Fourth ITU Workshop on FG DPM (19 July 2019)

Technical Specification D3.3

# Framework to Support Data Interoperability in IoT Environments

07. 2019.

Ilyoung Chong
HUFS

Hankuk University of Foreign Studies

# Presentation Outline

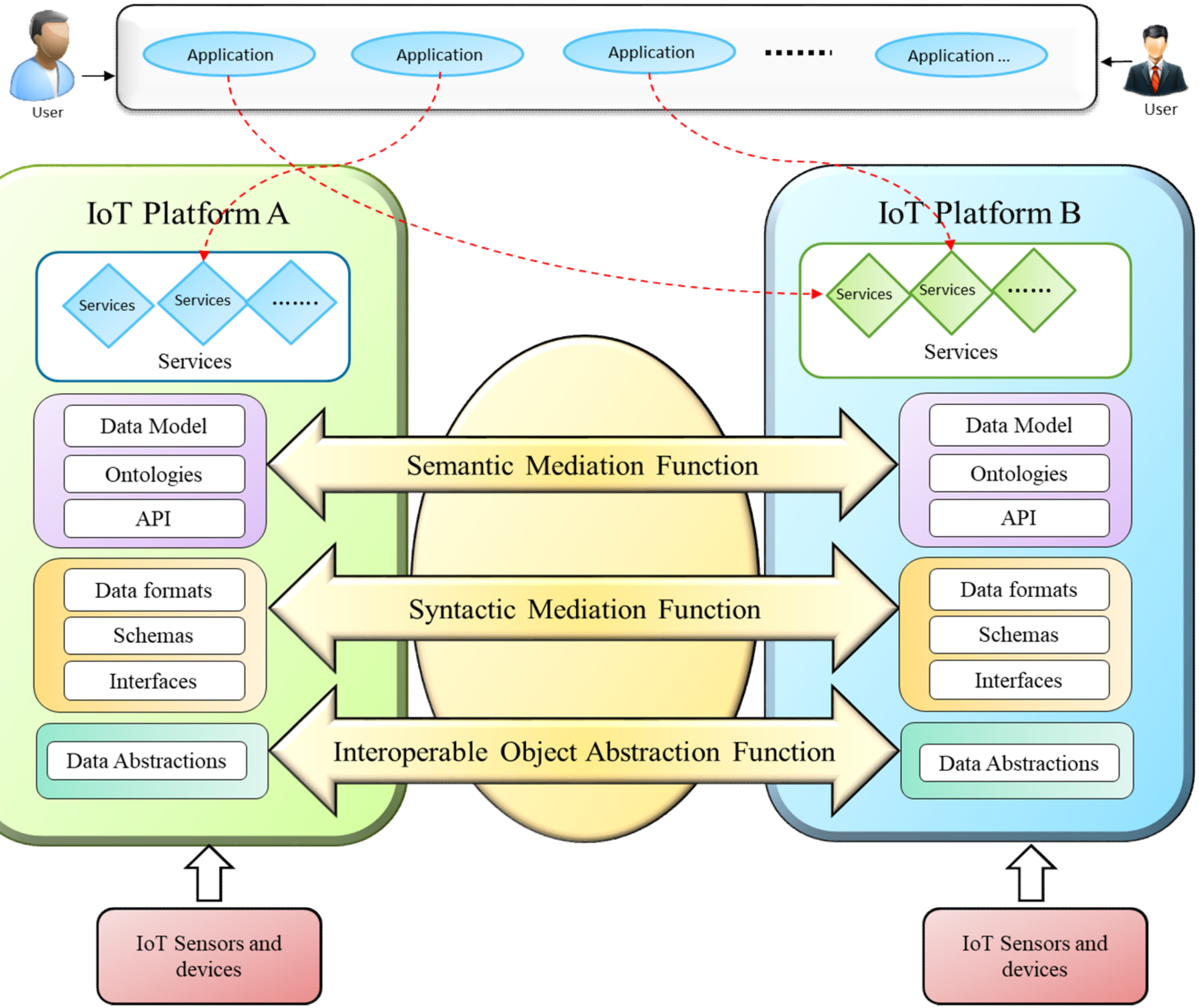
- Introduction
- Framework of Data Interoperability
- Web of Objects model to support semantic data interoperability provisioning
- Semantic interoperability provisioning using learning based approach.

# Introduction to Data Interoperability

# Issues on Data Interoperability

- Amount of data is
  - growing at an unprecedented rate in diverse application domains
  - generating billions of data streams in heterogeneous formats and semantics;
- The data has been processed and managed with
  - different data modals and
  - different approaches which result in heterogeneous information sources and
  - have become a major challenge in developing integrated applications;
- Data collected is
  - multi-modal, diverse, voluminous and
  - often supplied at high speed;
  - heavy challenges on data interoperability provisioning systems.

# View on Data Interoperability in heterogeneous environments




anLab

# Dimensions for Data Interoperability (1)

- **Semantic Interoperability**
  - Concerned with the meaning of data.
  - Consensus on meaning is required while exchanging the data across systems.
  - defines the true meaning of the contents that are generated by IoT devices and mutually agreed by a different system
  - enable different stakeholders to access and understand data unambiguously.

- **Syntactical Interoperability**
  - Heterogeneous IoT devices generate data that are stored and used in different formats.
  - is concerned with the data formats, syntax and coding methods.
  - Protocols used by IoT devices use standard syntax representation in diverse formats such as XML, JSON or HTML.

- **Object Abstraction Interoperability**
  - to support diverse object abstractions in terms of:
    - ✓Semantic and syntactic data representation, and
    - ✓Metadata description and coding
    - ✓Vocabularies

# Dimensions for data interoperability (3)

| | |
|---|---|
| **Semantic Interoperability** | • Data Model<br>• Semantic Schema<br>• Ontologies<br>• Semantic Annotations |
| **Syntactical Interoperability** | • Data Formats<br>• Syntax Coding<br>• Interfaces |
| **Object Abstractions Interoperability** | • Vocabulary<br>• Composite Data Object Abstractions<br>• Metadata descriptions |

# Requirements to Support Data Interoperability

# General requirements to support data interoperability

- **General Requirements:**
  - Use of standard vocabularies refer to provide a common understanding of data;
  - Validation mechanisms to evaluate data translation and conversion process;
  - Support of integration and sharing of data for services among processes of same and different organizations;
  - Classification and aggregation of data using diverse taxonomies;
  - Support of security and privacy → Additional data interoperability aspects need to be considered including interoperability of data.

# Requirements to support semantic data interoperability (1)

- **Semantic data modelling:**
  - the semantic representation of data to express a common understanding across systems.
  - semantic representation model to provide the conceptual understanding of data as well the relationship among entities.
- **Semantic integration and sharing:**
  - mechanism for the linking of data based on semantic ontology models
  - the linking mechanism to support dynamic integration and sharing of data
- **Semantic annotation of data:**
  - a semantic annotation mechanism to support the annotation of data coming from heterogeneous sources.
  - set of metadata to express the features diverse IoT data.
- **Semantic data management:**
  - a suit of well-defined services to manage the data allowing its access, retrieval and storage operations;

# Requirements to support semantic data interoperability (2)

- **Semantic Ontology alignment and mapping:**
  - provision of improved ontology alignment to support semantic interoperability.
  - ontology alignment techniques with enhanced accuracies can enable and improve interoperability across different systems.
- **Semantic representation of knowledge:**
  - In IoT providing rules with knowledge representation to support reasoning on the data which enhances its value.
  - Provision of information model semantically rich and expressive enough to represent different forms of the objects being maintained.
  - Information model enough to represent semantic information. For example, Ontologies in IoT to exchange the knowledge
- **Semantic data transformation:**
  - Mechanism to provide transformation service among heterogeneous semantic data models.

# Requirements to support data syntactic interoperability

- The syntactical format identification, registration and management mechanisms.
- The syntactical format description models to provide expressivity in definitions.
- Well defined syntactical templates to generate response object on the initial template instances.
- Syntactical translation mechanism to generate the transformation based on the provided templates.
- Syntactical formats registry to provide a repository of formats of diverse registered platforms.
- Well defined syntactical meta data schema and their mapping mechanisms
- Verification methods for format translation and conversion process to validate the effectiveness of translation mechanism.

# Requirements to support object abstraction interoperability

- Mechanism for creation and management of abstract data representations.
- Provision of semantics in the data representation model to maintain the same meaning across different data models.
- Uniform syntactic representation of data in standard formats.
- Description of metadata and their coding function to express diverse core data models.
- Provision of data and metadata profiles which can express the object abstract representations for different systems.
- Mechanism to generate object abstract representations profiles from different data.

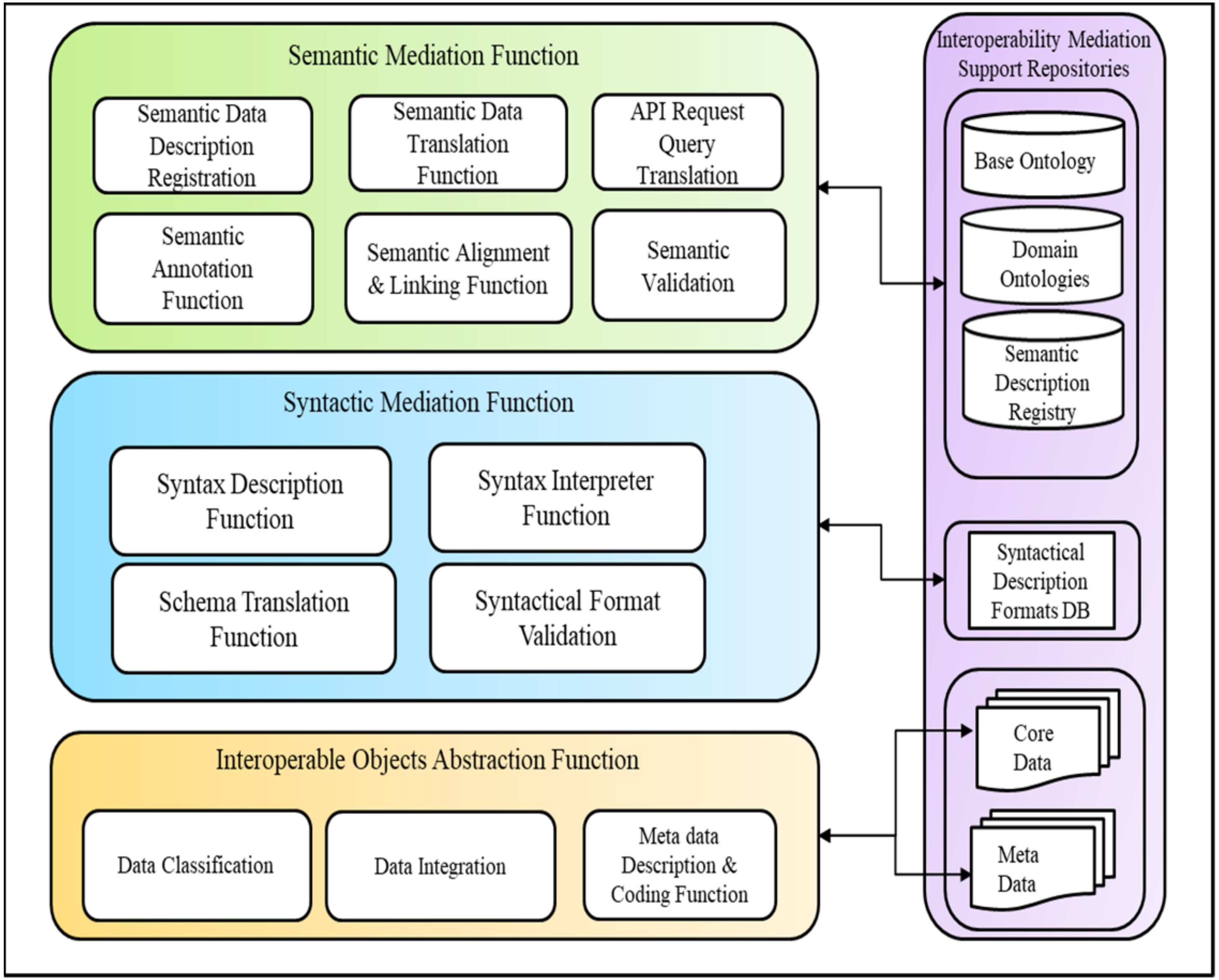
# Functional Model
# to support Data Interoperability

# Functional mediation model to support data interoperability

**Semantic mediation module**

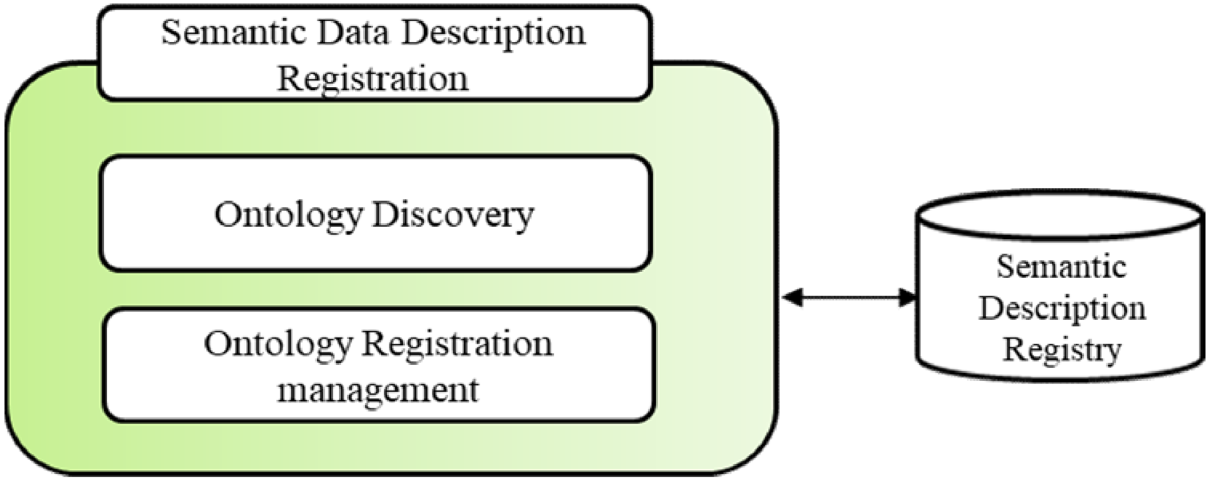
**Syntactic mediation module**

**Interoperable object abstractions**



# Semantic Mediation Functions (1):
## (1) Semantic data description registration

This function provides a registration capability, where platform level semantic data description formats are registered in the semantic description registry.

