ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS

kpler

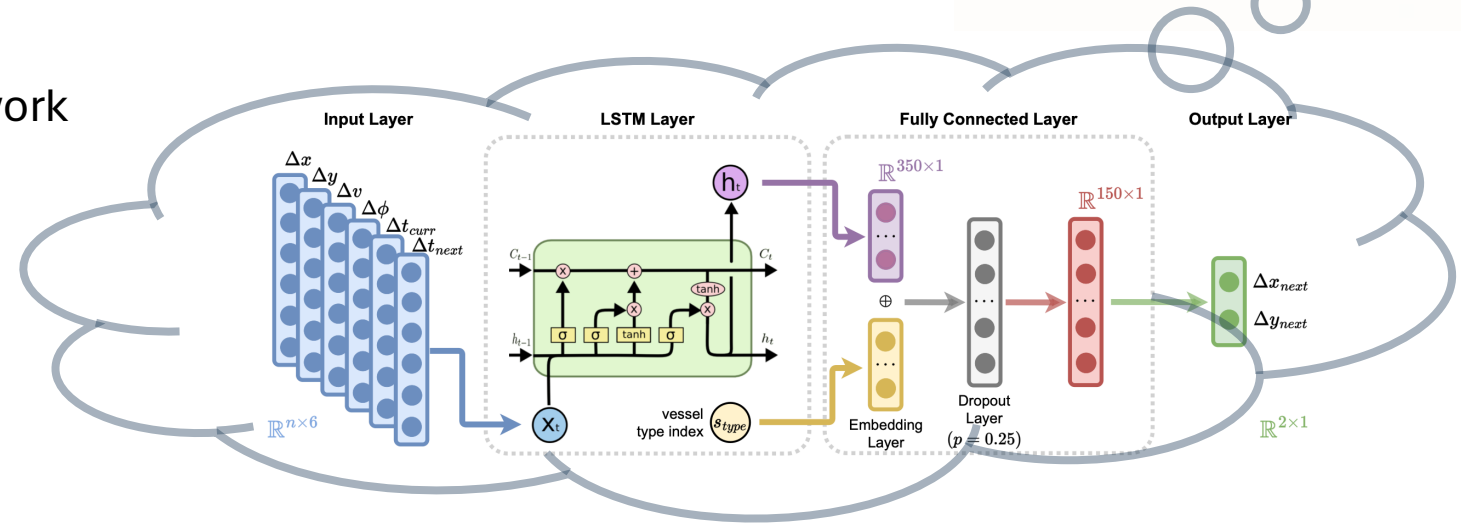
Nautilus in Motion: A Demo on Vessel Location Forecasting over Live Maritime AIS Streams

Andreas Tritsarolis
University of Piraeus

co-authors: Nikos Pelekis, Konstantina Bereta, Yannis Theodoridis

Introduction – Motivation - SotA

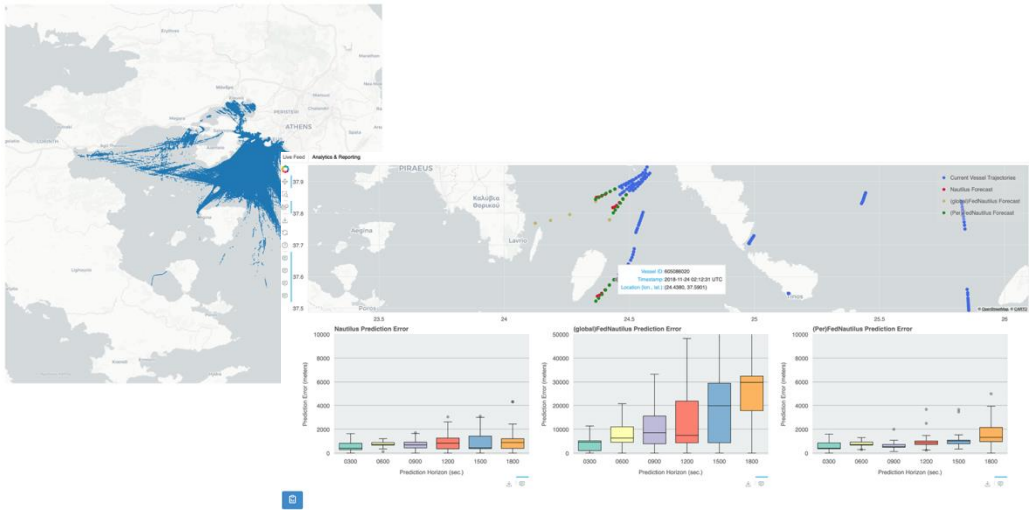
- Vast amounts of **AIS data** are generated daily by vessels worldwide
  - Vessels transmit AIS messages every few seconds
  - As such, **real-time tracking / analytics** is supported
- **Vessel Location Forecasting** (**VLF**) → critical task!
  - Contributes to maritime safety
  - Main challenges: maritime traffic conditions, privacy restrictions, etc.
- State-of-the-art on VLF → **[Fed]Nautilus**[1], a Neural-Network (LSTM)-based architecture
  - Exploits on the cross-silo **Federated Learning (FL)**

