

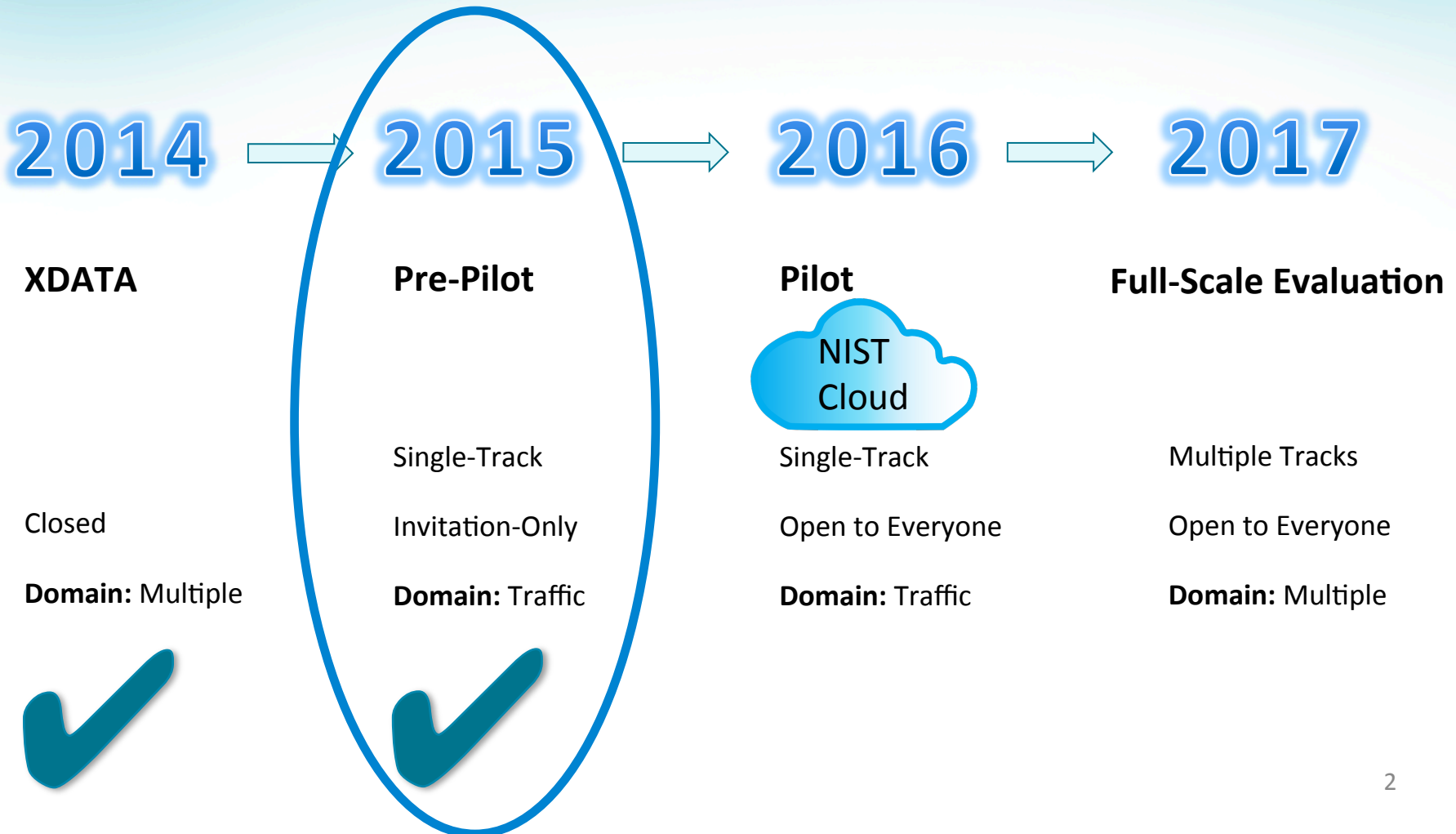


# NIST Data Science Evaluation Series Pre-Pilot Overview

Peter Fontana,  
March 17–18, 2016



# DSE - Schedule



# Pre-Pilot Data Sets

- Traffic Lane Detectors
- Traffic Events Listing
- Traffic Camera Video

# Other Publicly-Available Sets

- OpenStreet Maps
- U.S. Census Data, Including American Community Survey (ACS)
- NOAA Integrated Surface Hourly Weather Data Sets
- Publicly Available Sets welcome
- **Note: Cleaning Task restricts allowable data**



