# Enabling System Artifacts Reuse Through The Semantic Representation of Engineering Models: a Case Study of Simulink Models

Roy Mendieta[1], Eduardo Cibrián[2], Jose María Álvarez-Rodriguez[2], Juan Llorens[2]

roy.mendieta@reusecompany.com

[1]The Reuse Company, Madrid. Spain. [2]Computer Science and Engineering Department, Carlos III University of Madrid, Spain.

**Roy Mendieta** is System Engineer at The Reuse Company and a student of PhD Program of Computer Science and Technology at Carlos III University of Madrid, before these roles, Roy was IT Project Manager at Central Bank of Honduras where he was responsible for the implementation of systematization projects of economic statistics generation processes, using as a base, international methodologies for the generation of economic statistics.

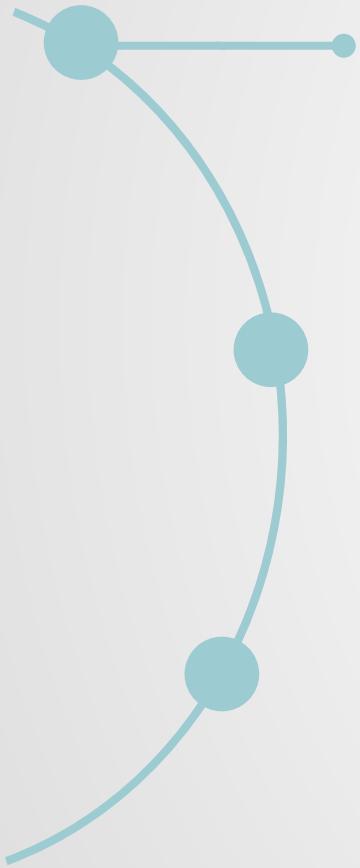
Roy holds a master's degree in Computer Science and Technology at Carlos III University of Madrid and a master's degree in Economics and Finance at Catholic University of Honduras, he also has a degree in Computer Science Engineer at Catholic University of Honduras.

Additionally, Roy has a Project Management Professional Certification (PMP) by Project Management Institute (PMI). His main lines of research are knowledge reuse and interoperability.

# Agenda

1. Introduction
2. Objectives
3. State of Art
4. Proposed solution
5. Case Study
6. Analysis of Results
7. Conclusions

# Introduction



**Emergence of Model-based Systems Engineering (MBSE)**

MBSE to support the whole specification process of a system (conceptual design, system requirements, design, analysis, verification or validation, etc.)

# Introduction

**Emergence of Model-based Systems Engineering (MBSE)**

MBSE to support the whole specification process of a system (conceptual design, system requirements, design, analysis, verification or validation, etc.)

**Necessity of reuse and understand models**

To reuse the knowledge generated methodologies (MDE), such as MBSE, it is necessary to understand the underlying concepts and relationships that allows us to make a semantic interpretation of the models.

# Introduction

**Emergence of Model-based Systems Engineering (MBSE)**

MBSE to support the whole specification process of a system (conceptual design, system requirements, design, analysis, verification or validation, etc.)

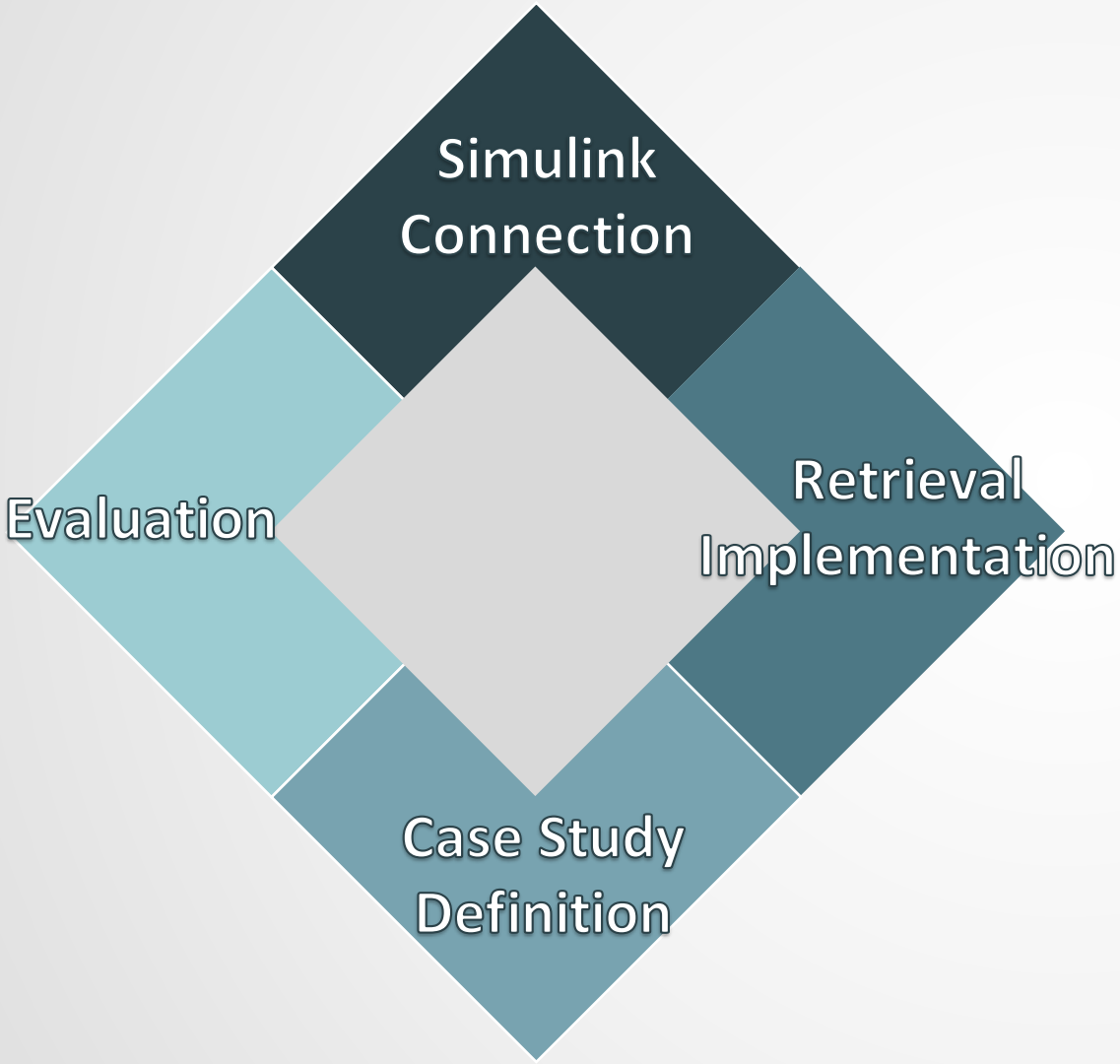
**Necessity of reuse and understand models**

To reuse the knowledge generated methodologies (MDE), such as MBSE, it is necessary to understand the underlying concepts and relationships that allows us to make a semantic interpretation of the models.

