

# Risk Assessment & Control Coordination for AI services : Case03 Power Line Inspection AI

Institute for Future Initiatives, The University of Tokyo Technology Governance Research Unit AI Governance Project



#### How to operate the RCModel - Risk Assessment & Control Coordination -





# Guide book and Case Studies of Risk Chain Model

AI Service and Risk Coordination Study Group

https://ifi.u-tokyo.ac.jp/en/projects/ai-service-and-risk-coordination/



Research Education People News Events Publications

#### How to use Risk Chain Model

Risk Chain Model (RCModel) Guide Ver1.0 최고

#### **Case Study**

\*These are fictional case studies below and don't raise issues or assure for any company or AI service.

Case01.Recruitment AI (2021/07) 译了



# Case Study



### **Case03 : Power Line Inspection AI**

- Define the "values and objectives to be achieved" by AI services -

This is a diagnostic service that automatically detects faulty parts on a transmission line using deep-learning image analysis technology.

Visual inspection of transmission lines is conducted using a high-magnification scope by maintenance personnel. However, for transmission lines in mountainous areas where visual inspection is difficult, skilled maintenance personnel are required to inspect moving images taken by a helicopter via slow playback, which requires long work time.

Against this background, AI services have been developed for automatic inspection of transmission lines and preparation reports. The image data were taken by a drone or helicopter. Although the determination is not made in real time, the faulty parts are predicted by the AI model immediately. Inspection reports are prepared automatically following the inspection of the AI system.

#### [Values & Objectives]

- Ensuring inspection quality (maintaining safety)
- Reducing the burden of inspection
- Corporate social responsibility

[Flow of Actual Operations Using AI Services]

- ① Drones capture the image data of transmission lines.
- ② The AI model predicts faulty parts via image analysis of the VTR data.
- ③ The AI system prepares automatic inspection reports, including analysis results.
- ④ The maintenance team checks the report.

The deep-learning AI model was developed by Co. X, built on its cloud environment managed. Co. X receives video data from Co. P, performs inference processing, drafts reports, and sends them to Co. P. The reports include images of transmission lines in areas that appear to be abnormal. Co. P checks the content and decides whether maintenance is necessary.

Note: although sequential learning is not performed, if the accuracy tends to deteriorate significantly after operation, Co. P will request Co. X perform additional learning or relearning as necessary.

## **Case03 : Power Line Inspection AI**

- System Overview -



## **Case03 : Power Line Inspection AI**

- Input & Output -

#### [Input Data]

Data	Purpos e	Collection Method	Data Manager	Including Privacy Data
Airborne data of power transmission lines (Video)		Co. P aerial photography with helicopters and drones	Co. X (Co. X Development Environment)	Yes (*)
Airborne data of power transmission lines (Video)	Product ion	Co. P aerial photography with helicopters and drones	Co. P (Co. P's internal management environment and Co. X's cloud environment)	Yes (*)

#### [Output]

[Output]	
Users	Power Service Maintenance Department, Co. P
Output	Abnormal locations of transmission lines (image)
Output Method	A report is automatically generated, and images (parts considered to be abnormal) are output at the locations where abnormalities are detected in the transmission lines
Expected Accuracy	False-negative: Within 1% *Probability of recognizing an abnormality as normal *However, false-positive results are also monitored (If it's too big, people will check it too often)
User judgment	Yes
Output of evidence information	Highlights of areas of concern are included in the report
Safety Risk	Yes (Power outage possible)
Connection with external system	No
Users	Power Service Maintenance Department, Co. P



# **Risk Assessment**



#### **Risk Assessment**

- Examination of significant risk scenarios -

	Values & Objectives		Service Requi	irement	Risk No.		Risk Scenario						
		1-1	Precision	Accuracy	R001	Unstable performance	Due to deterioration in AI's predictive performance, the fault parts are overlooked, and power cannot be supplied						
		T_T	performance	Robustness	R002	Impact of noise	Noise is mixed into the aerial image and the AI's judgment accuracy deteriorates						
	Inspection	1-2	Responding to changes in the	<ul><li>Robustness</li><li>IoT</li></ul>	R003	Changes in the external environment	Prediction performance is significantly reduced due to changes in the external environment, etc						
1	quality (Maintaining safety)		environment	- 101	R004	Changes in the IoT	Appropriate judgment cannot be made due to changes in image specifications (e.g., resolution, pixels, and format) due to changes in the imaging device						
	Salety)	1-3	1-3	Protection from external attacks	<ul><li>Robustness</li><li>Security</li></ul>	R005	Malicious contamination of abnormal data	Paint on power lines to make them misunderstand, and the AI changes the decision significantly					
													R006
		2-1	Appropriate inspection level	Accuracy	R007	Over detection	Over-identification of inspection points increases inspection costs						
2	Reduce burden for		Easy-to-	<ul> <li>Explainability</li> <li>Easy-to-</li> </ul>	R008	Appropriate judgment	If the judgment basis (marked part of the image) is not output, the person in charge cannot make judgment and explain, and appropriate maintenance cannot be performed						
	inspection	2-2	understand report	understand expression	R009	Excessive AI dependence	Depending on the content of the report, maintenance personnel may no longer doubt the results of AI decisions						
							Process	R010	Response to quality audits	Inability to provide appropriate explanations when subjected to quality audits			
3	Corporate social responsibilit	3-1	Accountability	description <ul> <li>Verifiability</li> </ul>	R011	Investigation at the time of trouble	When an external explanation is required due to the occurrence of abnormality or trouble, the cause and preventive measures cannot be considered and explained						
	У	3-2	Data Protection	Data protection	R012	Harmful rumors	Leakage of AI prediction (abnormality detection point) damages the reputation of a specific areas						

