



Stanford University
Human-Centered
Artificial Intelligence

October 2026

Beyond Pilots and PR: What Makes AI-for-Good Partnerships Work

Vera Lummis
Angela Aristidou
Rob Reich



Table of Contents

Authors	3
About Stanford HAI	4
Acknowledgments	4
Executive Summary	5
Introduction	8
1. What types of support can nonprofits access through these partnerships?	10
What are the benefits of different types of support?	10
2. How do AI labs and nonprofits form bespoke partnerships?	12
3. What motivates nonprofits and AI labs to form partnerships?	14
Key motivations for nonprofits	14
Key motivations for AI labs	15
4. What factors contribute to effective AI lab–nonprofit partnerships?	18
Nonprofit-side enabling factors	18
Lab-side enabling factors	20
Partnership design and management success factors	21
5. What causes friction in these partnerships?	23
6. What is the role of institutional funders?	26
Where funders are adding value	26
Where funders are falling short	27
7. Recommendations	29
Appendix A - Methodology and Limitations	31
Appendix B - A Diagnostic for Nonprofits on What Support to Seek	33
Appendix C - Critical Factors Enabling Success	35
Endnotes	37

Authors

Vera Lummis is a master's candidate in international policy at Stanford University, specializing in cyber and AI policy, and a member of the AI Policy Working Group at the Stanford Institute for Human-Centered AI (HAI). Her work focuses on shaping the partnerships, tools, and philanthropic investments that help organizations in low- and middle-income countries leverage data, technology, and AI to strengthen public systems and frontline service delivery. She previously led the digital innovation portfolio at GiveDirectly and worked at IDinsight, and has consulted on AI for good initiatives across the nonprofit, philanthropic, and startup sectors. She holds a bachelor's degree in public and international affairs from Princeton University.

Angela Aristidou speaks, writes, and advises about the real-world deployment of AI technologies. Her research program examines the deployment of AI-driven innovations and spans the contexts of health, higher education, nonprofit, humanitarian aid, and technology in the UK, United States, Canada, and several Asian countries. She specializes in strategy and entrepreneurship at University College London's School of Management, is a faculty fellow at the Stanford Digital Economy Lab and Stanford HAI, and holds degrees from Cambridge and Harvard.

Rob Reich is the McGregor-Girand Professor of Social Ethics of Science and Technology at Stanford University. His home appointment is in political science, and he has courtesy appointments in philosophy, law, the Graduate School of Education, and the Freeman Spogli Institute for International Studies. He is an associate director and senior fellow at Stanford HAI and director of the Ethics, Society, and Technology Initiatives at the Center for Ethics in Society. In 2024-25, he was on public service leave to work as senior adviser to the United States AI Safety Institute at the Department of Commerce. His scholarship in political theory engages with the work of social scientists and engineers. He has written widely about philanthropy and his current work is focused on governance of frontier science and technology.

About Stanford HAI

Stanford University's Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices. The views expressed in this white paper reflect the views of the authors.

Acknowledgments

The authors would like to thank Mahishaa Balraj, Alessandro Balzi, Malou van Draanen Glismann, Haolie Jiang, and Dominic Zappia for their thought partnership and support on formative research; Elena Cryst and Caroline Meinhardt for their editorial contributions; Joanna Burr for supporting data analysis for the study; and Alex Armira, Chris Ellis, Nancy King, Shana Lynch, and Michi Turner for preparing the publication. We are grateful above all to the interviewees who gave their time to this study.

Executive Summary

Frontier AI labs and nonprofit organizations are increasingly partnering to deploy frontier AI in support of nonprofits' missions to improve global health, climate, humanitarian, and development outcomes. Individual and institutional funders have invested hundreds of millions of dollars to support nonprofits in deploying AI for good, occasionally in formal partnership with AI labs. However, the rapidly growing AI for Social Good (AI4SG) ecosystem lacks a clear understanding of what these partnerships provide, how they form, and what drives or compromises their success.

Drawing on interviews with 36 individuals across 24 organizations (frontier AI labs, nonprofits, philanthropic funders, and independent experts) and four in-depth case studies, this report offers insights into how nonprofits and frontier AI labs partner to deploy AI to produce social impact and provides recommendations for how partnership and funding arrangements can be structured for success.

Key findings:

- 1. Nonprofits and AI labs partner by bridging overlapping yet, in successful collaborations, complementary goals.** Nonprofits primarily seek technical expertise to deploy and evaluate AI, insight into how AI can advance their mission and impact, and financial resources. AI labs primarily seek to demonstrate social impact, learn from real-world deployment in low-resource settings, improve their models, and drive scaled adoption of their technologies.
- 2. AI labs tend to form their most substantive partnerships with high-capacity nonprofits that stand to benefit from specialized support.** Brokered largely through existing relationships and funder introductions, these partnerships favor well-connected organizations that possess strong in-house product, engineering, and data capacity and a large geographic footprint. Because these attributes are relatively uncommon, this pattern widens the AI capacity and opportunity gap for the majority of nonprofits.
- 3. Contrary to the narrative that benefits flow primarily to nonprofits, AI labs also gain significant value from these partnerships.** Key benefits include access to users and real-world use cases, evaluations and benchmarks in last-mile settings, opportunities to advance the research frontier, and increased public and government trust. The presence of multi-directional benefits, rather than a one-directional donor-beneficiary structure, has important implications for how partnerships are governed.
- 4. Success factors and sources of friction cluster at the nonprofit, AI lab, and partnership levels.** On the nonprofit side, success is driven by in-house technical capacity and data readiness; a clearly owned problem definition and product vision; measurement and iteration; localization and participatory design; and the ability to advocate for needed support. On the lab side, success is driven by effective scoping of nonprofit needs, and

right-fit technical support on evaluations, cost, and model customization. Aligned or complementary values and designing for handoff of the AI solution from the outset further enable success. Leading sources of friction in these partnerships include fiscal uncertainty as subsidized support winds down; unclear or informal partnership terms and governance structures; technosolutionist pressure; misaligned commercial and reputational incentives; differing ways of working and time horizons; and reliance on individual “champions” over institutional processes.

5. Partnerships have significant room to strengthen AI governance, formalize expectations, and support nonprofits’ capacity to maintain solutions after the partnership ends. Many partnerships operate on trust-based or nonbinding agreements that understate the actual scope of the collaboration. Organizations infrequently formalize strong governance practices — such as external audits, public consultation, and transparency around data and AI use — that would help nonprofits adopt AI responsibly. In addition, nonprofits often bear the burden of ensuring partnerships include design features that support long-term sustainability of AI solutions.

6. Philanthropic funders play a vital role by convening partnerships, coordinating to reduce duplication and fund shared public goods, and investing directly. However, major foundations can in some cases concentrate resources on a smaller set of high-capacity organizations, invest in innovative tools rather than the enabling infrastructure and capabilities required for impact, and create funding strategies that can influence nonprofits’ choices, such as tying grants to specific AI lab partnerships.

Key recommendations:

This report aims to inform how AI labs, nonprofits, philanthropic funders, and intermediary organizations can more effectively design partnerships that support nonprofits in accessing, integrating, and evaluating AI for social good.

For AI labs and other organizations providing AI-related support

- Conduct in-depth scoping to understand organizational needs and technical capacity, reassessing as the partnership evolves.
- Prioritize nonprofit capacity-building over short-term deployments, devoting engineering time to durable tooling, data pipelines, and evaluation capabilities nonprofits can maintain independently.
- Establish plans from the outset to taper support and sustain solutions as partnerships conclude.
- Standardize partnership terms, including expected contributions and governance requirements.

For nonprofit and mission-driven organizations

- Build basic in-house technical capability before entering a bespoke partnership.
- Approach labs with a clearly defined problem, retaining control of the product vision and decision-making.
- Define success metrics, invest in context-specific evaluation datasets validated by domain experts, and monitor performance after deployment.
- Articulate vendor selection and performance criteria aligned with organizational objectives.
- Plan for the long-term costs of scaling, including commitments from institutional payers such as public-sector partners.

For institutional funders

- Direct capital toward data systems, internal engineering staff, and measurement expertise, rather than tools alone.
- Coordinate capital through funding coalitions, shared public goods, and databases of grant-funded use cases.
- Avoid arrangements that tie grant eligibility to specific proprietary solutions.
- Broaden resource allocation beyond exemplar organizations, while helping nonprofits leverage existing products and models.
- Mandate third-party evaluations and public reporting as standard conditions of grant funding tied to large commercial providers.