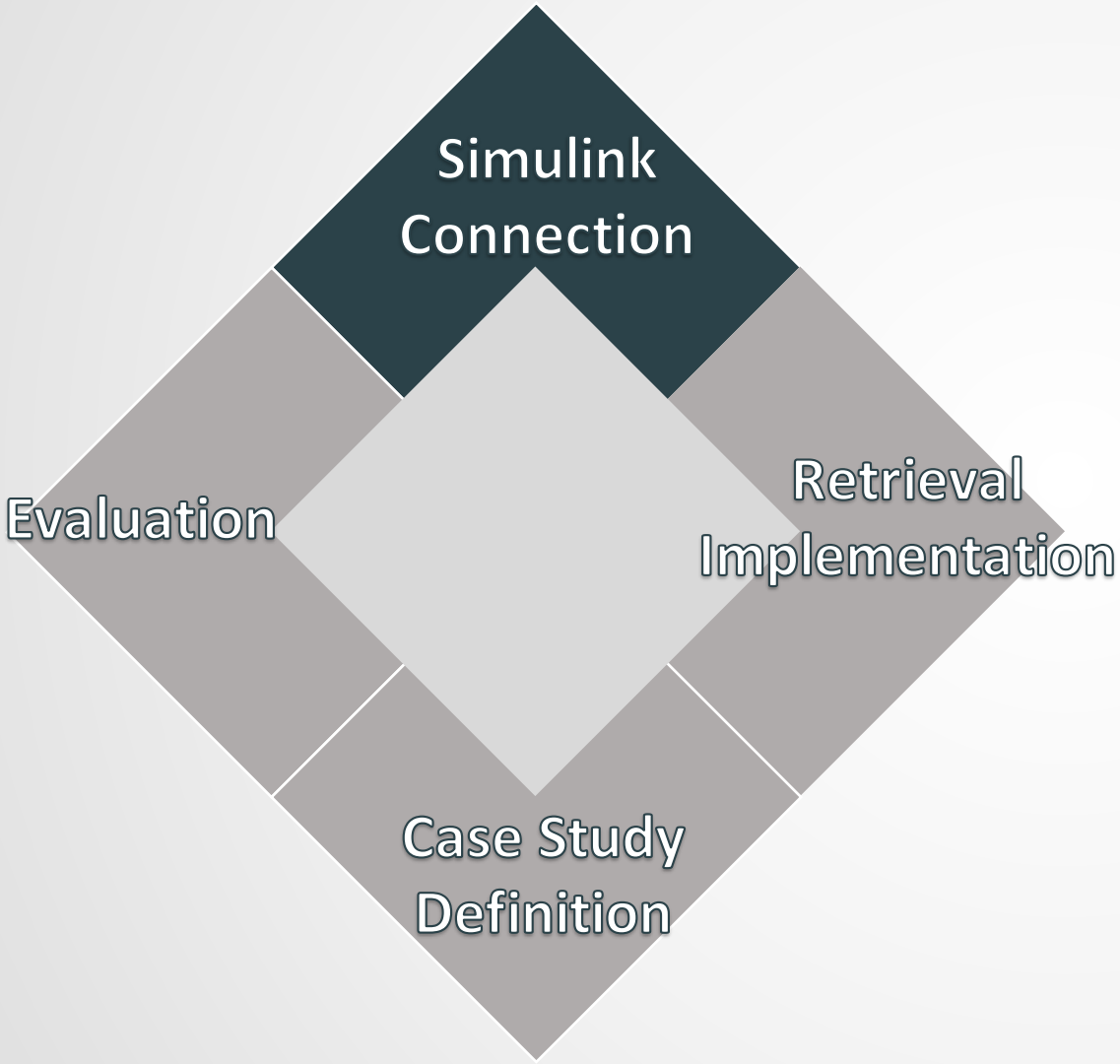
**Solution: RSHP2SIMULINK**

The solution makes use of an ontology-based approach for indexing and retrieving information following a meta-model, Information Representation Model Based on Relationships (RSHP), to reuse Simulink models

# Objectives


Simulink Connection
Retrieval Implementation
Case Study Definition
Evaluation

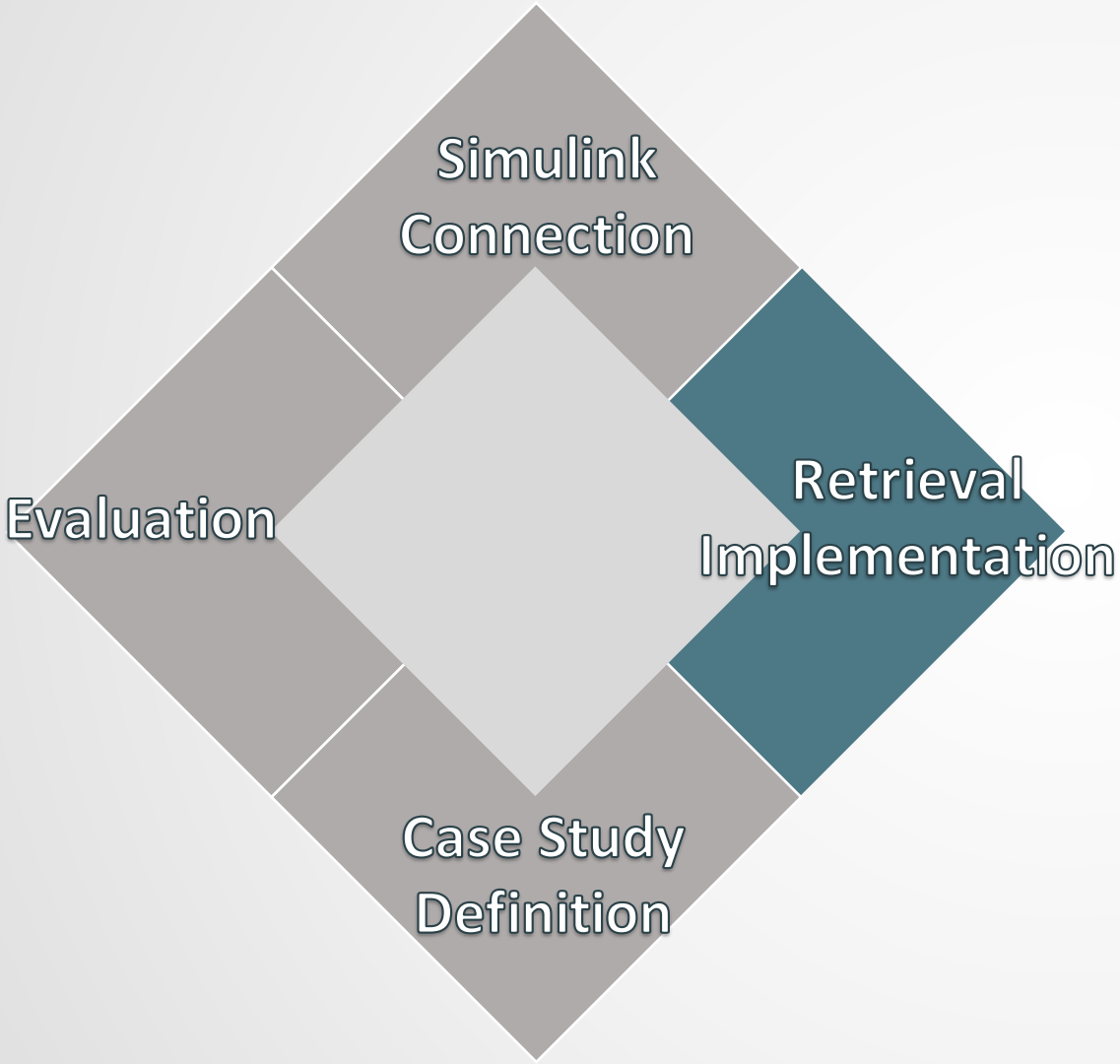
# Objectives




**Simulink Connection**

Design and implement a mapping between the Matlab Simulink meta-model and the RSHP meta-model is defined to represent and serialize analytical models in a repository.

