WHITE PAPER

RESPONSIBLE AI BY DESIGN, NOT FOR COMPLIANCE

Centering transparency and explainability, inclusion, and accountability through proactive, human-centered approaches to AI.

AUTHORS

Anna Swan
Qualitative Researcher
AnitaB.org

Yamelith Aguilar
Director, Research
AnitaB.org

ANITAB.ORG RESEARCH

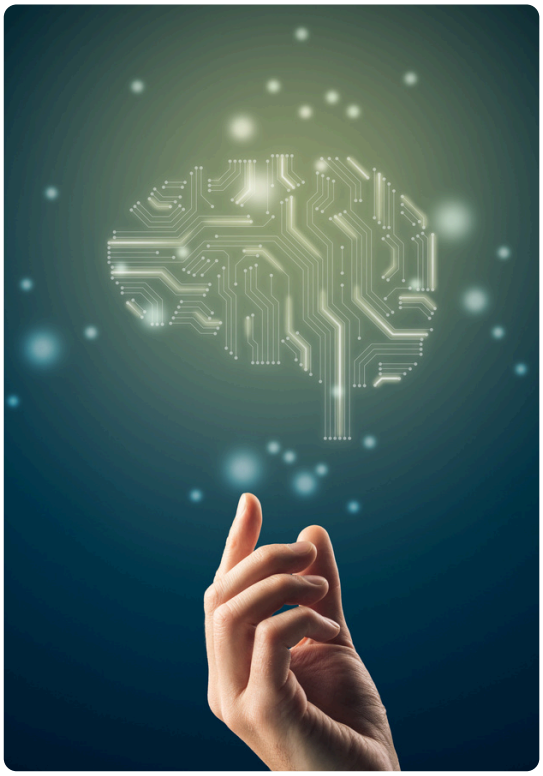
The AnitaB.org Research Team is dedicated to producing evidence-based insights into the lived experiences of those in tech. Through rigorous research and actionable data, we uncover systemic barriers, explore intersectional experiences, and identify strategies to advance workplace equity and structural change in tech.

Executive Summary

This white paper draws on insights from our AnitaB.org Tech Collaborative: The Responsible AI Forum, as well as 12 in-depth interviews with AI practitioners across sectors to examine what responsible AI means to them, how AI practitioners are implementing responsible AI, and what challenges and recommendations they have for others. It offers strategic recommendations to move beyond compliance and create AI systems rooted in transparency & explainability, inclusion, and accountability.

Key Responsible AI Principles

- **Transparency & Explainability:** Individuals should understand the inputs and subsequent outcomes of AI systems, be informed of how decisions are made by those systems, and have the ability to request additional explanation as necessary.
- **Inclusion:** Reflects the ethical evaluation of AI outcomes, including stakeholder involvement and data governance, and the integration of diverse perspectives in AI design, development, and deployment.
- **Accountability:** Reflects the obligation of all involved in AI to be responsible and liable for their actions and decisions.



Key Findings

1. AI Must Be Human-Centered to Be Truly Responsible
2. Advancing Responsible AI Requires Deep, Ongoing Interdisciplinary Collaboration
3. Bridging the AI Literacy Divide Is Critical for Responsible AI Adaptation

Recommendations

1. Reimagine responsible AI as an interdisciplinary practice, not just a technical one
2. Shift from reactive to proactive AI governance
3. Fund and integrate responsible AI literacy at all levels

Introduction

Artificial intelligence (AI) is rapidly shaping the future. As AI transforms our world, the stakes for fairness, accountability, and justice become increasingly high. However, the current guidelines and frameworks are fragmented; there is a lack of standardization and consensus on what responsible AI means in practice, and what guardrails should be in place to mitigate undue risk on all those affected by the implementation of AI systems.

This white paper is grounded in learnings from our AnitaB.org Tech Collaborative: The Responsible AI Forum, which convened leading AI practitioners, researchers, and advocates in March 2025. In addition to the event sessions, the AnitaB.org Research Team conducted 12 in-depth interviews with participating experts to further explore how responsible AI is being operationalized in real-world contexts, and conducted a literature review on responsible AI. The insights from the event sessions, interviews, and literature review form the foundation of this white paper.

