



**International Council on Systems Engineering**  
*A better world through a systems approach*

# Extracting Information from System Model as Graph Structure by Large Language Model in MBSE

Keisuke Sugawara  
Yutaka Komatsu  
Atsushi Wada

Japan Aerospace Exploration Agency



# Outline

---

- Background
- Objective
- Related work
- Method & Data
- Experiment
- Conclusion

# Background

- Problem in our MBSE activity
- Why use LLM?

# Background – MBSE

---

## MBSE: Model Based Systems Engineering

**“Model-based systems engineering (MBSE) is the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases.”**

INCOSE SE Vision 2020

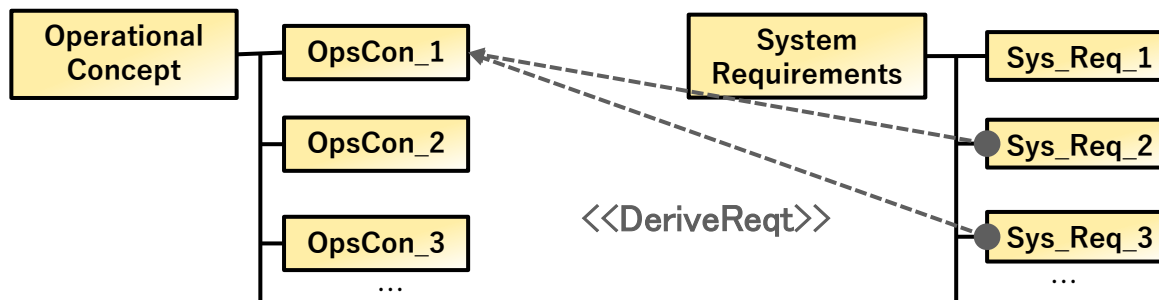
## MBSE activity for satellite development in JAXA:

- Demonstration of MBSE in RAISE (RApid Innovative payload demonstration SatellitE) series
- System model about system specification was built
- Its views were generated.

# Background – MBSE in JAXA

- **MBSE activity for RAISE-3:**
  - Views were generated from the system model and shown to reviewers
  - Conventional document-style system specification were also shown to them
  - → Few reviewers referred the view from system model to make review comment

## Built System Model (Sample)



## Generated View (Sample)

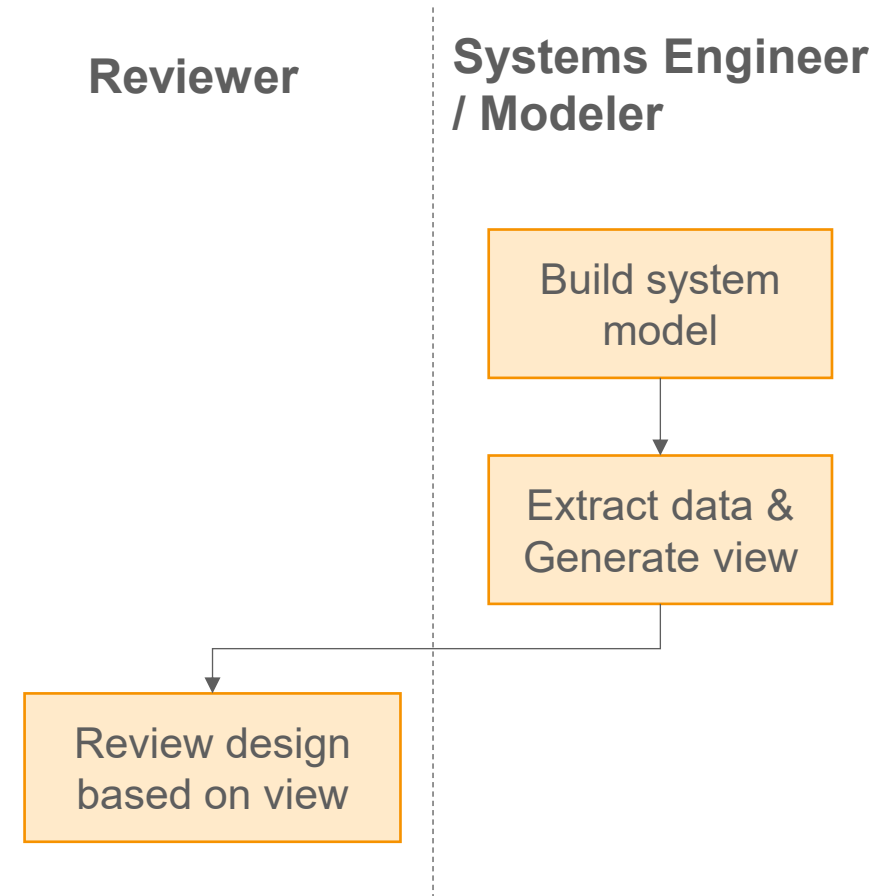
Operational Concept      System Requirement      ICS = Interface Control Specification      Rationale

Table 1. 表3. 上位文書と下位文書の対応表

運用コンセプト	総合システム開発仕様書	ICS	トレーサビリティ機能
・実証テーマ側で実験運用要求を作成するため、システム側より定期的に軌道情報を提供する。	7.3.1 軌道力学系機能 (1) 衛星システムからのGPSRデータを用いて軌道決定ができること。また、その軌道決定値を用いて、運用に必要な軌道情報（衛星軌道予測、アンテナ予報値等）が作成できること。 (2) 2週間後までの軌道予測、地上局可視解析が実施できること。また、それらの情報を必要とする実証テーマに提供できること。	N/A	MIMOチャネル推定実験Activity(衛星管制システム/地上局#1) D2-LEOMI-1 MIMOチャネル容量実験Activity(衛星管制システム/地上局#1) D2-LEOMI-2
・実証テーマ側で軌道情報とMIMO地上局の位置を考慮し、実験運用シーケンスを立案する。	N/A	LEOMI-5.2 作業分担表5.2-1 No22	MIMOチャネル推定実験Activity(LEOMI実験システム#1,2,3) D2-LEOMI-1 MIMOチャネル容量実験Activity(LEOMI実験システム#1,2,3) D2-LEOMI-2
・MIMO送信時には、衛星システムの姿勢制御により、衛星のMIMOアンテナをMIMO地上局（通信対象となる複数の地上局の中心位置）へ指向させる。	6.1.4 姿勢制御機能 (5) D-SAIL展開後の外乱トルクを補償し姿勢制御するため、リアクションホイール1台あたりの最大角運動量を0.3Nm以上、磁気トルクの1台あたりの発生磁気モーメントを15Am <sup>2</sup> 以上とすること	N/A	MIMOチャネル推定実験Activity(バス機器#3) D2-LEOMI-1 MIMOチャネル容量実験Activity(バス機器#4) D2-LEOMI-2

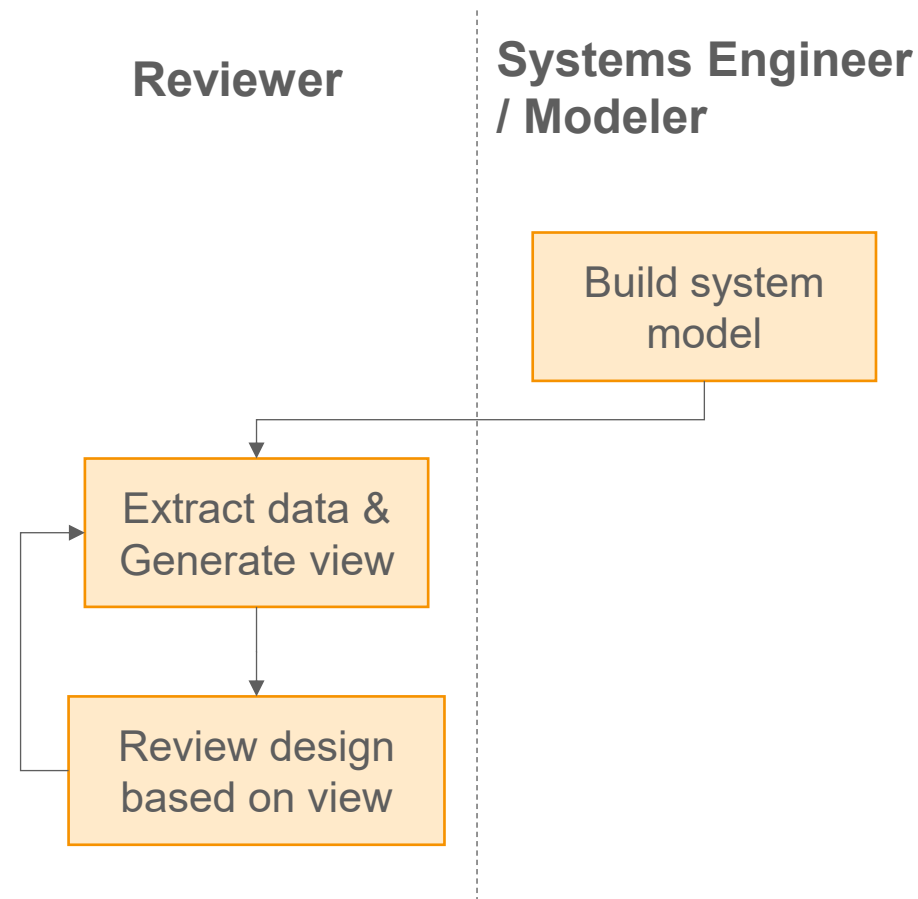
# Background – Model usage (as-is)

