Global Information Management

NASA's ATM Ontology: Semantic Integration and Querying across NAS Data Sources

Presented By: *Rich Keller, Ph.D.*

Date: *August 27, 2015*

Long Term Vision:

A Global Airspace Question-Answering System

current

"Identify all sectors within which any A320 aircraft is currently operating in US airspace "

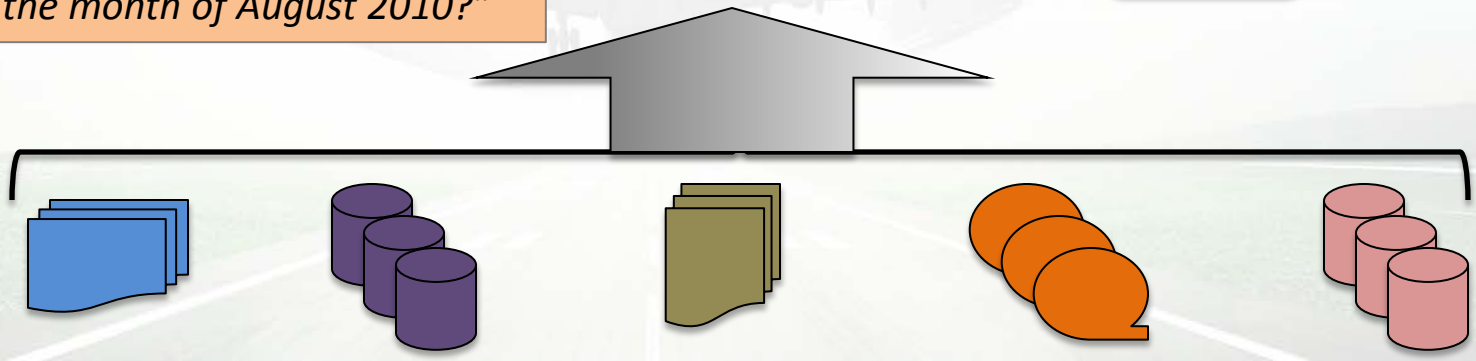
Airspace Oracle

ZTL sector 2
ZTL sector 10
ZOA sector 45
…

historical

"Which US carrier had the largest number of flights rerouted due to weather during the month of August 2010?"

UAL



Many Challenges!

- Question understanding
- Automated reasoning
- Information retrieval
- Natural language generation
- **Data exchange & integration**
  - **Data exchange**: How do you facilitate aviation data sharing and system interoperability?
    - Using *standards*: AIXM, FIXM, WXXM
  - **Data integration**: How do you take heterogeneous data from multiple sources and weave together a harmonized picture of global airspace operations?
    - Using *semantics*!

Some Small Steps Toward the Vision

NASA has developed a *semantics-based* data integration prototype capable of answering a limited set of queries about airspace operations

Outline

- Background and Motivation
- Semantic Integration Approach
- Prototype: Integrating and querying data for airspace operations at KATL on 2012/09/08

NASA Project Background

- NASA researchers need historical ATM data
  - NASA Ames conducts research on future ATM concepts
  - Researchers require data for analysis and concept validation
- NASA Ames' **ATM Data Warehouse** archives data collected from FAA, NASA, NOAA, DOT, industry
  - Warehouse captures:
    - live streamed data
    - published periodic data
  - Data holdings available back to 2009

A Sampling of Archived Data Warehouse Holdings

- ATCSCC Advisories
- Airline Situation Display to Industry (ASDI)*
- Air Route Traffic Control Center (flight plans & tracks)
- Corridor Integrated Weather Service (CIWS)
- Center-TRACON Automation System (CTAS)
- Exelis Commercial Track Feed
- METAR
- AIREP, PIREP
- Rapid Refresh (RR) Weather Forecast
- Terminal Aerodrome Forecast (TAF)
- Time-based Flow Management (TBFM)
- TRACON(flight plans & tracks)

**SWIM conversion underway for available sources*

ATM Data Warehouse: A microcosm of the NAS data environment

Problem: Non-integrated Data

- ATM Warehouse data is replicated & archived in its original format
- Data sets lack standardization
  - data formats
  - nomenclature
  - conceptual structure
- To analyze and mine data, researchers must write special-purpose code to integrate data for each new task

  ➔ Huge time sink!