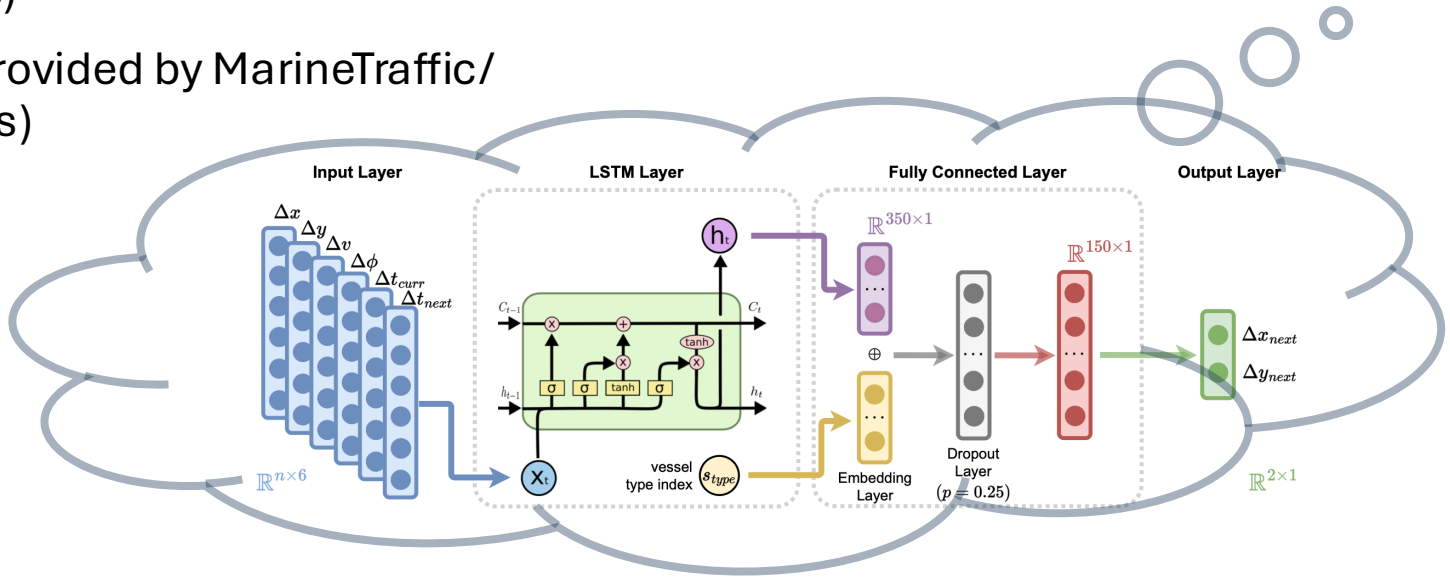


[1] A. Tritsarolis et al. (2024) On Vessel Location Forecasting and the Effect of Federated Learning. In Proceedings of the 25th IEEE MDM Conference.

Positioning of our work

- An end-to-end tracking and forecasting framework, powered by [Fed]Nautilus
- Demonstrated over two real-world AIS datasets:
  - **Piraeus** dataset (open[2], maintained by the Data Science Lab., Univ. Piraeus)
  - **Aegean** dataset (closed, provided by MarineTraffic/ Kpler for research purposes)

