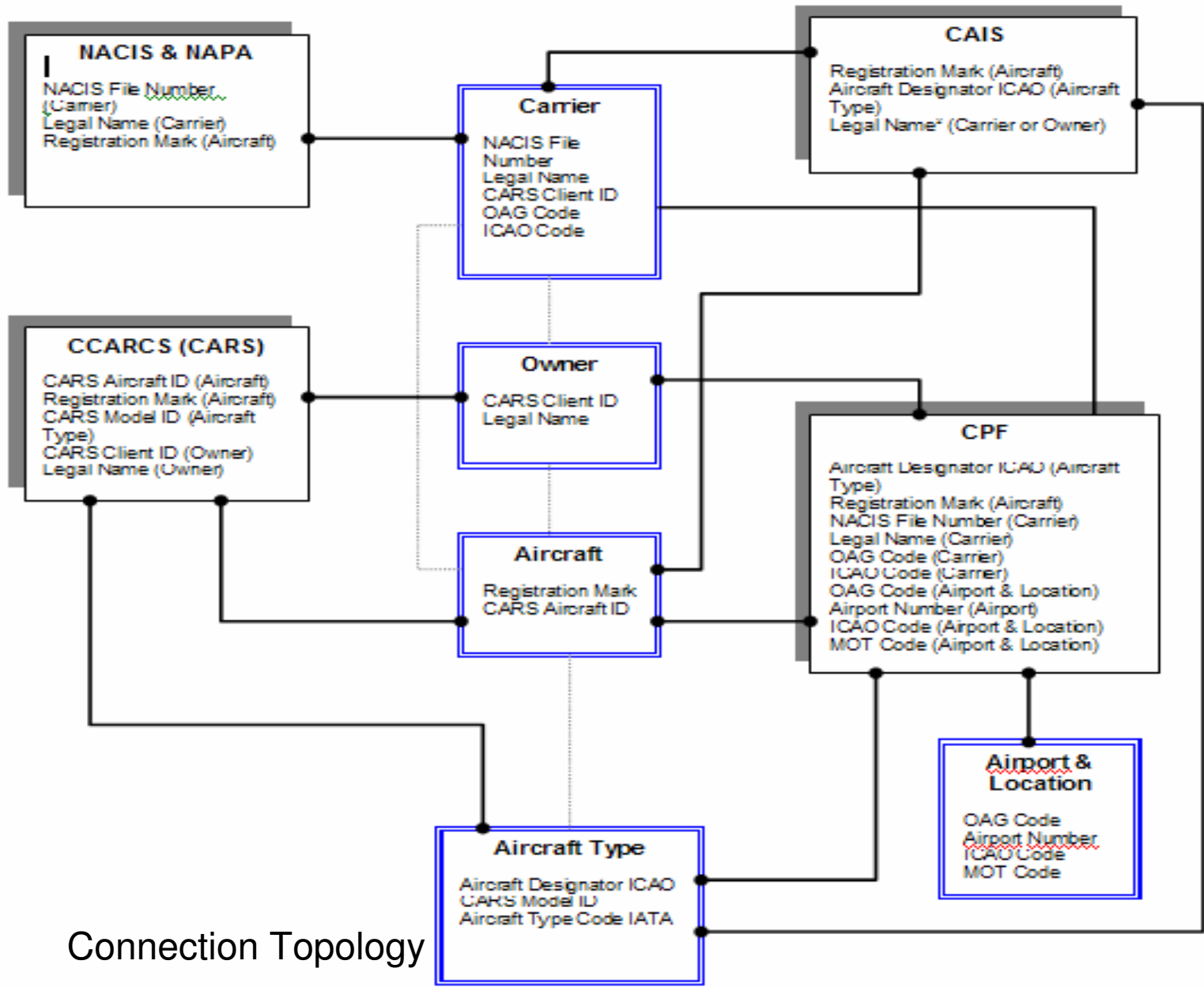
CCARCS	Canadian Civil Aircraft Register Computer System
CAWIS	Canadian Air Worthiness Information System
NACIS	National Aviation Company Information System
NAPA	National Aeronautical Product Approval Information System
CAIS	Computerized Airworthiness Information System
CPF	Cyclopedia Parameter Files System
ECATS	Electronic Collection of Air Transportation Statistics
AIRPORT- FINANCIALS	Financial information on Canadian Airports
GIFA	General Information From Airports
...and later to Other modes (Marine, Road, Rail, Dangerous Goods)	



Operational Problems:

- **Mappings not always available** (e.g. a given location may have an IATA code, and and ICAO code, sometimes only IATA, sometimes only CFS code, sometimes FAA code etc., hence cannot map across systems)
- **Inconsistent Semantics** (e.g. add to above:
 - inconsistent identification and different terminology for referencing the same concept
 - rounding errors for lat/long produced multiple airports at same lat/long, where in reality they were at other ends of runway or were 2 heliports at a hospital on adjacent buildings)



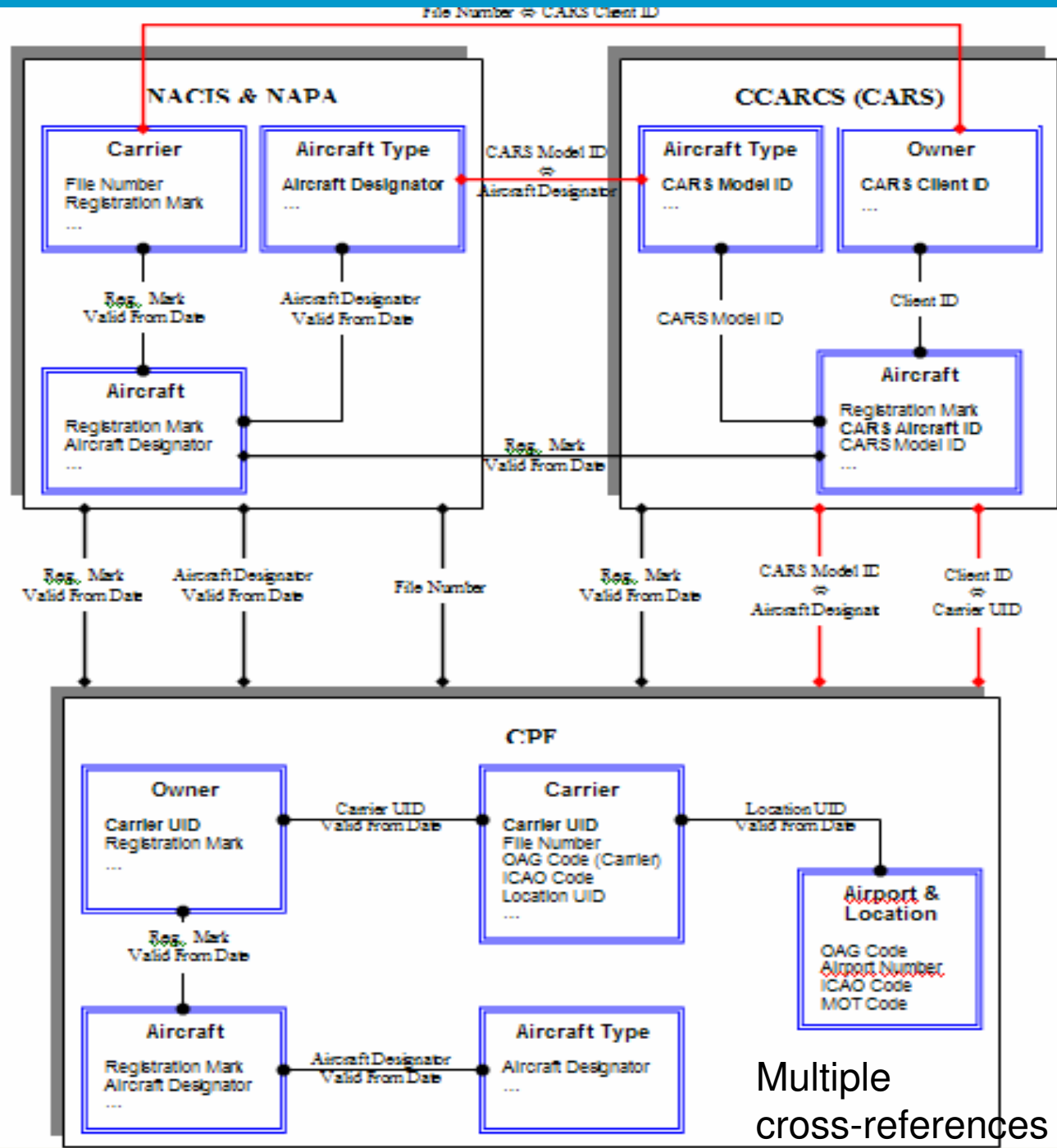


Connection Topology

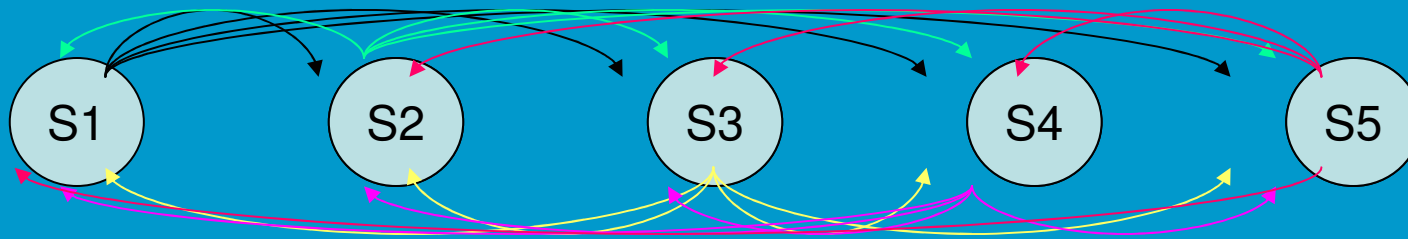
Previous Work:

- Originally proposed solution was to create cross-reference tables between each of the participating systems e.g. NACIS to CCARCS, NACIS to CARS and so on
- Depending on the concepts involved and the keys involved, this approach would exponentially increase mappings e.g. 5 systems will result in 20 mappings, one between each of the other systems





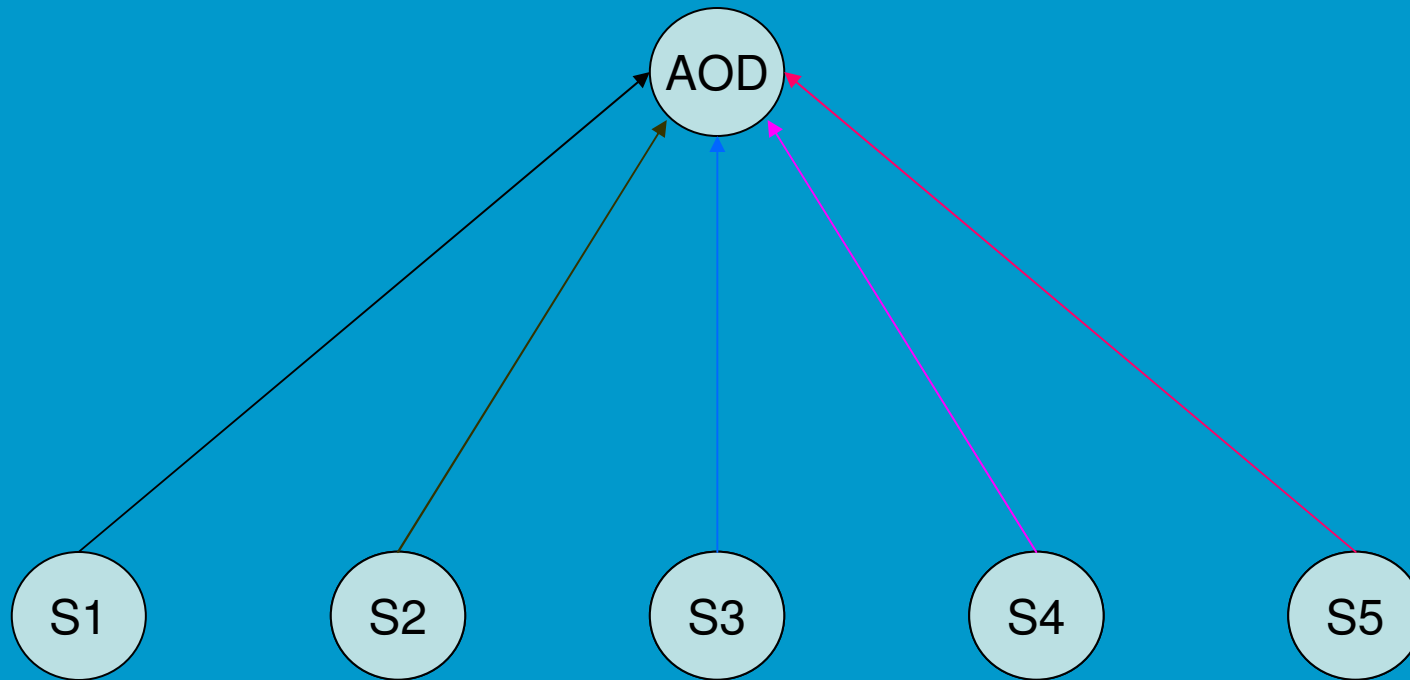
Mapping Requirements for 5 Systems – Direct Mapping (as per report)



MAPPINGS REQUIRED = $n(n-1) = 5(4) = 20$



Mapping Requirements for 5 Systems – AOD



MAPPINGS REQUIRED = 5



Solution: Apply knowledge based concepts

- Apply knowledge based concepts to create an Aviation Object Dictionary (AOD)
- The AOD challenge:
 - To define and merge all the disparate, often-conflicting 'identifying data set' of the main business entities into one authenticated master version.
 - To create common methods and processes to use the 'identifying information set' from within all the applications about the main business entities, and to maintain an accurate version of master data for identification to be able to provide the ability to cross-relate applications.”