- Systems engineers or modelers build system model and make view in advance.
- Reviewers review based on view.
- **Problem:**
  - Information that reviewers can get is limited by systems engineers or modeler
- If reviewers can extract information from system model by themselves, they can do better review



# Background – Model usage (to-be)

- However, to extract information from system model, knowledge of modeling language is required.
- Ideally, there should be a mechanism to extract information without knowing the modeling language



# Background – LLM

---

LLM: Large Language Model

- In recent years, performance has improved rapidly.
- They are able to perform a variety of tasks on behalf of humans

## General main use cases of LLM

Text Summarization

Search Information

Translation

Brainstorming

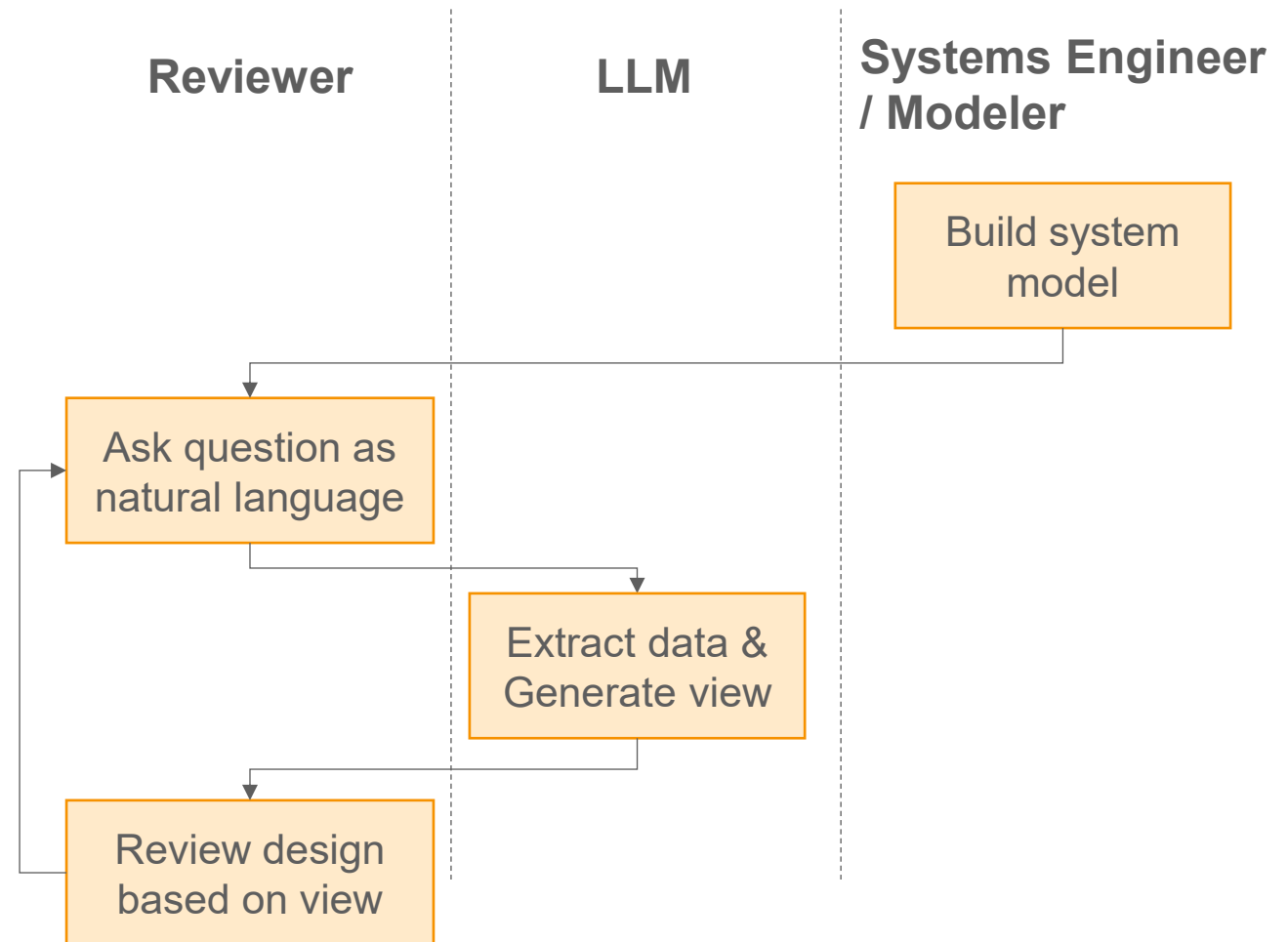
Role-play

Coding



# Objective

- Extract information from system model by using LLM
- Reviewers ask question about system model and LLM interprets it and extract information and gives it to reviewers
- Support reviewers

