MobiSpaces
new data spaces for green mobility

eSAAM 2024

The MobiSpaces Manifesto on Mobility Data Spaces

Christos Doulkeridis, Ioannis Chrysakis, Sophia Karagiorgou, Pavlos Kranas, Georgios Makridis and Yannis Theodoridis

Ioannis Chrysakis
Research and Innovation Project Manager
Netcompany-Intrasoft SA
Ioannis.Chrysakis@netcompany.com
Netcompany

Context

- Data spaces consist of **trusted frameworks** that manage the entire data lifecycle, encompassing various data models, metadata descriptors, ontologies for semantic interpretation, and data services for accessing, processing, and analyzing data.

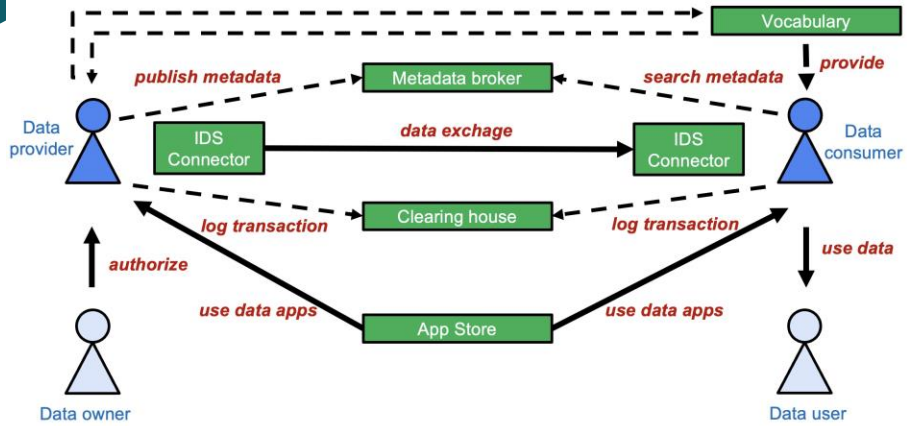
- Basic offerings of data spaces
  - **Data exchange**
  - Provision and consumption of **value-added services**

Fundamentals

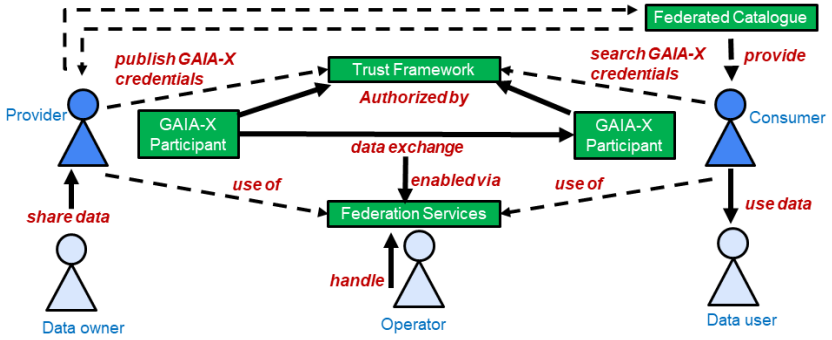
- State of the art initiatives (e.g., IDSA, GAIA-X)

- Basic concepts:
  - Participants in data spaces have typically a role
    - **Basic roles**: data providers or data consumers
    - **Intermediaries**: Identity authority, Vocabulary, Metadata Broker, IDS Connector, App Store, Clearing House

Fundamentals



Roles: Data Supplier, Data Customers, Intermediaries



Roles: Providers, Consumers, Operators (Federators)

MVDS

The MVDS is a tiny deployment of a data space. A typical implementation of a MVDS consists of: (i) two or more Connectors, (ii) the Certificate Authority granting X.509 certificates, and (iii) the Dynamic Attributes Provisioning Service (DAPS) to handle dynamic attributes and manage dynamic access tokens

The need

- As the mobility domain is one of the fundamental pillars of the modern digital economy worldwide, the **need for data spaces** tailored to the mobility domain is imperative.