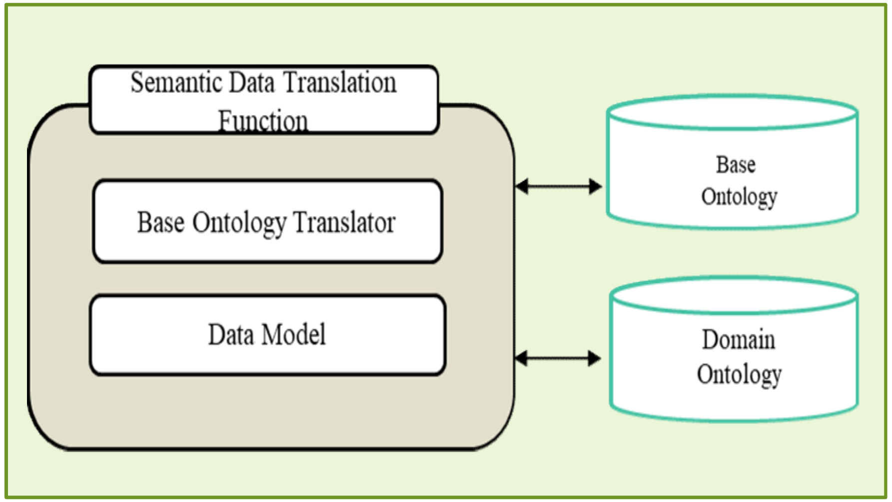



- Ontology discovery: provides the search and matching of ontology records in the semantic description registry.
- Ontology registration management: provides the functional capability to register and manage the semantic ontology models.

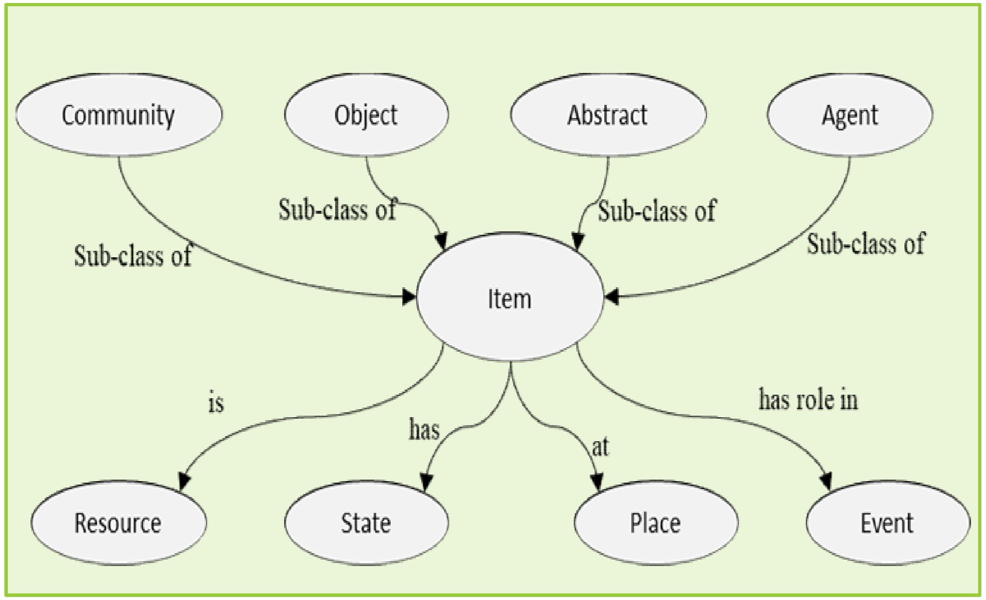
# Semantic Mediation Functions (2):
## (2) Semantic Data Translation

This function enables the translation of data formats to semantic formats registered by semantic registries;

- Base ontology translator: delivers the functional capability of translation of concepts from a domain ontology model to the base ontology model.
- Data Model: provides the capability to express the semantic meaning of the exchanged data using the information objects;



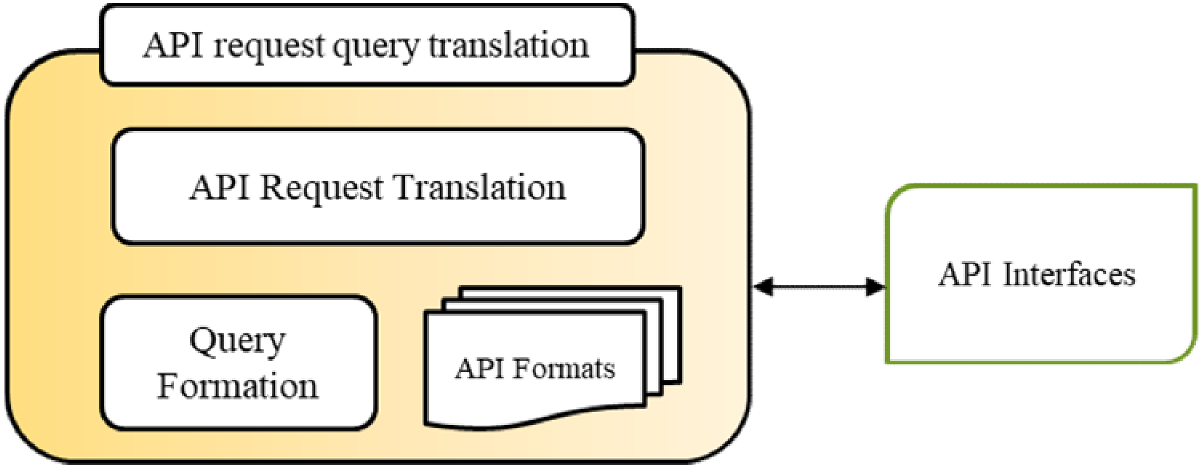

**Semantic Data Translation functional components**




**Semantic Data Model**

# Semantic Mediation Functions(3):
## (3) API request query translation

This function delivers an interface to receive requests for the data to be translated to particular registered semantic format.




- API Translator: provides the functional capability to translate the API request to a target request format.
- Query Formation: provides the functional capability to generate new formatted query
- API Formats: Provide the API formats of the registered platforms

# Semantic Mediation Functions (4):
## (4) Semantic annotation function

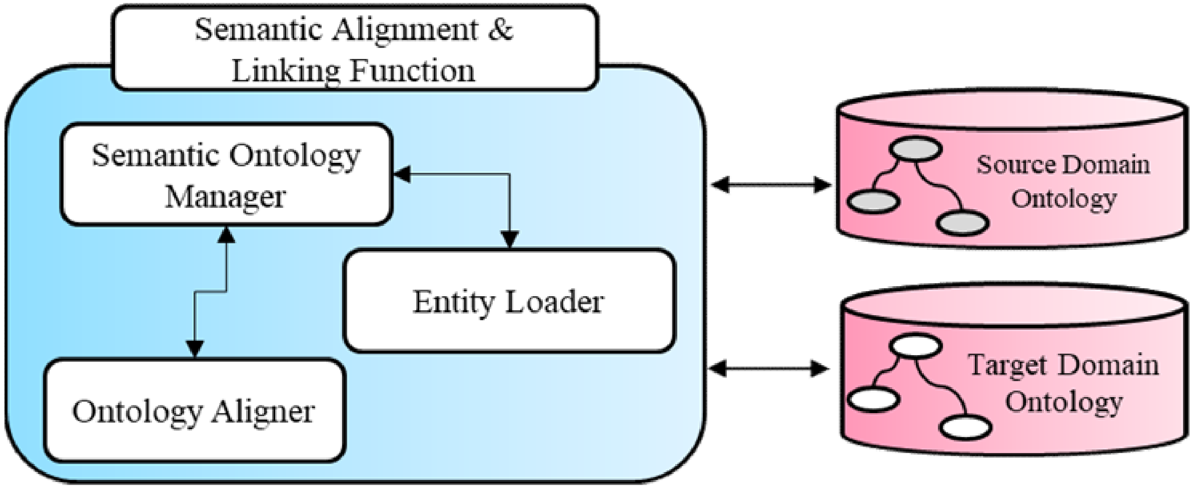
The **Semantic Annotation** in Semantic Mediation changes an isolated ontology into an ontology that can be interpreted, shared and reused by other ontologies.

**The Features of Semantic annotation are:**

- Describe the relationship between concepts and ontologies;
- Link information source to an ontology;
- Assign semantic concepts and properties to the target data.

# Semantic Mediation Functions (5):
## (5) Semantic Alignment and Linking function

This function enables the alignment and management of the source and the target semantic schemas.

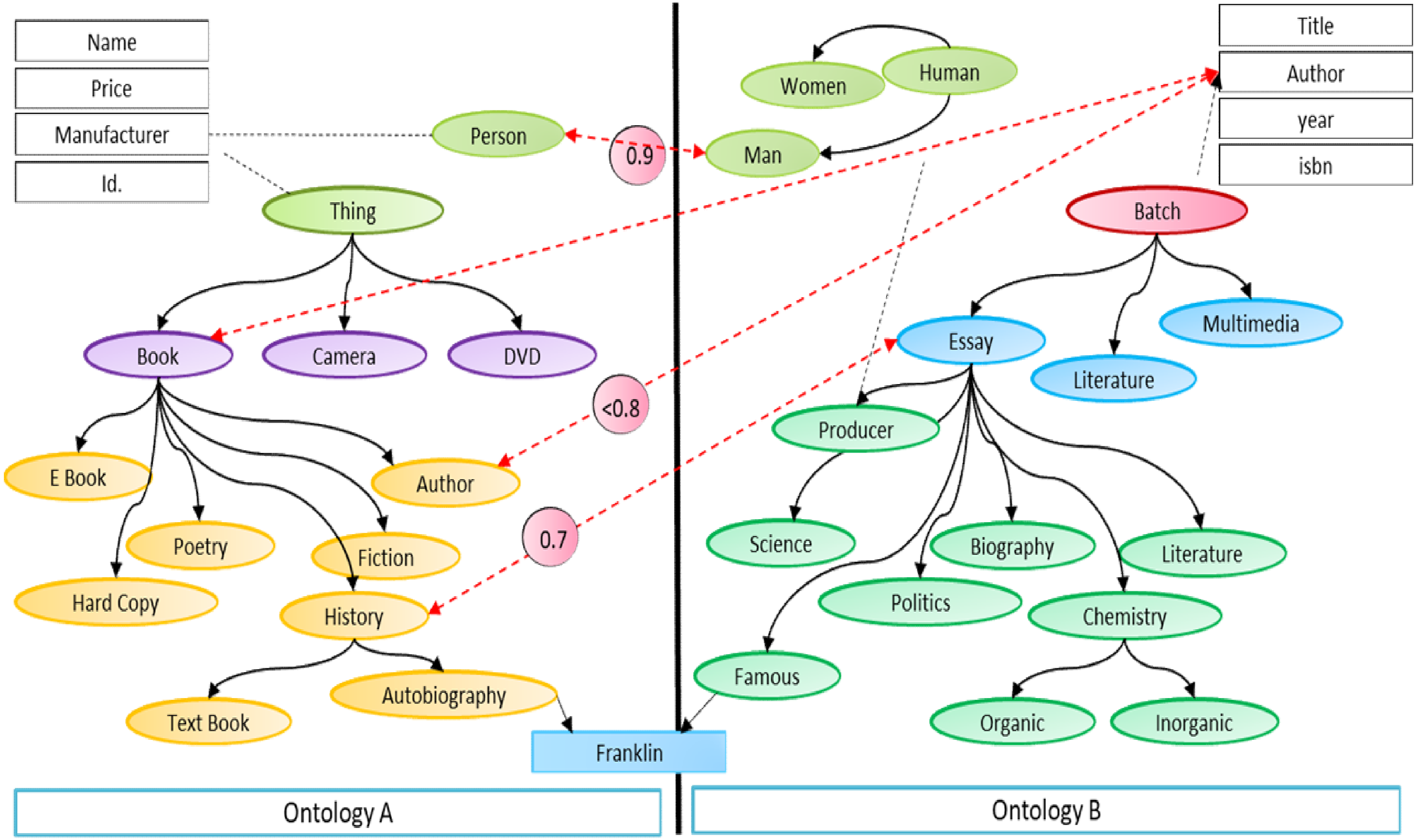



- Semantic Ontology Manager: includes the capability to manage the ontology aligner.
- Ontology Aligner: provides the capability for semantic ontology alignment. It takes the source and target ontology models and returns the alignment results.
- Entity Loader: provides the capability to load the entities from source and target ontologies.

# Semantic Mediation Functions (5-2):

## (5): Example: ontology alignment

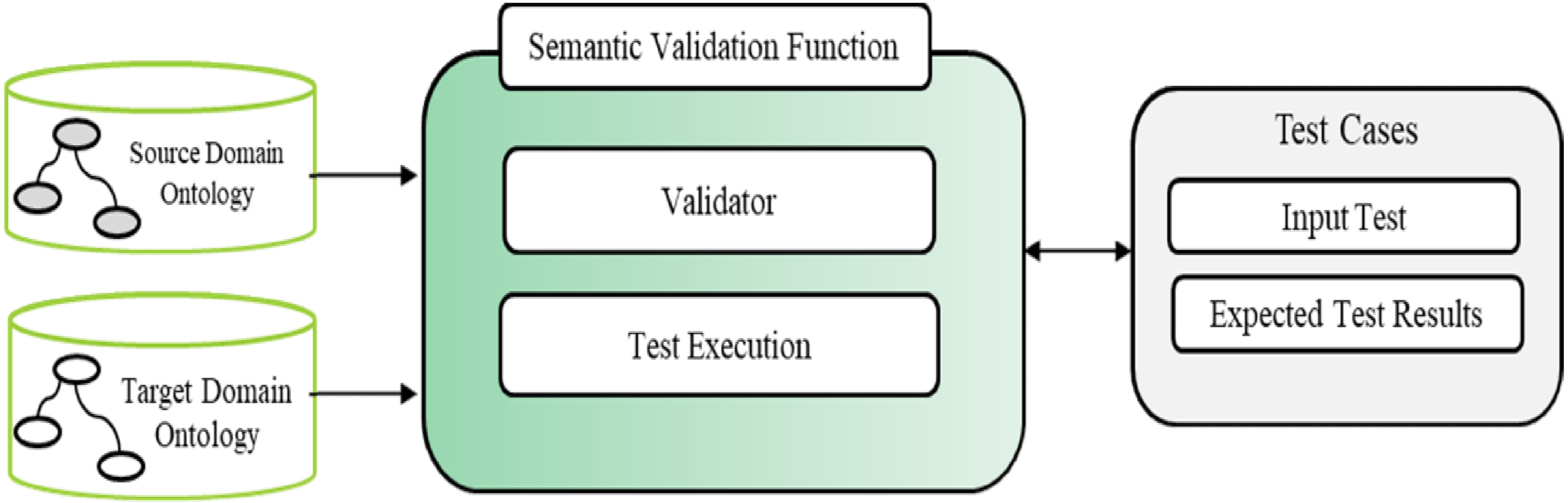
The mapping of matching techniques generates this similarity value. The higher the value between two entities the more exact match is assumed.




anLab

# Semantic Mediation Functions (6):
## (6) Semantic Alignment and Linking function

It constitutes the mechanism to validate the semantic structure of the data with validation test case defined on the bases of semantic ontology.




