Mission Assurance for AI software Systems

- Guidelines for developing AI application with a focus on uncertainty -

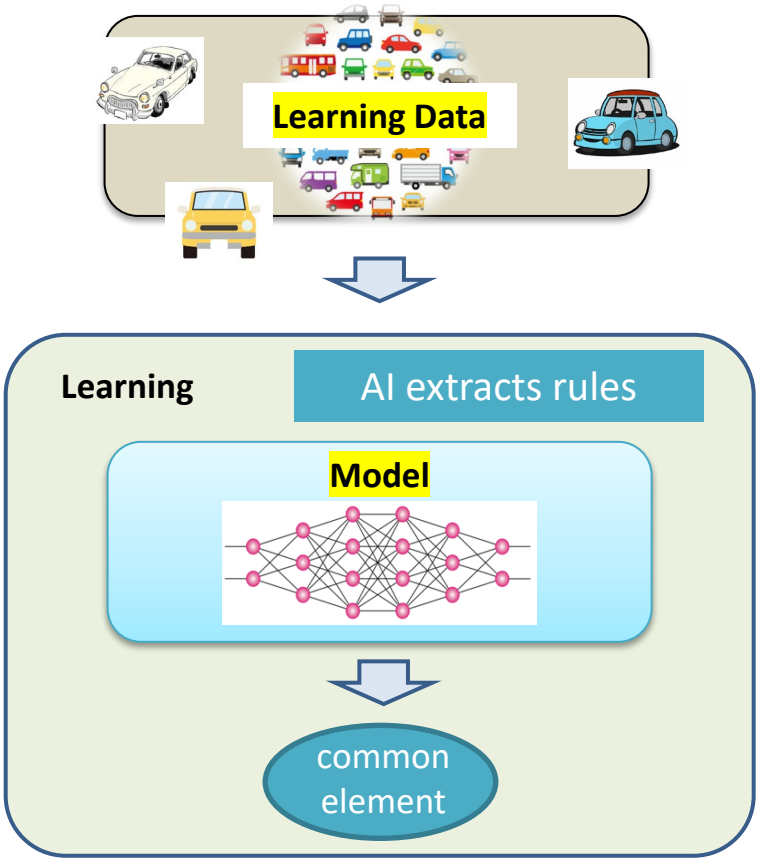
TRISMAC 2024
2024/6/24 - 26

Japan Aerospace Exploration Agency
Naoki Ishihama

Outline

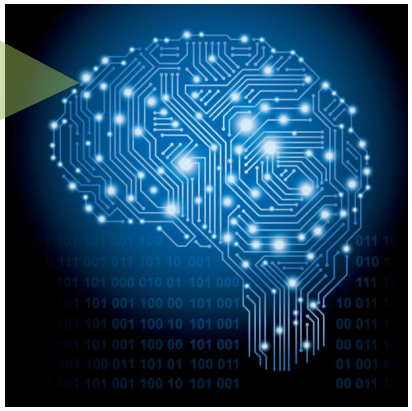
- About AI
- Different between legacy system and AI system
- Causes of False Reasoning of AI software
- Strategic Solutions to reduce uncertainty during reasoning
- AI system development process

About AI software
- Example: Deep Learning Software

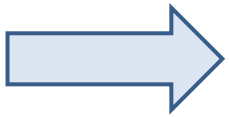


About AI software
- Example: Deep Learning Software

Deep Learning



...and...
Probability of "4" 60.6%.
...and...
Probability of "9" 39.4%.



