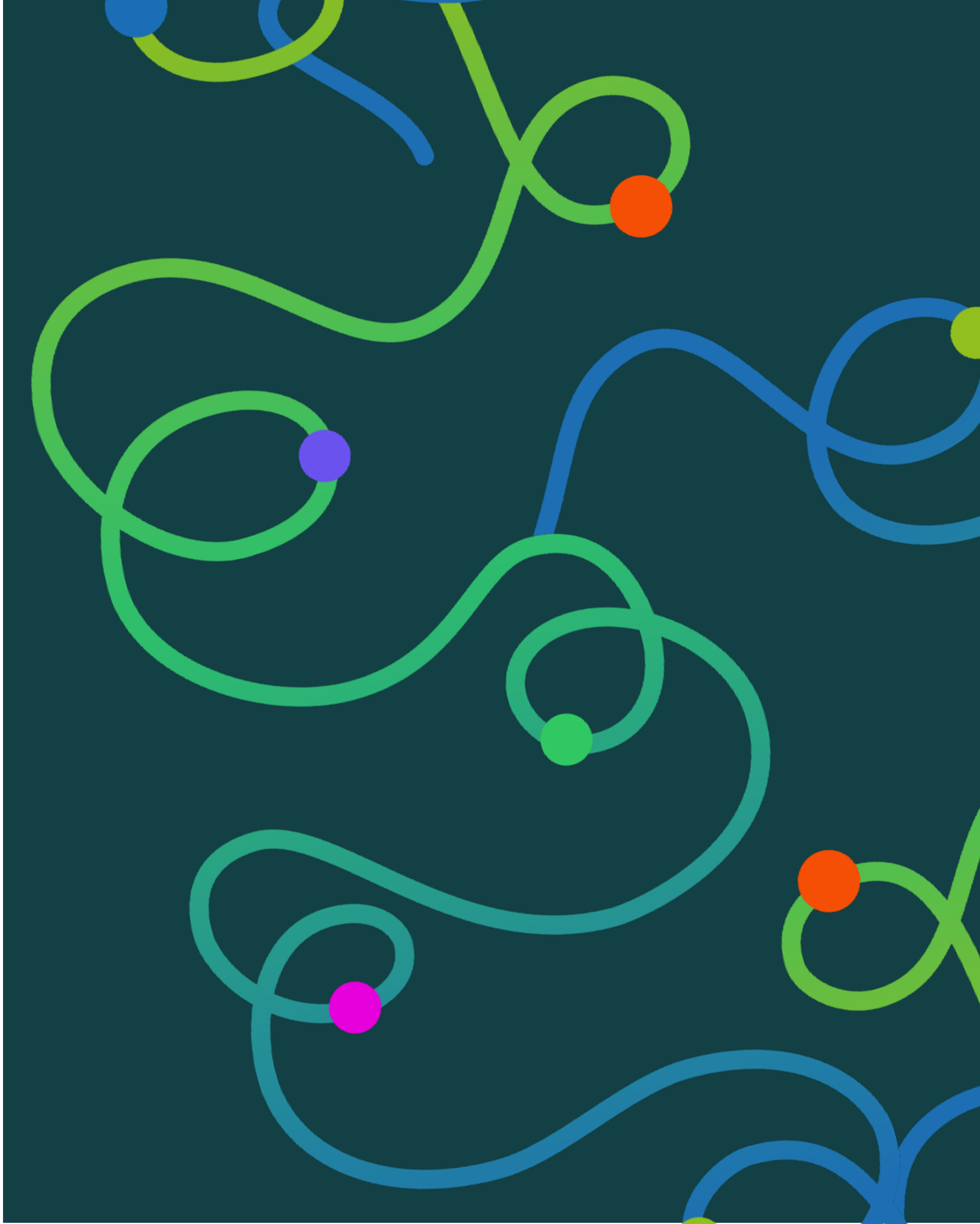
AI & EQUALITY

Annual Report
2024–2025



Letter from the CEO

In 2025, the gap between AI ethics rhetoric and ground-level accountability stopped being a technical problem and became a democratic one. Governments were deploying AI systems without procurement standards. Developers were optimizing for speed without benchmarks for human rights. Regulators were drafting frameworks without the tools to enforce them.

We responded by building infrastructure, not publishing more position papers. The 6-Part Framework, validated by SEFI. Regional toolboxes co-created with the African Center for Technology Studies ACTS, and Chile's National Centre for Artificial Intelligence CENIA, not exported from Geneva. The first Human Rights Benchmark for LLMs. Procurement frameworks that shift market incentives before systems are deployed. A global community of 750+ practitioners who don't wait for permission to reimagine what's possible.

This is the civic infrastructure layer of Digital Public Infrastructure. The accountability mechanisms that make AI sustainable rather than volatile. The missing foundation governments are building DPI without.
The future of AI isn't predetermined. We have the agency to shape it. This report shows what that looks like in practice.

