Enterprise Data Marketplaces

Establishing Data Marketplaces based on Semantic Technologies

Dr. Christian Mader

Enterprise Information Systems Department

Fraunhofer
IAIS

WHY DATA MARKETPLACES?

The Value of Data

Uber: the world's largest taxy company owns no vehicles

AirBnB: the world's largest accommodation provider owns no real estate

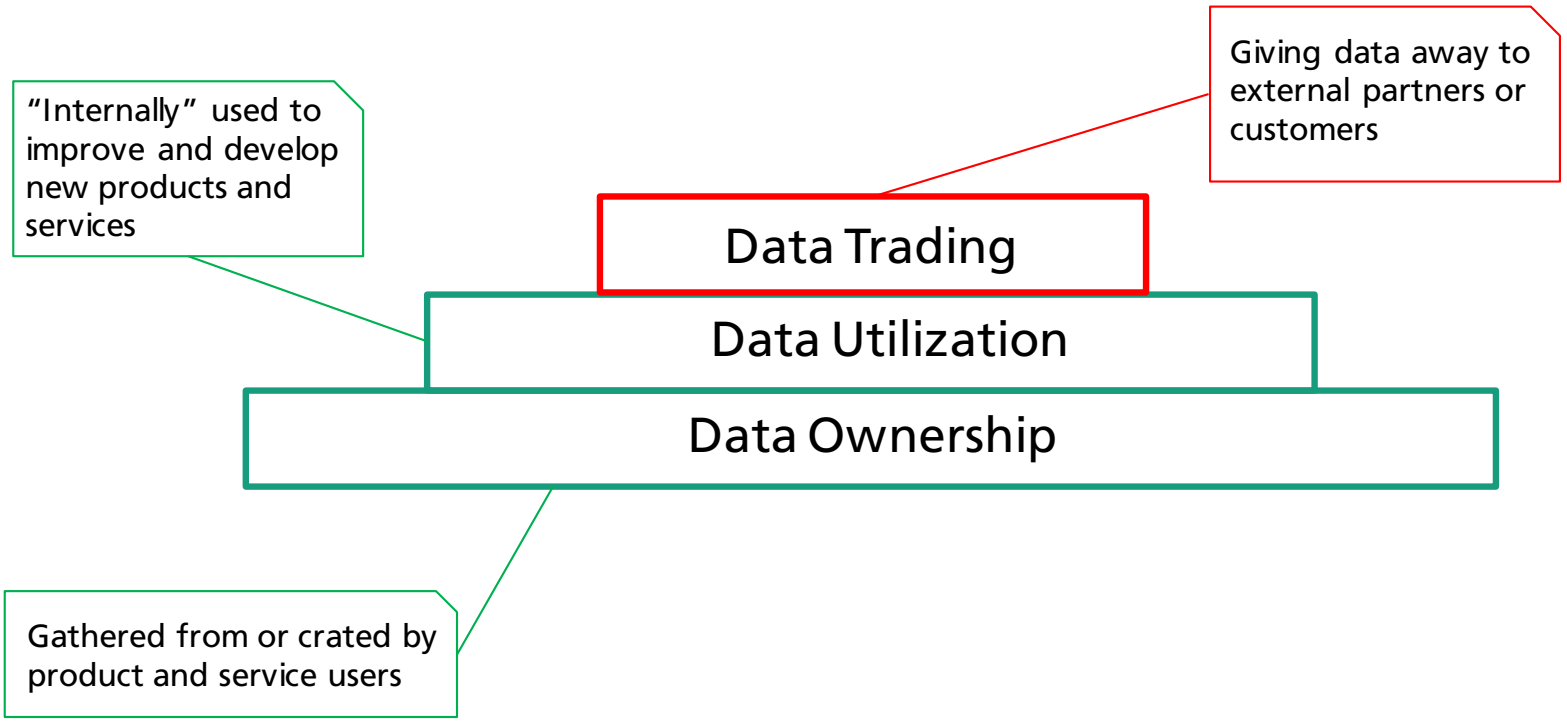
Alibaba: the most valuable retailer has no inventory

Facebook: the world's most popular media owner creates no content

Logos from the company websites

The Value of Data

Data becomes the **foundation** of enterprise business models



The Problem of Trading Enterprise Data

- Proprietary data exchange policies
  - Must be implemented by each business partner
- Restrictive compliance policies
  - No common set of supported tools
- Lack of data registries with wide industry support
  - Partners must know in advance what data exists and who it provides

The Problem of Trading Enterprise Data (2)

- Security Concerns
  - Ensure trust between parties
- Data Sovereignty
  - Make sure date is used in a specified way

The Vision of a Data Marketplace

1. Describe data you **have to provide**
   - What it **is about**
   - Define **who** and **how** to access it
     - Technical (protocols)
     - Organizational (access and usage control)
     - Economical (legal, pricing)
2. Describe data you **want to purchase**
3. Data Market takes care of the rest: **connects data providers and customers**

The Challenge

Define an **Information Model** to **describe** multiple aspects of data market places

- Data offers
  - Topic, usage policies, prices,…
- Data Services and Apps
  - Used algorithms, protocols,…
- Infrastructure
  - Security capabilities, operator,…
- Participants
  - Parties that exchange data, their roles, organization…