- **Validator:** this functional capability provides the validation function on the provided input test case to validate the semantic alignment
- **Test Execution:** provides the execution facility to perform the validation for the alignment function

anLab

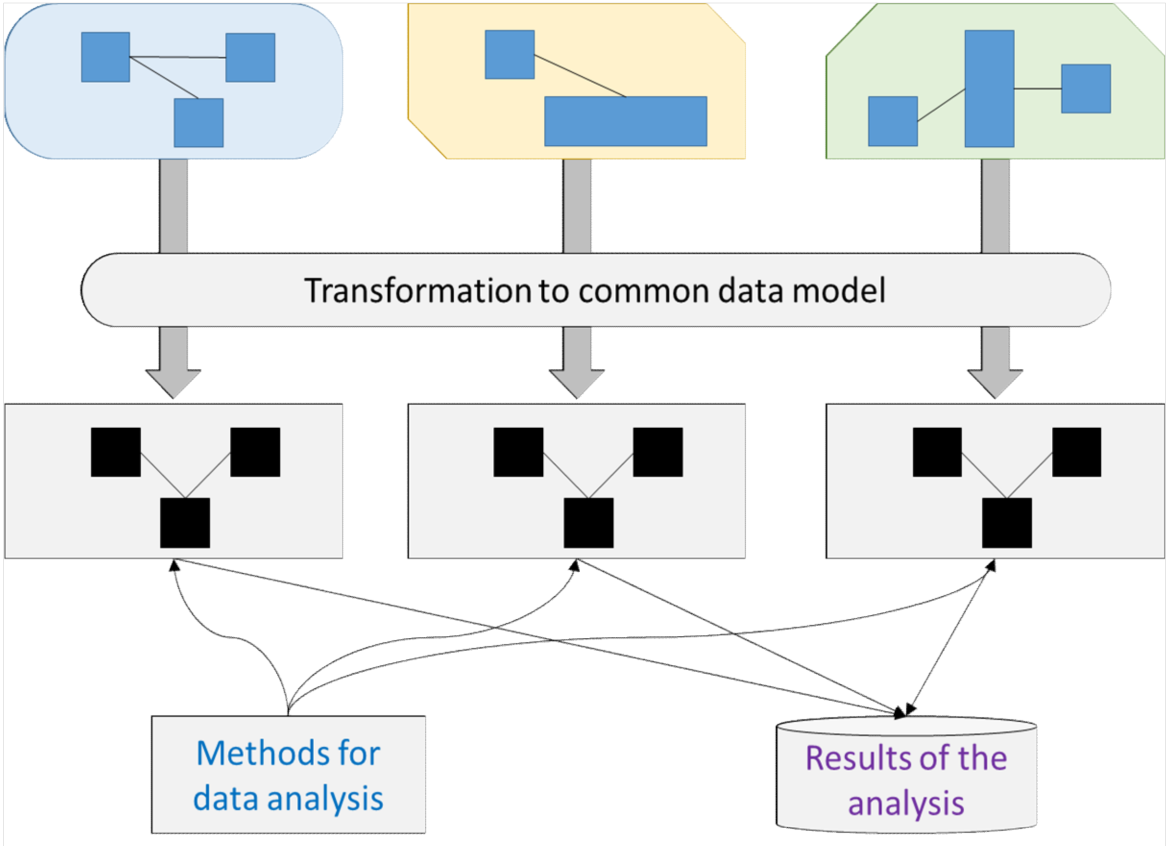
# Syntactic Mediation Functions (1)

## (1) Syntax description function

- Provides registration of syntaxes for the platforms.
- Enables useful records of syntactical formats for the platforms which enables interoperable data sharing.
- **The functional capabilities :**
  - Syntax registration interface: provides interface to record the platform syntax profiles in syntactical description formats DB.
  - Template discovery function: provides the functional capability to discover existing syntax templates.
  - Syntactical metadata management: provides the functional capability to manage metadata related to the syntactical models.

# Syntactic Mediation Functions (2)

- **<u>Syntax description function</u>** utilizes a common data model (CDM) to be mapped among different interoperating systems.
  - **Features of CDM:**
    - To allows transformation of data into a common format or CDM, common terminologies from standard vocabularies.



# Syntactic Mediation Functions

**(2) Syntax interpretation function:**

- Provides conversion among diverse data formats.
- The functional capabilities:
  - **API syntax convertor:** provides capability to translate the API syntax description to standard format.
  - **Profile manager:** delivers the capability to store and retrieve syntax profiles

**(3) Schema translation function**

- To support the translation mechanism at schema level.
- Different platforms make use of the conversion to interoperate the schemas at platform level.
- **The functional capabilities of Schema translation function are:**
  - **Core schema translator** to translate the core data schema with respect to syntactical conversion in CDM.
  - **Metadata schema translator** to convert metadata formats to describe core data into common Metadata formats.

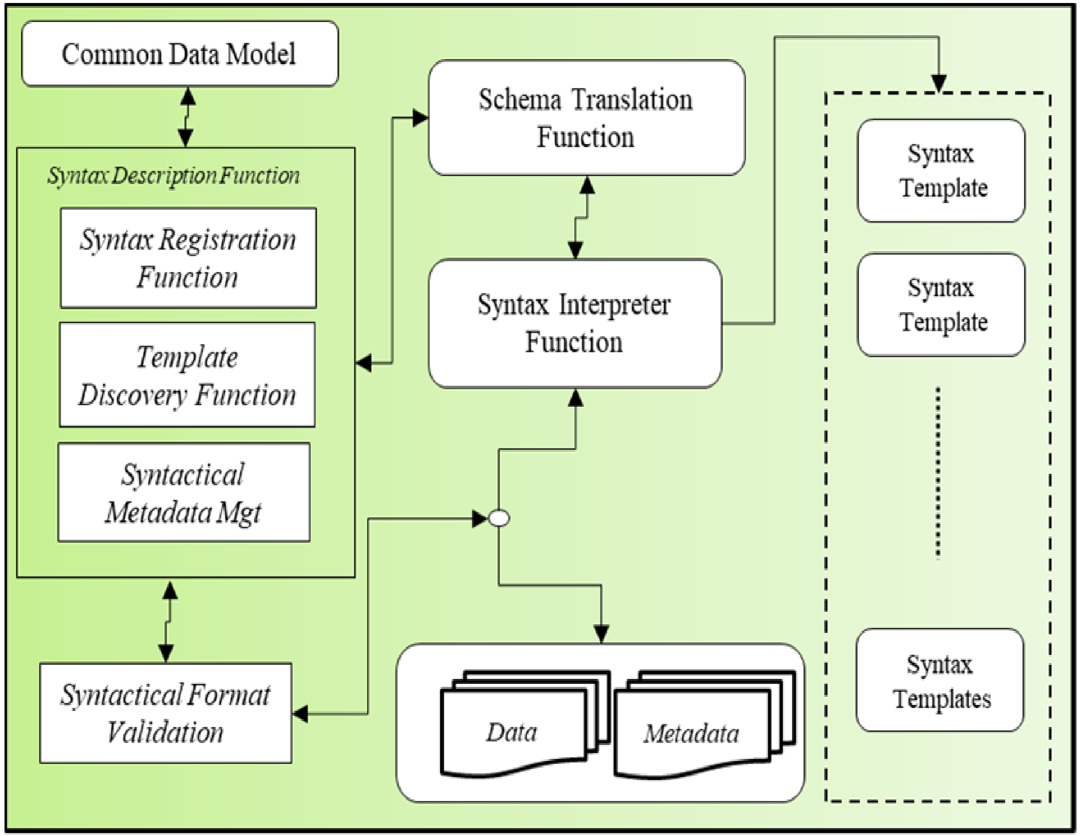
# Syntactic Mediation Functions

## (4) Syntactical validation function

- **Provides the mechanism to validate the syntactical structure of the data based on the defined core schema**
- **The functional capabilities of semantic validation function**
  - **Syntax format validator:** this capability delivers the syntax level validation function on the provided input syntax profile to validate the translation.
  - **Syntactical testing container:** provides the run time facility to execute the validation for the syntactical translation function.

# Syntactical mediation components Interaction System (SIS)

- **Syntax translation function** produce the conversion schema through the syntax interpreter service based on syntax descriptions.
- **Syntax Interpretation** is to generate the syntactical alignment based on the syntax that is chosen for the subject alignment.
- Syntax description and management services throughout the syntax translation perform three major tasks
  - Discovery of syntax template from the template repository;
  - Syntax registration, once a conversion has been performed successfully;
  - Syntactical Metadata management to support additional syntax level checking.



# Interoperable Object Abstraction Function (1)

## ■ Data Classification Function

- provides capability for the classification and categorization of data.
- enables the classified representation to be understandable through interoperable objects abstraction model.
- classifies the data based on the metadata
- Provides tagging of the data according to the object abstraction vocabulary.

## ■ Data Integration

- provides procedures to integrate the data from different sources
- based on the classified category, (categories are known);
- based on the metadata description;
- based on the fixed object abstraction ontology,

# Interoperable Object Abstraction Function

- **Metadata Description and Coding**
  - Provides the capability to assign new **metadata or assign additional metadata** to the converted data.

The **metadata** helps in several tasks such as:

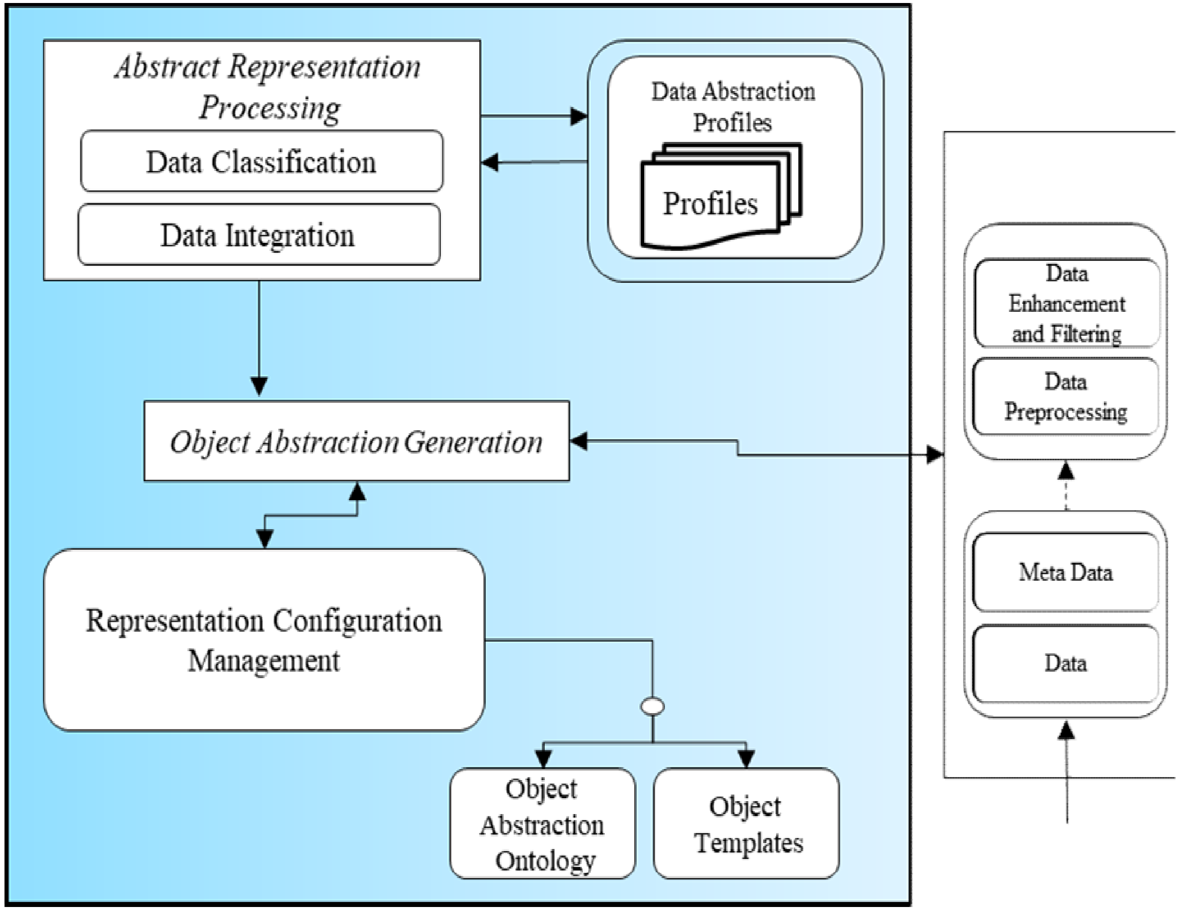
- Discovery, identification and classification of data;
- Describes the relationship and characteristics of data;
- Provides when the data is created and transformed and how, types of data file and other technical information are stored.

- The metadata description function uses a Metadata ontology as a library of metadata to describe and publish metadata.
- A metadata ontology allows users to search, refer and evaluate metadata, uses metadata standard or controlled vocabularies.

