Holistic AI comments on ‘NIST AI Risk Management Framework: Second Draft’

29th September 2022

About Holistic AI

Holistic AI is an AI Risk Management company, with a mission is to empower enterprises to adopt and scale AI with confidence. Our automated AI Risk Management platform audits, assures, and monitors AI systems' code, data, policies, and processes. This enables enterprises to maximise the value of their AI, minimise reputational, legal, and commercial risks, and accelerate innovation.

We have deep practical experience auditing AI systems, having assured over 100+ enterprise AI projects covering 20k+ different algorithms. Our clients and partners include Fortune 500 corporations, SMEs, governments, and regulators. We have a multidisciplinary team of AI/ML engineers, data scientists, ethicists, and law and policy experts.

Introduction

Holistic AI welcomes the opportunity to comment on the second draft of the National Institute of Standards and Technology’s (NIST) AI Risk Management Framework. We appreciate NIST’s open, transparent, and collaborative approach.

To remain competitive and agile in the modern economy, organisations have little choice but to adopt AI at scale. However, in doing so, they are exposed to serious reputational, legal, and commercial risks, which can have harmful knock-on effects for individuals and wider society. The only way to mitigate these risks is through robust AI risk management.

Given the rapid advancements in AI innovation, and the scale and speed at which AI risks can materialise, organisations should not rely solely on traditional risk management approaches when seeking to mitigate the novel risks posed by AI.

Holistic AI is therefore supportive of NIST’s leadership and work in this area. This is a crucial initiative for increasing awareness and knowledge of the practical and operational ways in which AI risks can be assessed, verified, and mitigated. Furthermore, this work meets the growing need for shared standards and industry benchmarks in this important area.

Comments on AI Risk Management Framework

AI systems require ongoing and dynamic monitoring

- We support NIST’s underpinning premise that AI poses novel risks, necessitating a distinct approach to AI Risk Management.
- The need for live, dynamic, and ongoing monitoring of AI systems is a key factor which differentiates AI Risk Management from more traditional forms of risk management.
- Given that AI systems continuously learn, adapt and evolve, they need to be carefully monitored on an ongoing basis.
- The higher risk a system, the more frequently it should be assessed and reviewed.
- The fact that AI performance tends to decay over time reinforces the importance of ongoing and dynamic monitoring.
- For organisations deploying many AI systems, this monitoring is virtually impossible without some form of scalable and automated solution.

- AI risks can emerge at or stem from any stage of the AI lifecycle, meaning every stage is relevant. We agree with NIST that the earlier in the lifecycle risks are assessed and surfaced, the better.

AI risk management requires technical assessment, testing and mitigation

- Meaningful AI Risk Management is not possible without performing technical, quantitative assessments and testing.
- There are a range of state-of-the-art tests which can be performed on the datasets and algorithmic models which AI systems are comprised of. These tests can identify issues relating to robustness, efficacy, bias, transparency, privacy, and other trustworthy and responsible AI verticals.
- Furthermore, there are technical mitigations which can be employed to address these issues.
- This is why we support the subcategory items listed under '*MEASURE 2: Systems are evaluated for trustworthy characteristics*'.
- Although the focus on governance related issues -- like accountability, roles and responsibilities, training, and culture is important and welcome -- without identifying and addressing technical issues in the code or data, AI risks cannot be adequately managed in the broadest and most robust sense.
- NIST should support organisations in understanding the technical state-of-the-art in this area. The technical resources we have provided below can contribute to this.

AI Risk Management accelerates innovation

- We strongly support NIST's pro-innovation stance and think that the AI RMF delivers on this objective.
- In our experience working with enterprises, we have found that robust AI Risk Management accelerates innovation and enhances business performance.
- This is because assessing, improving, and monitoring the performance of an organisation's AI systems is a core part of AI Risk Management. AI system performance can be measured in terms of both robustness and efficacy.
- Improving AI system performance is the only way to ensure that organisations are maximising ROI on the vast amounts spent on AI technologies and are harnessing AI's transformative potential.
- AI Risk Management processes also provide the guardrails to innovation. By boosting trust and increasing understanding of new technologies, different parts of the organisation are encouraged to adopt and scale AI use.
- There is ample evidence that many organisations are struggling to progress beyond pilots and scale AI use. AI Risk Management processes which support go/no-go decisions tackles this issue.
- NIST should advocate for the innovation-enabling role of AI Risk Management by supporting future work on this theme.

Empirical evidence and industry case studies are needed

- Although the AI RMF is use-case and industry agnostic, more case studies are needed, to highlight how organisations are applying the AI RMF (and other approaches) in their respective contexts.
- It would be helpful for NIST to encourage and support future work which highlights best practice and innovative approaches in industries and verticals which are particularly exposed to AI risks (e.g., HR, healthcare, insurance, banking etc.)
- We support NIST's outcomes-focused approach. The same AI technology can be low risk in one context but high risk in another, depending on what it is being used for. Also, some responsible AI verticals, like transparency and explainability, will be more important in some contexts than others.
- Therefore, it is important to understand the idiosyncratic practical and operational factors which organisations in specific industries or contexts must consider when applying the AI RMF.

- This can be elucidated through empirical evidence and industry-focused case studies.

Holistic AI resources

We note the request for resources and guidance to be contributed for the 'NIST Trustworthy and Responsible AI Resource Center'.

Technical resources

We have several open-source resources which technical teams seeking to develop and deploy trustworthy AI can use:

- **Holistic AI Risk Mitigation Roadmaps**: https://holisticai.gitbook.io/roadmaps-for-risk-mitigation/
- **Holistic AI Risk Mitigation Roadmaps GitHub repository**: https://github.com/holistic-ai/mitigation-roadmaps
    - A set of guides (with tutorials and examples) to mitigate some of the most common AI risks. Includes risks relating to system robustness and efficacy, bias, explainability, and privacy.

- **Holistic AI Library**: https://holisticai.readthedocs.io/en/latest/
- **Holistic AI Library GitHub repository**: https://github.com/holistic-ai/holisticai
    - An open-source tool to assess and improve the trustworthiness of AI systems. The current version of the library offers a set of techniques to easily measure and mitigate Bias across a variety of tasks.
    - Our long-term goal is to offer a set of techniques to easily measure and mitigate AI risks across five areas: Bias, Efficacy, Robustness, Privacy and Explainability. This will allow a holistic assessment of AI systems.
- **Measuring bias tutorials**: https://holisticai.readthedocs.io/en/latest/measuring_bias.html
- **Mitigating bias tutorials**: https://holisticai.readthedocs.io/en/latest/mitigating_bias.html

Non-technical resources

We have also published several research papers in the field of AI Risk Management, which may be of use:

- **Towards Algorithm Auditing: A Survey on Managing Legal, Ethical and Technological Risks of AI, ML and Associated Algorithms** (2021): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3778998

- **Systematizing Audit in Algorithmic Recruitment** (2021): https://www.mdpi.com/2079-3200/9/3/46

- **Practical risk management in AI** (2022): https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/1-practical-risk-management-in-ai

***

Please contact for any further information or follow-up on this submission.