The Challenge

The model should be

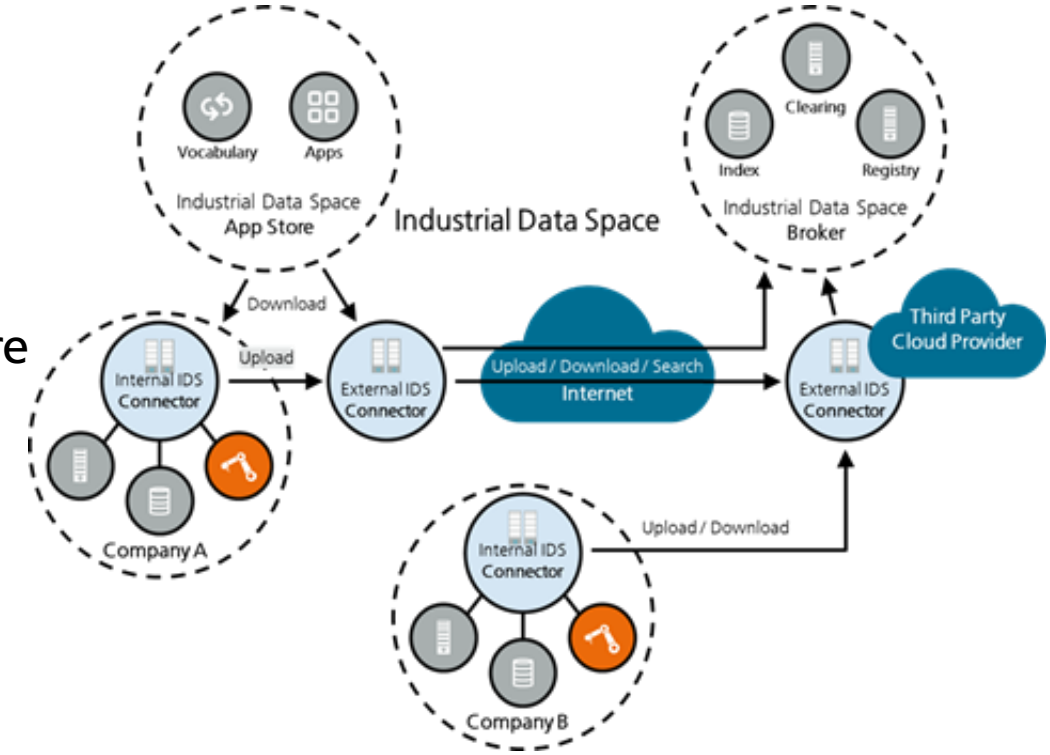
- Human readable
  - Easy formulation of queries
- Machine-processable
  - Drive search and retrieval use cases
- Technology agnostic
  - No vendor lock-in
- Extendable
  - Support future use cases

MARKETPLACE APPROACHES

Industrial Data Space (IDS)

INDUSTRIAL DATA SPACE

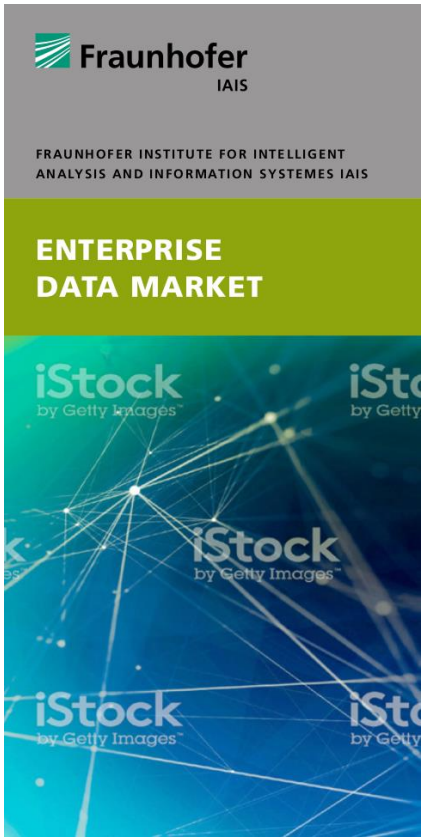
- BMBF-funded Project(s)
- Association with 80+ Industry Members
- Standardized Architecture
- Certification of Software and Participants
- Offers Software Components
- Implementation of (reference) Use Cases



Fraunhofer IAIS

SDIL Data Catalog

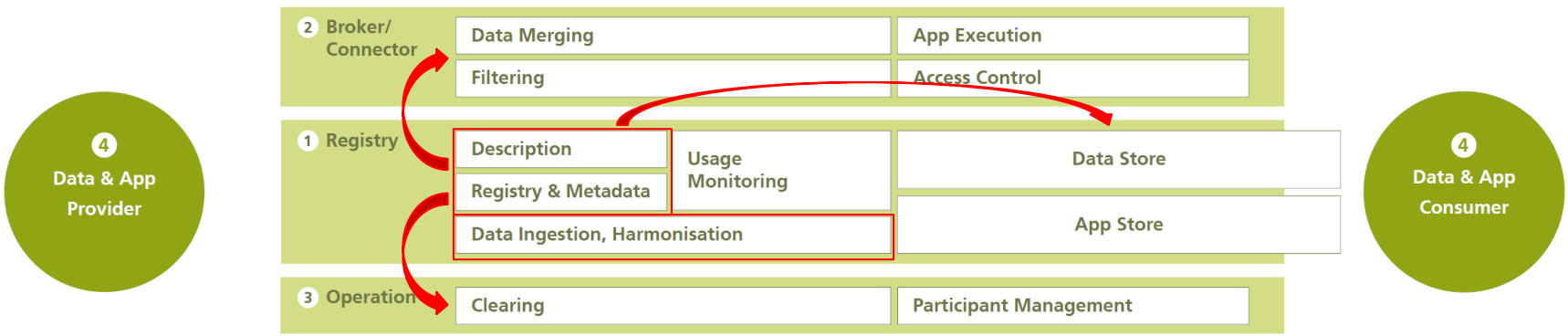
- Provided by Fraunhofer IAIS.EIS
- First Version currently rolled out
- Uses parts of the IDS Information Model to describe datasets
- Compatible with IDS Architecture
- In future act as an IDS Connector



SEMANTIC TECHNOLOGIES

What problems does it solve us?

