September 27, 2022

TO:
National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
Submitted electronically to AIframework@nist.gov

From:
Community of Deep Learning Evangelists

To Whom It May Concern:

Response to NIST Request for Public Comment regarding NIST AI Risk Management Framework Second Draft

Thank you for this opportunity to comment on the NIST AI Risk Management Framework Second Draft. We offer the following submission for your consideration to assist moving forward. We welcome future opportunities to continue supporting NIST's work and provide further input and assistance.

Public Comment #1:
Overall
We believe NIST is generally on the right track as it develops the AI RMF. This second draft of the AI RMF is highly complete. However, there are still many operational hurdles to the AI RMF, such as the need for knowledge background for operations, mixed understanding and interpretation, and optimistic and less optimistic views from different industries.

As the Community of Deep Learning Evangelists, the largest AI community in Japan with over 50,000 members working in collaboration with Japan Deep Learning Association, we would like to propose the following points that would make the AI RMF easier to use in the field, as a discussion from the viewpoint of for those who utilize it.

Public Comment #2:
2. Audience (p.6, L18-19)
Comment:
"Their actions can designate boundaries for operation (technical, societal, legal, and ethical)." stated. However, it is difficult to discern the divergence of complex and sophisticated responsibilities for operation (technical, social, legal, and ethical) which may involve legal responsibilities and, in some cases, a mixture of power relationships in discussions among stakeholders.

In order to make it work realistically, it is necessary to consider the establishment of an independent TIC center-like function to support the delineation of boundaries, and to develop a mechanism to improve corporate incentives, which may be part of the Playbook that will be reviewed and integrated semiannually in the future. We believe that further breakthroughs are needed.

With this in mind, first we detail specific proposals to further enhance the formal or semi-formal norms and guidance that support the operation of the AI RMF in Public Comment #3.

Public Comment #3:
3.2.1. Risk Measurement (p.8, L16-17)
Comment:
“Organizations will want to identify and track emergent risks and consider techniques for measuring them.” stated, Stakeholders on the AI ECO system should be appropriately aware of the risk components of AI services. They should also pay attention to social cases and incidents surrounding AI and recognize key risk scenarios. To this end, the risk chain on the AI ECO system should be visualized, risk control should be examined, and as an implementation mechanism for social systems and organizational management systems, it should be broken down as manageable risks, and the roles and responsibilities among stakeholders should be defined and consensus should be formed.

We do not necessarily believe that all risks need to be mitigated, but that appropriate controls based on some principle of proportionality should be established by stakeholders in the AI ECO system according to their responsibilities and roles, taking into account the magnitude of risk, technical difficulty, cost-effectiveness, continuity, and other factors.

Also, in the maintenance and operation of AI systems, there are likely to be many things that management must decide on even if they are non-engineers, such as the response to incidents (incident level determination, impact range prediction, initial response, response, reporting, and recurrence prevention). In this sense, it is very important to establish guidelines so that everyone's perspective is aligned from the initial stages of the project.

We believe this is the foremost lens through which the NIST Artificial Intelligence Risk Management Framework can provide more practical, actionable guidance and can be implemented across organizations through the life cycle.

To that end, we propose to include the following three practical resources available on the Internets as complementary resources to assist users of the AI RMF. One method for identifying and tracking emerging risks on the AI ECO system, measuring them, and facilitating common understanding, dialogue, updates and consensus with stakeholders on the AI ECO system; to help build, verify and explain the quality of AI systems from a technical perspective.
We believe that their inclusion will greatly enhance the value of the AI RMF for organizations and other users.

< Proposal #1 of Complementary Resource to the AI RMF：Japan Deep Learning Association Study Group Outputs （AI Governance and its Evaluation）: Chair: Arisa Ema, Associate Professor, Institute for Future Initiatives, University of Tokyo>

1）AI Governance Ecosystem : Trusted AI with Industrial Structure (July, 2021)

Theme/Issue: Various actors may govern the controlling and evaluation structures of AI systems. Explores possible governance frameworks to help build trustworthy AI systems.

Phase 1 report: AI Governance Ecosystem: Trusted AI with Industrial Structure (July 2021)
Recommended to establish “AI Governance Ecosystem,” where various external actors interact.

The report proposes addressing the following three perspectives, in the process of placing AI related principles such as safety, fairness, privacy, and transparency into practice. It emphasizes the importance for AI governance discussions and considers how AI services should be provided not only by single companies or organizations, but also with various other entities and actors such as external contexts and evaluation organizations.
1. AI governance ecosystem should be established
2. Reliability of AI should be ensured by taking industrial structure into account
3. Practical examples of Japan’s unique challenges and issues should be disseminated

Phase 2 report: AI Governance Ecosystem : Who Manages and Evaluates AI?（July 2022)
For the three recommendations specified in the phase 1 report to be put into practice, the AI governance ecosystem must have its functions enhanced and diversified by considering implemented cases. The following two areas were examined during the second phase period.

