Machine-in-the-Loop Process in Project Risk Management

Toshiki Mori, *Toshiba Corp.,* Naoshi Uchihira, *JAIST*

Abstract— **Despite the widespread recognition of the necessity and importance of project risk management, its effective implementation is often difficult. Much of the difficulty could be explained by the effects of cognitive biases related to heuristics when humans make decisions about probability and frequency. In this study, we propose a machine-in-the-loop process in which human and AI model cooperate in a complementary manner to mitigate the influence of cognitive bias in human decision-making and compensate for the lack of domain knowledge in prediction with the AI model. In a case study where the machine-in-the-loop process was applied to project risk management, we conducted an interview survey and confirmed the effectiveness of the machine-in-the-loop process with positive comments which support the reduction of uncertainty and cognitive bias.**

Index Terms*—*Artificial Intelligence, Machine Learning, Project Risk Management, Cognitive Bias, Dual System Theory.

I. Introduction

IN recent years, companies are facing drastic changes in the surrounding environment with increasing uncertainty due to rapid technological progress in the Internet and AI (Artificial Intelligence), diversification of customer needs, etc. Under these circumstances, in order for companies to establish a sustainable competitive advantage, it is necessary to build the ability to flexibly adapt in response to the changes by further strengthening project risk management. However, despite the widespread recognition of the necessity and importance of project risk management, its effective implementation is often difficult. In many organizations, the process has become a skeleton and is not operated correctly, resulting in repeated failures such as risks that later manifest themselves as major problems. According to the results of a survey on the maturity level of project management, the maturity level of risk management is the lowest among the nine areas of project management, and is reported to be the bottleneck of the project [1] [2].

Much of the difficulties in project risk management could be explained by the effects of cognitive biases related to heuristics when humans make decisions about probability and frequency. To reduce the impact of cognitive biases in human decision making, the complementarity between human and AI model can be utilized. In this study, we propose a machine-in-the-loop process in which human and AI model cooperate in a complementary manner. The AI model can help to mitigate the influence of cognitive bias in human decision-making, whereas human can help to compensate for the lack of domain knowledge in prediction with the AI model. We have applied the machine-in-the-loop process to project risk management and examined its effectiveness through interviews with practitioners.

This paper is organized as follows. In Section 2, we discuss the difficulties of project risk management. In Section 3, we propose a machine-in-the-loop process and its application to project risk management. Section 4 presents a case study of the machine-in-the-loop process as applied to project risk management, and summarizes the results of the interview survey. In Section 5, we survey related studies. Finally, Section 6 presents the conclusions.

Toshiki Mori is with Corporate Software Engineering & Technology Center, Toshiba Corporation, Kawasaki, Kanagawa, Japan. e-mail: toshiki1.mori@toshiba.co.jp, toshiki.mori@gmail.com

Naoshi Uchihira is with School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, Japan. e-mail: uchihira@jaist.ac.jp

II. Difficulties in Project Risk Management

The goal of project risk management is to identify and control risks that may lead to project failure in the early phases of the project. A risk is defined as an event or condition whose occurrence is uncertain and, if it occurs, will have a detrimental effect on the project. In Project Management Body of Knowledge (PMBOK) [3], a guideline developed by Project Management Institute (PMI), project risk management is one of 10 knowledge areas in project management, along with project integration management, project scope management, project schedule management, project cost management, project quality management, project resource management, project communication management, project procurement management, and project stakeholder management.

The risk management process is defined as follows [3].

Risk Management Planning. Define how to implement a risk management plan for a project.

Risk Identification. Determine which tasks will impact the project and document their characteristics.

Qualitative Analysis of Risks. Prioritize risks for subsequent analysis and action based on an assessment of their probability of occurrence and impact.

Quantitative Analysis of Risks. Numerically analyze the impact of the identified risks on the overall project goals.

Risk Response Planning. Develop options and measures to enhance opportunities and reduce threats to project goals. The risk response strategies include risk avoidance, risk transfer, risk mitigation, and risk acceptance.

Implementing Risk Response Plans and Monitoring Risks. Implement the risk response plan, track the identified risks, and monitor the remaining risks.

