EI

March 27, 2024

The Honorable Gina Raimondo
Secretary, Department of Commerce
1401 Constitution Ave. NW
Washington, DC 20230

RE: Openness in AI is critical for independent Responsible AI professionals

Dear Secretary Raimondo,

We, the undersigned AI ethics and governance practitioners collaborating under the banner of Ethical Intelligence, write to emphasize several key benefits of openness and transparency in AI models of particular importance to the startups and smaller businesses, companies outside the tech sector, and specialized innovators that make up most of our collective clients.

We appreciate the opportunity to provide comments to the National Telecommunications and Information Administration's Openness in AI Request for Comment. We applaud the NTIA's efforts in engaging a broad range of stakeholders on this important topic at this pivotal moment in AI changing the way we work and penetrating daily life all over the world.

Ethical Intelligence (EI) is an independent consulting group providing Responsible AI (RAI) ethics and governance services to diverse clients across industries, as well as support the global EI network community of RAI experts and practitioners. We leverage our cross-functional, multi-disciplinary, and international knowledge and expertise to guide the adoption and integration of AI responsibly throughout our client companies' business, culture and governance practices.

We are closely following the public debates and latest research on the advantages and disadvantages of openness in AI regarding cybersecurity, national competitiveness, and intellectual property. There are differing opinions within the EI network on these questions - we collectively contribute to bringing greater clarity to these issues and intend to continue doing so as the norms take shape and are standardized.

These comments are focused rather on the societal implications relating to development of a more level playing field for innovation in the AI ecosystem; the extensibility of AI systems for diversity and specialization; and the impartiality of AI ethics and governance oversight. Structural policy enabling greater access to knowledge is critical for the last mile of RAI implementation that constitutes our work.

In this letter, we would like to highlight three aspects in the discussion of openness in AI of critical importance to independent RAI professionals and the companies we work with: (1) *stimulating growth and innovation in small businesses*; (2) *facilitating customized solutions and diverse applications*; and (3) *enabling the independent evaluation of AI systems*.

1. **Openness in AI stimulates growth in smaller companies and allows for innovation in specialized use cases**.

The background document to this request for comment explains that "dual-use foundation models with widely available weights could play a role in fostering growth among less resourced actors, helping to widely share access to AI's benefits." As RAI practitioners working with enterprise companies outside of the dominant tech players, we understand the motivation to use open foundation models to enhance capabilities and productivity in their respective niche markets is one of the primary attractions of integrating AI into the way they do business. The customizability of open foundation models for more flexible adaptation methods and ability to fine-tune alignment interventions enable more viable down-stream applications.[1]

Building from their existing strengths in the market before the introduction of AI, the promise of AI is that it can make their business more efficient and strategic. Greater model access allows downstream developers in their own market context to optimize and perfect existing models to introduce innovation their own niche businesses, the nuances of which they are deeply familiar, instead of having to start from scratch.[2] Companies tend to adopt multiple different kinds of AI systems, which even if not initially, will eventually interact with each other as integration efforts scale throughout the organization.

2. **Openness in AI facilitates the customization of solutions and the diversity of applications.**

The NTIA request for comment astutely points out that "open foundation models can be readily adapted and fine-tuned to specific tasks." As we have repeatedly learned from our clients, some of the greatest value of integrating AI into a business is how those general purpose tools can be customized within their particular context and fine-tuned towards specifically tailored market solutions. The democratization of AI development with open foundation models widens the diversity of the community contributing to the development of new applications on top of the original open model architecture, which can then be integrated back into the original developer's ecosystem.

