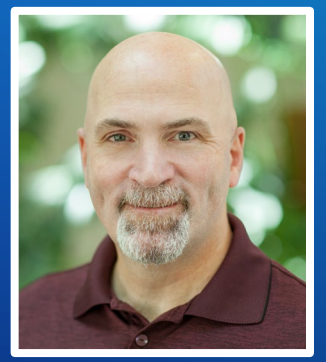
ArcGIS Velocity: Applying Real-Time and Big Data Analytics

Ken Gorton and Peter Nasuti



Ken Gorton
ArcGIS Velocity
Product Engineer
kgorton@esri.com

Peter Nasuti
ArcGIS Velocity
Product Engineer
pnasuti@esri.com

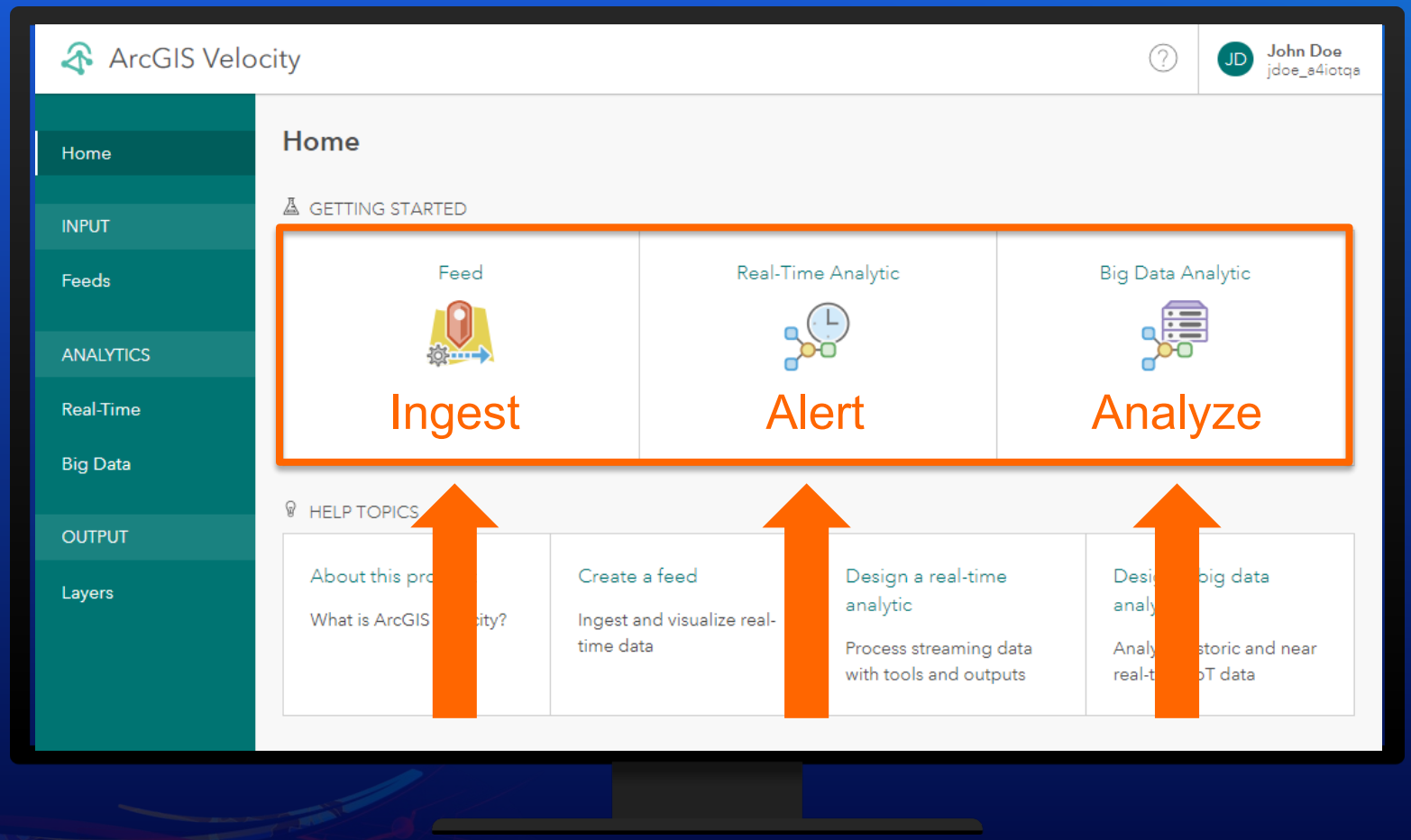
2021 ESRI DEVELOPER SUMMIT

ArcGIS Velocity

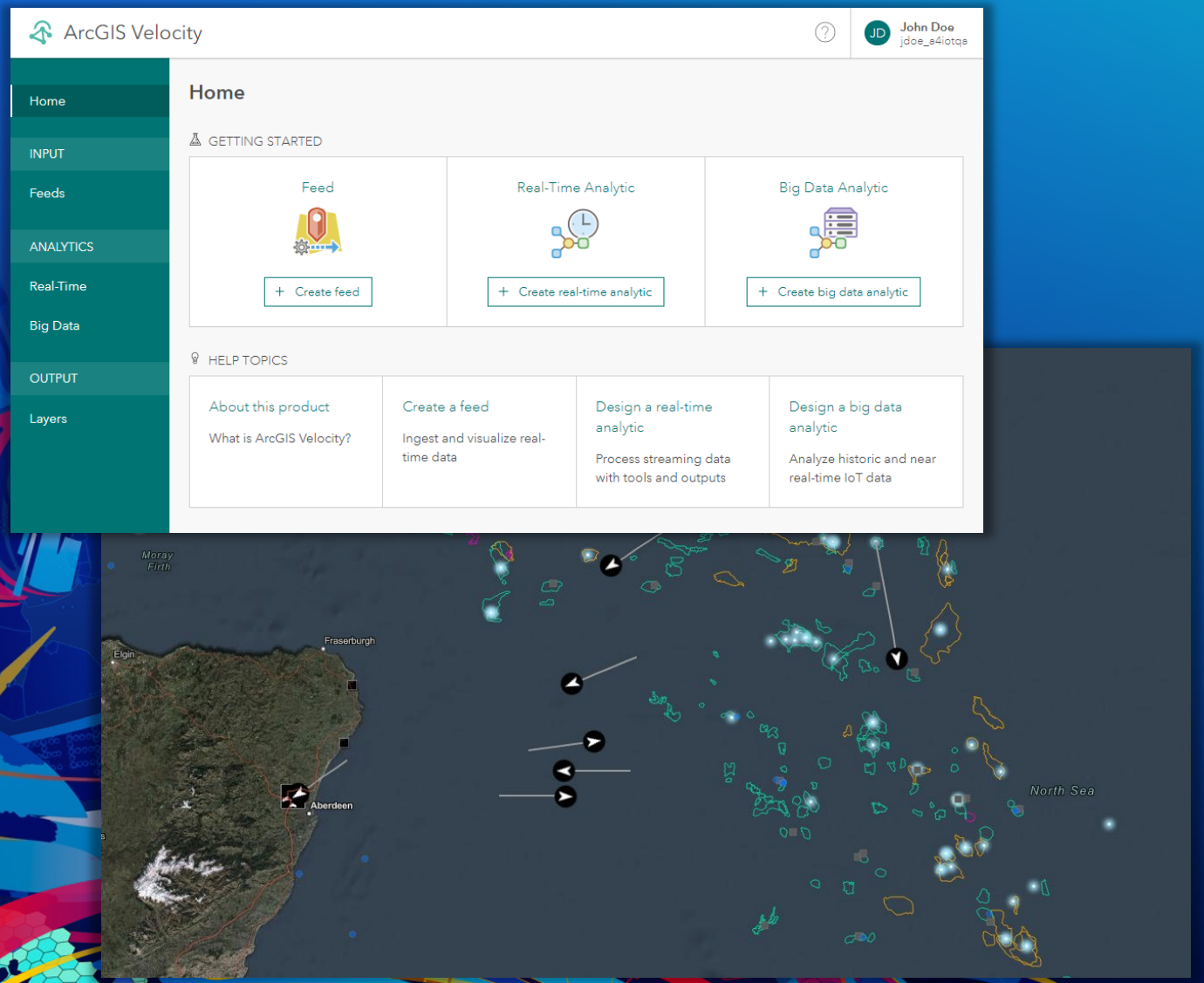
Real-time and big data capabilities for ArcGIS Online

- **Ingest real-time sensor data**
- **Analyze data to detect incidents**
  - In real-time