Apply Knowledge based concepts

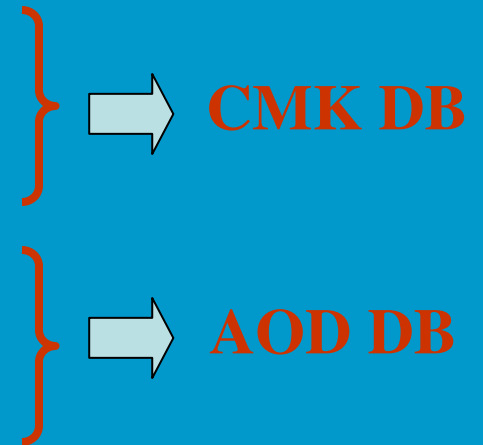
AOD Two pronged Approach:
to address: Mappings not available and
Inconsistent semantics

- 1) Mapping correlation of
 - a) Metadata
 - b) Values
- 2) Master Data Set
(for identification and usage)

Legend:

CMK → Common Mapping Key (for addressing Mapping problem)

AOD → Aviation Object Dictionary (for addressing Semantics problem)



AOD Approach

The overall approach of the AOD is to:

apply Semantic Web based techniques in enabling end-user applications to

- dynamically integrate,
- virtually combine,
- derive or deduce information

from:

- several heterogeneous application aviation databases
- *multi modal transport (rail/marine/road) databases (future).*



AOD Approach

- Define Semantic links to establish relationships between different sets of data based on common ontological terms
- Ontology
 - ‘what exists’, ‘what is’ ‘something that is’
 - an explicit specification of a conceptualization.



Semantic Model

- Describe the ontology by defining and associating a set of objects or entities
- Provide definitions that associate the entities in the Universe of Discourse

**Non
Lexical**

-
- Describe a set of representational terms
 - Provide definitions what the names are meant to denote

Lexical



Semantic Model

- NON-LEXICAL OBJECT level

- Describe the ontology by defining and associating a set of objects or entities
- Provide definitions that associate the entities in the Universe of Discourse

- LEXICAL OBJECT level

- Describe a set of representational terms for the non-lexical object
- Provide definitions what the names are meant to denote

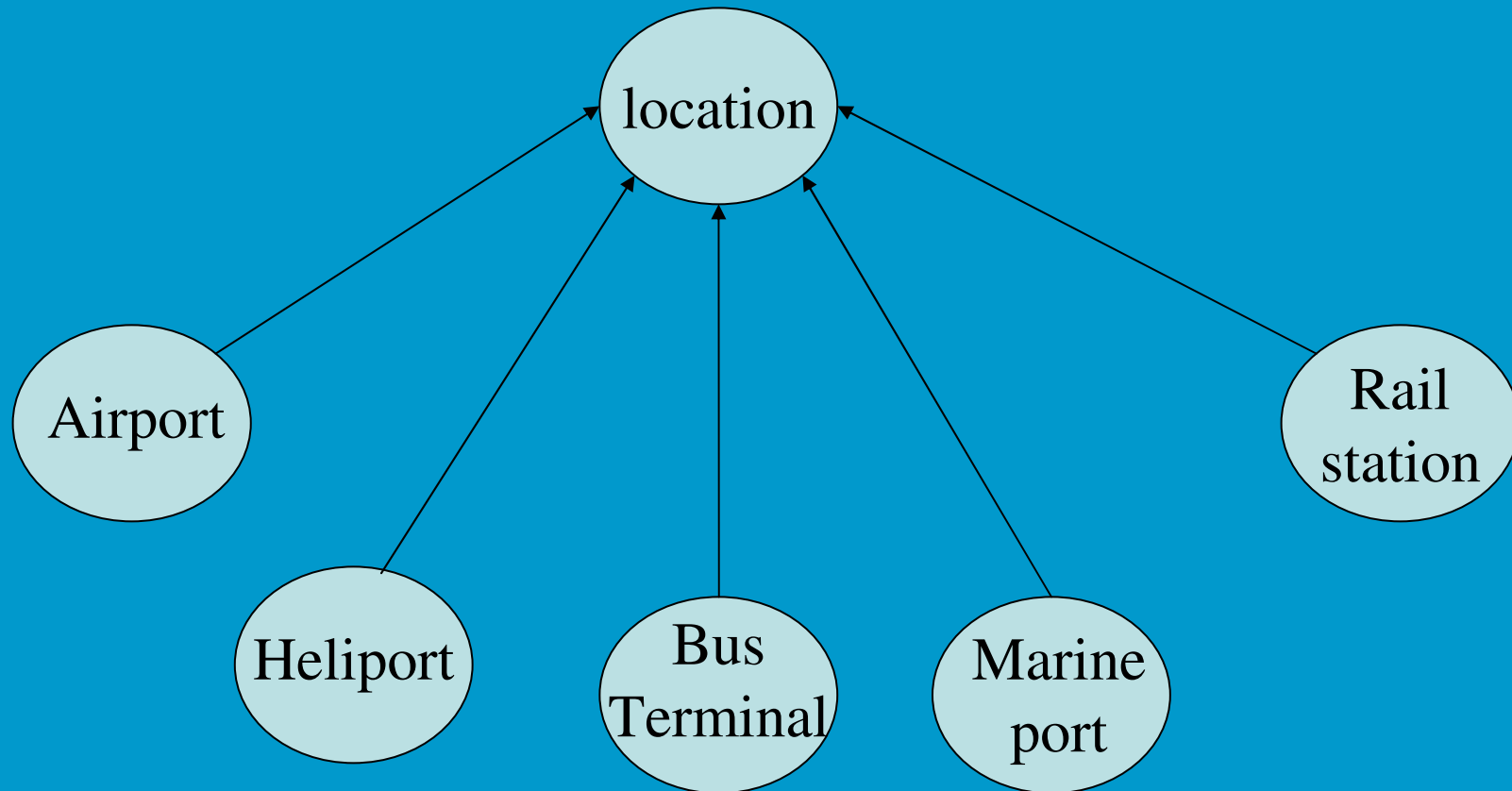


Relating objects:Non-Lexical Object Types

- **Case1: an ontology describes a hierarchy of concepts related by subsumption relationships;**
- Case2: suitable axioms are added in order to express other relationships between concepts and to constrain their intended interpretation



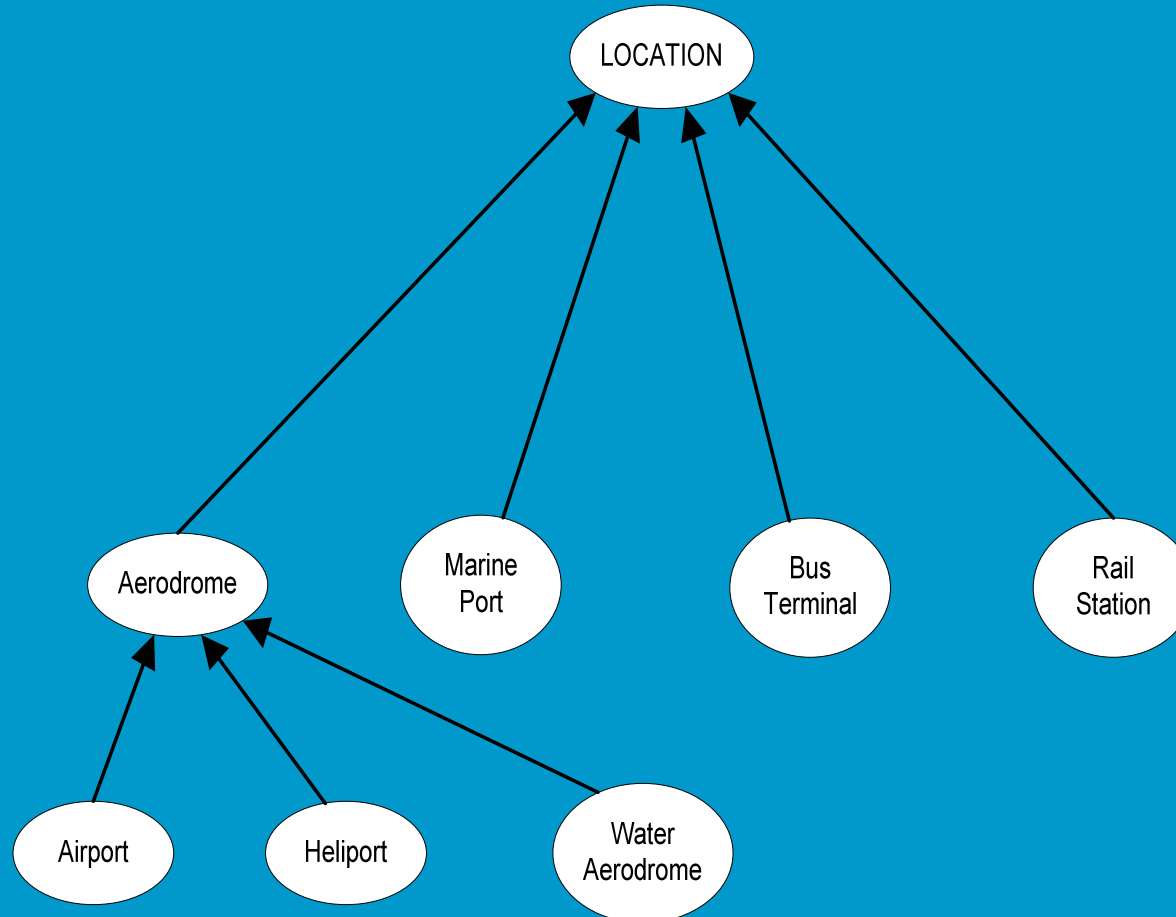
Case 1: Hierarchy of Concepts



Case 1: hierarchy



AOD – Ontology Topology for 'Location' as NIAM/ORM subtypes



Relating objects :Non-Lexical Object Types

- Case1: an ontology describes a hierarchy of concepts related by subsumption relationships;
- Case2: suitable axioms are added in order to express other relationships between concepts and to constrain their intended interpretation



Case 2 – relationships between concepts

Sentence Type 1:

Application CPF contains concept CARRIER with sub-concept IDENTITY having case ICAO-CODE

Sentence Type 2:

Schema CPFDB has Table CARRIERUID with column UID



Case 2 – relationships between concepts

Sentence Type 3:

[Sentence Type 2] is contained in [Sentence type 1]

i.e. [*Application* CPF contains *concept* CARRIER with *sub-concept* IDENTITY having *case* ICAO-CODE]

is contained in

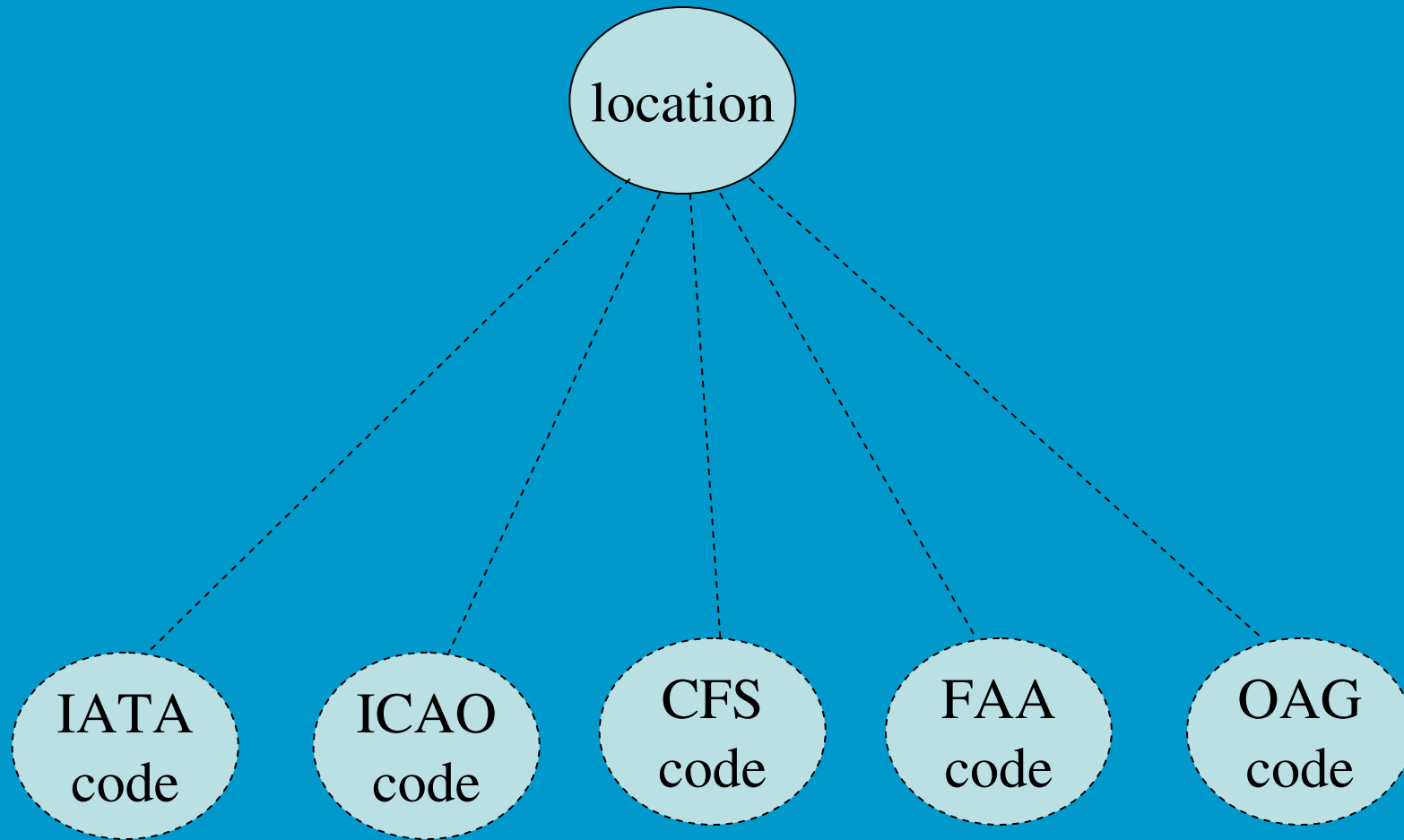
[*Schema* CPFDB has *table* CARRIERUID with *column* UID]