The NTIA background also further elaborates that openness in AI can "make it easier for system developers to scrutinize the role foundation models play in larger AI systems, which is important for rights- and safety- impacting AI systems (e.g. healthcare, education, housing, criminal justice, online platforms etc.)" A significant consideration in our work as independent RAI consultants is helping establish processes and other company norms that ensure that their tools do not introduce biases and perpetuate the legacy discrimination that undermines public trust their company and their industry. In addition to the ethnic, gender, and other sensitive

---

[1] Kapoor, Bommasani, et al., *On the Societal Impact of Open Foundation Models,* available at arXiv:2403.07918

[2] Seger, Dreksler, Moulange, Dardaman, Schuett, Wei, et al, *Open-Sourcing Highly Capable Foundation Models: An Evaluation of Risks, Benefits, and Alternative Methods for Pursuing Open-Source Objectives*, Centre for the Governance of AI, 2023, available at https://www.governance.ai/research-paper/open-sourcing-highly-capable-foundation-models.

protected classes that is of paramount concern, the reliability of the AI tools they've adopted to accurately and fairly account for differences is core to that trust.[3] Greater access to knowledge about open foundation models enable more robust adaptations that account for the precision of native language and the nuances of culture are essential to cooperative AI ecosystems of globalization and the interoperability of AI ethics and governance norms across jurisdictions.

3. Openness in AI enables the impartial and independent evaluation and continuous monitoring of AI systems.

Of fundamental significance to the professionalization of AI ethics and governance, that EI is helping lead, is the impartial ability to evaluate and monitor AI systems once they are deployed. External oversight plays a key part of the developing regulatory frameworks from international to local levels that have been proposed and that are being enacted. As the request for comment background document explains: "foundation models can allow for more transparency and enable broader access to allow greater oversight by technical experts, researchers, academics, and those from the security community."

The differences are vast between the ability to perform meaningful audits along the gradient from *black-box access* - which limit auditors to analyze system outputs - to white box and outside-the-box access, which give full access to the system and further information about the system's development and deployment.[4] Black-box methods are: not well-suited to develop generalized explanations; prevent system components from being studied separately; can produce misleading and unreliable results; and offer limited insights to help address failures. White box methods enable: more effective and efficient algorithms; stronger adversarial attacks; better assurances with an expanded attack toolbox; interpretability tools that aid in diagnostics; and an improved ability to explain specific AI system decisions.

Assurance audits for algorithmic systems are increasingly being included as requirements for RAI regulatory compliance. These accountability mechanisms are most reliably and systematically established with structural support for greater access to knowledge of AI systems. The societal implications for the integrity of these audits are profound, such as confronting the discrimination in hiring that algorithmic systems could even further embed into harmful employment practices that the New York City Local Law 144 is designed to address. Further guidance on ensuring the impartiality of audits - such as public disclosure, standardized audit criteria, and independent accountability - will in part depend on the protection of more open access to foundation models.[5]

---

[3] Surman, Bdeir, et al, *Accelerating Progress Toward Trustworthy AI*, Mozilla (Feb 2024), available at https://foundation.mozilla.org/en/research/library/accelerating-progress-toward-trustworthy-ai/whitepaper/

[4] Casper, Ezell, et al., *Black-Box Access is Insufficient for Rigorous AI Audits* (Jan 2024), available at https://arxiv.org/abs/2401.14446.

[5] Lam, Lange, et. al., *A Framework for Assurance Audits of Algorithmic Systems* (Jan 2024), available at https://arxiv.org/abs/2401.14908

Given the complex implications of AI that we are all only beginning to appreciate and understand, this external input is best done by a collaboration of experts trained in different disciplines, representing different life experiences, and subject to different jurisdictions. This is where EI is uniquely equipped to support companies dynamically in their transition to integrating AI into their business. We rely on openness in AI to do our job most effectively.

We thank the NTIA for this chance to contribute our perspective to its deliberations and remain available to provide further feedback.

Respectfully submitted,

Eddan Katz - AI Policy, Ethical Intelligence

Olivia Gambelin - Founder, Ethical Intelligence

Oliver Smith - EI Expert and Founder of Daedalus Futures

Flavio S. Correa da Silva - faculty member, University of Sao Paulo

Helena Ward - DPhil Candidate, University of Oxford

David Barnes - Founder, David Barnes, LLC

Eugene Fedorchenko - EI Expert, Strategist, Founder of IntFinite

Goda Mockute, AI Project Lead, Erasmus University Rotterdam

https://www.ethicalintelligence.co/