  - In *near* real-time
  - Over time
- **Send alerts and take action**

New Application – ArcGIS Velocity



…ingest, process, and analyze your real-time data with ArcGIS Online



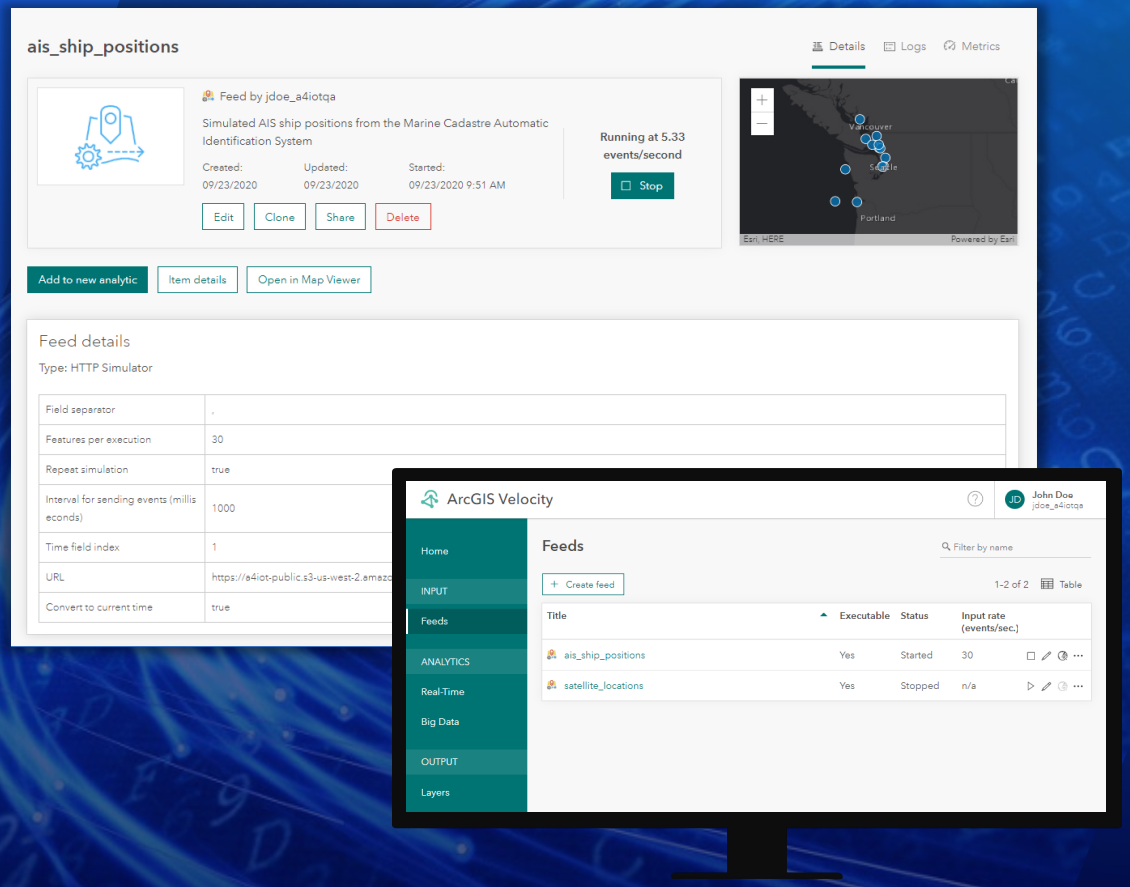
Demonstration

Feeds

Feeds

Ingest and visualize real-time data streams

- **Connect to data sources**
  - ArcGIS (Feature & Stream Layers)
  - Cloud IoTs (Amazon, Azure, Cisco)
  - Web & Messaging (HTTP, Kafka, RSS…)
- **Schema auto-discovery**
- **Behaves like a stream layer**
  - Immediate display of new data