- Mobility data **features** and **challenges**:
  - distributed data acquisition
  - imprecise data
  - recording of spatial data, error detection and cleaning of spatiotemporal records
  - semantic data representation is needed
  - analysis and learning of mobility patterns is useful for decision making purposes

The Contributions

- **MobiSpaces** contributes to the design of future mobility data spaces by:
  - offering data governance services for mobility data for data providers,
  - supporting data consumers via advanced data analysis services
  - providing these services in a data spaces compliant way
  - leveraging data connectors and the concept of data-app,
  - enabling trusted environment that respects the architectural design of data spaces.

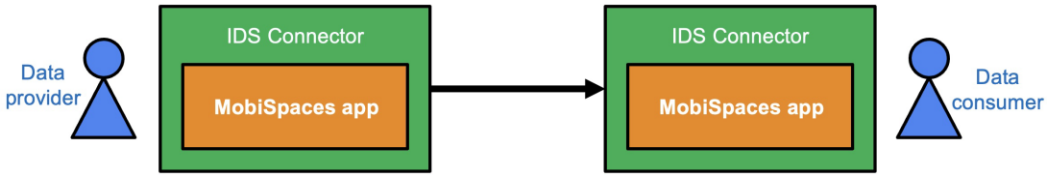
- To demonstrate the versatility of MobiSpaces two mobility domains are considered: the **urban domain** and the **maritime domain**.

The MOBISPACES Approach

- **For data providers**, the app offers an array of functionalities, encompassing data upload, management, and quality assurance, ensuring adherence to the FAIR principles.

- **For data consumers**, it is equipped with advanced analytical features, enabling the derivation of meaningful insights through sophisticated data processing methodologies, inclusive of AI-driven analytics and comprehensive visualization tools.

- The app's integration into the MobiSpaces ecosystem enhances **mobility data accessibility** and fosters **secure**, **seamless data sharing** with **value-added services**.



The **Data-App Connectivity** with MobiSpaces Ecosystem

MobiSpaces
new data spaces for green mobility

eSAAM 2024

Empowering the role of Data Provider

- The **curation of mobility data** by data owners entails data acquisition, data manipulation, error detection, cleaning and eventually storing this curated data in a desired format.

- MOBISPACES technology assist data owners to bridge the gap between its internal data management architecture with the architecture advocated in data spaces via providing:
  - **Data Governance Framework**
  - **AI-based Data Operations Toolbox**
  - **… and corresponding Value-added services**

MobiSpaces
new data spaces for green mobility

eSAAM 2024

Empowering the role of Data Provider Data Governance Framework

- The MOBISPACES Data Governance Framework is an assembly of **data** (and **infrastructure**) **services** that operate both in standalone mode as well as by interacting with each other to enable data governance in mobility.

- Data Governance Framework supports **data owners** of mobility data with functionality related to **data curation** and **data preparation**, thus facilitating the participation of a data owner in a data space by taking the role of a data provider.

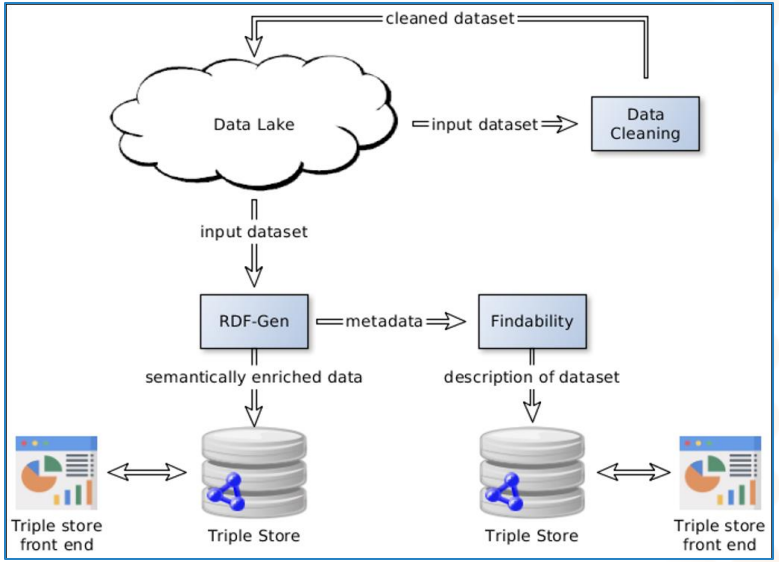
Empowering the role of Data Provider
AI-based Data Operations Toolbox

