February, 2021

JDLA

Study Group 'AI governance and its Evaluation'

Report on the Session #7

1. Introduction

The Japan Deep Learning Association establishes study groups as a forum for deepening knowledge and discussing domestic and international policy trends related to artificial intelligence (hereafter AI) and Deep Learning (hereafter DL). This study group, "AI Governance and its Evaluation," defines "governance" as a system of management and evaluation by various actors, and launched a study group in July 2020 to investigate what forms of governance are possible and conduct a year-long study to help build trustworthy AI systems.

In the 7th session (December 10, 2020), Mr. Hiroyuki Yamagishi and Mr. Masaru Ogawa, Engineering Planning Division, Railway Bureau, Ministry of Land, Infrastructure, Transport and Tourism (MLIT), and Mr. Takeshi Motomura, Iwata Godo Attorneys and Counsellors at Law ("Iwata Godo"), presented topics under the theme of Accident Investigation and Third-party Investigation Committees for AI governance.

This report is a reconstruction of the topical presentations and the discussions of the study group participants.

2. Status of studies on accident prevention for unmanned, automated railway track operations

In the first part, Mr. Hiroyuki Yamagishi presented a topic titled "Study on accident prevention of unmanned and automated railway track operation".

Accidents caused by Automated Train Operated by Yokohama Seaside Line[1] Went the Wrong Way

The Kanazawa Seaside Line[2], operated by Yokohama Seaside Line, is an unmanned, automated train service. On June 1, 2019, one of their automated trains went opposite direction to the one it was supposed to go when turning around at the last station, and collided with the stop sign at the end of the line, and stopped. In response to the accident, the Japan Transport Safety Board (JTSB) investigated the cause of the accident, while the Committee on Accident Prevention for Unmanned Automatic Railways (hereinafter

[1] https://www.seasideline.co.jp/company/company_profile/

[2] A railway line connects Shin-Sugita Station in Isogo-ku, Yokohama City, Kanagawa Prefecture, with Kanazawa-Hakkei Station in Kanazawa-ku, Yokohama City, Kanagawa Prefecture.

referred to as the Accident Prevention Committee) was taking the lead in preventing similar accidents and studying measures to prevent recurrence.

The Accident Prevention Committee was established on June 14, 2019, and is composed of experts, researchers, and companies engaged in unmanned automated railway operation. The Committee was held three times so far[3], to share information on the cause of the accident among the parties involved, and to consider measures to prevent the recurrence of similar accidents (see Fig. 1). On July 19, 2019, the following items were reported by the Committee as the results of the interim summary, regarding the confirmation of the presumed cause of the accident and the measures to prevent recurrence.

- ✓ **The presumed cause of the accident**

  The Accident Prevention Committee received explanations on the status of the investigation from JTSB and Yokohama Seaside Line, and confirmed the presumed cause of the reversal was due to the disconnection. In addition, the safety of this presumed causal event was also verified in other operators as well.

  As a result, the Committee confirmed that the system of the other operators does not allow the train to run even in the event of a disconnection.

- ✓ **Measures to Prevent Recurrence on the Yokohama Seaside Line**

  The recurrence preventive measures, as confirmed by the Accident Prevention Committee, are as follows.

