Data Integration using the NASA Air Traffic Management Ontology

Rich Keller*   Shubha Ranjan**

Mei Wei*   Michelle Eshow**
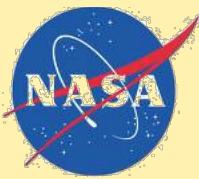
*Intelligent Systems Division   /   **Aviation Systems Division

NASA Ames Research Center

Point of contact:
    rich.keller@nasa.gov

Background:
NASA's ATM Data Warehouse

- NASA researchers require 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 ATM Data Warehouse Holdings

- Flight plans & tracks
  - Airline Situation Display to Industry (ASDI)*
  - Air Route Traffic Control Center tracks
  - TRACON tracks
  - Center-TRACON Automation System (CTAS)
  - Exelis tracks

  SWIM conversion underway for available sources

- Weather
  - METAR, TAF
  - Corridor Integrated Weather Service (CIWS)
  - AIREP, PIREP
  - Rapid Refresh (RR) Weather Forecast

- Traffic Management
  - Advisories & TMIs
  - Time-based Flow Management (TBFM)
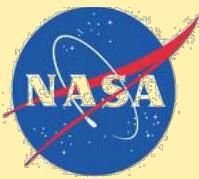
ATM Data Warehouse: A microcosm of the NAS data environment

Problem:
Non-integrated Data

- ATM Warehouse data is replicated & archived in its original format

- Data sets lack standardization
  - data formats
  - nomenclature
  - conceptual structure

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

- To analyze and mine data, researchers must write special-purpose code to integrate data for each new task
  - ➔ <u>Huge time sink</u>!
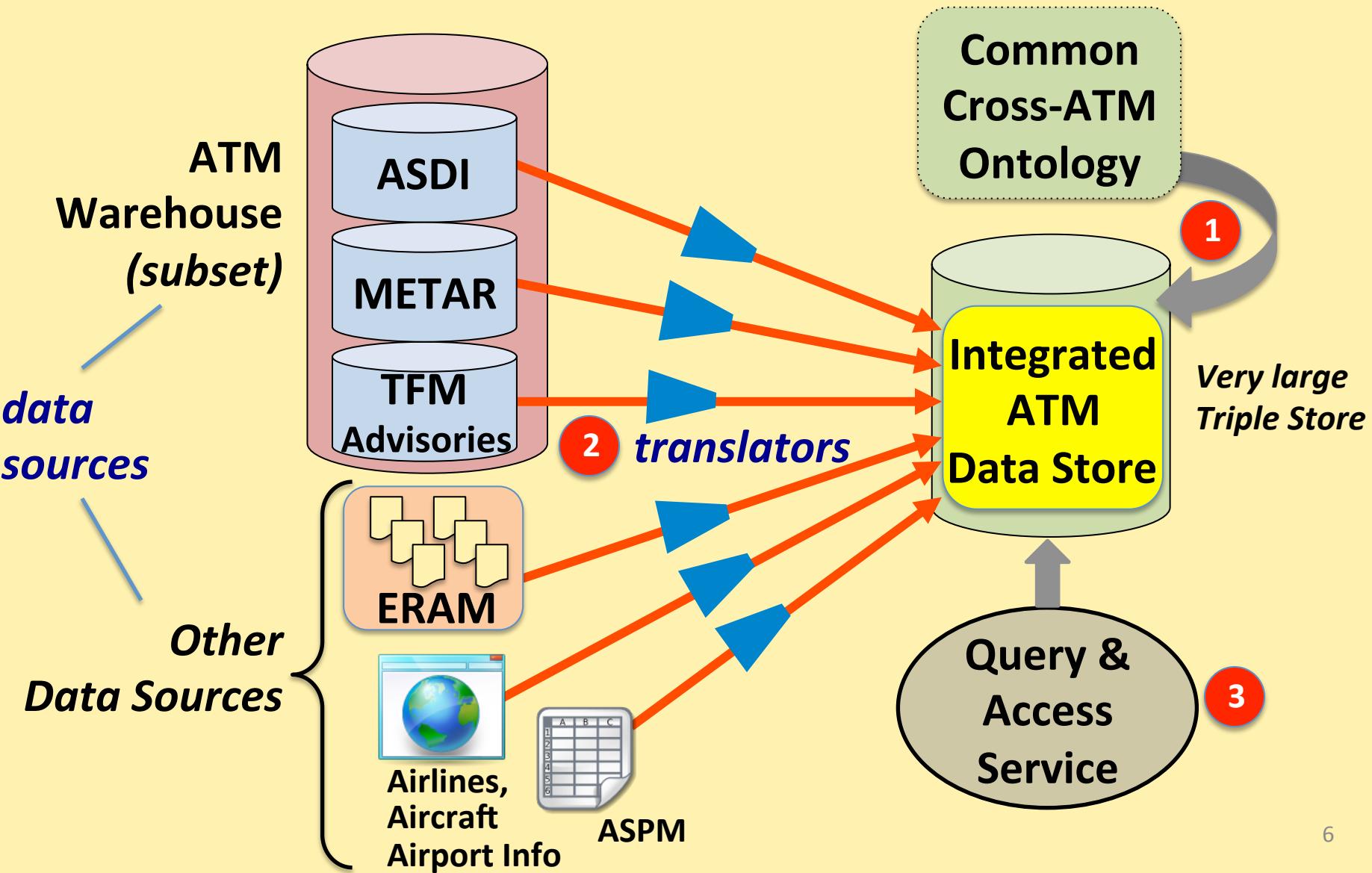
Proposed Solution