AI system development process

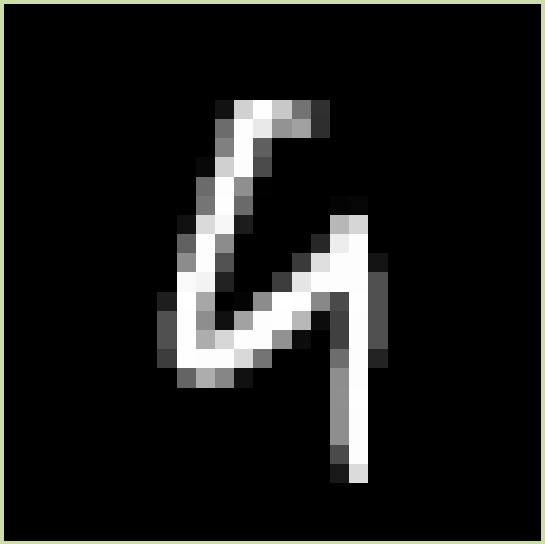
Phase	Summary
System Requirement Analysis Phase	Analyze use cases (system usage scenarios, external factors, etc.) and identify/clarify the system requirements.
System Design Phase	Design the architecture of the system (hardware and software).
AI Software Requirement Phase	Identify/clarify the **AI software functions/non-functions** requirements
AI Learning preparation phase	**Prepare and preprocess training data** based on use cases and requirements, etc.
AI learning phase	**Build the AI Software model and train** by using the preparing data until met to the requirement
AI test phase	Perform reasoning tests on the trained AI software model using test data and **validate the reasoning results**.
AI software release	AI software unit verification completed
System test phase	Do the system test including whole software and verify and validate for system requirement.
System deploy	Ships products
System operational phase	Monitor conditions during operation. When unexpected conditions occur, saved the data and do post hoc analysis.

AI software development phase

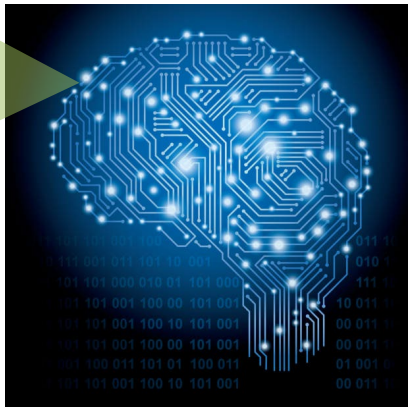
Difference between “legacy system” and “AI system” development process

Legacy system development process	AI system development process
System Requirement Analysis Phase	
System Design Phase	
Software Requirement Phase	AI Software Requirement Phase/ AI Learning preparation phase
Software Design Phase	AI learning phase 
Software Coding Phase	
Software testing phase	AI test phase
Software release	AI software release
System test phase	
System deploy	
System operational phase	

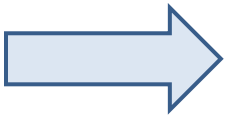
False reasoning …..