- **Possible cross-dataset mismatches:**
  - terminology
  - scientific units
  - temporal alignment
  - spatial alignment
  - conceptualization organization

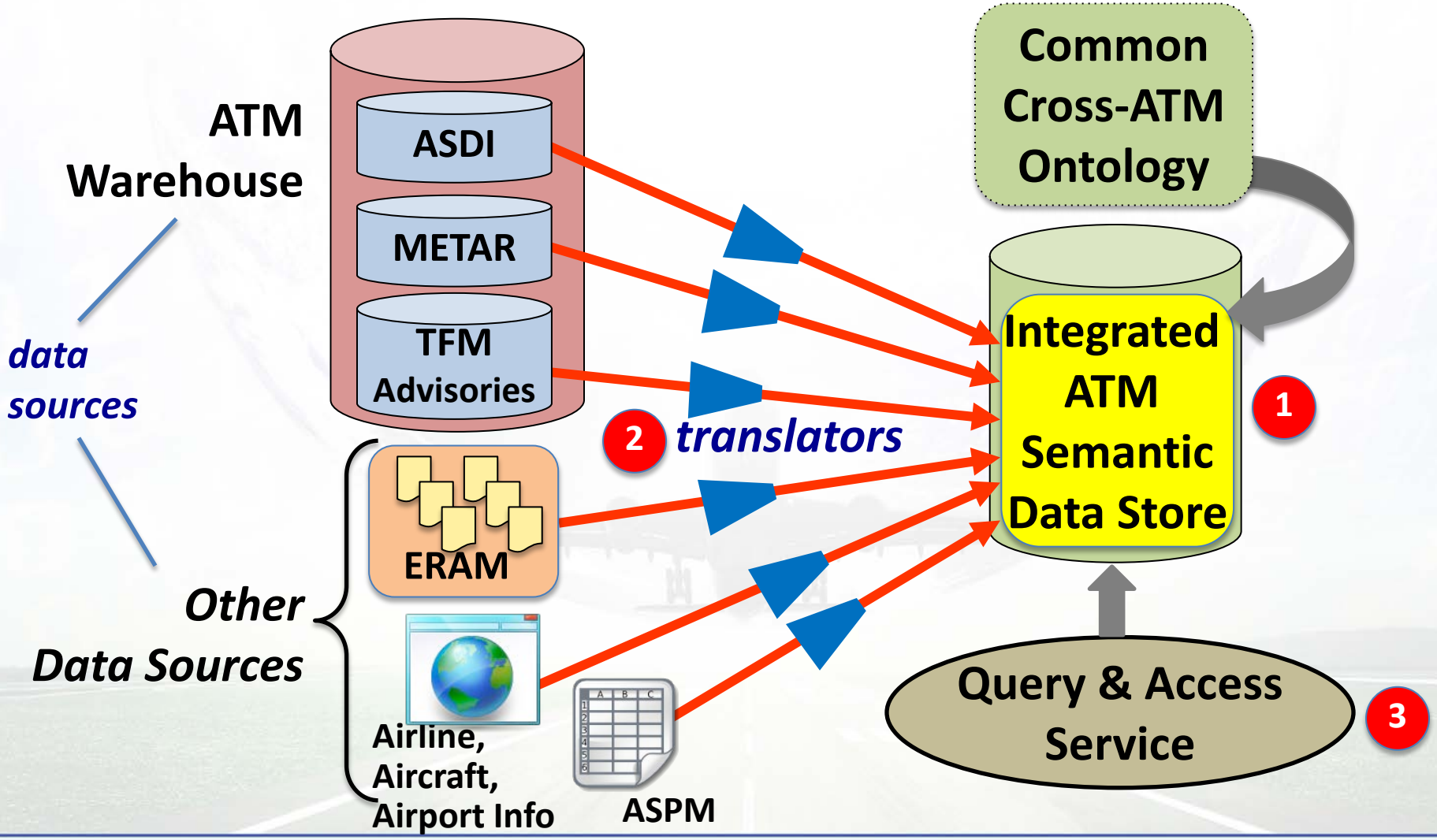
Proposed Solution

Relieve users of responsibility for integration!