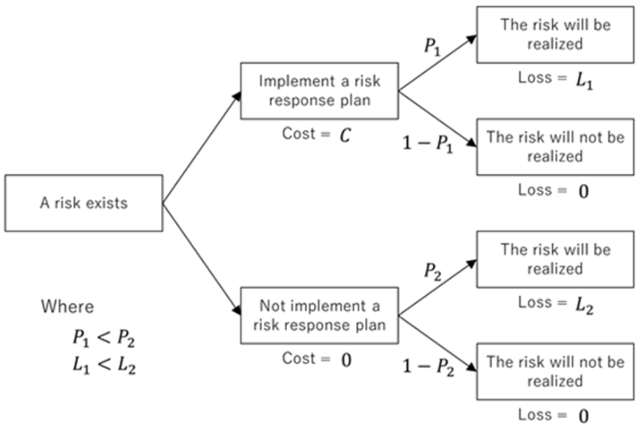

Fig. 1. A simplified decision-making model for project risk management.

Project risk management is a part of project management, but it is very different from other types of management activities in that it deals with probabilistic events or conditions. Project risks always exist, but they are not always recognized from the beginning, and even if they are recognized, they are not always realized, and they do not always remain the same throughout the project. If the risk is deterministic and it is known from the beginning that it will occur, then dealing with the risk should be incorporated into the project as a task in advance, and risk management should not be necessary in the first place. Risk management is necessary because the occurrence of risk is probabilistic and involves uncertainty. The fact that a risk is a probabilistic event or condition means that trade-offs must be made in deciding how to respond to the risk, where a trade-off is a situation or relationship in which pursuing one means sacrificing the other. The decision-making model in its simplest form in project risk management can be represented by a decision tree as shown in Fig. 1 [4]. If a risk response plan is implemented, the expected gain will be $E_1 = -C - P_1 L_1$, where the cost of the risk response plan is C and the probability and the loss that the risk will be realized is P_1 and L_1, respectively. If a risk response plan is not implemented, the expected gain will be $E_2 = -P_2 L_2$, where the probability and the loss that the risk will be realized is P_2 and L_2, respectively. If E_1 is higher than E_2, that is, $P_2 L_2 - P_1 L_1 > C$, the risk response plan should be reasonably implemented. However, it is not always possible to obtain the correct decision-making model in real projects. Part of the difficulty of project risk management can be explained by the difficulty of quantitative assessment of a risk.

Even if a correct decision-making model of risk is obtained, humans do not always make rational decisions and act accordingly. Bannerman points out that there is a large gap between the theory and practice of risk management, and cites the influence of human factors as one of the reasons for this gap [5]. Kutsch and Hall investigated deliberate indifference in risk management and categorized it into the following four types [6].

Untopicality. Intentionally ignoring risks that do not match past experience or preferences based on the intuition of managers and others.

Undecidability. Not being recognized as a risk by other stakeholders because the evidence for the risk is weak.

Taboo. Socially enforced indifference, such as not being allowed to know or touch something.

Suspension of Belief. Daring not to expend resources on a risk until it becomes apparent, hoping that the risk will resolve itself naturally.

Human decision making under uncertainty can be explained by prospect theory [7]. Prospect theory assumes general heuristics that humans use when making judgments about probability and frequency, and explains that human responses to probability are not linear, i.e., biases arise from rational judgments. Typical heuristics include reference dependence and loss aversion. Reference dependence is the property that people make decisions based on the potential gain or losses relative to their specific situation (the reference point) rather than in absolute terms. Loss aversion is the property that a loss is valued more strongly than a gain of the same amount. In other words, if there is a loss and a gain of the same amount, the dissatisfaction caused by the loss is perceived to be greater than the satisfaction caused by the gain of the same amount, which differs from expected utility theory

Fig. 2 shows the graph of the value function in prospect theory, where the origin of the graph is the reference point and the value (= satisfaction) is determined by the relative gain or loss from the reference point. According to loss aversion, even if the gain and loss from the reference point are the same, the negative value becomes greater than the positive value, indicating asymmetry. Applying this to the decision-making model shown in Figure 1, the baseline plan that does not include risk responses becomes the reference point. The left-hand side of $P_2 L_2 - P_1 L_1 > C$, which is the condition under which the risk response plan should be implemented, can be regarded as gain and the right-hand side as loss. Assuming that the loss is valued $k\ (> 1)$ times greater than the gain of the same amount, the condition becomes $P_2 L_2 - P_1 L_1 > k\, C$, which is biased toward avoiding the implementation of risk response plan, i.e., keeping the original plan as unchanged as possible. The intentional indifference to risk [6] may be strongly influenced by this cognitive bias.

