To: National Institute of Standards and Technology

From: OpenResearch

Subject: Feedback on the Initial Draft of the AI Risk Management Framework

About OpenResearch

[OpenResearch](#) is a not-for-profit research lab dedicated to complex long-term societal innovation — for example, work that requires an expansive time horizon, seeks to answer very open-ended questions, or develops technology that benefits humanity. The core goals of OpenResearch are to consider how technological advances affect social outcomes, to think and advise broadly about how to minimize the negative effects, and to promote equitable distribution of benefits across all individuals and communities.

OpenResearch has recently developed a new research initiative focused on better understanding the economic impacts of increasingly advanced artificial intelligence and evaluating policy options to inform decision making at leading AI labs and by public policy makers. The primary aim of this initiative is to use social science and public policy research on the economic impacts of AI to help foster a future where AI advancement leads to widely shared economic prosperity. In this capacity we humbly respond to the request for feedback on the NIST AI Risk Management Framework.

OpenResearch applauds the NIST's commitment to an open, multi-stakeholder process for soliciting feedback on the RMF and is grateful for the opportunity to provide input. Below, we offer brief comments on the initial draft for the NIST to consider.

RMF Feedback

In response to feedback prompt #1: "Whether the AI RMF appropriately covers and addresses AI risks, including with the right level of specificity for various use cases" and #7: "What might be missing from the AI RMF"

We encourage NIST to consider adding "Equity" to its list of socio-technical characteristics, and to include specific descriptions of the risks that advanced AI systems pose to inclusive economic prosperity in the absence of equity-focused governance and decision making. These risks include driving increased wage inequality (Acemoglu & Restrepo, 2021; Jaumotte et al., 2013; Moll et al., 2022), accelerating international economic inequality (Korinek et al., 2021), and exacerbating existing inequalities in access to

the technological tools and capacity required to benefit from state of the art AI systems (Weidinger et al., 2021). Numerous economists have identified AI's potential to exacerbate societal inequities as a leading economic challenge related to continued AI advancement (Autor et al., 2022; *Congressional Testimony of Erik Brynjolfsson*, 2019; Korinek et al., 2021). These effects, however, are not unavoidable impacts of technical AI progress. Rather, they can be mitigated through improved decisions about how to responsibly develop, deploy, and use AI systems as a means of generating inclusive economic growth and equitable access to opportunity (Acemoglu, 2021; Partnership on AI, 2021). Given the depth of scholarly evidence on the challenges that AI progress presents for economic and societal equity, we believe that grounding decisions related to increasingly advanced AI systems in concerns for equity is a critical aspect of responsible AI development.

Highlighting these risks with tangible descriptions of the mechanisms through which they could result in harmful societal impacts would be an additive contribution to the RMF. Explicitly articulating the impact that human decisions can have to either cause or mitigate increased economic inequality within the RMF could incentivize increased consideration for the societal impacts of specific AI applications. Furthermore, providing direction towards resources that can help users examine risks related to inequality as part of AI system deployment or system adoption decisions would increase the practical value of the RMF.

Conclusion

OpenResearch is grateful for the opportunity to provide feedback on the RMF related to the economic impacts of increasingly advanced AI systems and is eager to support future efforts by NIST to develop responsible AI frameworks and practical guides. Please contact Sam Manning with any questions.

References

Acemoglu, D. (Ed.). (2021). *Redesigning AI: Work, democracy, and justice in the age of automation*.

    Boston Review.

Acemoglu, D., & Restrepo, P. (2021). *Tasks, Automation, and the Rise in US Wage Inequality* (No.

    w28920; p. w28920). National Bureau of Economic Research. https://doi.org/10.3386/w28920

Autor, D., Mindell, D. A., & Reynolds, E. B. (2022). *The Work of the Future: Building Better Jobs in an Age*

    of Intelligent Machines. The MIT Press. https://doi.org/10.7551/mitpress/14109.001.0001

Congressional Testimony of Erik Brynjolfsson, (2019) (testimony of Erik Brynjolfsson).

    https://www.congress.gov/116/meeting/house/109981/witnesses/HHRG-116-SY15-Wstate-Brynjol

fssonE-20190924.pdf

Jaumotte, F., Lall, S., & Papageorgiou, C. (2013). Rising Income Inequality: Technology, or Trade and
Financial Globalization? *IMF Economic Review*, *61*(2), 271–309.
https://doi.org/10.1057/imfer.2013.7

Korinek, A., Schindler, M., & Stiglitz, J. E. (2021). *Technological Progress, Artificial Intelligence, and
Inclusive Growth* [IMF Working Paper].

Moll, B., Rachel, L., & Restrepo, P. (2022). *Uneven Growth: Automation's Impact on Income and Wealth
Inequality*. https://benjaminmoll.com/wp-content/uploads/2019/07/UG.pdf

Partnership on AI. (2021). *Redesigning AI for Shared Prosperity: An Agenda*. Partnership on AI.
https://partnershiponai.org/paper/redesigning-ai-agenda/

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B.,
Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J.,
Rimell, L., Hendricks, L. A., … Gabriel, I. (2021). Ethical and social risks of harm from Language
Models. *ArXiv:2112.04359 [Cs]*. http://arxiv.org/abs/2112.04359