For ecosystem intermediaries and advisers

- Create space for candid exchange between deployers, lab teams, and researchers.
- Develop transparent pathways to donor-supported partnerships, including with academic and open-source model providers.
- Align incentives across the ecosystem around governance frameworks, public goods, and the use of evidence.

Introduction

Over the past five years, the release of increasingly capable artificial intelligence models from OpenAI, Anthropic, Google, Alibaba, DeepSeek, and others has spurred extraordinary excitement about AI's immense potential, alongside growing concerns about its risks. Economists widely view AI as a general purpose technology — a platform for broader innovation that could affect every sector of the economy and change the way society functions.¹ Driven by such expectations, leaders across industry, government, and the nonprofit sector are seeking ways to deploy AI.

The social impact sector is no exception. As early as 2024, a McKinsey report identified approximately 600 AI use cases aligned with the UN Sustainable Development Goals — a threefold increase from 2018.² In the same year, a study by Stanford HAI and Project Evident found that 66% of nonprofits and 48% of philanthropic funders were already deploying AI in some form.³ When leveraged effectively, AI can help mission-driven organizations become more agile, data-driven, and target-focused;⁴ reach new beneficiaries; and amplify their impact.⁵

Although nonprofits could serve as important channels for delivering AI's benefits to society, constrained budgets and technical limitations often hinder adoption.⁶ In response, leading frontier AI labs have established dedicated philanthropic programs and specialized teams to support nonprofits, social enterprises, and public sector agencies with AI adoption. These initiatives provide the resources to strengthen internal operations, develop public goods (such as open datasets and domain-specific benchmarks), and build scalable solutions for social and health challenges.

Between 2017 and 2025, at least 35 major funding initiatives directed over \$410 million to AI for Social Good (AI4SG) projects.⁷ Since 2025, frontier labs have accelerated these investments, including by co-funding initiatives with major philanthropies. Prominent examples include the \$50 million OpenAI People-First AI Fund,⁸ a partnership between Anthropic and the Gates Foundation to invest \$200 million to deploy AI in global health, education, and agriculture,⁹ and \$60 million from Google.org's open funding calls in 2026.¹⁰ With the 2026 launch¹¹ of the OpenAI Foundation and an influx of tech-sector donors, billions¹² of dollars are expected to fund AI for good initiatives over the coming years.

Frontier AI labs play a central role in shaping how AI is deployed for social good. Yet, despite the growing importance of partnerships between model developers and nonprofits in determining both the downstream benefits and potential risks of AI, little empirical research explores how these collaborations actually function — and how they enable or hinder a nonprofit's ability to deploy frontier AI responsibly and effectively, and thereby to produce social impact.

Our research addresses this gap. Drawing on lessons from earlier technology-for-good movements, we conducted a qualitative study into a nascent but rapidly growing ecosystem: how frontier AI labs and nonprofit organizations partner in practice to deploy AI for good. In this report, we share findings on what motivates organizations to form these partnerships, how partners define success and thereby establish expectations about what makes for successful partnerships, and common sources of friction. Based on our findings, we offer timely, practical insights to help organizations

design, manage, and resource partnerships between frontier AI labs and nonprofits. Ultimately, these insights aim to support the responsible, sustainable deployment of frontier AI capabilities in ways that benefit communities. Accordingly, this report is tailored to guide strategic decision-making for four primary stakeholders:

- **AI labs:** To help structure and support technical partnerships that deploy AI responsibly and effectively for social good.
- **Nonprofit and mission-driven organizations:** To help organizations navigate, negotiate, and manage collaborations with frontier labs to ensure successful outcomes.
- **Institutional funders:** To provide actionable guidance for identifying enabling conditions and structuring support for success from the outset.
- **Intermediaries and advisors:** To assist in designing programs that expand the social sector's capacity to access, evaluate, and integrate frontier AI.

Our research draws on in-depth interviews conducted between April and July 2026 with 36 key personnel from 24 organizations involved in partnerships between frontier AI labs and nonprofits. Roughly a third of interviewees each came from nonprofits, mission-driven units at five frontier AI labs, and donor and advisory organizations (see Appendix A for a detailed methodology). We began by conducting a landscape assessment to map partnership models across the AI4SG ecosystem. We then identified four case studies of partnerships, conducting retrospective interviews with three to five key personnel on each. Because we sought to understand what enables nonprofits to effectively deploy AI through these partnerships, we focused on cases that demonstrate how success can be achieved. Finally, we interviewed key informants to capture a comparative set of perspectives. Our data

sources, including interviews and archival material, cover organizations operating in South Asia, sub-Saharan Africa, and North America across agriculture, health, education, climate, and humanitarian response.

1. What types of support can nonprofits access through these partnerships?

Our landscape research found that AI labs offer support to nonprofits through a range of mechanisms and models, including:

- Prize-based programs
- Incubator and accelerator programs, often in partnership with philanthropic foundations
- Engagement in cross-sector coalitions around a certain topic, such as the Global AI for Learning Alliance (GAILA)¹³
- Direct bespoke partnerships with nonprofit organizations
- Embedded staffing models or volunteer programs, such as the Google.org Fellowship Program¹⁴ or Claude Corps¹⁵

These partnerships vary significantly in their design, duration, depth, and goals. The majority of experts we interviewed noted that some of these partnerships involve deep, bespoke, and closed-door collaborations — the focus of our four case studies. These take one of two forms:

1. Bilateral partnership between an AI lab and a single nonprofit
2. Cohort-based model where an AI lab brings several nonprofits together to collaborate

These deep partnerships typically last between six and 24 months, occasionally longer, and can involve more intense periods of collaboration (such as weekly calls over several months). The AI lab supports the nonprofit in its objectives to test, adapt, and deploy its AI

technology in its products, services, or programs. The nonprofit, in turn, receives support of various forms as described below.

What are the benefits of different types of support?

In these bespoke, closed-door partnerships, almost every interviewee categorized the support they received into three broad areas: programmatic and technical support, financial support, and external relations.

Programmatic and technical support. Organizations consistently emphasized that access to technical expertise — particularly product and AI engineering staff time — was the most valuable support they received, directly targeting their expertise gaps in AI deployment, evaluation, and cost optimization. In cases where nonprofits had insufficient staffing, arrangements where labs embedded staff members into nonprofits were highly valued, especially when they focused on the nonprofit's specific needs rather than applying a generic technical solution. Notably, several interviewees stressed the importance of finding a lab counterpart who could broker access to the relevant internal teams. These individuals — referred to as “technical focal points” or “cohort managers” — helped nonprofits navigate labs' internal structures, connect with research and engineering staff, and access additional support not formally allocated to the partnership.

Financial support. Through these partnerships, nonprofits accessed financial support, including cash grants and cloud compute and API credits, provided

directly by AI labs, their philanthropic arms, or other external funders. Cloud compute credits and API credits were the most common form of funding, with packages ranging from \$100,000 to \$300,000 in credits. However, the value of compute and API credits depended on a nonprofit's existing technical capacity and platform compatibility. For example, one resource-constrained nonprofit successfully used a lab's compute credits to train a deep learning model during the partnership. Once the collaboration ended, however, it discontinued using the lab's cloud platform because the nonprofit's broader technical infrastructure was built on a different platform.

Flexible grants were less common but highly sought after, enabling nonprofits to hire staff, pay for software, and fund pilot research. Some nonprofits secured direct bilateral grants from labs, while others received funds through accelerator programs. Several interviewees noted that combining grants with technical expertise was most effective, helping nonprofits extend their planning horizons and strengthening their capacity to integrate and sustain new technologies.

External relations. In some cases, working with AI labs helped nonprofits engage with important external stakeholders. A key stakeholder was the government and, in some cases, foundations and other industry convening bodies. In one case, the AI lab partner had a policy team and strong government relationships. It contributed its own government engagement experience to help the nonprofit expand outreach to new high-level government stakeholders. This joint approach helped the nonprofit secure government backing for its proposed AI-enabled solution, which was designed for integration with the government's community-focused programming.

Overall, nonprofits benefited from different types and combinations of support based on their organizational needs. To help nonprofits identify the right support for their goals, Appendix B provides a diagnostic questionnaire.