- Common and standardized description of not only **Data** and **Apps** but all aspects of a Data Marketplace
- Allows **reuse** of existing information models
- Machine readability improves **search** and enables advanced **queries**
- Basis for **advanced** functionality (e.g., security, usage control)

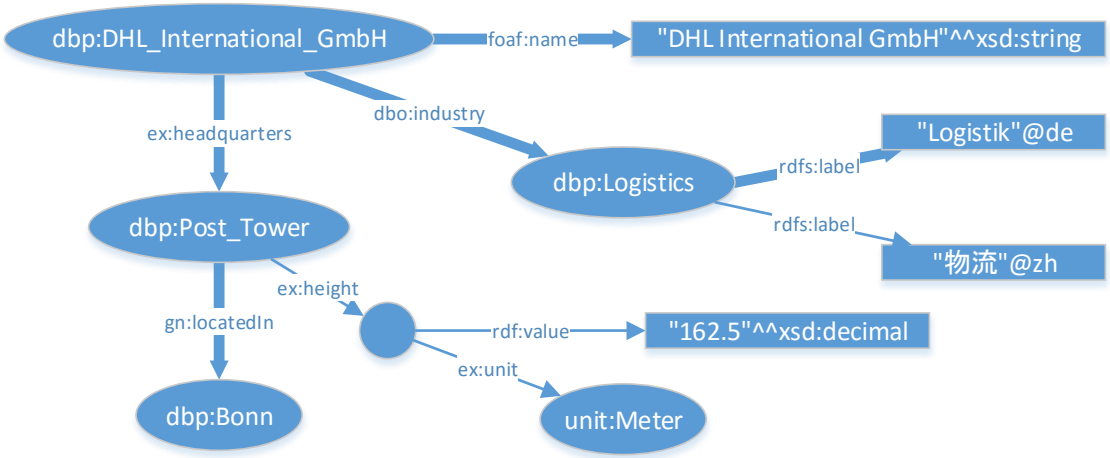


Fraunhofer
IAIS

Data Integration with Semantics

- Build a **knowledge graph** to describe things and their relations
- **Triple**-based form

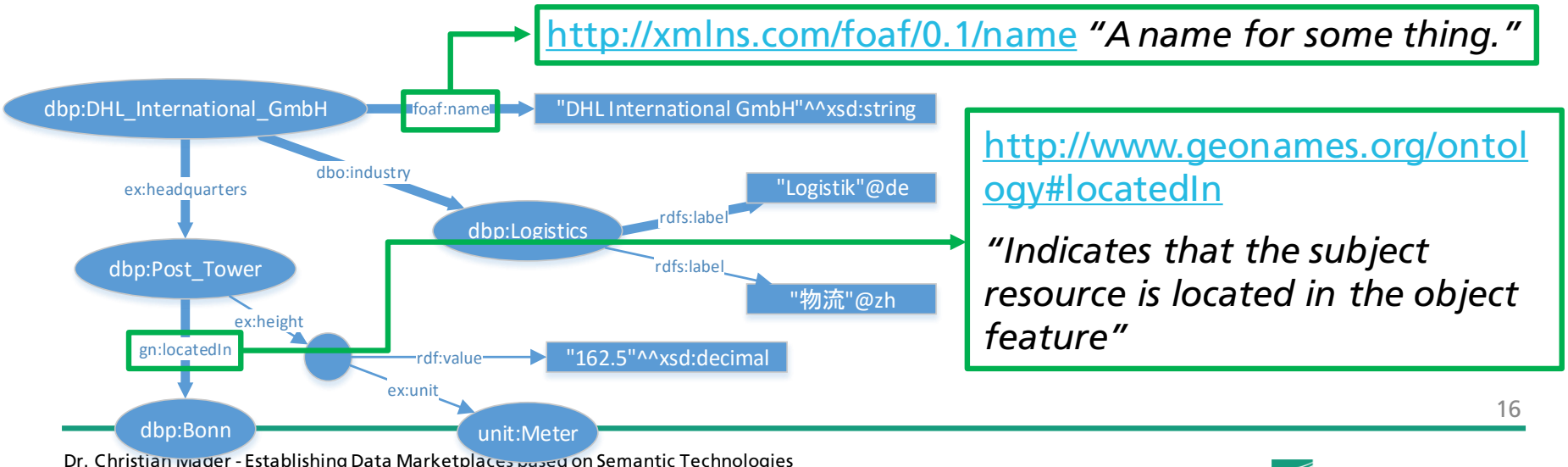
- All triple elements are **URLs**



Fraunhofer IAIS

Data Integration with Semantics (2)

- Triple elements can be **resolved**
- Can and should be **reused**
- Relations now have a **meaning** (semantics)
- We gain a **common understanding** of their meaning and the resources that are described