'Pre-integrate' the Warehouse data sources using **Semantic Integration**

1. Develop an integrated data repository based on a common semantic data model ("an ontology")
2. Write translators to transform data from the original sources into an integrated common data repository
3. Expose integrated repository, not individual sources, to users for query and access

Semantic Integration Approach:



What is an Ontology??

- Ontology = data model + database
  - **data model:** provides a unified framework for describing, interrelating, and reasoning about different types of ATM data

The data model provides a basis for integrating heterogeneous ATM data from multiple sources

  - **database:** contains integrated air traffic management information from multiple sources, stored as per data model

This database can be queried like a conventional database. But it can also draw inferences from the data and generate new data using inference rules.

  - Plays similar role as UML, but adds inference and reasoning

What is modeled by the NASA ATM Ontology?

Covers selection of concepts found in the AIXM, FIXM, WXXM conceptual models

❖ 150+ object types

•Flights •Aircraft and manufacturers •Airlines •Airports and physical infrastructure •NAS facilities •Air traffic management initiatives •Surface weather conditions and forecasts •Airspace sectors, fixes, routes, airways •Flight plans and paths

❖ 150+ object properties

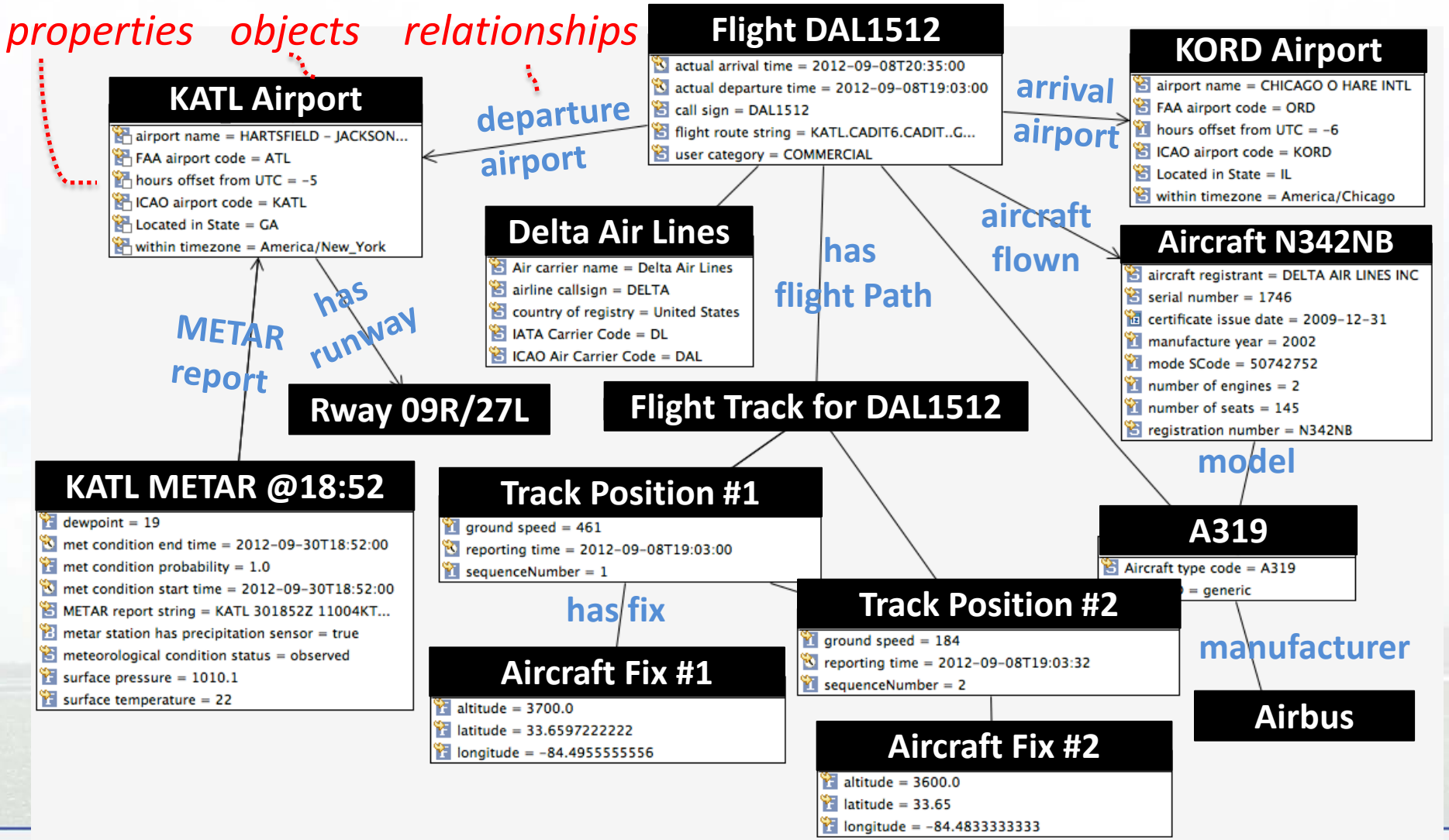
•actualDepartureTime •actualArrivalTime •airportArrivalRate •cloudType •dewpoint •EDCTarrivalHold •equipmentCode •groundSpeed •heading •hourlyPrecipitation •IATAcarrierCode •issuedTime •manufactureYear •maxVisibility