“Organizations can have technical support and in-kind credits, but if [they] don’t have cash to pay staff, the program’s not going to scale. [A lab’s] value add is going to be around the deep technical partnership and right-sizing it for what these organizations need.”

– AI lab employee

2. How do AI labs and nonprofits form bespoke partnerships?

Across our study, bespoke partnerships between frontier AI labs and nonprofits formed through personal networks rather than public processes — relying on ecosystem intermediaries, brokering by philanthropic funders, and pre-existing employee or executive relationships. Several interviewees said these networks formed around shared ideas for building and testing AI-enabled products for social good. Exceptions included competitive accelerator, volunteer, or fellow programs, which are sometimes openly advertised. Even for open opportunities, labs often notified specific organizations early on. Consequently, partnerships were seen as accessible mainly to nonprofits already connected to philanthropic and tech-for-good networks, leaving equally capable organizations at a disadvantage.

Across all four case studies, partnerships formed through executive-level relationships and top-down agreements. Ecosystem advisers similarly emphasized that pre-existing relationships and participating in well-networked circles — such as speaking engagements at industry events or strong personal ties — helped organizations broker direct partnerships with AI labs.

Several interviewees observed that funders are well positioned to introduce commercial AI labs to nonprofits in their portfolios, using the funding relationship to help establish and de-risk initial partnerships. AI lab interviewees described relying on these recommendations to identify and screen potential nonprofit partners, deferring to funders' view of which organizations were most likely to succeed.

Finally, some partnerships originate through direct outreach from AI labs. This pathway appeared to work for larger, established nonprofits with high-profile executives.

We noted several key themes in how AI labs select partners and define partnership success (see Table 1). However, most interviewees observed that labs lacked structured processes for evaluating whether partners actually met their criteria. Nonprofit partner scoping and evaluation was seldom standardized across an AI lab, relying instead primarily on individuals' or specific teams' experiences and objectives. A few teams formalized fit assessment through “research-led diligence” — a systematic review of impact goals, strategic alignment, technical needs, and the lab's capacity to provide appropriate support prior to final executive approval. Some teams took a more ad hoc approach: one interviewee described opening conversations with a range of organizations and narrowing the pool based on contextual factors, especially technical capacity.

Table 1. AI lab interviewee descriptions of key criteria for prioritizing nonprofit partnerships showing substantial overlap between criteria across all five labs included in the study.

X = explicitly not a criteria; ● = a named criteria; ○ = not mentioned

Criterion	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5
1. Has technical capacity (e.g., at least one engineer on staff, ideally with prior exposure to the lab's AI technologies) and able to integrate and maintain AI solutions	●	●	●	●	●
2. Is a well-known, high-performing organization , with potential to become a “lighthouse” success story	●	○	●	●	●
3. Provides access to people with problem, domain, and local knowledge the lab lacks	●	●	●	●	●
4. Addresses a widespread, critical social challenge and is likely to deliver large-scale positive outcomes and mission alignment	●	●	●	●	●
5. Targets gaps left by public and private market failures where the lab's support can be uniquely additive	●	×	●	○	×
6. Offers research value for frontier AI by letting the lab learn from evaluations, new data, and process or tech improvements	●	●	●	○	●
7. Aligns with the lab's strategic priority domains and distinctive capabilities	●	●	●	●	●
8. Has the potential to be sustained and adopted at scale , especially via government	●	○	●	●	●

3. What motivates nonprofits and AI labs to form partnerships?

Across our study, clear patterns emerged in how interviewees described the motivations behind lab and nonprofit partnerships. This section examines how nonprofits describe their goals in these partnerships and the criteria labs use to prioritize partners and define success.

Key motivations for nonprofits

Building internal AI expertise, not just an AI solution.

Nonprofit motivations were largely pragmatic, focused on accessing the technical expertise and financial support needed to adopt and evaluate AI-enabled solutions to help them achieve greater impact, efficiency, and reach. More strategically oriented and tech-savvy nonprofits aimed to leverage these partnerships to develop durable expertise in deploying and evaluating frontier AI technology and to gain access to new models with which to innovate.

“The conversation that we had at the time [was] to build the capacity internally ... If costs don’t work out, do we have the know-how to then transfer over [to other models]? Our teams didn’t used to run evals at all before we started this partnership, and today we run evals for pretty much everything.”
– Nonprofit employee

A funder noted that initial experimentation with AI often exposed an organization’s need for basic capacity building in data systems and measurement. Organizations with strong data systems and a culture of measurement and experimentation were better positioned to build AI capacity through these partnerships. Those without relied on partnerships to develop core data engineering and measurement capacity, alongside AI deployment.

In one particularly compelling case, a nonprofit collaborated with an AI lab’s philanthropic arm to co-design and co-build AI applications, enabling the nonprofit to develop its own solutions. This partnership ultimately resulted in IP that the lab’s philanthropic arm transferred to the nonprofit.

Developing public goods for the ecosystem. In some cases, partnerships focused on building public goods for the wider AI4SG ecosystem. For example, one organization partnered with a lab’s engineering team to build a knowledge graph others could use through protocols such as the Model Context Protocol, which is an open standard for connecting AI models to data sources and tools. Another partnership aggregated public curricula, learning outcomes, and education guidelines from different countries into a single platform that AI models could access.

Reputational benefits. Several interviewees said that nonprofits sought these partnerships for perceived reputational benefits rather than a genuine need for AI solutions. Partnering with a frontier AI lab was seen as helpful to fundraising efforts.

Key motivations for AI labs

AI labs often approach “AI for good” partnerships with philanthropic intent that is closely tied to technological or commercial objectives. Interviewees repeatedly named three factors motivating AI labs to form partnerships.

1. Social impact

Translating technology into societal impact through widespread adoption. Many labs said they were motivated to support nonprofits in order to deliver positive social impact through wide-scale AI adoption. Several interviewees noted that labs judged partnership success accordingly — by the scale and reach of their technology. One AI lab stood out as an exception: Its nonprofit partnerships almost always produced open-source public goods, including datasets, models, and platforms.

One interviewee said that teams often sought opportunities to develop reusable — and therefore scalable — applications of their tools. Their guiding metrics were “how many people are potentially directly or indirectly impacted” and “how critical and urgent” a social problem was. There was a clear preference among labs to optimize for partnerships that would deliver scale in number of users and geographic reach — even if that meant working with a smaller group of large-scale, well-resourced organizations. One AI lab interviewee noted: “If the options are to work with 100 organizations to reach a million people or work with one organization to reach a million people, we would love to probably optimize towards the latter just for ease.” Even so, we found that some employees encouraged their teams to focus on the depth of impact and not simply the magnitude.

While social impact was a primary driver, several interviewees observed that labs often showed little interest in the simpler, fit-for-purpose applications often needed by under-resourced public and social sector organizations. Instead, labs tended to prioritize frontier research or advanced applications that these organizations could struggle to absorb.

“If you were exclusively impact-focused ... you would be doing free and or subsidized deployments and development for someone who is a government functionary ... whereas what we would care more about is can you develop an agentic workflow for a researcher for disease modeling.”
– AI lab employee

Addressing market failures. Lab employees often described prioritizing partnerships that address problems the commercial market is unlikely to solve — focusing on “public good issues or collective action problems where commercial incentives don’t crowd in,” including in “underserved environments,” demonstrating genuine philanthropic intent. Frequent examples included using AI to address challenges faced in low- and middle-income countries related to food security, agriculture, climate change, disease surveillance, humanitarian response, education, and health.

2. Research to improve their technology

Learning about technology use in real-world settings. Lab interviewees almost unanimously cited real-world testing as a primary motivation for partnering with nonprofits, using these deployments to improve their models or technology. This highlights the value labs gain from such collaborations — including user access, opportunities to build trust with government, and insights from last-mile settings — allowing them to advance technological alongside social impact goals.

“The partnerships are created because these frontier models want to leverage the nonprofit organizations to improve their models ... to work well for everybody in the world, every use case in the world... The challenge is they have limited access to data.”
– **Philanthropic funder**

One AI lab regional lead noted that its partnership with a nonprofit organization deploying voice AI in low-resource settings “gave us more evals on something that we already knew” — referring to weaknesses in voice latency, multilingual capability, and local language support. These collaborations also surfaced insights that were difficult to obtain from enterprise deployments, such as model performance in areas with “limited cellular networks” or “on a particular kind of low-end phone.”