Relating objects:Lexical Object Types

- Case1: relationships between concepts and their representations;
- Case2: suitable axioms are added in order to constrain their intended representational interpretation





Case 1: Representations



“The Thing”...and ”The name of the Thing”

- Popular OO models, ER models, Attribute Based Models (UML, IDEF1X, Oracle CASE etc.) -→ do NOT formally support this distinction
- This difference is a fundamental premise in NIAM or ORM
- Even Ontology structures mix up these concepts....see next example slide

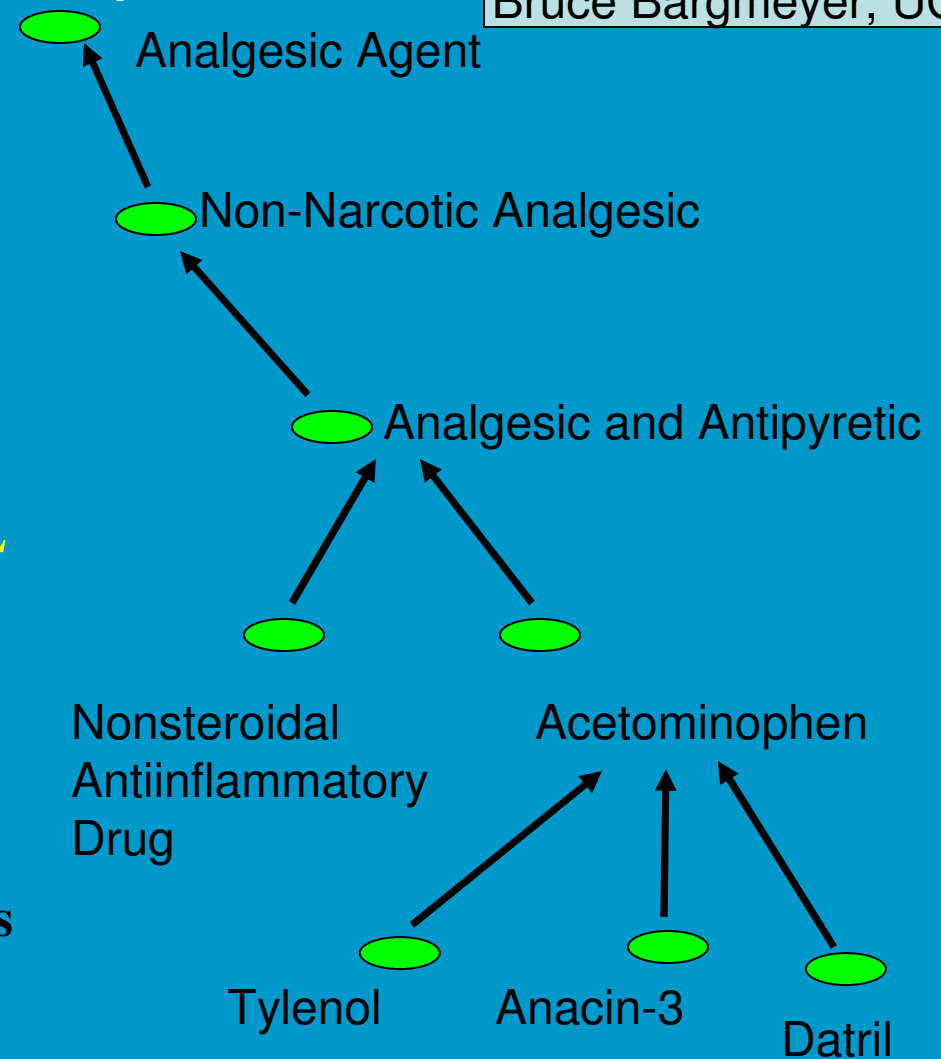


Reasoning: Relationship metadata can be used to infer non-explicit data

Slide: Courtesy
Bruce Bargmeyer, UCLA

For example...

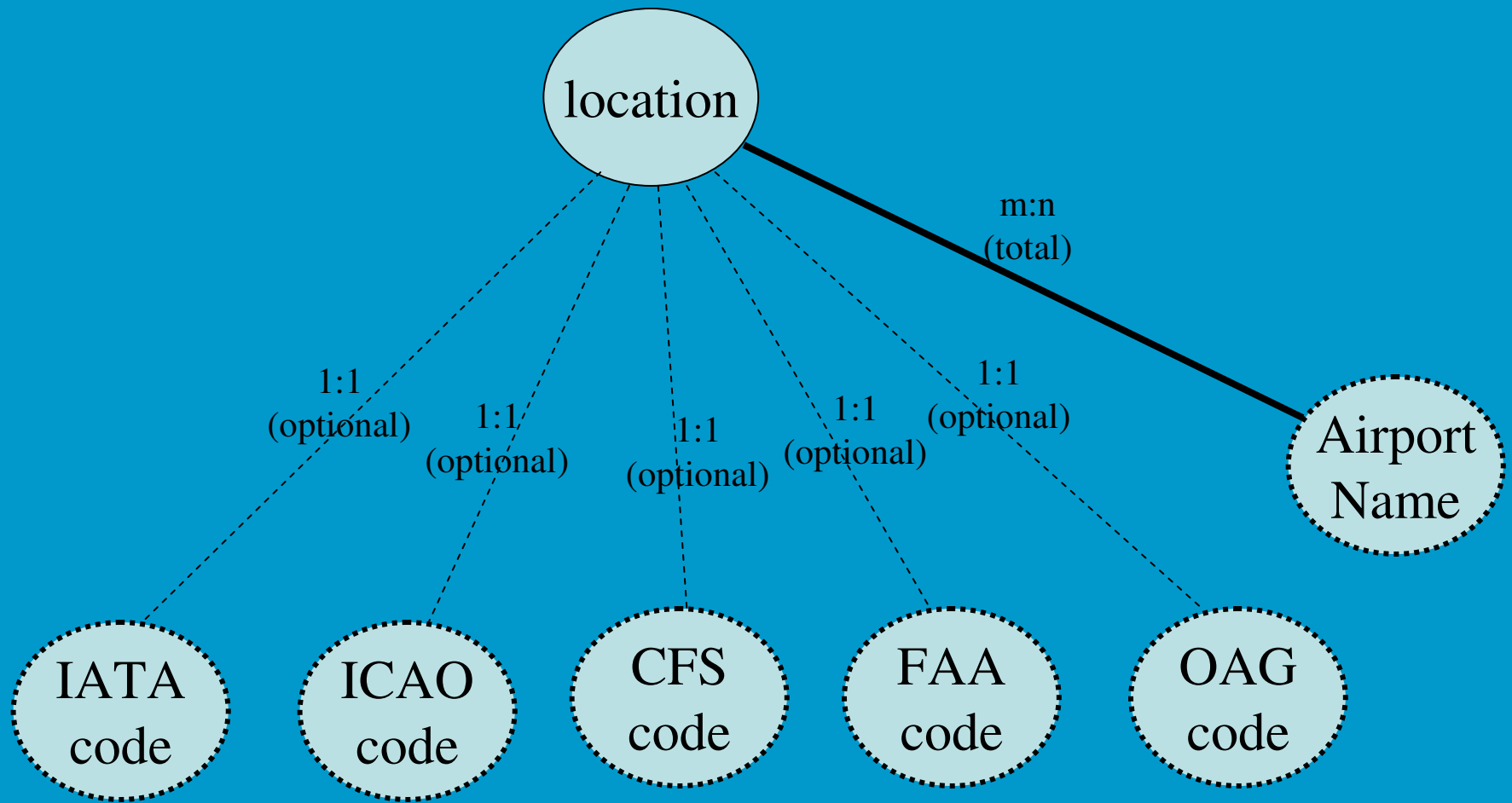
- (1) patient data on drugs currently being taken contains brand names (e.g. Tylenol, Anacin-3, Datril,...);
- (2) concept system connects **different drug types and names** with one another (via is-a, part-of, etc. relationships); [Note: NO FORMAL DISTINCTION ...how is this connection made? – Baba comment]
- (3) so... patient data can be linked and searched by *inferred* terms like “acetaminophen” and “analgesic” as well as trade names explicitly stored as text strings in the database



Relating objects:Lexical Object Types

- Case1: relationships between concepts and their representations;
- Case2: suitable axioms are added in order to constrain their intended representational interpretation