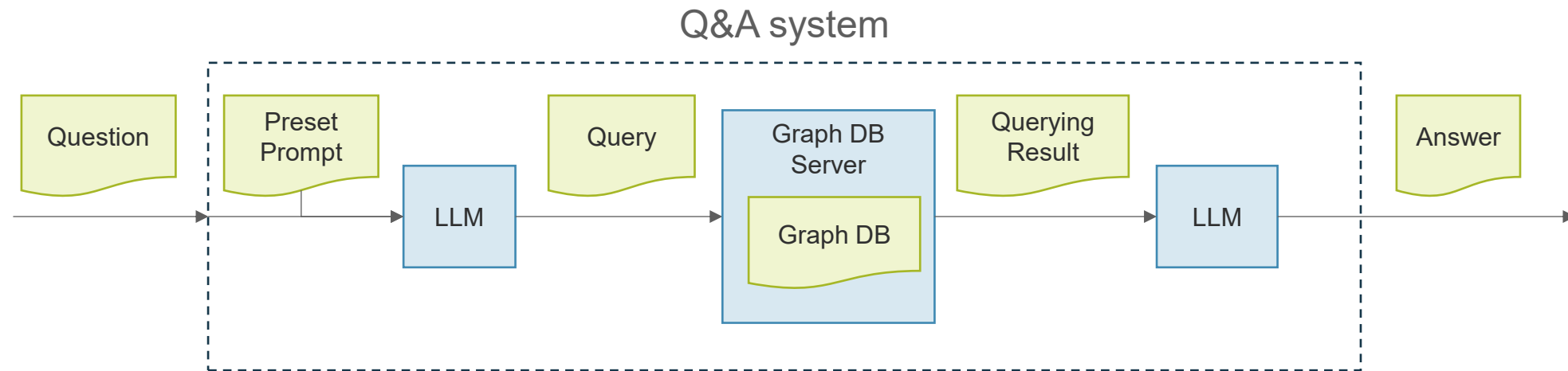


# Related Work

- RAG & Graph RAG

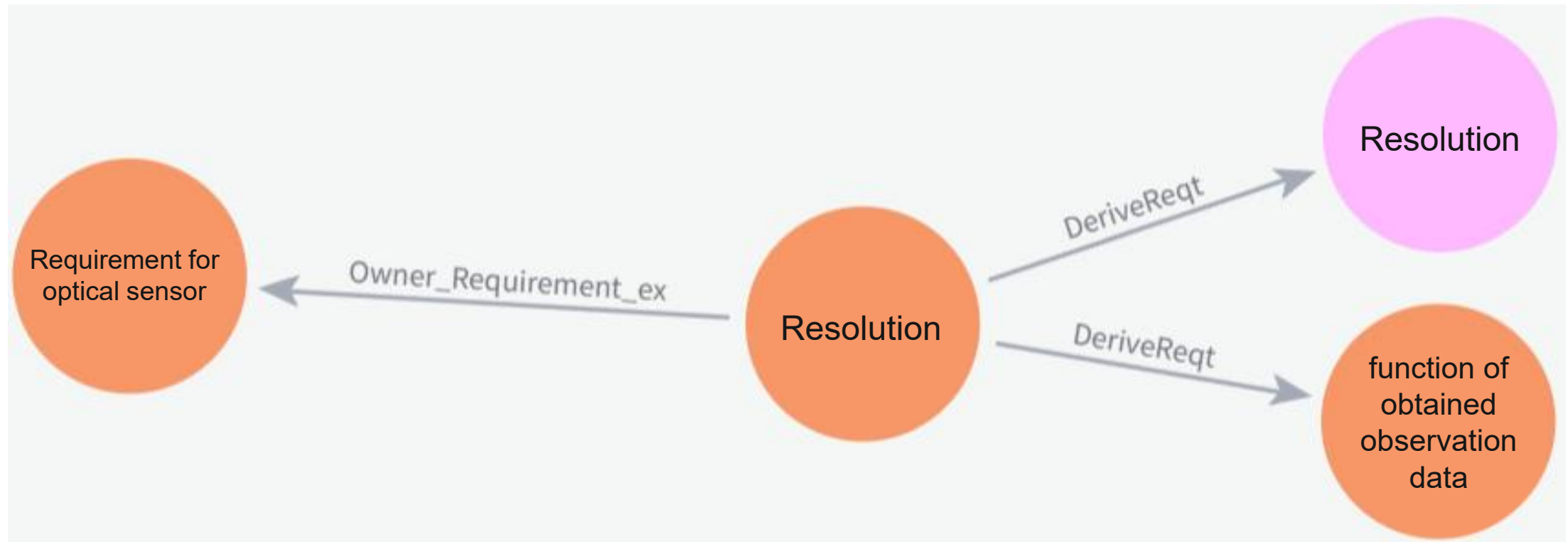
# Related work – RAG

- **RAG** : Retrieval Augmented Generation
- **Graph RAG** Framework:
  - LLM converts a question from a user to a query
  - Querying is executed and a result is given to LLM
  - LLM converts the result into natural language and give it to the user



# Related work – Query from Graph DB

- Neo4j is used as graph database
  - Popular graph database
  - Powerful querying function by Cypher query



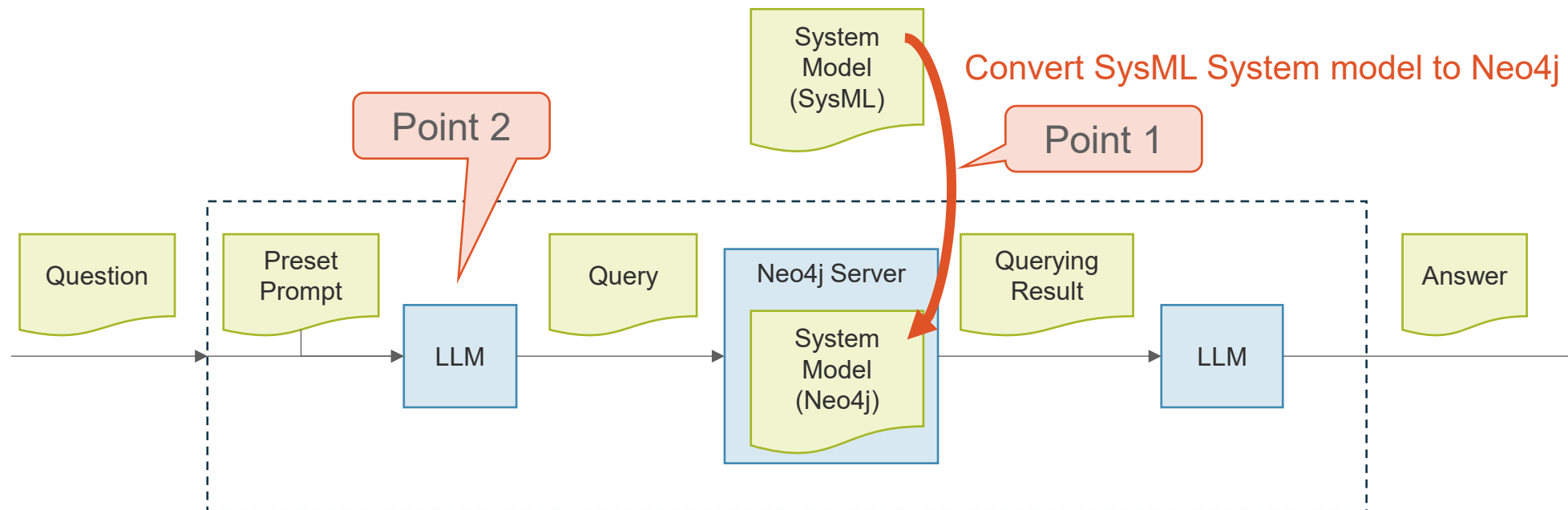
# Method

Incorporate SysML into RAG framework

# Proposed Method

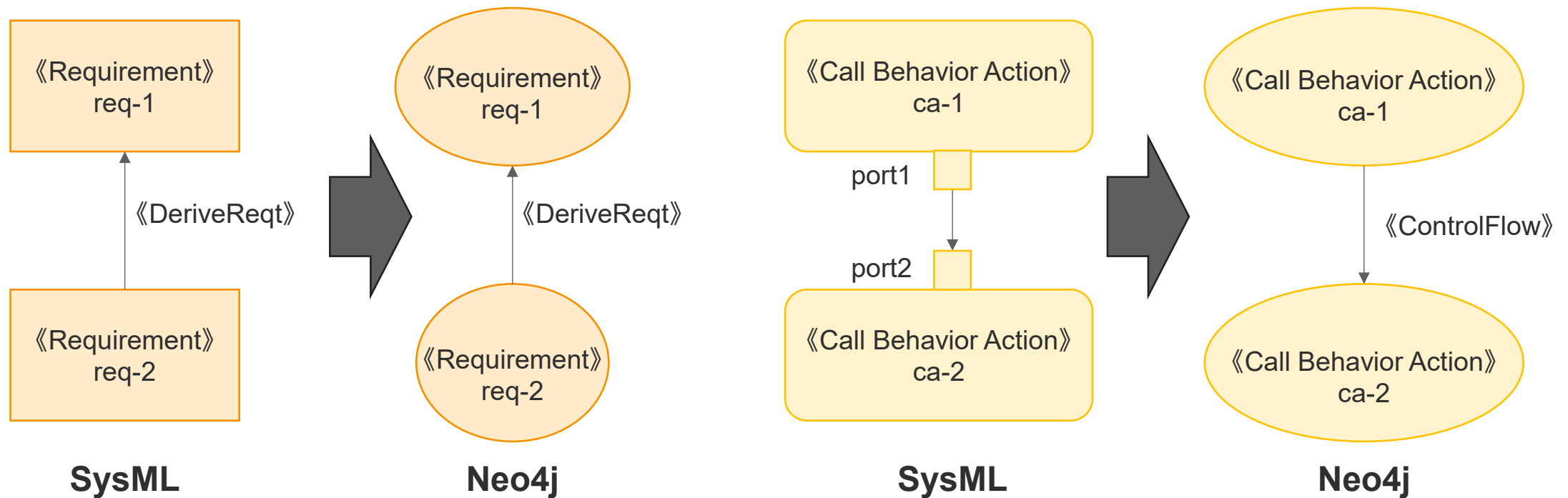
Key points of proposed method:

- 1. Convert SysML model to Neo4j graph
- 2. Convert question from a user as natural language to a cypher query with preset prompt of metamodel



# 1. Convert SysML model to Neo4j graph

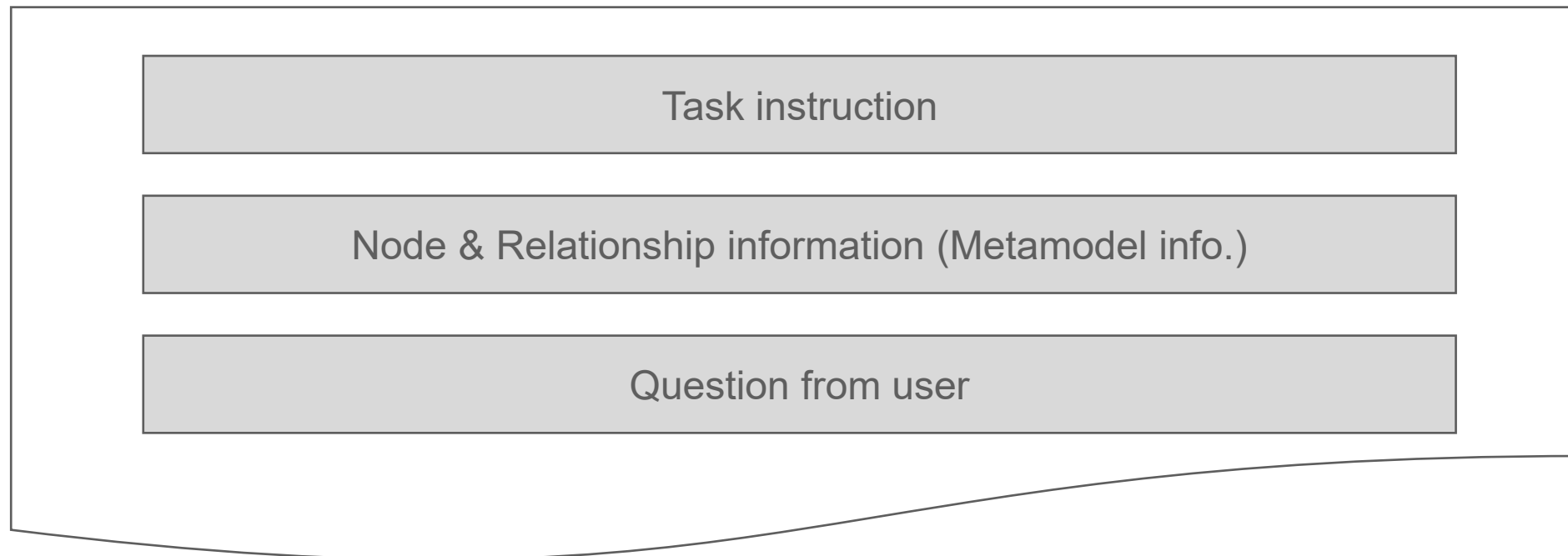
- An element and a relationship of SysML are converted to a node and a relationship of Neo4j
- After exporting SysML model as Generic Table, this convert can be done automatically by Python
- In this study, port information is ignored



