



Future of Life Institute
PO Box 706
Allston, MA, 02134

March 24, 2020

SUBJECT: Support for H.R. 6216, the National Artificial Intelligence Initiative Act of 2020

Dear Chairwoman Johnson, Ranking Member Lucas, and Representatives McNerney, Olson, Lipinski, and Weber,

On behalf of the Future of Life Institute (FLI), we commend you for the introduction of the National Artificial Intelligence Initiative Act of 2020 (H.R. 6216). Since our inception in 2014, FLI has led and participated in numerous efforts to improve the governance of AI technologies nationally and internationally, and in both the private and public sectors. With this perspective, **we are very pleased to offer our full support for the National AI Initiative Act.** While also supporting American leadership in AI research, we believe this legislation will help ensure the safe and ethical research, development, and use of AI technologies to maximize their benefits for all of humanity. There are many commendable provisions in the National AI Initiative Act, but we are particularly thankful that you had the foresight to include the following.

Support for AI safety research and development. FLI has managed two successful rounds of philanthropically-backed grant funding for AI safety research. We are therefore intimately aware of how significant federal investment in AI research and development can improve the safety and trustworthiness of AI systems, not just to advance the raw capabilities of AI systems. Thankfully, this legislation appropriately prioritizes AI safety research across the National Institutes of Standards and Technology (NIST), the National Science Foundation (NSF), and the Department of Energy (DOE). For each agency, this is most readily apparent in Sections 301(a)(1)(D), 401(c)(1)(F), and 501(b)(7), respectively, which identify for funding support key research areas of the AI safety field, including “assurance, verification, validation, security, and control.” Advancing the field of AI safety through federal investments will help fulfill the stated purpose of the National AI Initiative to have America “lead the world in the development and use of trustworthy artificial intelligence systems in the public and private sectors” (Sec. 101(a)(2)).

Ethics review requirements. FLI fully agrees with the sense of Congress, as provided in the legislation, that “the incorporation of ethical, social, safety, and security considerations into the research design and review process for Federal awards may help mitigate potential harms before they happen” (Section 401(c)(3)(A)(ii)). Thus, we are very supportive of requiring ethical statements to be submitted by researchers seeking AI research funding from the NSF or by DOE (Sections 401(c)(3)(B) and 501(e)(1), respectively). As intended, we believe these sensible requirements will help AI researchers and policymakers reflect on potential risks associated with



Future of Life Institute
PO Box 706
Allston, MA, 02134

AI, and therefore encourage the safe and ethical development of AI capabilities *before* harm is caused.

A mandate for a NIST risk framework. FLI strongly supports evaluating AI systems' benefits and costs through a thoughtful, proactive risk-based approach. Without proactive risk management, the development of several industries could be hindered because of the loss of public trust following accidents or failures of AI systems. However, it is unfortunately *not* a simple task to evaluate the potential risks of increasingly capable AI systems. As has been noted by the Office of Management and Budget (OMB) and the Office of Science and Technology Policy (OSTP), one of the more challenging aspects of AI is that "there is a risk that [an AI system's] pursuit of its defined goals may diverge from the underlying or original human intent and cause unintended consequences—including those that negatively impact privacy, civil rights, civil liberties, confidentiality, security, and safety."¹ Therefore, we are especially grateful for your inclusion of Section 301(b) in this bill. This provision requires NIST to develop a "voluntary risk management framework for the trustworthiness of artificial intelligence systems" that will include, among other elements, "standards, guidelines, best practices, methodologies, procedures, and processes for assessing the trustworthiness of, and mitigating risks to, artificial intelligence systems." There are many important potential benefits that will come from this risk framework, but to highlight two benefits of greatest interest to FLI, it will a) assist *any* public or private sector AI researchers as they consider the ethics of their research, even if they are not required to do as a condition of receiving grant support, and b) support federal agencies' compliance with the draft guidance from OMB on regulating AI applications, which depends heavily on the ability of federal agencies to use risk assessment methodologies that will be researched and developed by NIST under this provision.

Thank you again for your introduction of this important legislation. Please contact Jared Brown at jared@futureoflife.org should you require any further assistance.

Sincerely,

A handwritten signature in black ink, appearing to read "Max Tegmark", written in a cursive style.

Prof. Max Tegmark
President, Future of Life Institute

¹ Office of Management and Budget. (2020). Guidance for Regulation of Artificial Intelligence Applications (DRAFT), p. 12, available at <https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf>.