1. Updating the AI governance ecosystem: The AI governance ecosystem diagram has been updated to illustrate the phase 2 topics and discussions.
2. AI governance ecosystem case studies: The study group has examined how the AI governance ecosystem functions regarding specific service areas such as human resources technology.

Reference
See: https://www.jdla.org/en/en-document/

2）AI Governance Ecosystem Database
This database provides a collection of information sources on related initiatives/practices by players composing the AI Governance Ecosystem. The list will be further developed over the period of study. (For more about the AI Governance Ecosystem, please see the report of the JDLA study group “AI governance and its evaluation”)

Reference
See: https://www.jdla.org/en/en-document/en-ai-governance-eco-system/

<Proposal #2 of Complementary Resource to the AI RMF：Policy Recommendation “RCModel, a Risk Chain Model for Risk Reduction in AI Services”: Takashi Matsumoto (Deloitte Tohmatsu Group/The University of Tokyo) ,Arisa

Ema (The University of Tokyo/RIKEN AIP Center) >

With the increasing use of artificial intelligence (AI) services and products in recent years, issues related to their trustworthiness have emerged and AI service providers need to be prepared for various risks. The AI Governance Project of the Technology Governance Policy Research Unit, Institute for Future Initiatives, the University of Tokyo has been studying a framework for risk assessment and control of AI services. This recommendation is one of the outputs of the research.

In this policy recommendation, we propose a risk chain model (RCModel) that supports AI service providers in proper risk assessment and control. We hope that RCModel will contribute to the realization of trustworthy AI services. In the future, we plan to systematize the framework through joint research with various stakeholders.

<Overview of the Risk Chain Model>
1) Organizing and structuring risk components
There are many factors that may cause risk in the provision of AI services (hereinafter referred to as "components"). The RCModel identifies them as follows

(1) Technical components of AI systems, and

(2) Components related to the service provider's code of conduct (including communication with users), and

(3) components pertaining to the user's understanding, behavior, and usage environment.

2) Identification of risk scenarios and components that are risk factors
Risk scenarios pertaining to AI services, such as unfair judgment, accidents due to loss of control, etc., were identified. Then, for the risk scenarios that should be prioritized for consideration, we identified the components that are risk factors.

3) Visualization of risk chain and consideration of risk control
Since it is difficult to sufficiently reduce risks by individual components, AI service providers can consider step-by-step risk reduction by visualizing the relationship among components related to risk scenarios (risk chain). This allows the location of risk factors and effective and efficient controls to be examined.

Reference
See: https://ifi.u-tokyo.ac.jp/en/project-news/4828/

< Proposal #3 of Complementary Resource to the AI RMF: Machine Learning Quality Management Guideline : National Institute of Advanced Industrial Science and Technology (AIST)>

The "Machine Learning Quality Management Guideline" deals with quality management throughout the lifecycle of AI systems using machine learning, and systematically outlines the necessary efforts and inspection items to fulfill the quality requirements for the provision of AI system services. In particular, by establishing standards and achievement targets for the quality of software components (machine learning elements) implemented with machine learning that are included in the system, the Guidelines aim to help companies measure and improve the quality of the AI systems they have built, and to reduce accidents and economic losses caused by misjudgment of AI.

The quality management required by the guidelines covers not only the development of machine learning elements, but also the entire system lifecycle from the definition of quality requirements for AI systems incorporating machine learning elements, demonstration tests, system development, and maintenance and operation. Depending on the development circumstances of each AI system and other factors, it is expected that the division of roles among service providers, system developers (SI vendors, engineers), and other stakeholders will be defined to meet these quality management requirements. It is also envisioned to be used for consensus building in the ordering and outsourcing of development work and for setting acceptance inspection conditions.

Reference
See: https://www.digiarc.aist.go.jp/en/publication/aiqm/

Respectfully submitted,

(General Lead of the AI legal group of the community)
Hiromu (Kit) Kitamura, Evangelist for Artificial Intelligence, QMS and Legal
AI Legal Group,
Community of Deep Learning Evangelists

(Lead for Proposal)
Hiromu (Kit) Kitamura, Evangelist for Artificial Intelligence, QMS and Legal
AI Legal Group,
Community of Deep Learning Evangelists

(Experts contributed)
Atsuko Miura (Community of Deep Learning Evangelists)
Saiko Kumao (Community of Deep Learning Evangelists)
Yuka Koseki (Community of Deep Learning Evangelists)
Takashi Matsumoto (Deloitte Tohmatsu Group/The University of Tokyo/ Japan Deep Learning Association)
Arisa Ema (The University of Tokyo/ Japan Deep Learning Association)

CC) Community of Deep Learning Evangelists/ Japan Deep Learning Association