# Traffic Camera Video



# Traffic Event Listing

- **Event\_ID:**  
MDOT\_CHART\_a000a2de48a600680055fa2ec4235co  
a
- **Description:** "Incident @ I-495 INNER LOOP PRIOR TO EXIT 33 MD 185 CONNECTICUT AVE [Collision, Personal Injury]"
- **Creation Time:** 5/29/15 13:50
- **Incident Type:** accidentsAndIncidents, injury accident,
- **Location (lat, lon):** 39.00536,-77.077752,5,3

# Traffic Detector Data

Lane_id	Measurement start	Speed	Flow
156053	2015-05-18 10:23:20.033-04	29.80	8
156054	2015-05-18 10:23:20.033-04	24.20	22
156056	2015-05-18 10:23:20.033-04	37.20	27

# Pre-Pilot Data Information

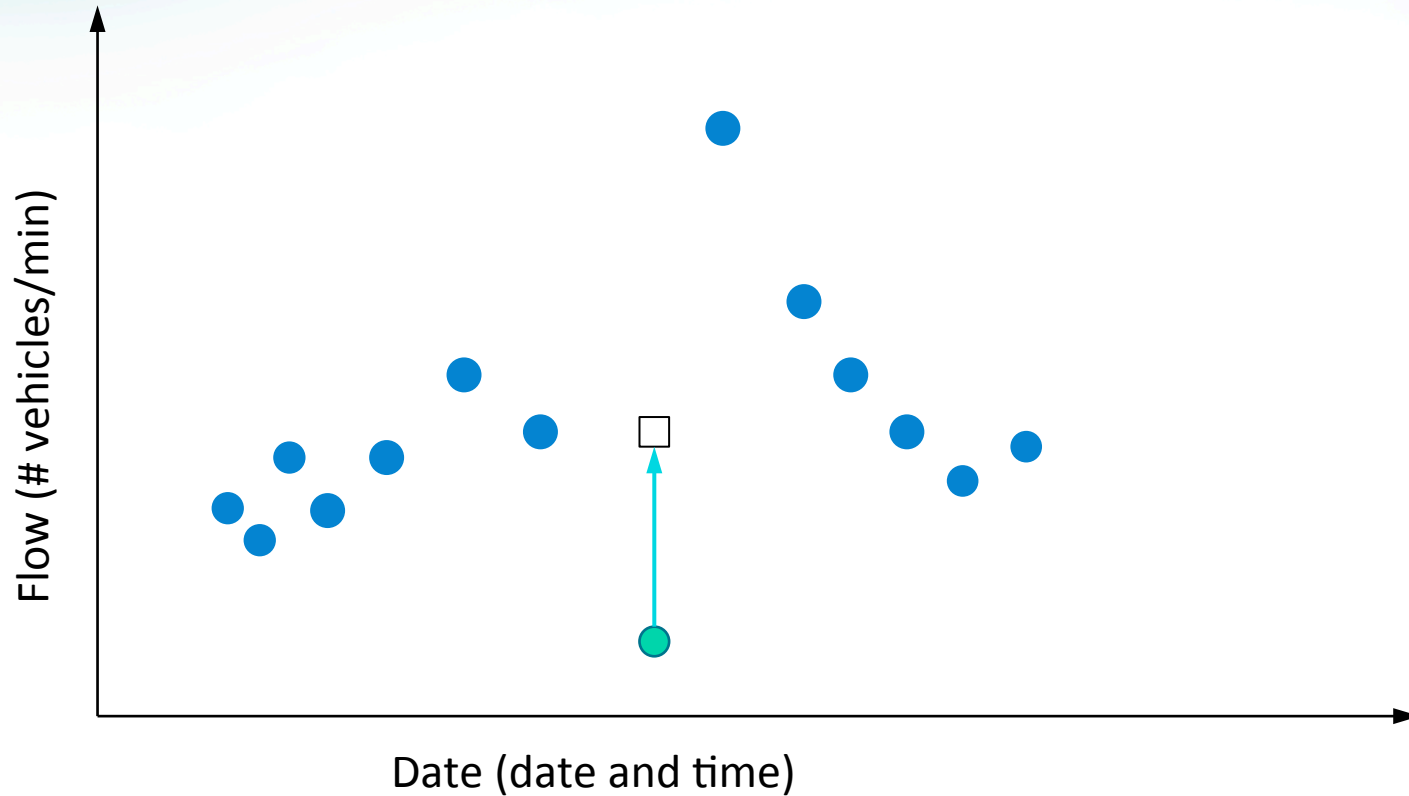
- **Traffic Lane Detectors:** Traffic Detectors on Major routes in the MD-DC-VA area from September 2006–August 2015, 150 GB
- **Traffic Events Listing:** Traffic Events in the MD-DC-VA Area from October 2003–August 2015, 200 MB.
- **Traffic Camera Video:** Traffic Camera Video Feeds for 20 MD Traffic Camera, June 2015–August 2015, 4TB