Relieve users of responsibility for integration!

Integrate Warehouse data sources

on the server side

using **Semantic Integration**

Semantic Integration Approach:
Prototype System Diagram

ATM Warehouse *(subset)*

ASDI

METAR

TFM Advisories

data sources

2 *translators*

Other Data Sources

ERAM

Airlines, Aircraft Airport Info

ASPM

Common Cross-ATM Ontology

1

Integrated ATM Data Store

*Very large Triple Store*

3

**Query & Access Service**
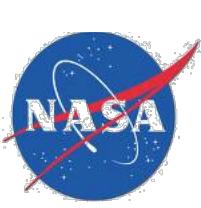
What is modeled in the NASA ATM Ontology?
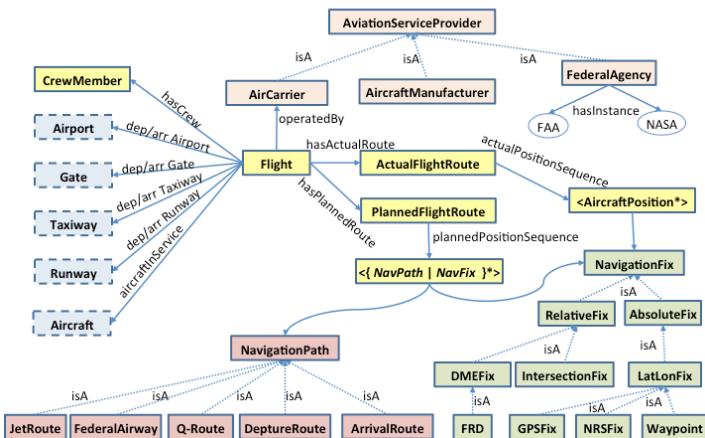
❖ **150+ object types:** •Flights •Aircraft and manufacturers •Airlines •Airports and physical infrastructure •NAS facilities •Air traffic management initiatives •Surface weather conditions and forecasts •Airspace sectors, fixes, routes, airways •Flight plans and paths