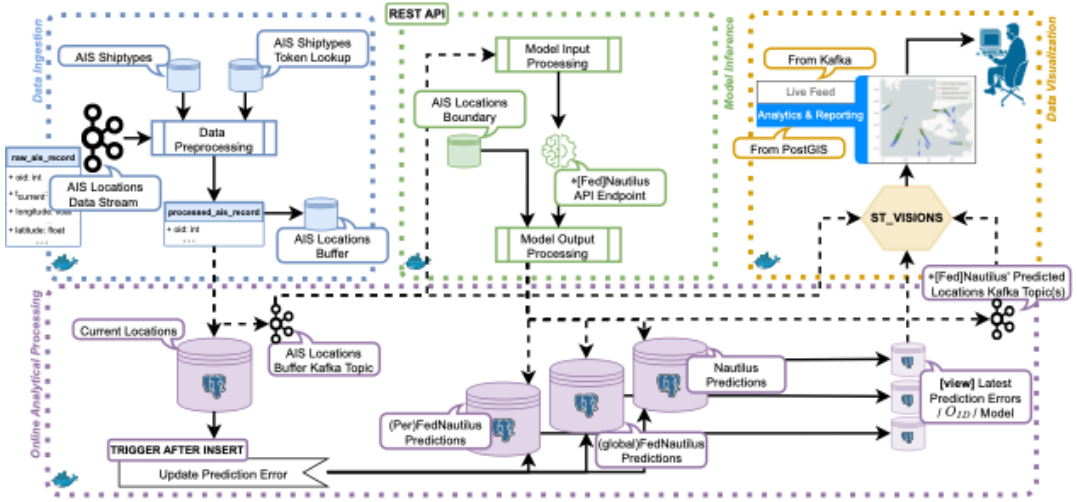


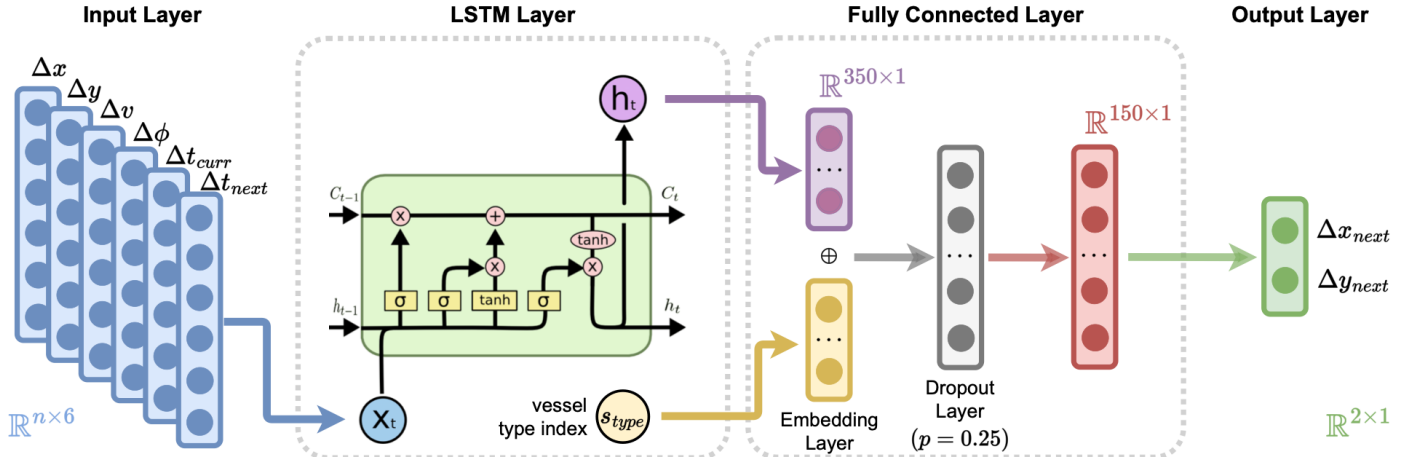


[2] A. Tritsarolis et al. (2022) The Piraeus AIS dataset for large-scale maritime data analytics, Data in Brief, 40; AIS data-stream live @ https://www.datastories.org/unipi-ais/

The Nautilus Framework

- Four modules:
  - **Data Ingestion**: preprocesses the raw AIS data-stream
  - **Model Inference**: deploys pre-trained [Fed]Nautilus models
  - **Online Analytical Processing**: Stores the processed AIS data-stream, infers [Fed]Nautilus predictions; performs analytics & reporting
  - **Data Visualization**: Visualizes live and predicted AIS data-streams and provides visual analytics