For the purposes of this white paper, we define responsible AI (RAI) as an applied and values-driven approach to AI development and deployment that is human-centered, just, transparent and explainable, and grounded in ethical decision-making. Responsible AI requires interdisciplinary and community-engaged oversight, assessment of impacts, and iteration to ensure technology promotes innovation for human flourishing. We recognize that the term “AI” does not capture the required nuance—however, we use the term “AI” very broadly throughout this white paper to capture the wide-ranging scope of artificial intelligence and for the sake of brevity. We encourage practitioners to reflect on the context and specificity that makes sense to them.

The findings presented here center on three essential pillars for advancing responsible AI: human-centeredness, interdisciplinary collaboration, and AI literacy. These pillars emerged as core values that should guide every step in the process of responsible AI design, development, deployment, governance, and the consistent assessment of impact. While responsible AI has been most often discussed in terms of guiding principles, three of which we also discuss as salient components of RAI based on findings gathered through field observation and interviews, these principles do not always translate into practice. This white paper offers strategic guidance on how the themes discussed can be translated into action, ensuring AI actively advances social progress and delivers meaningful benefits for all.

Responsible AI Principles

Responsible AI is typically defined by guiding principles that ensure AI is developed in accordance with both ethical and business-oriented goals[1], and that harm is minimized[2]. Key informants at the Responsible AI Forum were asked what principles they followed working with AI responsibly. Of all the principles discussed, the following were the most frequently mentioned: transparency and explainability, inclusion, and accountability. We define these principles below, drawing on both existing RAI literature and excerpts from interviews with AI practitioners.

Transparency & Explainability

Transparency and explainability are intertwined, enabling the clarity necessary to ensure human rights are protected and individuals are treated fairly. According to UNESCO's Recommendations on the Ethics of Artificial Intelligence[3], individuals should understand the inputs and outcomes of AI systems, be informed of how decisions are made by those systems, and have the ability to request additional explanation as necessary.

"Responsible AI requires a governance-first cultural shift. Explainability is fundamental to building trust for long-term customer adoption"

- Usha Jagannathan, Director of AI Products

Inclusion

Although not always identified as a core principle of responsible AI, inclusion (or inclusiveness) is a key component of a human-centric approach to AI for social good[4]. Inclusion is "the process of 'proactively involving and representing the most relevant humans with diverse attributes, those who are impacted by and have an impact on the AI ecosystem context'"[5]. It reflects the ethical evaluation of AI outcomes, including stakeholder involvement and data governance, and the integration of diverse perspectives in AI design, development, and deployment.

"Businesses that are developing AI products...should have folks at the table that are representatives of the communities who are being affected by those tools."

- Jordan Loewen-Colón, Adjunct Assistant Professor of AI Ethics and Policy

Accountability

Accountability reflects the obligation of all involved in AI to be responsible and liable for their actions and decisions. In the context of RAI, accountability is comprised of four primary goals: compliance (ensuring AI systems align with legal and ethical standards), report (properly recording and justifying its use), oversight (evaluation of conduct), and enforcement (determining consequences according to evidence)[6].

"Who is responsible if something goes wrong? Nobody wants to be the point person. But in any organization the buck needs to stop somewhere."

- Mellini Monique, Founder and Voice Consultant

Key Finding

AI Must Be Human-Centered to Be Truly Responsible

Human-centered AI (HCAI) leverages data to support social progress and human thriving. It goes beyond legal compliance, and considers social, historical, and cultural context to ensure that AI is created with values-driven purpose[7]. The design of HCAI serves the needs of all stakeholders[8] and embeds human perspective throughout and beyond the development of AI[9]. It demands a shift in power, away from a narrow group of technical experts and towards a broader, more inclusive form of participation and accountability.