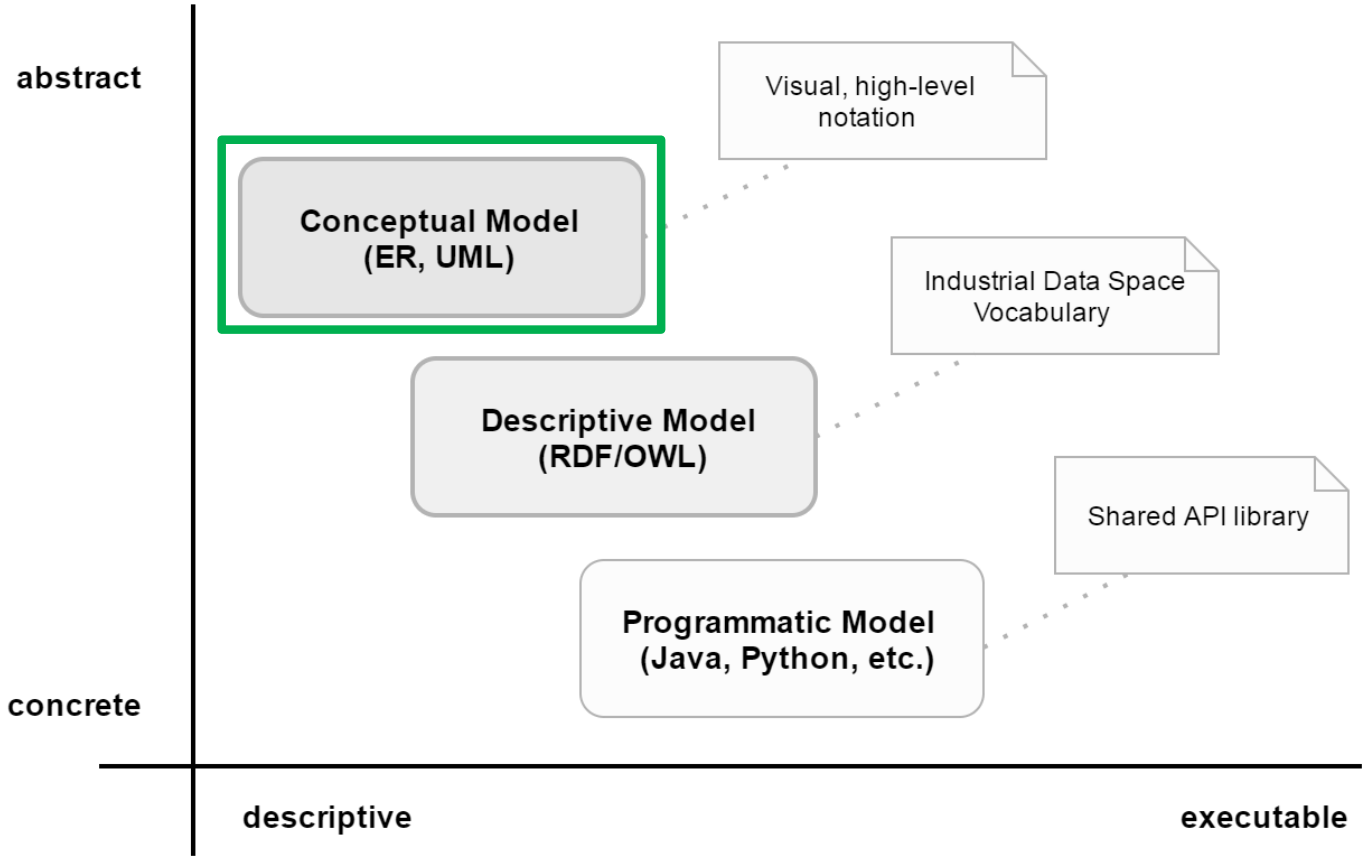


Data Integration with Semantics (3)

- **Shared understanding**
  - Meaning of data is clear to humans **and** machines
- **Existing models** and knowledge for various domains, e.g.,
  - People and their relations, Organizations
  - Business Processes
  - Policies
- Easy to **integrate and combine** knowledge graphs
  - Query multiple data sources "at once"

A SEMANTIC INFORMATION MODEL FOR THE IDS

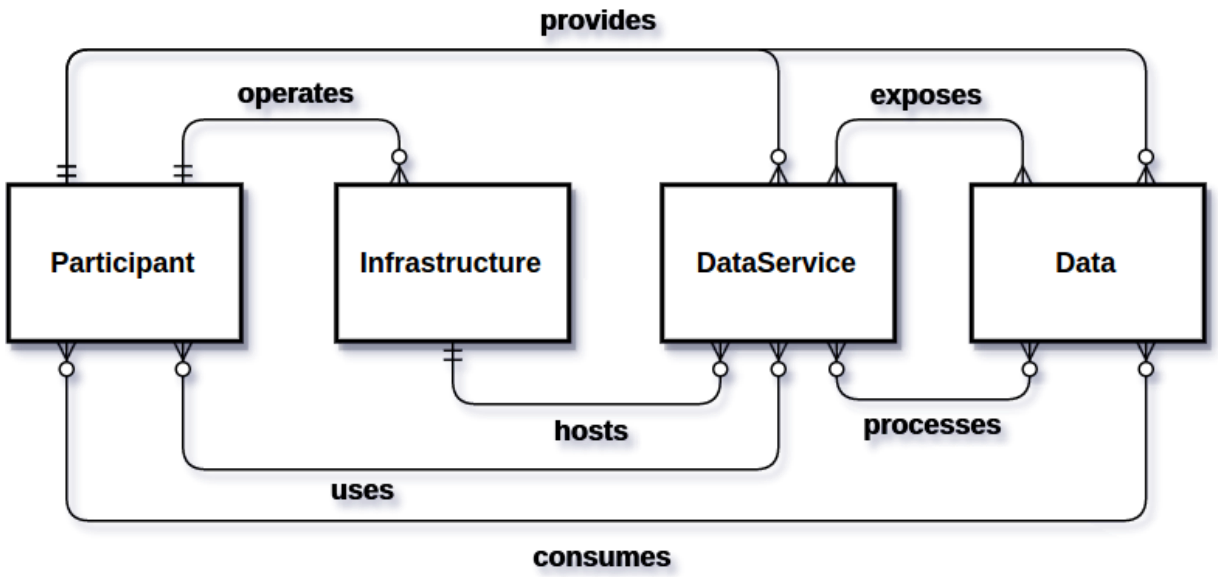
A Semantic Information Model for Data Marketplaces