## 2. Convert question from a user as natural language to a cypher query with preset prompt of metamodel

- Intension of the metamodel is written in a preset prompt to LLM
- The preset prompt and a user question are merged and given to LLM
- Claude 3.7 sonnet is used as LLM

### Prompt Structure





## 2. Convert question from a user as natural language to a cypher query with preset prompt of metamodel

- Intension of the metamodel is written in a preset prompt to LLM
- The preset prompt and a user question are merged and given to LLM
- Claude 3.7 sonnet is used as LLM

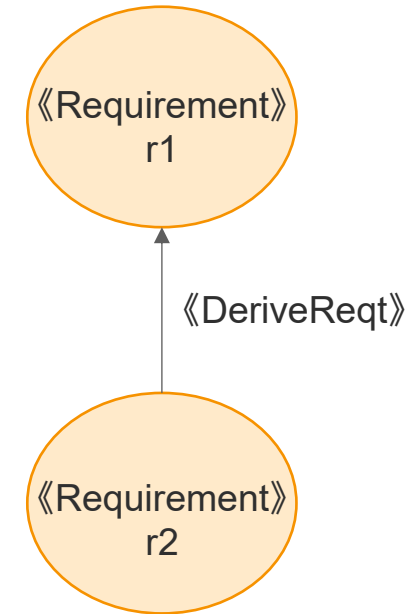
### Prompt Template

```
#Task: Make Cypher query for Neo4j database.

# Additional information
The node 'Requirement' means a requirement for the system. The
property 'requirement_target' indicates the system to which the requirement is allocated.
...(about other nodes) ...

the relationship '(r1:Requirement)-[:DeriveReq*]->(r2:Requirement)' means that r1 is
derived from r2, or r1 is also a requirement to realize r2.
... (about other relations) ...

# Question
{question}
```

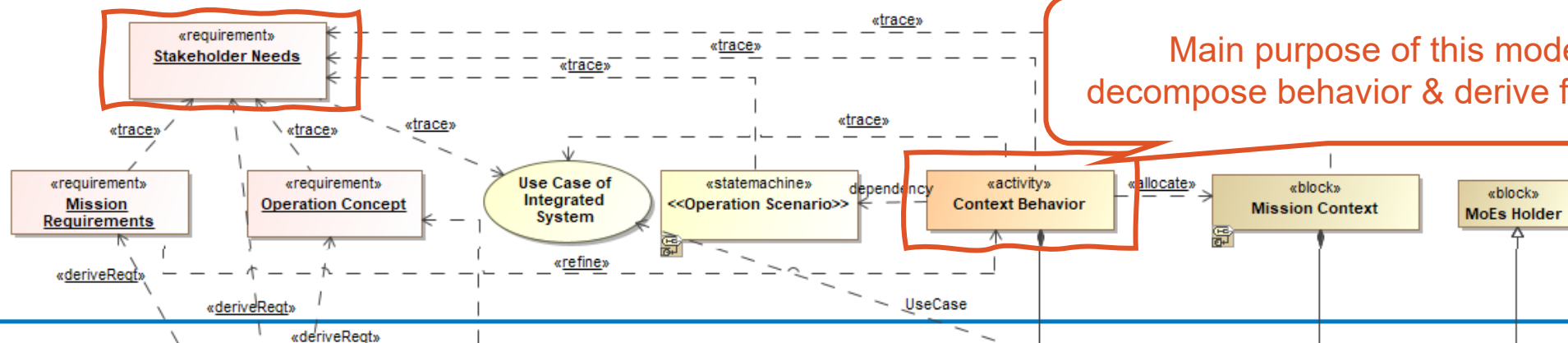


# Experiment

# Target System Model – Metamodel



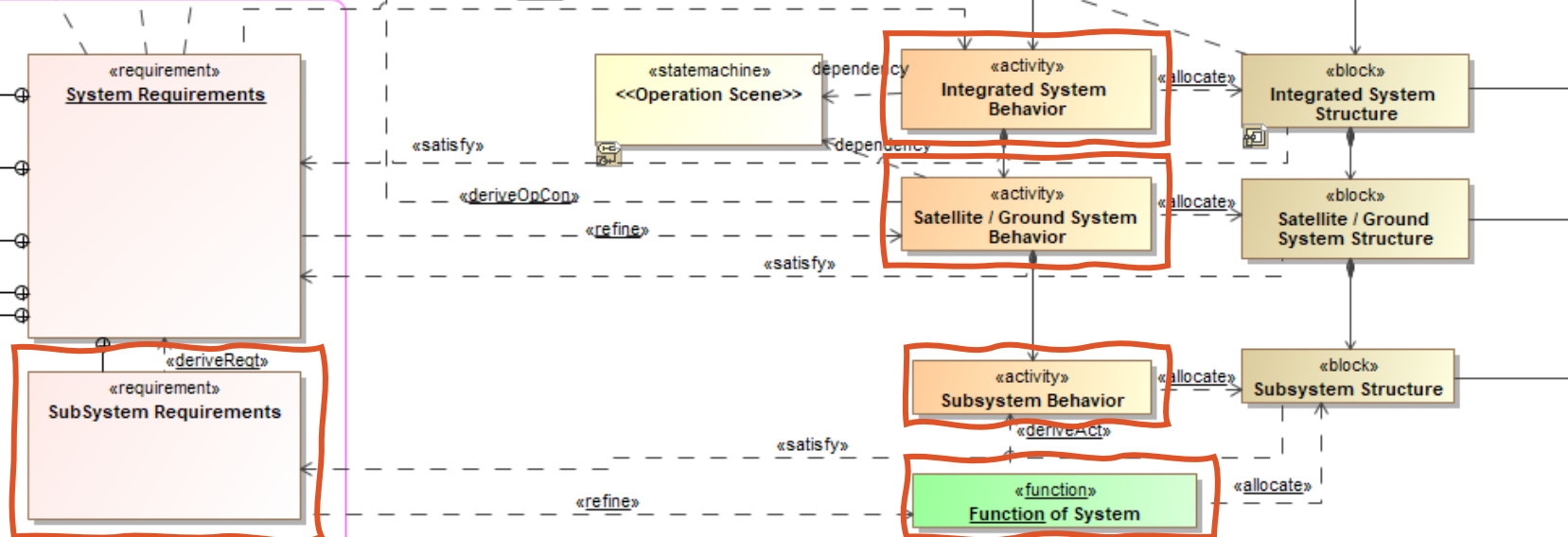
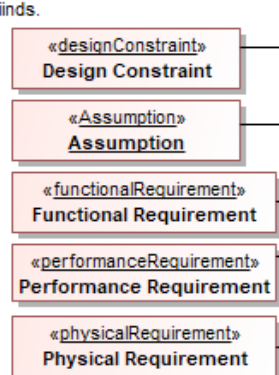
PROBLEM  
(BLACK BOX)