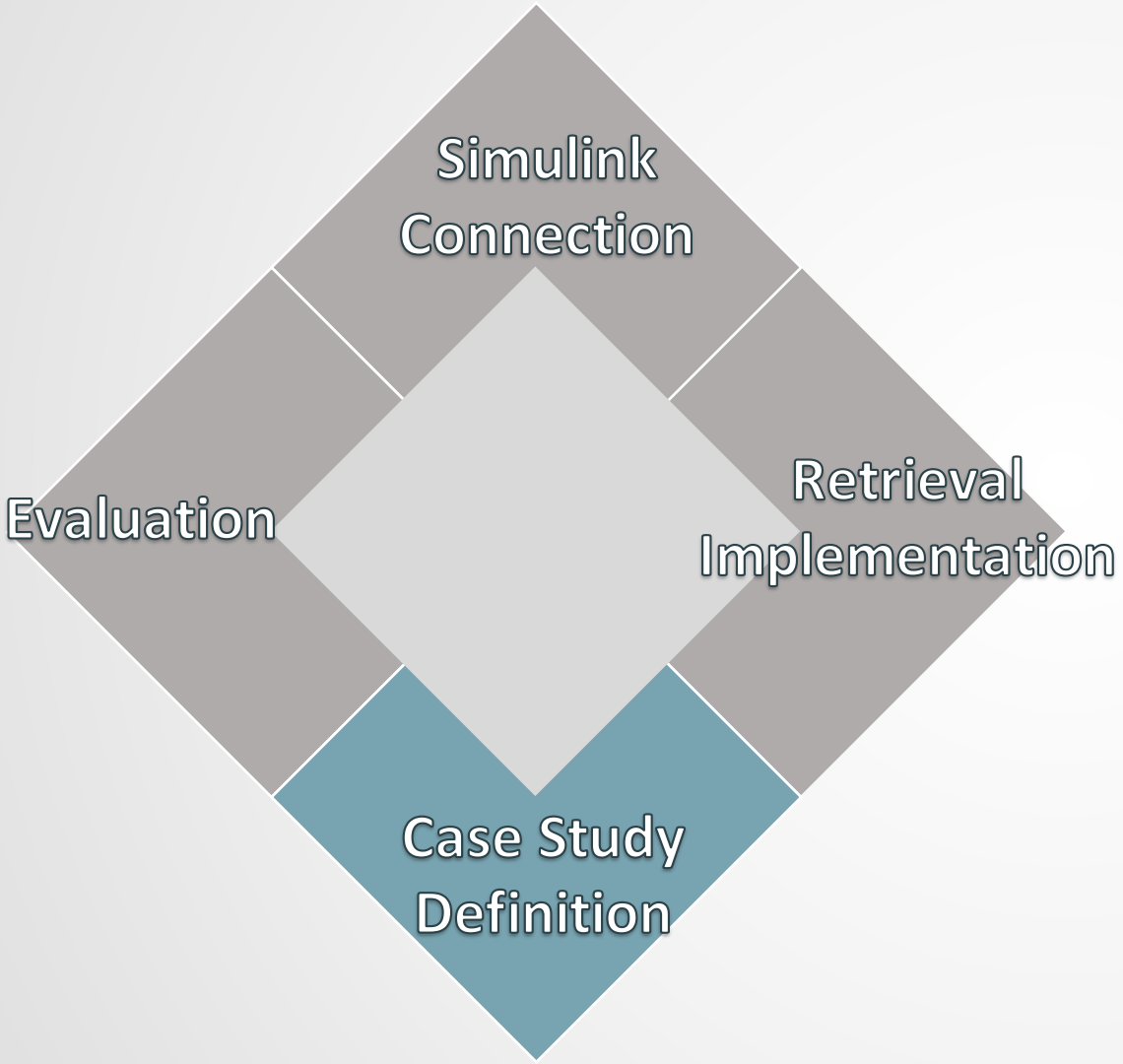
# Objectives




**Retrieval Implementation**

A retrieval process is implemented on top of this repository to allow users to perform text-based queries and look up similar artifacts.