❖ 100+ relationship types

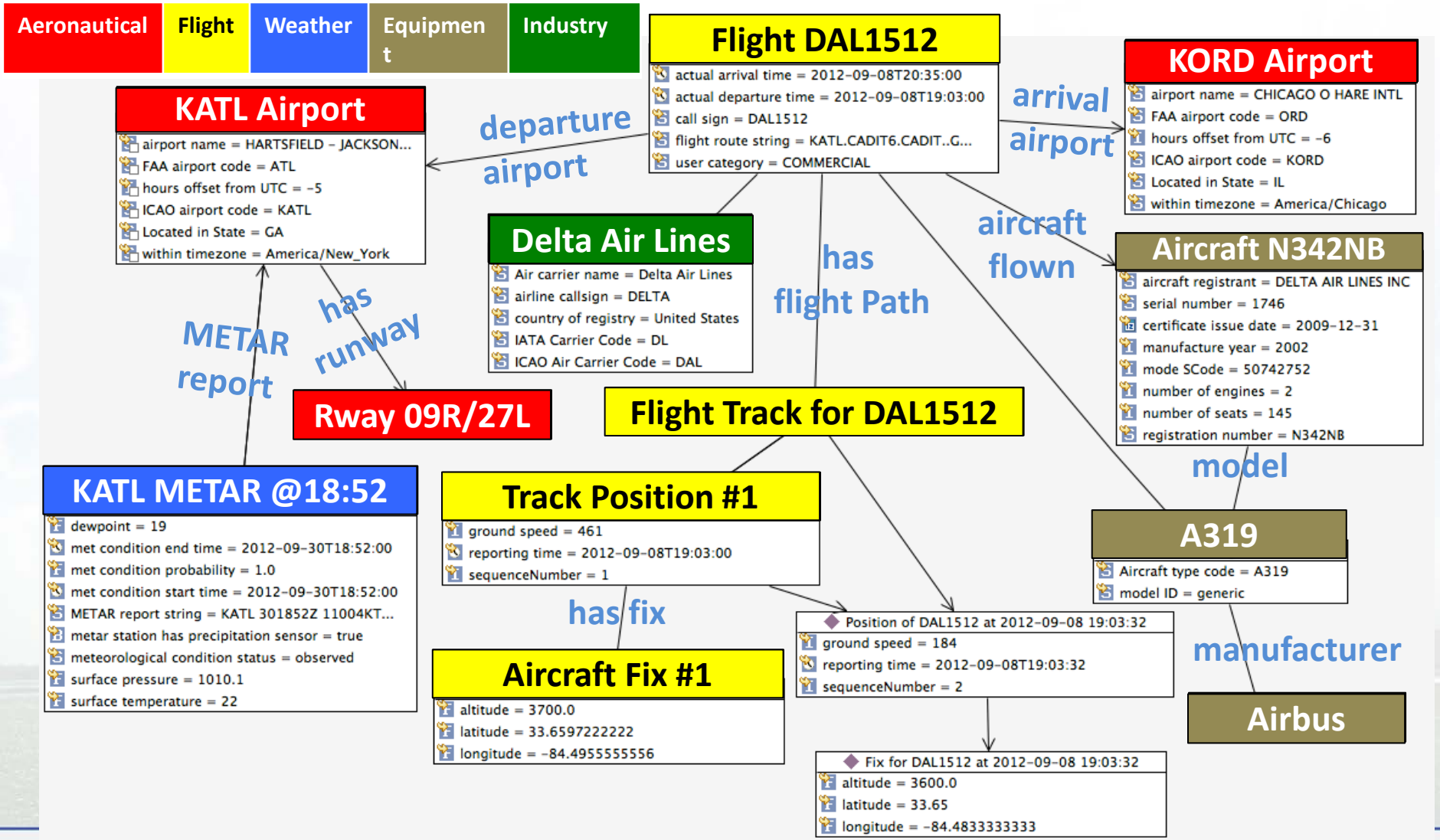
•hasRampTower •hasRunway •operatedBy • locatedInSector •manufacturedBy •hasSurfaceWindCondition •hasLOAwith •exemptedAFP •departureScope •ADLday •adjacentSector •aircraftFix •aircraftFlown •arrivalRunway •reRouteConstraint