Fraunhofer IAIS

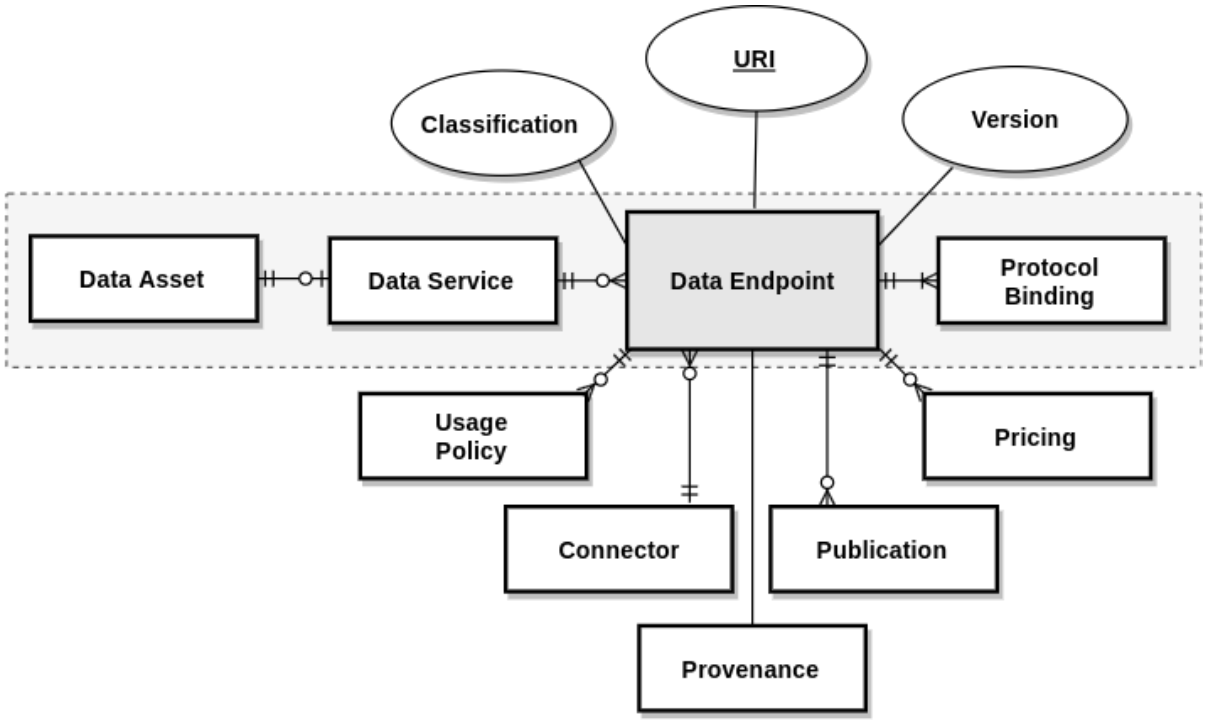
A Semantic Information Model for Data Marketplaces – Conceptual Level

- Information Model of the Industrial Data Space

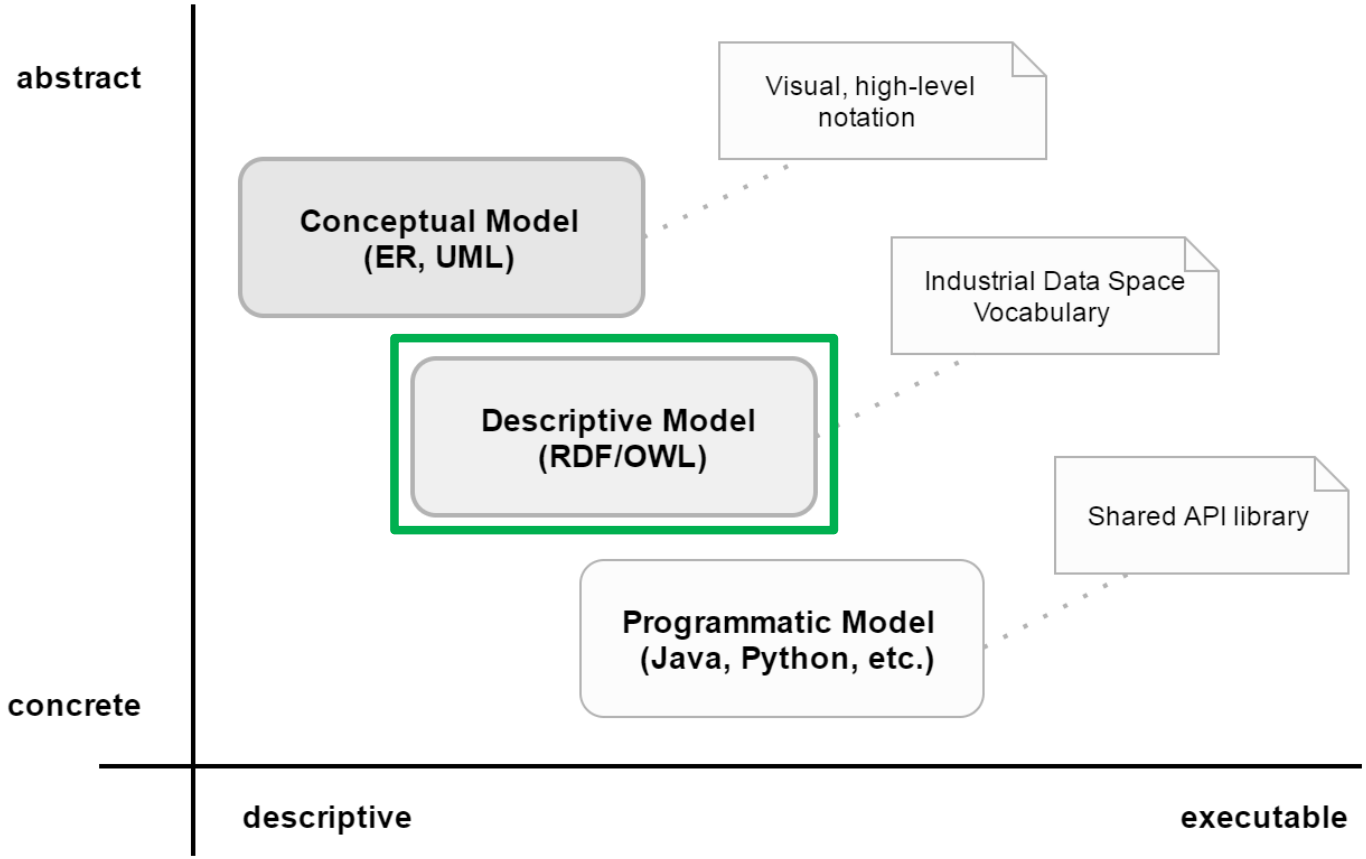


A Semantic Information Model for Data Marketplaces – Conceptual Level (2)