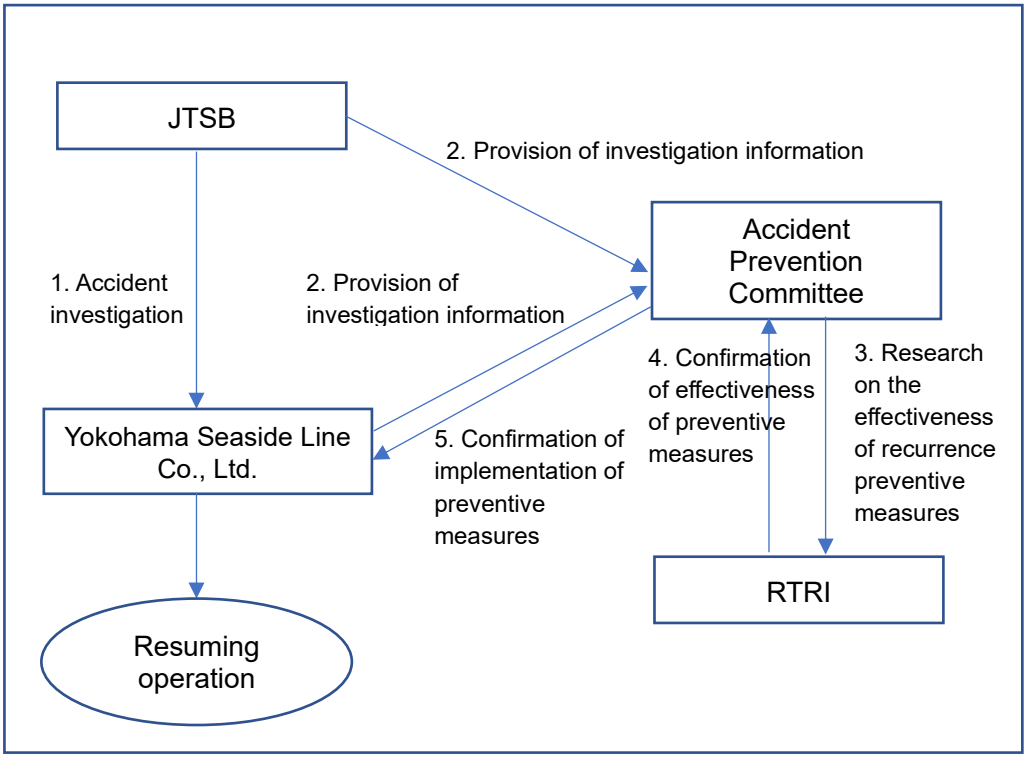
  - If the command line that conveys the direction of travel to the motor control unit becomes unpressurized due to a broken wire, etc., the train will not depart, and the circuit shall be changed so that the Automatic Train Operation (ATO) system on the ground side can detect the broken wire, etc.
  - Modify the software for automated operation on the mechanism of motor force control.
  - Modify the software so that the emergency brake operates when the Automatic Train Stop (ATS) system becomes unpressurized due to circuit breakage, etc.
  - To eliminate the risk of reverse running due to malfunction of the relay for backward correction of stopping position, the relay shall be removed and the position shall be corrected manually.

The Accident Prevention Committee determined that the above-mentioned recurrence

[3] as of December 10, 2020

preventive measures to be effective from two perspectives: "The results of the investigation by Railway Technical Research Institute (RTRI) showed that the measures were effective," and "The measures have a proven track record, as similar measures have already been taken by other operators in the same field". The Yokohama Seaside Line underwent repair work in accordance with the recurrence preventive measures identified by the Accident Prevention Committee on July 19, 2019. After these preventive measures were appropriately taken and safety was confirmed, a decision was made on whether to resume automated operation, and they resumed their operation on August 31, 2019[4].

Figure 1: Overview of the process for investigation of the Yokohama Seaside Line reversing accident and consideration of recurrence preventive measures[5]



3. Study Group on Automated Train Operation Technology

In the second part, Mr. Masaru Ogawa gave a presentation titled "Study Group on Automated Train Operation Technology: FY2019 Summary".

---

[4] Yokohama Seaside Line Co., Ltd., "Press Release by Yokohama Seaside Line Co., Ltd." (August 27, 2019)

[5] Compiled from "Progress Report on the Investigation of Railway Personal Injury Accidents" (February 27, 2020) by the JSTB, MLIT, and the contents of the announcement as shown in footnote 2.

Status of Study on Automated Train Operation Technology

Currently, Japan is facing a depopulating society, making it difficult to secure train operators and maintenance personnel for the railway businesses. Especially in rural area, the shortage of staff is a serious problem, and there is a need to introduce automated train operation that does not require an operator to be on board in order to maintain the regional railway business. One of the issues in introducing automated train operation is the form of operation. Conventional automated train operation has been realized in new transportation systems that were originally built on the premise of automated operation, such as elevated structure that prevent people and others from easily entering the line. However, in order to spread the use of automated train operation in the future, existing railways will have to be converted to automation, but the structure and location of general railway[6]make it difficult to introduce automated operation from the perspective of safety and stable transportation. Especially for regional railways, it is difficult to make large-scale capital investments for the introduction of automated operation, and it is necessary to study the technical requirements for automated operation on general lines.