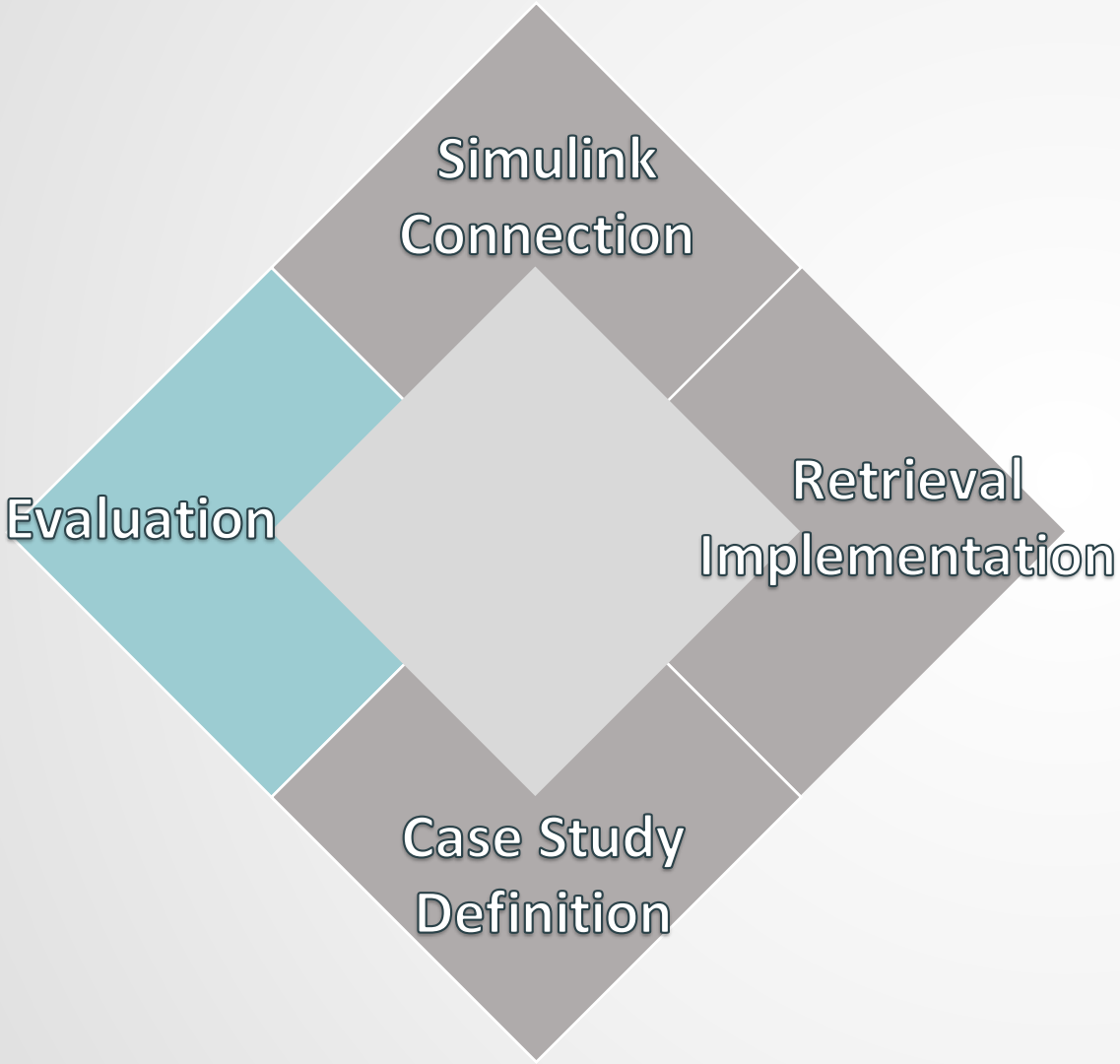
# Objectives




**Case Study Definition**

38 Simulink models have been indexed and 20 real user queries have been designed to retrieve models

# Objectives




**Evaluation**

Study the effectiveness od the proposed solution, in terms or precision and recall metrics.

# State of Art

## MODEL REUSE

- An effective retrieval approach of 3D CAD models for macro process reuse [1]
- Towards an ontology-based retrieval of UML class diagramsPlaceholder text [2]

## SIMULINK MODEL REUSE

## RSHP APPLICATIONS

# State of Art

## MODEL REUSE

- An effective retrieval approach of 3D CAD models for macro process reuse [1]
- Towards an ontology-based retrieval of UML class diagramsPlaceholder text [2]

## SIMULINK MODEL REUSE

- Structuring Simulink models for verification and reuse [3]
- Submodel pattern extraction for Simulink models [4]

## RSHP APPLICATIONS

# State of Art

## MODEL REUSE

- An effective retrieval approach of 3D CAD models for macro process reuse [1]
- Towards an ontology-based retrieval of UML class diagramsPlaceholder text [2]

## SIMULINK MODEL REUSE

- Structuring Simulink models for verification and reuse [3]
- Submodel pattern extraction for Simulink models [4]

## RSHP APPLICATIONS

- Reuse of physical system models by means of semantic knowledge representation: A case study applied to Modelica [5]
- Towards effective SysML model reuse [6]

# State of Art - Conclusions

→Reuse based on specific features of the domain knowledge or text comparison

# State of Art - Conclusions

→Reuse based on specific features of the domain knowledge or text comparison

→The proposed solution aims to improve the reuse of the embedded information in the Simulink models by providing:

1) A semantic representation of Simulink models using an existing meta-model like RSHP
2) A retrieval process based on comparing the underlying graphs of a query against a repository of Simulink models

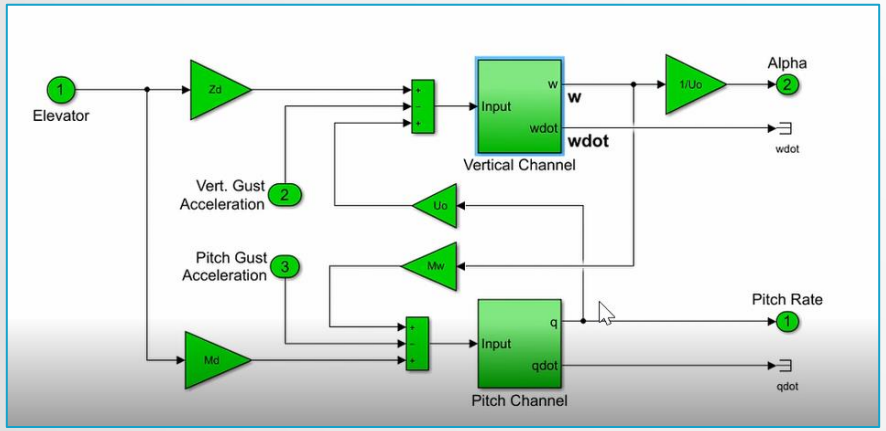
# Approach

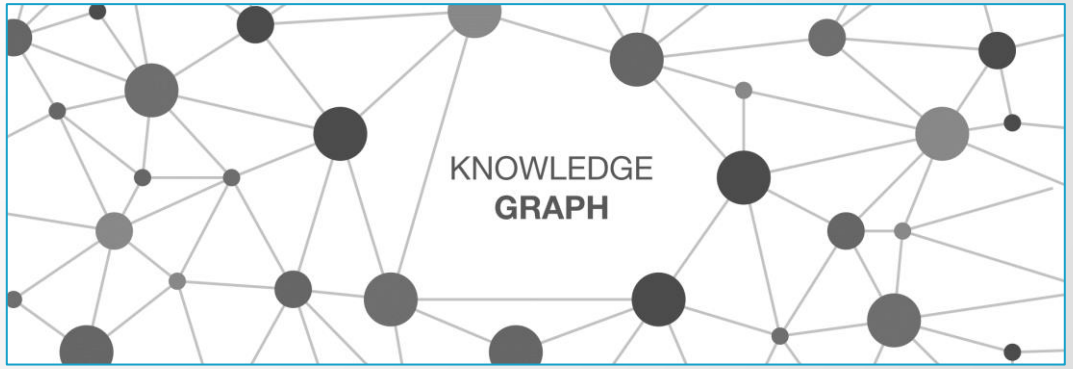
- Reuse semantic information needs to deal with a lot of factors that have to be considered in reuse techniques.

# Approach

- Reuse semantic information needs to deal with a lot of factors that have to be considered in reuse techniques.