#### **Risk Assessment&Control Summary**

- Organize the roles of each stakeholder based on the examination of each risk chain -

	Values &	Risk	Dick Cooperie	Uncerta	Environ	Caused	RC		Control Summary	
	Objectives	No.	Risk Scenario	inly	mental change	by user	RC	AI System	AI service provider	User
		R001	Unstable performance	0			•	Ensuring a sufficient accuracy rate Recording the results of AI decisions	Confirmation of power supply status Periodic image verification relearning	Switching rule for alternative operation Alternative investment system
	Inspection	R002	Impact of noise	0	0		•	Camera specification definition Image noise correction Robustness of the model	Maintaining the shooting distance Image verification at the same point Relearning	
1	quality (Maintaining safety)	R003	Changes in the external environment	0	0		•	Development of multiple models Parallel execution of multiple models	Multiple model development system Performance monitoring	
		R004	Changes in the IoT	0	0		•	*Same as R002	*Same as R002	*Same as R002
		R005	Malicious contamination of abnormal data	0	0		•	*Same as R002	*Same as R002	*Same as R002
		R006	Security protection					Security management	Investigation and improvement of causes	
2	Reduce burden for	R007	Over detection	0			•	Ensuring a sufficient accuracy rate Overdetection warning	Defining Inspection Levels Reference information for the report Investigating the cause of overdetection	Rules for selecting inspection points Demand for model improvement
	inspection	R008	Appropriate judgment	0		0	•	*Include in the R009	*Include in the R009	*Include in the R009
		R009	Excessive AI dependence	0		0	•	Output of the rationale	The basis for judgment and reference information are included in the report.	Ensuring knowledge of inspection services Consideration of misjudgment
3	Corporate 3 social	R010	Response to quality audits	0			•	Record data comprehension Record the performance of the model Recording the results of AI decisions	Internal audit Response to external audits Access control	
	responsibility	R011	Investigation at the time of trouble	0				Save log data	System operation monitoring Fault handling	Manual alternative operation
		R012	Harmful rumors					Data protection	Education on professional ethics	Education on professional ethics

Step5

### Organization

- Organize the roles of each stakeholder based on the examination of each risk chain -

Co. P) Top Management			Ministry of Land, Infrastructure, Transport and Tourism
<ul> <li>Values and Objectives</li> <li>Approval risk controls</li> </ul>			
Co. P) Legal Dept.	Co. P) Internal Audit Dept.		Electric power company
	<ul> <li>Internal audit</li> <li>Response to external audits</li> </ul>		
- AI Service Provider - Co. P) Power Service Maintenance Dept.	Co. X) AI Dev Dept.	Co. P) Power Line Management Dept.	- User - Co. P) Maintenance Staff
<ul> <li>Confirmation of power supply</li> <li>Defining inspection levels</li> <li>Maintaining the shooting distance</li> <li>Periodic image verification</li> <li>Image verification at the same point</li> <li>The basis for decision and reference information are included in the report.</li> <li>Multiple model development system</li> <li>Performance monitoring</li> </ul>	<ul> <li>Securing the accuracy rate of AI models</li> <li>Camera specification definition</li> <li>Image noise correction</li> <li>Robustness of the model</li> <li>Decision basis output of the model</li> <li>Alert upon overdetection</li> <li>Development of multiple models</li> <li>Parallel execution of multiple models</li> </ul> Co. X) Cloud Service Dept.	Information coordination of power supply status	<ul> <li>Knowledge and skill for inspection</li> <li>Consideration of misjudgment</li> <li>Inspection rules for overdetection</li> <li>Switching rule and system for alternative operation</li> <li>Demand for model improvement</li> </ul>
<ul> <li>Relearning</li> <li>Response to internal audits</li> </ul>	<ul> <li>Recording the results of AI decisions</li> <li>Save log data</li> </ul>		