Humans should be centered at every stage of development and deployment[10], and this includes acknowledging how biases shape AI. AI systems have demonstrated extreme value in areas such as healthcare, where they have been leveraged to achieve more accurate and efficient diagnoses and treatments. However, datasets used to train algorithms may replicate patterns of discrimination in medicine, therefore further disadvantaging some patients[11]. One example of bias in action was an algorithm used to assess patient need. Because the training data was based on past spending, and Black patients historically had both less to spend on healthcare and more barriers to access, the algorithm indicated that White patients were in need of more dedicated care[12]. This was corrected, but cases like these indicate the need to preemptively mitigate harms; leaders in AI development must acknowledge how their own perspectives shape what, how, and for whom they build AI tools.

HCAI also means adapting models to be localized and culturally relevant. Machine translation facilitates communication worldwide. However, low-resourced languages are often left behind in training datasets, as they are perceived as less valuable and thus inform fewer AI models. Speakers of these low-resourced languages have pointed out that machine translation has the potential to promote the preservation of their languages, cultures, and identities. For translation services to be human-centered, AI must be designed with all local communities in mind, particularly those that are often overlooked[13].

Human centeredness requires a continuous evaluation of real-world impact. For example, before its purchase and transformation, Twitter users raised concerns about Black individuals being excluded from images created by the platform's cropping algorithm. To address these concerns, Twitter invited researchers, activists, and developers to participate in a community audit or "bias bounty" competition, resulting in the reassessment of the training data embedded in their saliency model[14]. While there are limitations to this type of intervention, the platform re-centered human values to counter harmful unintended outcomes of an AI tool.

"Many want to remain elitist, but if we continue to leave so many communities out, I believe that the use of AI is going to implode, and won't be in our best interest. And shouldn't we be taking cues from other industries already?"

- Mellini Monique, Founder and Voice Consultant

Advancing Responsible AI Requires Deep, Ongoing Interdisciplinary Collaboration

Closely related to the value of human-centeredness, interdisciplinary collaboration emerged as an essential component of RAI. Featured sessions, breakout sessions, and interviews that occurred at the Responsible AI Forum all conveyed the need to reach beyond technical expertise and invite non-technical voices to conversations around the design, development, deployment, governance, and assessment of the long-term impact of AI. The impacts of AI systems are not limited to technical domains; they ripple across entire ecosystems.

Historically, discussions of responsible AI have targeted developers. But prioritizing technical solutions can miss the broader picture; for AI to be responsible, multiple stakeholders from a variety of domains should be engaged to fully address more nuanced and structural challenges[15]. For example, an AI tool designed to rank job applicants may overvalue candidates who attended elite institutions, simply because past successful hires were from those schools. This reinforces historical biases, missing potentially promising candidates with less prestigious educational backgrounds. Without collaboration with social scientists, ethicists, or the individuals impacted by the AI tool early and often, biases like this may go unnoticed before launch. It is not that the developers or the tech are wrong, but building the system should require input from a broader pool of stakeholders.

The need for interdisciplinarity also extends to AI governance. Recently, researchers have called for increased sociotechnical approaches to governance. They emphasize that experts in the social sciences and humanities can provide perspective on “how AI-powered systems might interact with one another, with people, with other processes, and within their context of deployment in unexpected ways.”[16] In addition to auditing AI systems for things like technical safety and legal compliance, sociotechnical audits can be implemented to evaluate social and systemic impacts; audits should be conducted by both engineers as well as philosophers, environmental activists, legal scholars, and/or industry experts[17,18]. Additional efforts to mitigate risk and incorporate diverse perspectives include oversight boards and committees to advise on and translate RAI principles into practice. Microsoft’s AI, Ethics, and Effects in Engineering and Research (AETHER) Committee has assessed the internal design of AI interventions and provided recommendations to ensure harm is mitigated [19]; external boards, non-profits, and consultancies can provide similar insight.

When done intentionally, interdisciplinary collaboration surfaces hidden risks earlier, limits real-world harm, and becomes a catalyst for innovation rather than a liability or a barrier to progress. And as a wide-reaching technology, the rapid advancement of AI also requires global cooperation[20] and cultural sensitivity. The greater variety of input and oversight, the greater our chances will be to see more successes than failures.