- The toolbox consists of a batch layer and a real-time layer.
  - **Batch Layer** uses a declarative "**SQL on everything" approach** for querying all data sources, similar to Presto. Users can perform complex data processing via workflows or SQL, with support for various storage types—raw files, relational, non-relational, mobility, and encrypted data
  - **Real-Time Layer** handles **online data aggregation** (e.g., compression) and offers a library of efficient mobility operators, **deployable at the Edge**, supporting both data ingestion and in-situ querying.

- The toolbox contains a set of infrastructure tools:
  - **Intelligent Resource allocator**: optimizes resource allocation based on input constraints, modeling it as a constrained optimization problem
  - Resources are assigned for orchestration via the **Execution Orchestrator**, targeting Kubernetes clusters.
  - **Workflow Builder**: offers a GUI for users to create complex, AI-based batch and ETL workflows using a customizable, easy-to-use canvas-and-palette interface.

Empowering the role of Data Provider
Value-added services

- Error Detection and Cleaning
- Semantic Representation and Reasoning
- Data Interlinking
- Data Provenance
- **Fairification**



- **Mobility-aware Edge Processing** by providing tools for edge data management, optimizing the deployment of data acquisition services.
  - This is often overlooked in data spaces, which typically emphasize secure data sharing and standardized data exchange.

MobiSpaces
new data spaces for green mobility

Empowering the role of Data Consumer

- At the data consumer side, MobiSpaces offers a **set of application services** for the analysis of mobility data.

- This is mainly provided by means of the **Edge Analytics Suite (EAS).**

Empowering the role of Data Consumer
Edge Analytics Suite (EAS)

- EAS offers diverse functionalities:
  - Distributed edge analytics
  - Federated Learning (FL) for mobility data
  - Explainable Artificial Intelligence (XAI) methods for interpreting mobility models
  - Visual analytics with interactive map-based visualization

Empowering the role of Data Consumer
Edge Analytics Suite (EAS) value-added services

- MobiSpaces offers a diverse array of mobility specific **data analytics methods and techniques** offered as a service through EAS.

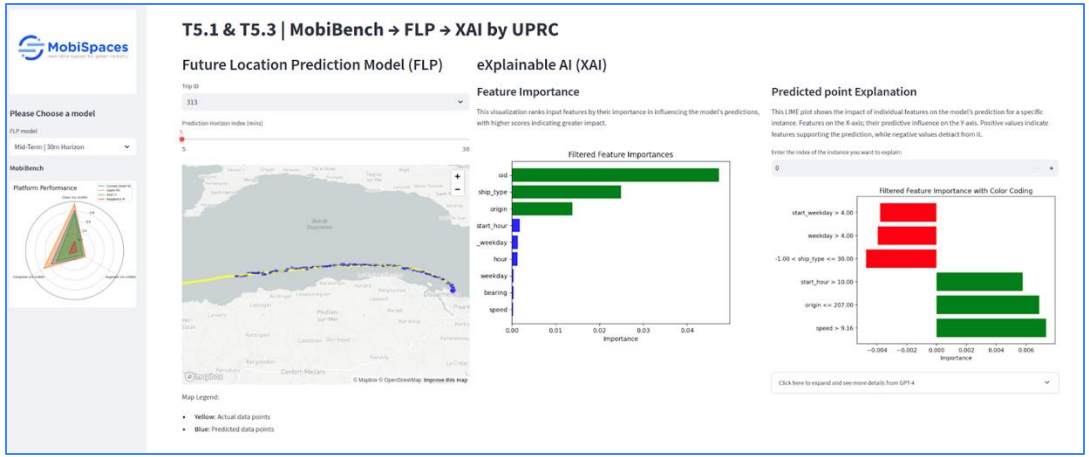
- Providing a **suite of ML algorithms** designed for mobility data to cover essential tasks such as trajectory cleaning and segmentation, trajectory forecasting, and anomaly detection.