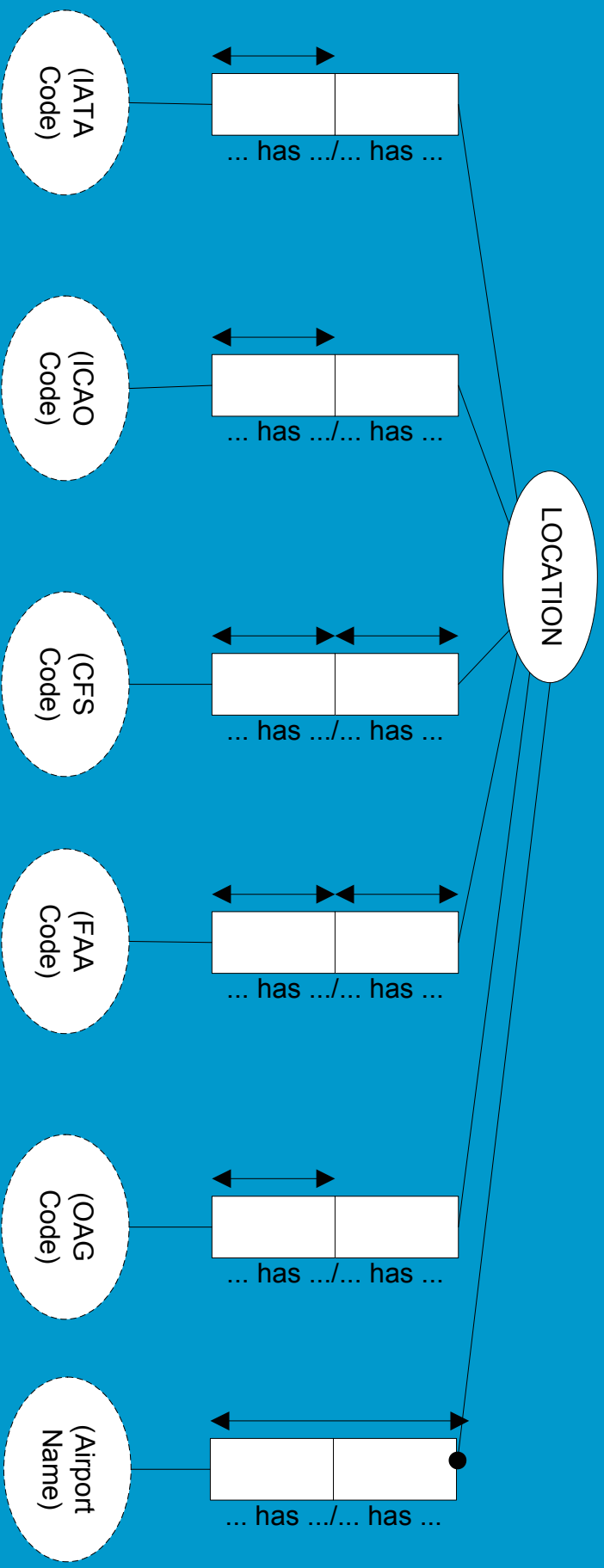


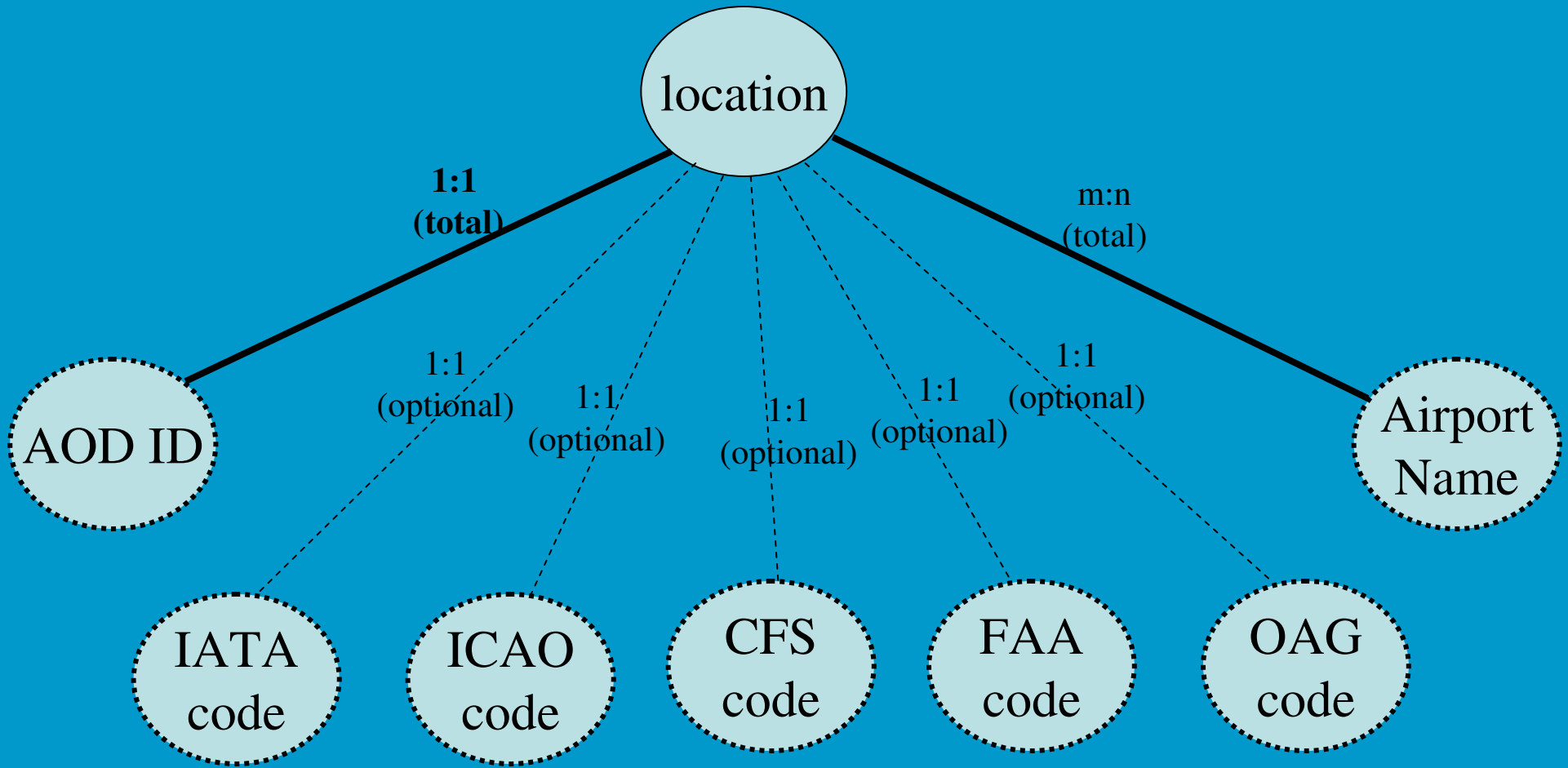


Case 2: Representations
Multiple or Non Total
"ORM 2007"



Multiple or non-total references for an object

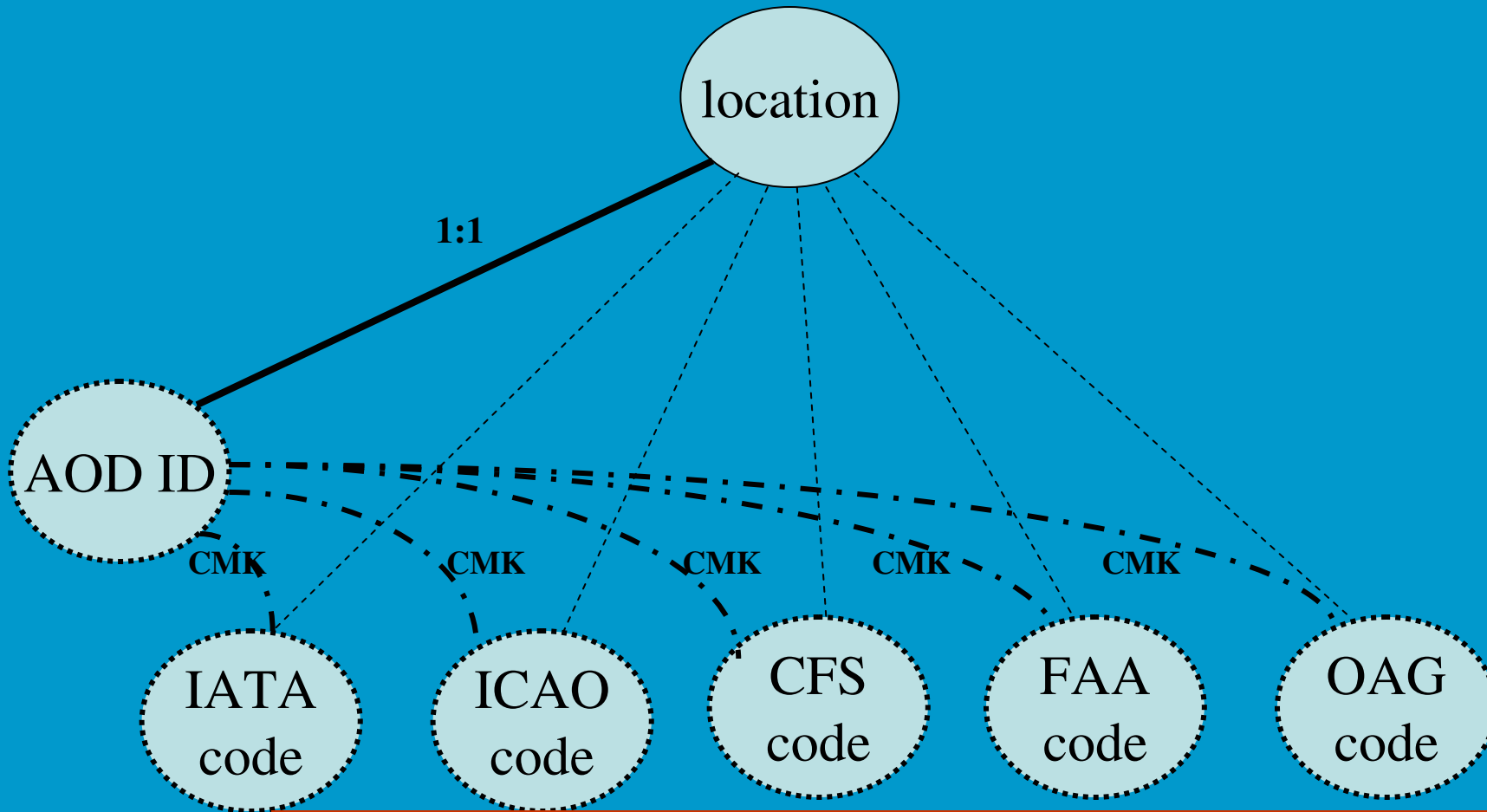




Case 2: Representations Adding a Total 1:1 reference

“ORM 2007”