Deep Learning



...and...
Probability of "4" 92.2 %.
...and...
Probability of “9" 7.8%



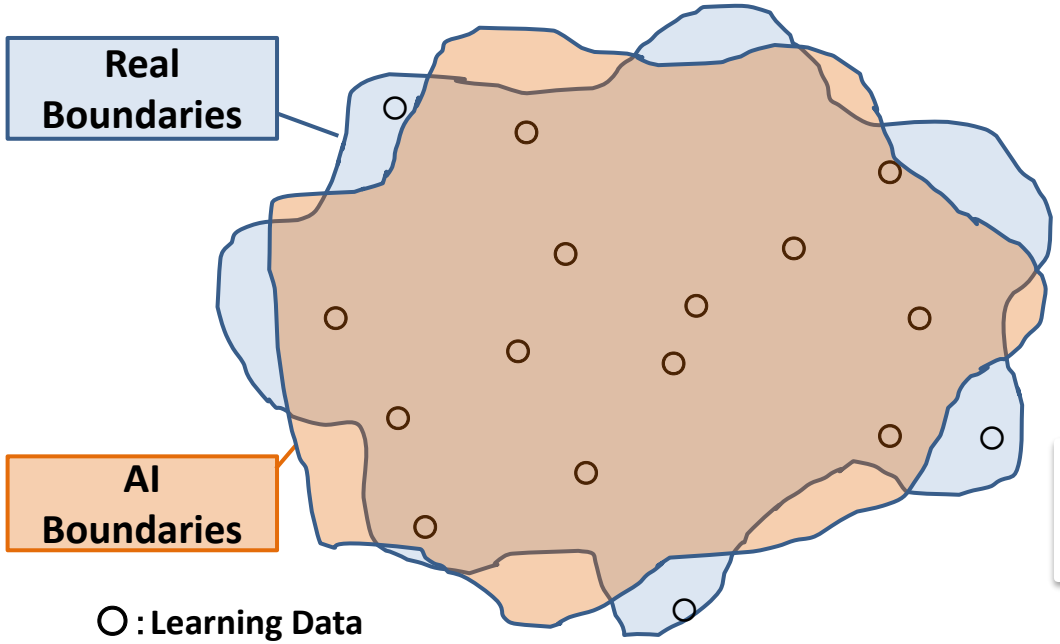
False Reasoning

False Reasoning type and reason

False Reasoning Type:

- AI answers what it knows as something else.
- AI cannot identify what it knows.
- AI answers what it does not know as what it does know.

Why do make false reasoning?



What is the Causes of False Reasoning??

Causes of False Reasoning

Cause A: Insufficient training data => **Model and Data Uncertainty**	Cause A1: **No training data** for reasoning object
	Cause A2: **Insufficient variation** in training data for reasoning object
	Cause A3: **Insufficient training** data for reasoning object taking into account **external and internal factors**
	Cause A4: **Insufficient quantity or poor quality** of training data for reasoning object
Cause B: Includes **different training data** from the inference target => **Data Uncertainty**	
Cause C: **Insufficient model** => Model **Uncertainty**	
Cause D: Incorrect implementation	

Key: Reduce the uncertainty

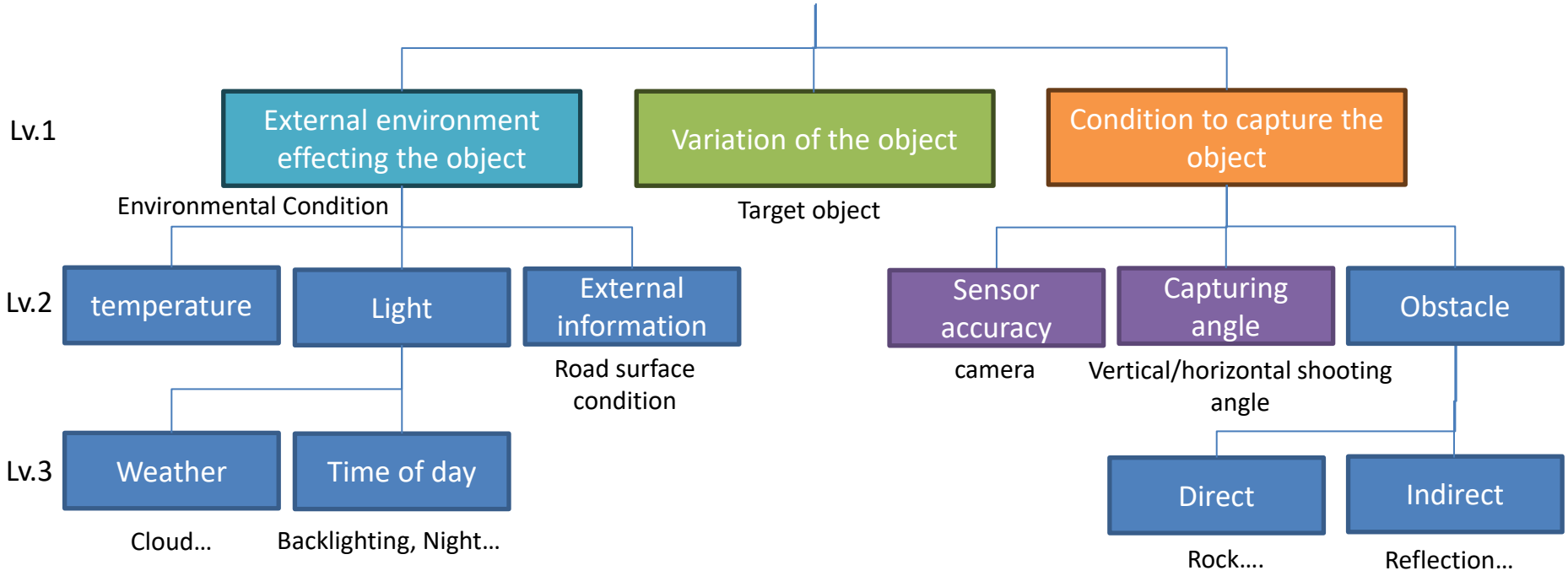
Define the "Cause" and "Solution"