The above-mentioned issues and background led to the establishment of the "Study Group on Automated Train Operation Technology" (hereinafter referred to as the "Automated Operation Study Group"), which has been held six times in the past, starting with the first session held on December 3, 2008. The Automated Operation Study Group has been studying technical requirements[7] for maintaining and improving safety and convenience for the introduction of automated train operation, which will also contribute to the productivity revolution in the railway sector, for general lines, broadly classified into urban and regional railways.

A Model Case Study of Automated Urban Railways

In examining a model case for automated operation in urban railways, the conditions of the rail line section and the requirements of Grade of Automation (GoA) are defined as preconditions (see Fig. 2).

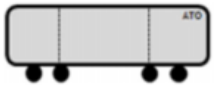
Figure 2: Requirements of Grade of Automation (GoA)[8]

---

[6] The "general railways" referred to in this study group are those that are (1) level (non-elevated structure), (2) with level crossings, and (3) without platform doors.

[7] The Automated Operation Study Group is considering the use of the latest technologies such as sensor technology, ICT, and wireless train control technologies, etc.

[8] Excerpted from "Study Group on Automated Train Operation Technology: FY2019 Summary" (July 3, 2020). The level of automated train operation are defined by the IEC as Grade of Automation. Automation level "GoA1" is non-automated train operation, "GoA2" is semi-automated train operation,

Automation levels (defined by AUGT standard)	Image of the crew system (Descriptions in [ ] is the main work of the attendant.)	Introduction Status in Japan
GoA0 On Sight Train Operation (TOS)	 An operator (and conductor)	Tramways
GoA1 Non-automated Train Operation (NTO)		General lines with level crossings
GoA2 Semi-automated Train Operation (STO)	 An operator [Train start-up, door handling, emergency stop operation, evacuation guidance, etc.]	Certain subways, etc.
GoA2.5 Automated operation with attendant to perform emergency stop operations, etc.	 An attendant in the cab of the first car [Emergency stop operation, evacuation guidance, etc.]	N/A
GoA3 Driverless Train Operation (DTO)	 An Attendant [Evacuation guidance, etc.]	Certain monorails
GoA4 Unattended Train Operation (UTO)	 No Attendant on board	Certain new transportation, etc.

As the conditions of the rail line section, the line must be double-track, use Automatic Train Control (ATC)[9] and Automatic Train Operation (ATO)[10] systems, and have movable platforms, railway crossing (Type 1), tunnels and bridges. As for the grade of automation, the study considers the realization of automated operation by raising the preconditions step by step, assuming that the stage where an attendant will be on the operator's cab of the first car (GoA2.5), the final stage will be the realization of Driverless Train Operation (DTO)[11](GoA3) or Unattended Train Operation (GoA4).

One of the main considerations at this time for GoA3 or GoA4 is to deal with obstacles on the tracks, and since there is no attendant in the cab of first car, the following was summarized.

- ✓ Besides platforms and railway crossings, strengthen the protective fences

"GoA3" is driverless train operation with an attendant, and "GoA4" is unattended automated train operation. In addition, although not defined in IEC and JIS, automated operation with an attendant who performs emergency stop operations, etc., is classified as "GoA2.5".

[9] ATC (Automatic Train Control) is a device that continuously transmits signals and speed information from the ground to the train, and ensures a high level of safety by constantly checking the train speed on the train side and automatically activating the brakes, or relaxing the brakes when the speed drops below the speed limit. It is a device that ensures a high level of safety.

[10] ATO (Automatic Train Operation) is a device that automatically performs train departure, acceleration, deceleration, and fixed position stop control.

[11] DTO stands for Driverless Train Operation.

between stations to prevent people, vehicles, cargo, and other obstacles from entering the tracks.