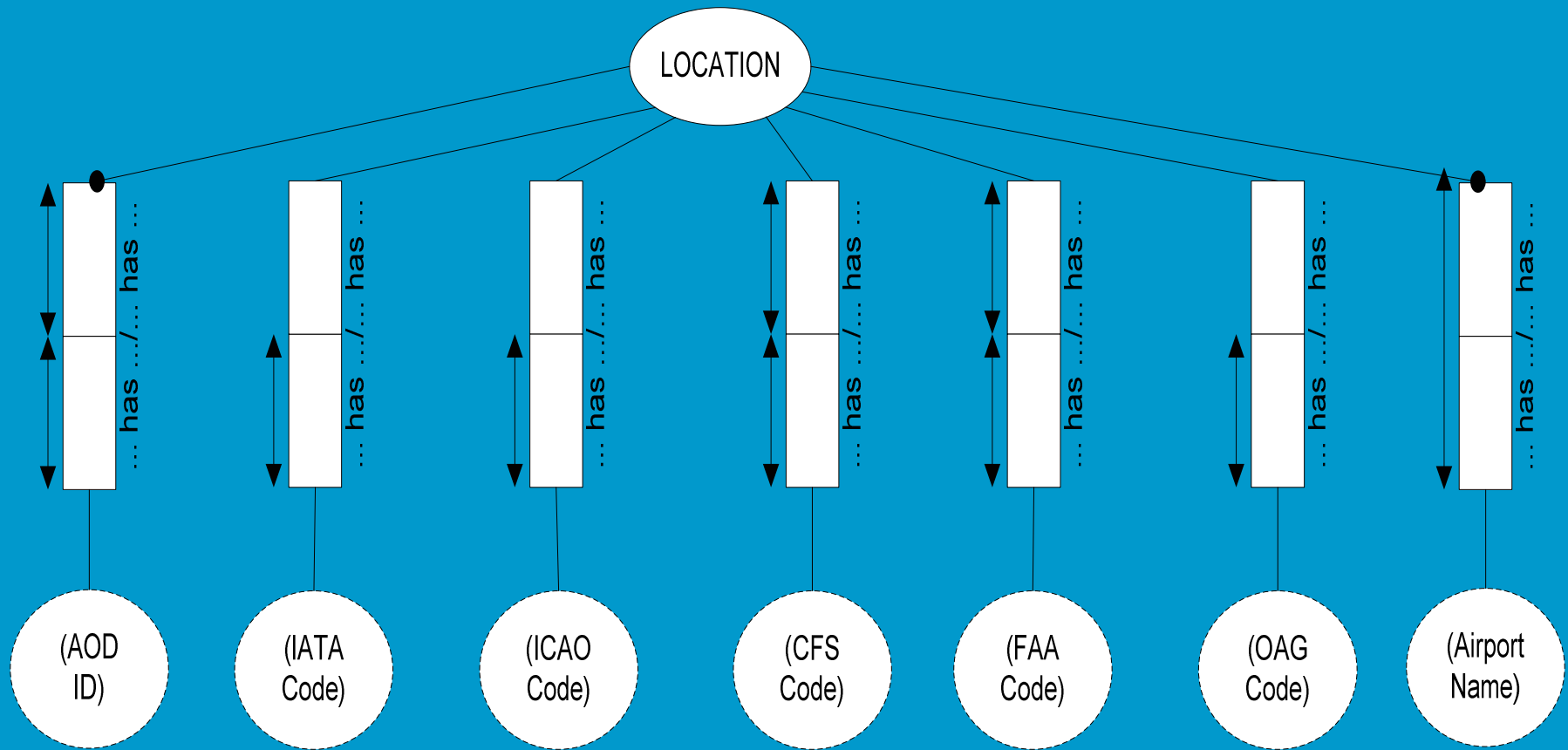


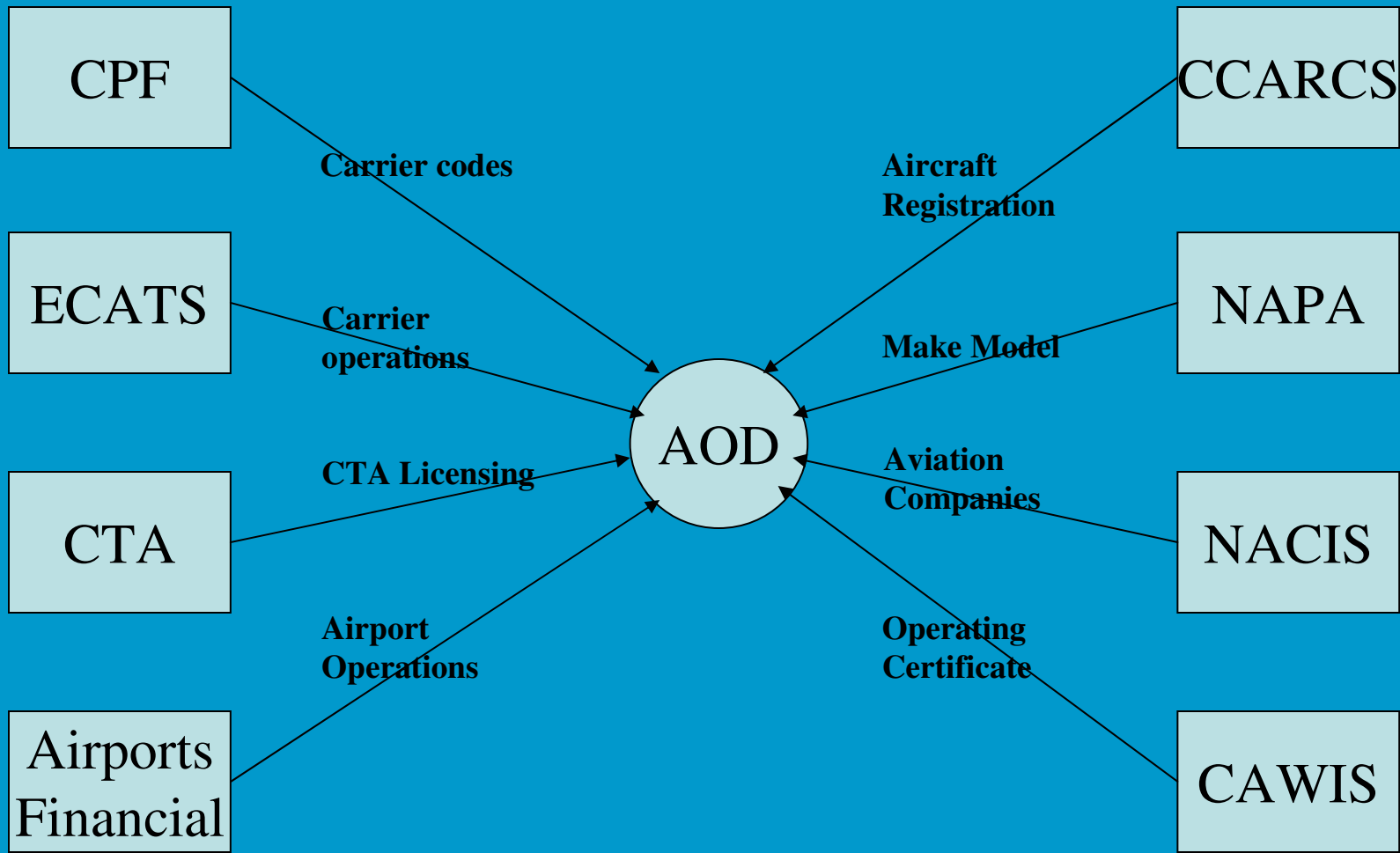


**Representation Mappings to support Concepts
i.e. Non lexical object to Lexical object (data element)**



Adding a total 1:1 reference via AOD

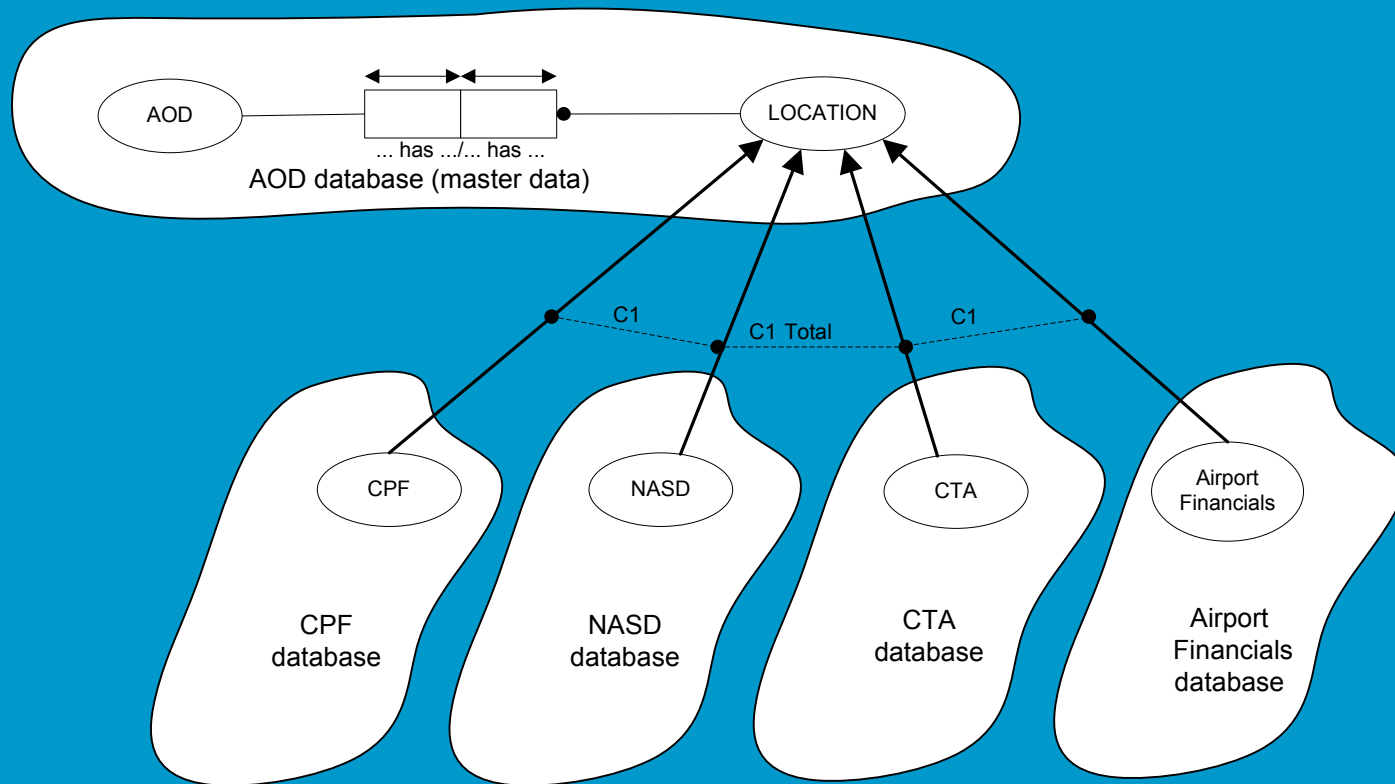




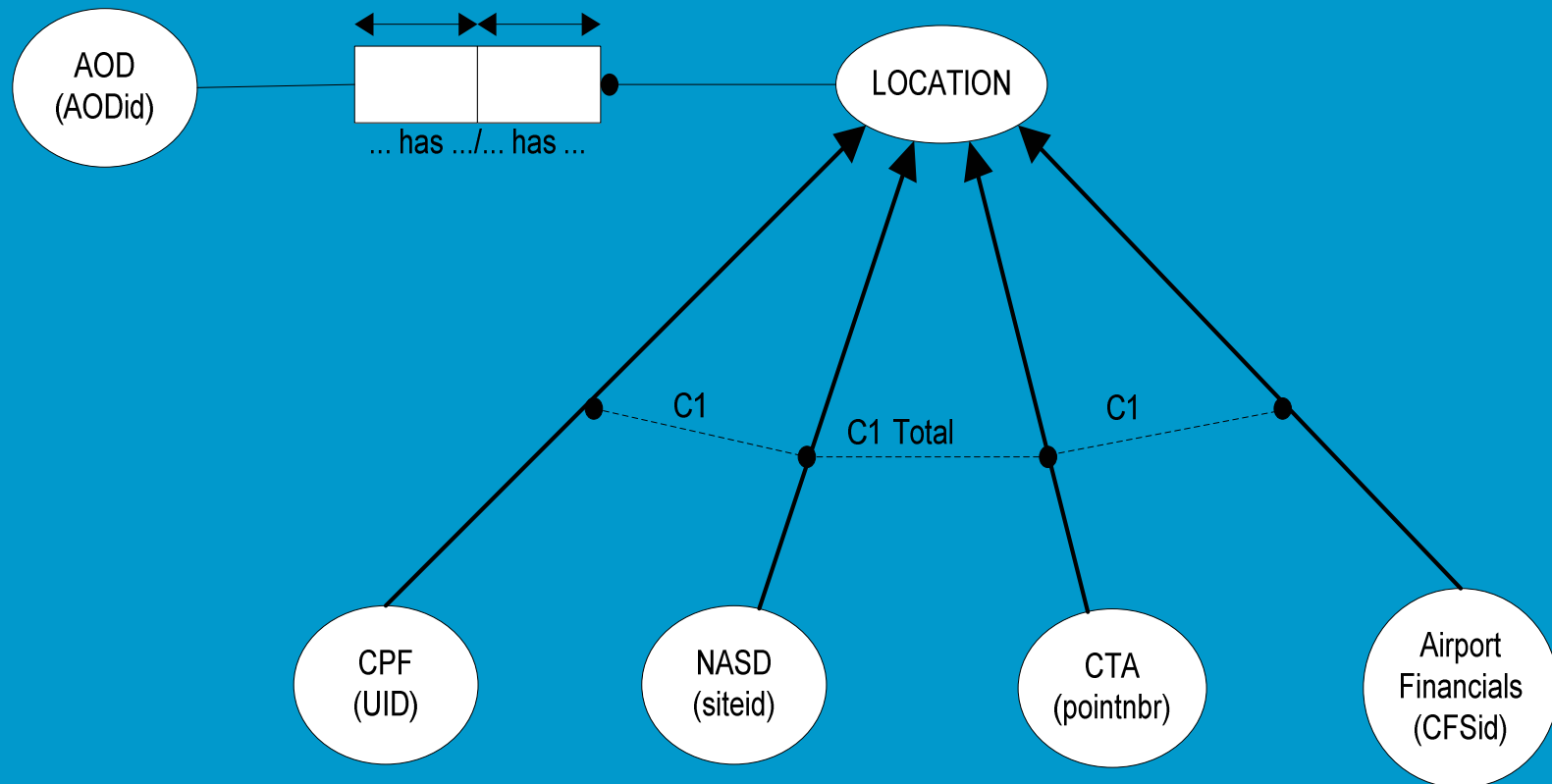
AOD – permits dynamic integration based on ontology related concepts



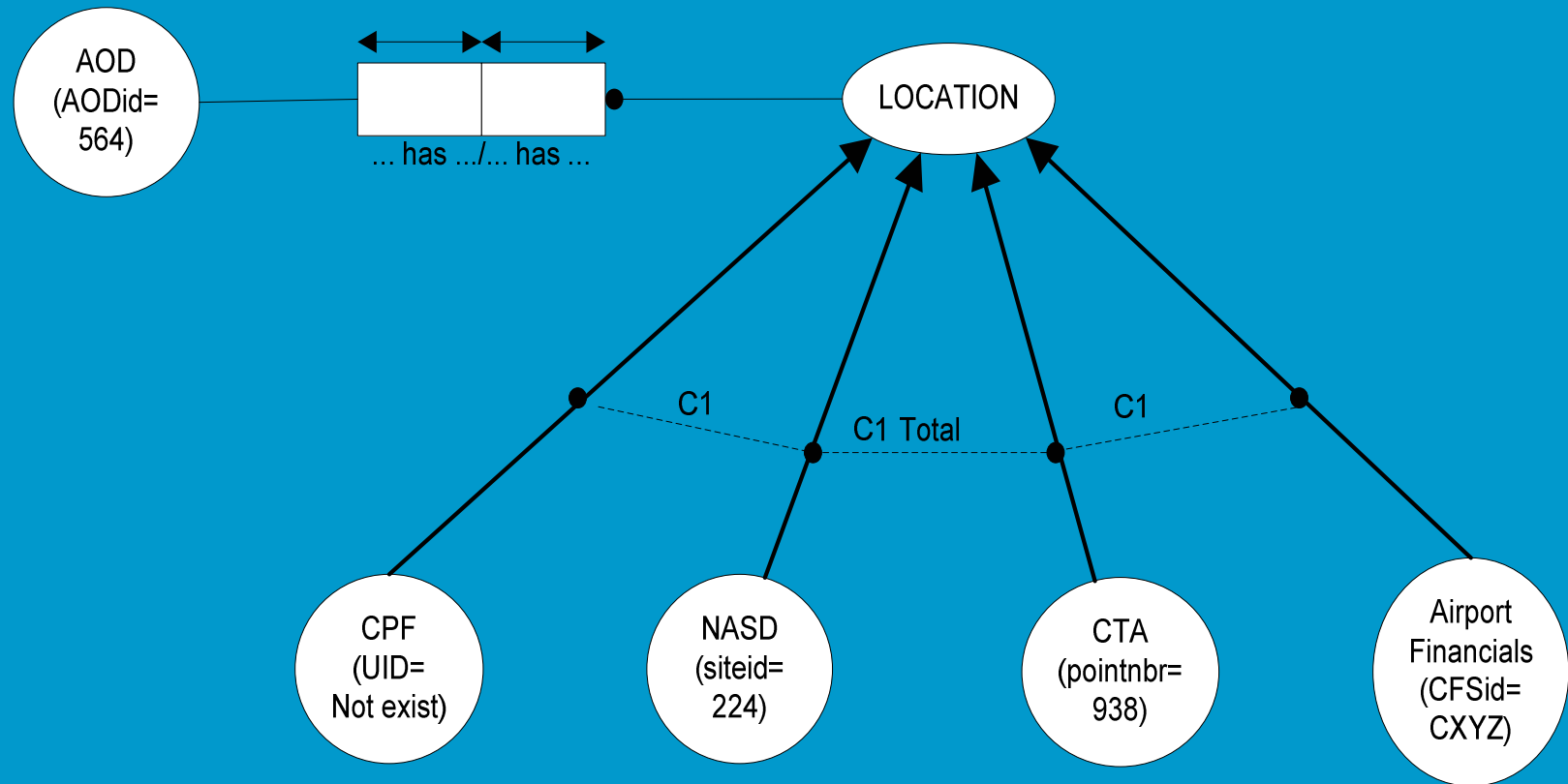
Alternate concept references logically treated as a subtype configuration aligned to physical databases