# Interoperable Object Abstraction Functions

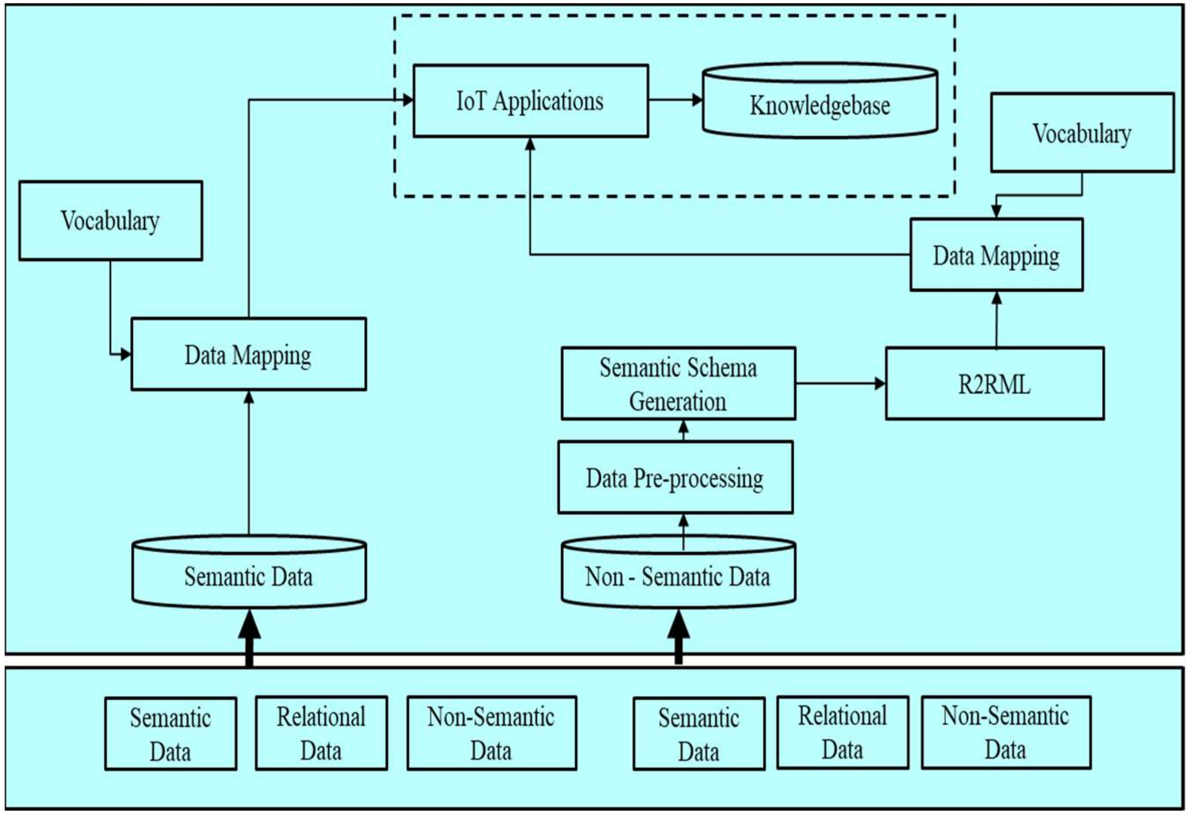
## Interaction of object abstraction functions

- Data processing and enhancement is to enable the instantiation of abstraction from existing profile;
- Data abstraction profiles are a set of profiles P = { P1, … , Pn }, each represents a unique profile based on the type of data;
- Catalog of profiles is maintained in central profile repository;
- Interoperable representation configuration discovers the appropriate template



# Data processing for semantic and non-semantic data

- Consider both data from semantic and non-semantic sources;
- **Transformation of semantic and non-semantic data** (e.g. relational), integration and mapping to RDF triples;
- **For transformation, integration, and mapping,** data requires schema;
- **Semantic data source**
  - Knowledge base;
  - semantic ontology to express semantic relationships among data;
  - Data in RDF format and stored in OWL;
  - Knowledge base composed of domain terminology (concept definition) and their instances.

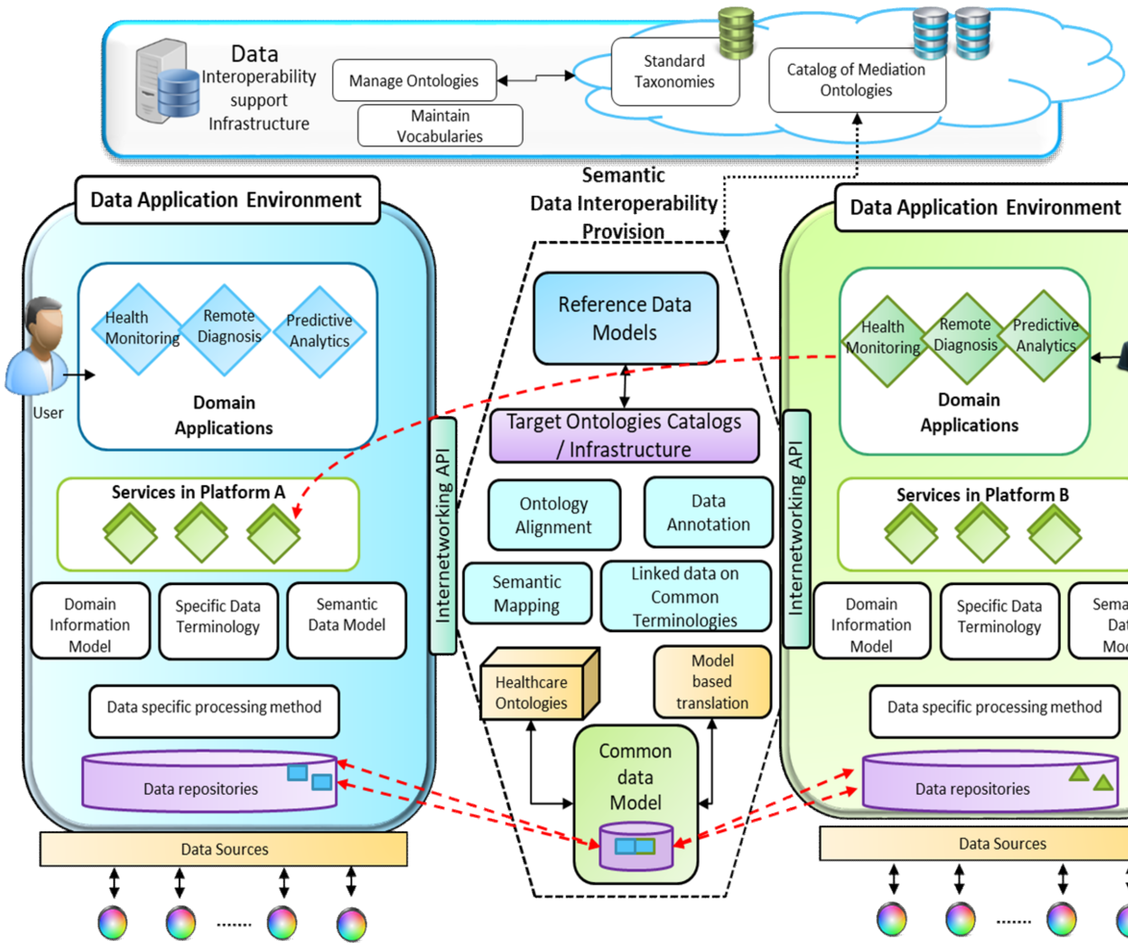


# Experimental Result:

Semantic Data Interoperability Provisioning through WoO Framework indicated in [ITU-T Y.4452]

# WoO enabled Semantic Interoperability Provisioning Model

- **Data interoperability provisioning in heterogeneous environments**
  - WoO Provides crucial components to enable semantic data sharing among heterogeneous data applications.
  - Catalog of base ontologies to support semantic mediations;
  - Supported with Data interoperability infrastructure based on microservices to provide interoperability features;
  - Provides an interface to map the elements in ontology catalog.



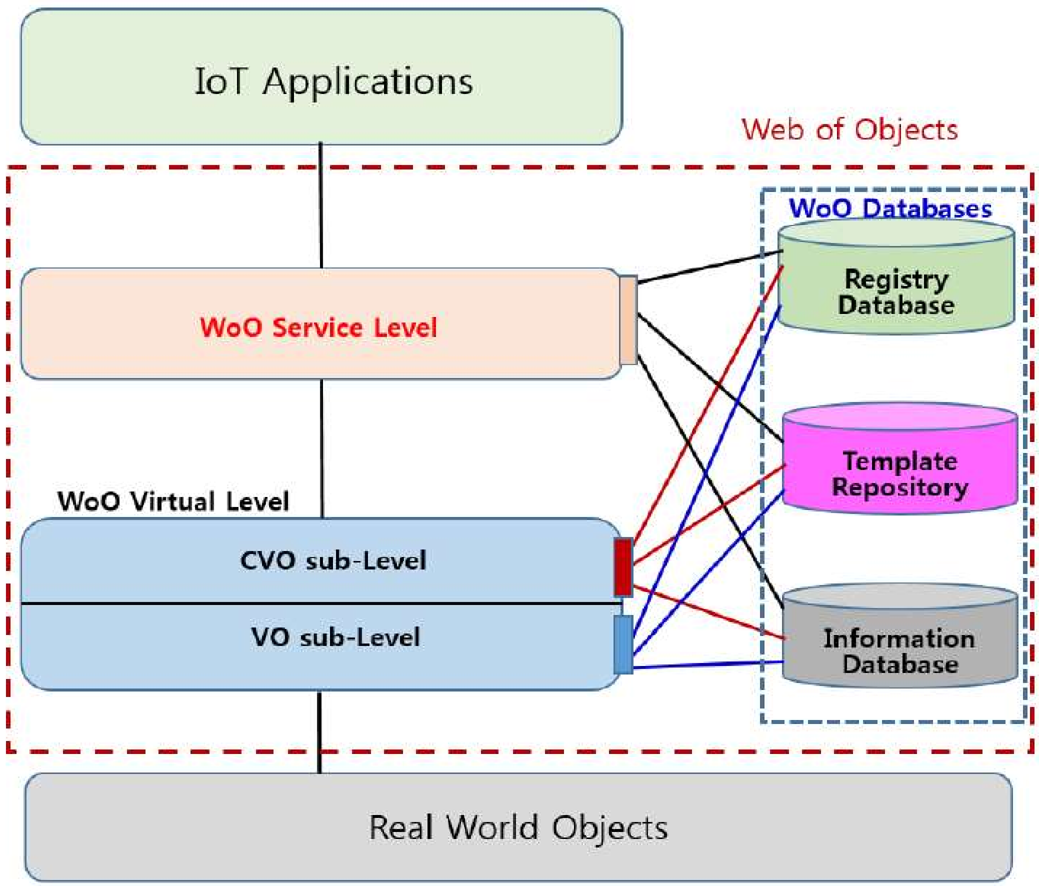

WoO

anLab

# WoO enabled Semantic Interoperability Provisioning Model

Web of Objects (WoO) (in [ITU-T Y.4452]) based architecture is used to support "Common  Data Model (CDM) for data interoperability provisioning

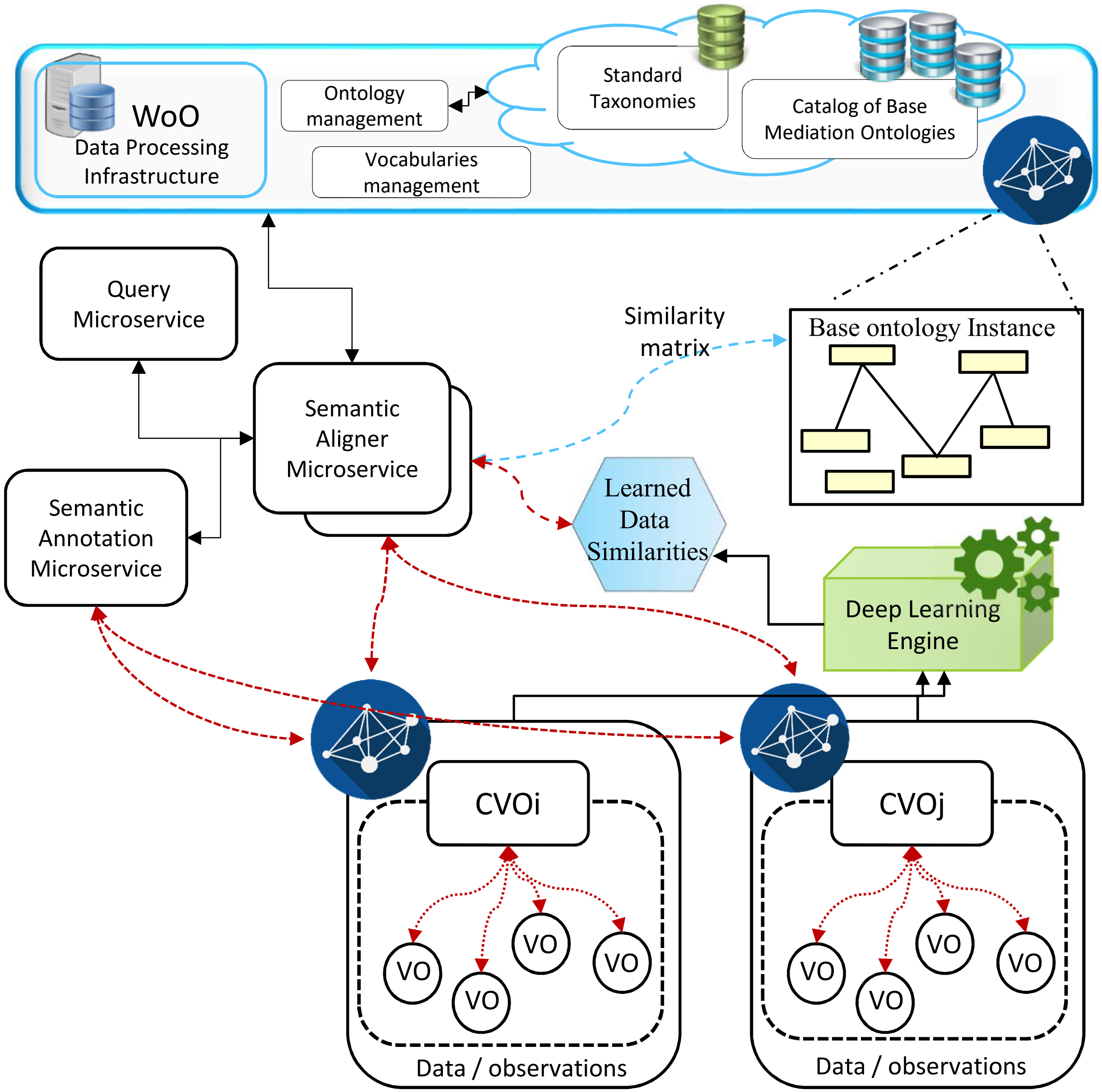
- WoO provides an efficient and comprehensive infrastructure to support data interoperability with layered model
- In particular, Semantic Mediation functionalities are mapped in the "Service Level" of WoO framework to allow :