- These application services are seamlessly **integrated within the data space**, allowing data consumers to access and utilize these tools efficiently.

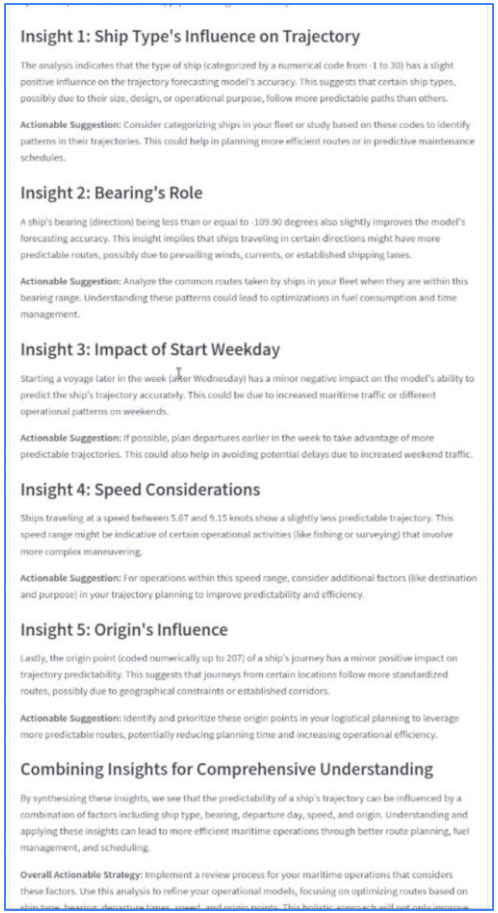
Empowering the role of Data Consumer Edge Analytics Suite (EAS) innovation

- **Federated AI** to address the complexity of trajectory data

- **Explainable AI** tailored for mobility domain
  - Vessel Route Forecasting (VRF) - A GPT-Based Model
  - Future Location Prediction Model (FLP)



The EAS In Action



Insight 1: Ship Type's Influence on Trajectory

The analysis indicates that the type of ship (categorized by a numerical code from -1 to 30) has a slight positive influence on the trajectory forecasting model's accuracy. This suggests that certain ship types, possibly due to their size, design, or operational purpose, follow more predictable paths than others.

Actionable Suggestion: Consider categorizing ships in your fleet or study based on these codes to identify patterns in their trajectories. This could help in planning more efficient routes or in predictive maintenance schedules.

Insight 2: Bearing's Role

A ship's bearing (direction) being less than or equal to -109.90 degrees also slightly improves the model's forecasting accuracy. This insight implies that ships traveling in certain directions might have more predictable routes, possibly due to prevailing winds, currents, or established shipping lanes.

Actionable Suggestion: Analyze the common routes taken by ships in your fleet when they are within this bearing range. Understanding these patterns could lead to optimizations in fuel consumption and time management.

Insight 3: Impact of Start Weekday

Starting a voyage later in the week (after Wednesday) has a minor negative impact on the model's ability to predict the ship's trajectory accurately. This could be due to increased maritime traffic or different operational patterns on weekends.