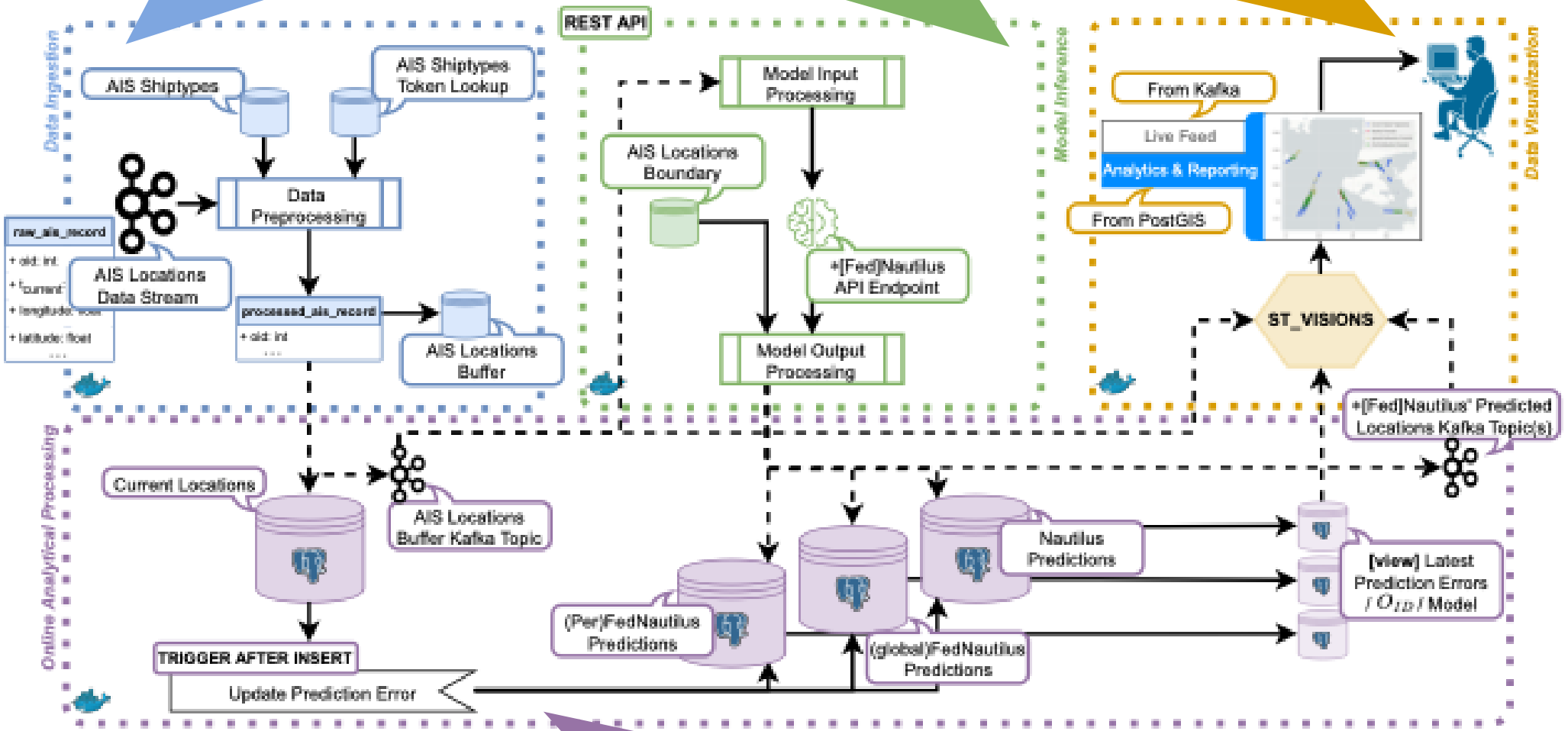




Preprocesses the raw AIS data-stream

Deploys pre-trained *[Fed]Nautilus models

Visualizes live and predicted AIS data-streams and provides visual analytics

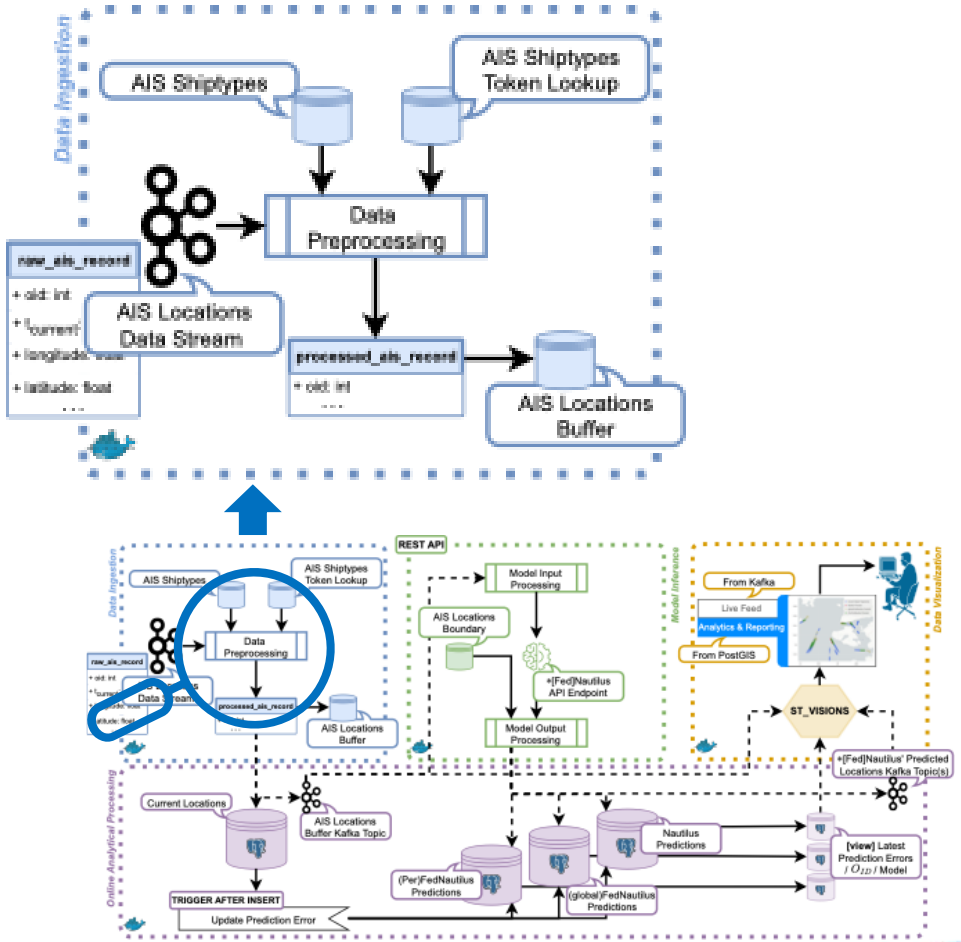


Stores the processed AIS data-stream, infers [Fed]Nautilus predictions; performs analytics & reporting

The [Fed]Nautilus architecture (1/4)

Data Ingestion

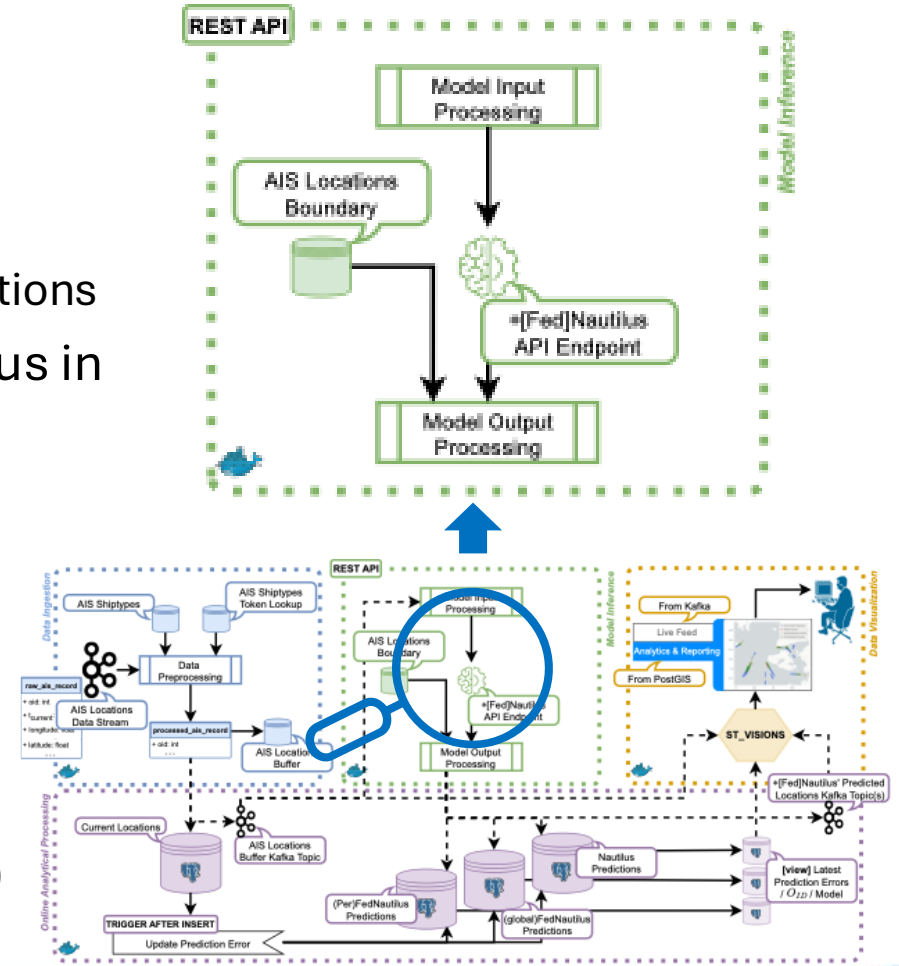
- Preprocessing & temporary storage
  - From raw AIS data-stream to valid AIS messages
- Data cleansing (noisy / duplicate / erroneous record elimination)
  - Noisy → first three MMSI digits (MID) …or location outside coverage area
  - Duplicate → $\Delta t_{min} = 10$ sec
  - Erroneous → speed $\notin [1, 50]$ knots
- Valid records are inserted to
  - Buffer (for model inference)
  - Kafka topic
  - PostGIS database (table: Current_Locations)



The [Fed]Nautilus architecture (2/4)