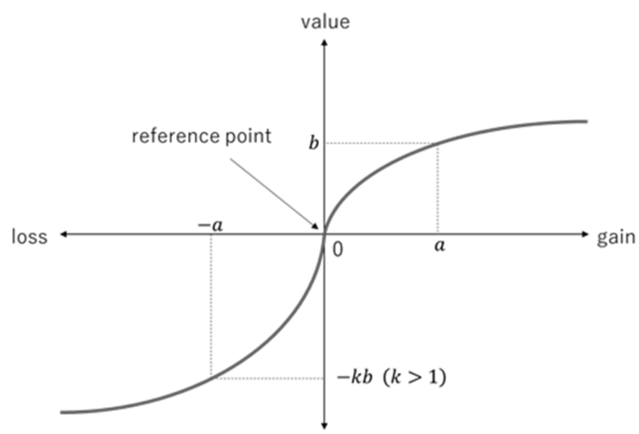

Fig. 2. The value function in prospect theory.

Prospect theory suggests that the heuristics that are the source of cognitive biases are assumed to be intuitive by System I in dual process theory [7]. Dual process refers to the two information processing systems that humans possess. The one is called System I, which is characterized by being intuitive, associative, fast, automatic, emotional, parallel processing, and effortless. The other is called System II, which is characterized as analytical, controlled, serial, rule-dominated, and labor-intensive. System I and System II have a complementary relationship, with System I quickly finding answers to problems, while System II monitors the quick decisions made by System I and approves them or makes corrections and changes as needed. In order to deal with the difficulty of project risk management caused by cognitive biases, it is important to know how to effectively coordinate System I and System II. Croskerry argues that in order to make optimal decisions, it is important to have the right balance of the two systems in a "mixed " is important for optimal decision making [8]. However, in reality, there are many situations in which System II cannot necessarily modify System I. Project risk management, which requires quick decision making with difficult trade-offs for various uncertain events, would be one of the most typical examples.

III. Machine-in-the-loop Process in Project Risk Management

The human decision-making process can be generalized as follows: (1) understand and realize the current situation based on input information from the environment, (2) formulate hypotheses and plan possible actions, (3) evaluate multiple options, (4) make a decision and take action, (5) update knowledge (if necessary) by checking the results of feedback from the environment. System I, which governs human daily thinking, tries to simplify decision-making by providing various shortcuts while following the above steps. For example, setting simplistic behavioral rules tied to specific situations, subjective evaluation of options based on personal preferences, etc. While this can greatly streamline the decision-making process, it can also be a breeding ground for biased thinking. In particular, the temptation to take shortcuts with System I is even stronger in project risk management, where decisions must be made within a limited amount of time, cost and resources. Therefore, it is necessary to deliberately activate an objective and logical System II and have a mechanism to monitor System I on a regular basis. System I and System II, so to speak, play a role like the accelerator and brake of a car. With these two systems in place, we can drive safely on a risky road.

Artificial Intelligence (AI) technology can be used to reduce the impact of cognitive biases in human decision making. Humans and AI models have a complementary relationship. On the one hand, humans can use the prediction and inference results of AI to raise awareness and eliminate human preconceptions and biases, leading to rational decision making. On the other hand, AI models can supplement missing background knowledge and other data with human insights and feedback (inconsistencies with reality, differences with experience and senses, etc.) to improve their ability to respond to new changes and unexpected events. The complementarity between the human and AI models promotes effective coordination between System I and System II. The human decision-making process is activated by System I by default, but the intervention of AI evokes System II and promotes rational decision-making in which humans and AI cooperate. As a result, the influence of cognitive bias in human decision making can be expected to be reduced. Furthermore, through the feedback of action results to the AI model and the teaching of domain knowledge, efficient action rules based on human experience and knowledge will gradually be accumulated in the AI model. Bengio states that "AI is currently only able to play the role of System I, and will need to acquire System II capabilities in the future. [9]" However, this is only in the "machine-dominated world", and in the "human-dominated world", machines (AI) can play the role of evoking System II in humans without necessarily acquiring System II capabilities.