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

- Solution1: Systematic context analysis

Cause	Solution
Cause A: Insufficient training data	**Solution1**: **Prepare necessary and sufficient training and test data** using systematic context analysis methods



Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

- Solution1: Systematic context analysis

Cause	Solution
Cause A: Insufficient training data	**Solution1**: **Prepare necessary and sufficient training and test data** using systematic context analysis methods

….. Difficult to prepare 100% of the necessary and sufficient data

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

Cause	Solution
Cause A: Insufficient training data	**Solution1**: **Prepare necessary and sufficient training and test data** using systematic context analysis methods

….. Difficult to prepare 100% of the necessary and sufficient data

Cause	Solution
Cause A: Insufficient training data	**Solution2: Identify the incorrect reasoning results**

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

- Solution2: Identify the incorrect reasoning results

Cause	Solution
Cause A: Insufficient training data	**Solution2: Identify the incorrect reasoning results**

Breakdown the "Cause" and "Solution"

Cause	Solution
Cause A1: **No training data** for reasoning object	Solution2-2: Identify **the change of input data trend**
Cause A2: **Insufficient variation** in training data for reasoning object	Solution2-2: Identify **the change of input data trend** Solution2-3: Identify **incorrectly captured features** Solution2-4: Identify **ambiguous** reasoning results
Cause A3: **Insufficient training** data for reasoning object taking into account **external and internal factors**	Solution2-2: Identify **the change of input data trend** Solution2-4: Identify **ambiguous** reasoning results
Cause A4: **Insufficient quantity or poor quality** of training data for reasoning object	Solution2-1: Identify **misalignment and inconsistency in data distribution** Solution2-2: Identify **the change of input data trend** Solution2-3: Identify **incorrectly captured features** Solution2-4: Identify **ambiguous** reasoning results

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

- Solution2: Identify the incorrect reasoning results

Solution
Solution2: Identify the incorrect reasoning results
Solution2-1: Identify **misalignment and inconsistency in data distribution**
Solution2-2: Identify **the change of input data trend**
Solution2-3: Identify **incorrectly captured features**
Solution2-4: Identify **ambiguous** reasoning results