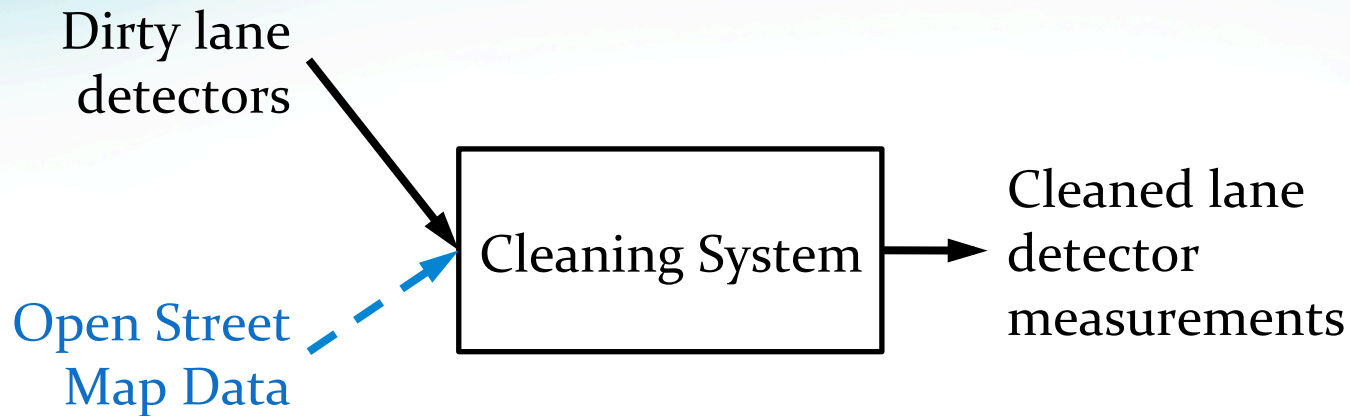
# Pre-Pilot Tasks Overview

- **Cleaning:** Clean traffic lane detector data flow values where a portion of its measurements were manipulated to be incorrect
- **Alignment:** Match traffic events with the traffic video segments containing those events
- **Prediction:** Predict the number and types of traffic incidents in a region over a time period
- **Forecasting:** Provide future traffic flow values