Main purpose of this model:  
decompose behavior & derive function

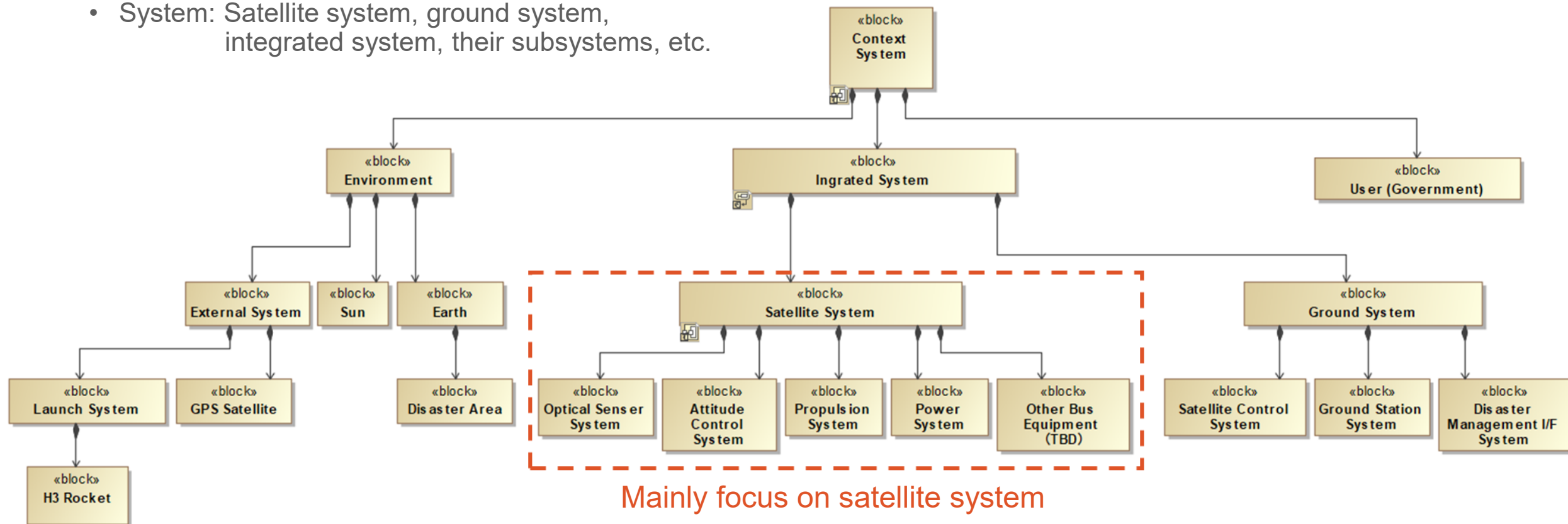
PROBLEM  
(WHITE BOX)

System Requirements/SubSystem Requirements are divided into these kinds.



# Target System Model – Description

- Fictitious satellite system
- Mission: Provide optical image observed by satellite
- System: Satellite system, ground system, integrated system, their subsystems, etc.



# Test Case (Questions about the System Model)

---

- Assumed situation
  - Design review.
  - Reviewers are professionals of SE and intend to ask questions about design of the system.
- Only questions about the fact in the system model are covered.
  - e.g., List the all mission requirements.
- Questions requiring inference are **not covered**.
  - e.g., What is the most important mission requirement?
- The number of questions is 51.

# Test Case (Questions about the System Model)

- Characterize questions with three types of features
- It is possible to know the response performance for each feature, enabling deeper analysis
- Hypothesis: The more features a question has, the easier to answer correctly

No.	Feature	Meaning
f1	Directness of reference	Must be answerable by direct reference to the target data. There is no need to refer to it via two or more relationships, such as element A -> element B -> element C.
f2	Uniqueness of expected elements	The expected answer is narrowed down to one. Correct example: To which system is functional requirement "A" assigned? Incorrect example: List all the behaviors that use function "B".
f3	No spelling inconsistencies or typographical errors in the question text.	Question text does not include any spelling inconsistencies or typographical errors.

# Example 1: Answered Correctly

- A user asked a question about previous LL (lessons learned).
- The system model has an element defined by “LLs”.
- The query searched the word “LL” and was able to find the proper element.



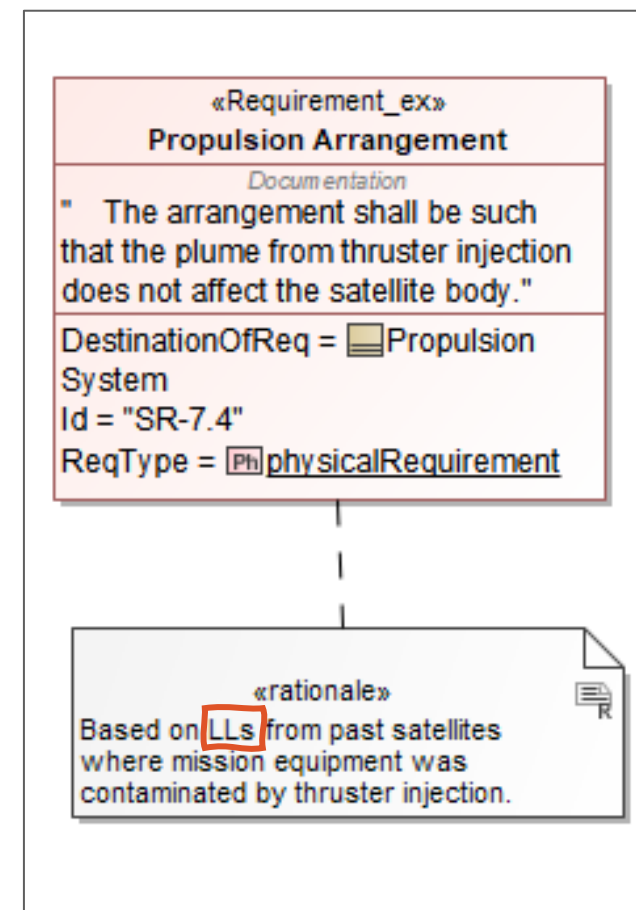
**[User question]** Show a list of requirements defined based on previous LLs.

**[Cypher query]**

```
MATCH (r:Requirement_ex)
WHERE r.documentation =~ '.*LL.*' OR r.rationale =~ '.*LL.*'
RETURN r.name, r.documentation, r.rationale, r.requirement_target, r.requirement_type
```

**[Answer to the user]**

One requirement defined based on past LLs is the requirement for thruster placement. This requirement requires that the plume from thruster injection be placed in such a way that it does not affect the main body of the satellite. This is based on lessons learned from previous satellites where thruster injection contaminated mission equipment. This requirement is classified as a physical requirement for the propulsion system.



# Example 2: Answered Correctly

- Questions that require tracing multiple relationships (requirement, function and structure)
- LLM was able to generate Cypher and query



## [User question]

Show a list of the functional requirements assigned to the attitude control system.

LLM

## [Cypher query]

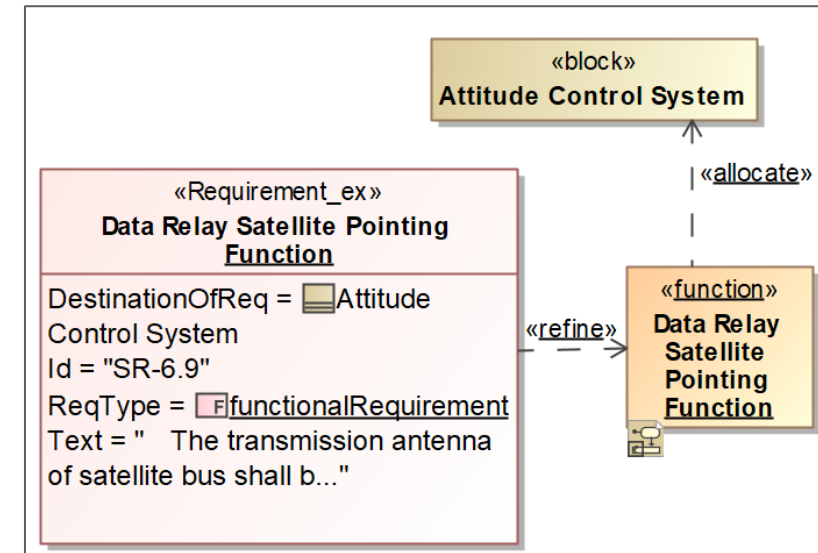
```
MATCH (r:Requirement)-[:Refine]->(f:function)-[:Allocate]->(s:Structure)
WHERE r.requirement_type = 'functionalRequirement [Class]'
      AND (s.name =~ '.*attitude.*' OR s.name =~ '.*ADCS.*')
RETURN DISTINCT r.name, r.documentation, r.requirement_target
```

## [Answer to the user]

The following four functional requirements are assigned to the attitude control system

- (1) Data relay satellite pointing function: The satellite ...
- (2) Ground station pointing function: The satellite bus transmission antenna ...
- (3) Attitude maneuver function: The satellite system can maneuver ...

...





# Example 3: Answered Incorrectly

- A user asked a question about previous lessons learned.
- The system model has an element defined by lessons learned, but the expression “LLs” is used instead.
- The query doesn’t match anything and failed.



## [User question]

Show a list of requirements defined based on previous **lessons learned**.