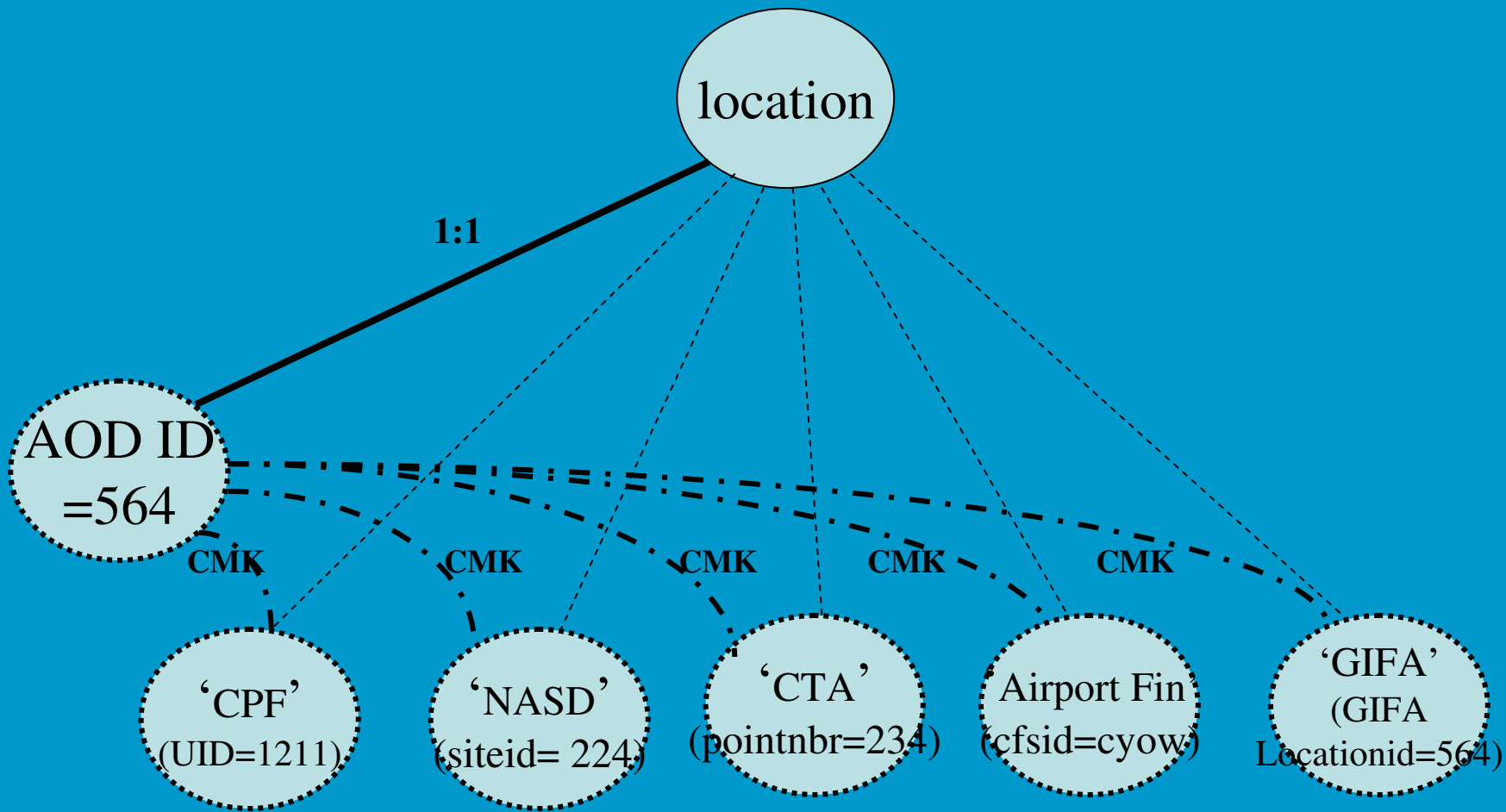


Mapping at the metadata level for columns in participating databases



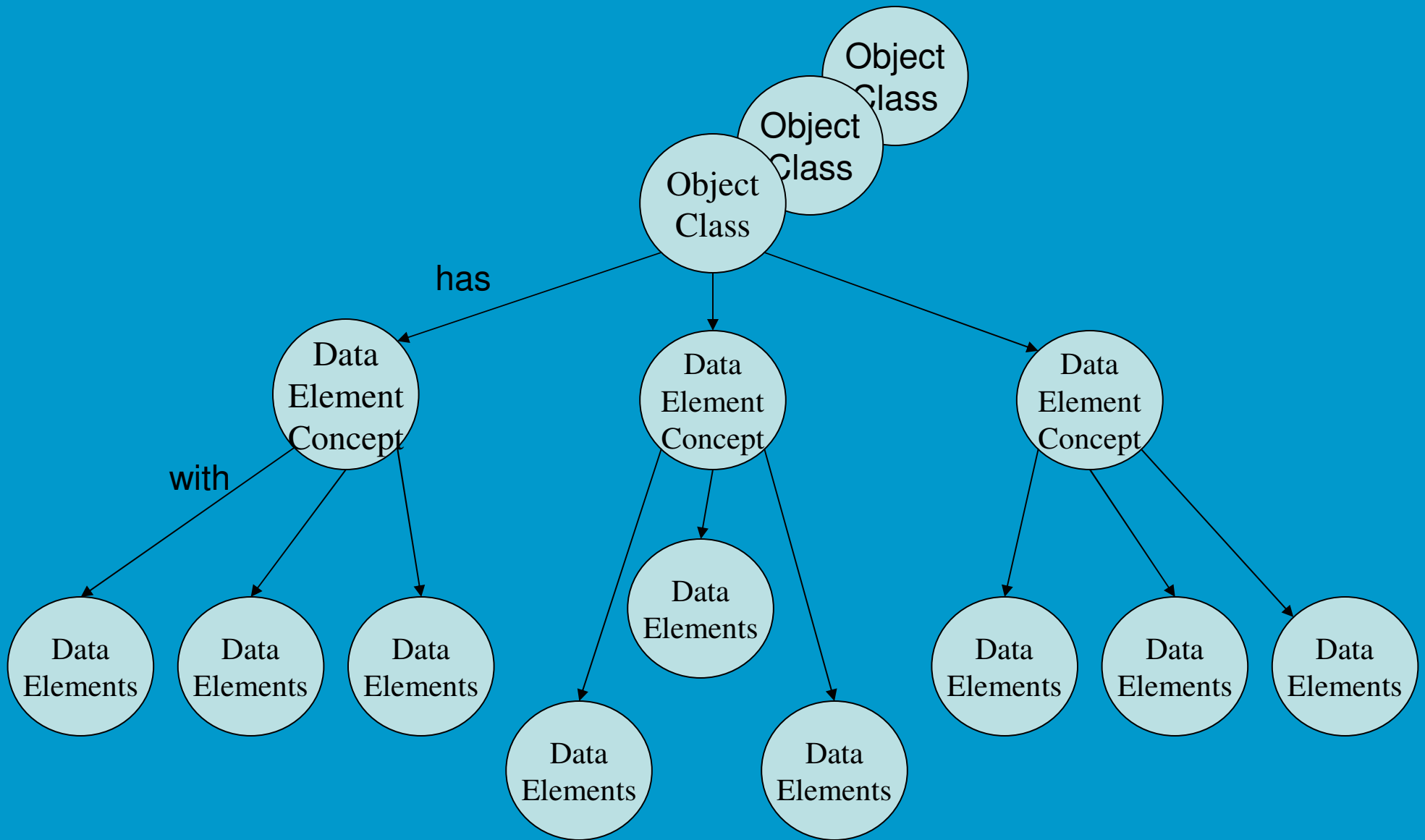
Mapping at the value-based level for columns in participating databases





Representation Mappings applied to Physical level 2 (value)





AOD CMK: ObjectClass – Data Element Concept – Data Elements Hierarchy



ISO 11179 based mappings used to derive the Common Mapping Key constructs - Location

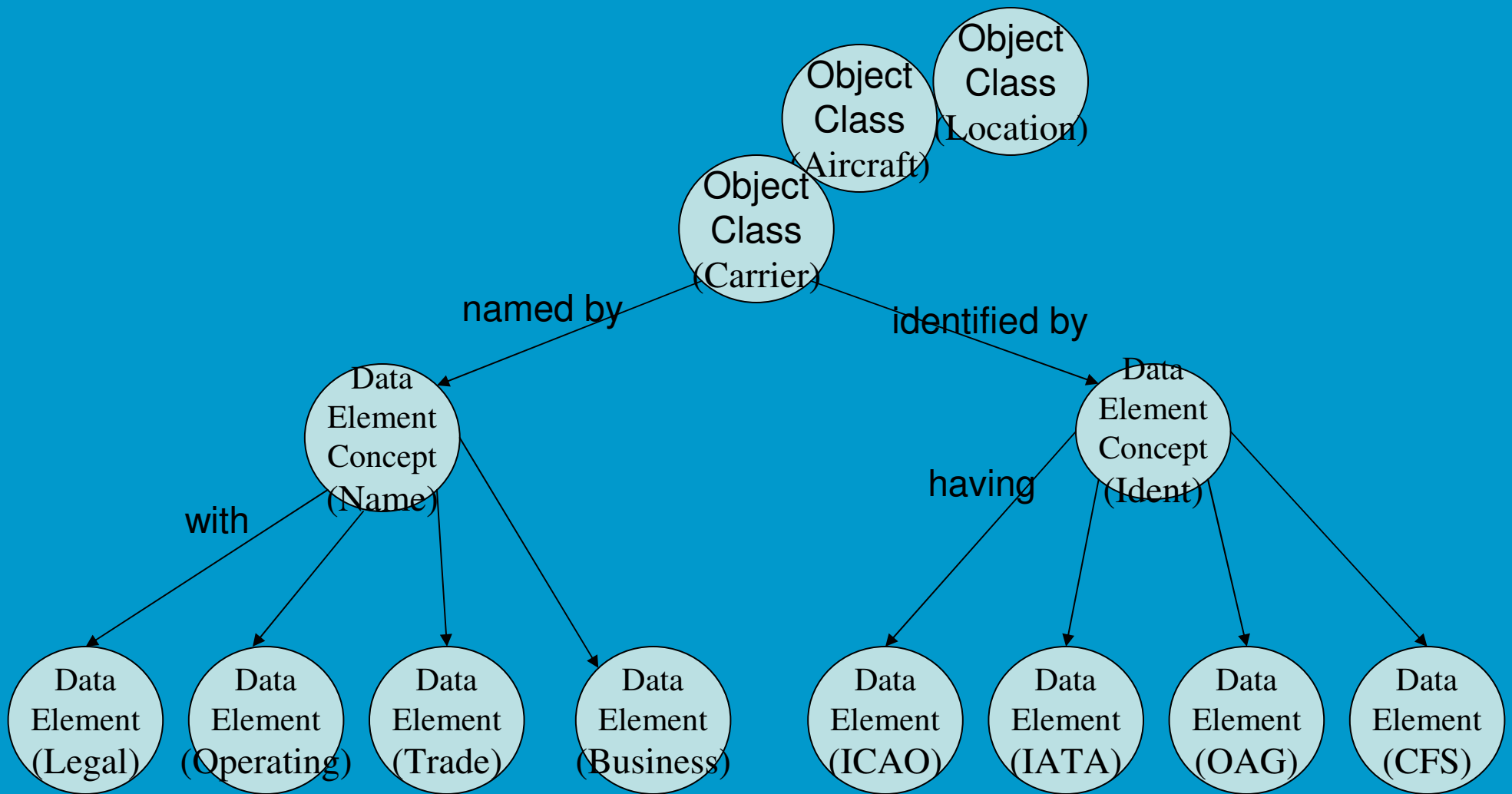
Object Class	Data Element Concept	Data Elements
Location	Identity	IATA Code (Location)
Location	Identity	OAG Code (Location)
Location	Identity	ICAO Code (Location)
Location	Identity	CFS Code (Location)
Location	Identity	FAA Code (Location)
Location	Identity	AOD (Location id)



ISO 11179 based mappings used to derive the Common Mapping Key constructs - Carrier

Object Class	Data Element Concept	Data Elements
Carrier	Name	Legal name
Carrier	Name	Operating name
Carrier	Name	Trade name
Carrier	Name	Preferred name
Carrier	Name	Carrying on Business name
Carrier	Identity	IATA Code (Carrier)
Carrier	Identity	OAG Code (Carrier)
Carrier	Identity	ICAO Code (Carrier)
Carrier	Identity	AOD Party id





AOD CMK: Object Class – Data Element Concept– Data Element Hierarchy with Example instantiations



CMK Sentence Types

Sentence Type 1:

*Application **CPF** contains object class **CARRIER** with data element concept **IDENTITY** having data element **ICAO-CODE**;*

Sentence Type 2:

*Schema **CPFDB** has Table **CARRIERUID** with column **UID**;*

Now bringing the parts together,

Sentence Type 3:

[Sentence Type 2] is contained in [Sentence type 1]

*i.e. [Application **CPF** contains object class **CARRIER** with data element concept **IDENTITY** having data element **ICAO-CODE**] is contained in [Schema **CPFDB** has Table **CARRIERUID** with column **UID**];*

Promulgating these sentences to the 1:1 total reference to AOD:

*[Application **AOD** contains object class **CARRIER** with data element concept **IDENTITY** having data element **PARTY_ID**] is contained in [Schema **AODDB** has Table **AOD_PARTY** with column **AOD_ID**];*



CMK Sentence Types

Sentence Type 3:

[Sentence Type 2] is contained in [Sentence type 1]

i.e. [*Application* **CPF** contains *object class* **CARRIER** with *data element* concept **IDENTITY** having *data element* **ICAO-CODE**]

is contained in

[*Schema* **CPFDB** has *Table* **CARRIERUID** with *column* **UID**];

Promulgating these sentences to the 1:1 total reference to AOD:

[*Application* **AOD** contains *object class* **CARRIER** with *data element* concept **IDENTITY** having *data element* **PARTY_ID**] is contained in
[*Schema* **AODDB** has *Table* **AOD_PARTY** with *column* **AOD_ID**];



AOD Work Plan to date...