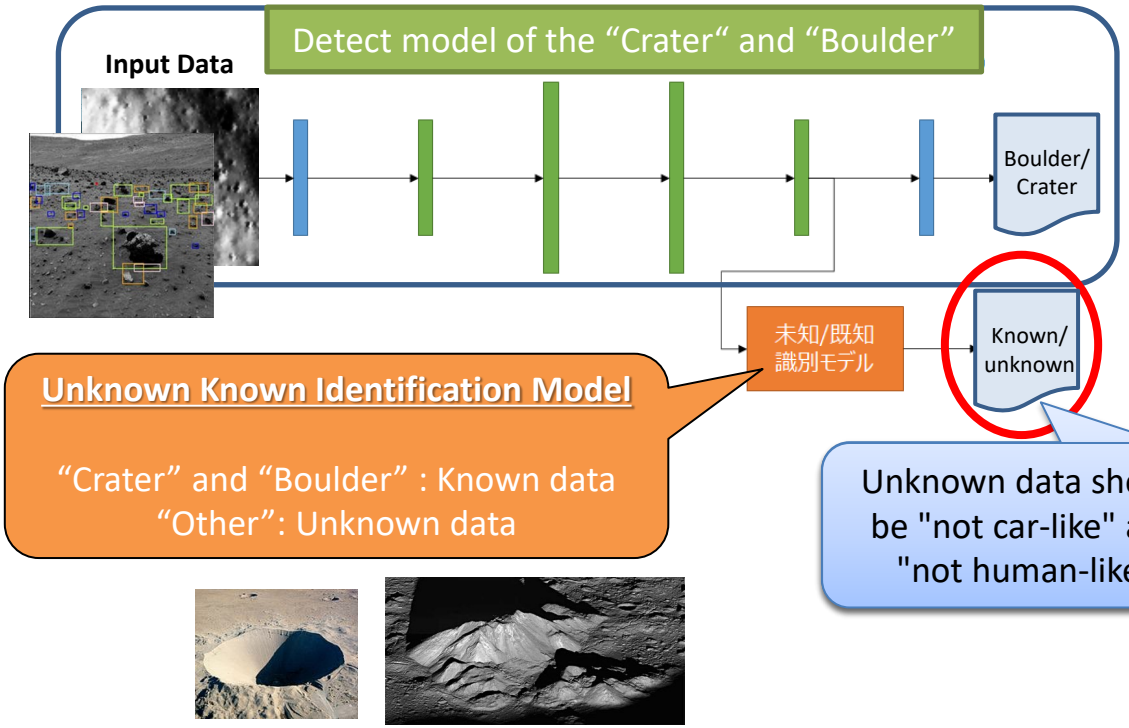
Strategic Solutions to reduce uncertainty during reasoning

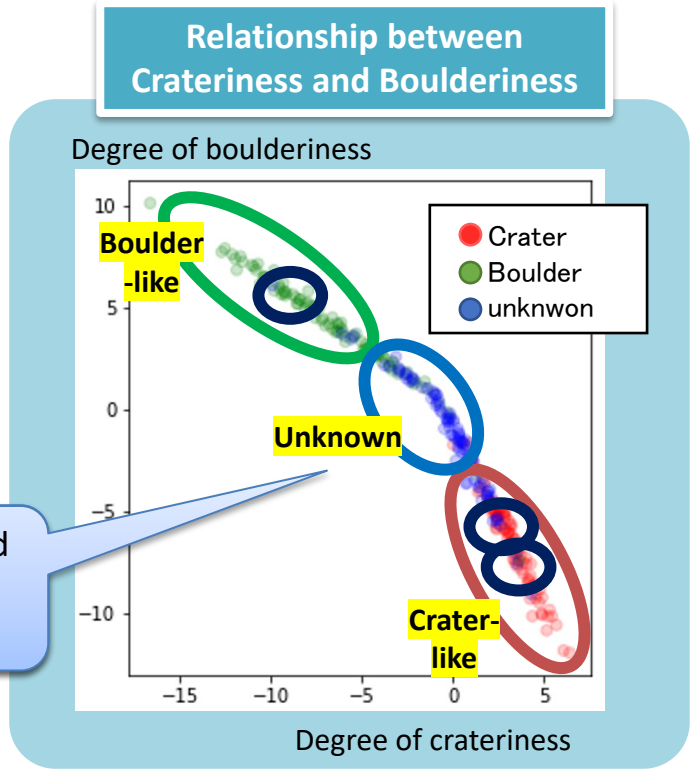
- Cause A: Insufficient training data

- Solution2-2: Identify the change of input data trend

AI is answering what it doesn't know
- Identifies Crater and Boulder and Rims of different shapes and sizes

Example method:





Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

- Solution2-3: Identify **incorrectly captured features**

Recognize the incorrect way of remembering

- The AI is answering by looking at the characteristics of the surrounding environmental information, not the object to be identified.