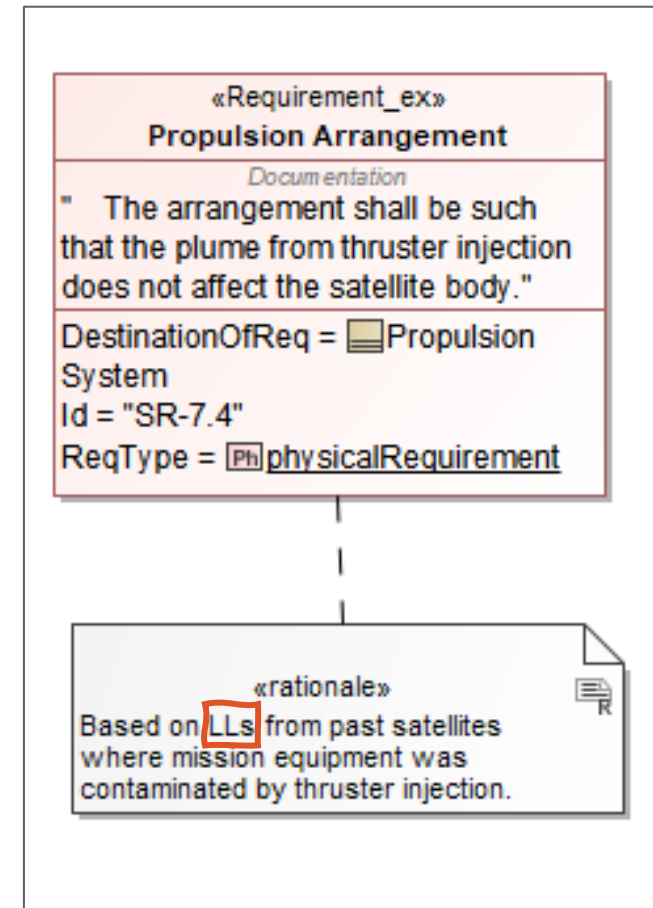
LLM

## [Cypher query]

```
MATCH (r:Requirement_ex)
WHERE r.documentation =~ '.*lessons learned.*' OR r.rationale =~ '.*lessons learned.*'
RETURN r.name, r.documentation, r.requirement_target,
r.requirement_type, r.rationale
```

## [Answer to the user]

Sorry, but we do not have information on the list of requirements defined based on previous lessons learned. We do not have the necessary data to provide a specific answer to this question.



# Evaluation

---

- An answer from LLM is scored on a 3-levels scale.
  - 100: The answer by LLM is consistent with the correct answer.
  - 50: The answer by LLM is partially consistent with the correct answer.
  - 0: The answer by LLM is hardly consistent with the correct answer.
- Overall score
  - The mean score for all questions was 78.1 points.
- And questions are categorized and evaluated according to their combination of features.
  - Three types of features → the combination is  $2^3 = 8$

# Evaluation

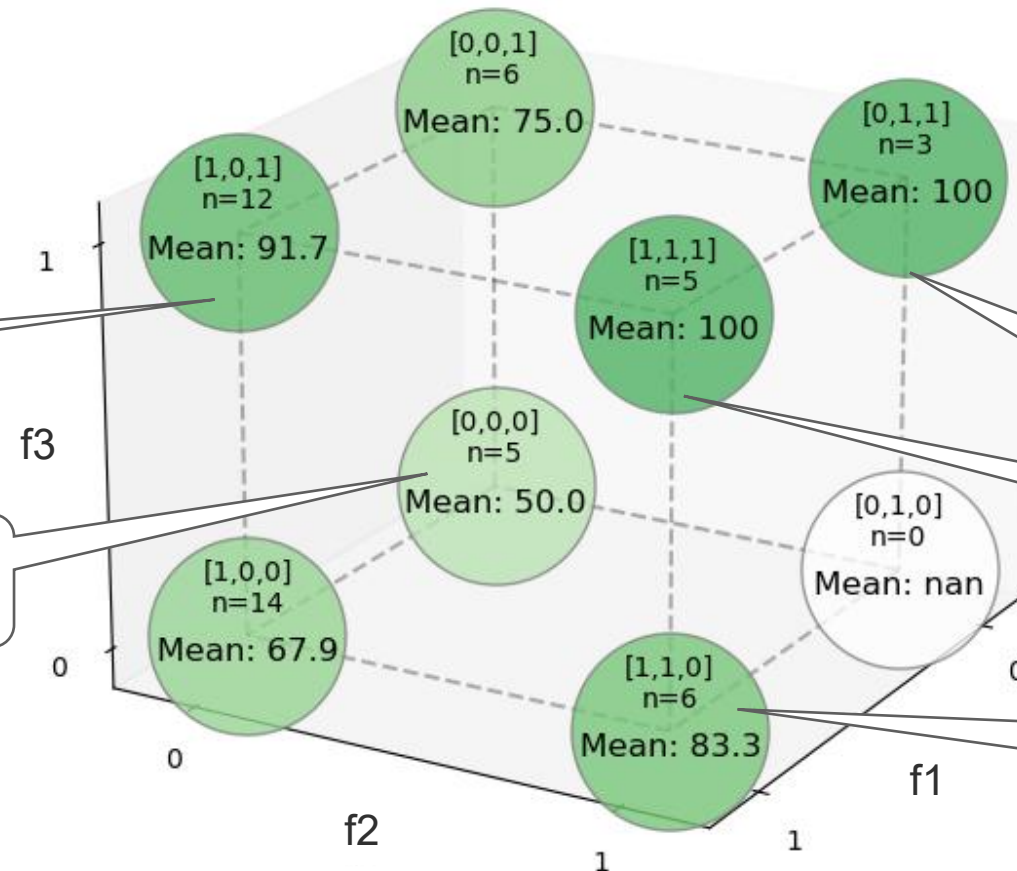
- **Result:** The more features a question has, the higher the mean score of their answers.
  - The proposed method is vulnerable to typographical error or spelling inconsistencies

[f1, f2, f3]  
n=The number of question.  
Mean: Mean score.

No.	Feature
f1	Directness of reference
f2	Uniqueness of expected elements
f3	No spelling inconsistencies or typographical errors in the question text.

[1,0,1] doesn't have Uniqueness of expected elements

[0,0,0] Questions that satisfy no features (the most difficult cases)



[0,1,1] doesn't have directness of reference

[1,1,1] Questions that satisfy all features (the easiest cases)

[1,1,0] has typographical error or spelling inconsistencies

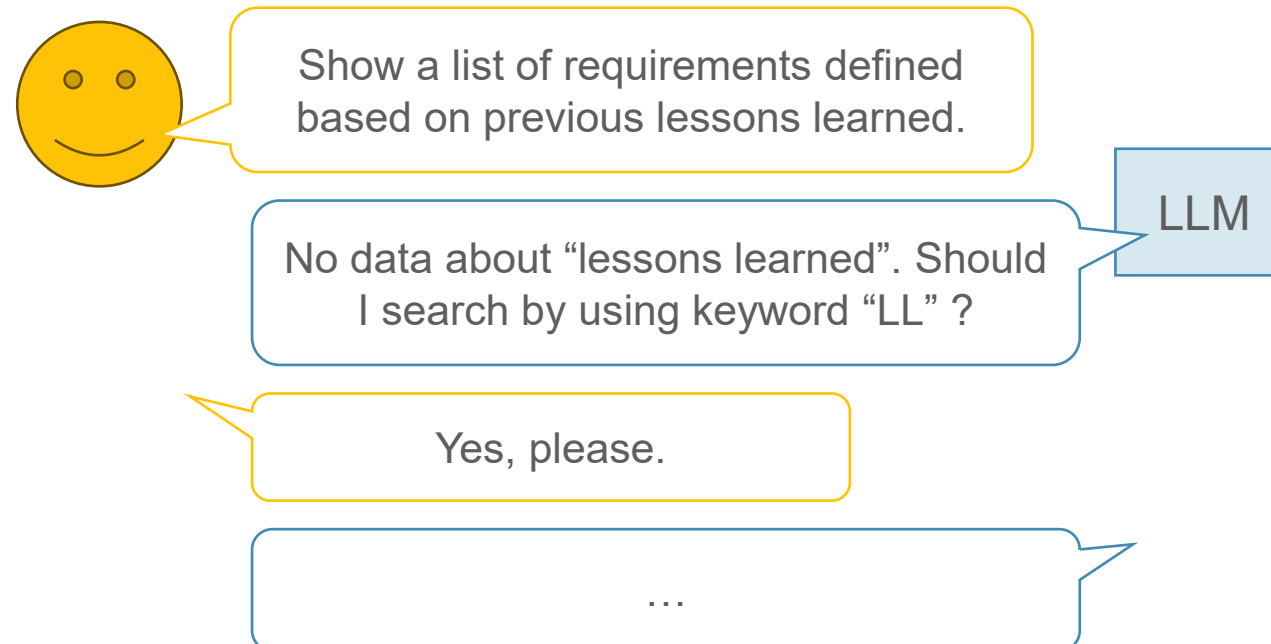
# Discussion – (1)Weakness and Countermeasure

---

- Questions which satisfies many features are easy to be answered correctly
- Consideration for each feature:
  - f1: Directness of reference
    - Probably the metamodel information is written in the prompt, thus LLM was able to understand it and make query accurately
  - f2: Uniqueness of expected elements
    - It should be because Cypher can make query that list all element satisfying given conditions
  - f3: No spelling inconsistencies or typographical errors in the question text.
    - **The proposed method is vulnerable to typographical error or spelling inconsistencies**
    - It should be because Cypher query is easy to perform keyword searches
      - If a given keyword has typo, a query to be made will be also has typo.