Bridging the AI Literacy Divide Is Critical for Responsible AI Adaptation

There is a growing need for targeted AI literacy efforts that address two distinct but equally urgent gaps. Technical and non-technical stakeholders need to be equipped with the knowledge to engage responsibly and effectively with AI.

Non-technical professionals across industries must be equipped with foundational AI literacy to confidently and critically engage with AI tools. It is important that non-technical professionals understand the capabilities and limitations of AI tools and ultimately, make informed decisions about their use. This includes upskilling in areas such as prompt design, data awareness, and recognizing potential harms.

These same skills need to be taught to all non-technical individuals, starting with young students. AI knows no age limits, and young people are avid users of AI tools. A recent study found that 51% of 14–22-year-olds reported using AI[21]. Embedding AI literacy early is critical to prepare young people not only to use and understand AI, but also to engage critically with its risks, ethics, and societal impacts. This prepares the tech field for a future workforce equipped to develop and deploy AI responsibly.

At the same time, AI is rapidly changing and technical practitioners need to be upskilled to understand and inform the rapidly changing AI models. However, the greater gap in AI for technical practitioners is that developers, engineers, and data scientists often lack structured exposure to AI ethics and frameworks that educate them on societal impacts. Technical practitioners have a deep understanding of how to build AI systems, but they lack training to consider the broader consequences of those systems to critical populations, public trust, or civil liberties. Responsible AI training should be foundational and integral throughout all aspects of learning and not a singular unit in their education. Unfortunately, most available technical trainings in AI do not include responsible AI practices. That said, some organizations are taking steps toward equipping technical employees with more well-rounded skills—Google has formally trained over 30,000 employees in AI Principles and has piloted the “Moral Imagination” workshop where product teams walk through potential impacts of AI products [22]. To ensure AI is both ethical and responsible, organizations should provide a wealth of opportunities to upskill their workforce.

> *“We should learn from our past mistakes. We cannot just teach people how to use resources that are created by technology companies. We need to teach students how to create that technology and understand how it works.”*
>
> ***-Key Informant***

Challenges

Through conversations with practitioners working on and with AI, three main challenges emerged when discussing RAI. The challenges reflect systemic issues when trying to design, deploy, and govern responsible AI.

NAVIGATING MISALIGNMENT AND LACK OF STAKEHOLDER BUY-IN

Responsible AI requires collaboration across entire ecosystems. The effective implementation of responsible AI does not occur in isolation. However, the most significant challenge practitioners and experts working on AI raised is the lack of alignment around what responsible AI looks like and why it matters. One interviewee shared that it's less about convincing stakeholders to agree that Responsible AI is necessary, but rather to get the buy-in around what responsible AI practices should be incorporated. In addition, interviewees note that there is a sense of risk aversion when discussing responsible AI, including a fear of slowing down innovation, making irreparable decisions, or opening the door to scrutiny and liability that can impact stakeholder buy-in.

BALANCING INNOVATION AND RESPONSIBILITY

One interviewee posed the question, *"How do we balance innovation with responsibility?"* This sentiment was widely shared throughout both the interviews with practitioners and the breakout sessions. At the current intersection of AI, we find ourselves in a world where the key principles of AI innovation are to move fast, deliver, and disrupt. Implementing responsible AI practices is often seen as a pause in development, innovation, and profit generation. The problem with this line of thinking is about the outcomes this progress has on those left behind, the bias that isn't being accounted for, and the harm these developments are causing. How do we balance the need to innovate fast with the need to be responsible? There must be a middle ground where responsible AI is proactive and flexible, embedding transparency, inclusion, and accountability into the creative process itself. All AI practitioners and organizations building and using AI technologies need to shift how they measure success, fund research, and ultimately reward ethical insight—not just technical innovation.