Advancing the AI research frontier. For lab staff, the most valuable partnerships were those that enabled the lab to advance its technology through research. This included partnerships that helped labs access “evals, benchmarks, validations, new data sets” and “things that could still get you a scientific publication and have some impact along the way.” One interviewee noted that a core principle driving these partnerships was technological improvement on novel challenges: “Are we pushing the frontier forward? And then is there a bidirectional value here where not only are we able to see ... how people are using our technology in new and interesting ways, but in the process of doing that deeper collaboration, have we been able to upgrade the technology to be able to do that?”

According to another lab interviewee, because labs are fundamentally research-driven organizations, shared goals tied to research advancement were a key characteristic of the most successful nonprofit-lab partnerships: “Where I’ve seen some of the strongest partnerships has been around shared research goals... The models are so incredible and so powerful, and a lot of the benefit is not going to be from more incredible, powerful models. It’s going to be from helping people use and deploy those models.”

3. Reputational and commercial benefits

Commercial benefits. The majority of interviewees observed that commercial benefits were another key motivation to form partnerships. In some cases, partnering with high-reach nonprofits could help labs access underserved markets and bolster their reputations. One lab interviewee said that their social impact partnerships made them consider “what do we want to be known for” and align those goals with the company’s strategic advantage. Another lab interviewee noted that it was important to teams to “make some noise about it, create some buzz about it, because surely for commercial purposes that kind of marketing is useful.”

Building public goodwill through demonstrating positive use cases of AI. Many labs leverage these partnerships to identify positive use cases for AI. One lab interviewee said the objective was to build “not just trust in the company” but “trust in the technology and trust in this future.” To achieve this, several interviewees observed that labs — and occasionally funders — concentrate their most substantial partnerships on a small number of organizations that can deliver large-scale impact and demonstrate what is possible with AI in priority sectors, such as health and education. One funder described this approach: “When they do partner, they do it very well ... [The lab] invests the time, the effort, provides resources, provides compute ... because they want exemplars ... to show that it’s possible. Presumably so that others can then get into [the lab] stack ... [Y]ou don’t need a thousand examples, you just need five to 10.”

“This is a solution which the government needs to take over and run at scale ... When we want the government to formally start using our products ... it can’t be in pilot stage.”
– AI lab employee

Nearly half of interviewees noted that these partnerships could build goodwill or trust with governments and the public by demonstrating the positive impact of AI. One lab interviewee stated that partnering with governments was a core objective, and nonprofit partnerships were a critical channel for supporting government service delivery. Another emphasized that partnering with a nonprofit was aimed at adapting AI technology for public sector needs, leading to future adoption.

4. What factors contribute to effective AI lab–nonprofit partnerships?

The interviews revealed common patterns in what helps nonprofits in these partnerships successfully integrate, adapt, and deploy AI in ways that benefit the people they serve. This section shares key lessons

on both sides of the partnership — nonprofit and lab — and at the partnership level (see Table 2 for a summarized overview and Appendix C for a more detailed guide).

Table 2. Enabling factors for success

● critical factor ○ supporting factor

Enabling factor	Nonprofit-driven	AI lab-driven	Partnership level
Technical capacity and readiness	●		
Choosing the right problem and product vision	●		
Right-fit measurement and iteration	●		
Localization and participatory design	●		
Advocating for needs throughout the partnership	○		
Effective scoping of nonprofit needs		●	
Right-fit technical support, especially on evaluations, cost and optimizing model deployments		●	
Spending time with end users and in context		○	
Aligned motivations and values			●
Designing for handoff of a solution from the outset			●

Nonprofit-side enabling factors

Establishing technical readiness and capacity. The majority of interviewees across all four case studies identified in-house technical capacity (product, user research, data science, and AI engineering) and robust data infrastructure (including proprietary, context-specific data) as essential to being able to effectively design and evaluate AI tools.

In one case, a nonprofit’s product lead and AI engineer met weekly with a dedicated engineer from the AI lab, where they could ask specific questions and receive on-demand support. The interviewee noted that a “capable in-house team” of five to six engineers “who could extract that support was a necessity for a successful engagement.”

In another case, the nonprofit had spent several years developing in-house data systems and datasets relevant

to their use cases and had experimented with natural language processing technology, enabling them to “pick up from there and build [AI] on top of it.” According to a funder who supports technical assistance engagements with nonprofits deploying AI, “Having an engineering team in place is a faster win for us,” adding that otherwise, “they don’t have anyone critically inspecting the help.”

Several interviewees observed that it is fairly uncommon for nonprofits to have the ideal technical expertise. One ecosystem adviser said these partnerships tend to “benefit organizations that are already ahead of the curve tech-wise.”

Choosing a high-impact problem and developing a vision for where AI can drive value. The majority of interviewees noted that it was crucial that nonprofits identify a problem for which AI can help, establish a strong vision for their AI product, and retain control over the solution design throughout the partnership: “If you depend too much on the [lab] ... continuing on the product after that team leaves becomes difficult, and the product becomes dead.” Choosing the right problem required access to high-quality data, such as ground-truth measurements for weather forecasting or exemplary pedagogical exchanges for a chatbot tutor. It also required staff or users capable of acting on those AI outputs effectively.

Across our case studies, nonprofit teams approached partnerships with a clear understanding of beneficiary needs, a defined product vision, and an explicit plan to leverage the AI lab primarily for scientific and engineering support: “This is not a co-brainstorming session ... [Y]our nonprofit needs to be very, very clear. This is what we’re building, and this is why we’re building it.” Because the lab was primarily “there for engineering,” the product owners at the nonprofit opted to “stay on every single call ... to guide that, because things can derail.”

In two cases, the nonprofits chose to start on small but important problems and build from there. One nonprofit began its generative AI efforts with a tightly scoped use case that enabled the product team to retain full control over vision rather than starting with a more ambitious, complex use case that would involve extensive input from senior leadership. Another nonprofit identified an AI use case that was closely aligned with government priorities, envisioning a platform that could eventually be adapted to support other high-impact public sector initiatives.

Measuring key metrics early and iterating with results.

Many interviewees stressed the importance of defining a right-fit measurement approach for the success of the AI-enabled solution at the start of the partnership. Teams that agreed on success metrics before building, updated their evaluation methods throughout the collaboration, and continually refined their product based on results deployed AI more effectively.

It was equally important for partners to measure AI performance based on custom evaluation datasets developed by the nonprofit, with domain-expert and ecosystem input. In one case, a nonprofit was building automated evaluations against a rubric that tested not only whether a conversation was meaningful but “whether it evoked elements of curiosity; whether it evoked elements of a teacher.”

Nonprofits with successful deployments also built human-led, post-deployment evaluations and human-in-the-loop oversight into their workflows, adapting them over time. In one instance, a nonprofit ran “manual evaluations of at least 100 conversations every week” to continuously refine its prompts and evaluation practices.

Experts noted that nonprofits succeed when they tailor evaluation metrics to their stage of design and AI deployment. At earlier stages, it was viewed as critical

that teams prioritize model and product evaluation to understand tool performance and whether users engaged with it successfully. At later stages, user and impact evaluation were important to measure the changes in behavior, knowledge, and outcomes. In one notable example, a nonprofit conducted an impact evaluation of its AI-enabled solution before establishing product-user fit. After receiving null results, the team had to backtrack and adjust its product choices.

Focusing on beneficiary and community needs through localization and participatory design.

Most nonprofit interviewees stressed the importance of centering the needs and preferences of end users. Because large language models (LLMs) can underperform in certain low-resource languages and contexts, some nonprofits switched to locally developed models in cases with superior performance. They also found ways to ensure models performed in line with their users' needs by refining prompts and integrating local examples.

Developing custom, context-rich datasets was another key focus, even when data collection proved resource-intensive.

“LLMs are not built for [our] user base ... we’ve had to build our own speech audio datasets just to be able to get reading assessment right... we’re a long way away from having AI be useful for everyone, let’s say, at the base of the pyramid.”

– Nonprofit employee

Maintaining clarity on needs and advocating for them throughout the partnership.

Several interviewees noted that nonprofits succeed when they articulate and advocate for the necessary support, particularly given the inherent power imbalances in these partnerships.