Model Inference

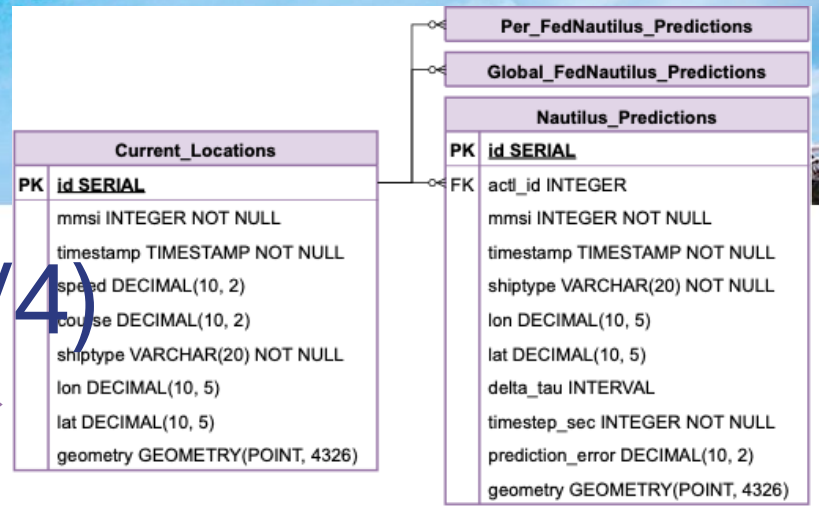
- #records$_{ID}$ satisfies transition constraints → eligible for location forecasting
  - $\text{length_min} = 18$ transitions; $\text{length_max} = 32$ transitions
- Query Nautilus, (global)FedNautilus, (Per)FedNautilus in parallel
  - Distinct REST API endpoints
- Inference process
  - Load parameters of requested model
  - Construct model input for requested horizon(s)
- Model output → Inserted to
  - Kafka topic
  - PostGIS database (table: ∗[Fed]Nautilus_Predictions)

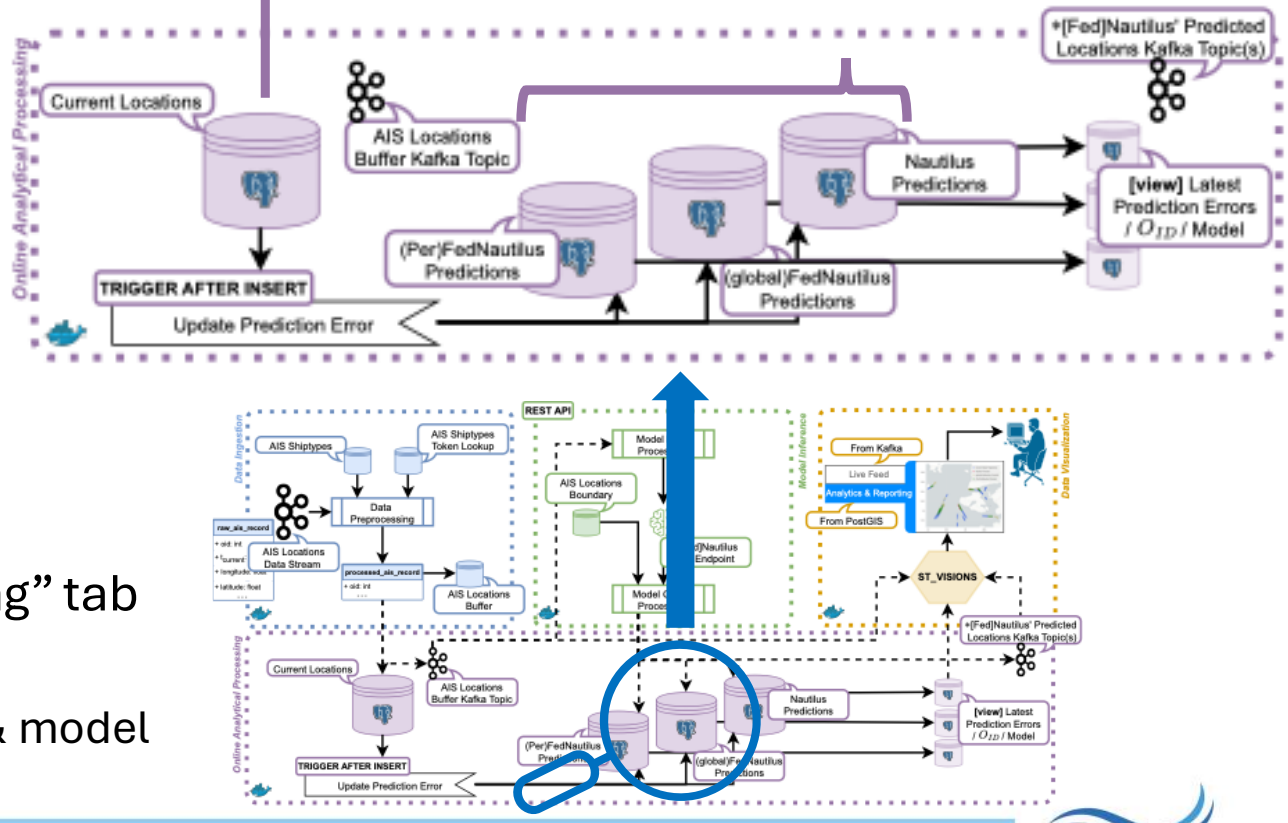


The [Fed]Nautilus architecture (3/4)