MANAGING UNCERTAINTY

Experts at the Responsible AI Forum shared a growing challenge of navigating uncertainty within the responsible AI space. This uncertainty lies within the unpredictable future of AI development, as well as the lack of regulations developed by government authorities—and more importantly, the extent to which these regulations should be followed. Interviewees shared the importance of being flexible in these situations, taking comfort in not knowing everything and the need to be agile and able to adapt rapidly. They also discussed the need to be forward thinking to anticipate future challenges and mitigate what can be mitigated. Lastly, they touched on the notion of being aware of mistakes that have happened in past uses of AI to ensure those mistakes don't get repeated.

Calls to Action from Practice

Interviewees emphasized clear and urgent actions needed to move responsible AI forward. The following calls to action reflect their insights, committments, and challenges to the field.

BE PROACTIVE AND DELIBERATE

RAI principles should be embedded in practice from the beginning—not just at the start of development, but at the inception of an idea. Being proactive means working to anticipate challenges, obstacles, and impacts early on, and taking necessary steps to mitigate harm. Being deliberate means leveraging technological solutions to be problem- and context-specific. As one interviewee said: *"Why are we adding AI? Are we adding to a process? Are we improving something? Or are we just doing something to say we did something? So being deliberate about it, not just falling for the hype of doing things because it's cool."*

CREATE REGULATIONS AND GUARDRAILS

Without rules, even technology designed "responsibly" can quickly get out of control and require backpedaling to protect people, the planet, and our data. But we don't have to reinvent the wheel, we can learn and adapt from existing models of regulation. The EU AI Act[23] provides the first large-scale comprehensive legal framework for regulating AI, and some U.S. states have begun to adopt something similar. Utah's Office of Artificial Intelligence Policy (OAIP) is the first state-level agency to formalize work that prioritizes AI transparency, accountability, and garners public trust in AI[24]. Guardrails should never stifle innovation—a point that was reiterated over the course of the Forum—but should be used to ensure strategic direction, scope, consumer protection, and foster impact that is foreseeable and manageable.

Recommendations

REIMAGINE RESPONSIBLE AI AS AN INTERDISCIPLINARY PRACTICE, NOT JUST A TECHNICAL ONE

AI must be built by more than just technical practitioners; it must be co-created with ethicists, organizers, educators, and the communities that will be impacted by AI. To do so, organizations can establish interdisciplinary teams, incorporate non-technical employees in design review, support ethics and oversight boards, and implement sociotechnical audits. Establishing interdisciplinary AI collaboratives within companies or ecosystems ensures the key principles of responsible AI are deeply embedded in the design and not just thought about at the final review of the tool. This adjustment shifts responsible AI from merely being about compliance to center co-creation.

DESIGN FOR THE FUTURE: SHIFT FROM REACTIVE TO PROACTIVE GOVERNANCE

Responsibility should not be an afterthought once an AI tool has been developed. It should not be a simple checklist that is completed to acquire a level of compliance with a set regulation. Companies should invest early in structures and practices that prioritize reducing harm, promote societal wellbeing, and ultimately, that listen to the voices of those most impacted. Responsible AI is not about fixing what is broken, but rather about building systems that don't harm individuals or communities.

FUND AND INTEGRATE RESPONSIBLE AI LITERACY AT ALL LEVELS

Democratizing AI begins with democratizing the basic understanding of AI. Both non-technical and technical professionals should be educated on the fundamental responsible AI principles. There is no age limit or requirement to gain a basic understanding of AI, how to use it, how to understand the risks, and how to make ethical decisions about its use.

CONSIDER CONTEXT AND DESIGN DELIBERATELY

Organizations at the forefront of AI design and development must define what responsible AI means to them. Translating responsible AI principles into practice requires careful evaluation of context—both the context in which AI is created and deployed. Organizations should reflect on the scope of their work, considering what principles they will abide by and what each principle means in practice. Organizations must also critically engage their AI workforce in ensuring they understand and can successfully implement the principles into every aspect of their work. The impacts of these practices should then be measured, assessed, and iterated upon.