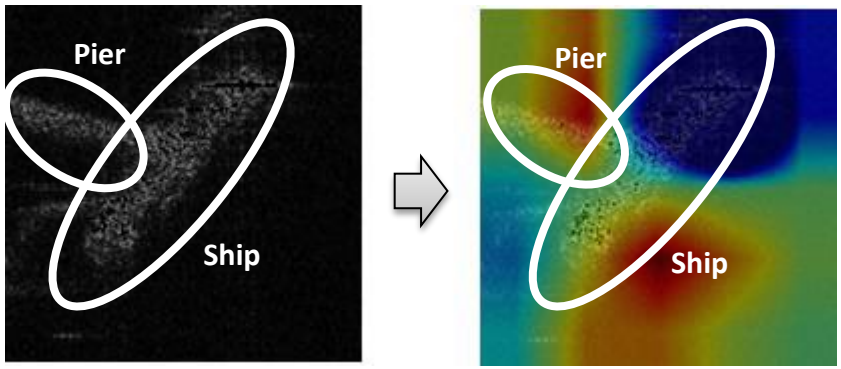
Ex: The AI sees a pier as a feature and answers "ship".

=> Reason: Since the image of the vessel shows a pier, the AI extracted the pier as a common feature

Example method:

Grad-CAM: This method visualizes where the AI focused its attention and gave an answer.

Ex: Classification task (Ship)



SAR Image (Classification task)

Grad-CAM

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

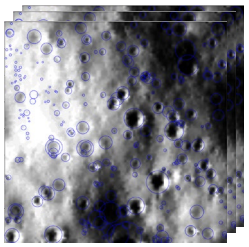
- Solution2-4: Identify **ambiguous** reasoning results

Detect that AI made a mistake because it remembered something halfway.

Ex: Detection task (“Crater” and “Boulder”)

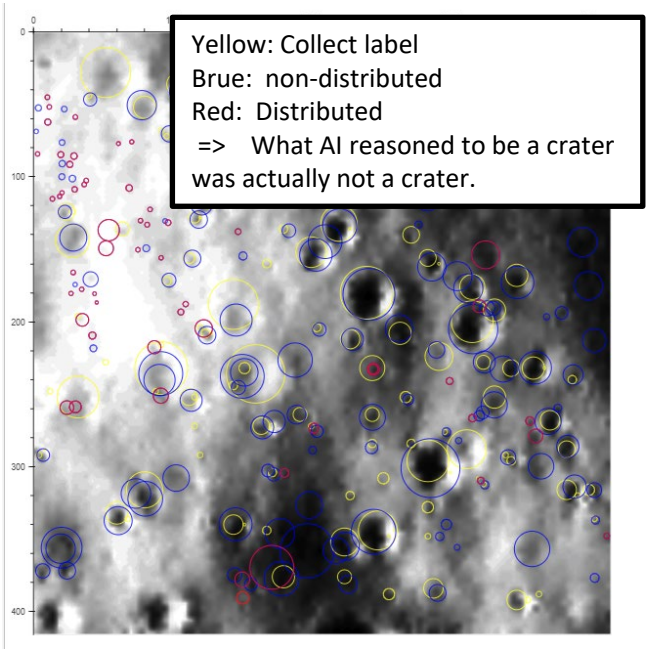
Test/Operation Data

Detection



Detection Result
(N times sampling)

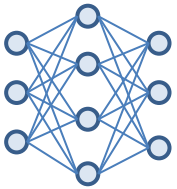
Statistical processing



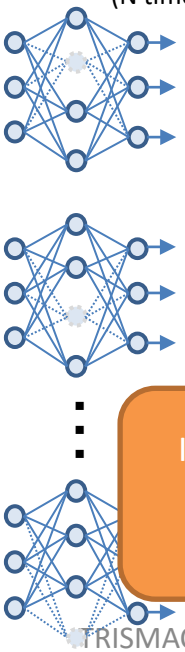
Yellow: Collect label
Brue: non-distributed
Red: Distributed
=> What AI reasoned to be a crater was actually not a crater.