❖ **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

❖ **2M+ triples in prototype** (one day of ops at one major airport)
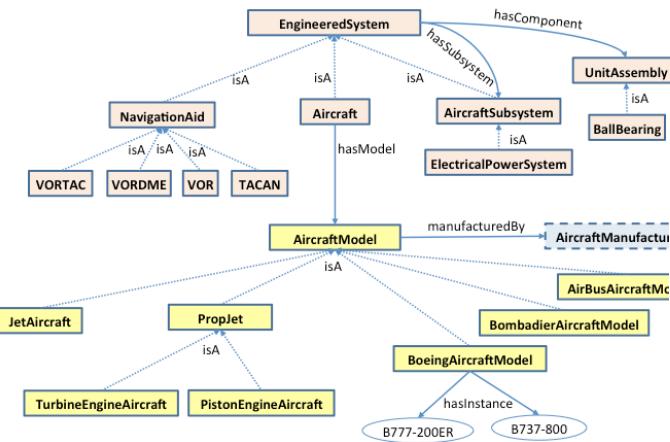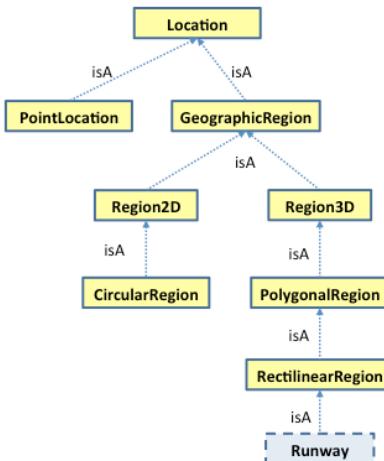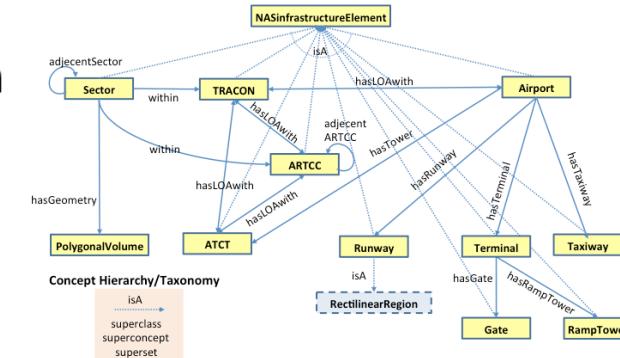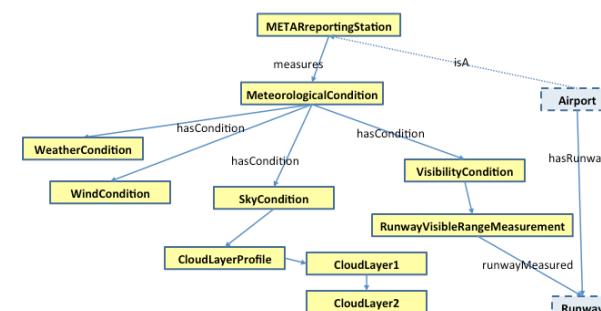