Resources

Collectives Fostering Interdisciplinary Work on AI

- Algorithmic Justice League
- Data & Society
- Responsible AI Institute
- AI Alt Lab

AI Literacy

- For K-12
  - MIT – Responsible AI for Social Empowerment and Education
- For people in Industry (both technical and non-technical)
  - Princeton – GradFUTURES Responsible AI Learning Cohort

Conclusion

Although Responsible AI has been a topic of research and recommendation for several years, we find ourselves at a pivotal moment. For decades in books and films, we speculated about intelligent machines that could operate independently, either serving or destroying humanity. As computing technologies boomed, we agonized over the future of work, and we continue to fixate on which jobs are replaceable. But the imaginations of artificial intelligence are no longer in the distant future; that future is now. AI can be —and is—a powerful tool that augments our capacity to create a world that works for everyone, where technological innovation does not further division and lead to inequality.

AI practitioners must no longer think about Responsible AI as a final audit or a regulatory checklist. From the start, AI systems must be built to recognize the dignity of the people it will touch, to innovate with human-centeredness rather than with speed and to integrate interdisciplinary co-creators into each step. In a time of declining public trust and growing harms, responsible AI is the only path toward a future where AI serves equity and advances collective human progress.

We invite all practitioners to not only apply these insights within their own spheres, but to stay actively engaged in shaping the future of responsible AI.

Explore AI in Action at GHC

Want to dive deeper into the future of responsible AI? Join us at Grace Hopper Celebration for expert-led breakout sessions and a dedicated AI track designed to spark critical conversations, showcase real-world applications, and connect you with leaders shaping the field.

Register for GHC today and be part of the movement driving ethical, inclusive innovation in tech.

References

[1] Jagannathan, U. (2024). Responsible AI: By Design. In Deep Tech Unveiled: A Coffee Table Book for CEOs. Network Science.

[2] Lynch, T. (2022 December 15). What is the Difference Between AI Ethics, Responsible AI, and Trustworthy AI? We ask our Responsible AI Leads. The Institute for Experiential AI, Northeastern University. Text

[3] UNESCO. (2022). Recommendation on the Ethics of Artificial Intelligence. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137

[4] Sargiotis, D. (2024). Fostering Ethical and Inclusive AI: A Human-Centric Paradigm for Social Impact. Available at SSRN 4879372.

[5] Zowghi, D., & Bano, M. (2024). AI for all: Diversity and Inclusion in AI. AI and Ethics, 4(4), 873-876.

[6] Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: what it is and how it works. AI & Society, 39(4), 1871-1882.

[7] Capel, T., & Brereton, M. (2023, April). What is human-centered about human-centered AI? A map of the research landscape. In Proceedings of the 2023 CHI conference on human factors in computing systems (pp. 1-23).

[8] Sun, J., Liao, Q. V., Muller, M., Agarwal, M., Houde, S., Talamadupula, K., & Weisz, J. D. (2022, March). Investigating explainability of generative AI for code through scenario-based design. In Proceedings of the 27th International Conference on Intelligent User Interfaces (pp. 212-228).

[9] Friedrich, J., Brückner, A., Mayan, J., Schumann, S., Kirschenbaum, A., & Zinke-Wehlmann, C. (2024). Human-centered AI development in practice—Insights from a multidisciplinary approach. Zeitschrift Für Arbeitswissenschaft, 78(3), 359-376.

[10] Gray, J., & Witt, A. (2021). A feminist data ethics of care for machine learning: The what, why, who and how. First Monday.

[11] Hoffman, S., & Podgurski, A. (2019). Artificial intelligence and discrimination in health care. Yale J. Health Pol'y L. & Ethics, 19, 1.

[12] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453.

[13] Costa-Jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., ... & NLLB Team. (2022). No language left behind: Scaling human-centered machine translation. arXiv preprint arXiv:2207.04672.

[14] Lopez, P. (2023). Power and Resistance in the Twitter Bias Discourse. Algorithmic Regimes: Methods, Interactions, and Politics.

[15] Ryan, M., De Roo, N., Wang, H., Blok, V., & Atik, C. (2024). AI through the looking glass: an empirical study of structural social and ethical challenges in AI. AI & SOCIETY, 1-17.