- Description of a Service that **offers** data



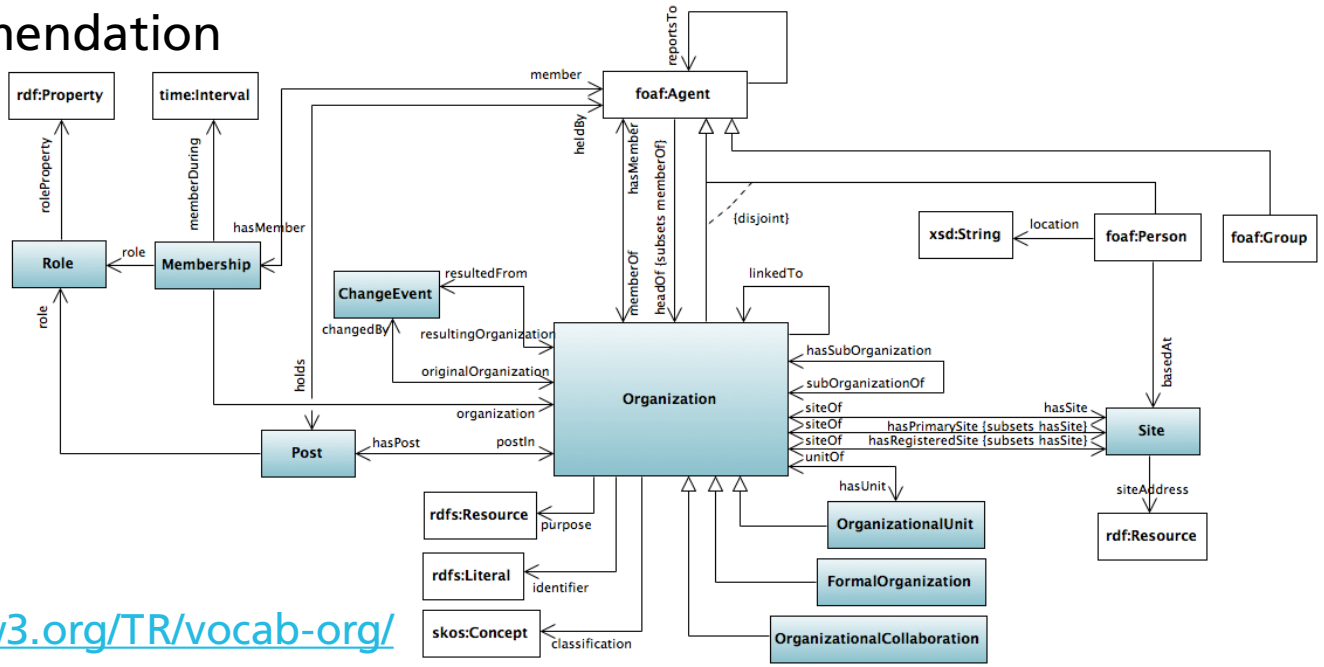
A Semantic Information Model for Data Marketplaces



Fraunhofer IAIS

A Semantic Information Model for Data Marketplaces – Descriptive Level

- The Organization Ontology (ORG)
  - A core ontology for organizational structures
  - W3C recommendation

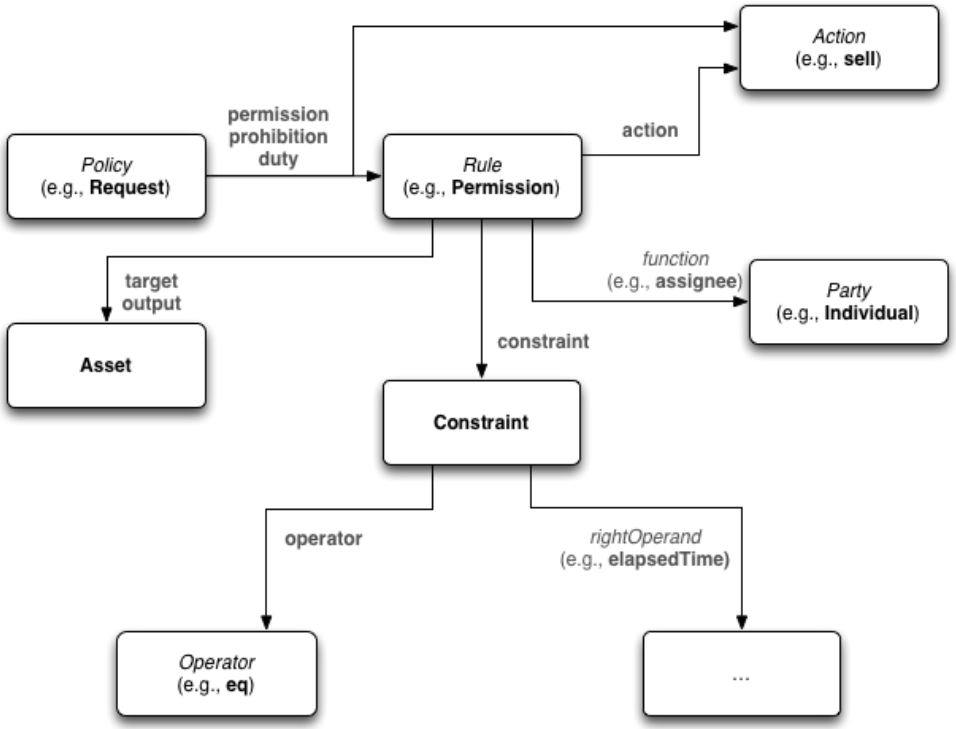


- https://www.w3.org/TR/vocab-org/

Fraunhofer IAIS

A Semantic Information Model for Data Marketplaces – Descriptive Level (2)

- Open Digital Rights Language (ODRL)
  - support traditional rights expressions for commercial transaction, open access expressions for publicly distributed content, and privacy expressions for social media
  - https://www.w3.org/ns/odrl/2/ODRL21



Fraunhofer IAIS

A Semantic Information Model for Data Marketplaces – Descriptive Level (3)