Online Analytical Processing

- Maintains four tables: current locations vs. ∗[Fed]Nautilus predictions
- Performs OLAP via an SQL Trigger
  - Matches actual location with a set of predictions within Δt_{match} time
  - Calculates the displacement error (field: prediction_error)…
    …temporal difference (field: delta_tau)
    …prediction horizon (field: timestep_sec)
- Provides results to the “Analytics & Reporting” tab
  - SQL Views
  - Fetch the latest prediction error per vessel & model

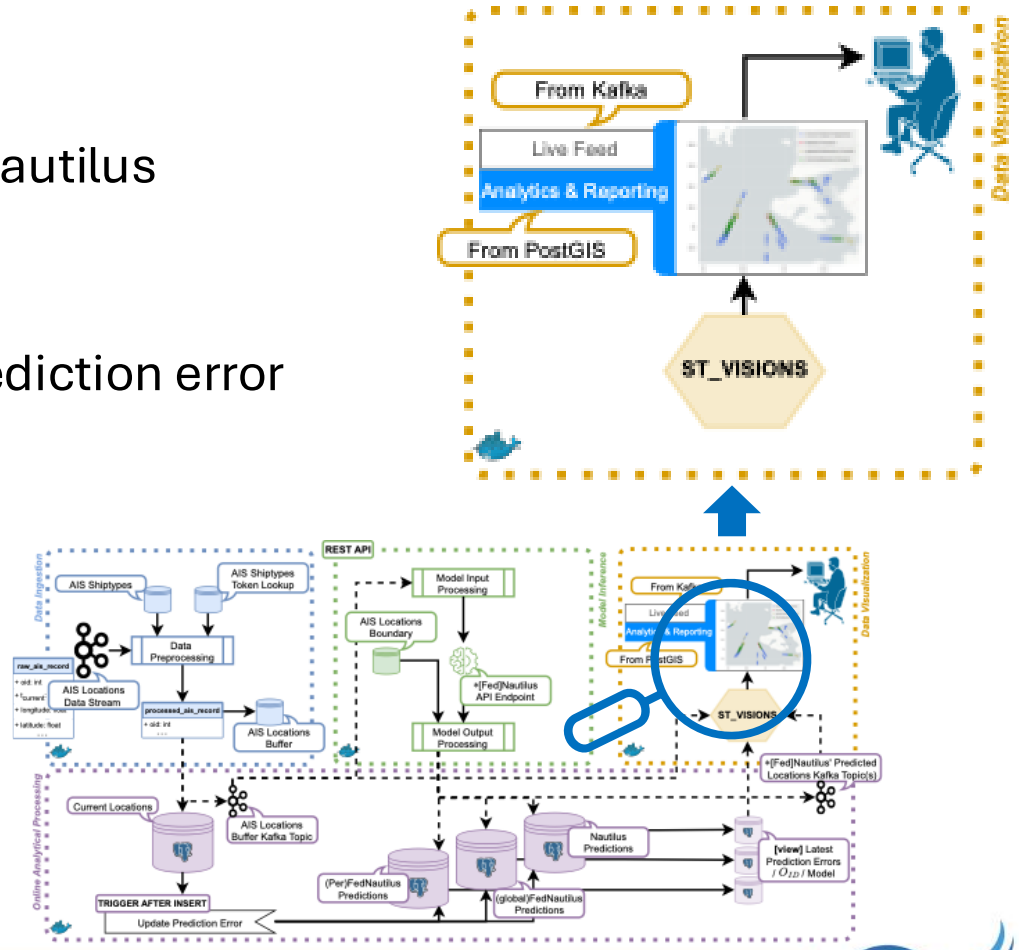




The [Fed]Nautilus architecture (4/4)