# Pre-Pilot Cleaning Task



# Pre-Pilot Cleaning Task



**Metric (cost)** = Mean Absolute Error in Flow

# Pre-Pilot Alignment Task

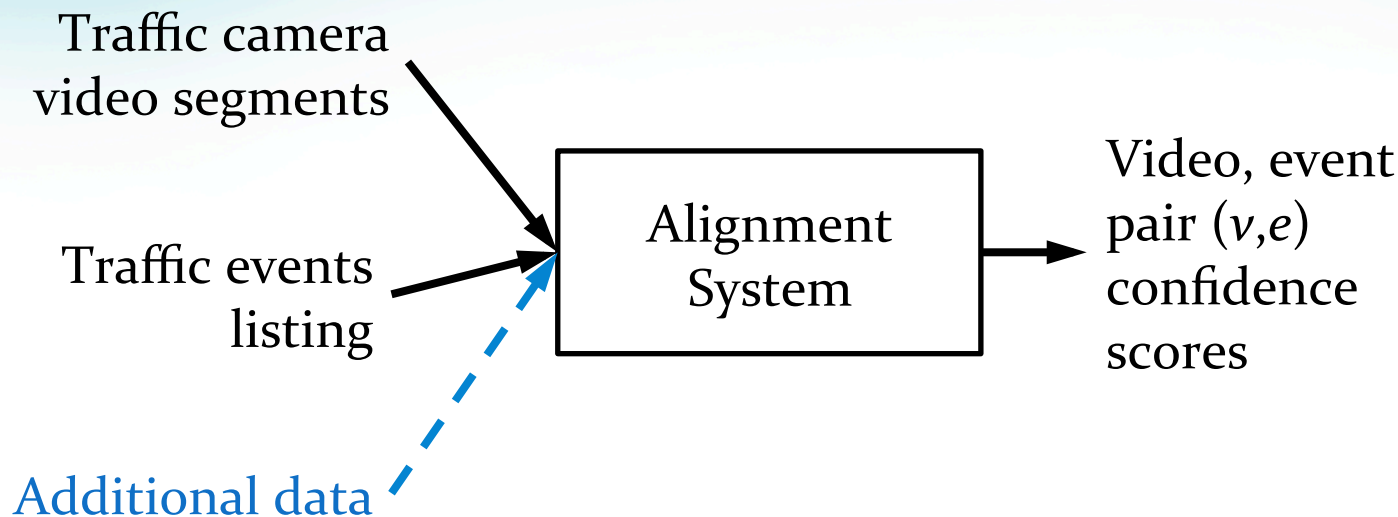


MDOT\_CHART\_a000a2de48a600680055fa2ec4235c0a,"Incident @ I-495 INNER LOOP ... **90%**

MDOT\_CHART\_400122fc49a700680055fa2ec4235c0a,"Incident @ I-495 OUTER LOOP ... **3%**