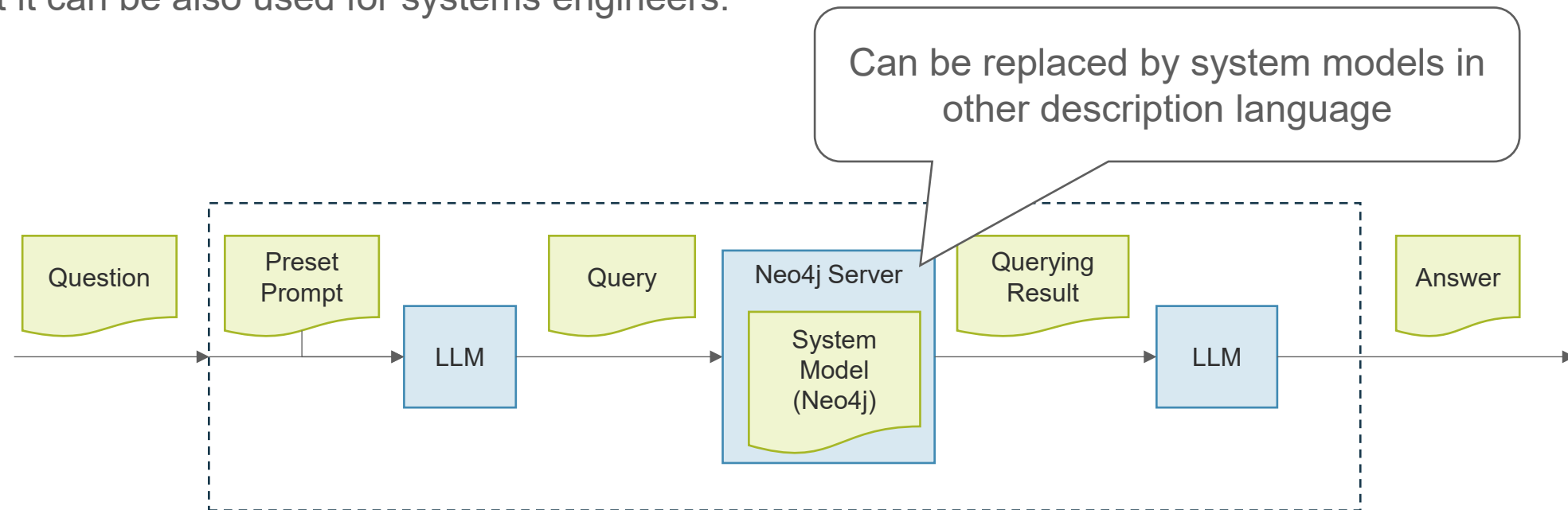
# Discussion – (1) Weakness and Countermeasure

- Remaining issues and countermeasures
  - Vulnerableness to typographical error or spelling inconsistencies
- Countermeasures: **Interactive conversation approach**



## Discussion – (2)Extensibility

- Our proposed framework is applicable for other modeling languages and other LLM.
  - Overall performance depends on performance of LLM
- In this study, we assumed reviewers as users.  
But it can be also used for systems engineers.



# Conclusion

---

- **Background:**
  - SysML model is hard to understand and extract information for users who is unfamiliar to it
- **Objective:**
  - Extract information from system model by question as natural language by using LLM
- **Proposed method:**
  - converting SysML model to Neo4j graph,
  - converting users' questions into Cypher query with preset prompt of metamodel,
  - and converting querying result into an answer to the user
- **Result:**
  - LLM can convert simple questions to Cypher query and extract information correctly
  - Vulnerable to typographical error or spelling inconsistencies
- **Future work:**
  - Interactive conversation approach

# Appendix



# Detail of the Target System Model

---

- The number of the elements (nodes in Neo4j): 297
- The number of the elements (nodes in Neo4j): 470
- Elements
  - Requirement, Block, Activity, Rationale, function
- Relationships
  - DeriveReqt, Allocate, Refine, ObjectFlow, ControlFlow. Trace, Association

# Prompt (1 of 2)

---

Template for the prompt:

#Task: Make Cypher query for Neo4j database.

#schema:

{schema}

#Note:

Use only the relation types and properties provided in the schema.

Do not use other relation types or properties not provided.

Responses should be in Japanese. Do not include explanations or apologies in your answers.

Do not answer questions that ask for anything other than the creation of a Cypher statement.

If the query results empty, simply state that you cannot answer the question.

Do not include text other than the generated Cypher query.

User questions may contain typographical errors or distortions. Attempt to perform a partial match search for multiple properties, rather than a full-text search. Splitting words and using wildcard search are effective.

Consider that the question is related to artificial satellites and aerospace, create a query after correcting any words that appear to be misspelled as necessary.

There is no need to limit the number of characters output. If multiple records are generated as query results, do not reduce the number of outputs.

# Prompt (2 of 2)

## #Additional information:

The target graph database represents design information for satellite and ground systems.

Simple name is written in the property "name" of each node. The details of the design information are described in the property "documentation".

The node "Requirement\_ex" represents a requirement for the system. It includes mission requirements and stakeholder needs. The property "requirement\_target" indicates the system to which the requirement is made. The property "re-quirement\_type" can be "functionalRequirement [Class]" meaning a functional requirement or "performanceRequirement [Class]" meaning a performance re-quirement. The property "rationale" is a rationale of the requirement.

The relationship (r1:Requirement\_ex)-[:DeriveReq\*]->(r2:Requirement\_ex) means that r1 is derived from r2. It can also be said that r1 is a requirement needed to realize r2.

The relationship (r1:Requirement\_ex)-[:Trace]->(r2:Requirement\_ex) means that r1 is a mission requirement and r2 is a stakeholder needs, and r1 is derived from r2.

The node "function" means a function at the satellite subsystem level.

The "Refine" from "Requirement\_ex" to "function" means that the function is redefined as a requirement. Note that not all Requirement\_ex have a relationship to functions.

The node "Activity" means behavior, action, or operation. The property "context" of "Activity" means the subject of the behavior.

"DeriveAct" from 'function' to 'activity' means that the function is used in the activity.

"Allocate" from 'function' to 'Structure' means that the block executes the function or the function is assigned to the block.

"Association\_block2block" from 'block' to 'block' means that it is a part of the block.

"Association\_Activity2Activity" represents hierarchy relationship between "Activity" to "Activity". It means that the latter Activity is executed in the former Activity.

"ControlFlow\_Activity2Activity" from 'Activity' to 'Activity' means that the latter Activity is executed after the former Activity.

"ObjectFlow\_Activity2Activity" from 'Activity' to 'Activity' means that the latter Activity is executed after the former Activity, with passing an object. The property "Object" is it.

Some items contain TBD (To Be Determined) information.

## Example of a question and query:

"What are the requirements that can be derived from a requirement related to 'AAA'?"