Several nonprofits with highly successful AI lab collaborations were able to clearly advocate for their needs from the outset, securing critical support including funding, API credits, and engineering assistance. In one case, a nonprofit refused partnership offers from AI labs that did not offer access to engineering staff: “The non-negotiable by then was engineering investment from the other side.” Another nonprofit forged partnerships with multiple AI labs, through which they accessed different types of support. To manage these partnerships, the nonprofit interviewee stressed the importance of clearly communicating its organizational priorities to each lab and relying on internal benchmarks to assess model performance.

Lab-side enabling factors

Effective scoping of nonprofit needs. Interviewees reported inconsistent approaches in how AI labs scope nonprofit technical needs. Successful partnerships prioritized in-depth scoping to uncover nonprofit needs before allocating resources — either at the start or throughout the partnership. Notably, interviewees emphasized engineer skillsets and responsiveness to nonprofits' ad-hoc requests over the number of engineering hours provided.

Scoping took many forms across different labs. One spent months in discussions to understand a nonprofit's challenges and to confirm that AI was the best solution. An interviewee from another lab observed that while scoping had been inconsistent

across nonprofits, and in some cases superficial, years of experience showed that “research-led due diligence” was essential: “what do they actually need versus what is the thing that we are best set up to do.” A third lab set up ongoing communication channels with the nonprofit to address questions as they arose.

Focusing technical support on evaluations, cost, and optimizing model deployments. Interviewees highlighted three technical areas that were crucial to nonprofits’ ability to deploy AI effectively: evaluations, optimizing model architecture for performance and efficiency, and reducing operating costs at scale.

“[The labs’] major contribution was to highlight the importance of evals ... [A] lot of orgs weren’t getting improvement in initial products, because they didn’t know how to do evals.”

– Philanthropic funder

In rare cases, labs intentionally designed technical support to help nonprofits build capabilities. For example, one engineer set up initial evaluation pipelines and handed the tools to the nonprofit partner, offering guidance while empowering them to iterate independently: “Once that competence is built locally... the knowledge proliferates within that ecosystem in a far more natural and organic way instead of just planting somebody, having them do this, and leave.”

Another AI lab interviewee noted that the majority of nonprofits needed support on cost optimization, such

as prompt caching and identifying the smallest models capable of completing the task at hand.

Spending time with end users in their local context. Several interviewees from both nonprofits and AI labs emphasized that lab staff were far more effective when they visited nonprofit staff and their communities. Bridging knowledge and cultural gaps required intentional investment — including, as one AI lab interviewee noted, hiring staff based in the same countries as their mission-driven deployments or recruiting employees with prior local experience.

Partnership design and management success factors

Aligning motivations and values. Several interviewees observed that partnerships were more likely to succeed in cases where both sides were genuinely aligned on shared definitions of success regarding social impact. In one case, a nonprofit and AI lab aligned early on their vision to partner with the government, collaborating on the design and pitch they brought to public officials, and using this vision to guide resource allocation and project milestones. In another, a lab staff member helped the nonprofit investigate if AI was the right solution for beneficiary needs, asking: “Is AI really needed in this particular problem?... [L]et’s not just build AI for the sake of AI.”

For some AI labs, nonprofits’ partnership experiences depended more on whether they identified lab staff with shared values. A lab interviewee noted that staff approaches varied widely even within the same organization: while some staff members might engage in more superficial, PR-driven collaborations, others were motivated to engage in more substantive, impact-driven work.

Designing for long-term solution handoff from the outset. In rare instances, labs and nonprofits planned for sustainability after their partnership by designing handoff strategies for AI solutions and capabilities. The most common strategies included clear documentation describing the solution and how to maintain it, sustainable post-subsidy pricing models, and capacity-based milestones that phased out external technical assistance as nonprofit teams developed the necessary skills. One interviewee from an AI lab's philanthropic arm, drawing on years of oversight experience, designed a new partnership model to ensure nonprofits maintain long-term deployment capacity: "There is a tapered support. That means we will provide capital, but [the lab] will be on the hook to provide fractional product management or engineering capacity. That will go down in a one-to-three-year time horizon as the capacity of the receiving organization goes up."

Across the ecosystem, however, partnerships rarely plan for a sustainable handover. When partnerships lack a fixed end date, they seldom establish clear criteria for when and how to conclude their collaboration. In one instance, a nonprofit partnered with an AI lab over five years to develop an open-source model. Once the partnership ended and the project was made publicly available on GitHub, the nonprofit team found that only two staff members — and very few data scientists and engineers external to the collaboration — could effectively implement the software. The nonprofit subsequently hired a consulting firm to turn the model into a usable product, and it now defines at the outset whether an AI lab partnership aims to meet research goals or build products that the nonprofit or a broader technical community will maintain.

Some interviewees noted that accelerator and/or employee volunteer programs were particularly short-

lived and lacked sufficient focus on capacity-building for the nonprofit. A funder shared an example: "[A lab] brought in ... six engineers to one of our partners ... They did all of the evals. When they left, there was no culture of AI evals in the organization."

5. What causes friction in these partnerships?

Difficulty financing AI at scale. Most nonprofits interviewed aimed to scale their AI solutions, but acknowledged the challenge and uncertainty around how to transition away from subsidized lab support. One nonprofit planned for this outcome by shaping their work as a government partnership.

“We want at least the compute costs to be absorbed by the government.”
– **Nonprofit employee**

An ecosystem adviser viewed even bespoke partnerships between nonprofits and AI labs as time limited, raising the possibility of “a two-tier system where [you] use the cheaper models for sub-Saharan Africa, the good models for those who can afford it.” Another ecosystem adviser noted: “Despite all the conversation about the falling costs of AI, for many of these organizations, their AI solutions are not financially viable at scale without concessionary credits.”

Lack of clear partnership terms and governance structures. Formal agreements covering the terms of the partnership varied widely, with some lacking formal contracts entirely. Even in extended collaborations lasting six months to one year, several organizations operated with minimal formal contractual agreements in place. When formal agreements did exist, they typically took the form of NDAs, letters of intent, or memorandums of understanding, and occasionally more extensive partnership and data-sharing

agreements. The variation in governance structures suggests that many partnerships were trust-based commitments. Without written agreements or documented plans, several interviewees cautioned that nonprofits have little recourse when actual support falls short of verbal promises and when deciding how and when to end the collaborations. An AI lab interviewee also noted that ambiguous partnership terms can amplify negative power dynamics, prompting nonprofits to overprepare for meetings in ways the labs view as unnecessary.

When asked if partnerships incorporated AI governance mechanisms — such as mandatory third-party evaluations, civil society involvement or participatory design processes, or transparency requirements — interviewees reported that such practices were largely absent. Where governance practices were in place, several interviewees from nonprofits said they took the lead — for example, by ensuring participants were aware of AI use and minimizing data shared with the AI model. Ultimately, the responsibilities and risks of responsible AI deployment fell largely on the nonprofits.

Influence of technosolutionist perspectives. Some interviewees observed a tendency of upstream actors — including funders, AI labs, and government partners — to promote technosolutionist approaches to social impact, driving nonprofits to pursue advanced AI tools even when they did not suit their organizational needs. This risk was most prevalent in partnerships that bypassed user research and contextual design, especially since labs provided engineering expertise

rather than design and user experience support. One philanthropic funder observed: “What you run the risk of ... [is] just trying to shove a technical solution down people’s throats ... as opposed to ... [taking] the time to understand what problem they’re facing, how they’re currently addressing it, and then building from that fact base.”

Several interviewees also described partnerships in which a nonprofit received funding tied to a specific AI solution. In some cases, nonprofits later discovered that the solution was not needed for the problem they had, but they lacked the flexibility to switch to another technological solution. One nonprofit interviewee said the process was like “using a tank to kill the chicken.”

Differing commercial and reputational incentives.

Several interviewees mentioned that partnerships suffered when commercial objectives conflicted with social impact objectives. One nonprofit employee attributed this tension to incentives some lab employees had, describing a counterpart whose “role is to basically get organizations like us to use [the lab’s] products”: “I’m not questioning her sincere intentions in terms of having a positive impact, but what she needs to happen is that we like the [the lab’s] products and we use them.”

Interviewees from AI labs noted that public relations and marketing teams pushed for “big numbers” when describing these partnerships, with team engagement waning once external recognition for their work faded. In some cases, these reputational incentives pressured nonprofits to exaggerate partnership results. One nonprofit interviewee described feeling “forced to find numbers” to showcase its research collaboration with an AI lab on a Demo Day, even though they ultimately chose not to adopt the AI technology due to poor performance.