- ✓ To detect and respond to obstacles in front by means of sensors, etc., installed on the trains.
- ✓ For railway crossings, measures should be taken to improve visibility to motorists, detect abnormalities in the control rods, detect obstacles, provide push-buttons to warn of obstacles, and stop trains when abnormalities occur.
- ✓ For the platform section, prevent obstructions and ensure safety by installing movable platform fences on the platform bridge on the track side, measures to prevent falls from the edge of platforms other than the track side, and devices to prevent falls between trains and platforms.

The following issues will be further discussed in the future.

- ✓ Requirements for protective fences, such as height and structure
- ✓ Requirements for frontal obstacle detection system, such as the size of the obstacle that needs to be detected, detection distance, etc.
- ✓ Actions and necessary functions at the railway crossing after detection of abnormalities
- ✓ How to respond to derailments and train collisions between inbound and outbound lines
- ✓ What to do in case of a fire on the platform or in the car.
- ✓ Evacuation guidance methods
- ✓ Division of roles and duties between commanders and operators/attendants
- ✓ Qualification requirements for operators/attendants

etc.

A Model Case Study of Automated Train Operation in Regional Railways

As in the model case for urban railways, preconditions have been defined for regional railways. The conditions for rail line section are that they use single track and ATS or ATO system, that they do not have movable platform fences, and that they have railway crossings, tunnels, and bridges classified as Type 1, 3, or 4. As for the Grade of Automation, the realization of GoA2.5 is being considered, in which an attendant is assigned to the operator's cab of the first car to perform emergency stop operations and evacuation guidance in case of emergency.

The main considerations at this time for GoA2.5 are the system requirements and the handling of obstacles on the railway tracks. The system requirements were summarized as follows.

- ✓ To ensure the safety of the entire system by adding security functions to the ATO system in addition to the pattern-type ATS system and making it more sophisticated
- ✓ Control by ATO system to avoid exceeding the maximum speed or speed limit

It was also decided to take the following actions regarding obstacles on the railway tracks.

- ✓ When an obstacle is found on the track, the GoA2.5 personnel shall press the emergency stop button to make an emergency stop.
- ✓ In the event of a train derailment or other failure that requires the stopping of a train, GoA2.5 personnel shall use signal flame tubes, train protection radios, etc. to immediately stop the train involved.

Other responses are also expected to include the need for GoA2.5 personnel to evacuate passengers in the event of an accident or disaster.

The following issues will be further discussed in the future.

- ✓ Details about the work to be done by GoA2.5 personnel, including whether a train driver's license is required and the use of directives.
- ✓ A function to move the train at low speed to a safe place when the train stops at a point where evacuation guidance is impeded, such as in a tunnel or on a bridge.
- ✓ In the case of a system that cannot obtain control information before passing the first ATS ground station, such as when departing from a station, how to ensure safety by the system until the control information is received.

4. Third-party Investigation Committees (Third-Party Committees) in corporate scandals

In the third part, Mr. Takeshi Motomura presented a topic titled "Practices and Issues of Third Party Committee (Investigation of Misconduct)".

Outline of the Third-party Committee[12]

In Japan, Third-party Committee is mentioned in the guideline publicized by Japan Federation of Bar Associations (hereinafter referred to as JFBA), but the instructions are not legally enforceable and there is no legal basis for them. However, JFBA Guidelines

[12] The "Third-party Committee" referred to in this study group refers to an entity that is required to play the role of finding the root causes of an event by conducting an investigation from a neutral and independent standpoint that has no previous involvement with the parties involved, including the requesting party, and to present improvement measures, including recurrence preventive measures at the requesting organization.

are usually followed when companies establish an investigation committee under the name of a “Third-party Committee” in case of scandals such as accounting fraud, data falsification, or quality perjury, and formulate measures to prevent recurrence. Third-party Committees are also sometimes formed under the names of “External Investigation Committees” or “Special Investigation Committees”. In JFBA Guidelines, it is positioned as a committee that "investigates facts, finds facts of frauds/wrongdoings, and analyses the causes," and is distinguished from the Legal Liability Assessment and Pursuit Committee and the Management Advisory Committee.