❖ Object/property/relationship <u>instances</u> also stored in ontology

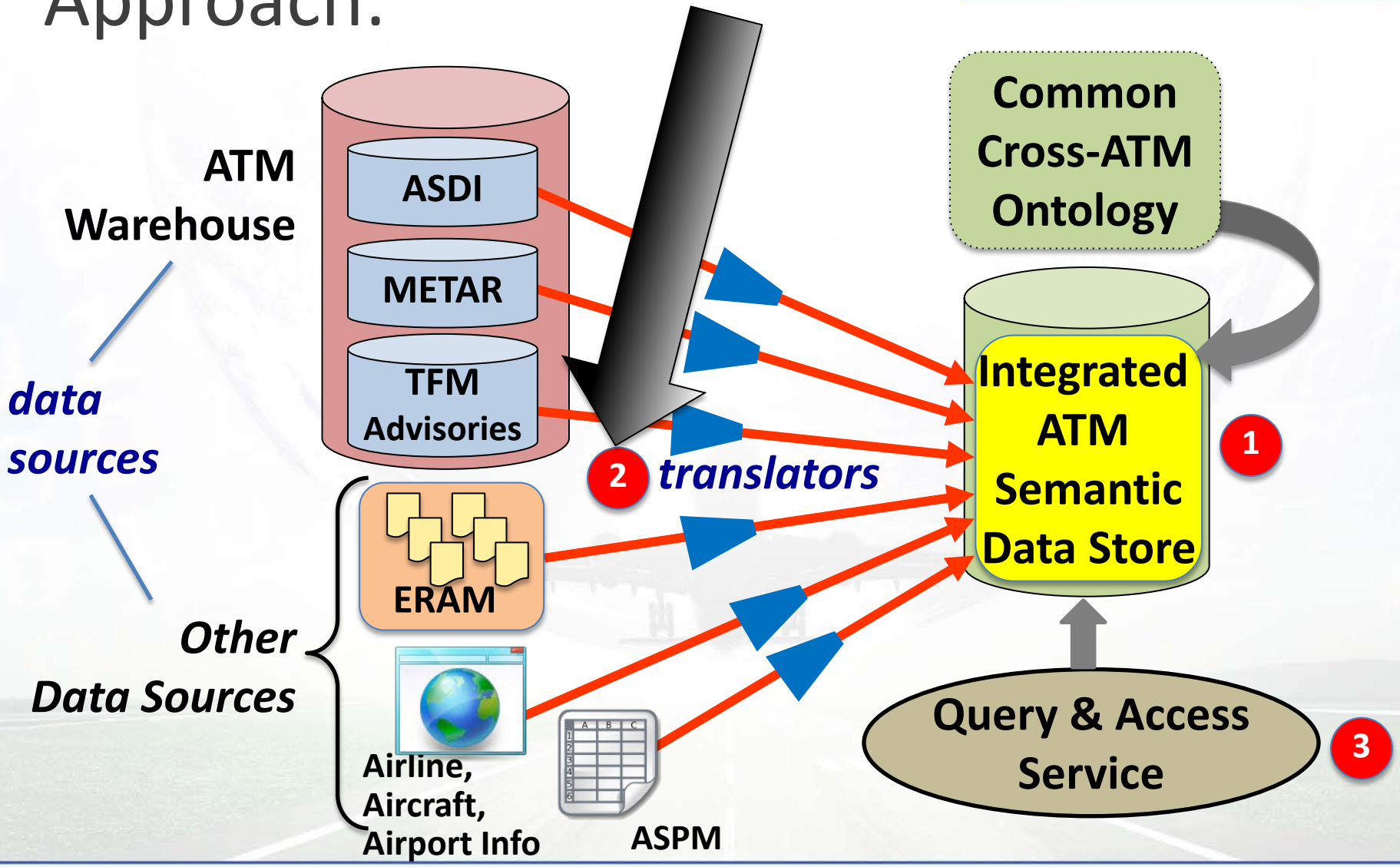
Ontology Representation of a Flight (viewed as graph)