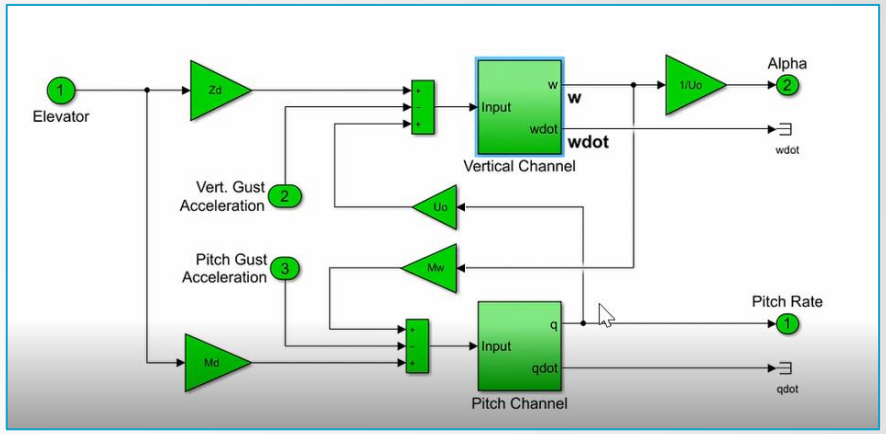
- One of the main objectives is to identify, classify, organize and represent Simulink models using semantics:
  - This solution applies a domain ontology to model, such information and build a retrieval information process based not only on calculating the similarity of the underlying graphs between two artifacts



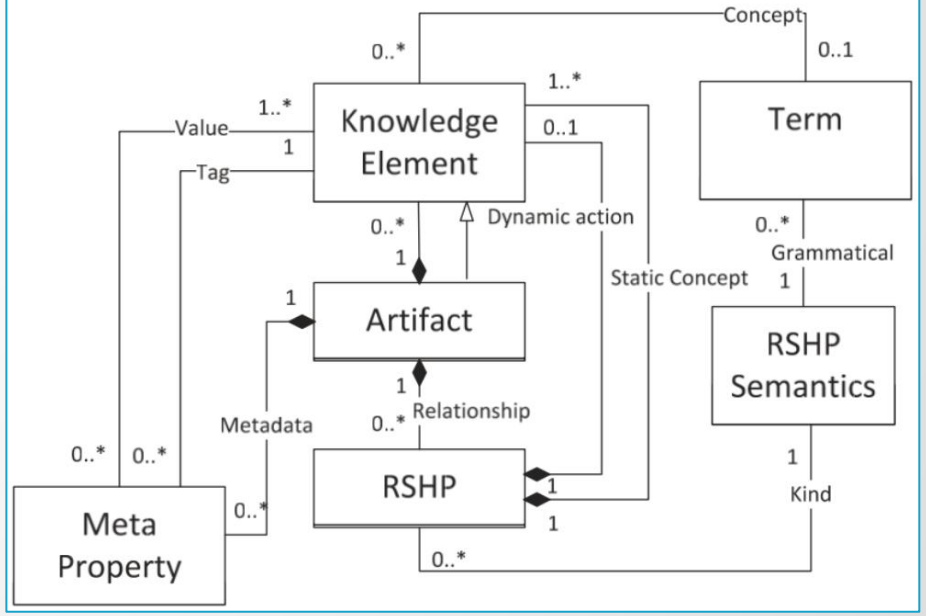



# Representation process

| Simulink Element | RSHP Element |
|---|---|
| Model | Artifact |
| Block | Artifact |
| Block Type | Artifact Type |
| Block Name | Artifact Name, Noun Term |
| Block Properties | Metaproperties |
| Line | RSHP |




Metamodel of RSHP representation



# Proposed solution

- An application developed in Visual Studio .Net 2019 with framework 4.8

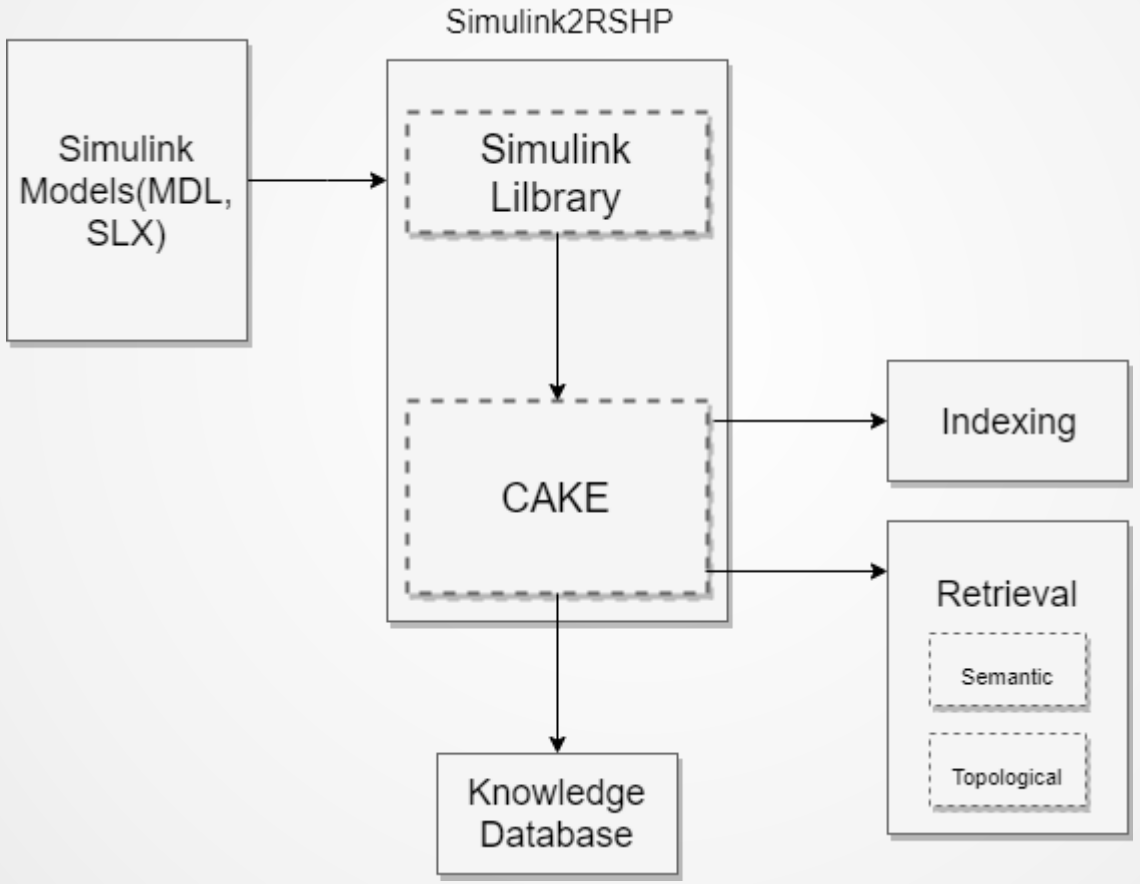
# Proposed solution

- An application developed in Visual Studio .Net 2019 with framework 4.8
- The application allows to parse Simulink files using a Simulink software library for Java and IKVM creating a semantic representation of the Simulink models using the CAKE-RSHP model

# Proposed solution

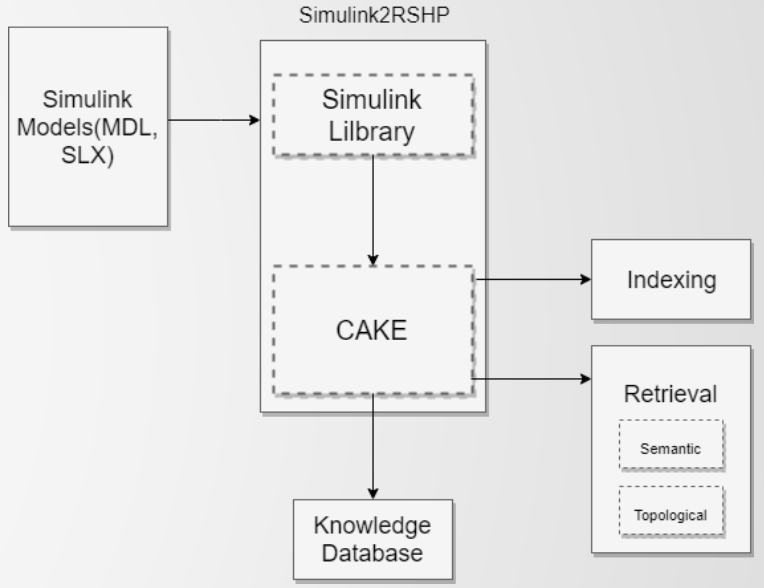
- An application developed in Visual Studio .Net 2019 with framework 4.8
- The application allows to parse Simulink files using a Simulink software library for Java and IKVM creating a semantic representation of the Simulink models using the CAKE-RSHP model
- Using CAKE-RSHP framework, it is possible to use the built-in mechanisms already available for indexing and retrieving information.