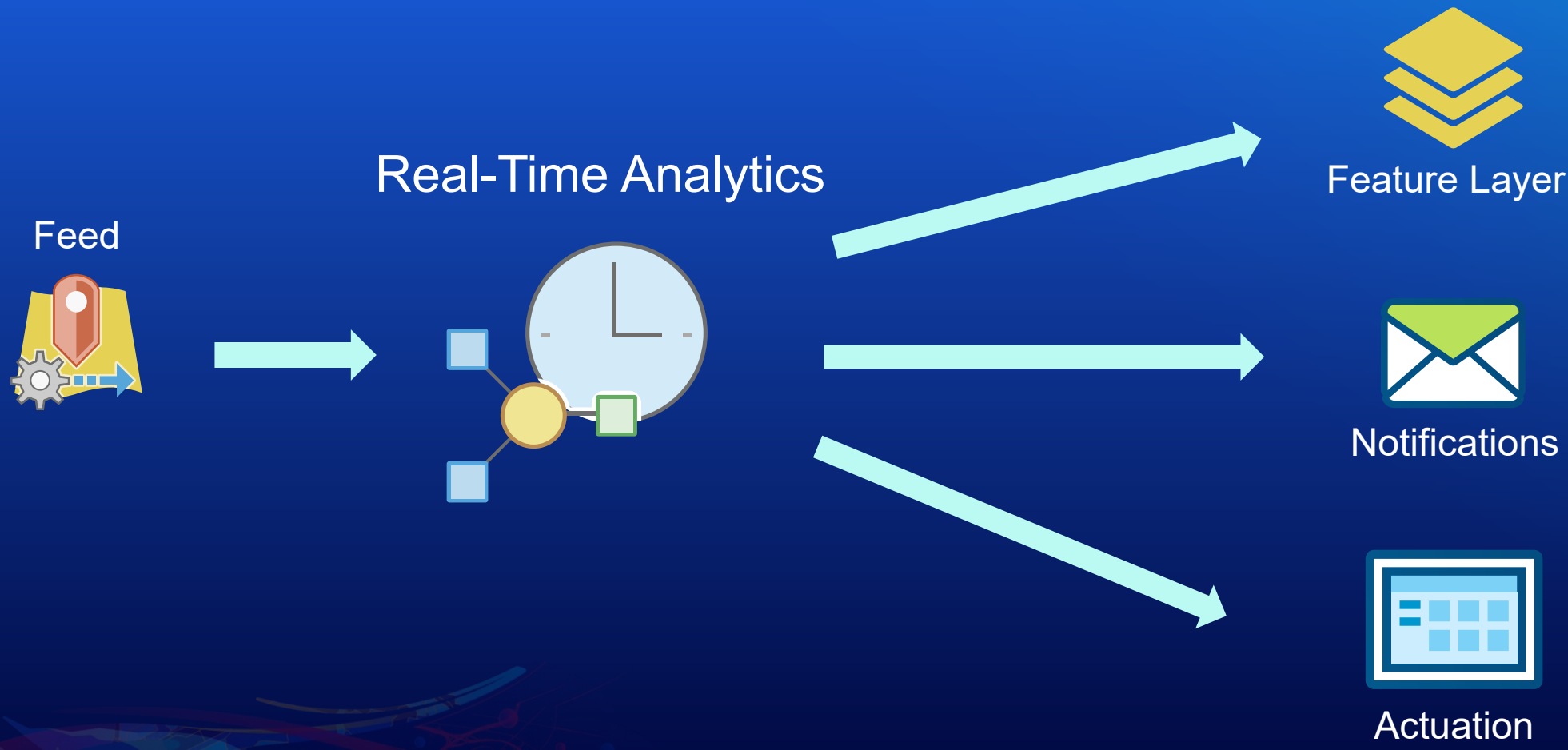


…discover and share real-time data streams as items in ArcGIS

Real-time Analytics

Real-time analysis

Process and analyze real-time data streams



Input & output types

Supported types and formats

Feeds and data sources

Feature Layer

Azure Event Hub

HTTP Poll

Stream Layer

Azure Service Bus

AWS IoT

HTTP Receive

Kafka

RabbitMQ

Cisco Kinetic EFM

