



Northrop Grumman
2980 Fairview Park Drive
Falls Church, Virginia 22042-4511

northropgrumman.com

April 28, 2022

Ms. Elham Tabassi
National Institute of Standards and Technology
100 Bureau Drive, Stop 200
Gaithersburg, MD 20899
Alframework@nist.gov

Dear Ms. Tabassi,

Thank you for the opportunity to provide feedback on the draft Artificial Intelligence Risk Management Framework (AI RMF). Our Northrop Grumman Responsible AI Working Group reviewed the document and were pleased with the progress made since the concept paper. We humbly offer the following feedback in response to some of your key questions.

- **Whether the AI RMF appropriately covers and addresses AI risks, including with the right level of specificity for various use cases.**
 - Page 7 Lines 5-7: The statement “Ideally, members of all stakeholder groups would be involved or represented in the risk management process, including those individuals and community representatives that may be affected by the use of AI technologies” does not align with the need to be adaptable to different use cases. As a defense company, most of our AI systems will never interact with the general public, so they would not be stakeholders for our use cases. Recommend rewording this sentence to “members of all *affected* stakeholder groups.”
- **Whether the functions, categories, and subcategories are complete, appropriate, and clearly stated.**
 - Page 9 Lines 6-9: The statement “Accuracy indicates the degree to which the ML model is correctly capturing a relationship that exists within training data” assumes that the training data itself is an accurate representation of ground truth. Recommend addressing the potential for data bias in this section, referring to the more complete discussion of bias in later sections.
 - Page 9 Lines 18-19: The statement “Reliability indicates whether a model consistently generates the same results, within the bounds of acceptable statistical error” is not entirely accurate within the context of AI. Given new data, we would not expect a model to produce exactly the same result. Recommend rewording to account for the non-deterministic nature of AI systems.

- Page 13, Lines 12-13: Recommend changing the sentence “Absence of harmful bias is a necessary condition for fairness” to “Mitigation of harmful bias is a necessary condition for fairness.” In some cases, harmful bias may exist in historical data sets, but removing those data sets completely may be counterproductive. Rather, it is important to understand the sources of bias and actively work to mitigate the harmful effects to ensure fairness in the end product.
- **What might be missing from the AI RMF.**
 - Page 5 Lines 7-12: The definition listed for AI seems more like a definition for Machine Learning. Recommend editing the definition to something broader. The definition provided by Darrell West and John Allen of the Brookings Institution could be used: “Artificial Intelligence refers to algorithms or systems that operate in an intentional, intelligent, and adaptive manner.” (Source: [How artificial intelligence is transforming the world \(brookings.edu\)](https://www.brookings.edu/blog/algorithmic-accountability/2019/04/11/how-artificial-intelligence-is-transforming-the-world/)).

Northrop Grumman continues to be committed to the development of Responsible AI, and we look forward to further engaging with the NIST team.

Sincerely,

Dr. Amanda C. Muller
Northrop Grumman Responsible AI Lead