# Proposed solution - Architecture



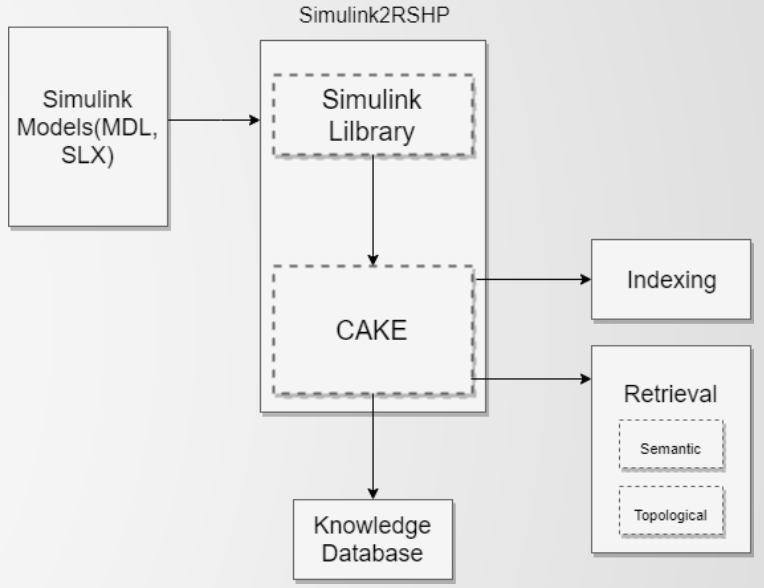
# Proposed solution - Components

- **Simulink2RSHP:** This component groups Simulink Library and CAKE, basically it allows us to semi-automatically apply the mappings between the Simulink elements and the RSHP meta-model creating an underlying semantic graph based on the domain ontology.



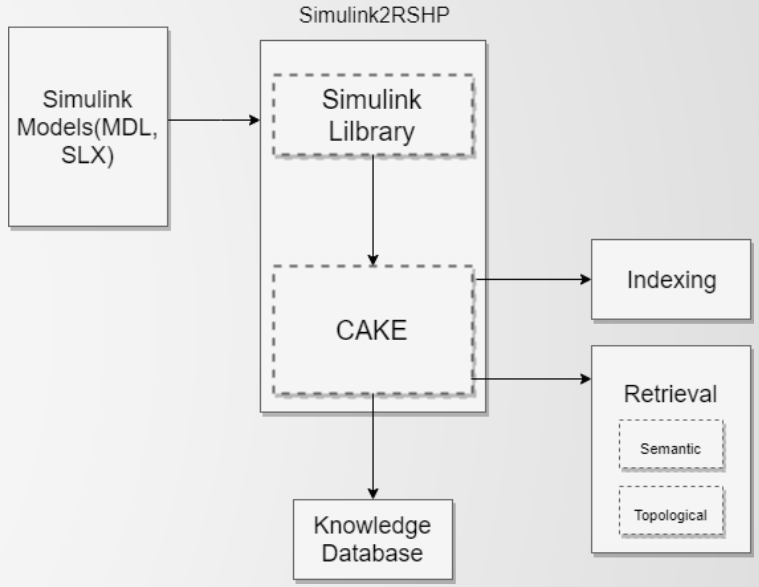
# Proposed solution - Components

- **Simulink2RSHP:** This component groups Simulink Library and CAKE, basically it allows us to semi-automatically apply the mappings between the Simulink elements and the RSHP meta-model creating an underlying semantic graph based on the domain ontology.
- **Simulink Library:** This component allows mapping the objects that are obtained from invoking the reading processes of the Simulink library. Once the information is obtained from the files, it is represented using the CAKE API.



# Proposed solution - Components

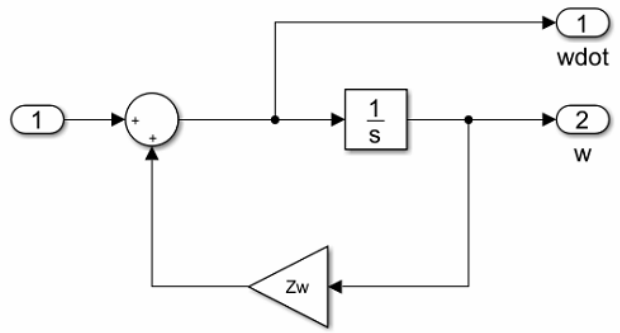
- **Simulink2RSHP:** This component groups Simulink Library and CAKE, basically it allows us to semi-automatically apply the mappings between the Simulink elements and the RSHP meta-model creating an underlying semantic graph based on the domain ontology.
- **Simulink Library:** This component allows mapping the objects that are obtained from invoking the reading processes of the Simulink library. Once the information is obtained from the files, it is represented using the CAKE API.
- **CAKE.** Once the information is represented in the RSHP language, it is possible to use the built-in capabilities for information retrieval and indexing. The CAKE API internally implements a pattern matching algorithm between graphs that returns a value of similarity.