The query for the question is "MATCH (r1:Requirement\_ex)-[:DeriveReq\*]->(r2:Requirement\_ex) WHERE r1.name =~ '.\*AAA.\*' RETURN DISTINCT r2.name, r2.documentation, r2.requirement\_target, r2.requirement\_type"

## #Question:

{question}

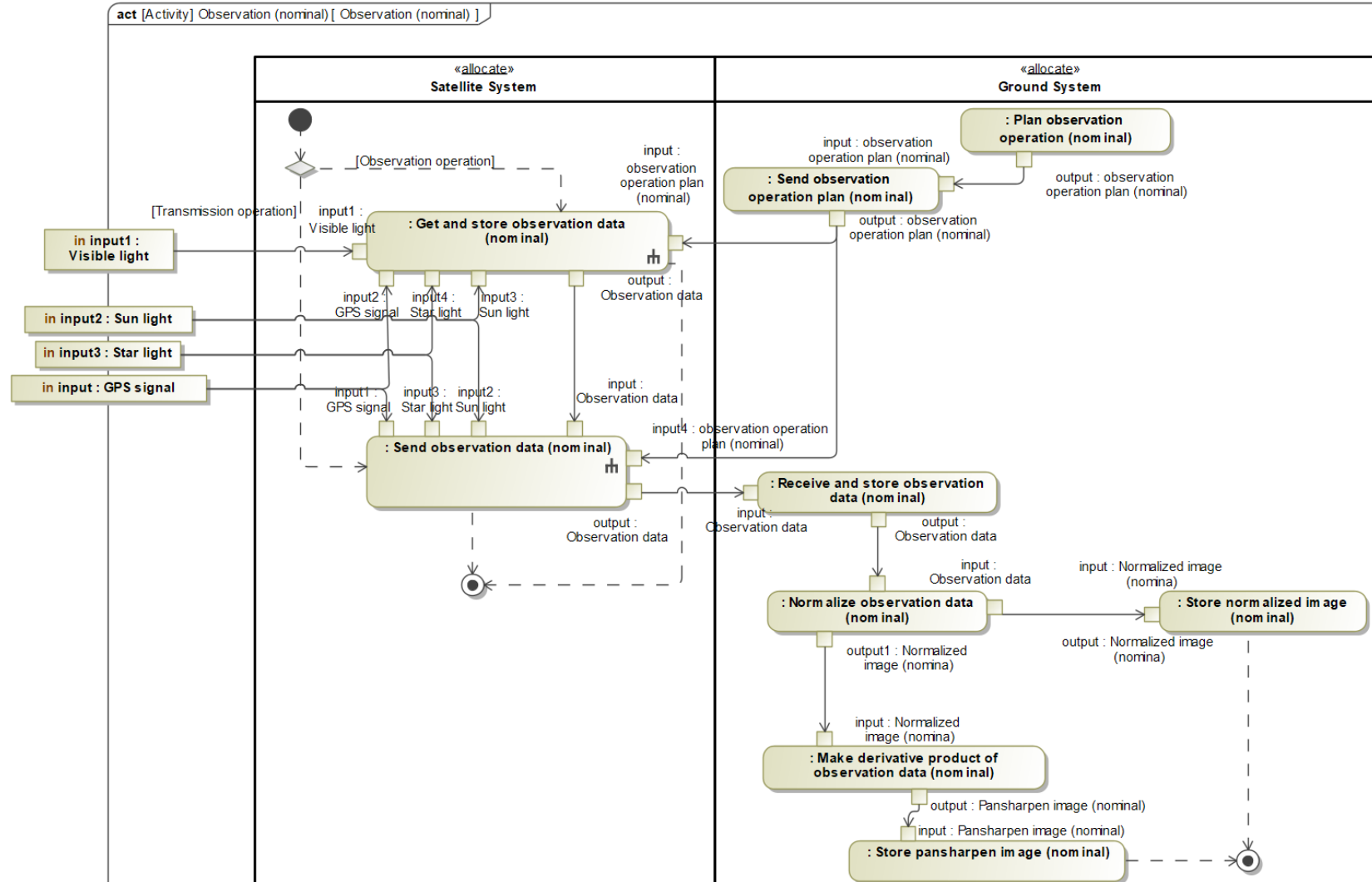
# Test Case (Questions about the System Model)

- Assume reviewers' intention and derive questions from it

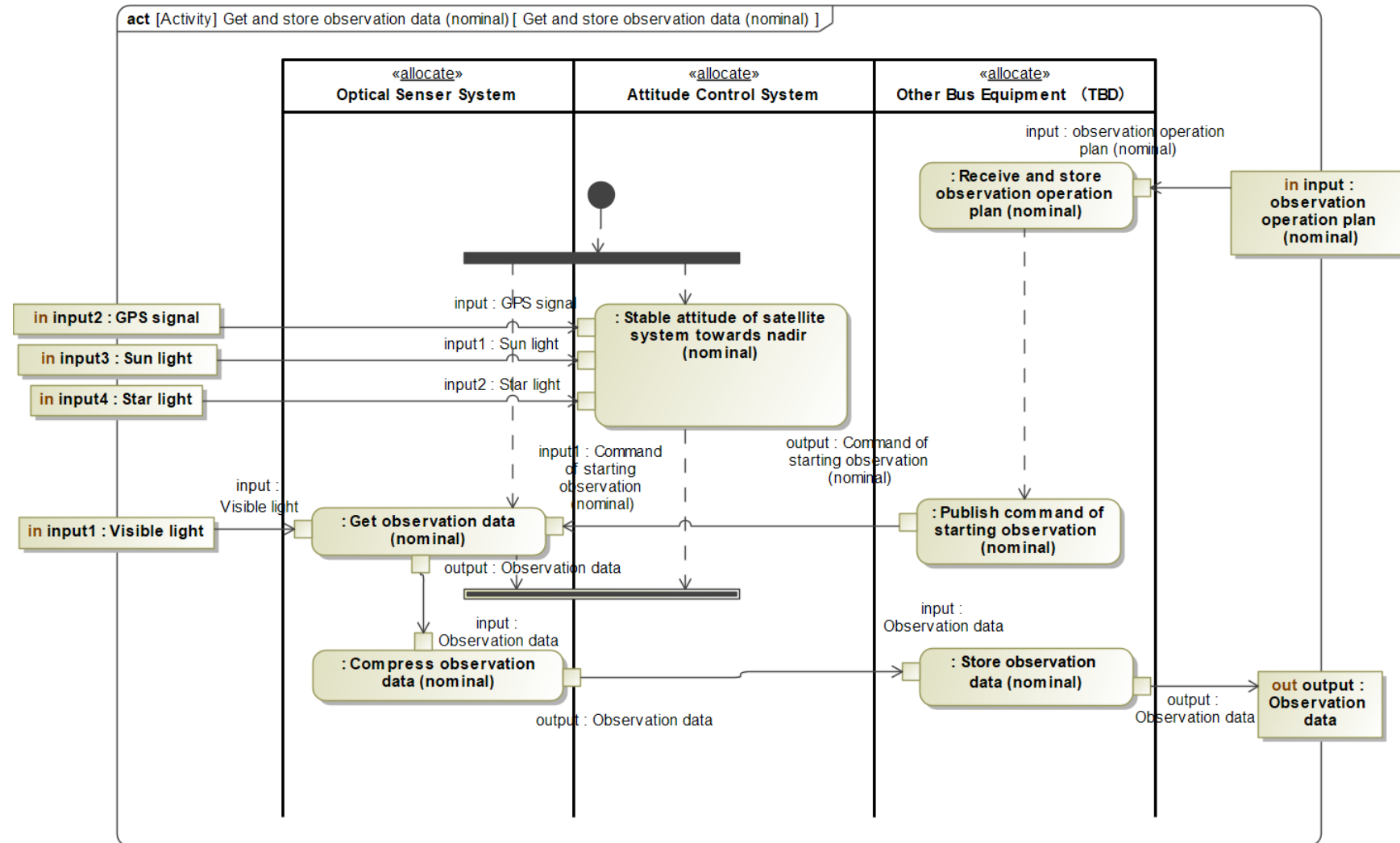
## Intension of question list

No.	Background and intension of question
A	Designers and reviewers check the list of items included in the TBD (to be determined), and designers consider whether there are any items that can be finalized by refining the design. Based on past experience, the reviewers check for items that should not be left undecided (e.g., items with rework risk).
B	Describe in natural language the flow of the scenario modeled as an activity diagram or state transition diagram (assuming the reviewers are not particularly familiar with SysML).
C	Ensure that the functions are broken down sufficiently to accomplish the mission.
D	Answers are used as evidence to review whether a certain requirement is appropriate for a higher-level requirement or not. The model is assumed to be referenced by tracing from lower-level functions to upper-level requirements via "Derive".
E	Ask whether all performance, functions, etc. are defined as requirements without omission.
F	Ask which functions are performed by which systems/subsystems (assuming reviewers are particularly concerned about the allocation of functions between satellite and ground systems).
G	Question the design rationale and determine its validity. Assume reference to "Rationale" tied to functional requirements; if no Rationale, assume reference to other relevant elements.

# Activity Diagram (Integrated System Level)



# Activity Diagram (Integrated System Level)



# Q&A about System Behavior

Question	True answer	LLM's answer
Show a list of activ-ity done by attitude control system in activity of the inte-grated system "Ob-servation (nominal)"	<ul style="list-style-type: none"> <li>- Point antenna to data relay satellite (nominal)</li> <li>- Point antenna to the ground station (nominal)</li> <li>- Stable satellite system attitude towards nadir (nominal)</li> </ul>	<p>The following three behaviors are performed by the attitude control subsystems in the integrated system behavior "Observation (nominal)":</p> <ul style="list-style-type: none"> <li>- Point antenna to data relay satellite (nominal)</li> <li>- Point antenna to the ground station (nominal)</li> <li>- Stable satellite system attitude to-wards nadir (nominal)</li> </ul> <p>These behaviors play an important role in the satellite's ability to make nomi-nal observations and transmit data properly.</p>

# Requirement and Rationale