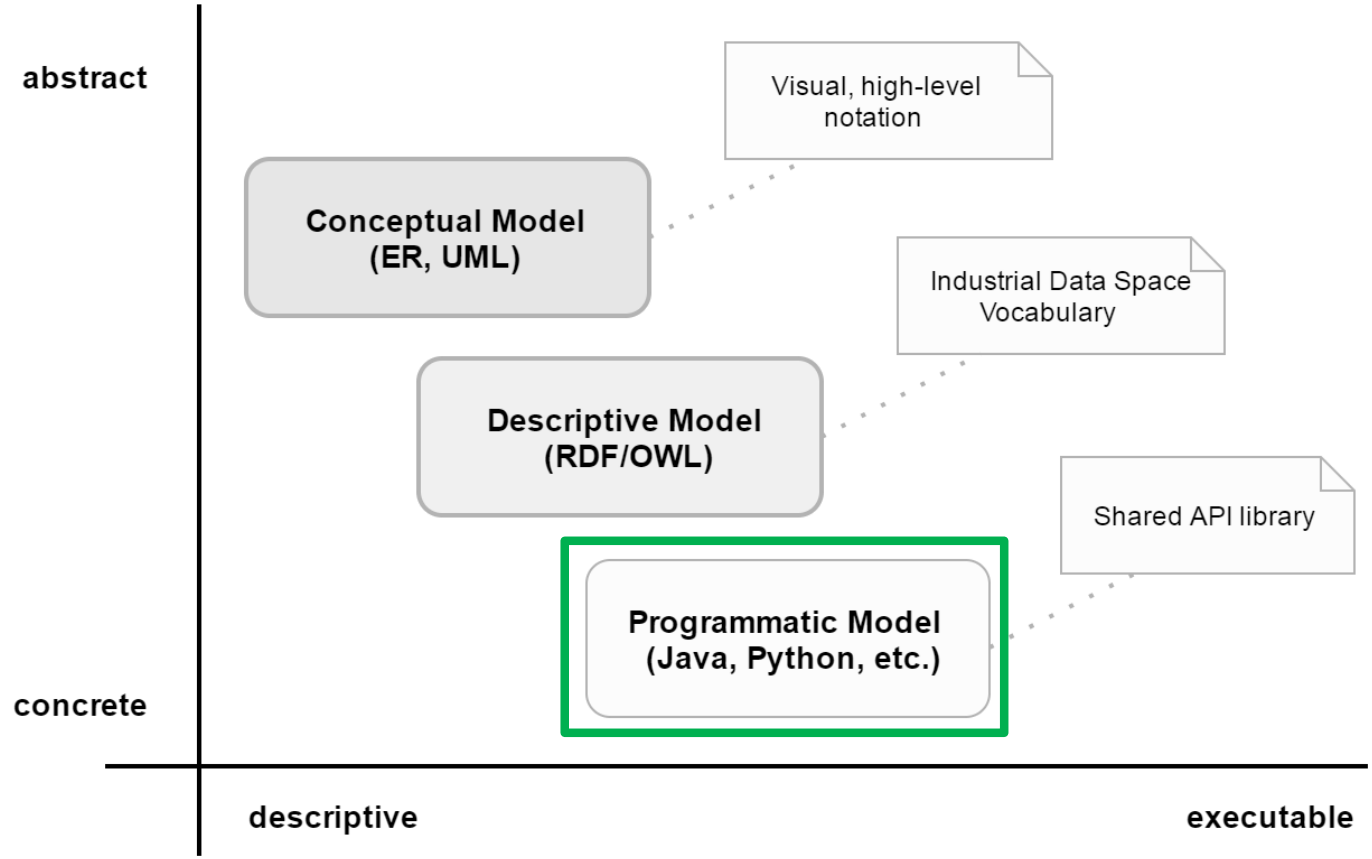
Classes

```
ids:Connector rdfs:subClassOf
ids:InfrastructureComponent;
    a owl:Class;
    rdfs:label "Connector"@en ;
    rdfs:comment "Connector for
hosting data services."@en;

    ids:validation [
        owl:onProperty ids:provides;
        valid:relationType
"@OneToMany";
    ].
```

Properties

```
ids:provides a owl:ObjectProperty;
    rdfs:domain ids:Connector;
    rdfs:range ids:DataEndpoint;
    rdfs:label "provides"@en;
    rdfs:comment "The DataEndpoints
provided by a Connector."@en.

ids:securityProfile a
owl:ObjectProperty;
    rdfs:domain ids:Connector;
    rdfs:range ids:SecurityProfile;
    rdfs:label "securityProfile"@en;
    rdfs:comment "The SecurityProfile
supported by the Connector."@en.
```