Detection results with confidence level
(Red circles: False positives detected)

MC-Dropout



Reason N times with different nodes to Dropout.



SoftMax

Obtain variance and mean of reasoning results

If the value of variance is large, the reasoning result may be wrong

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

Cause	Solution
Cause A: Insufficient training data	**Solution1**: **Prepare necessary and sufficient training and test data** using systematic context analysis methods

….. Difficult to prepare 100% of the necessary and sufficient data

Cause	Solution
Cause A: Insufficient training data	**Solution2: Identify the incorrect reasoning results**

….. Difficult to identify 100% of the incorrect reasoning results

Strategic Solutions to reduce uncertainty during reasoning

- Cause A: Insufficient training data

Cause	Solution
Cause A: Insufficient training data	**Solution1**: **Prepare necessary and sufficient training and test data** using systematic context analysis methods

….. Difficult to prepare 100% of the necessary and sufficient data

Cause	Solution
Cause A: Insufficient training data	**Solution2: Identify the incorrect reasoning results**

….. Difficult to identify 100% of the incorrect reasoning results

Cause	Solution
Cause A: Insufficient training data	**Solution3: Take measures at the system level**

Strategic Solutions for uncertainty during reasoning

Cause	Solution
Cause B: Including the incorrect training data	Solution4: Identify no mislabeled data and no non-inferred data included => Solution2-1: Identify **misalignment and inconsistency in data distribution**

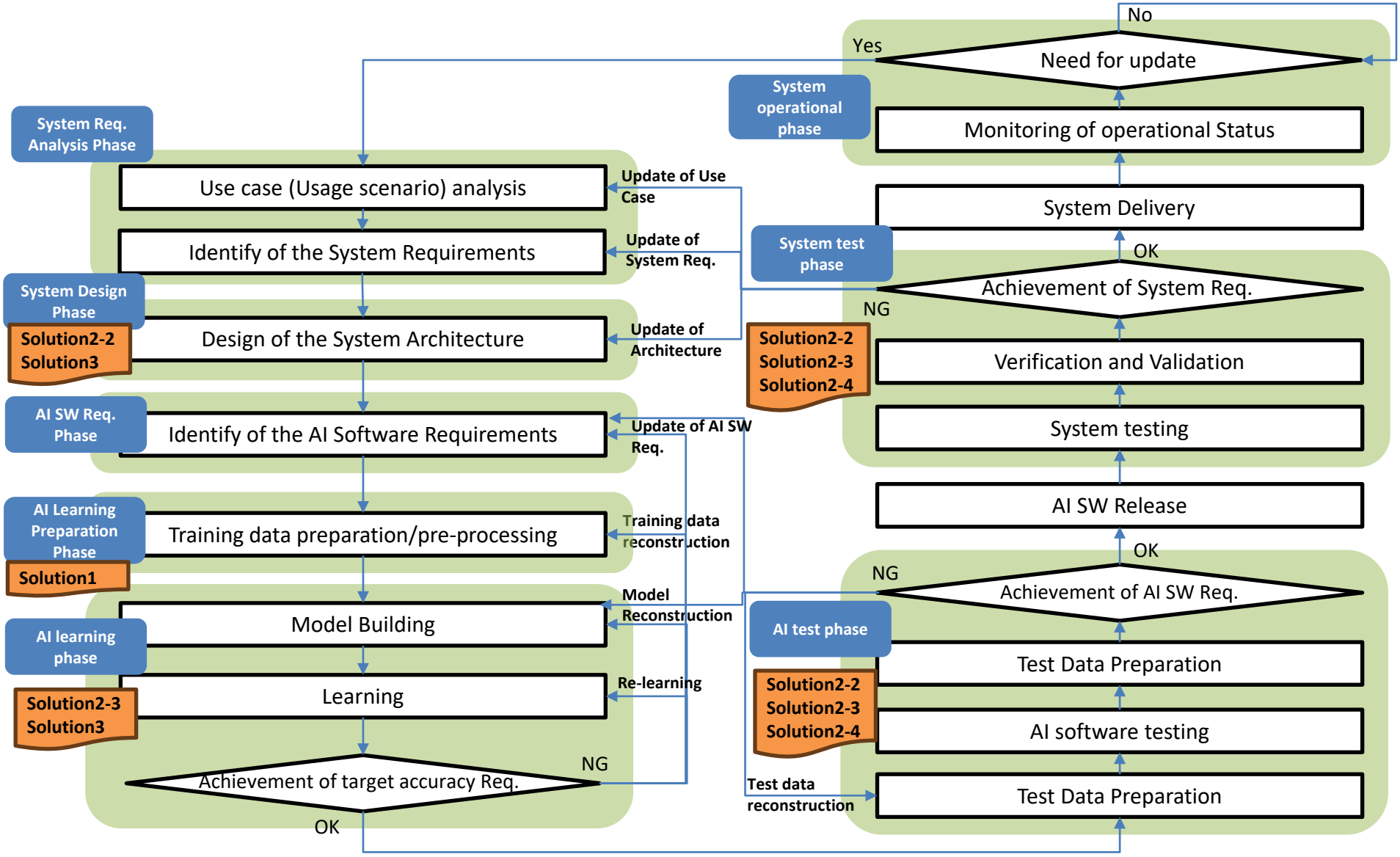
….. Difficult to identify 100% of the incorrect reasoning results

