From: Onwubiko, Cyril
To: aiframework
Subject: AI Risk Management Framework // Comments
Date: Tuesday, August 24, 2021 1:36:05 PM
Attachments: Outlook-bey5xvwl.png
AI RMF RFI Response - Cyril Onwubiko.pdf

Dear AI RMF Team,

I understand that NIST intends to extend the deadline for submitting comments to the AI Risk Management Framework to September 15, 2021, and in this respect, please find attached my comments to the AI Risk Management Framework.

I would be available for clarification of any of my comments should you need to, and to collaborate on this AI RMF development moving forward.

==========

Responses to Specific Request for information (pages 11,12, 13 and 14 of the RFI)

Response: "NIST is among the institutions addressing these issues. NIST aims to cultivate the public's trust in the design, development, use, and evaluation of AI technologies and systems in ways that enhance economic security, and improve quality of life." We would request you to change "economic security" to "public confidence".
Rationale: AI risk should not be focused on "economic security" rather on ensuring "public trust and confidence" in the way AI applications are designed, deployed and used.

Suggested Change: NIST is among the institutions addressing these issues. NIST aims to cultivate the public's trust in the design, development, use, and evaluation of AI technologies and systems in ways that enhance public confidence, and improve quality of life.

==========

2. How organizations currently define and manage characteristics of AI trustworthiness and whether there are important characteristics which should be considered in the Framework besides: accuracy, explainability and interpretability, reliability, privacy, robustness, safety, security (resilience), and mitigation of harmful bias, or harmful outcomes from misuse of the AI;

Response: AI trustworthiness should include 'data representativeness'.
Rationale: If the dataset on which an AI model is trained is not 'representative' of the problem domain, then the model can still be explainable and interpretability, and still yet harmfully biased. It will be biased on the type of data it's trained upon. For example, if an AI model is trained against dataset of a certain demographic, then it's extremely likely to be biased. So, data representativeness should be a mandatory requirement for trustworthiness.

Suggested Change: AI trustworthiness characteristics
should include: accuracy, explainability and interpretability, representative (data), reliability, privacy, robustness, safety, security (resilience), and mitigation of harmful bias, or harmful outcomes from misuse of the AI.

==========

4. The extent to which AI risks are incorporated into different organizations' overarching enterprise risk management – including, but not limited to, the management of risks related to cybersecurity, privacy, and safety;

Response: It should include the management of risks related to social and racial.

Suggested Change: The management of risks related to cybersecurity, privacy, safety, social and racial.

==========

6. How current regulatory or regulatory reporting requirements (e.g., local, state, national, international) relate to the use of AI standards, frameworks, models, methodologies, tools, guidelines and best practices, and principles;

Response: New regulations are being enacted to address AI specific risks, such as responsible use, avoidance of harmful bias, social and racial injustice. For example, the European Union has proposed an AI Regulation, to address High-Risk AI applications. The regulation can be found here https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN

Suggested Change: AI Regulation for trustworthiness, responsible use, transparency, fairness, equity and accountability

==========

8. How organizations take into account benefits and issues related to inclusiveness in AI design, development, use and evaluation – and how AI design and development may be carried out in a way that reduces or manages the risk of potential negative impact on individuals, groups, and society.

Response: Organisations are demanding diversity and inclusiveness in AI design, development, use and evaluation to reduce risk of potential negative impact on individuals, groups and society as follows:
1) Diversity in AI Teams: AI Teams should include "citizen-stakeholders" (that is, non-AI personnel/experts), who are knowledgeable in other domains, e.g., Product Managers, Data Owners, Architects, Engineers, Linguists to the AI Team, so that it's a multidisciplinary team.
2) Develop AI Training Package/Program for staff who are interested so that more people are involved.
3) Ensure under-represented communities and people (e.g., Women, Blacks, Asian, Hispanics etc) are part of the AI Team.
4) Form an AI Ethics Board, as an accountability structure, to hold both the organisation and the AI Team accountable.

Suggested Change: AI Ethics Board

==========

11. How the Framework could be developed to advance the recruitment, hiring, development, and retention of a knowledgeable and skilled workforce necessary to perform AI-related functions within organizations.

Response:
1) Develop AI Trainings Package/Program for staff who are interested so that more people are involved

2) Diversity in AI Team - include 'citizen-stakeholders' (that is, non-AI personnel), but who are knowledgeable domain experts in their own areas, e.g., Product Managers, Data Owners, Architects, Engineers to the AI Team, so that they can train and learn on the job (up-skill and re-skilling)

3) Ensure under-represented communities and peoples (e.g., Women, Blacks, Asian, Hispanics etc) are part of the AI Team to encourage interest, and motivate uptake.