We propose a machine-in-the-loop process in which the human decision-making process and the prediction and inference by AI models mutually complement each other. Figure 3 shows an image of the machine-in-the-loop process. The left side of Fig. 3 represents the human decision-making process, and the right side of Fig. 3 shows the cycle of prediction, estimation, and model updating of the AI model. The solid arrows represent the control flow, and the dotted arrows represent the data flow. The dotted arrows from the AI model to the human indicate (A) support for awareness through the presentation of new information and (B) the provision of decision-making materials based on prediction and estimation results. In these linkages, the interpretability of the AI model becomes very important because it is necessary for humans to interpret the prediction/estimation results and reflect them in decision making. The dotted arrow from the human to the AI model indicate model updates by feeding back the results of decision-makings and the gap between the prediction/estimation results and the reality, which contributes to continuous performance improvement of the AI model.

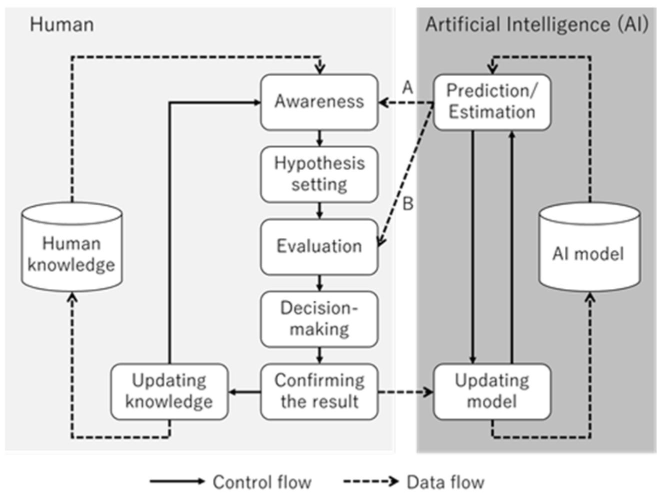

Fig. 3. An image of the machine-in-the-loop process.

In the application to project risk management, human interpretation of the prediction/estimation results of AI models will enable recognition of the influence of unexpected factors on project failures and unknown similarities between different projects. It will help to make decisions about how to respond to ongoing projects based on failure probabilities. Human experience and knowledge can be also reflected in the AI model to continuously improve the accuracy of the AI model and the efficiency of model building work.

IV. Case Study

In this section, we describe a case study of machine-in-the-loop risk management. It should be noted that the case is a modified version of an actual case for the purpose of explanation. The target organization has an organizational database that accumulates process data, such as project characteristics, development scale, progress in each process of development, cost, and quality. After clarifying the definitions of project success/failure, a prediction model of project failure was constructed. We used Naive Bayes classifier [10] [11] for the prediction model. Naive Bayes is a relatively simple machine learning model based on Bayes' theorem, which simplifies probability calculation by placing conditional independence assumptions between variables, and realizes an algorithm for sequential updating of posterior probabilities by adding new information. This mechanism of probability updating is called Bayesian updating, and is considered to be relatively close to the human decision-making process in risk management [12]. Naive Bayes is also known to be computationally efficient and robust to noise. Missing values in the data can also be handled without contradiction by assuming that no new information was added in the Bayesian update.

In the following, we will show how the human (Project Manager: PM) and the machine (AI) interact with each other in this case study. Fig. 4 shows the output of the machine (AI) on the importance of the explanatory variables, which indicates the magnitude of impact on project failure. The explanatory variables on the X-axis are arranged in the order of the process. According to the graph, we can observe that the later the process, the more the impact on project failure appears in the explanatory variables, which is consistent with the human (PM) sense. On the other hand, the impact of some explanatory variables includes those that are not consistent with the intuition of the human (PM). Additional research will divide the response into two cases: one where the human (PM) bias is corrected, and the other where new knowledge is put into the machine (AI) to build a new model.

Fig. 5 shows the output of the machine (AI) on the risk transition of an ongoing project by process, where the red line is the project that the human (PM) is currently focusing on. According to the graph, we can observe that the project is relatively risky, as the failure probability gradually increases and eventually reaches a high probability, but the intuition of the human (PM) often differs from the observation. Further investigation including continuous monitoring of the project will modify both the human (PM) bias and the machine (AI) modeling.