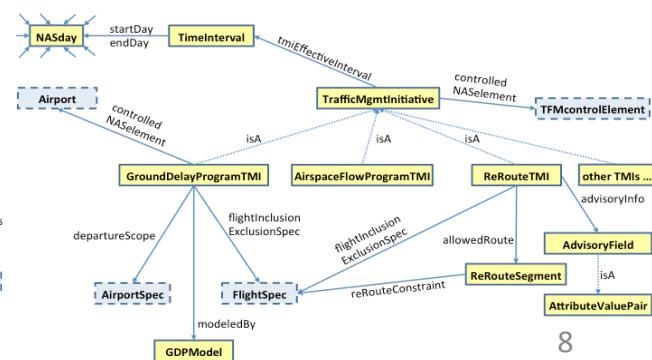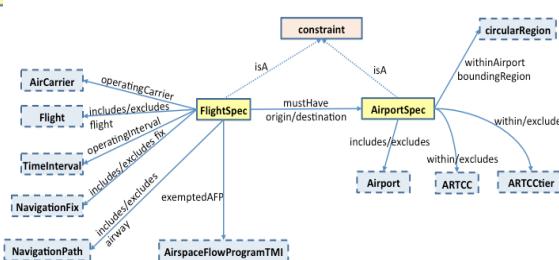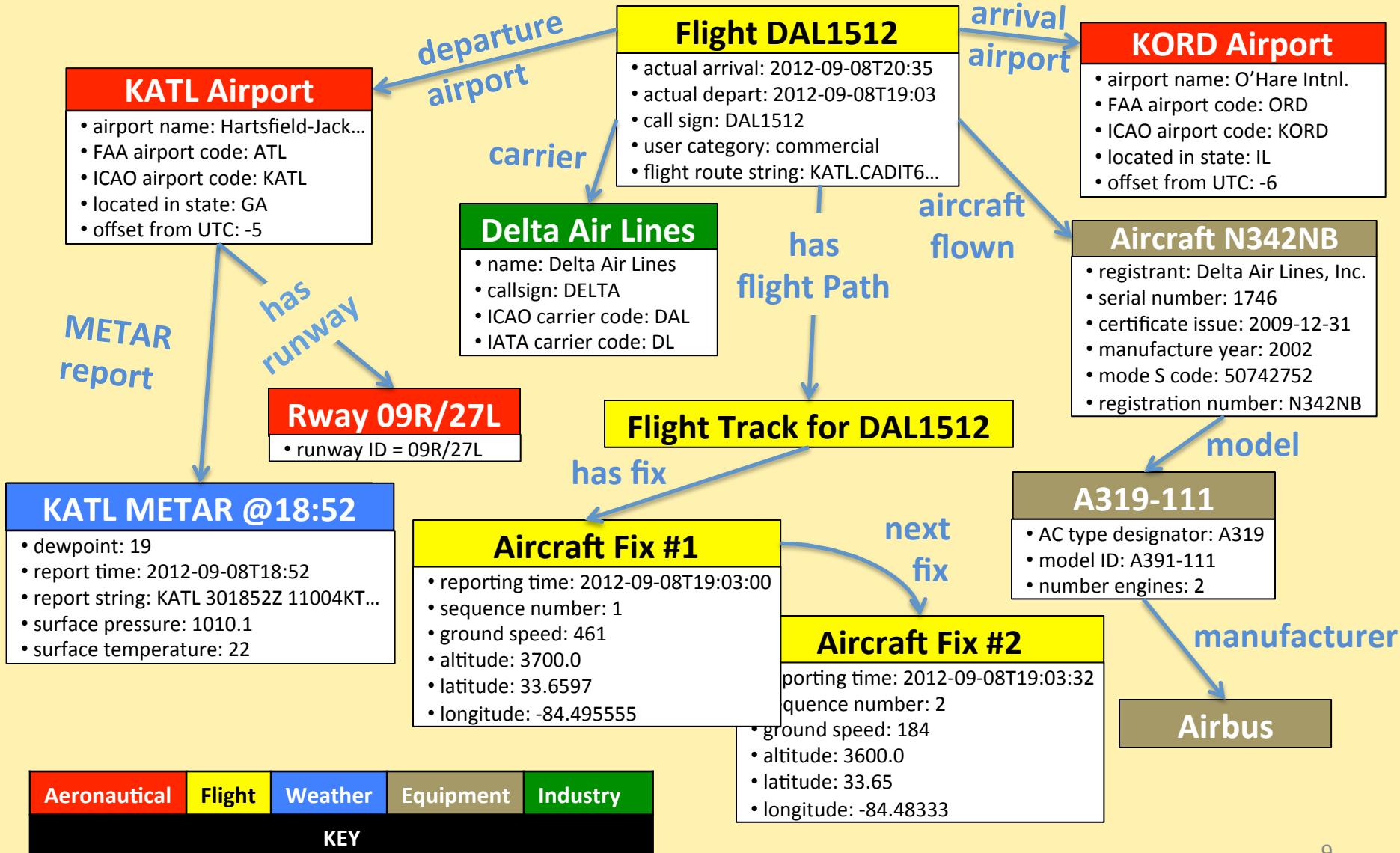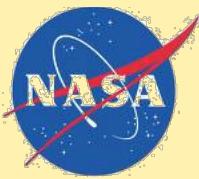
Ontology Subsets
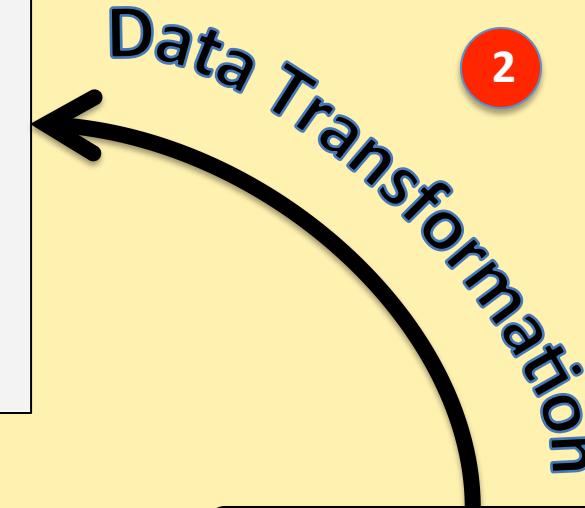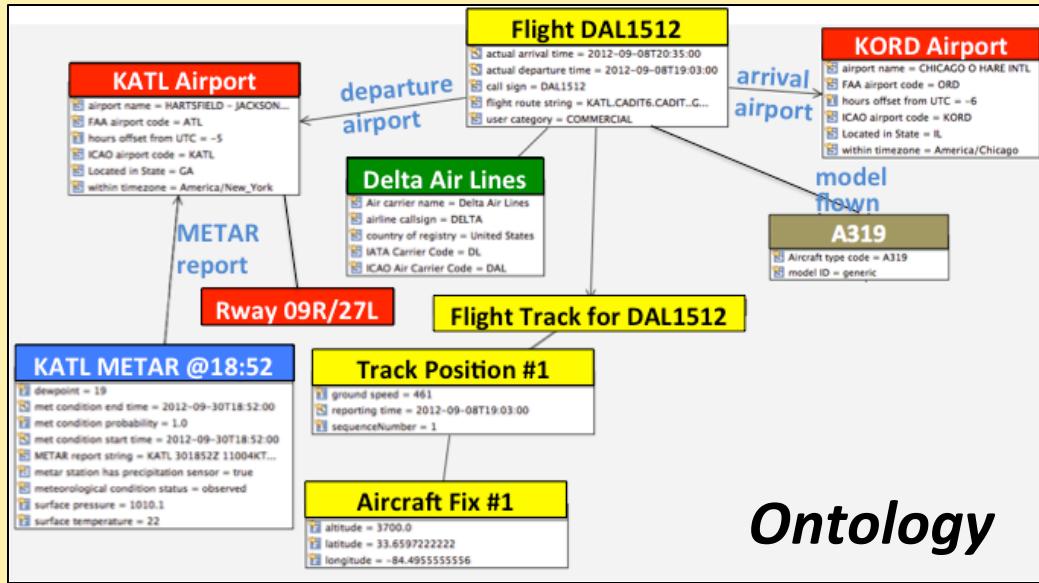
Ontology Representation of a Flight