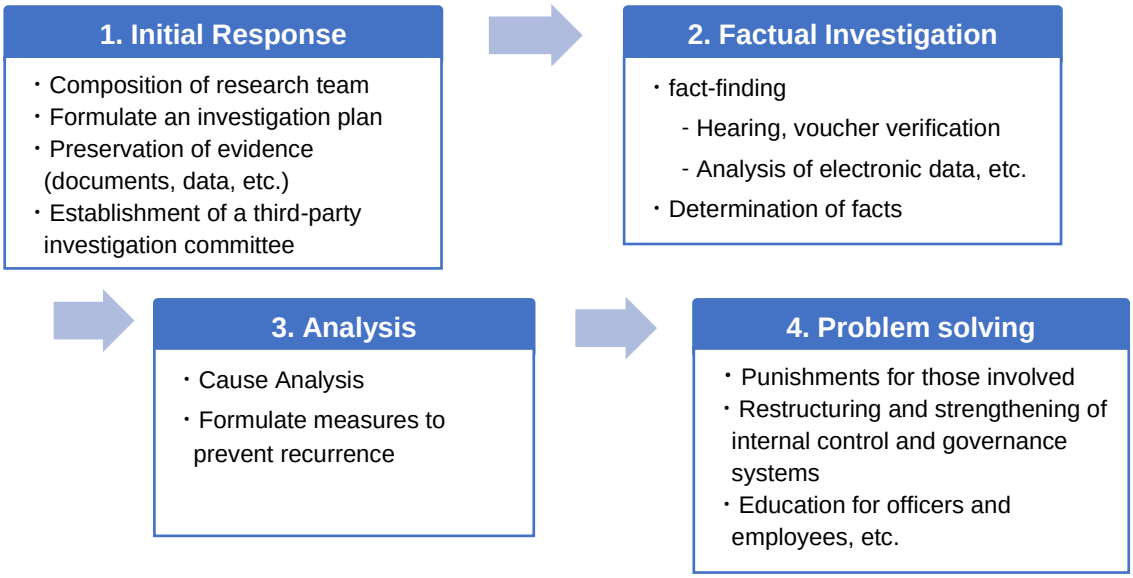
Flow of response to the discovery of misconduct and injustice

A Third-party Committee is formed as an initial response when misconduct or injustice is uncovered through internal or external whistleblowing. During the initial response, an investigation team is formed by lawyers, accountants and etc., to formulate an investigation plan. The preservation of evidence such as documents and data is also carried out at the same time.

After the investigation plan is formulated, a fact-finding investigation is conducted, which includes interviews with the parties involved, analysis of electronic data (digital forensics), analysis of accounting information, and verification of relevant vouchers, to determine the facts such as the amount of damage and correction, the parties involved, and the impact.

After a factual investigation, the causes are analyzed and measures to prevent recurrence are discussed, and actual preventive measures are formulated, including disciplinary action against those involved, such as management changes, restructuring of internal controls, strengthening of governance systems, and education of executives and employees.

Figure 3: Flow of response to the discovery of misconduct and injustice[13]



In parallel with the above-mentioned response process, the company will also deal with its shareholders, business partners, affiliates, regulatory authorities, investigative agencies, audit firms, and other stakeholders, but the entire process will be carried out in a short period of time, and after the recurrence preventive measures formulated by the Third-party Committee are implemented, the management of the business will return from "contingency" risk management to "normal" risk management, and normal business activities can be restored.

Issues for Third-Party Committees

In the practice of Third-party Committees, independence is a very important aspect. First of all, the independence of a Third-party Committee can be broadly classified into formal independence and substantial independence.

- ✓ The issue of formal independence is the issue of the qualifications of the committee members, i.e., "what kind of committee members have independence and neutrality," and also the issue of the method of appointment, i.e., "who has the authority to appoint the committee members?".
- ✓ The issue of substantial independence involves the issues of impartiality, objectivity, and neutrality of the investigation content, which is how to achieve an investigation that ensures independence. Specifically, it is a question of what matters should be included in the contract to ensure the investigation, whether

[13] Prepared from materials released by this study group

the report should be disclosed in advance and the opportunity for rebuttal should be given, and to what extent the Third-party Committee should be obligated to explain the results and method of the investigation.