Save log data

Step5





- Examine the risk chain (relation of risk factors) for each important risk scenario -

R001

#### Maintenance of predictive performance

Due to deterioration in AI's predictive performance, the fault parts are overlooked, and power cannot be supplied



- Consider risk control according to the risk chain -

R001

#### Maintenance of predictive performance

Due to deterioration in AI's predictive performance, the fault parts are overlooked, and power cannot be supplied

	Risk Control	
AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
<ol> <li>[Accuracy] Ensuring a sufficient accuracy rate of models when learning (AI Dev Dept., Co. X)</li> <li>[Traceability] Stores information on AI judgment results when using (Cloud Service Dept. Co. X)</li> </ol>	<ul> <li>③ [Safety] Obtain regular information from the power transmission management division on whether there is any abnormality in the power supply (Power Line Management Dept., Co. P / Power Service Maintenance Dept., Co. P)</li> <li>④ [Auditability] Periodically and comprehensively verify aerial images to make sure that AI does not overlook areas to be inspected (Power Service Maintenance Dept., Co. P)</li> <li>⑤ [Agility] Asks for relearning of AI models to ensure sufficient accuracy (Power Service Maintenance Dept., Co. P / AI Dev Team, Co. X)</li> <li>⑥ [Consensus] In case that relearning cannot be done in time, an alternative operation by human should be decided (Power Service Maintenance Dept., Co. P).</li> </ul>	<ul> <li>⑥ [User Responsibility] In case that relearning cannot be done in time, make arrangements for manual alternative operation (Maintenance Staff in Co. P)</li> <li>⑦ [Controllability] Control system to confirm aerial images manually when relearning is not in time (Power Service Maintenance Dept., Co. P / Maintenance Staff in Co. P)</li> <li>⑧ [Proper Use] When relearning cannot be done in time, substitute operation by human (Maintenance Staff in Co. P)</li> </ul>

- Examine the risk chain (relation of risk factors) for each important risk scenario -

R002

#### Impact of noise

Noise is mixed into the aerial image and the AI's judgment accuracy deteriorates



- Consider risk control according to the risk chain -

R002 <b>Impact of noise</b> Noise is mixed into the aerial im	age and the AI's judgment accuracy deterior	ates
	Risk Control	
AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
② [Stability] Clarify the required specifications of the camera and perform regular maintenance (AI Dev Dept., Co. X /Power Service Maintenance Dept., of Co. P)	① [Sustainability] Maintains the shooting distance and speed of helicopters and drones so that the image information does not change significantly depending on the shooting date (Power Service Maintenance Dept., Co. P)	
<ul> <li>③ [Data Quality] Degradation of image data by noise correction, etc (AI Dev Dept., Co. X)</li> <li>④ [Robustness] Learning to enhance model</li> </ul>	<ul> <li>② [Auditability] Compare the image information of the judgment result of the same point, and confirm whether the focus of the image has</li> </ul>	
robustness (AI Dev Dept., Co. X) (5) [Interpretability] Outputs the judgment basis of the model (points of interest in the image) (AI Dev Dept., Co. X)	<ul> <li>clearly changed (Power Service Maintenance Dept., Co. P)</li> <li>(8 [Agility] Asks for relearning of AI models to ensure sufficient accuracy (Power Service Maintenance Dept., Co. P / AI Dev Team, Co. X)</li> </ul>	

6 [Traceability] Stores information on AI judgment results when using (Cloud Service Dept., Co. X)

Step4

- Examine the risk chain (relation of risk factors) for each important risk scenario -

R003

#### Changes in the external environment

Prediction performance is significantly reduced due to changes in the external environment, etc



- Consider risk control according to the risk chain -

R003

#### Changes in the external environment

Prediction performance is significantly reduced due to changes in the external environment, etc

	Risk Control	
AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
<ul> <li>③ [Capability] Prepare a development environment for each model (AI Dev Dept., Co. X)</li> <li>④ [Data Balance] Prepare sufficient learning data for each model (AI Dev Dept., Co. X)</li> <li>⑤ [Accuracy/Robustness] Learning to ensure predictive accuracy and robustness of models (AI Dev dept., Co. X)</li> <li>⑥ [Availability] Parallelizing and predicting multiple models at the same location (For example, a model developed for each season and weather is used to predict abnormal locations at the same location) (AI Dev Dept., Co. X)</li> <li>⑦ [Traceability] Stores information on AI judgment results when using (Cloud Service Dept., Co. X)</li> </ul>	<ul> <li>(Power Service Maintenance Dept., Co. P)</li> <li>② [Scalability] Ensuring sufficient development structure for service provision according to the number of models, etc (Power Service Maintenance Dept., Co. P / AI Dev Team, Co. X)</li> <li>⑧ [Auditability] Periodically verify the performance of each model (Power Service Maintenance Dept., Co. P)</li> <li>⑨ [Agility] Asks for relearning of AI models to ensure sufficient accuracy (Power Service Maintenance Dept., Co. P / AI Dev Team, Co. X)</li> </ul>	