anLab

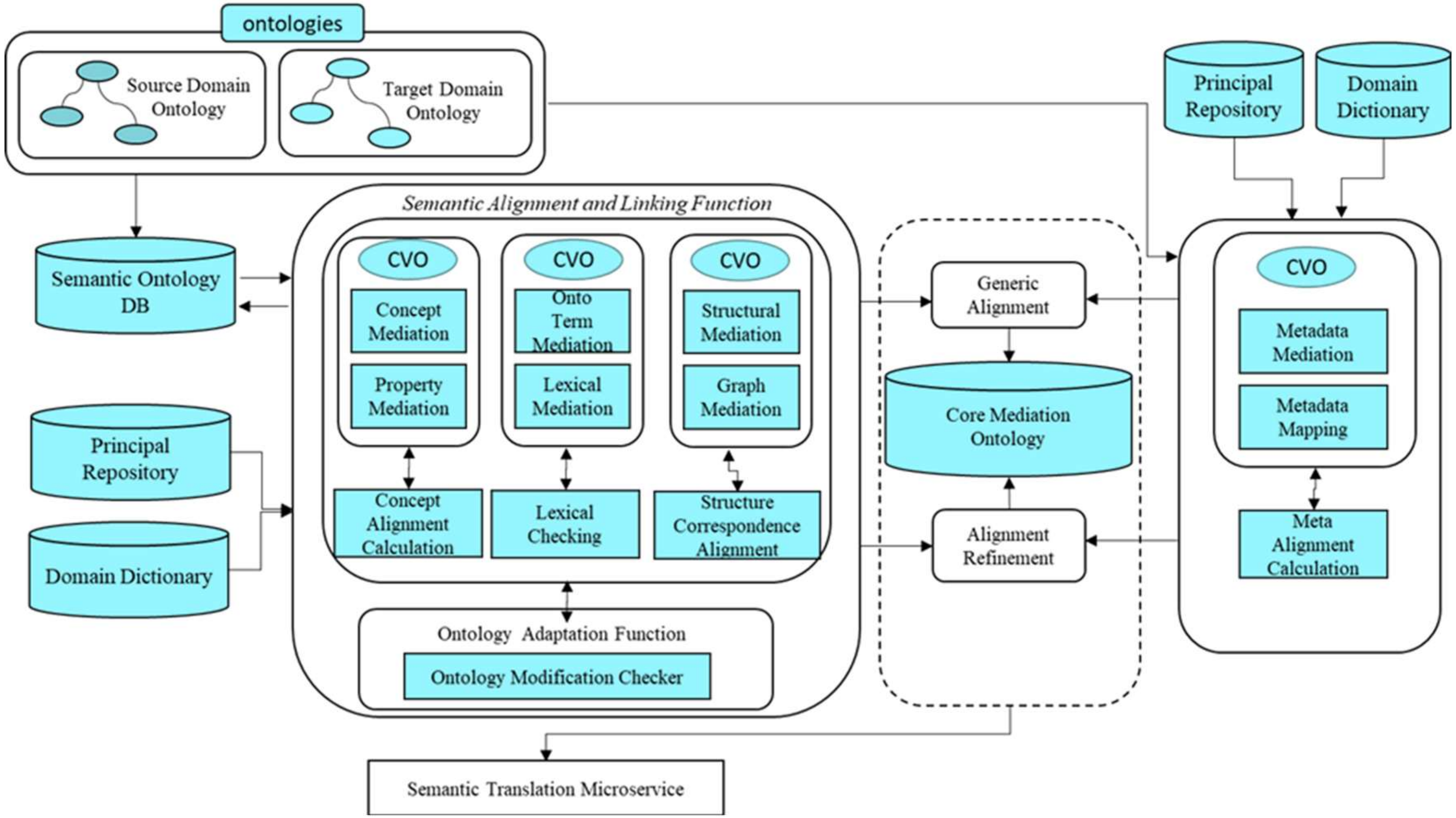
# WoO enabled Semantic Interoperability Provisioning Model

- **Semantic alignment based on WoO environment**
  - Semantic aligner microservice provides ontology alignments based on the learned similarities among ontologies;
  - Maintains the similarity alignments matrix in the base ontology instances;
  - Query microservice searches the similar terminologies based on the extraction specification.
  - URI of data element specification searched in target ontology catalogs



# Procedure for semantic alignment

- Functional components designed in CVO and data maintained in VOs;
- Mediation and alignment of source and target ontology concepts in RDF;
- Mediation of the data properties and object properties in the ontologies;
- Process of concept alignment performed in RDF formats.



# Semantic Interoperability Provisioning Model

- **Semantic Annotation of concepts in the model**
  - Annotation performed by enriching the contents from a domain object by linking the information with the semantic ontology;
  - Annotation of information helps the content to be machine-readable and semantically interoperable;
  - A multi step annotation process has been developed for healthcare concepts.

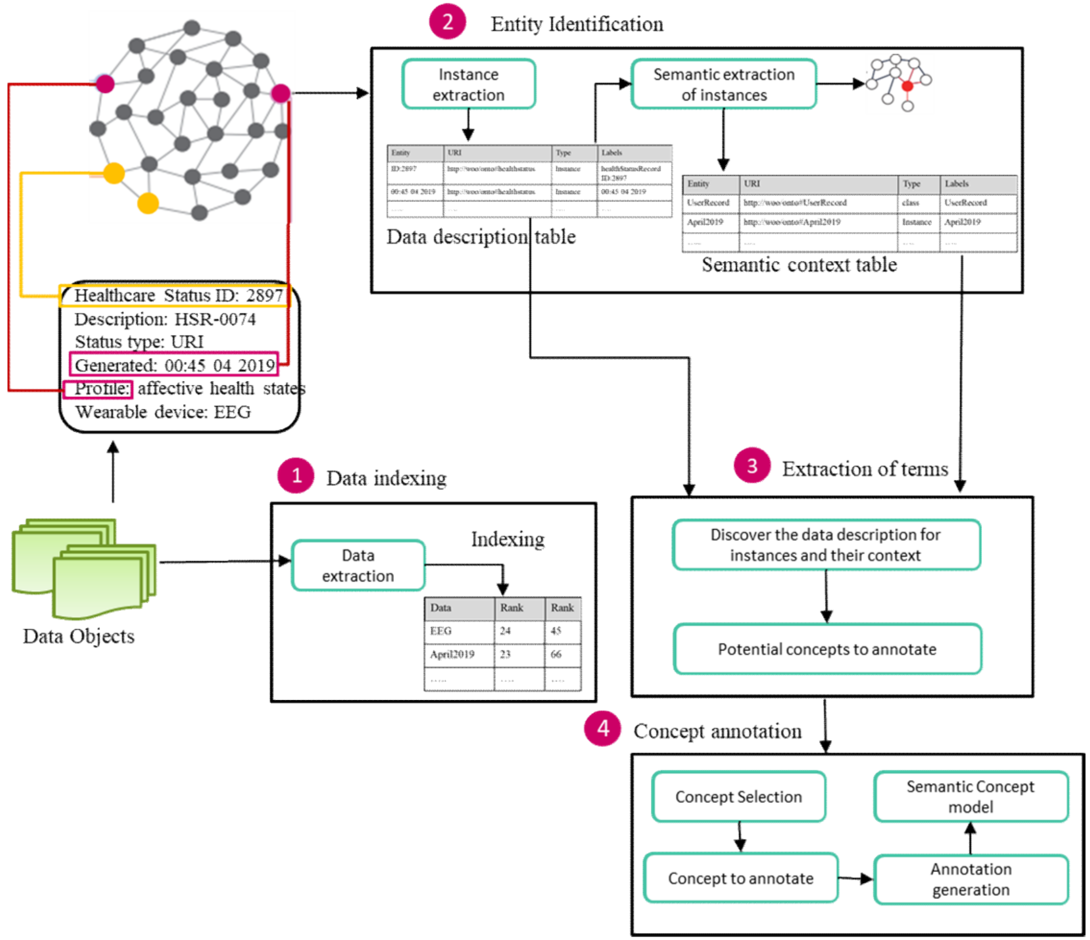
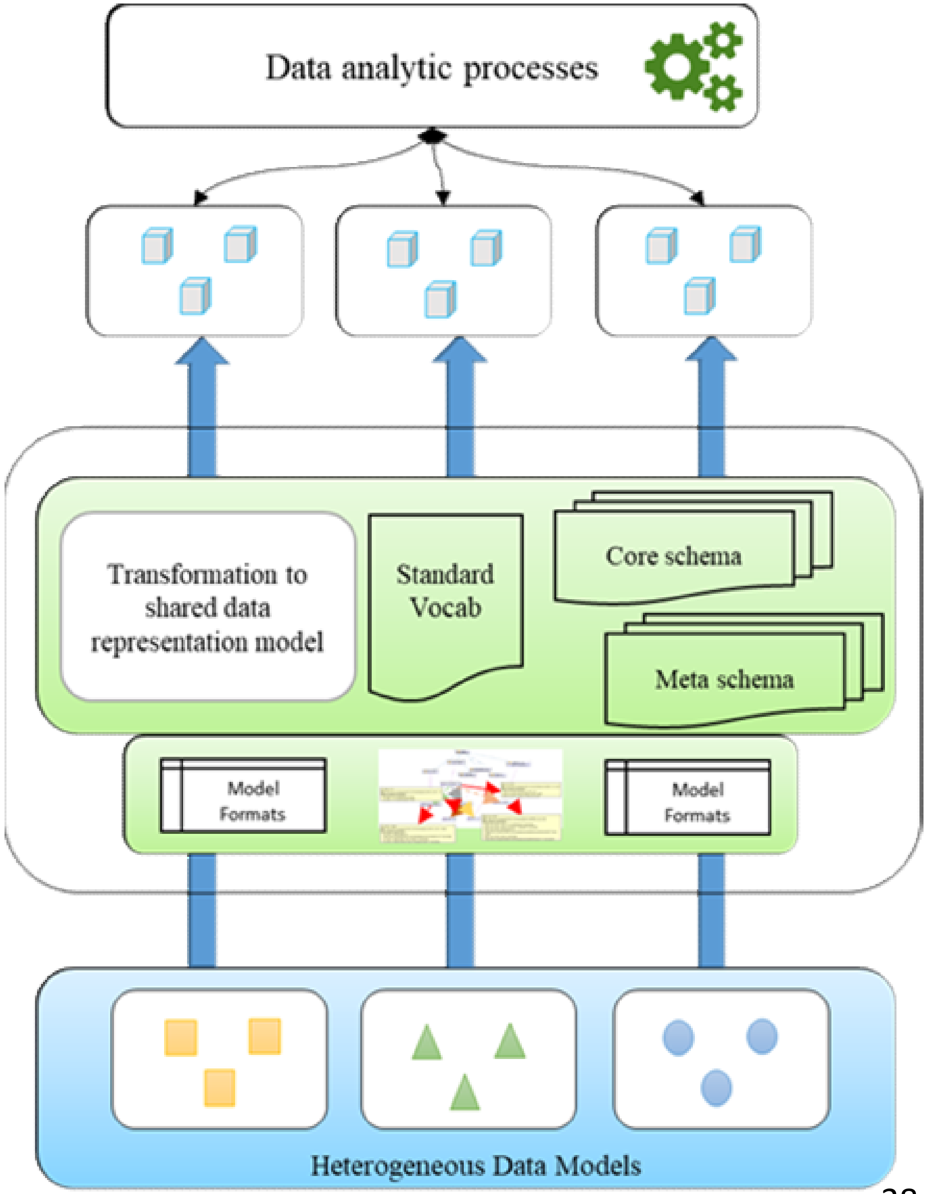



Figure. WoO based Semantic annotation for data interoperability

# Common Data Model (CDM)

## Common Data Model

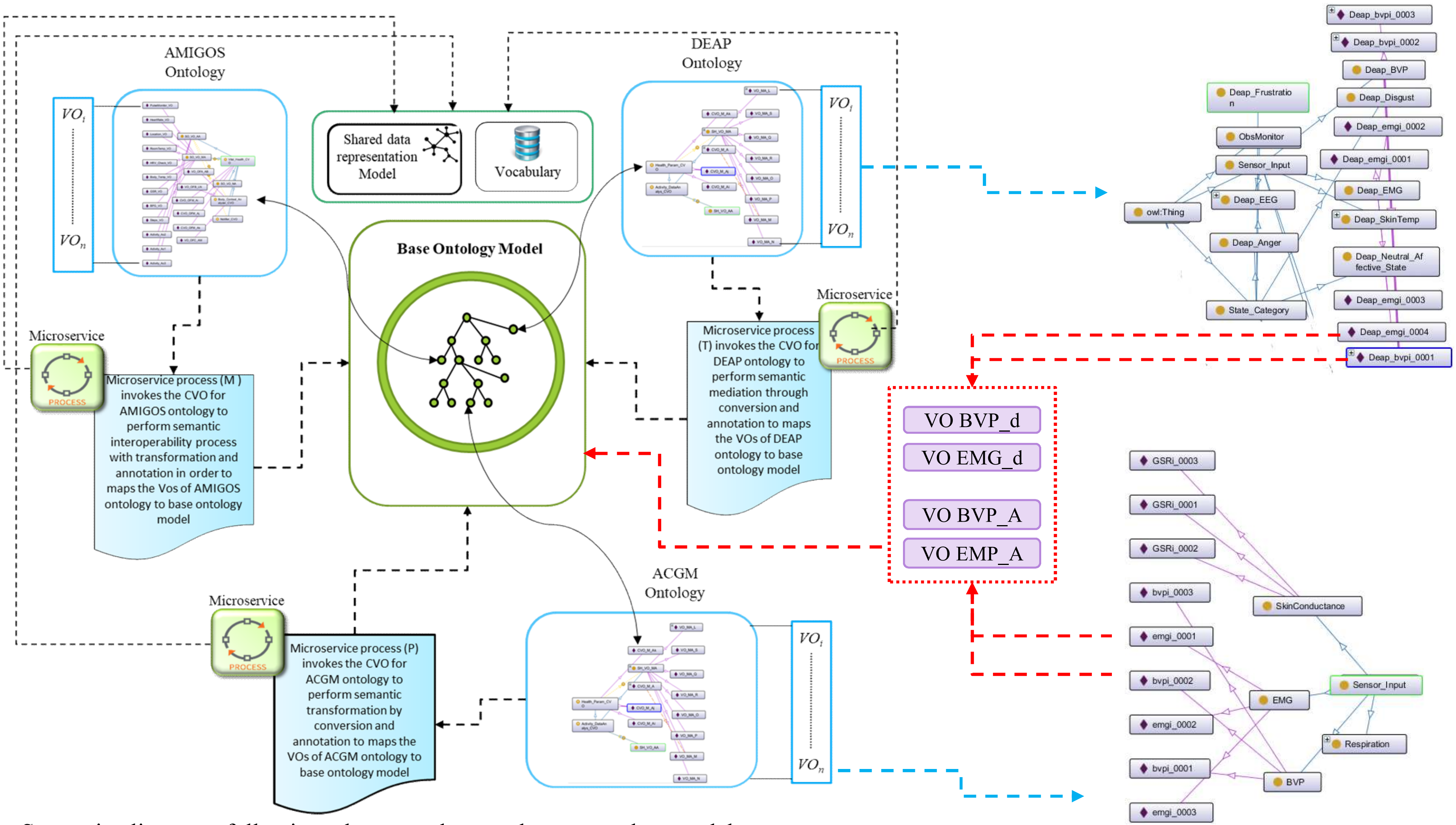
- **CDM model** involves standard vocabularies, core schema, meta-data schema and the common data format.
- **Data elements** follow core schema (a collections of reference concepts) and meta schema (a set of reference metadata) based on standard terminology.
- **Features**
  - Expressive description of data elements;
  - Commonly understood meaning of data across diverse domains;
  - Reusability of data among different applications;
  - Uniform transformation of data into standard formats.