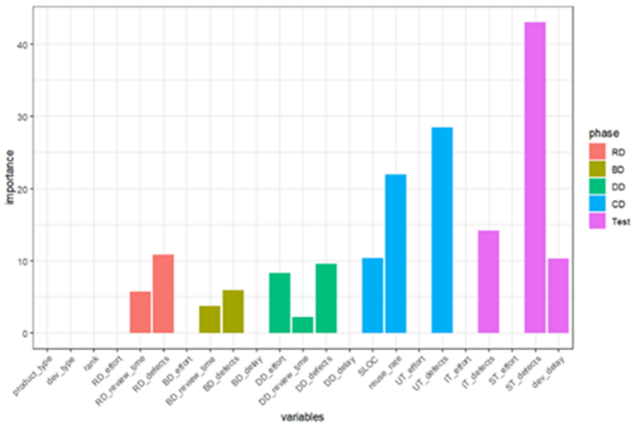

Fig. 4. The importance of the explanatory variables.

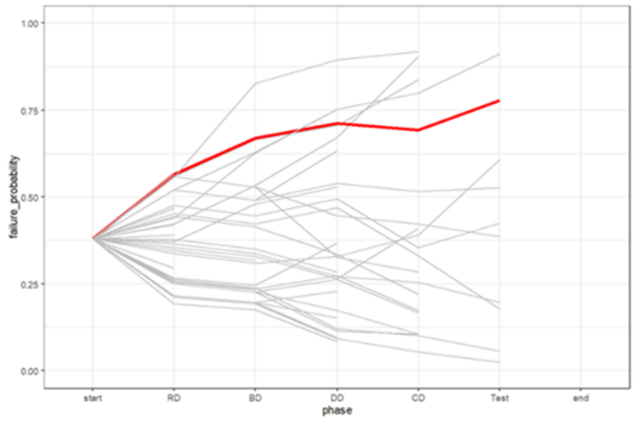

Fig. 5. The risk transition of an ongoing project by process.

Fig. 6 shows the output of the machine (AI) on the similarity between projects calculated using the intermediate data of the prediction model, and the cluster analysis performed on it. The vertical axis represents the distance between the projects; the closer the distance, the more similar the prediction results are. The projects in each cluster span different sectors, which may lead to new insights into the unexpected similarity of the projects. The red boxes show the projects that had similar predictions but different actual results (success or failure). By examining where the differences in actual results originated in these projects, the human (PM) may be able to understand the essential mechanisms of project failure and obtain hints for new data collection for the machine (AI).

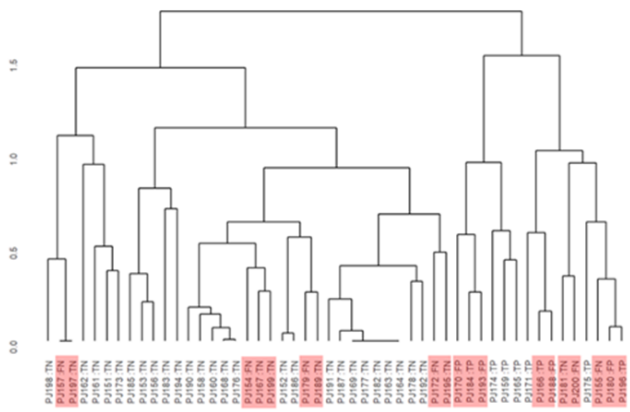

Fig. 6. The similarity between projects calculated using the intermediate data of the prediction model.

We conducted an interview survey with two practitioners in the applied department of machine-in-the-loop risk management. In the interviews, the following questions were asked about the effects of the machine-in-the-loop process.

Q1. Has uncertainty and ambiguity been reduced?

Q2. Have preconceptions and biases been corrected?

Q3. Did it promote consensus building among stakeholders?

All interviews were recorded and coded by topic. As a result, 24 codes were obtained. We categorized all the obtained codes into positive and negative opinions, as shown in Fig. 7. In Q1 and Q2, about 70% of the opinions were positive, supporting that the machine-in-the-loop process is effective in reducing uncertainty and cognitive bias. On the other hand, in Q3, there are more negative opinions than positive ones, which may indicate that consensus building among stakeholders is related not only to cognitive bias but also to transaction costs [13], and that it takes a long time to see the effects because it requires continuous improvements of organizational culture.