In addition to the "issue of independence" required for the investigation by the Third-party Committee, there are various other issues such as "the issue of publication of the investigation report and waiver of attorney-client privilege", "the issue of non-sharing of relevant materials prepared and collected evidence in the investigation", "holding those involved accountable", and "the role of outside directors". These are issues that need to be considered in the future.

5. Discussion points in the question and answer session

In the 7th session, AI and accident investigation/Third-party Committee was discussed. The following questions and answers were raised based on the topics discussed.

- Transformations and new perspectives arising from the advancement of unmanned technologies
  - ✓ As unmanned technology progresses and systems become more complex, it is possible that the conventional way of managing organizations will be transformed, including the transformation of the personnel who participate, such as business operators, local governments, and experts.
  - ✓ The issue has not yet reached the point of automated train operation using AI at this point though, it is necessary to examine the boundary of responsibility for cases in which deep learning is used for obstacle detection.
- Why the Yokohama Seaside Line accident could not be prevented
  - ✓ One of the reasons why the Yokohama Seaside Line accident could not be prevented was that the design concept was left as it was during manned operation, and as a result, the company was not able to take measures to prevent accidents, which other operators were able to do.
- New problems arising from accident preventive measures
  - ✓ In the case of AI, overfitting may occur as a result of measures specific to the cause of the accident, which may cause further incidents.
- Assessing the safety of automated operation
  - ✓ There is a need to ensure the same or higher level of safety for automated train operation as for conventional general manned train operation.
- Scope of injustice to be covered by the Third-party Committee
  - ✓ Systematic injustice and harassment are broadly subject to investigation by the

Third-party Committee; in the case of AI-related issues, systematic harassment due to the use of AI is subject to investigation by the Third-party Committee.

- ✓ If incidents arise as a result of failure to detect AI problems or failure to stop the release of AI products, investigation by the Third-party Committee may be applied to the entire development organization or to the business using the AI.
- ✓ In cases where the responsibility is unclear, such as collusion between AIs, it is possible that only a Third-party Committee is likely to be able to clarify the issue.

➢ Investigation methods of Third-party Committees and use cases of AI

- ✓ In the case of whistleblowing, not only the reporting itself, but also collecting data through monitoring will provide evidence. In future cases where AI will be used, such as for preventive purposes such as investigations, or in the accumulation and acquisition of information to the extent that it does not become surveillance, the privacy of employees, staff and customers may become a point of contention.
- ✓ As for the use of AI for employee email review during fraud investigations, it is not a problem if the review is conducted as part of preventive and detective controls, just like the traditional human review, because the labor contract allows for verification of e-mails during work hours. However, as for personal cell phones, it is necessary to obtain the written consent of the individual before conducting the email review because of privacy issues.

➢ Incentives for cooperation with Third-party Committees

- ✓ As an incentive for cooperation in investigations, it is possible for the Third-party Committee to reduce the range of incentives allowed under the guidelines, but since criminal penalties are the purview of the police and prosecutors, providing incentives for the purpose of cooperation in investigations is restricted.

We will continue to discuss AI governance in Japan and abroad through this study group.

Written by Ayuto Makiguchi
Translated by Michiko Shimizu

<Outline of the 7th Session of the Study Group>

Date & Time: Thursday, December 10, 2020, 16:00-18:00 (Zoom)

Agenda:

- Topical presentations:
  - "Study on accident prevention of unmanned and automated railway track operation" provided by Mr. Hiroyuki Yamagishi (Engineering Planning Division, Railway Bureau, Ministry of Land, Infrastructure, Transport and Tourism (MLIT))
  - "Study on Automated Train Operation Technology " provided by Mr. Masaru Ogawa (Engineering Planning Division, Railway Bureau, Ministry of Land, Infrastructure, Transport and Tourism (MLIT))
  - "Practices and Issues of Third Party Committee (Investigation of Misconduct)" provided by Mr. Takeshi Motomura, Iwata Godo Attorneys and Counsellors at Law ("Iwata Godo")
- Question and answer session / discussion