- Define a generic ontology based data model
 - Identify master data identifications
 - Identify specializations vis-à-vis each of the participating systems e.g.
 - CCARCS, NACIS/NAPA, CAWIS, CPF, ECATS
 - Develop mappings for each specialization
 - Develop basic interface mechanisms for use...
- [Eureka: we have a user community already....]

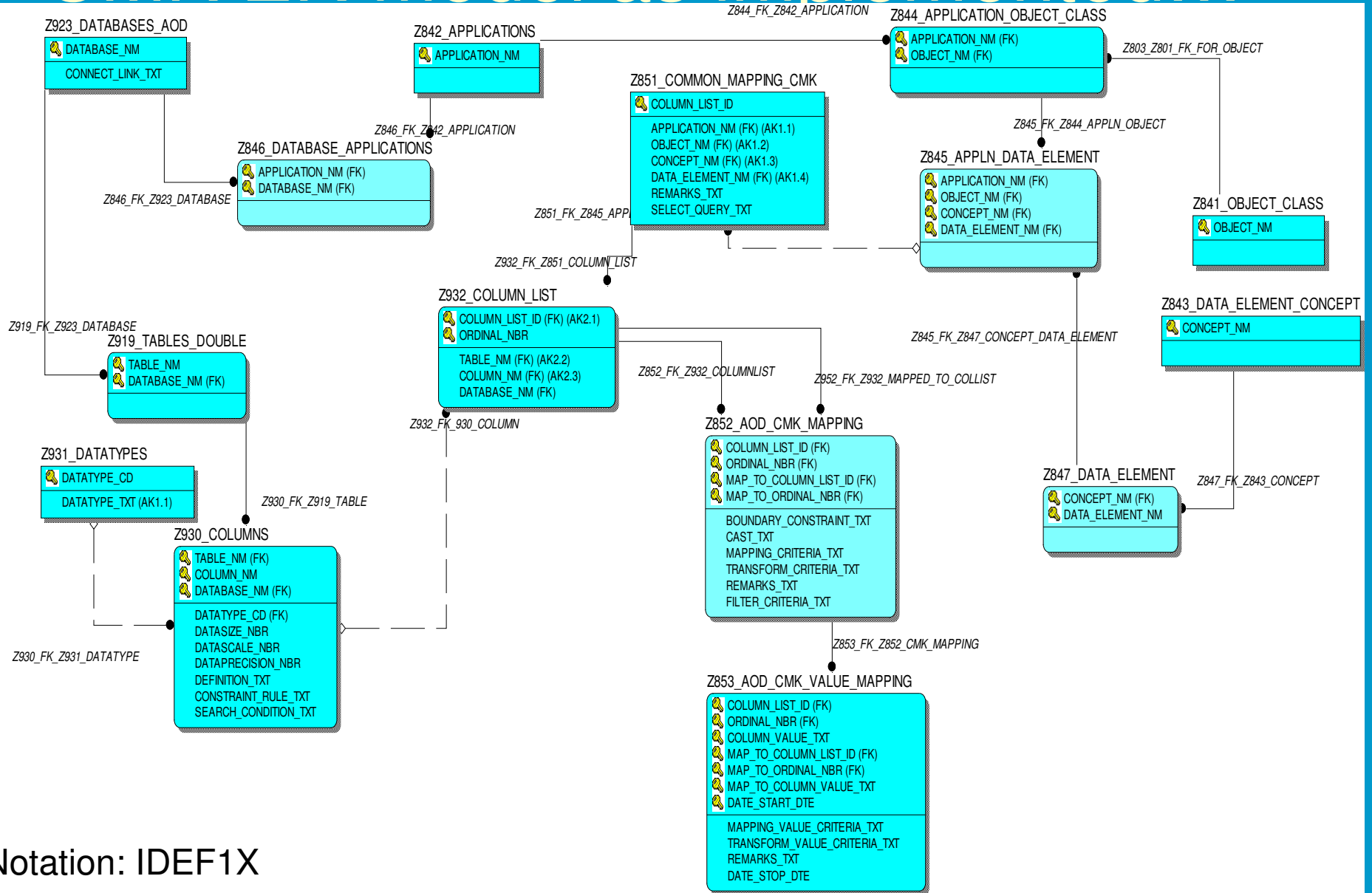


AOD DB - CMK

- Provides a **2 level facility** using **CMK** (Common Mapping Key) for:
 - **Level 1: Metadata mapping** between the column identifiers (names) across multiple systems
 - **Level 2: Value based mapping** between column values across multiple systems

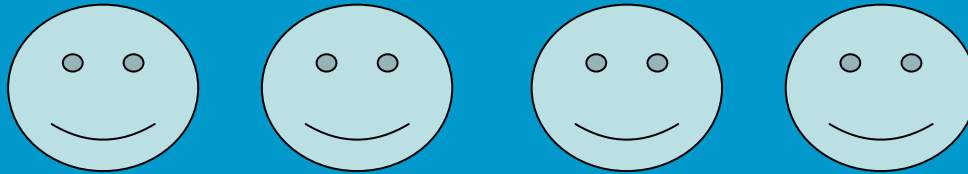


CMK ER model as implemented...



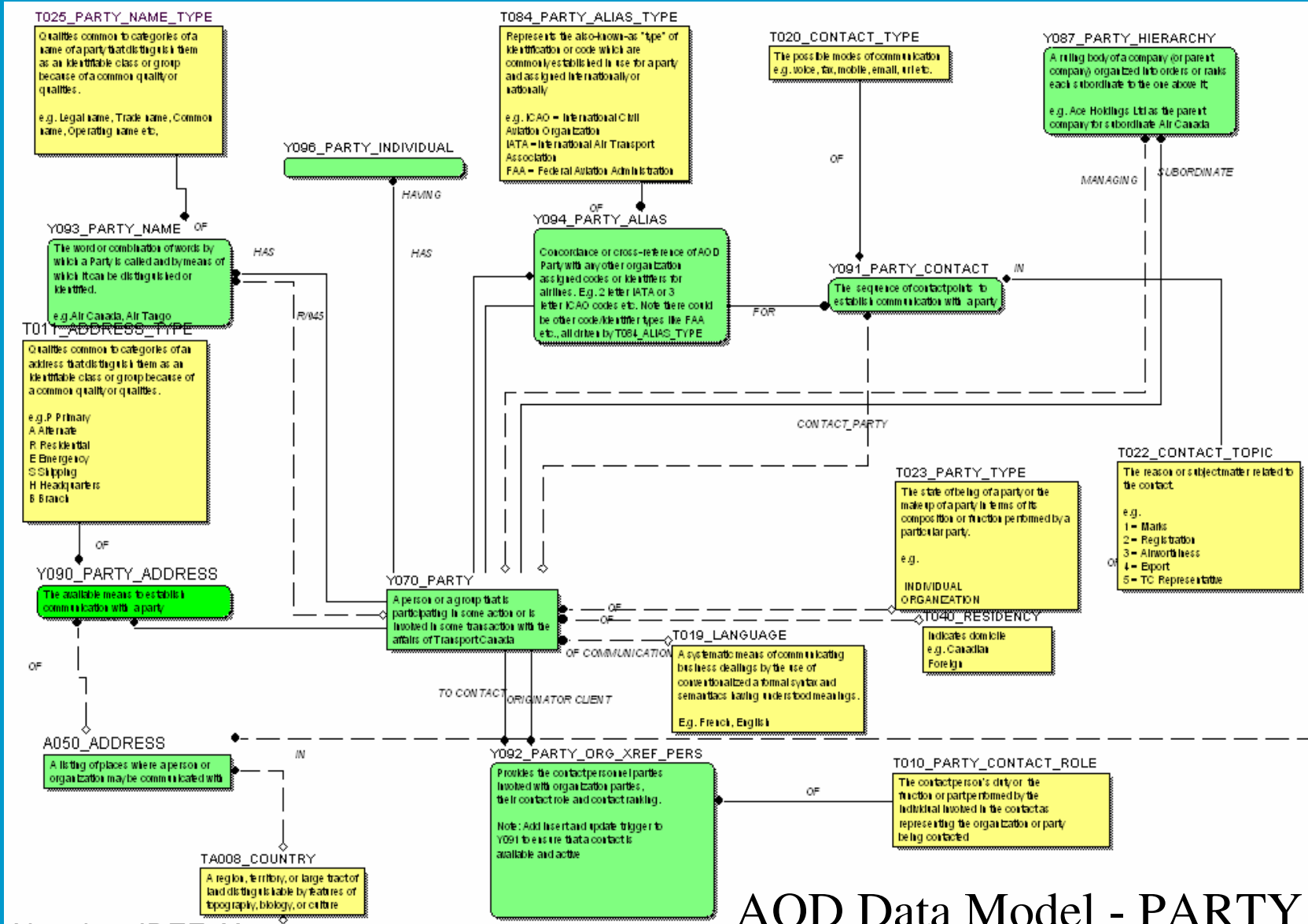
Notation: IDEF1X

AOD



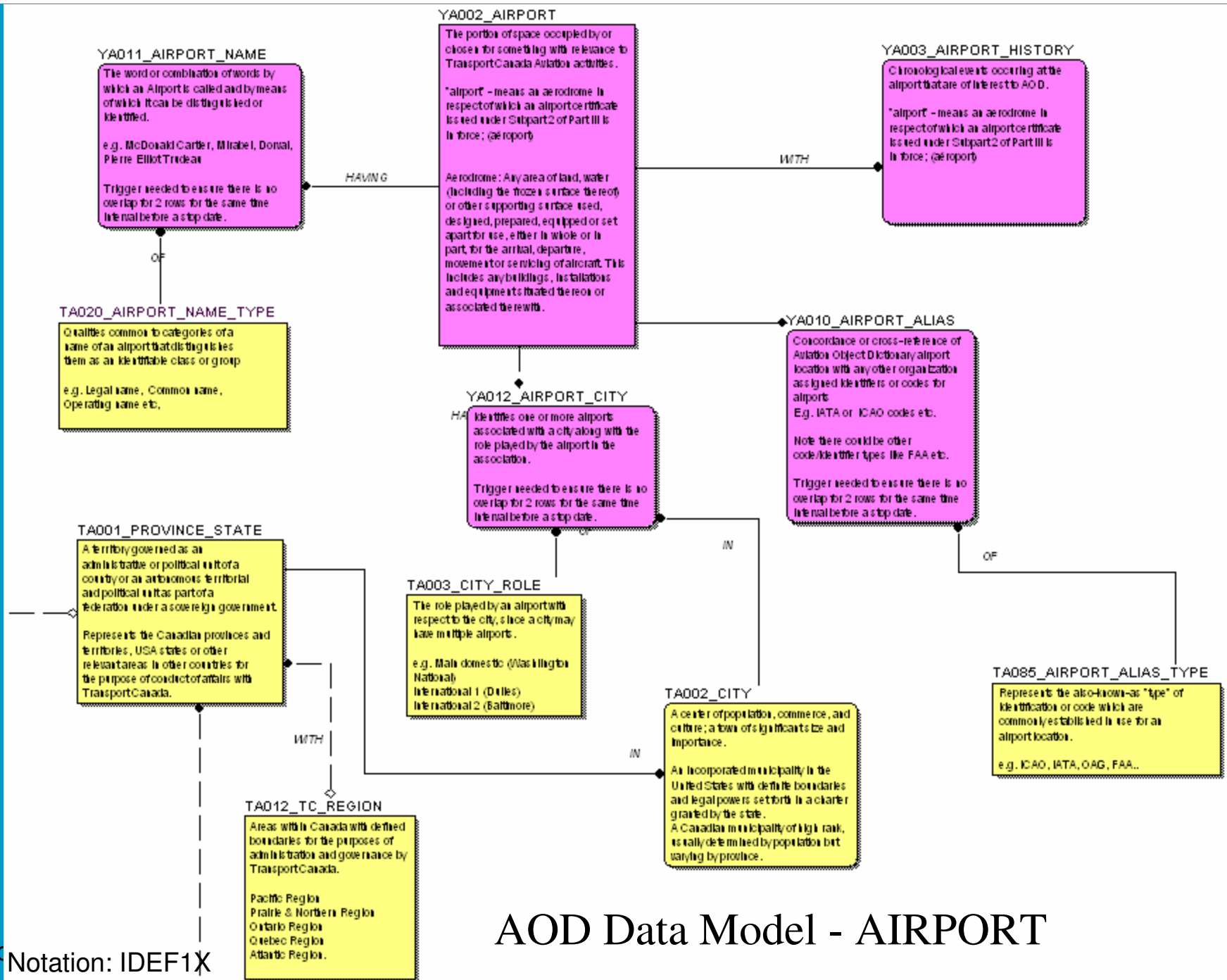
- Common Mapping Key (CMK) allows a **mapping** from AOD to other participating system
- Enables **navigation** from one system to another
- AOD DB – CMK contains Multiple **Entry Points**(MEP)
- MEPs define “data entry points” that can be used to query or “enter” the AOD-DB to use the CMK
- E.g. OPERATOR ontological entry points are:
 - Operating Certificate Number
 - Operating Certificate Type (Fixed Wing/Rotorcraft)





AOD Data Model - PARTY

Notation: IDEF1X



AOD Data Model - AIRPORT

AOD

E.g. OPERATOR multi domain ontological based entry points are:

- Service Type (Domestic, Intl Sched, Intl Non-Sched, Aerial Work)
- Legal Name (can be duplicated since usually unique only within Provinces)
- Common Name
- Aerial Work (Aerial Advertising, Aerial Construction, Aerial Inspection & Surveillance, Aerial Harvesting, Aerial Mapping, Aerial Photography, Aerial Sightseeing, Aerial Spraying, External Load, Fire Fighting, Flight Training, Forest Fire Management, Glider Tower, Heli-logging, Human Organs, Parachute Jumping, Wild Life Management)
- Canadian / Foreign Operator
- On-Behalf Operator
- Aircraft Owner
- Specific Aircraft Mark
- (next slide...)