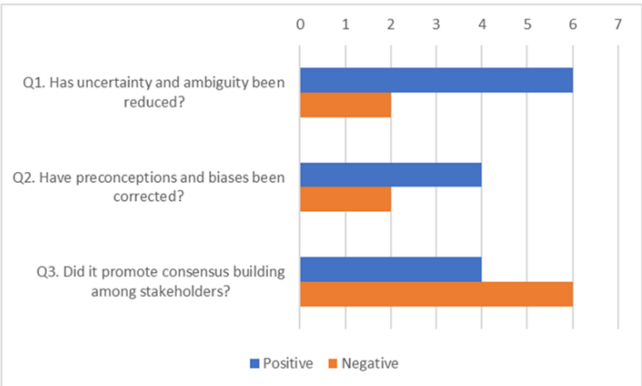

Fig. 7. The results of categorizing all the obtained codes into positive and negative opinions.

V. Related Work

Several related studies have been conducted on the application of machine learning to risk management. Takagi et al. conducted a questionnaire survey of 32 project managers to characterize risky software development projects [14]. Lee et al. present a Bayesian network risk management framework for large-scale engineering projects [15]. Two types of Bayesian network models were developed based on the results of questionnaire surveys, and sensitivity analysis and IF-THEN analysis of major risks were conducted. However, these studies did not explicitly discuss the complementarity between humans and AI.

Interpretability of AI or machine learning has been studied extensively in recent years in connection with research such as Explainable AI [16] [17]. Lipton considers interpretability not as one simple concept but as a combination of several different concepts, and broadly categorizes it into transparency about the internal workings of a predictive model and a post-hoc interpretability about external outputs [18]. Many tools that implement post-hoc interpretability have recently been introduced. Local Interpretable Model-agnostic Explanations (LIME) provides model-agnostic explanations by locally approximating the model around a given prediction [19]. As an extension of LIME, the SHapley Additive exPlanations (SHAP) method was proposed based upon the Shapley value concept from game theory [20].

Decision support systems (DSS) have been studied for a long time [21] [22], but they differ slightly from the purpose of this research in two respects: decision making is only a part of the problem-solving process, and DSS are mainly intended to provide unidirectional support from machine to humans. As for the bidirectional cooperation between humans and AI, human-in-the-loop machine learning has been proposed for applications in the medical field [23]. In medical applications, it is important for humans to be actively involved in the model building process to prevent errors, because there are special situations in which data sets contain events that occur only rarely and decisions are made based on the high level of human expertise. Human-in-the-loop is a concept that focuses mainly on the machine learning side, while we are more interested in the human side. Accordingly, the concept of Human AI teaming has been introduced in recent years. Bansal et al. state that for human-AI collaboration to achieve better performance, merely high accuracy is not enough; human mental models of AI capabilities are important [24]. Lai and Tan conducted experiments to see how human-AI coordination affects performance at several levels, from decisions made by humans alone to full automation by AI, and found that AI's presentation of predicted outcomes with explanations remarkably improves the performance [25].

VI. Conclusion

The difficulty of project risk management is thought to be largely influenced by cognitive biases related to heuristics when humans make decisions about probability and frequency. In order to reduce the influence of cognitive bias in human decision making, we proposed a machine-in-the-loop process in which the human decision-making process and the prediction and inference by the AI model cooperate in a complementary manner.

We conducted an interview survey in a case study where the machine-in-the-loop process was applied to project risk management, and found that the majority of comments were positive about the reduction of uncertainty and cognitive bias, confirming the effectiveness of the machine-in-the-loop process. On the other hand, the ratio of negative comments on consensus building among stakeholders was relatively high. This may be because consensus-building among stakeholders requires a long time to be effective, since it is related not only to cognitive bias but also to transaction costs, and requires continuous improvements of organizational culture.

We believe that the concept of machine-in-the-loop process, which will correct cognitive bias and extrapolation problem by the complementary use of humans and AI models, can be applied not only to project risk management but also to various other fields.

References

[1] Ibbs, C. W., Kwak, Y. H.: Assessing project management maturity. *Project Management Journal* 31(1), 31-43 (2000).

[2] Grant, K. P., Pennypacker, J. S.: Project management maturity: An assessment of project management capabilities among and between selected industries. *IEEE Transactions on engineering management* 53(1), 59-68 (2006).

[3] Project Management Institute: *A guide to the project management body of knowledge (PMBOK guide) 6th edn*. Project Management Institute, Inc., (2017).