Some nonprofits sought out partnerships for their reputational or financial benefits. One AI lab interviewee mentioned participants in a cohort program who “are there more for the recognition of being part of those programs, but really not leveraging all the resources that we make available for them.”

Power asymmetries. Several interviewees said that the power asymmetry between AI labs and nonprofits resulted in labs often setting the partnership terms, particularly around access to scarce resources such as compute.

“[The labs] are the one doing the reaching out [with] the terms of engagement ... You take it or you leave it. We did a study trying to understand the various modalities of people using compute in Africa ...70 to 80% use compute that is donated by the big providers ... So you really cannot afford to say no.”

– Philanthropic funder

However, established nonprofits or those with experience developing tailored partnerships with AI labs typically had greater success ensuring their specific needs were addressed in partnership agreements.

Divergent organizational timelines and ways of working. At the partnership level, the most consistent challenge related to the differences in the time

horizons needed by AI labs and nonprofits to achieve their shared objectives, and clashes in how each side worked. Labs' focus on fast-paced work, building, and iterating, and their general commercial orientation sometimes conflicted with nonprofits' emphasis on deliberate process, the need to work through coalitions, or the slower timelines associated with public sector work. One nonprofit partner recalled an AI lab describing its approach as "building the plane while flying it," adding that the lab was "not clear enough about its vision ... to establish how much time this is going to ask of [the nonprofits], what they're going to get out of it."

Another nonprofit interviewee described a partnership where a lab and funder set unrealistic timelines, expecting the organization to conduct research, consult with country offices and government, operationalize an AI product, and measure impact within one year: "All this is supposed to happen in a year. To me, that doesn't make sense."

Tension between mission-driven individuals and institutional priorities. Several interviewees observed that when values were clearly aligned between nonprofit and AI lab partners, it was due to particular staff members rather than the lab as a whole: "I'm yet to see if [the lab] as a whole, what their perspective will be." Indeed, some lab staff reported a clash between their institution's interests and their individual goodwill. Interviews revealed several mission-driven champions within AI labs who actively worked to overcome institutional barriers to deliver better support to nonprofits — such as designing new forms of engagement and aligning and rallying colleagues. One lab employee said that his vertical within the company had "mission-oriented people," but because management rewards "an interesting blog post or speaking engagement," it is a "management culture

issue, not an organizational culture issue." Another lab employee said: "That full systems view is difficult to have. I'm hoping that people like me within these organizations will be able to bring it, but sometimes it's a voice into the abyss."

6. What is the role of institutional funders?

Alongside AI labs, philanthropic foundations are the most important upstream actors in the AI4SG ecosystem. Our study showed that funders play a critical role beyond allocating resources.

Where funders are adding value

Convening and brokering partnerships. Several interviewees praised the convening role that funders play, using their relationships to connect nonprofits with AI labs and with one another. As one philanthropic funder put it, labs come to them to identify and bring together promising organizations because “we are a strong convener, and ... [labs] are not,” though it is “not a ... door that we control.” An AI lab interviewee highlighted funder-organized convenings as vital mechanisms for ecosystem-wide learning and coordination among nonprofits.

Pooling and aligning funding across foundations. Several funders recognized that resource-intensive AI development demands coordinated investment. One described forming “a coalition of funders” around thematic areas, such as health and agriculture. Rather than issuing duplicative grants, these coalitions established a shared “center of excellence” to pool capital, and expand access to data and high-performing models and products.

Another funder noted the importance of collaboration at the proposal design stage, particularly given that investments related to AI applications for social impact could often be shared among organizations: datasets, tooling, and standards. Nevertheless, this interviewee noted that funders seldom collaborate in ways that reduce duplication across the ecosystem, suggesting

that even if cross-organizational coordination costs \$500,000 per year, the resulting efficiency would generate significant social returns on investment.

Providing hands-on technical or convening support.

In one particularly effective philanthropic model, a funder convened an accelerator program with multiple nonprofit organizations and an AI lab. Relying on light-touch advisory support from the AI lab, the funder provided in-depth support to nonprofits in the accelerator to strengthen data systems, design measurement approaches, implement data analytics to support experimentation and learning, and design and implement AI evaluation pipelines. Because the lab could not engage directly with every nonprofit, this funder stepped in to deliver the bespoke support each of its grantees needed to start innovating with AI. In addition, the funder incentivized and supported its grantees to build capabilities to conduct A/B testing, to develop the capacity to iterate on AI products.

In another case, after learning that a lab’s research team was building advanced models to predict food insecurity, the lab’s philanthropic arm convened a partnership coalition with nonprofits to guide development. This arrangement allowed the lab to co-design and evaluate the model with partners, ensuring they built the technical capacity to deploy the model independently. After a portfolio review, the lab found that “collaboration was the common denominator” across every AI4SG initiative that reached meaningful scale.

Reducing duplication through shared resources. To prevent anyone from having to reinvent the wheel, two funders are building shared, openly available resources. One is assembling a non-proprietary funder playbook

alongside an open use-case repository so funders can see which organizations are working on a problem and avoid redundant grants.

Other funders are financing public goods — funding open-source training datasets and building the benchmarks and evaluation frameworks that improve model performance.

Where funders are falling short

Concentrating capital in high-profile big bets.

Funders were sometimes criticized for clustering their AI4SG investments around a small group of prominent nonprofits. As one funder observed, these “darling organizations” become “flushed with resources,” while less visible — and potentially more impactful — initiatives remain unfunded. One ecosystem adviser suggested an alternative approach that focuses on supporting the wider ecosystem.

“I really welcome these efforts to channel that work, not maybe directly to one or two lighthouse organizations, but ... to more of an ecosystem-based approach ... more community-led and based on good research as to what people actually need ... [It] can’t really be designed in Seattle or ... California ... or an enclave in Bellagio.”

– Ecosystem adviser

Funding ill-equipped nonprofits and fueling

duplication. According to one funder, the recent boom in AI-related philanthropic funding has driven nonprofits to build their own new AI tools, even when adopting existing tools would serve them better. In another case, a foundation interviewee believed that funders too often supported “boots on the ground” nonprofits to experiment with AI without evaluating whether they had the technical talent to understand and manage the technology.

Funding “shiny” products over sustainable

AI adoption. Several funders acknowledged that foundations often gravitated toward flashy, technologically advanced products — with less attention paid to the investments that help organizations build fit-for-purpose, sustainable, scalable solutions.

“We want that shiny thing to be able to point to and say, our money led to this shiny tool ... when sometimes what actually needs to be done is ... unsexy things that are most important to actually get beyond that valley of death to scale.”

– Philanthropic funder

Another funder noted a pattern where foundations are drawn to co-funding initiatives that amass large amounts of capital but often dilute project specificity and reduce the investment’s overall effectiveness.

Instead, interviewees stressed that rather than funding AI-enabled innovations, AI4SG funding must be grounded in a more nuanced understanding of broader ecosystem readiness and structural gaps — such as organizations’ data infrastructure and AI governance capabilities. A deeper problem, one funder observed, is a lack of technical and AI expertise among most philanthropic institutions: “95% of ... program officers ... are not equipped to evaluate the basics of a technology strategy or proposal ... don’t even know the difference between a product manager and a data engineer.”

Tying funding to specific AI lab partnerships. A few interviewees from AI labs and foundations observed that foundations actively pursue partnerships with the largest commercial AI labs, driven in part by the prestige of aligning with major industry players. Fewer pursue partnerships with open-source providers and alternative AI solution providers.

Several nonprofit interviewees pointed out that when funders convene partnerships with frontier AI labs and tie grants to these partnerships, they influence which AI vendors nonprofits choose.

“There is an implicit, hey, we’ll fund these projects because [a particular vendor] will get used ... [T]here’s less funding for those kinds of [agnostic to provider] initiatives.”
– **Ecosystem adviser**

7. Recommendations

In our research, we sought to map the landscape of AI4SG initiatives and to interview key personnel across successful partnerships. Based on these insights, we offer provisional prescriptions for how AI labs, institutional funders, and intermediary organizations can more effectively design partnerships with nonprofit organizations to help them access, integrate, and evaluate AI for social good. Below, we present key recommendations for each stakeholder group.