Flight DAL1512
- actual arrival: 2012-09-08T20:35
- actual depart: 2012-09-08T19:03
- call sign: DAL1512
- user category: commercial
- flight route string: KATL.CADIT6…

departure airport

KATL Airport
- airport name: Hartsfield-Jack…
- FAA airport code: ATL
- ICAO airport code: KATL
- located in state: GA
- offset from UTC: -5

arrival airport

KORD Airport
- airport name: O'Hare Intnl.
- FAA airport code: ORD
- ICAO airport code: KORD
- located in state: IL
- offset from UTC: -6

carrier

Delta Air Lines
- name: Delta Air Lines
- callsign: DELTA
- ICAO carrier code: DAL
- IATA carrier code: DL

aircraft flown

Aircraft N342NB
- registrant: Delta Air Lines, Inc.
- serial number: 1746
- certificate issue: 2009-12-31
- manufacture year: 2002
- mode S code: 50742752
- registration number: N342NB

has flight Path

METAR report

has runway

Rway 09R/27L
- runway ID = 09R/27L

Flight Track for DAL1512

model

KATL METAR @18:52
- dewpoint: 19
- report time: 2012-09-08T18:52
- report string: KATL 301852Z 11004KT…
- surface pressure: 1010.1
- surface temperature: 22

has fix

Aircraft Fix #1
- reporting time: 2012-09-08T19:03:00
- sequence number: 1
- ground speed: 461
- altitude: 3700.0
- latitude: 33.6597
- longitude: -84.495555

next fix

Aircraft Fix #2
- reporting time: 2012-09-08T19:03:32
- sequence number: 2
- ground speed: 184
- altitude: 3600.0
- latitude: 33.65
- longitude: -84.48333

A319-111
- AC type designator: A319
- model ID: A391-111
- number engines: 2

manufacturer

Airbus

KEY

| Aeronautical | Flight | Weather | Equipment | Industry |

9

Example: Mapping an ASDI Departure Message onto Ontology

Ontology

Data Transformation

②
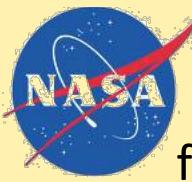
ASDI Departure Msg

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

Original Source Format

Representative SPARQL Queries ③

from benchmark set of 17 queries for evaluating performance on scale-up

- 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 Index (WITI)
  - W1: Calculate hourly (High Wind, Low Ceiling, Low Visibility) WITI values

- Flight Delay Data
  - A3: Compare hourly Airport Acceptance rate with Arrival Demand at an airport
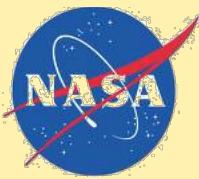
Status

- Right now, ATM Ontology is just a prototype
  - Ontology contains integrated data corresponding to a single day of NAS operations at a major airport (9/8/12 @ ATL) using a subset of ATM sources

- Two commercial triple stores test-deployed on local server:

  - AllegroGraph (from Franz)

  - GraphDB (from OntoText)

- Within NASA Ames, triple stores can be queried via HTTP as a SPARQL endpoint

Future Plans

- Increase scale (a key challenge!)
  - 2.4M triples for one day; 36M for 30 days
  - only flights arriving/departing one airport

- Increase scope: additional data sources

- Build tools and services on top of triple store:
  - data browser
  - data query interface
  - data download service

Long Term Goal

To build the world's largest repository of
Linked Open Data
describing the Global Airspace System

   – A queryable resource for aviation applications,
research, analysis, and public policy decision-making

Featuring interconnected data about…

- Flights
- Airports
- Airlines
- Runways/taxiways
- Terminals/Gates
- Airspace control facilities
  (ARTCCs, TRACONs, towers)
- Air traffic management initiatives

- Weather
- Aircraft
- Aircraft mechanical systems
- Aviation safety data
- Aircraft manufacturers
- Airspace topology
  (sectors, fixes, routes)
- Departure/Arrival routes