MQTT

Azure Blob Store

Web Socket

Simulator

Azure Cosmos DB

AWS S3

Outputs

Azure IoT Hub

Feature Layer

Stream Layer

Kafka

AWS S3

Email

Text Message

RabbitMQ

Azure Blob Store

HTTP Post

Formats

Delimited

JSON

Feature JSON

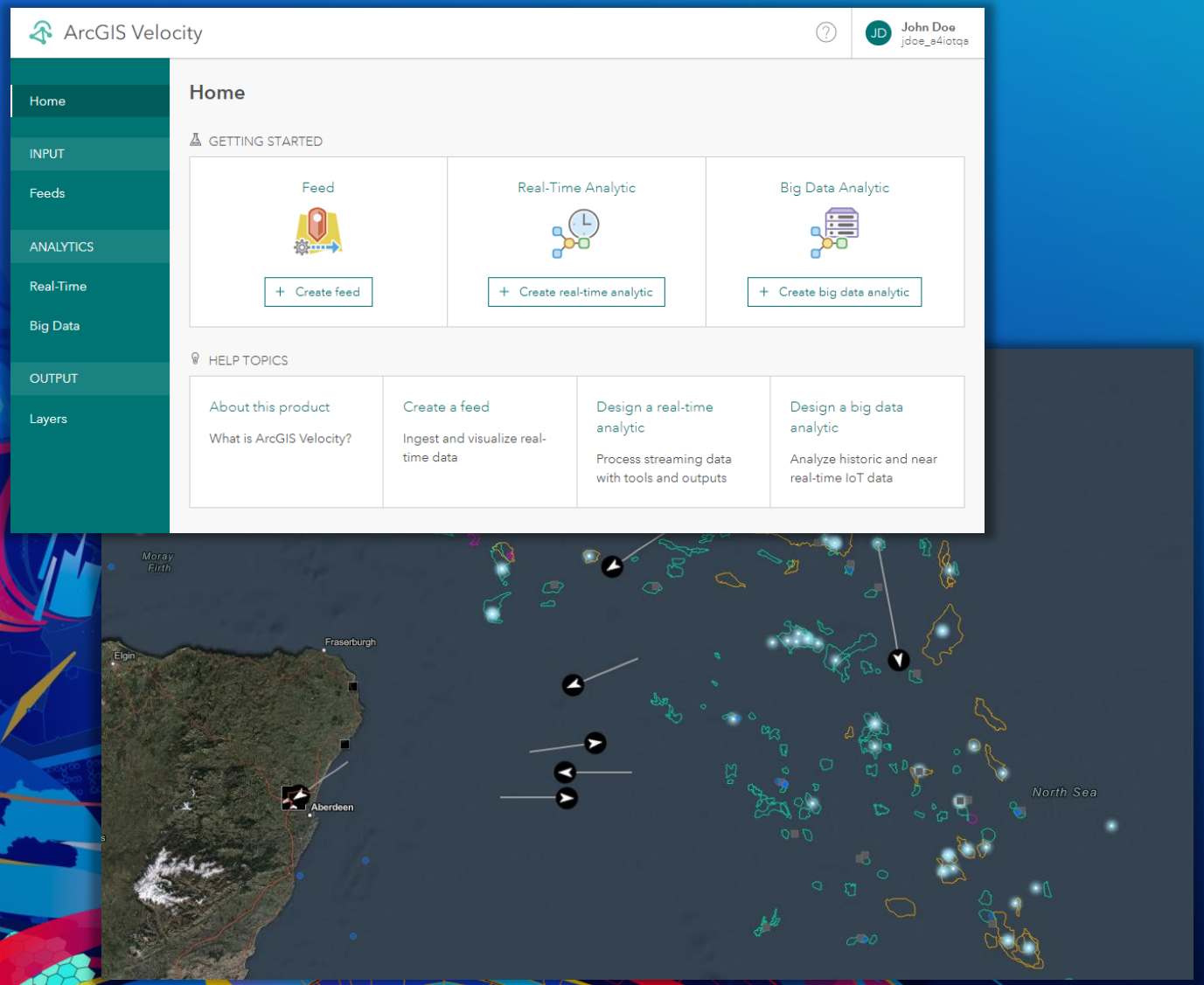
Parquet

GeoJSON

RSS

XML

Shapefile

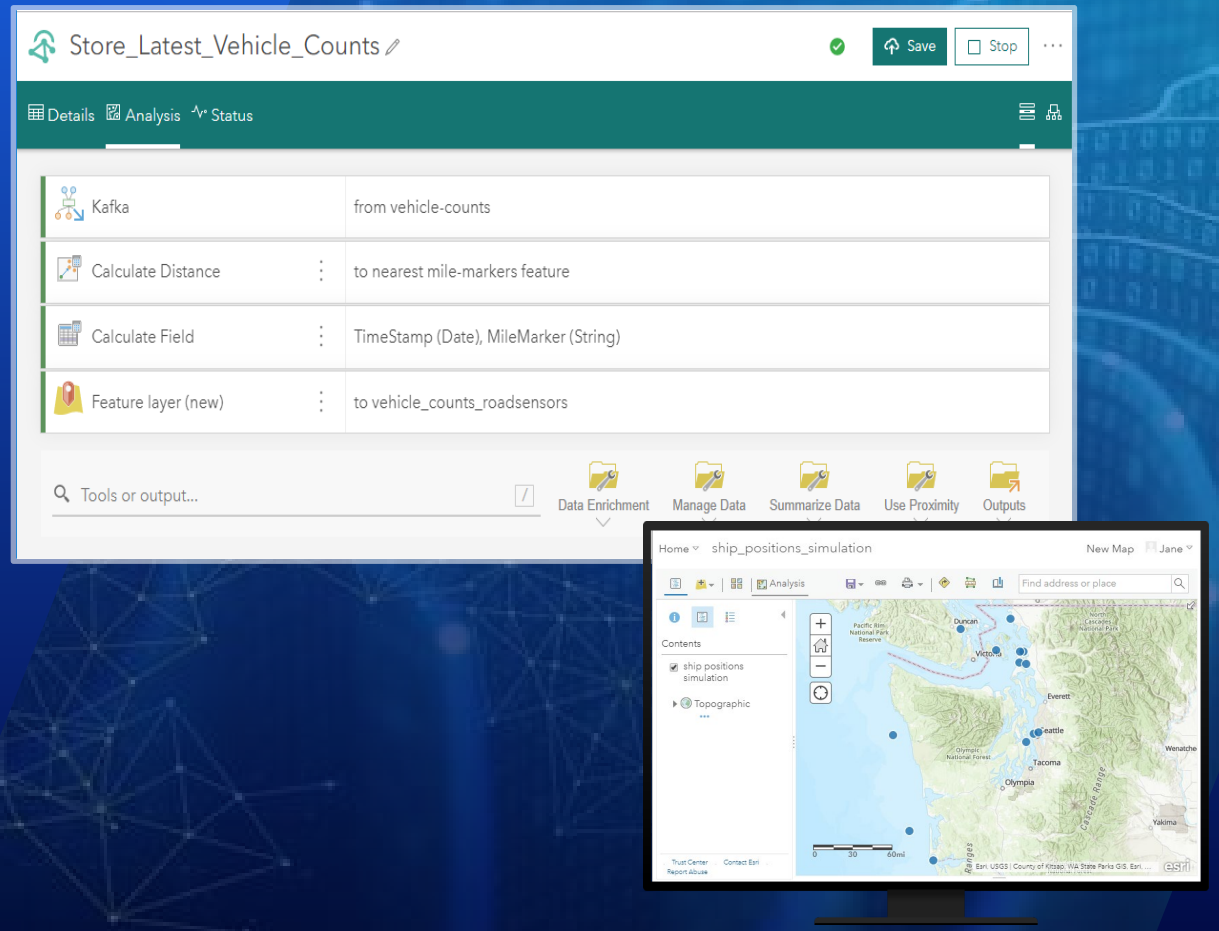


Demo: Real-time processing and alerting of aircraft ETA