For AI labs and other organizations providing AI-related support (“support providers”)

- **Conduct in-depth scoping to understand organizational needs and capacity:** Invest in an in-depth scoping phase at the start of a partnership to understand what mix of support to provide, and then continually assess and match support to nonprofit needs as the partnership evolves. In particular, support providers should rigorously evaluate organizational readiness for data and AI capacity and consider forward-deployed engineering support.
- **Prioritize nonprofit capacity building over short-term deployments:** Ensure engineers help nonprofits develop durable tooling, data pipelines, and evaluation capabilities related to AI adoption and ongoing maintenance.
- **Establish plans to sustain solutions after partnerships conclude:** Outline clear plans from the start of the partnership to taper off financial and technical support as nonprofits build the capacity to independently sustain solutions over time. Labs should collaborate with deployers and

philanthropic funders to identify long-term funding mechanisms — such as government absorption of compute costs or phased licensing models — well before subsidized support concludes.

- **Standardize partnership governance and terms:** Create explicit, standardized partnership terms — including expected contributions, effort levels, and governance requirements such as data privacy and minimization and public knowledge sharing. For open-ended partnerships, set norms and deliverables for the partnership and how it will conclude. Clear partnership terms and governance practices reduce power asymmetries and avoid operational friction.

For nonprofit and mission-driven organizations

- **Ensure in-house technical capacity:** Build foundational data engineering and product management capabilities before entering bespoke partnerships with AI labs. Internal technical capacity is the strongest determinant of an organization’s ability to direct lab support, evaluate AI systems, and absorb technical assistance.
- **Retain ownership of problem definition, product vision, and raw datasets:** Approach labs with a clearly defined, user-grounded problem that AI is appropriate for solving. Nonprofits should retain control over product vision and decisions related to user experience, as well as raw datasets obtained through their operations, and leverage lab expertise primarily for specialized engineering and scientific challenges.

- **Establish right-fit evaluation practices:** Define clear success metrics from project inception. Deployers should invest in context-specific, domain-expert-validated evaluation datasets, monitor products post-deployment, and transition to rigorous experimental evaluations (such as randomized controlled trials) once products are mature.
- **Define criteria to guide vendor selection:** Articulate clear performance metrics aligned with organizational priorities, such as cost or quality, and evaluate model providers against these criteria.
- **Plan for the long-term costs of scaling:** At the start of a partnership, plan for how your organization will transition from funder or lab subsidies. For example, secure commitments from institutional payers, such as public-sector partners, to fund scaled AI solutions after funder or lab subsidies end.

For institutional funders

- **Fund nonprofits to develop AI-related capacity and infrastructure:** Direct capital toward helping nonprofits build foundational AI infrastructure and the capabilities to sustain AI solutions — including strong data systems, internal engineering staff, measurement and evaluation expertise, and transitioning from subsidized support from AI labs.
- **Coordinate capital to prevent duplication:** Establish funding coalitions, invest in public goods within key sectors, such as health and education, and maintain up-to-date resources on grant-funded use cases and products to harmonize grantmaking strategies and prevent redundant development.

- **Ensure funding is vendor-agnostic:** Avoid funding arrangements that tie grant eligibility to specific proprietary lab ecosystems or vendors.
- **Broaden resource allocation beyond “lighthouse” organizations:** Invest in incubating and accelerating a wider ecosystem of AI-ready nonprofits, while helping organizations leverage existing products and models.
- **Mandate governance mechanisms in philanthropically-funded AI partnerships:** Require or incentivize organizations to conduct third-party performance evaluations and openly report on AI deployment outcomes as standard conditions of grant funding tied to partnerships with large commercial providers.

For ecosystem intermediaries and advisers

- **Support cross-organizational learning:** Create dedicated spaces where deployers, lab teams, and researchers can candidly exchange learnings.
- **Make philanthropically-funded AI partnerships and technical assistance more accessible:** Develop transparent pathways for nonprofits to enter into funder-supported partnerships with AI labs and other types of providers, such as academic institutions and open-source model communities.
- **Coordinate and drive change across the ecosystem:** Align goals and incentives among stakeholders around core objectives: advancing governance frameworks, funding public goods, and promoting evidence-based practices.

Appendix A - Methodology and Limitations

Selection Criteria

Because our aim was to understand how nonprofits can effectively deploy AI to advance their missions, and generate real-world impact through collaboration, we focused on case studies of successful partnerships. These examples offered the most instructive insights into achieving meaningful outcomes. We selected cases defined by the following key features:

- Bespoke partnerships, where AI labs and partnering nonprofits formed close collaborations to develop customized approaches, products, or processes
- A completed period of regular collaboration (e.g., weekly or biweekly engagement) between the AI lab and partner organization(s) over at least six months, within the past two years
- Efforts by an AI lab to provide technical expertise as part of the partnership
- The nonprofit's intent to integrate the lab's frontier AI technology into a tool, product, or service to advance its social mission

Each case involved a frontier AI lab and a nonprofit serving last-mile communities in low- and middle-income countries, supported by philanthropic funding, pro-bono assistance from an AI lab, or both. Structurally, two cases were bilateral partnerships between an AI lab and a single nonprofit, while the other two were part of a broader accelerator program.

Interview Approach

In total, our research consisted of 5 preparatory interviews and 31 formal research interviews with 36 individuals across 24 unique organizations:

- 13 individuals from 9 nonprofit organizations deploying frontier AI technology across five sectors — agriculture, health, education, climate, and humanitarian response — with operations primarily concentrated in South Asia, sub-Saharan Africa, and North America.
- 11 individuals working in “mission-oriented” units across 5 frontier AI labs
- 5 individuals from 4 philanthropies funding AI4SG partnerships (including philanthropic arms of AI labs)
- 7 experts from 6 organizations providing advisory and evaluation support to AI4SG partnerships

For each case study, we conducted semi-structured retrospective interviews with key personnel across partner organizations, averaging interviews with 3 to 5 key personnel per case. To capture broader sector-level insights, we also interviewed ecosystem-level experts and representatives from comparative cases. All interview data was supplemented with internal metrics reports, project documentation, and publicly available online data. Interviews were conducted between April and July 2026 — with the majority completed by May — and lasted 30 to 60 minutes each.

Overview of interview sources	
Total engagement	5 preparatory / 31 formal interviews
Nonprofits / deployers	13 individuals / 9 organizations
Frontier AI labs	11 individuals / 5 organizations
Philanthropies / funders	5 individuals / 4 organizations
Technical advisers / intermediaries	7 individuals / 6 organizations

Limitations

Our research relies primarily on self-reported data from interview participants. We intentionally selected case studies of successful partnerships because they were highly instructive, though these cases also yielded insights into sources of friction and potential risk factors. In addition, we interviewed individuals across other partnerships to capture broader ecosystem dynamics. While our sample is diverse and extensive, it is not fully representative of all such initiatives; findings should be interpreted accordingly. To protect proprietary information, all data was anonymized and aggregated, with specific company names and identifying details removed or generalized.

Appendix B - A Diagnostic for Nonprofits on What Support to Seek

Nonprofits can use the following diagnostic checklist and questionnaire to determine what types of support they should seek based on internal gaps.