anLab

# Scenario of Semantic Interoperability Provision

- Physiological health data from heterogeneous sources are processed and integrated
  - Microservices invoke the CVOs to extract VO instances to integrate in base ontology reference;




Semantic alignment following a base ontology and common data model

anLab

# Scenario of Semantic Interoperability Provision Procedure

- Semantic Alignment (with RDF mapping) for enhancing the semantic sharing;
- Annotation of shared data with mapping on standard vocabulary;
- To support the interoperability of these different annotation schemes a common annotation based on the base ontology model

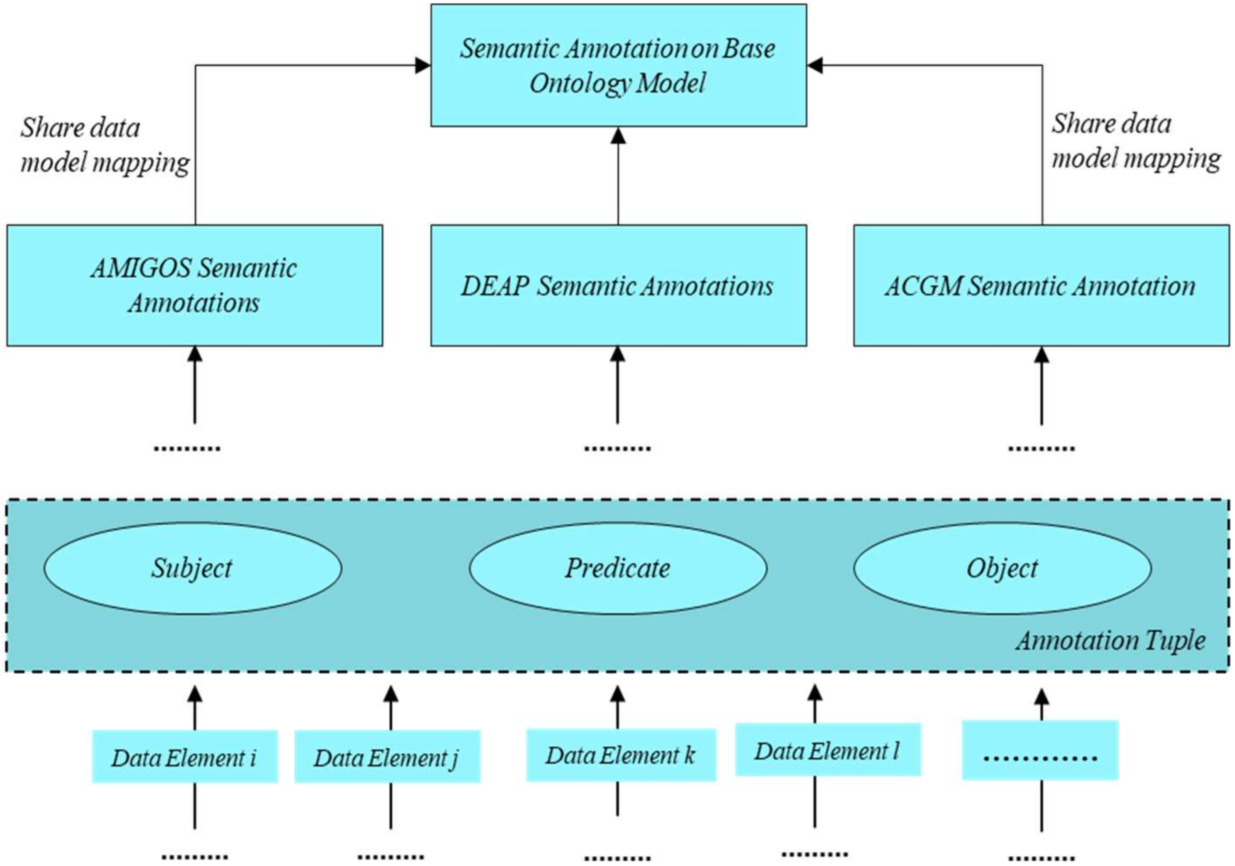



Figure. Example semantic annotation using RDF

# Semantic Interoperability Provisioning Model Components

## WoO based microservices for Semantic interoperability.

- Microservice representation model provides distinct meta-data properties for description of semantic objects;
- Provides meta-data required to interact with the CVOs.
- Meta-data related with the functions describe the tasks of a microservice and policy attributes describe the level of accessibility;

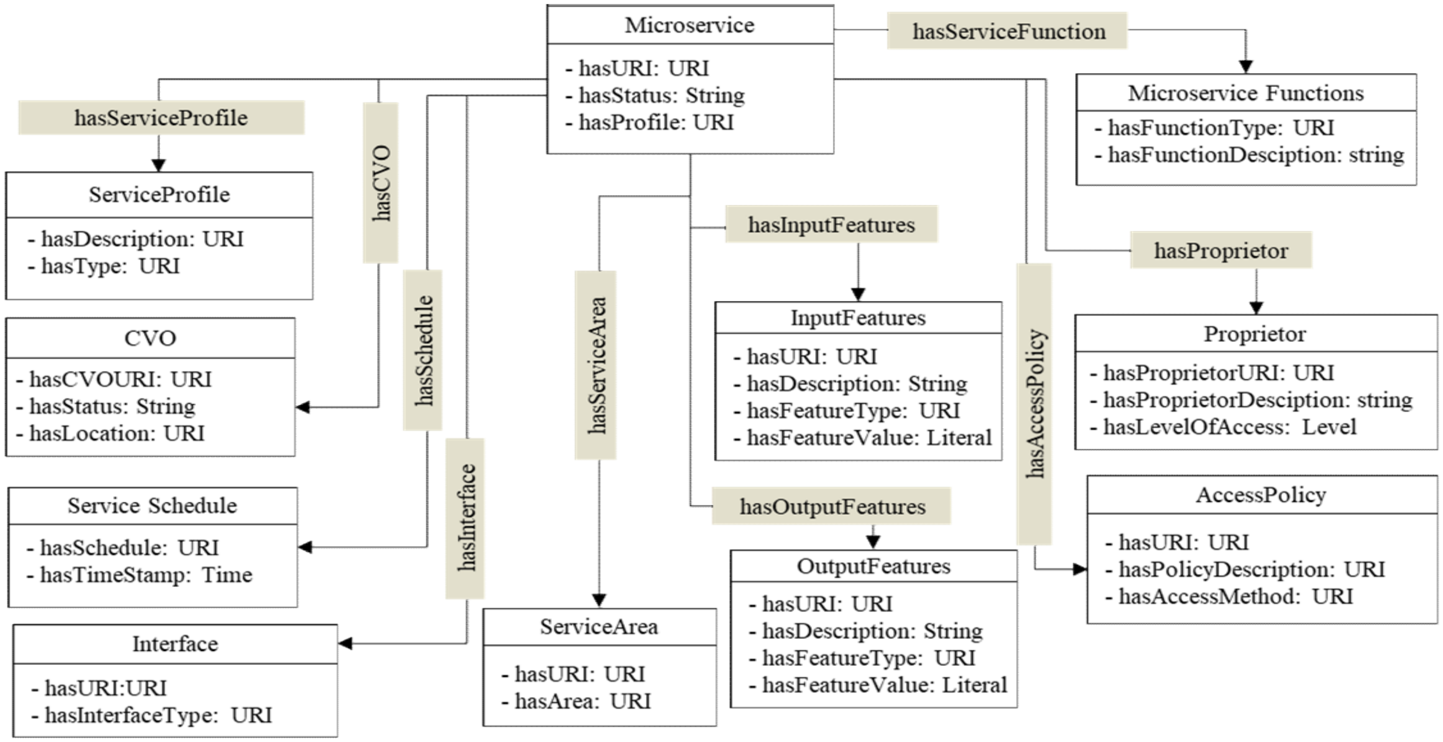



Figure. Microservices representation model

anLab

# Semantic Interoperability Provisioning Model Components

- Generic model of microservices has been designed to deal with the heterogeneity of objects;
- Process of instantiation in microservices includes choosing the most appropriate CVOs from the list of CVOs that are available the repository and a mashup of the VOs;
- If no match for CVO to satisfy the generated request; a new instance of CVO is instantiated based on the primary template of the CVO;

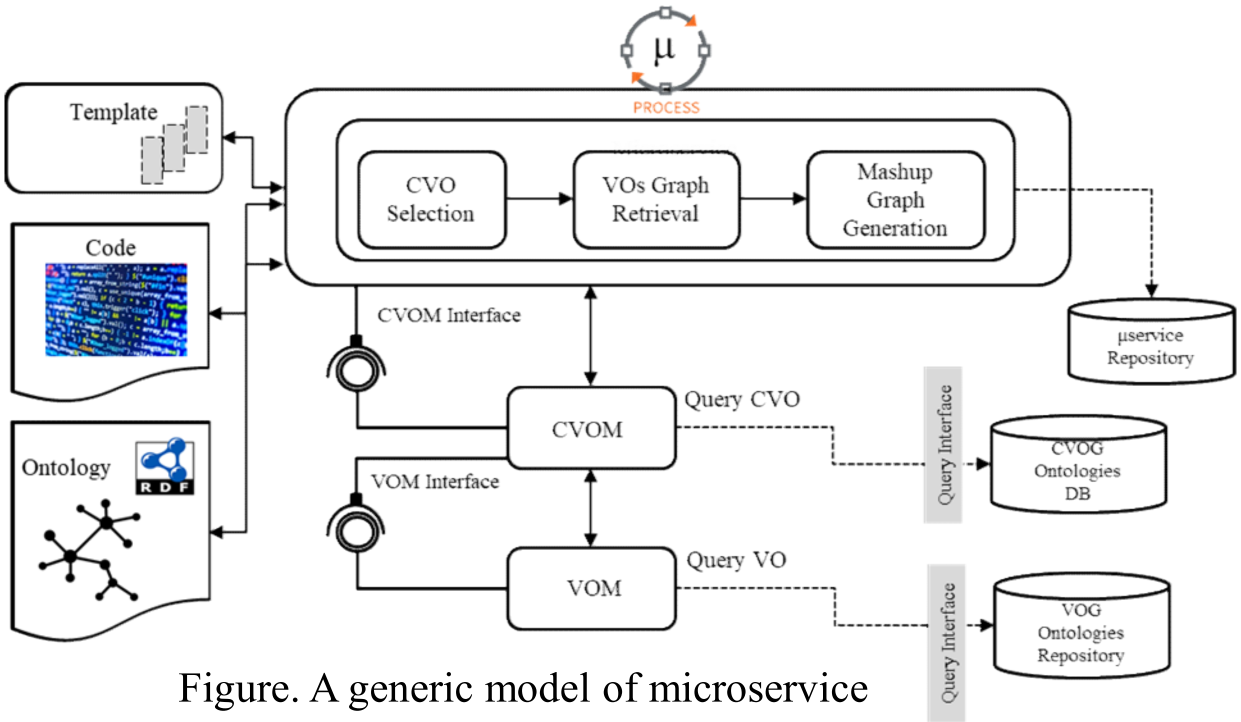



Figure. A generic model of microservice

# WoO Semantic Interoperability Provisioning Model Implementation Settings

**Implementation architecture for data interoperability;**

- Semantic annotation with microservices has been modeled;
- Semantic alignment of data with mediation ontology has been achieved;
- Base ontology catalog has been designed;
- Liking with standard terminologies provided;
- IPS provided to retrieve the data elements;
- Query processing services formed for standard Sparql queries.

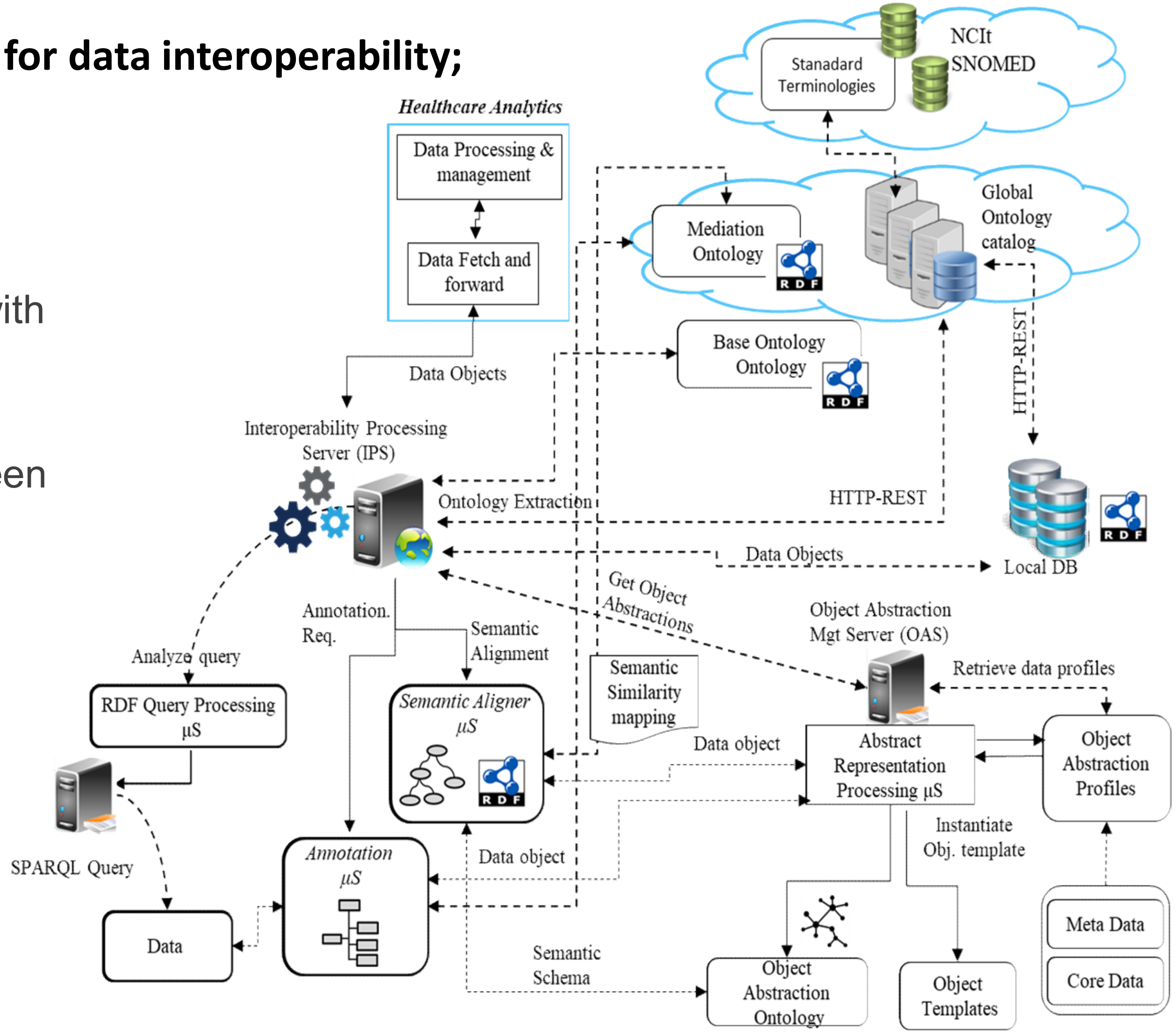



Figure. WoO based Semantic Interoperability Implementation architecture

# Semantic Interoperability Provisioning using Deep Learning

- Semantic alignment based on deep representation learning
  - (a) Learn the representations of source and target ontology graphs;
  - (b) Provide semantic alignment