[4] Mori, T., Uchihira, N.: An Integrated Approach of Machine Learning and Knowledge Creation in Project & Program Risk Management. *Journal of International Association of P2M* 14(1), 415-435 (2019).

[5] Bannerman, P. L.: Risk and risk management in software projects: A reassessment. *Journal of Systems and Software* 81(12), 2008, 2118-2133 (2008).

[6] Kutsch, E., Hall, M.: Deliberate ignorance in project risk management. *International journal of project management* 28(3), 2010, 245-255 (2010).

[7] Kahneman, D: *Thinking, fast and slow.* Straus and Giroux, New York (2011).

[8] Croskerry, P.: A universal model of diagnostic reasoning. *Academic medicine* 84(8), 1022-1028 (2009).

[9] Bengio, Y.: From System 1 Deep Learning to System 2 Deep Learning. In: *Neural Information Processing Systems*, (2019).

[10] Lewis, D. D.: Naive (Bayes) at forty: The independence assumption in information retrieval. *European conference on machine learning*, Springer, Berlin, Heidelberg (1998).

[11] Mori, T., Uchihira, N.: Balancing the trade-off between accuracy and interpretability in software defect prediction. *Empirical Software Engineering* 24(2), 2019, 779-825 (2019).

[12] Mori, T., Tamura, S., Kakui, S.: Incremental estimation of project failure risk with Naive Bayes classifier. In: *7th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, pp. 283-286. IEEE, Baltimore (2013).

[13] Williamson, O. E.: The economics of organization: The transaction cost approach. *American journal of sociology* 87(3), 548-577 (1981).

[14] Takagi, Y., Mizuno, O., Kikuno, T.: An empirical approach to characterizing risky software projects based on logistic regression analysis. *Empirical Software Engineering* 10(4), 495-515 (2005).

[15] Lee, E., Park, Y., Shin, J. G.: Large engineering project risk management using a Bayesian belief network. *Expert Systems with Applications* 36(3), 2009, 5880-5887 (2009).

[16] Gunning, D.: Explainable artificial intelligence (XAI). Defense Advanced Research Projects Agency (DARPA), (2017).

[17] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. *ACM computing surveys* 51(5), 1-42 (2018).

[18] Lipton, Z. C.: The Mythos of Model Interpretability. In: *Proceedings of the ICML 2016 Workshop on Human Interpretability in Machine Learning,* (2016).

[19] Ribeiro, M. T., Singh, S., Guestrin, C.: Why should I trust you?: Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144. (2016)

[20] Lundberg, S. M., Lee, S.: A unified approach to interpreting model predictions. In: *Proceedings of the 31st international conference on neural information processing systems*, (2017).

[21] Arnott, D., Pervan, G.: A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87 (2005).

[22] Arnott, D., Pervan, G.: A critical analysis of decision support systems research revisited: the rise of design science. *Journal of Information Technology*, 29, 269–293 (2014).

[23] Holzinger, A.: Interactive machine learning for health informatics: when do we need the human-in-the-loop?. *Brain Informatics* 3(2), 2016, 119-131 (2016).

[24] Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., Horvitz, E.: Beyond accuracy: The role of mental models in human-ai team performance. In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 7(1), (2019).

[25] Lai V., Tan, C.: On human predictions with explanations and predictions of machine learning models: A case study on deception detection. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*, (2019).

Toshiki Mori is Expert of Corporate Software Engineering & Technology Center at Toshiba Corporation. Before he joined Toshiba in 1992, he received B.E. degree in precision mechanical engineering in 1990 and M.S. degree in information engineering in 1992, both from the University of Tokyo. He was a Visiting Industrial Associate from 1996 to 1999 at the Center for Design Research and the Manufacturing Modeling Laboratory of Stanford University. He received his Ph.D. from School of Knowledge Science, Japan Advanced Institute of Science and Technology, in 2020. His current research interests include the application of data mining and knowledge science to software engineering and project management.

Naoshi Uchihira is Professor of Japan Advanced Institute of Science and Technology (JAIST). He received Dr. Eng. degrees from Tokyo Institute of Technology, and Ph.D. (Knowledge Science) from JAIST in 1997 and 2010, respectively. He worked at the corporate R&D center of Toshiba Corporation from 1982 to 2013. His current research interests include project management and digital innovation design.