Actionable Suggestion: If possible, plan departures earlier in the week to take advantage of more predictable trajectories. This could also help in avoiding potential delays due to increased weekend traffic.

Insight 4: Speed Considerations

Ships traveling at a speed between 5.67 and 9.15 knots show a slightly less predictable trajectory. This speed range might be indicative of certain operational activities (like fishing or surveying) that involve more complex maneuvering.

Actionable Suggestion: For operations within this speed range, consider additional factors (like destination and purpose) in your trajectory planning to improve predictability and efficiency.

Insight 5: Origin's Influence

Lastly, the origin point (coded numerically up to 207) of a ship's journey has a minor positive impact on trajectory predictability. This suggests that journeys from certain locations follow more standardized routes, possibly due to geographical constraints or established corridors.

Actionable Suggestion: Identify and prioritize these origin points in your logistical planning to leverage more predictable routes, potentially reducing planning time and increasing operational efficiency.

Combining Insights for Comprehensive Understanding

By synthesizing these insights, we see that the predictability of a ship's trajectory can be influenced by a combination of factors including ship type, bearing, departure day, speed, and origin. Understanding and applying these insights can lead to more efficient maritime operations through better route planning, fuel management, and scheduling.

Overall Actionable Strategy: Implement a review process for your maritime operations that considers these factors. Use this analysis to refine your operational models, focusing on optimizing routes based on

Example of GPT output for human-centric XAI using OPenAI

MOBISPACES and Trust

- Trust is **the basis for data spaces**, where each participant is evaluated and certified before being granted access to the trusted business ecosystem.

- Necessary for mobility data spaces to support federated and non-monopolised internet services with many concurrent cloud providers and edge nodes both acting as consumers and producers of data.

- In MOBISPACES we elaborate on the trust and security concepts using different **Attribute-based Access Control (ABAC) mechanisms**.

MOBISPACES and Trust
Attribute-based Access Control (ABAC) mechanisms

- ABAC supports **context-aware** (e.g., based on location, time, etc.) and **multiple levels of trust and security per user role**, service and Application Programming Interface (API).

- **Identity management** enriched with attributes for location aware edge services, trust management, data access control, and data usage control.

- The mechanisms are supported by a **Single Sign-On (SSO) solution** deployed through the SSO as a Service (SSOaaS) paradigm.

- The mechanisms assign subject attributes to subjects and object attributes to objects, developing **policies that define access rules** for each object within the environment.

Conclusion and Future Work

- We present how the MOBISPACES research project leverages innovative technology designed for **future mobility data spaces**.
- It offers a wide variety of **data services** that empower the roles of **data provider** and **data consumer** in data spaces.
- By adopting the MobiSpaces perspective, data owners can embark in mobility data spaces more easily and can exploit multiple **ready-to-go tools** and techniques for mobility data to harness the merits of **data spaces**.
- In the future, we aim to explore how to **leverage established data spaces**, both in mobility and other domains, to maximize value across **larger data ecosystems**.

MobiSpaces
new data spaces for green mobility

eSAAM 2024

Project Info

- Maximum grant amount is EUR 8.808.063,00
- Duration: 36 months, from September 1st, 2022 to August 31st,2025
- 25 Partners:

Large Industries

GFT, Netcompany, kpler, FUJITSU, BOSCH, AMT, Atos, FREQUENTIS, SIEMENS, eng

SME

Trust-IT Services, LEANCALE, UBITECH, OKYS, netU, emisia, DigiSys 4eU, UNPARALLEL, White Label Consultancy

Research Centers

UNIVERSITY OF PIRAEUS RESEARCH CENTER, AALBORG UNIVERSITY, AIT, ULB UNIVERSITÉ LIBRE DE BRUXELLES, Geodatastyrelsen

Thank you

For more information:
Ioannis.Chrysakis@netcompany.com
https://www.chrysakis.eu
Research and Innovation Development Department
Netcompany-Intrasoft SA.

Netcompany