Real-Time Analytics

Process and analyze real-time data streams

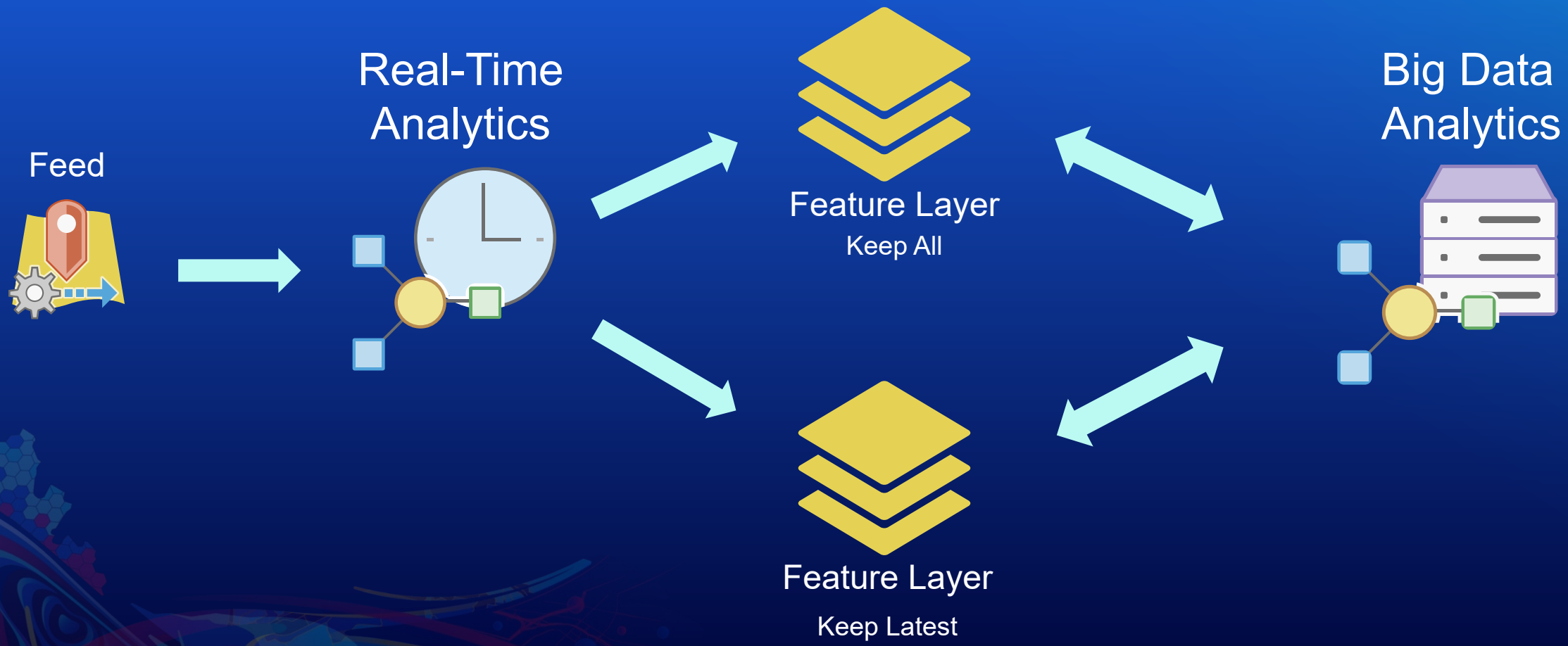
- **Input Sources**
  - Feeds (one or more)
  - Static data sources (enrichment, joins)
- **Analytic Tools**
  - Build a pipeline of zero to many….
  - Analyze individual observations
- **Outputs**
  - ArcGIS Layers (Map & Feature)
  - Alerts & Notifications (Email, Actuations)