Suggested Change:

1) Develop AI Trainings Package/Program for staff who are interested so that more people are involved

2) Diverse Multidisciplinary Cross-Functional AI Team: The AI Team should include 'citizen-stakeholders' (that is, non-AI personnel), but who are knowledgeable domain experts in their own areas, e.g., Product Managers, Data Owners, Architects, Engineers to the AI Team, so that they can train and learn on the job (up-skill and re-skilling)

3) Ensure Under-Represented Communities and Peoples (e.g., Women, Blacks, Asian, Hispanics etc) are part of the Multidisciplinary AI Team to encourage interest, and motivate uptake.

==========

12. The extent to which the Framework should include governance issues, including but not limited to make up of design and development teams, monitoring and evaluation, and grievance and redress.

Response:
AI Ethics Board is required.

AI Ethics Board is an independent Forum of both internal and external stakeholders, not necessarily AI Experts, but a mix of AI experts, non-AI persons, legal, privacy advocates, ethics and people from communities that are under-represented.

The AI Ethics Board should perform the oversight function of reviews and assessment, and impact analysis of how a new AI application or requirements for a new AI application may impact individuals, groups and communities; and/or their uses or purpose of use may impact individuals, groups and communities.

They should also advise on how best to design, develop and use the AI application in order to avoid harmful bias, discrimination and social and racial injustice.

Suggested Change: AI Ethics Board

Regards,
Cyril

Dr. Cyril Onwubiko
Director, Enterprise Security Architecture
Chief Information Security Office (CISO)
Pearson Plc
80, Strand, London, United Kingdom, WC2R 0RL

Personal: https://www.c-mric.com/cyril

Learn more at pearson.com

Planned leaves:

All comments will be made public as-is, with no edits or redactions. Please be careful to not include confidential business or personal information, otherwise sensitive or protected information, or any information you do not wish to be posted.

Comment Template for Responses to NIST Artificial Intelligence Risk Management Framework Request for Information (RFI)

Submit comments by August 19, 2021:

General RFI Topics (Use as many lines as you like)	**Response #**	**Responding organization**	**Responder's name**	**Paper Section (if applicable)**	**Response/Comment (Include rationale)**	**Suggested change**
			Cyril Onwubiko, PhD			

Responses to Specific Request for information (pages 11,12, 13 and 14 of the RFI)	Page 12				"NIST is among the institutions addressing these issues. NIST aims to cultivate the public's trust in the design, development, use, and evaluation of AI technologies and systems in ways that enhance economic security, and improve quality of life." We would request you to change "economic security" to "public confidence". Rationale: AI risk should not be focused on "economic security" rather on ensuring "public trust and confidence" in the way AI applications are designed, deployed and used.	NIST is among the institutions addressing these issues. NIST aims to cultivate the public's trust in the design, development, use, and evaluation of AI technologies and systems in ways that enhance public confidence, and improve quality of life.
1. The greatest challenges in improving how AI actors manage AI-related risks – where "manage" means identify, assess, prioritize, respond to, or						

communicate those risks;						
2. How organizations currently define and manage characteristics of AI trustworthiness and whether there are important characteristics which should be considered in the Framework besides: accuracy, explainability and interpretability, reliability, privacy, robustness, safety, security (resilience), and mitigation of harmful bias, or					AI trustworthiness should include 'data representativeness'. Rationale: If the dataset on which an AI model is trained is not 'representative' of the problem domain, then the model can still be explainable and interpretability, and still yet harmfully biased. It will be biased on the type of data it's trained upon. For example, if an AI model is trained against dataset of a certain demographic, then it's extremely likely to be biased. So, data representativeness should be a mandatory requirement for trustworthiness.	AI trustworthiness characteristics should include: accuracy, explainability and interpretability, representative (data), reliability, privacy, robustness, safety, security (resilience), and mitigation of harmful bias, or harmful outcomes from misuse of the AI.

harmful outcomes from misuse of the AI;						
3. How organizations currently define and manage principles of AI trustworthiness and whether there are important principles which should be						

considered in the Framework besides: transparency, fairness, and accountability;						
4. The extent to which AI risks are incorporated into different organizations' overarching enterprise risk management – including, but not limited to, the management of risks related to cybersecurity, privacy, and safety;					It should include the management of risks related to social and racial.	The management of risks related to cybersecurity, privacy, safety, social and racial.