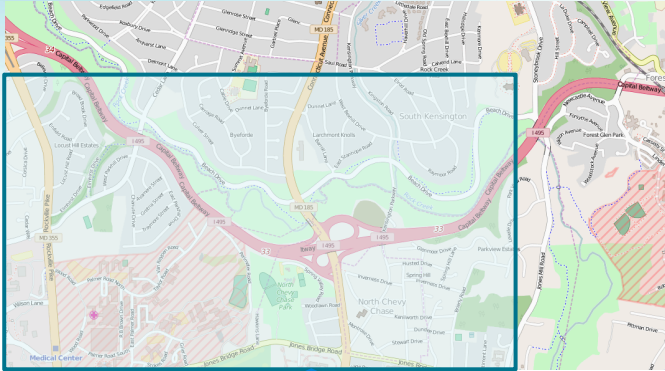
# Pre-Pilot Alignment Task



**Metric (cost)** =  $(Miss\ Rate) + 10^*(False\ Alarm\ Rate)$

Decision threshold selected to **minimize** metric

# Pre-Pilot, Prediction Task



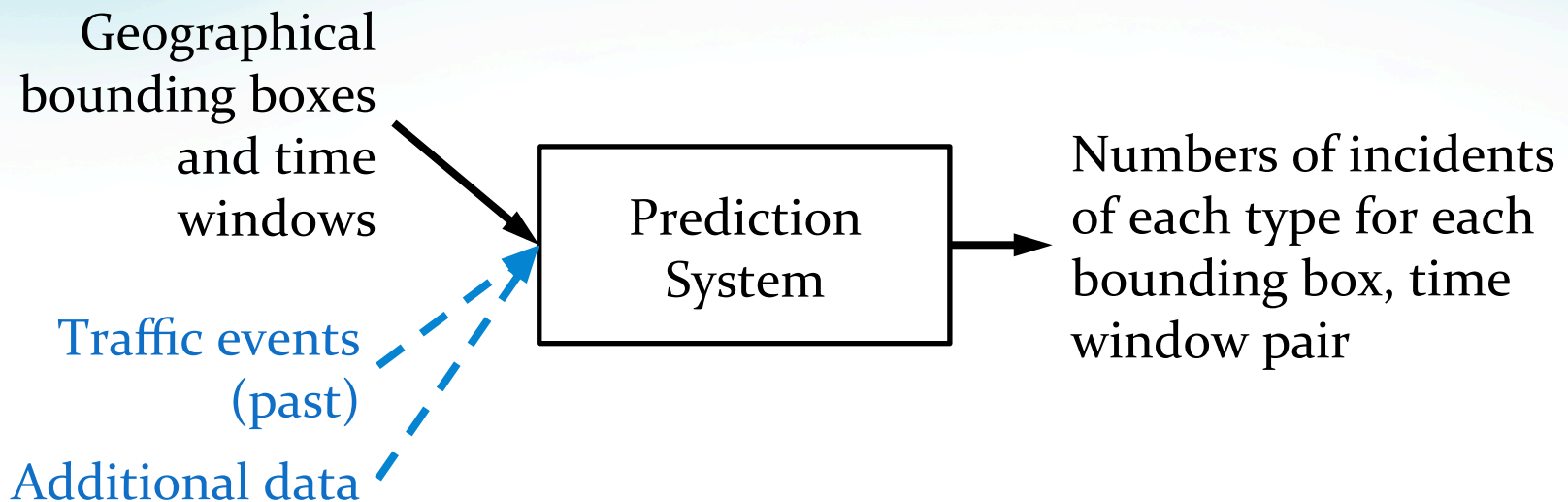
March 1, 2015

April 1, 2015



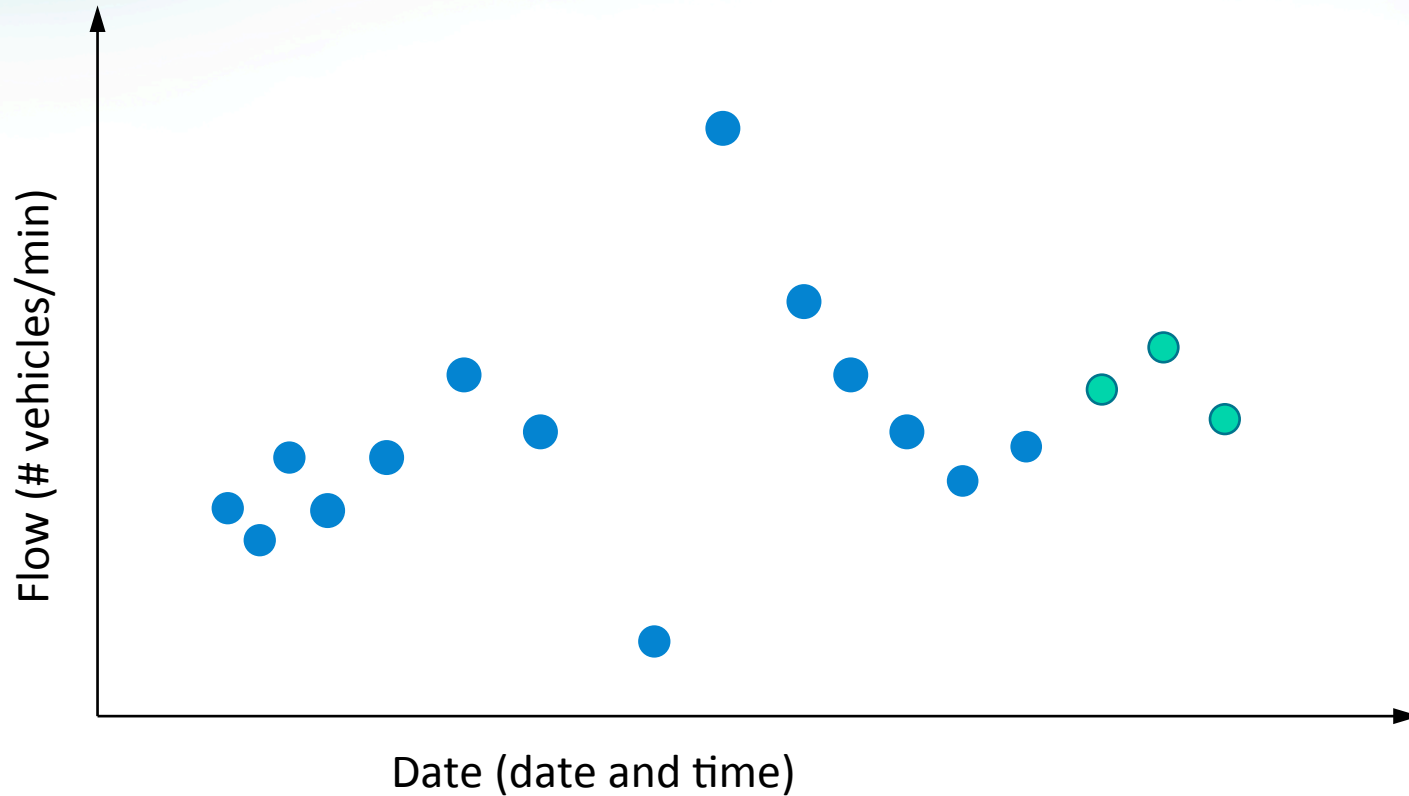
10 accidentsAndIndicents  
3 obstructions  
No other incidents