Cause	Solution
Cause B: Including the incorrect training data	**Solution3: Take measures at the system level**

Cause	Solution
Cause C: insufficient model	**Solution5: Rebuild the model to meet the requirement** => general method

Cause	Solution
Cause D: Incorrectly implemented	**Solution6: Review and test as legacy software development** => legacy method

AI system development process:
- development flow including requirements for avoiding false reasoning



AI system development process: (Example) System design phase

Activity Requirement:

- Architectural design shall be performed based on the system requirements and constraints.
- To identify functional and non-functional requirements for each sensor, actuator, computer system, and other hardware based on the driving scenario, and to design the placement of each sensor and actuator.
- The following requirements for learning and test data preparation shall be identified without omission.
  - Architectural design information (e.g., placement), Hardware characteristics, External environmental conditions affecting the hardware
- The architectural design shall meet safety requirements.

Requirement for uncertainty during reasoning:

- A function that can detect changes to input data shall be provided, since changes in usage scenarios, changes in identification targets, and changes in environmental conditions may occur after shipment.
- After shipment, a function shall be provided to monitor the status during operation and save data in which unexpected situations occur.
- AI software may output unexpected inference results, so the design shall take this into account.

Input:

- Use Case
- Identified target
- System Req. Specification

Output:

- Architecture Design Specification
- Sensor/Hardware specification

Summary

- Causes of False Reasoning of AI application
- Strategic Solutions for uncertainty during reasoning
- AI system development process
  - Activity Requirement
  - Requirement for uncertainty during reasoning
    - Including example method to identify the false reasoning
  - Input and Output


なぜ人は宇宙をめざすのだろう？
僕たちはいったいどこに向かうのだろう？
Thank you for your attention!
人類の
未知への挑戦を。
HUMAN POSSIBILITIES
さあ、次の冒険へ。
Our Story
HUMANS IN SPACE
はじめる
JAXA