Ontology crosses AIXM, FIXM, WXXM boundaries

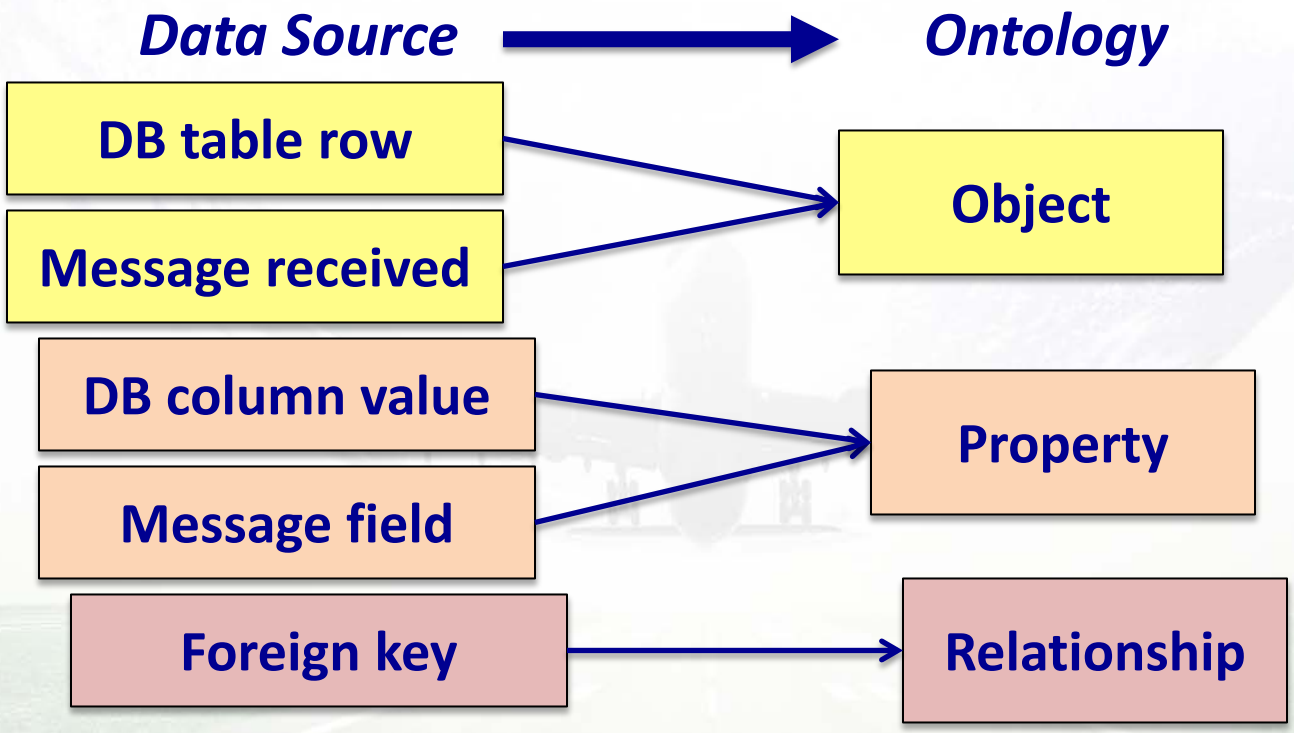


Semantic Integration Approach:



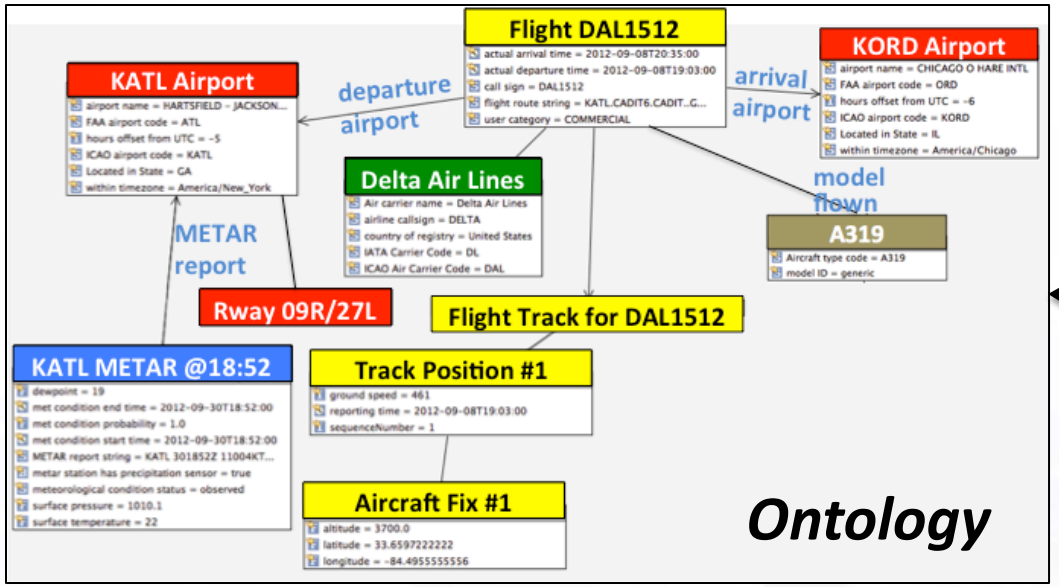
Data Translators

- How is data mapped from the source schemas into the ontology schema?
  - custom translator is written for each data source