AOD

E.g. OPERATOR multi domain ontological based entry points are...:

- Region (for Canadian Operator only)
- Location of Operator (main base, sub base, Airport, Aerodrome)
- PARTY organization changes
- Party Roles (Individual, Single Org, Consortia, Fleet, Govt, Regulatory Agency)
- Party Contact (Postal address, email, phone, URL)
- Operator Support Personnel
- Operator Support Organization
- Operator Support Personnel License Type

ANY ENTRY POINT CAN BE USED WITH ANY VALUE TO QUERY ANY COMBINATION TO RETRIEVE APPLICABLE SET OF CMKs via ONTOLOGICAL DOMAIN BRIDGING BASED ON NON-LEXICAL OBJECT RELATIONSHIPS



TC-AOD's realization of Ontological Bridges (Standards based)

Topic: Multi- Metadata Models Framework

Standards/Specification:
ISO 19763
OMG MOF

- Reference metamodels
- Reference Frameworks

Topic: AOD Ontology &

Standards based dynamic integration

- Concepts
- Definitions
- Designations
- Data Elements

- Multi Metadata models
- Export/Import XML XMI

Topic: Data

Standards:
ISO 11179 Metadata Registries

Topic: IM, IT, DW

Standards and Specification:
ISO 9075

OMG Common Warehouse Metamodel

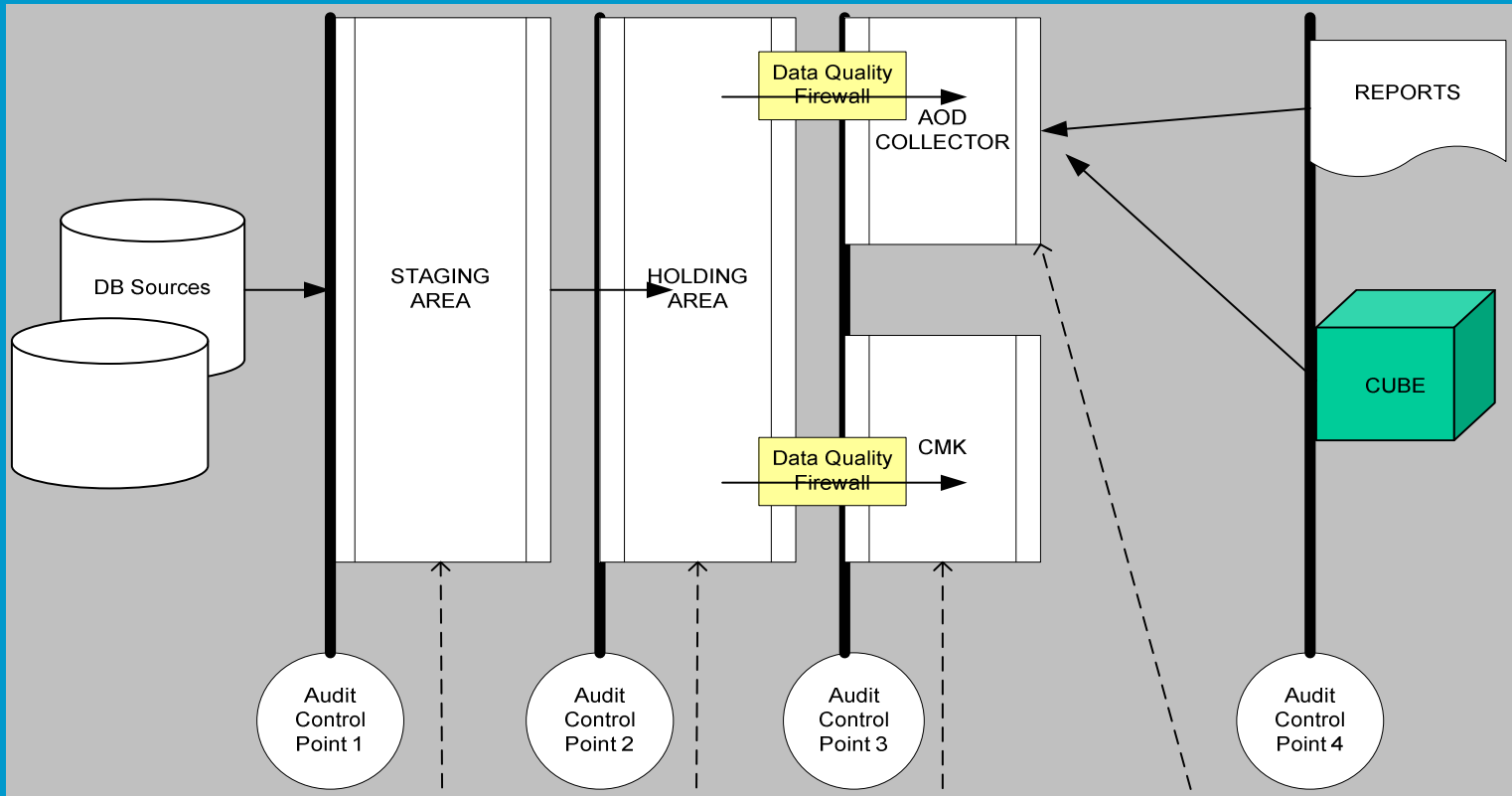
- Mappings
- DB Schemas, Keys/Indexes
- Relational DB Structure
- Record Structure
- Transformation Tools
- Extraction Tools
- Data Quality Filters



AOD DB – CMK...Ontology, ORM and Standards Based...

- AOD Data Model ontological concepts and domains based on ISO 11179 and driven by NIAM & ORM
- AOD-CMK Data Model based on ISO 19763 and MOF concepts establishing ontology based mappings and domain bridges across multi-MOF levels
- AOD Engine and Interface Model based on ISO SQL 9075
- AOD Data Quality Firewall based on ISO 11179-19763-9075 combo using Object Role Modelling across multi-MOF levels (both vertical and horizontal parallel pairs)
- AOD Services Model based on BPMN and sub processes aligned to Service Oriented Architecture





- Source Data Structure
- Source Data Types
- No constraints
- Subset of source
- Allow data discovery
- Preview Data Quality to make go-nogo decision

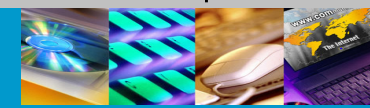
- Data Type Harmonization
- Data Conversion
- No constraints
- Data Model structures same as Collector
- Source data keys mapped to AOD keys for main concepts
- Audit Control
- Keeps error rows from dataloads

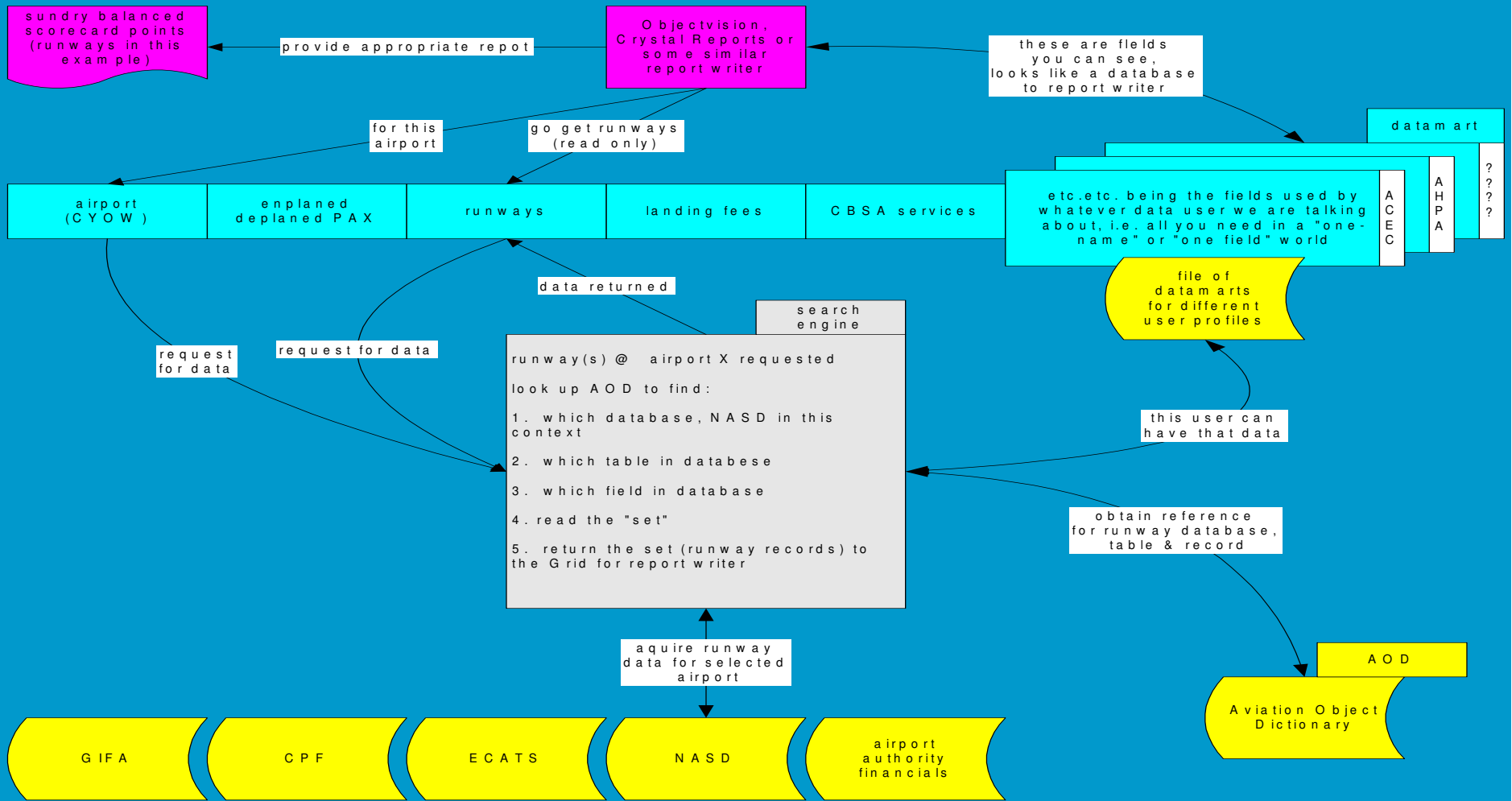
- CMK
- Provides mapping between source systems key data and AOD
- Cross-ref metadata
- Cross-ref key data values
- Concept knowledge based search
- Link to entry points in AOD

- AOD
- +3rd Normal Form
- Fully constrained
- Data Model structures for completing identification via probes
- Data validated via Data Qual Firewall
- Provides concept entry points and probes
- History data structures

AOD - FLOW OF INFORMATION

Courtesy: Madhu Hundal, AOD Project, Transport Canada





User interpretation of ontological concept based bridging across domains (slide produced by User)--courtesy Iain Henderson, TC



Ontology based AOD – BENEFITS...

- AOD represents a virtual on-demand integration Central Registry for TC Corporate data
- Brings together disparate data from heterogeneous systems
- Provides path for AOD based dynamic integration yet allowing current systems to continue
- No forced conversion---implement at individual pace!
- Portable skill set for personnel migration
- DBMS independent
- TC Corporate Standards, ISO Metadata Standards and ISO SQL
- NIAM and ORM as a foundation of data model based on ontological concepts used to solve a complex situation