# Pre-Pilot Prediction Task



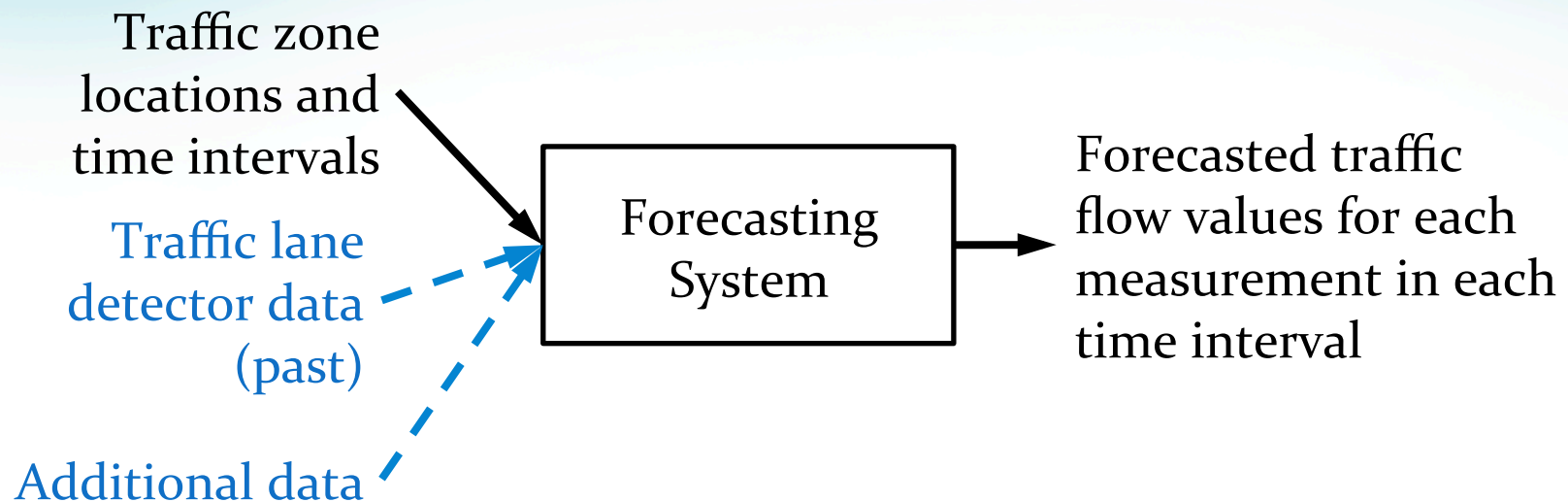
**Metric (cost)** = Mean Root Mean Square Error

# Pre-Pilot Forecasting Task





# Pre-Pilot Forecasting Task



**Note:** here, traffic lane detectors are aggregated traffic into **traffic zones**.

**Metric (cost)** = Mean of Mean Absolute Percentage Error

# Pre-Pilot Participants

Measure	Count
Number of Registrants	5
Number of Participant Sites	1
Cleaning Task Submissions	3
Prediction Task Submissions	7

Participants selectively invited (some accepted)

Submissions received only for cleaning and prediction tasks

3 NIST baseline systems for cleaning task

1 NIST baseline system for prediction task



# Pre-Pilot Results Now Available

[datascience@nist.gov](mailto:datascience@nist.gov)