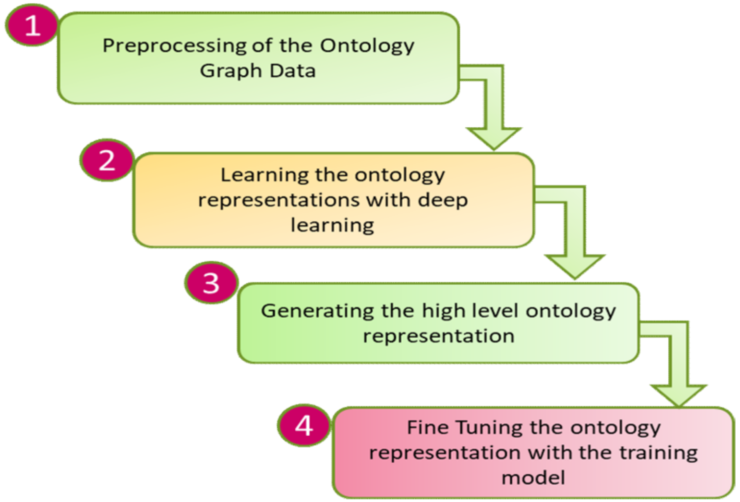



Figure. Deep representation learning process

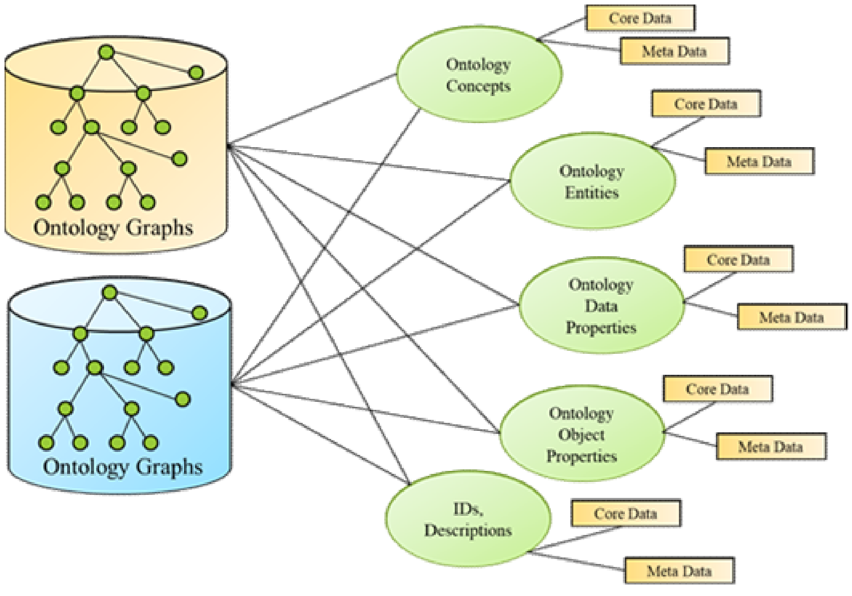



Figure. Ontology graph alignment data descriptions

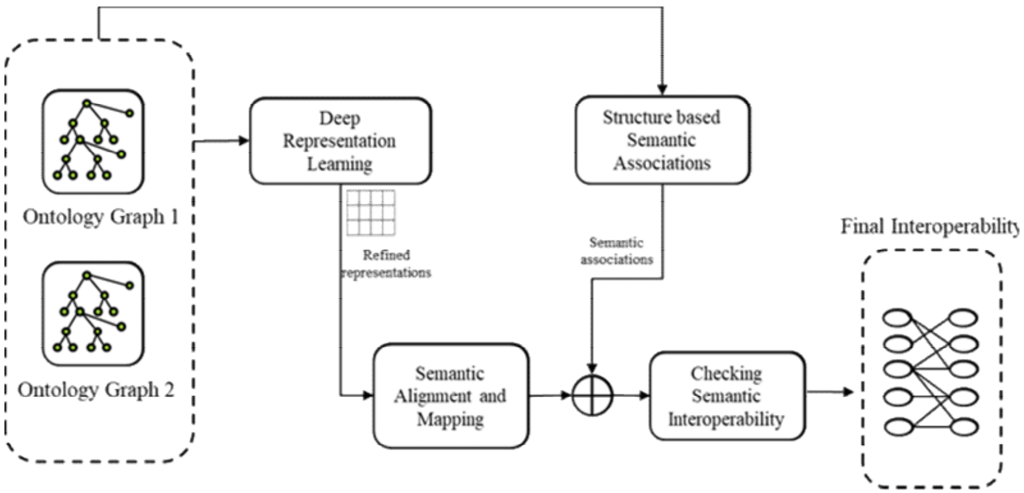



Figure. Composite semantic alignment process

# Data Interoperability Provisioning using Deep Learning

- Deep learning based semantic alignment computing procedures based on representation learning;

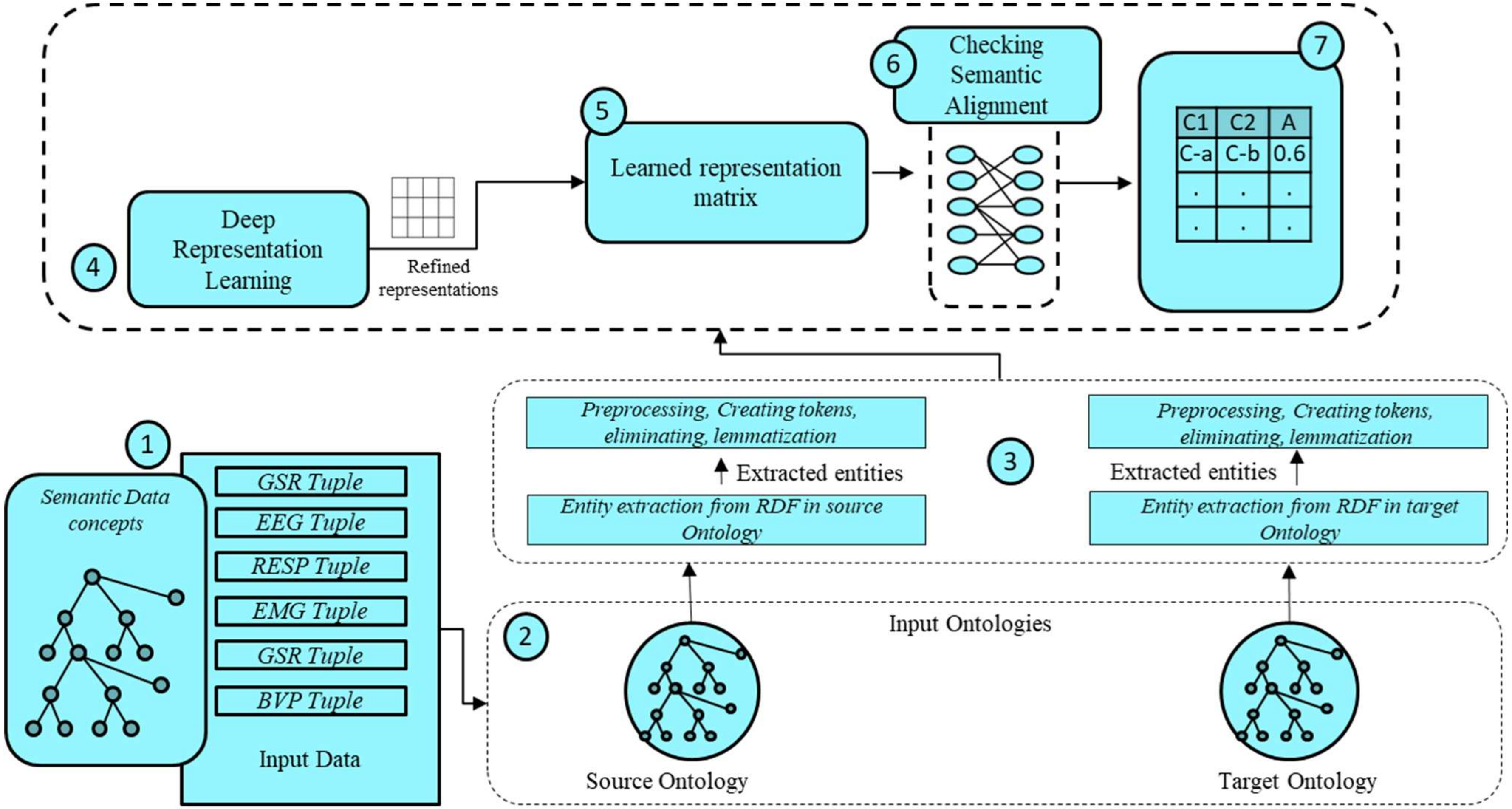



Figure. Ontology data preparation and deep representation learning process steps

# Semantic Interoperability Provisioning using Deep Learning

- Deep learning model for ontology alignment;
- Input data and preprocessing procedure for computing semantic similarities of ontology models

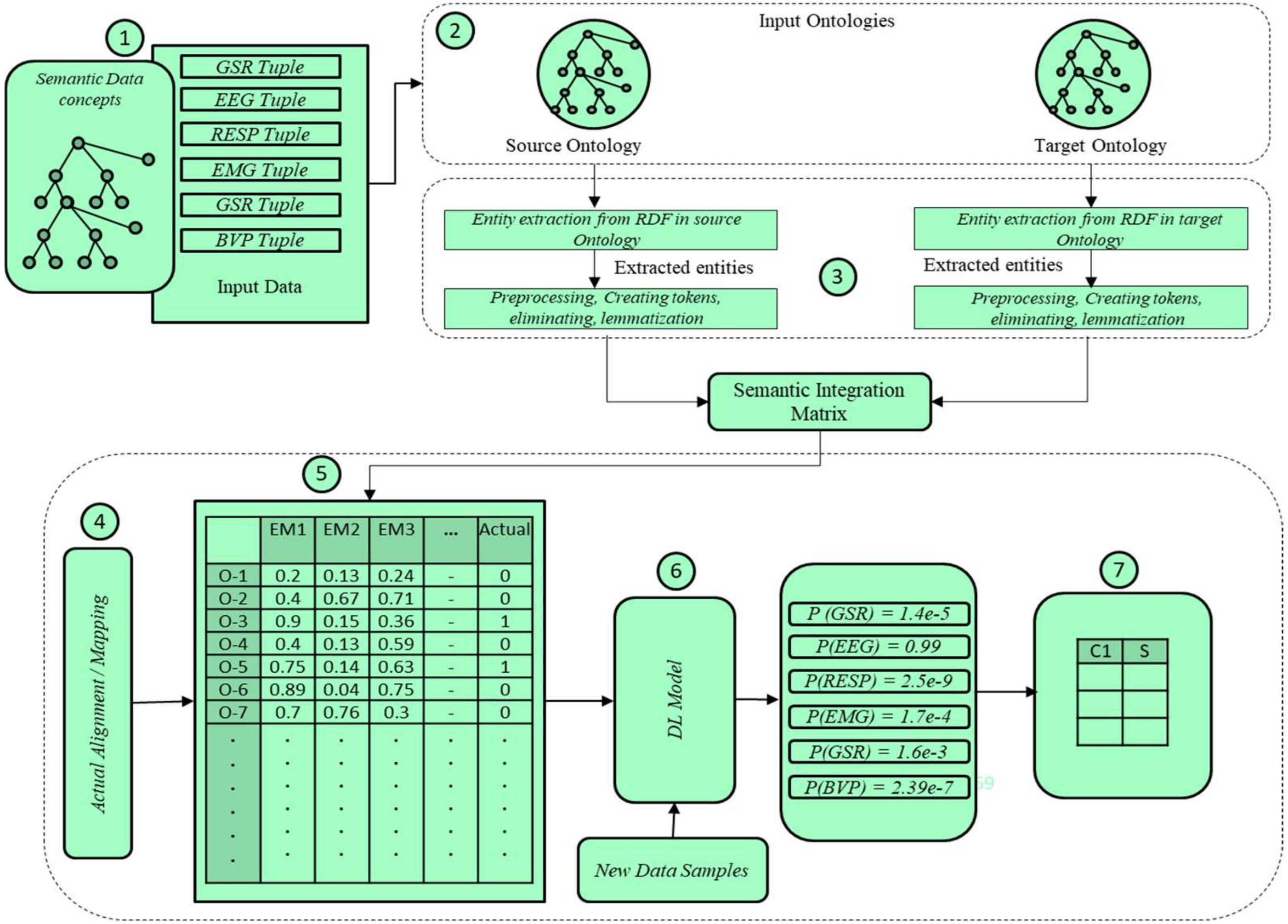
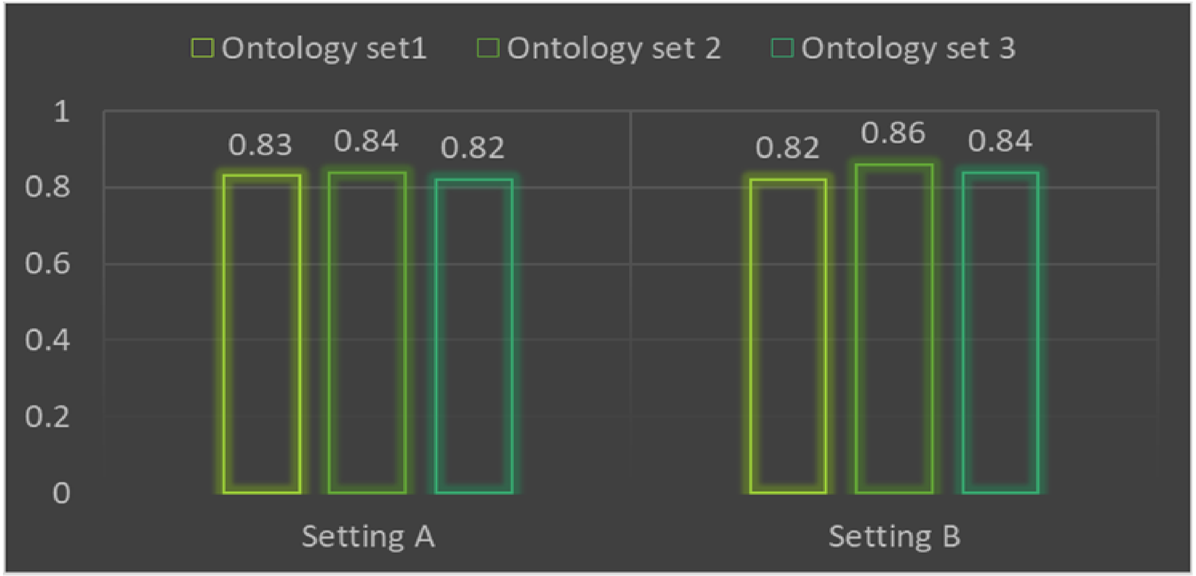



