# esaam





Funded by the European Union

# Introducing an Enhanced Metadata Broker for Manufacturing Data Spaces

Alexandros Nizamis<sup>1,2</sup>, Dimosthenis Ioannidis<sup>2</sup>, Panagiotis Gkonis<sup>1</sup>, and Panagiotis Trakadas<sup>3</sup>

<sup>1</sup>Department of Digital Industry Technologies, National and Kapodistrian University of Athens, Greece <sup>2</sup>Information Technologies Institute, Centre for Research and Technology Hellas, Greece <sup>3</sup>Department of Port Management and Shipping, National and Kapodistrian University of Athens, Greece

# **Presentation Outline**



### **Presentation Outline**

- Problem Statement and Concept Overview
- Ontology and Semantic Framework to Model and Manage Metadata for Industry 4.0/5.0
- Metadata Broker for Manufacturing Data Spaces
  - Architecture
  - User Interfaces
- Discussion / Conclusions
- Q&A

## **Problem Statement and Concept Overview**

CONTRACTOR CONTRACTOR



### **Problem Statement and Concept Overview**

#### **Problem Definition:**

- Today collaborative ecosystems and value networks have been established based on data sharing mechanisms and principles
  coming from concepts like Data Spaces. However, this data-centric approach has also increased the need for effective metadata
  management that enables entities participating in data sharing scenarios to find and trust available data.
  - FAIR Principles
  - EU AI Act
- Initiatives like IDSA, GAIA-X etc. have introduced concepts such as metadata brokers and catalogues.
  - These components are mainly considered as optional ones in a Data Space
  - There are no concrete guidelines, applications and examples of metadata brokers that can adopt FAIR principles and further support and accelerate the research in specific domains by connected data with possible AI solutions (that today is most commonly use of data).

#### Our Approach:

A metadata broker for manufacturing related Data Spaces is introduced. It is based on Data Spaces principles and artefacts that it extends by enabling semantic-based modeling and search capabilities. It targets to accelerate the use of data by data scientists and AI



# Ontology and Semantic Framework to Model and Manage Metadata for Industry 4.0/5.0

### Data Sp4ce Metadata Registry (DS4MR) Ontology

Available approaches cannot effectively describe some core aspects related to manufacturing domain and related Data Spaces

- Current modeling approaches related to manufacturing domain are mainly focused on the detailed representation
  of manufacturing processes and on the description of available data.
- On the other hand, the modeling approaches related to AI include data description aspects, however they start from AI perspective and metadata injection seems that have to be done from AI experts and not just by data providers.
   Ontologies for data mining field have been highlighted as the most promising for metadata description.

Data Sp4ce Metadata Registry (DS4MR) Ontology version 1 has been implemented using protégé tool and it is available in RDF/OWL format.

This first version includes 59 classes, 14 object properties and 13 data properties. The aim of this ontology is to represent:

- a metadata broker with its associated self-description including Data Space related concepts such as tokens etc.
- a dataset with associated metadata such as the data format, size, characterization, usage, related quality processes etc.



### Data Sp4ce Metadata Registry (DS4MR) Ontology

Class Name	Description		
Metadata Broker	This class describes the Metadata Broker entity as a core software component		
Self Description	It represents the data space related info and metadata that are connected to a registration in the Broker		
Dataset	This class is the central concept related to the resource that all the metadata info is attached to		
Data Characterization	It represents the characterization of a dataset as labeled, structured etc.		
Data Distribution	This class describes the distribution of a dataset		
Data Format	Data Format class includes a series of sub-classes to represent the format of dataset (e.g. csv)		
Data Owner	This class holds information related to the owner of the dataset		
Data Size	This class characterize a dataset regarding to its size (e.g. Big Data)		
Dataset Type	It is a class that represent the type/features of a dataset (e.g. timeseries)		
Initial Data Source	It represents the initial/original data source type that the dataset was stored (e.g. a database)		
Quality	It holds information related to a quality procedures were applied in a dataset(e.g. pre-processing)		
Usage	This class represents the different types of solutions/applications that a dataset used for		



### **Semantic Framework**

A semantic framework to enable the application of CRUD operations and reasoning to the ontological resources has been developed. This framework is used as the core back-end part of the introduced Metadata Broker. It provides a set of web services that represent specific SPARQL queries and semantic rules to enable CRUD and reasoning operations over DS4MR

#### **Implementation Details:**

- For handling the ontology (RDF/OWL file) the Owlready2 Python package was used.
- REST calls are translated to SPARQL queries and relevant answers are provided by the ontological resources stored in a triple store. The REST API provided by the Semantic Framework is available by using Flask framework that enables to setup a Python web server.
- GraphDB was the selected triple store for storing the ontological resources.
- The framework, that is a Python application using Owlready2 and Flask was containerized using Docker. The triplestore, so GraphDB was deployed as a separate container.



# Metadata Broker for Manufacturing Data Spaces

### Metadata Broker - Architecture

The proposed Metadata Broker architecture was based on IDSA concept and principles. It is also adopting the IDSA RAM v04 guideline that a broker/registry of metadata has to keep common characteristics with Data Space Connectors.

 In particular, it has to be a Connector that contains endpoint(s) for the registration, publication and querying of metadata connected to a data resource



### Metadata Broker – User Interfaces

- Regarding the UIs of the proposed Metadata Broker, guidelines of IDSA RAM 4.0 were followed.
  - website/web interface with full text and facet search capabilities is proposed.
- Design approaches from AI Model Marketplaces coming from research were considered as well

#### **Data Space Metadata Broker**

A metadata registry for manufacturing Data Spaces



Dataset Start Date			Dataset End Date		
Please : 🛩	Please : 👻	Please : ¥	Please : ¥	Please : 🗸	Please : 🗸
fonth	Day	Year	Month	Day	Year

Data Format	Data Type	
Please Select 🗸	Please Select	~
please select the format of data that are available	please select the type of data	

Data U V Plea

AR/V

Auto Defe Dem

IOT A

Pred Prod

Simu Supp Othe

sage	Data Source		
se Select	Please Select ~		
R Applications nomous Robots Applications	please select the original data source type		
ct Detection	Data Quality Process		
and Forecasting	Please Select ~		
ictive Maintenance	please select if any data quality process were applied		
uction Scheduling & Optimization lations & Digital Twins	Data Size Characterization		
ly Chain Management	Please Select V		
r	please characterize the size of data		

# **Discussion / Conclusions**

### **Discussion / Conclusions**

#### Contributions:

- The main needs for keeping detailed metadata for specific domains such as Industry 4.0/5.0 has been discussed
- Data Space Metadata Broker for these domains was introduced by extending the available data broker reference concepts and implementations.
  - The core addition of this work was done by introducing a new ontology that was designed for the modeling of metadata related to Industry 4.0/5.0.
  - An initial version of Semantic Framework to enable the as-a-service handling and management of this ontology by Data Space Connectors back-ends or corresponding UIs was introduced.
  - $\circ$  A dedicated UI for this Metadata Broker was also proposed and presented as design ideas.

#### Next Steps:

- Further development of DS4MR Ontology and the enrichment of the Semantic Framework with more services that will include new SPARQL and reasoning scripts to support more complex registration and search functionalities
- Uls' implementation is a future plan as well, alongside with the testing of the fully implemented Metadata Broker in an operational Manufacturing Data Space.

# **Questions?**





# Thank you

Alexandros Nizamis alnizami@dind.uoa.g

> This work is co-funded from CyclOps project under Grant Agreement 101135513 of Horizon Europe Program