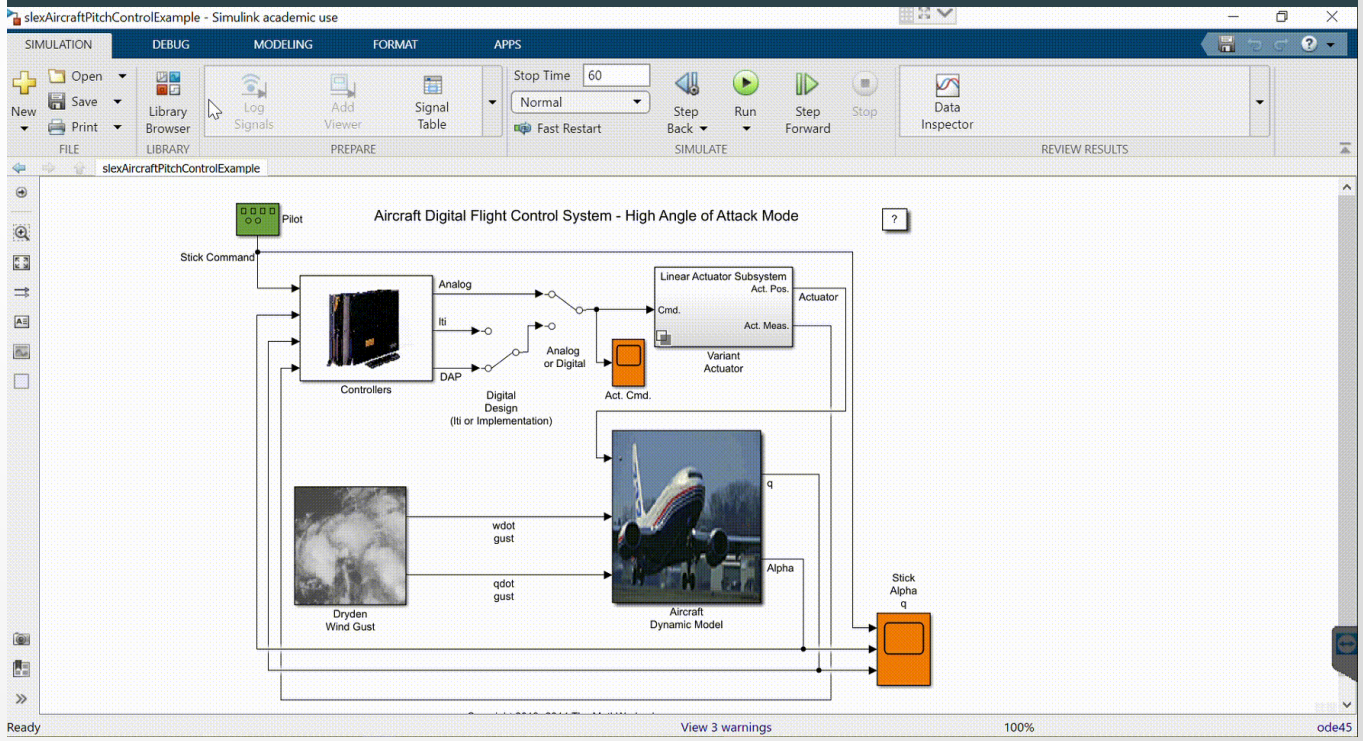
# Case Study - Motivation

Where is the Vertical channel located in an Aircraft dynamic model?

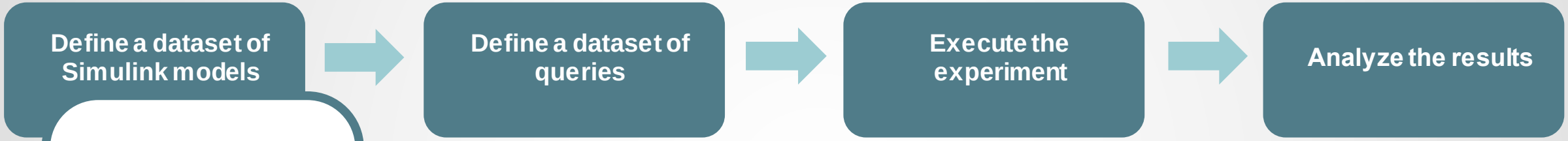
## Query Format Example




## Model Example



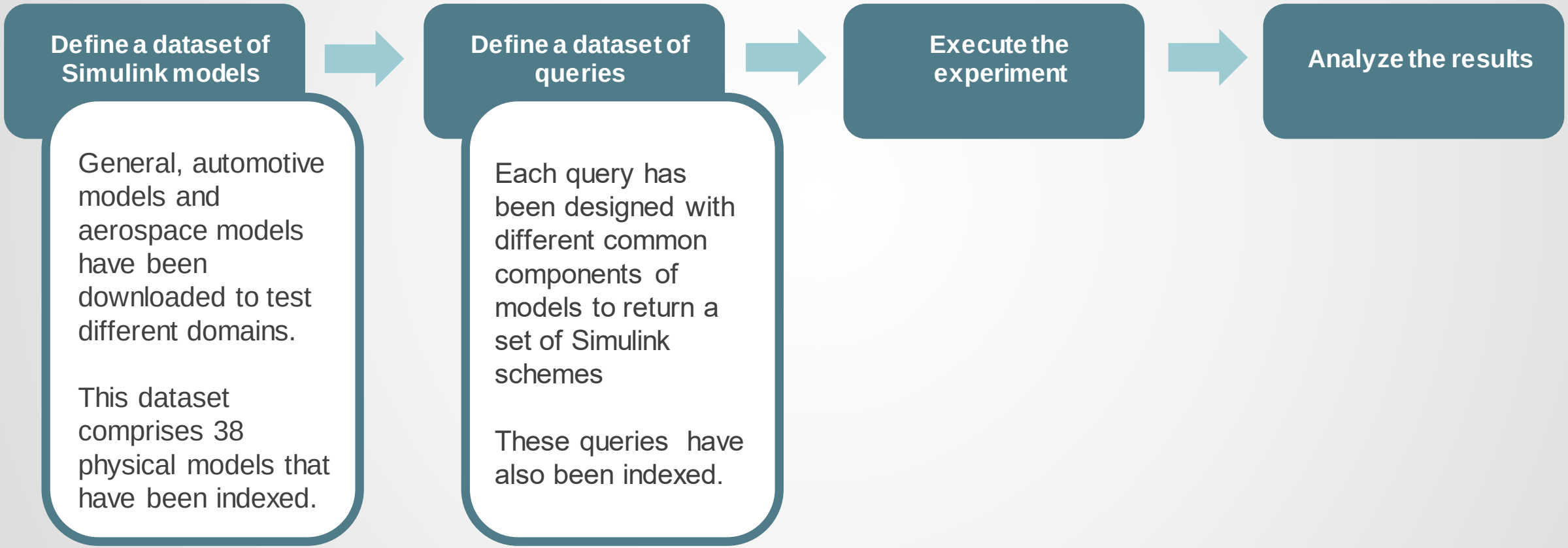
# Case Study




General, automotive models and aerospace models have been downloaded to test different domains.

This dataset comprises 38 physical models that have been indexed.

# Case Study


Define a dataset of Simulink models
General, automotive models and aerospace models have been downloaded to test different domains.
This dataset comprises 38 physical models that have been indexed.
Define a dataset of queries
Each query has been designed with different common components of models to return a set of Simulink schemes
These queries have also been indexed.
Execute the experiment
Analyze the results

# Case Study

**Define a dataset of Simulink models**

General, automotive models and aerospace models have been downloaded to test different domains.

This dataset comprises 38 physical models that have been indexed.

**Define a dataset of queries**

Each query has been designed with different common components of models to return a set of Simulink schemes

These queries have also been indexed.

**Execute the experiment**

For each query defined in the previous step, analyze the models retrieved by Simulink2RSHP taking into account all the semantic information represented into the dataset.

**Analyze the results**

# Case Study

**Define a dataset of Simulink models**

General, automotive models and aerospace models have been downloaded to test different domains.

This dataset comprises 38 physical models that have been indexed.

**Define a dataset of queries**

Each query has been designed with different common components of models to return a set of Simulink schemes

These queries have also been indexed.