Figure. Leaning semantic similarities based on integrated matrix for enhanced integration

# Results

- Evaluating the deep learning model on standard benchmark dataset of three ontology sets from Ontology Alignment Evaluation Initiative (O.A.E.I) [3].
- Two settings has been defined.
  - Setting-A configuration of deep learning model (2 learning layers with 400+ epochs)
  - Setting-B configuration of deep learning model (3 learning layer with 600+ epochs)
- Improvement in accuracy achieved
  - 12.4% with Setting -A
  - 11.5% with Setting -B

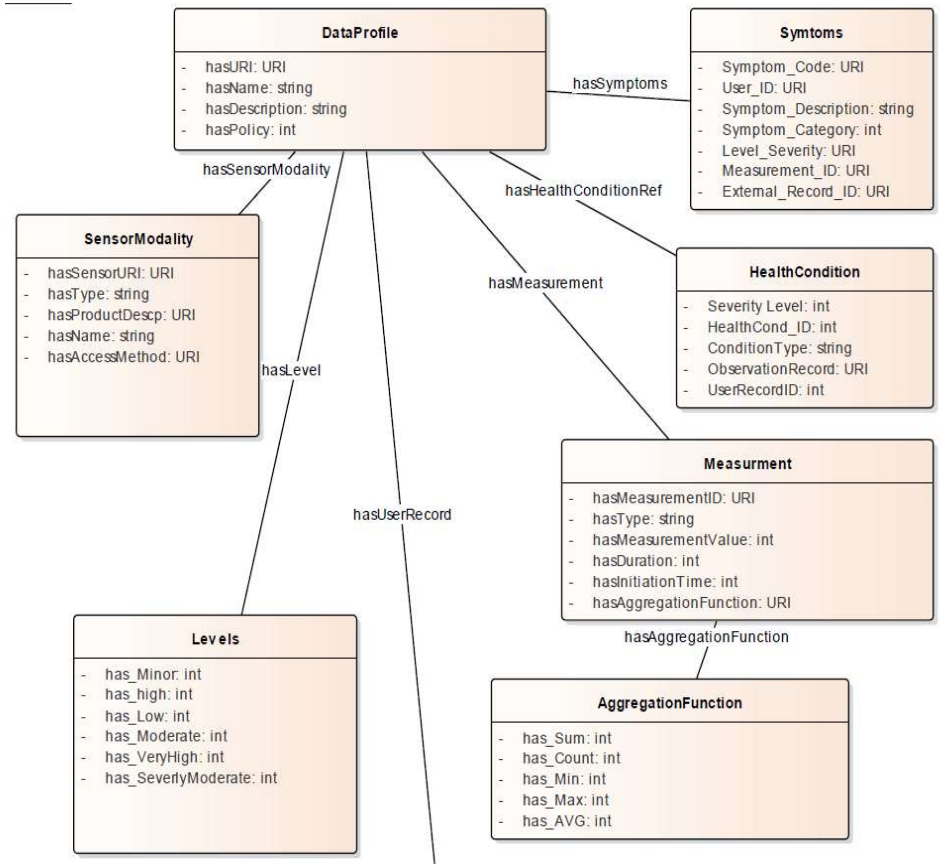


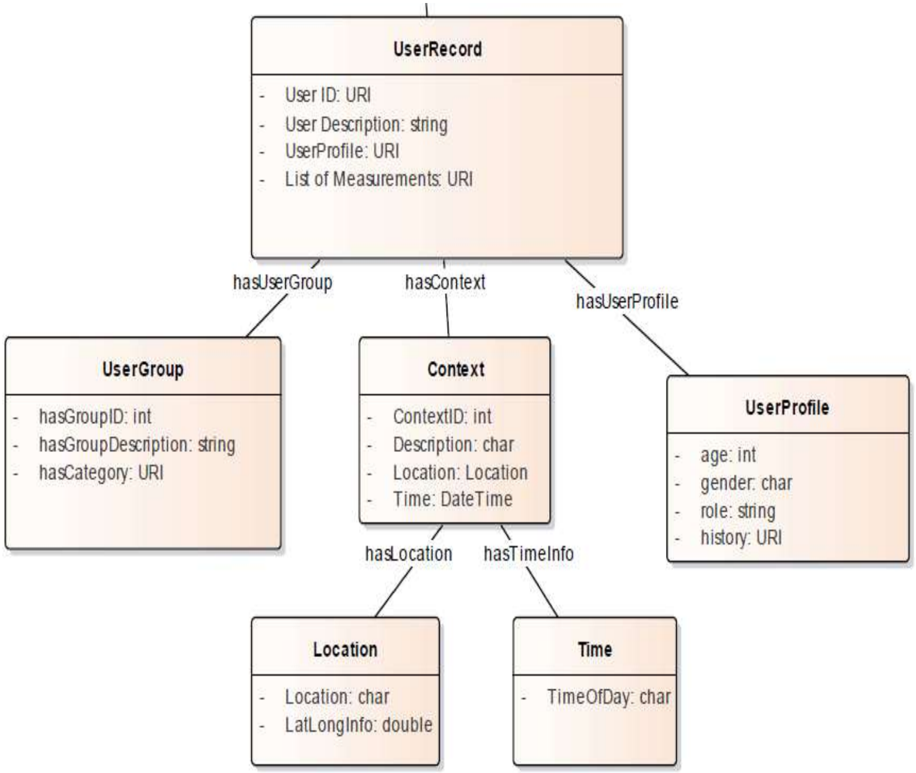

Accuracy results on ontologies dataset

# Results

## Data Model Features and Entities

- Model defines related data entities (Subset includes Data Profiles, Symptom, Context, …)
- Entities – Measurement, User, Sensor Modality, Severity Level, Location …



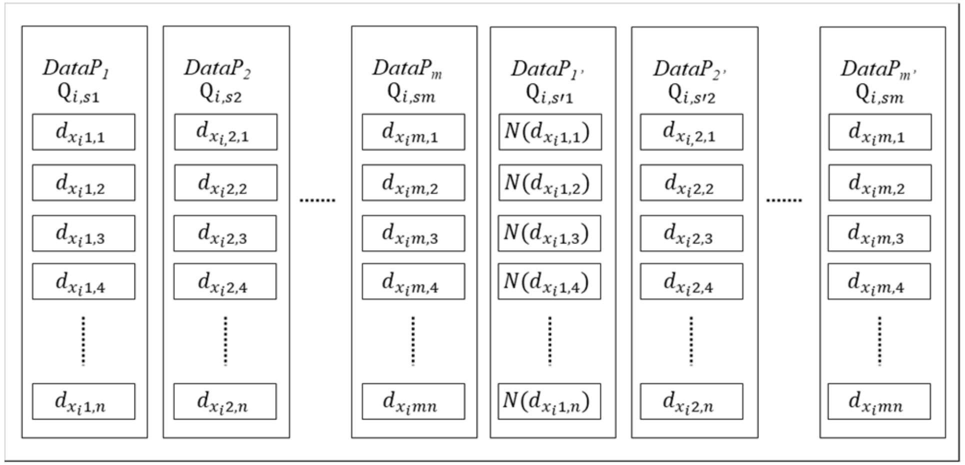
# Results

## Dataset Samples

- Data from diverse sources need to cope with the interoperable data model
- A mapping function has been defined
- Data sources contain set of diverse data

$$D = \{D_1, D_2, \dots \dots \dots, D_n\};$$

- Set of n samples from the interoperable data model to be provided to learning algorithm.




Data point in the interoperable data model

| | 1 | 2 | 3 | 4 | | 16 | 17 |
|---|---|---|---|---|---|---|---|
| 1 | 7.2580 | 17.2332 | -1.0822 | 1.2097 | | 8.5340 | 3.5703e+05 |
| 2 | 22.6929 | 17.5208 | 1.4111 | 7.4782 | | 10.7094 | 3.5695e+05 |
| 3 | 28.8802 | 8.5710 | 3.8632 | 14.7412 | | 12.2313 | 3.5694e+05 |
| 4 | 27.0729 | 3.9403 | -1.1982 | -3.9591 | ..... | 12.5192 | 3.5702e+05 |
| 5 | 22.1502 | 1.8372 | -5.3109 | -16.0571 | | 12.3148 | 3.5703e+05 |
| 6 | 14.4856 | 4.1697 | -0.4160 | 10.0973 | | 10.4600 | 3.5691e+05 |
| 7 | 10.3496 | 5.9229 | 4.3663 | 16.1543 | | 9.9888 | 3.5682e+05 |
| | | ⋮ | | | | ⋮ | |
| 12224 | -1.2032 | 2.5154 | 2.3336 | 3.2183 | ..... | 22.1964 | 4.0205e+05 |
| 12225 | 6.6603e-06 | 6.9112e-06 | -9.8753e-06 | -9.3206e-06 | | 17.8857 | 4.0211e+05 |

AMIGOS dataset [5] 12225x17 matrix

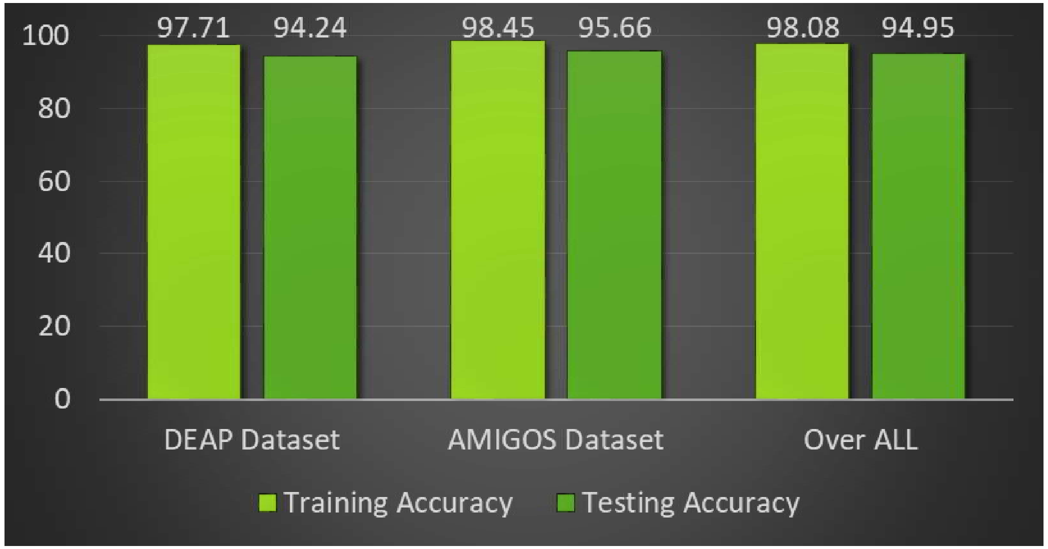
| Datasets / Features | DEAP dataset | AMIGOS dataset |
|---|---|---|
| Physiological Sensor Modalities | Electroencephalogram (EEG), GSR, Blood-Volume-Pressure (BVP), Temperature of the participant skin, Electromyogram (EMG). | Electrocardiogram (ECG), Electroencephalogram (EEG) readings, Galvanic Response (GSR) reading. |

DEAP dataset [4] and AMIGOS dataset [5] types of data

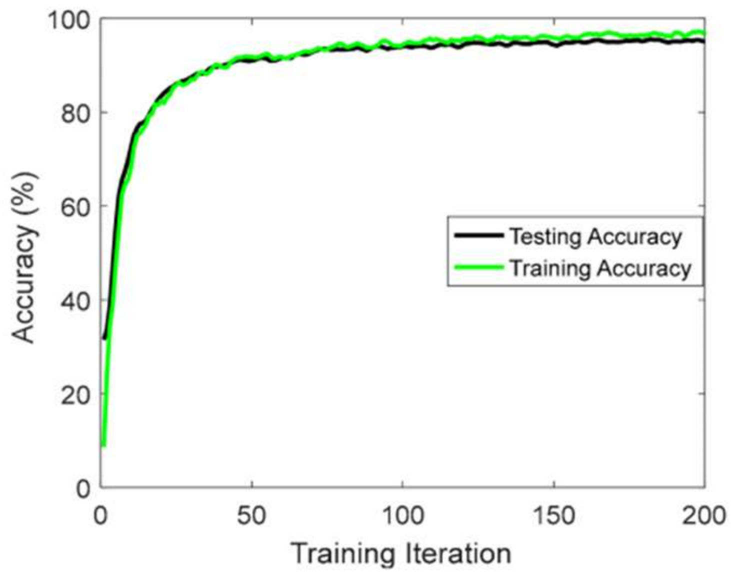
anLab

# Results

- The WoO based CDM is capable of data interoperability leveraging single classification model that can be used with diverse data sets;
- The proposed model can be utilized to enable sharing of data from diverse sources and to enhance learning accuracy;
- With CDM based interoperable data model better accuracy results have been achieved.




Achieved Accuracy (Training and Testing with respect to DEAP Dataset [4], AMIGOS Dataset [5] )




Overall Accuracy (Training and Testing dataset)

# Thank you