…generating alerts from and triggering actions to IoT devices in real-time

Accumulating feature layer data for future analysis

Evaluate incidents, trends, and patterns against any decision timeline



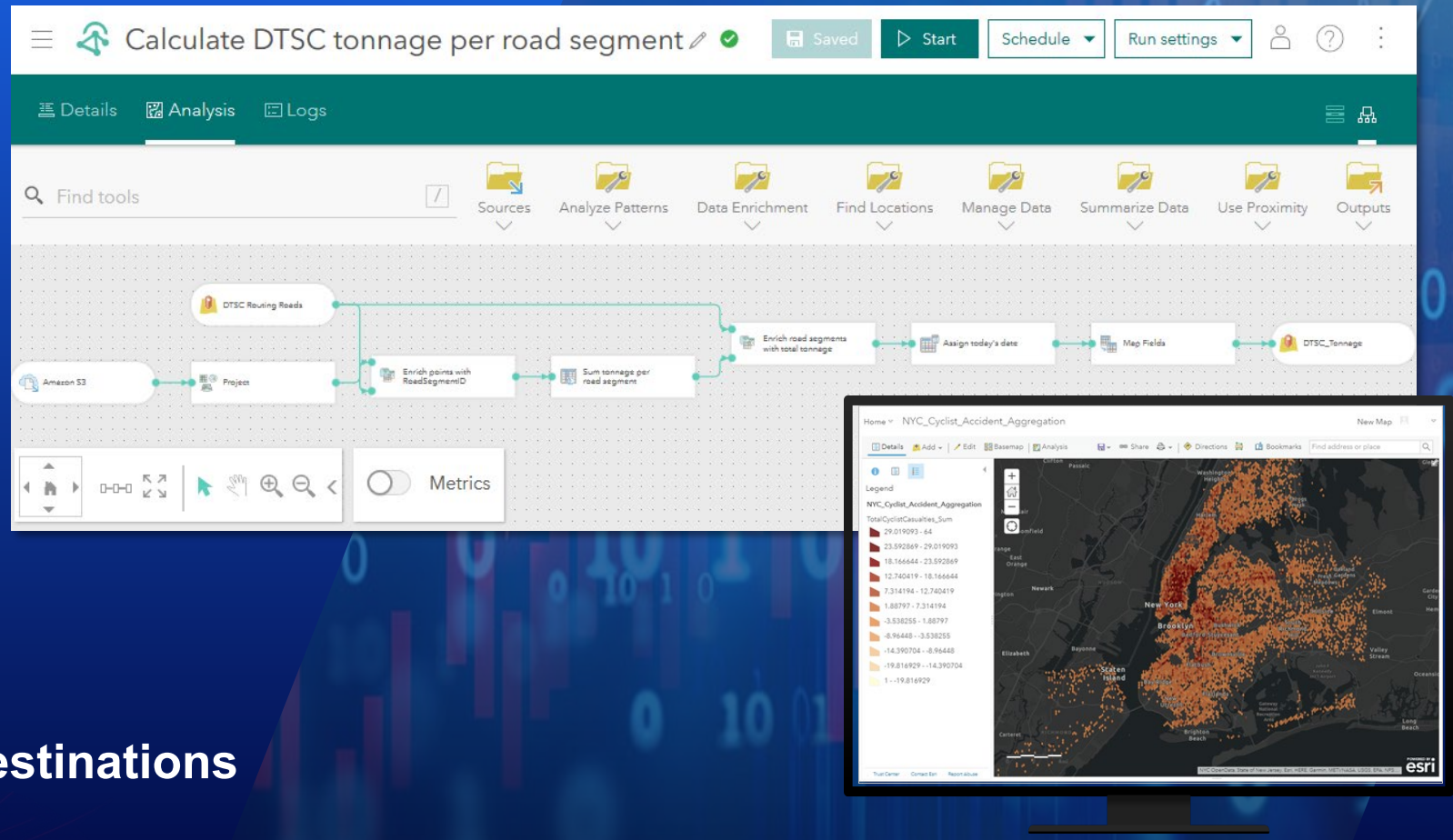
Big Data Analytics

Create analytic pipelines to process batches of data

Big Data Analytics

Perform batch analysis on stored big data

- **Input Sources**
  - ArcGIS (Feature Layers)
  - Cloud (Amazon S3, Azure Blob, …)
  - Web & Messaging (HTTP, RSS…)
- **Analytic Tools**
  - Analyze Patterns
  - Find Locations
  - Manage, Summarize, Enrich Data
- **Output results to one or more destinations**



…big data analysis pipelines on demand or scheduled for recurring execution

Source & output types

Supported types and formats for big data analytics

Sources

- Feature Layer
- Amazon S3
- Azure Blob Store
- Azure Cosmos DB
- HTTP Poller
- RSS Poller

Outputs

- Feature Layer
- Stream Layer
- Amazon S3
- Azure Blob Store
- Azure Event Hub
- Email
- Text Message
- HTTP
- Kafka

Formats

- Delimited Text
- EsriJSON
- GeoJSON
- JSON
- RSS
- Parquet
- Shapefile
- XML

Analytic tools

Only available in big data analytics

Analyze Patterns

Calculate Density
Calculate a magnitude-per-unit area from point features within a neighborhood around ea...

Calculate Journeys
Calculate polylines representing periods in which a track entity moved consistently or ...

Find Hot Spots
Given a set of features, identify statistically significant hot spots and cold spots us...

Find Point Clusters
Finds clusters of point features in surrounding noise

Data Enrichment

Calculate Motion Statistics
Calculates motion statistics and measures for event features including distance, time s...

Find locations

Detect Incidents
Creates a unique incident for each continuous set of observations that satisfy conditions

Detect Gaps
Detect the absence of events from streaming data using a Track ID and feature Start Tim...

Find Dwell Locations
Find locations where moving objects have stopped, or dwelled, using given time and dist...

Find Similar Locations
Identify candidate features most similar or dissimilar to input features

Manage Data

Calculate Field
Calculate new field values for features

Dissolve Boundaries
Aggregate features based on specified attributes

Filter by Expression
Filter events or features by attribute values

Filter by Time Range
Filter events or features by time range

Filter by Geometry
Filter features spatially in relation to another dataset

Map Fields
Transform data from an input schema to an output schema

Merge
Combine features of the same type and schema from two datasets or pipelines into a sing...

Overlay Layers
Overlay features from two datasets via Intersect or Erase

Project
Project input data to a spatial reference

Select Fields
Retain a subset of the layer's fields

Summarize Data

Aggregate Points
Use area features to summarize a set of point features. Boundaries from the area featur...

Summarize Within
Calculate statistics in areas where an input layer is within or overlaps a boundary layer

Summarize Attributes
Calculate summary statistics for attributes in a dataset

Join Features
Join attributes from one layer to another based on spatial, temporal, and attribute rel...