- Examine the risk chain (relation of risk factors) for each important risk scenario -

R007

#### **Over detection**

Over-identification of inspection points increases inspection costs



- Consider risk control according to the risk chain -

R007	<b>Over detection</b> Over-identification of inspection	points increases inspection costs Risk Control	
	AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
accuracy (A	I Dev Dept., Co. X)	inspection level (Power Service Maintenance Dept.,	⑥ [User Responsibility] Understand that the final inspection point is determined by the inspection team (Power Service Maintenance Dept., Co. P)
4 [Process ]	Integrity] Display an alert when an		

Co. P)	team (Power Service Maintenance Dept., Co. P)
(4) [Process Integrity] Display an alert when an	
excessive number of inspection points are 2 [Sustainability] Review appropriate inspection	⑦ [Controllability] Rules for selecting final
detected (AI Dev Dept., Co. X) levels (model performance) based on	inspection points (Maintenance Staff in Co. P)
maintenance costs (Power Service Maintenance	
Dept., Co. P)	8 [Self-Defense] Require model improvements if
	too many inspection points are detected
⑤ [Understandability] Information such as the	(Maintenance Staff in Co. P)
inspection history of inspection points and the	
weather since the last inspection is output on the	
report (Power Service Maintenance Dept., Co. P)	
(9) [Auditability] Verify the performance of each	
model and confirm the reason for over-detection	
(Power Service Maintenance Dept., Co. P)	
(1) [Acility/] Acive for released at models to	
① [Agility] Asks for relearning of AI models to	
ensure sufficient accuracy (Power Service Maintenance Dept., Co. P / AI Dev Team, Co. X)	
Maintenance Dept., Co. P / AI Dev Tealit, Co. X)	

- Examine the risk chain (relation of risk factors) for each important risk scenario -

R009

#### Excessive AI dependence

Depending on the content of the report, maintenance personnel may no longer doubt the results of AI decisions



- Consider risk control according to the risk chain -

#### R009

#### **Excessive AI dependence**

Depending on the content of the report, maintenance personnel may no longer doubt the results of AI decisions

	Risk Control	
AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
① [Interpretability] Outputs the judgment basis of the model (AI Dev Dept., Co. X)	② [Understandability] Outputs reference information such as judgment grounds of inspection points and inspection history on the report (Power Service Maintenance Dept., Co. P)	<ul> <li>③ [User Responsibility] Understand that the final inspection point is determined by the inspection team (Power Service Maintenance Dept., Co. P)</li> <li>④ [Expectation] Understanding the accuracy of model judgments (Power Service Maintenance Dept., Co. P)</li> <li>⑤ [User Ability] Education of knowledge and skills necessary to determine inspection points (Power Service Maintenance Dept., Co. P / Maintenance Staff in Co. P)</li> <li>⑥ [Awareness] Finding places where AI seems to have made mistakes (Maintenance Staff in Co. P)</li> <li>⑦ [Proper Use] Determine appropriate inspection points (Maintenance Staff in Co. P)</li> </ul>

- Examine the risk chain (relation of risk factors) for each important risk scenario -

R010

#### Response to quality audits

Inability to provide appropriate explanations when subjected to quality audits



- Consider risk control according to the risk chain -

R010

#### Response to quality audits

Inability to provide appropriate explanations when subjected to quality audits

Risk Control		
AI System (AI dev dept., Co. A)	AI Service Provider (HR dept., Co. A)	User (Person in HR dept., group A)
<ul> <li>② [Data Quality/Data Balance] Record data distribution and understanding (AI Dev Dept., Co. X)</li> </ul>	① [Consensus] Organize the division of roles and information to be disclosed when responding to audits (Power Service Maintenance Dept., Co. P)	
③ [Accuracy/Robustness/Generalization] [Interpretability] Record Model Performance (AI Dev Dept., Co. X)	⑤ [Auditability] Conduct internal audits and respond in advance (Internal Audit Dept., Co. P)	
④ [Traceability] Record the results of AI decisions (Cloud Service Dept., Co. X)	⑥ [Transparency] Organize information to be disclosed to external auditors (Power Service Maintenance Dept., Co. P)	
	⑦ [Accessibility] Set necessary access rights for external auditors (Power Service Maintenance dept., Co. P)	
	⑧ [Accountability] Disseminate appropriate response to external audits (Power Service Maintenance Dept., Co. P)	
	⑨ [Sustainability] Responding to issues discovered during the audit (Power Service Maintenance Dept., Co. P)	