A Semantic Information Model for Data Marketplaces



Fraunhofer IAIS

A Semantic Information Model for Data Marketplaces – Programmatic Level

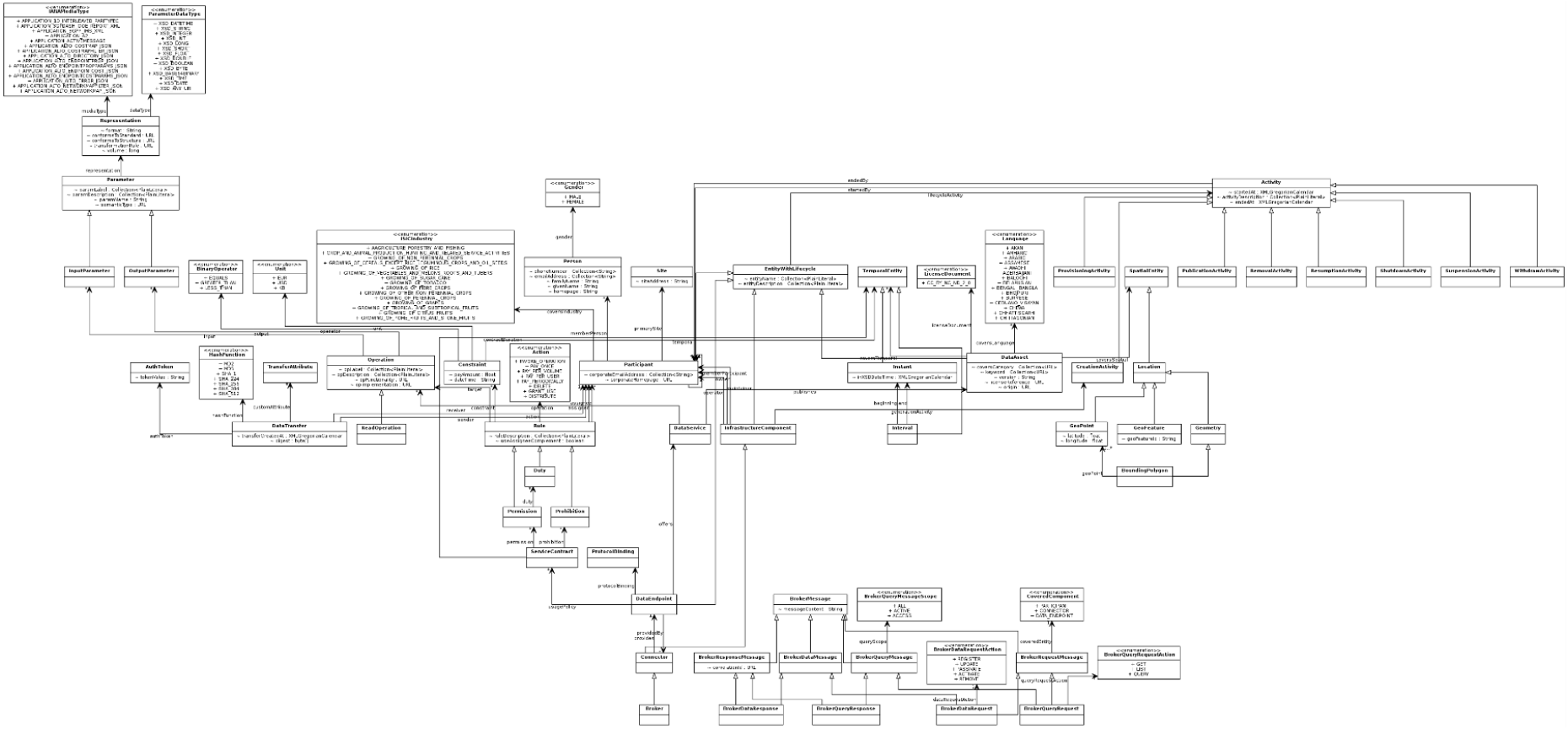
- Participants should describe their assets as Linked Data using the IDS Information Model, e.g.,
  - Connectors
  - Usage Policies
  - Services
  - Data Assets
  - Apps
  - ...

A Semantic Information Model for Data Marketplaces – Programmatic Level (2)

- Different ways of creating these descriptions:
  - **Directly:** using plain-text (editor, RDF syntax checker)
  - **Tool-supported:** e.g., Protege, TopBraid, VoCol
  - **Programmatically:** using RDF libraries (e.g., Jena, RDF4J)

- Valid solutions, **but**
  - require **knowledge of RDF** of any IDS participant
  - teaching RDF **does not scale** for a high number of participants

A Semantic Information Model for Data Marketplaces – Programmatic Level (3)



A Semantic Information Model for Data Marketplaces – Programmatic Level (4)

- **Automatic generation** from the descriptive model:
  - **Java Libraries**
    - Artifacts: sources (RDF + Java classes), binaries, documentation (javadoc, class diagram) and tests (+source for additional documentation of usage)
  - **Class diagram** of the model (previous slide)
  - **Other Documentation**

A Semantic Information Model for Data Marketplaces – Programmatic Level (5)

```
Connector connector = new ConnectorBuilder(new
    URL("http://example.org/companyA/connector"))
    .operator(participant.getId())
    .maintainer(participant.getId())
    .owner(participant.getId())
    .generationActivity(creationActivity)
    .lifecycleActivities(Arrays.asList(creationActivity))
    .entityNames(Arrays.asList(new
PlainLiteral(„Connector")))
    .entityDescriptions(Arrays.asList((new
PlainLiteral(„IDS Connector for classified business
data.", "en"))))
    .build();
```

```
connector.toRdf()
```

A Semantic Information Model for Data Marketplaces – Programmatic Level (6)

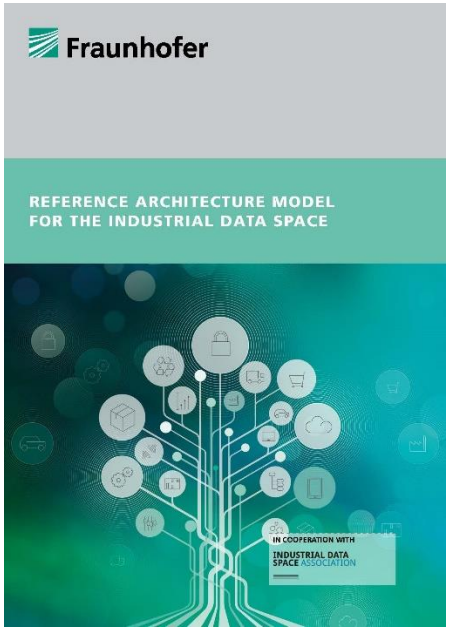
```
http://companyA.com/IDS/connector1 a ids:Connector ;
  ids:entityName [
    a skos:Collection ;
    skos:member "Official IDS Connector of companyA"@en] ;
  ids:operator
<http://industrialdataspace.org/participants/companyA> ;
  ids:owner
<http://industrialdataspace.org/participants/companyA> ;
  ids:provides [ a
    skos:Collection;
    skos:member
<http://industrialdataspace.org/connector1/endpoint1>
  ];
```

...

MODEL APPLICATIONS

Industrial Data Space (IDS)

- Installable components available for association members
  - Connector, Broker, App Store
- Central element: **IDS Information Model**
  - Reference Architecture Document
  - Under Apache 2.0 license
  - Developed on GitHub
  - https://github.com/IndustrialDataSpace



Fraunhofer IAIS

Conclusions

- Benefits of a Semantic Data Model
  - Builds a **common** human-readable and machine-processable **understanding** of a Data Marketplace
    - Data and Apps
    - Infrastructure and Participants
  - Easy **adoption** of existing models
  - **Extensible** for future requirements
  - **Usable** also by non-RDF experts

Thank You!

- Dr. Christian Mader
  - christian.mader@iais.fraunhofer.de
- Fraunhofer IAIS.EIS
  - https://www.iais.fraunhofer.de/de/institut/abteilungen/enterprise-information-systems.html
- Industrial Data Space
  - http://www.industrialdataspace.org/