  - similar in spirit to data warehouse ETL tools

Example: Mapping an ASDI Departure Message

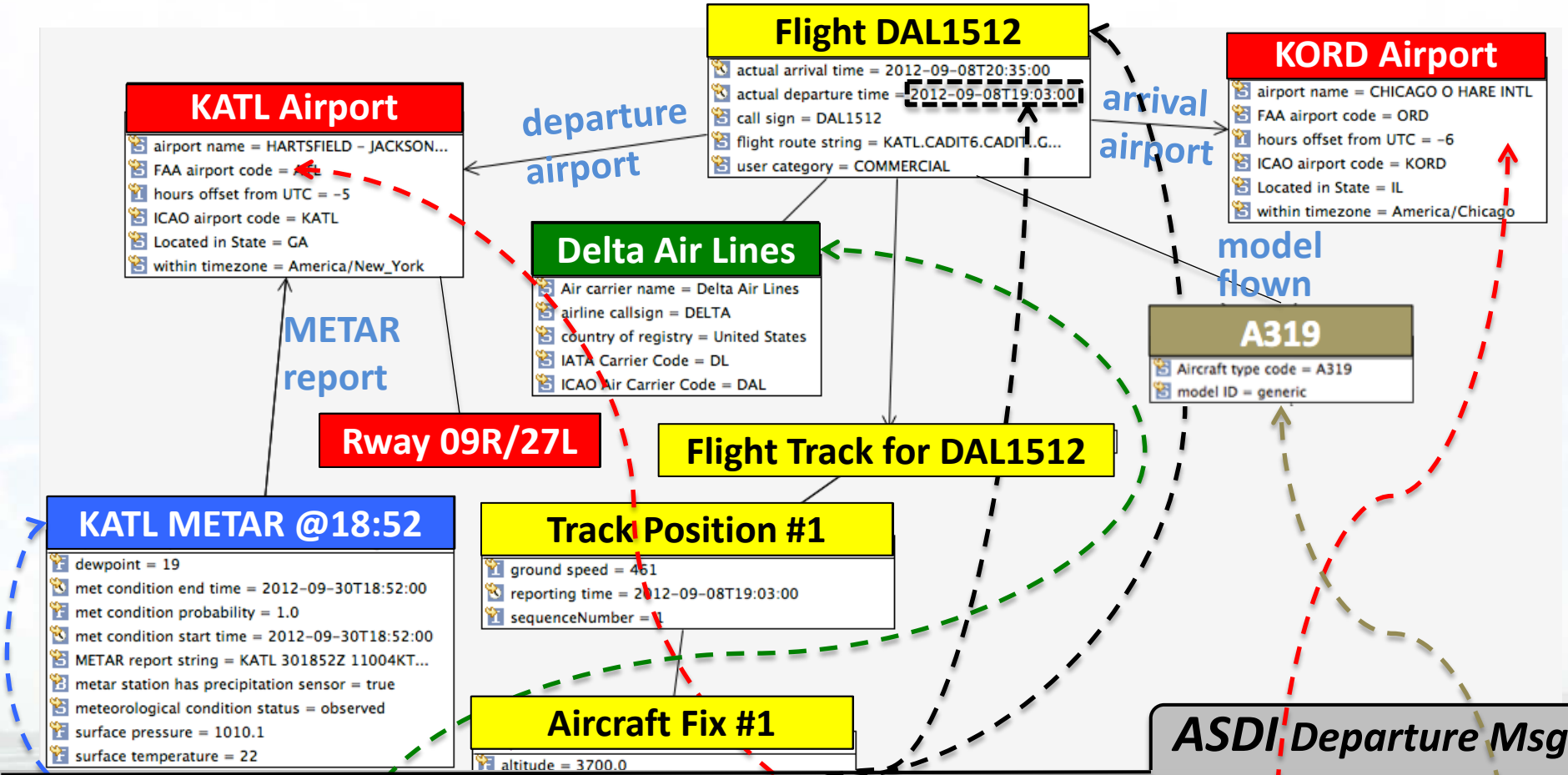


Data Transformation

ASDI Departure Msg

Message-Time-UTC	AC-ID	Departure-Time-UTC	Departure-Named-Fix	Arrival-Named-Fix	AC-Type
2012-09-08 19:02:35	DAL1512	2012-09-08 19:03:00	KATL	KORD	A319

Example: Mapping an ASDI Departure Message



ASDI Departure Msg

Message-Time-UTC	AC-ID	Departure-Time-UTC	Departure-Named-Fix	Arrival-Named-Fix	AC-Type
2012-09-08 19:02:35	DAL1512	2012-09-08 19:03:00	KATL	KORD	A319

Querying the Ontology

- Querying = graph-matching:
  - Each query represents a graph pattern
  - The pattern is matched against the ontology network and all possible matches are returned
- SPARQL: W3C standard ontology query language (uses SQL-like syntactic constructs)
- Benchmark Queries:
  - Set of 17 queries developed to evaluate query performance as ontology scales up
  - Query solutions all require integrated data; none can be answered using a single data source alone

Representative Queries

(restricted to flights on 9/8/12, arriving/departing KATL)

- Flight Demographics:
  - F1: Find Delta flights using A319s departing ZTL airports
  - F3: Find flights with rainy departures from ATL
- Sector Capacity:
  - S4: Find which sector controlled the most flights during a given hour
  - S6: Find the busiest sectors in the NAS on a given day, aggregating hourly
- FAA Advisories / TMIs
  - T1: Find flights that were subject to GDP Advisories
- Weather-Impacted Traffic (WITI) Calculation
  - W1: Calculate hourly WITI values (High Wind, Low Ceiling, Low Visibility)
- ASPM (Flight Delay) Data
  - A3: Compare ASPM AAR with Arrival Demand on an hourly basis at an airport

Status

- Right now, ATM Ontology is just a prototype
  - Includes over 380K instances of ATM objects/properties
- Working to deploy a test version @ NASA
- Initial results promising, but scale-up will be challenging
- Key tasks ahead:
  - Increase scale
  - Increase scope
  - Develop query interface

Collaborators and Funding

Rich Keller Mei Wei
Intelligent Systems Division

Shubha Ranjan Michelle Eshow
ATM Data Warehouse Group
Aviation Systems Division

NASA Ames Research Center
Contact: rich.keller@nasa.gov

Funded by NASA Aeronautics Research Mission Directorate
Aviation Operations & Safety Program

?? Questions ??

?? Comments ??