Reconstruct Tracks
Connect time-sequential points to tracks and summarize features within the track

Use Proximity

Create Buffers
Buffer incoming features by a specified distance, generating output buffer polygons

Calculate Distance
Calculate the distance from a target feature to the nearest feature in another dataset

Snap to Network
Snap event features to a street segment based on proximity and other measures

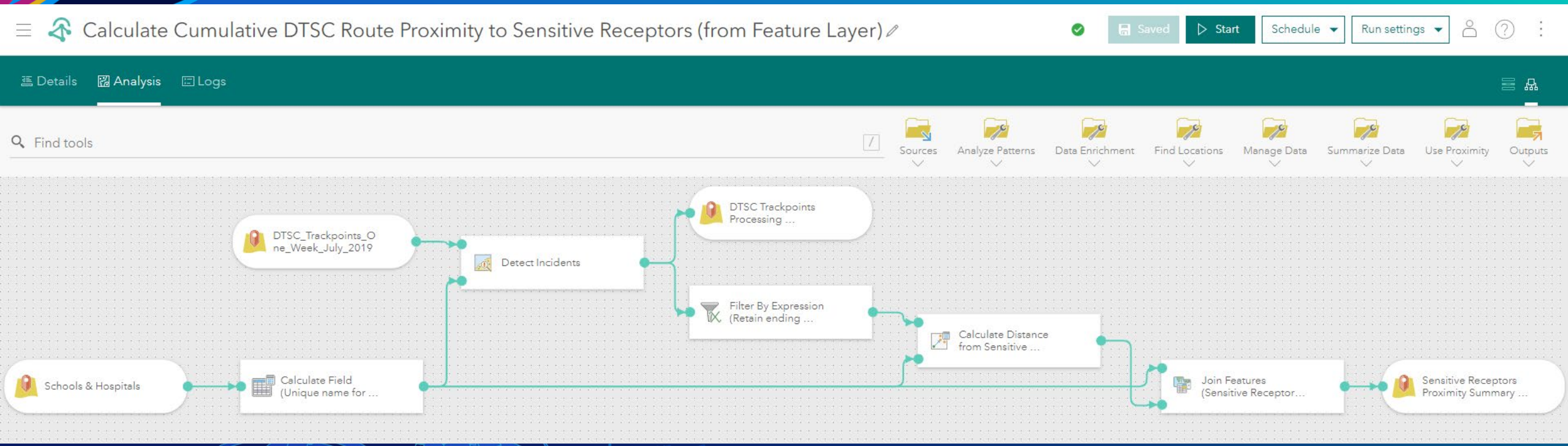
Demo: Analyzing toxic material proximity to schools and hospitals

- Goal of analysis: What schools and hospitals in California are at the greatest risk from the transport of toxic substances?

- To answer this question, we must understand:
  1. Where do trucks go near schools and hospitals?

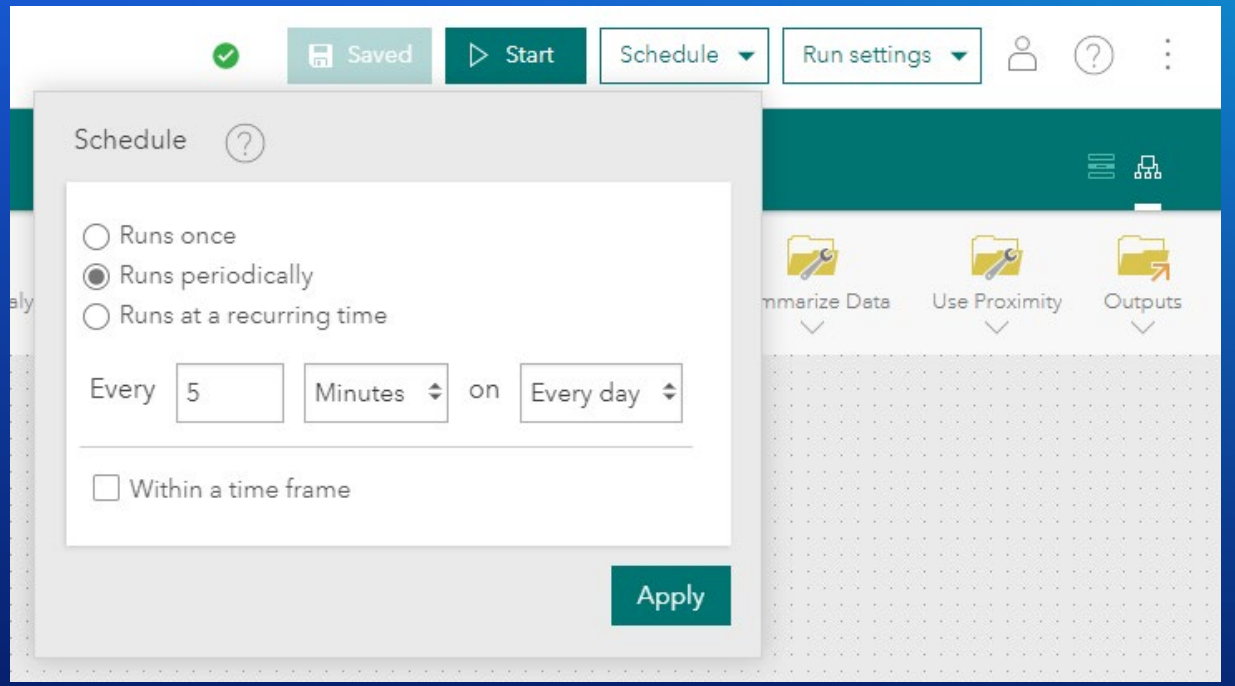
  2. For locations where trucks go near a school or hospital, what is the total cumulative time that trucks have spent within a given distance of a school or hospital?