5. Standards, frameworks, models, methodologies, tools, guidelines and best practices, and principles to identify, assess, prioritize, mitigate, or communicate AI risk and whether any currently meet the minimum attributes described above;						
6. How current regulatory or regulatory reporting requirements (e.g., local, state, national, international) relate to the use of AI standards, frameworks, models, methodologies, tools, guidelines					New regulations are being enacted to address AI specific risks, such as responsible use, avoidance of harmful bias, social and racial injustice. For example, the European Union has proposed an AI Regulation, to address High-Risk AI applications. The regulation can be found here https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN	AI Regulation for trustworthiness, responsible use, transparency, fairness, equity and accountability

and best practices, and principles;						
7. AI risk management standards, frameworks, models, methodologies, tools, guidelines and best practices, principles, and practices which NIST should consider to ensure that the AI RMF aligns with and						

supports other efforts;						

8. How organizations take into account benefits and issues related to inclusiveness in AI design, development, use and evaluation – and how AI design and development may be carried out in a way that reduces or manages the risk of potential negative impact on individuals, groups, and society.					Organisations are demanding diversity and inclusiveness in AI design, development, use and evaluation to reduce risk of potential negative impact on individuals, groups and society as follows: 1) Diversity in AI Teams: AI Teams should include "citizen-stakeholders" (that is, non-AI personnel/experts), who are knowledgeable in other domains, e.g., Product Managers, Data Owners, Architects, Engineers, Linguists to the AI Team, so that it's a multidisciplinary team. 2) Develop AI Training Package/Program for staff who are interested so that more people are involved. 3) Ensure under-represented communities and people (e.g., Women, Blacks, Asian, Hispanics etc) are part of the AI Team. 4) Form an AI Ethics Board, as an accountability structure, to hold both the organisation and the AI Team accountable.	AI Ethics Board

9. The appropriateness of the attributes NIST has developed for the AI Risk Management Framework. (See above, “AI RMF Development and Attributes”);						

10. Effective ways to structure the Framework to achieve the desired goals, including, but not limited to, integrating AI risk management processes with organizational processes for developing products and services for better outcomes in terms of trustworthiness and management of AI risks. Respondents are asked to identify any current models which would be effective. These could include – but are not limited to – the NIST Cybersecurity Framework or Privacy						

Framework, which focus on outcomes, functions, categories and subcategories and also offer options for developing profiles reflecting current and desired approaches as well as tiers to describe degree of framework implementation; and						

11. How the Framework could be developed to advance the recruitment, hiring, development, and retention of a knowledgeable and skilled workforce necessary to perform AI-related functions within organizations.					1) Develop AI Trainings Package/Program for staff who are interested so that more people are involved 2) Diversity in AI Team - include 'citizen-stakeholders' (that is, non-AI personnel), but who are knowledgeable domain experts in their own areas, e.g., Product Managers, Data Owners, Architects, Engineers to the AI Team, so that they can train and learn on the job (up-skill and re-skilling) 3) Ensure under-represented communities and peoples (e.g., Women, Blacks, Asian, Hispanics etc) are part of the AI Team to encourage interest, and motivate uptake.	1) Develop AI Trainings Package/Program for staff who are interested so that more people are involved 2) Diverse Multidisciplinary Cross-Functional AI Team: The AI Team should include 'citizen-stakeholders' (that is, non-AI personnel), but who are knowledgeable domain experts in their own areas, e.g., Product Managers, Data Owners, Architects, Engineers to the AI Team, so that they can train and learn on the job (up-skill and re-skilling) 3) Ensure Under-Represented Communities and Peoples (e.g., Women, Blacks, Asian, Hispanics etc) are part of the Multidisciplinary AI Team to encourage interest, and motivate uptake.

12. The extent to which the Framework should include governance issues, including but not limited to make up of design and development teams, monitoring and evaluation, and grievance and redress.					AI Ethics Board is required. AI Ethics Board is an independent Forum of both internal and external stakeholders, not necessarily AI Experts, but a mix of AI experts, non-AI persons, legal, privacy advocates, ethics and people from communities that are under-represented. The AI Ethics Board should perform the oversight function of reviews and assessment, and impact analysis of how a new AI application or requirements for a new AI application may impact individuals, groups and communities; and/or their uses or purpose of use may impact individuals, groups and communities. They should also advise on how best to design, develop and use the AI application in order to avoid harmful bias, discrimination and social and racial injustice.	AI Ethics Board