[16] Bogen, M., & Winecoff, A. (2024). Applying a sociotechnical approaches to AI governance in practice. Center for Democracy and Technology. https://cdt.org/wp-content/uploads/2024/05/2024-05-23-AI-Gov-Lab-applying-sociotechnical-approaches-to-ai-governance-in-practice-final.pdf

[17] Mökander, J. (2023). Auditing of AI: Legal, ethical and technical approaches. Digital Society, 2(3), 49.

[18] Schiff, D. S., Kelley, S., & Camacho Ibáñez, J. (2024). The emergence of artificial intelligence ethics auditing. Big Data & Society, 11(4), 20539517241299732.

[19] Newman, J. C. (2020). Decision Points in AI Governance: Three Case Studies Explore Efforts to Operationalize AI Principles. UC Berkeley, Center for Long-Term Cybersecurity. https://cltc.berkeley.edu/wp-content/uploads/2020/05/Decision_Points_AI_Governance.pdf

[20] Coral, L. & Tavenner, E. (2024, November 20). To Ensure a Responsible AI Future, We Need Global Cooperation and Multidisciplinary Dialogues To Complement and Harness the Power of the International Network of AI Safety Institutes. New America. https://www.newamerica.org/oti/blog/to-achieve-global-ai-cooperation-we-need-multidisciplinary-dialogues-in-addition-to-the-international-network-of-ai-safety-institutes/

[21] Common Sense Media. (2024). Teen and Young Adult Perspectives on Generative AI: Patterns of Use, Excitements, and Concerns. Common Sense Media, Hopelab, Center for Digital Thriving at the Harvard Graduate School of Education. https://digitalthriving.gse.harvard.edu/wp-content/uploads/2024/06/Teen-and-Young-Adult-Perspectives-on-Generative-AI.pdf

[22] Croak, M. & Gennal, J., (2022, July 6). An update on our work in responsible innovation. Google: The Keyword. https://blog.google/technology/ai/an-update-on-our-work-in-responsible-innovation/

[23] European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 12 July 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Official Journal of the European Union, L 234, 1–83. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2024:234:TOC

[24] Loewen-Colón, J., Odubela, A., & Jordan, J. (2025). Implementing an AI evaluation framework. Aspen Policy Academy. https://aspenpolicyacademy.org/project/implementing-an-ai-evaluation-framework/

ACKNOWLEDGEMENTS

Volunteer Researchers

- Rajashree Dahal, Ph.D. Student, Natural Language Processing Lab, University of Illinois Chicago
- Chelsi Miller, Graduate Student, Northwestern University

Interviewees

- Jen Boyce, Staff Applied AI/ML Scientist, Sprout Social
- Patricia M. Clay, Associate Vice President for Information Technology and CIO, Hudson County Community College
- Dr. Usha Jagannathan, Director of AI Products, IEEE SA
- Kathy Jordan-Baker, Principal, Mavisworks, LLC
- Tina Hinch, Co-founder & Co-lead, Vision 1948 – Bringing Gender Parity Back to Tech
- Jordan Loewen-Colón, Adjunct Assistant Professor of AI Ethics and Policy at Queen's University's Smith School of Business
- Mellini Monique, Voice Consultant, Humans & AI; Founder, Vocal Culture Garden / Ambassador, Open Voice Trusmark Initiative
- Sailakshmi Santhanakrishnan , Head of ESG and Sustainability , Amazech (Dallas , Texas)
- Katie Steen-James, Senior U.S. Policy Manager, Open Source Initiative
- Kelsey Steven, Instructor at Ada Developers Academy

About AnitaB.org

At AnitaB.org, we envision a future where economic opportunity is accessible to all, and leadership reflects the full spectrum of talent and experience. As a solution-oriented convener, we design opportunities that fuel human potential, ensuring that individuals are seen, heard, and equipped to thrive in evolving workplaces.

What sets us apart is our unwavering commitment to disrupting the status quo while honoring the full humanity of every individual. We cultivate spaces where people can be vulnerable, seen, and supported without judgement. Because progress starts with truth, empathy, and the courage to build something better for all.