Data Visualization

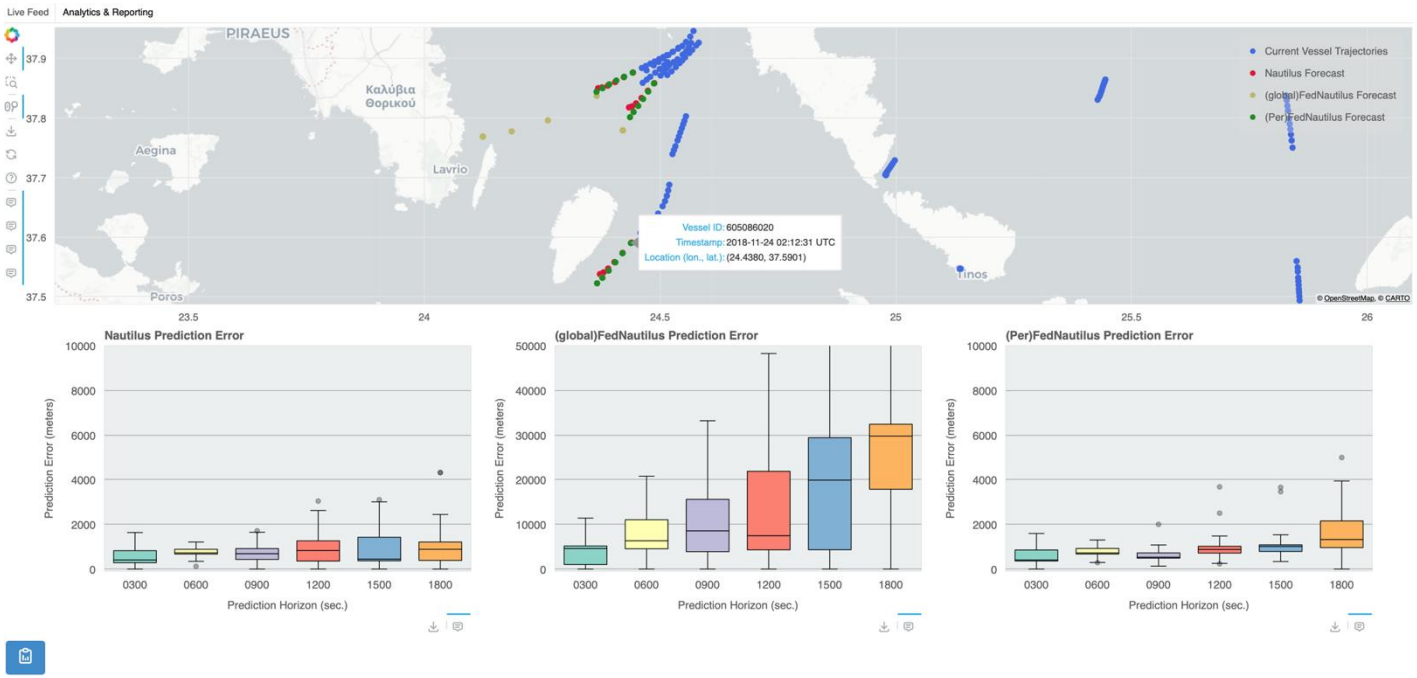
- "Live Feed" of vessels' latest position(s) and ∗[Fed]Nautilus prediction(s)
  - Prediction error boxplots
- "Analytics & Reporting", based on ∗[Fed]Nautilus prediction error tables
- VA-oriented features
  - Toggle layers' / boxplots' visibility
  - View vessels' attributes on-demand
  - Adjust detail level (pan/zoom)
- Based on ST-Visions[3]



[3] A. Tritsarolis et al. (2021) ST_VISIONS: A Python Library for Interactive Visualization of Spatio-temporal Data. In Proceedings of the 22nd IEEE MDM Conference.

Demonstration Scenario

- **Piraeus** dataset, $\Delta t_{match} = 35$ sec.
  - **Aegean** dataset scenario in the paper



Snapshot of the Piraeus dataset

Live Feed
Analytics & Reporting
Tab #1
“Live Feed”
Tab #2
“Analytics & Reporting”
Current Vessel Trajectories
Nautilus Forecast
(global)FedNautilus Forecast
(Per)FedNautilus Forecast
Toggle layer visibility
Map refresh rate: 1500 msec.
Table refresh rate: 2000 msec.
Toggle boxplot visibility
2x
Eleusis
Skaramangas
Megara
Νέα Πέραμος
ATHENS
Paiania
PIRAEUS
Aianteio
Agia Marina
38.05
38
37.95
37.9
37.85
23.3
23.4
23.5
23.6
23.7
23.8

Dive deeper into our research agenda in the maritime domain

DATA STORIES

Source-code available @GitHub:

Thank you for your attention!

I'll be glad to answer your questions!

MobiSpaces new data spaces for green mobility

This work was supported in part by the Horizon Framework Programme of the European Union under grant agreement No. 101070279 (MobiSpaces; mobispaces.eu). In this work, Kpler provided the Aegean AIS dataset and the requirements of the business case.