**Execute the experiment**

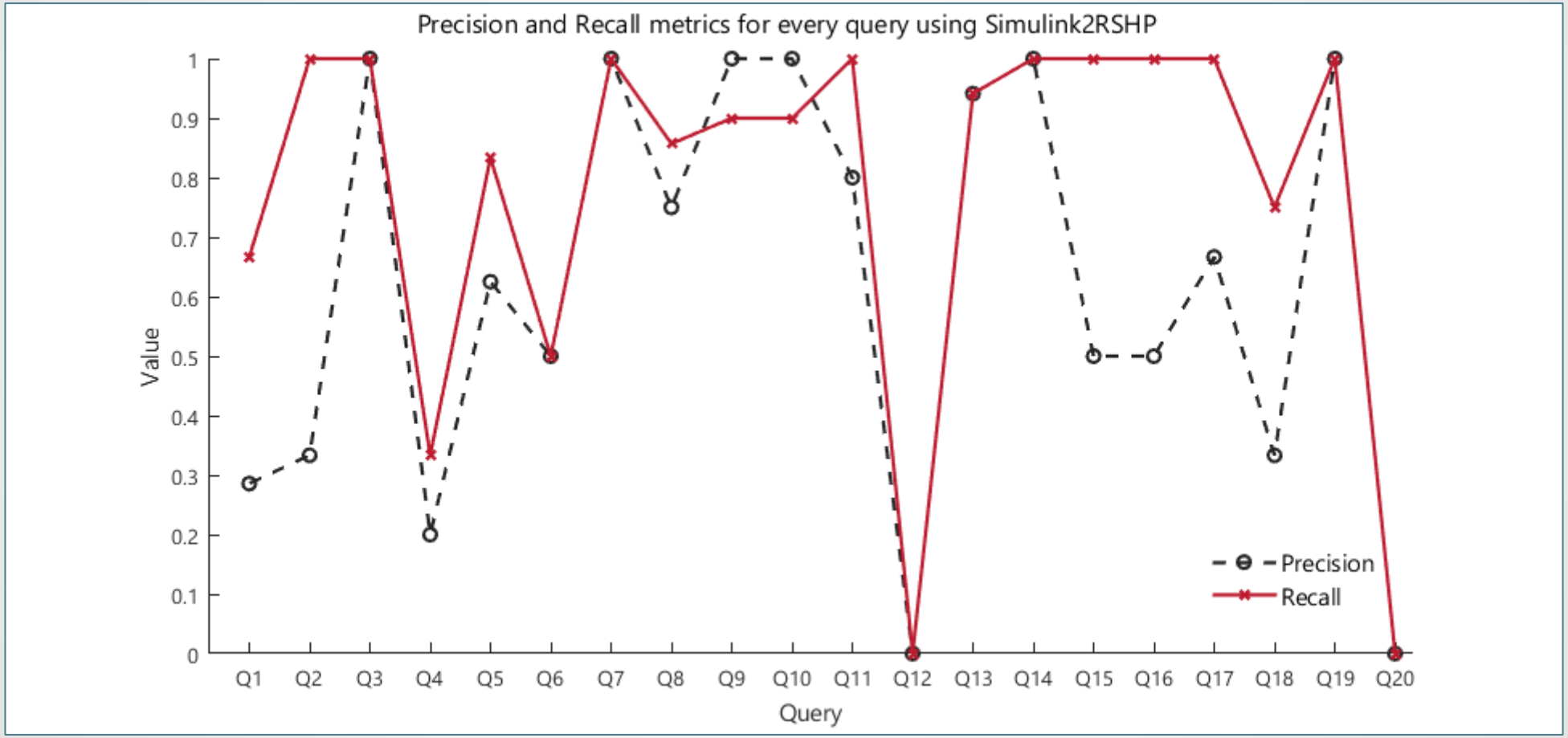
For each query defined in the previous step, analyze the models retrieved by Simulink2RSHP taking into account all the semantic information represented into the dataset.

**Analyze the results**

Extract measures of:

1) Precision
2) Recall
3) F1 score

# Analysis of results



# Analysis of results - Conclusions

## SIMULINK2RSHP

- The global average of the metrics was excellent for precision and good for recall, since they are above 60% and 70% respectively[1].
- 10% of the queries obtained a value of precision below acceptable and, in the case of recall, just 20% of the queries, because the queries had components with incomplete semantic information.

[1]Considering the schema proposed in: J. H. Hayes, A. Dekhtyar, and S. K. Sundaram, "Improving after-the-facttracing and mapping: Supporting software quality predictions," IEEEsoftware, vol. 22, no. 6, 2005, pp. 30–37.

# Paper Conclusions

- Authors have used a Simulink API to propose a process of semantic interpretation of models and have developed Simulink2RSHP which performs the mapping between elements of Simulink models and CAKE components.
- Simulink2RSHPdetermines the similarity using a combination of semantic and topological algorithms.
- The results obtained in the experimentation demonstrate the feasibility of the approach. It is possible to build indexing and retrieval engines for physical models using a semantic representation

# Future Work

- Improvements in the representation of system artifacts are planned, including terminology, thesaurus and semantic clusters.
- Other types of models will be also included in the experimentation, such as those supported in the Modelica language.
- In terms of experimentation, this small setting is representative to demonstrate the feasibility of the approach. However, larger settings including real user needs are completely required to provide a more significant and realistic validation.

# Acknowledgement

The research leading to these results has received funding from the project H2020-ECSEL Arrowhead Tools under grant agreement nº 826452 and from specific national programs and/or funding authorities.

# References


[1] B. Huang, S. Zhang, R. Huang, X. Li, and Y. Zhang, “An effectiveretrieval approach of 3d cad models for macro process reuse,” TheInternational Journal of Advanced Manufacturing Technology, vol. 102,no. 5, 2019, pp. 1067–1089.

[2] K. Robles, A. Fraga, J. Morato, and J. Llorens, “Towards an ontology-based retrieval of uml class diagrams,” Information and SoftwareTechnology, vol. 54, no. 1, 2012, pp. 72–86.

[3] M. W. Whalen, A. Murugesan, S. Rayadurgam, and M. P. E.Heimdahl,“Structuringsimulinkmodelsforverificationandreuse,” in Proceedings of the 6th International Workshop on Modeling in Software Engineering - MiSE 2014. Hyderabad,India:ACMPress,2014,pp.19–24. [Online]. Available:http://dl.acm.org/citation.cfm?doid=2593770.2593776

[4] J. Cordy, “Submodel pattern extraction for simulink models,” ser. SPLC’13. ACM, 2013, pp. 7–10.

[5] E. Gallego, J. M. Alvarez Rodriguez, and J. Llorens, “Reuse of physicalsystem models by means of semantic knowledge representation: A case study applied to modelica,” Sep. 2015, pp. 747–757.

[6] R. Mendieta, J. L. de la Vara, J. L. Morillo, and J. M.´Alvarez-Rodr´ıguez, “Towards effective sysml model reuse.” in MODELSWARD,2017, pp. 536–541.

# Thanks For Your Attention


roy.mendieta@reusecompany.com