Checklist for Nonprofits

Programmatic and technical support	
	Engineering staff time: hours dedicated by lab and nonprofit engineers to the partnership
	Product manager staff time: nonprofit product manager assigned to the partnership, or external product management support
	Embedded staff: external technical staff embedded in the nonprofit's team
	Brokered internal access: lab staff member connecting the nonprofit to relevant internal expertise, including research staff not formally assigned to the partnership
	AI strategy support: periodic advisory sessions offering guidance on where the technology can realistically be deployed
Financial support	
	Compute credits: free access to compute to train and run models
	API credits: free access to a lab's AI model API enabling free model usage
	Cash grants: direct cash funding to the nonprofit or partnership
External relations	
	Government engagement: direct lab or funder engagement with government counterparts acting on the partnership's behalf

Questionnaire for Nonprofits

Area	Diagnostic question	Support to seek
1. Skills	Do we have the product, AI engineering, and measurement expertise in house to build the AI solution — including the software to integrate and localize the model and to run evaluations — and maintain it after support ends?	If yes: lab product and/or engineering staff advisory calls to address expertise gaps. If no: cash grants to hire technical staff or consultants, plus lab product or engineering staff and/or advisory calls, to address expertise gaps.
2. Time	Can we allocate the in-house technical and program staff to this partnership at the capacity required to build and test the AI solution?	If yes: proceed, and explore lab product and/or engineering staff advisory calls to address expertise gaps. If no: embedded staff members from the lab, plus cash grants to help the nonprofit allocate resources to support the work of embedded staff members.
3. Infrastructure	Do the compute and/or API credits offered align with our technical architecture?	If yes: proceed. If no: cash grants in addition to API and/or compute credits to support switching and adaptation costs.
4. Data	Do we have the data, pipelines, and related governance mechanisms to implement the AI solution — and the right to use it for this purpose?	If yes: proceed. If no: engineering staff advisory calls to build data foundations before model work begins, plus cash grants for data collection, labeling, or licensing.
5. Ability to navigate lab support	Do we know how to navigate the partnership and reach the teams we need within the lab?	If yes: proceed. If no: a lab-side relationship owner who can broker internal access.
6. Sustainability	Do we have a clear handover plan and the resources to pay for the AI product after the partnership ends?	If yes: proceed. If no: documented technical handover designed to meet capacity-building objectives, government engagement support, and an identified source of funding external to the lab.

Appendix C - Critical Factors Enabling Success

Enabling factor	Critical or supporting	Driven by	How this appears in practice
Nonprofit-side factors			
Technical capacity and readiness	Critical	Nonprofit	• In-house product, user research, data science, and AI engineering capacity • Robust data infrastructure, including proprietary, context-specific data • A team able to extract on-demand support as needs arise
Choosing the right problem and a product vision the nonprofit owns	Critical	Nonprofit	• A problem identified before the partnership begins • A product vision, and control over solution design, maintained throughout • An explicit plan for what the AI lab is there to provide • A tightly scoped first use case • Enough data to represent what good looks like, such as ground-truth measurements or documented examples of high quality practice • Staff or users able to act on the AI outputs produced
Right-fit measurement and iteration	Critical	Nonprofit	• Measurement agreed before building and updated throughout • Custom evaluation datasets developed with domain-expert and ecosystem input • Human-led review after deployment, watching for over-reliance • Intermediate outcomes, such as adoption and repeat engagement, measured including through A/B tests • Metrics matched to deployment stage, with impact evaluation reserved until there is evidence of product-user fit
Localization and participatory design	Critical	Nonprofit	• Outputs adapted and localized through prompting and local examples • Smaller, locally developed models substituted for frontier models that underperform • Custom datasets built to reflect users and languages, even when resource intensive
Advocating for needs throughout the partnership	Supporting	Nonprofit	• Support needs identified from the outset, including financial support, API credits, and engineering time • Offers declined when access to engineering talent was absent • Priorities, objectives and key results, and non-negotiables stated to all partners • A model-agnostic stack
AI lab-side factors			
Effective scoping of nonprofit needs	Critical	AI lab	• A scoping phase before support is allocated • Research-led diligence • Responsiveness to ad hoc needs, rather than a fixed quantity of engineering hours • A dedicated project manager and a regular communication channel, with engineers brought in as needed

Appendix C - Critical Factors Enabling Success (cont'd)

Right-fit technical support, especially on evaluations, cost, and optimizing model deployments	Critical	AI lab	- Evaluations - Model architectures optimized for performance and efficiency - The cost of running a solution at scale reduced through prompt caching and smaller models - Tooling handed over with guidance to ensure that capability remains with the nonprofit
Spending time with end users and in context	Supporting	AI lab	- AI lab staff visiting communities alongside nonprofit staff - Staff hired in the countries of deployment, or those with regional experience - Capabilities adapted to the nonprofit's existing operations
Partnership design and management			
Aligned motivations and values	Critical	Both parties	- Values alignment between the teams staffed to the work - A counterpart willing to ask whether AI is actually needed for the problem - A shared objective beyond the build itself, such as a joint approach to government engagement
Designing for handoff of a solution from the outset	Critical	Both parties	- Documentation describing the solution and guidance for maintaining it - Sustainable pricing models for when subsidized or free credits end - Capacity-driven milestones in the agreement, such as staffing support that winds down on a defined schedule - Tapered support, where lab-provided product management or engineering capacity reduces over one to three years as the nonprofit's capacity rises - Clear definition at the outset of whether the partnership aims to achieve research goals or to build a product the nonprofit can maintain

Endnotes

1. Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock, “GPTs Are GPTs: Labor Market Impact Potential of LLMs,” *Science* 384, no. 6702 (2024): 1306–8, <https://doi.org/10.1126/science.adj0998>.
2. Medha Bankhwal et al., “AI for Social Good: Improving Lives and Protecting the Planet,” McKinsey & Company, May 10, 2024, <https://www.mckinsey.com/capabilities/quantumblack/our-insights/ai-for-social-good>.
3. Sarah Di Troia et al., “Inspiring Action: Identifying the Social Sector AI Opportunity Gap,” Stanford Institute for Human-Centered Artificial Intelligence, February 8, 2024, <https://hai.stanford.edu/policy/inspiring-action-identifying-the-social-sector-ai-opportunity-gap>.
4. Soraya Sedkaoui and Rafika Benaichouba, “Generative AI as a Transformative Force for Innovation: A Review of Opportunities, Applications and Challenges,” *European Journal of Innovation Management* 29, no. 3 (2026): 677–701, <https://doi.org/10.1108/EJIM-02-2024-0129>.
5. Marie-E. Godefroid, Ralf Plattfaut, and Björn Niehaves, “Identifying Key Barriers to Nonprofit Organizations’ Adoption of Technology Innovations,” *Nonprofit Management and Leadership* 35, no. 1 (2024): 237–59, <https://doi.org/10.1002/nml.21609>; Angela Aristidou, Andrew Duncelman, and Sam Fankuchen, “How AI Can Deepen Nonprofit Relationships,” *Stanford Social Innovation Review* 23, no. 4 (Fall 2025): 73–74, <https://doi.org/10.48558/yq8c-9h56>.
6. Matthias Fabian Gregersen Trischler and Jason Li-Ying, “Exploring the Relationship between Multi-Dimensional Digital Readiness and Digital Transformation Outcomes,” *International Journal of Innovation Management* 26, no. 3 (2022): 2240014, <https://doi.org/10.1142/S136391962240014X>; Francisco do Adro et al., “Social Entrepreneurship Orientation and Performance in Non-Profit Organizations,” *International Entrepreneurship and Management Journal* 17, no. 4 (2021): 1591–1618, <https://doi.org/10.1007/s11365-021-00748-4>.
7. Hongjin Lin et al., “Funding AI for Good: A Call for Meaningful Engagement,” *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (April 2026), 1–24, <https://doi.org/10.1145/3772318.3790374>.
8. OpenAI, “A People-First AI Fund: \$50M to Support Nonprofits,” September 8, 2025, <https://openai.com/index/people-first-ai-fund/>.
9. Anthropic, “Anthropic Forms \$200 Million Partnership with the Gates Foundation,” May 14, 2026, <https://www.anthropic.com/news/gates-foundation-partnership>.
10. Google.org, “Google.org Impact Challenge: AI for Government Innovation,” accessed September 17, 2026, <https://www.google.org/impact-challenges/ai-government-innovation/>; Google.org, “Google.org Impact Challenge: AI for Science,” accessed September 17, 2026, <https://www.google.org/impact-challenges/ai-science/>.
11. OpenAI, “Update on the OpenAI Foundation,” March 24, 2026, <https://openai.com/index/update-on-the-openai-foundation/>.
12. Joel Khalili, “These AI Barons Are Ready to Give Away Their Fortunes,” *Wired*, August 9, 2026, <https://www.wired.com/story/ai-billionaires-are-pledging-their-wealth-good-or-bad/>.
13. Dalberg, “Global AI for Learning Alliance (GAILA): Coordinating Collective Action to Advance Learning Outcomes,” April 2026, <https://dalberg.com/our-ideas/global-ai-for-learning-alliance-gaila-coordinating-collective-action-to-advance-learning-outcomes/>.
14. Google.org, “Google.org’s Global Programs,” accessed September 17, 2026, <https://google.org/impact-report/global-programs>.
15. Anthropic, “Claude Corps,” accessed September 17, 2026, <https://www.anthropic.com/claude-corps>.