Demo: Analyzing toxic material proximity to schools and hospitals



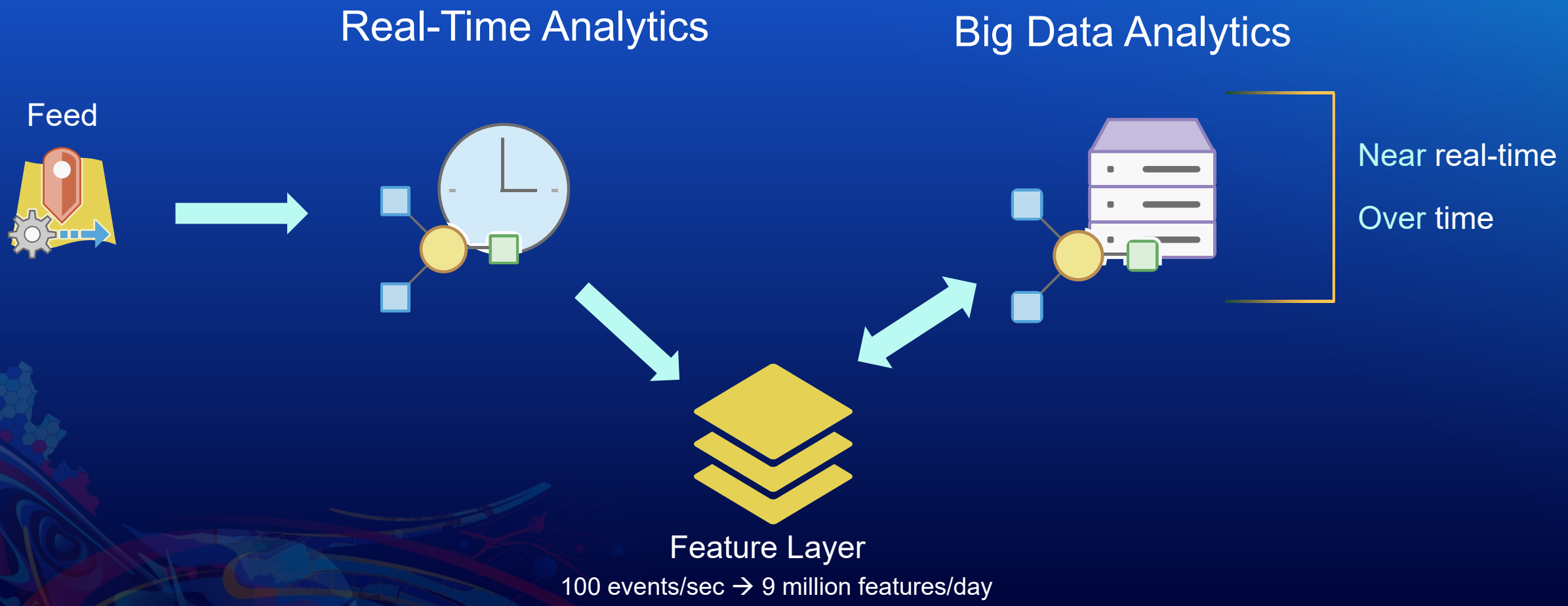
Analytic scheduling

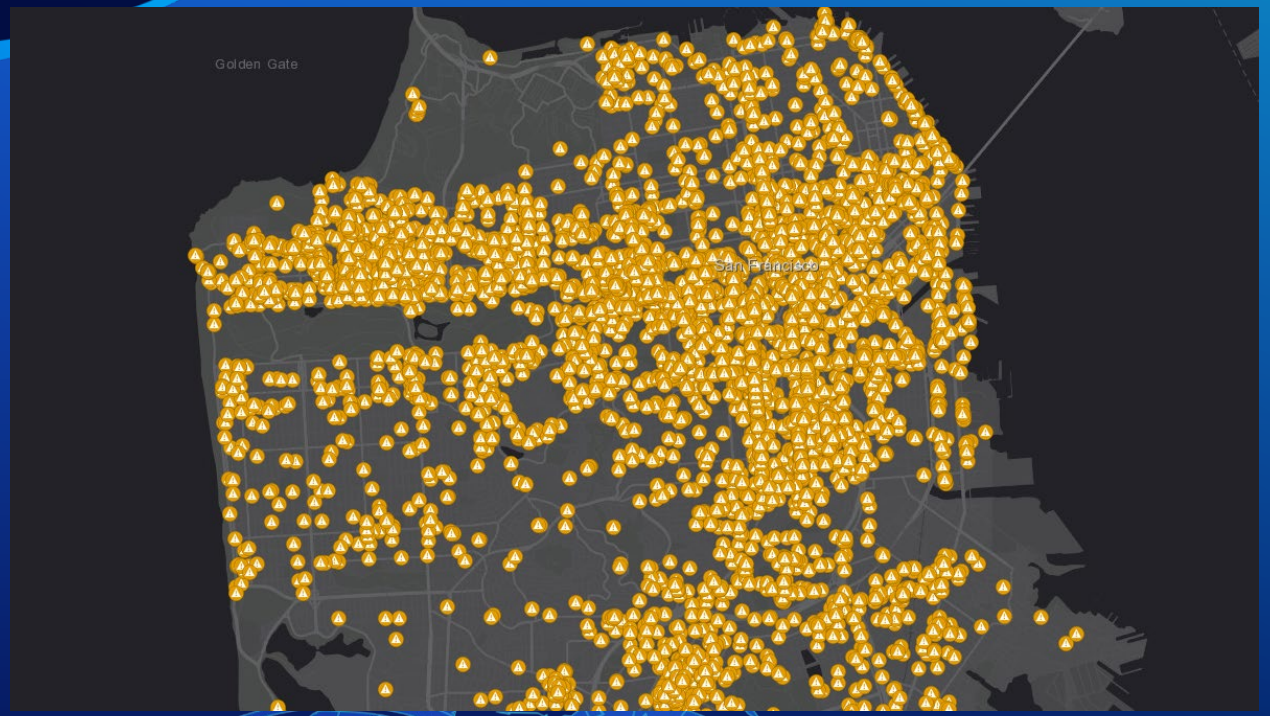
- **Run once**
  - **Runs when the user clicks Start**
- **Run periodically**
  - **i.e. every 2 min, every 5 min, every hour**
- **Run at a recurring time**
  - **i.e. every morning at 6am**



Near-real-time analysis

Leverage big data analytics to process most recent features





Demo: Near-real-time analysis of 311 data

Please provide your feedback for this session by clicking on the session survey link directly below the video.