Caitlin Kraft-Buchman
CEO & Founder, AI & Equality by Women at the Table
Geneva, Switzerland

Introduction:
Rules Into Tools

The year 2025 marked a turning point: **the gap between AI ethics rhetoric and ground-level implementation shifted from technical oversight to democratic threat.** The <AI & Equality> Initiative by Women at the Table exists to close this divide, transforming universal “Rules” into working “Tools” that are legally grounded, technically rigorous, and feminist by design.

We’ve built an ecosystem that pairs a global community with technical innovations and human rights-based design frameworks. Through regional co-creation across Africa, Europe, and Latin America, we provide the infrastructure needed for genuine accountability—ensuring the future of AI is co-authored by those most impacted by its deployment, not just those building it.

As governments worldwide build Digital Public Infrastructure, we provide the civic infrastructure layer—the accountability mechanisms that transform DPI from technical architecture into democratic stabilization. By centering equality, we’re building the foundation that makes digital governance sustainable.

We’ve built an ecosystem that pairs a global community with technical innovations and human rights-based design frameworks.

The Framework: **From Ethics to Engineering**

We replace vague principles with rigorous, academically validated methodology that governs the entire AI lifecycle.

Our open-access tools make **human rights-based AI development practical**, not aspirational.

Our 6-Part Framework

Formally validated by the European Society for Engineering Education, intervenes before bias is built into models, not after. It pushes developers, buyers and builders to question power structures at 6 critical stages:

1. **Objective & Team**
2. **Requirements**
3. **Data**
4. **Modeling**
5. **Testing**
6. **Monitoring**

White paper

The technical foundation sits in our White Paper, ***Integrating Human Rights Considerations Along the AI Lifecycle***, authored by Emmta Kallina, Sofia Kypraiou, and Caitlin Kraft-Buchman. It bridges international legal standards and engineering requirements, arguing that human rights must be design constraints, not afterthoughts.

Assessment tool

Practitioners drown in regulatory compliance, so we built the **Human Rights Impact Assessment (HRIA) Workbook**—a strategic 8-page distillation of the Alan Turing Institute and Council of Europe's 100-page HUDERIA methodology. It strips academic jargon while maintaining legal rigor, letting teams pivot from "vibes-based" ethics to documented accountability. Teams can identify affected rights-holders, assess risks, and implement mitigations in real time.

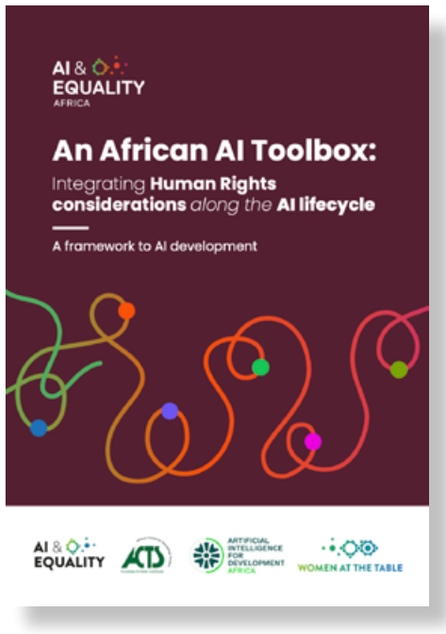
Regional Co-Creation: **Africa and Latin America**

Most AI is designed by people who will never use it in the contexts where it matters most. We champion regional co-creation, ensuring the tools that govern AI are built by the communities they serve.

AI & EQUALITY AFRICA

The African AI & Equality Toolbox

Led with the African Centre for Technology Studies (ACTS), the **African AI Toolbox** applies our Framework across six strategic sectors: Agriculture, Climate, Health, Education, NLP, and Technology-Facilitated Gender-Based Violence (tfGBV). Rooted in feminist, decolonial, and JEDI principles, it curates groundbreaking African case studies proving that innovation is inseparable from human rights.



The work speaks for itself: Makerere AI Lab's 3D-printed smartphone adapters turn standard microscopes into offline malaria diagnostics; sensors.AFRICA empowers citizens to monitor air quality, transforming environmental data into accountability; tfGBV research from Code for Africa provides frameworks for protecting vulnerable users in digital spaces; NLP initiatives like KenCorpus ensure African languages aren't erased from the global AI landscape.

Available as a **standalone website** and **comprehensive library** with video interviews and case studies, the Toolbox is a living document proving that **Design by Inclusion isn't ethics, it's engineering necessity.**

IA & EQUIDAD LATINOAMÉRICA

The Latinoamérica AI & Equality Toolbox

Most AI ethics courses are taught in English. For Latin America's 600+ million Spanish speakers, nothing comprehensive existed for free.

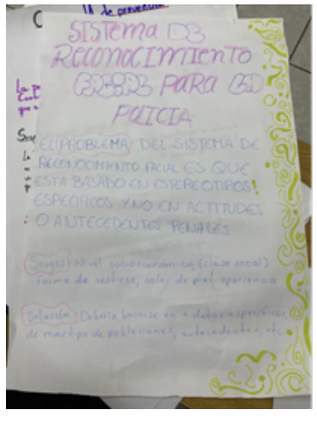
Through a collaboration between Women at the Table, CENIA (Chile's National Centre for Artificial Intelligence), and the University of Chile, led by our strategic designer and Latin America lead Pilar Grant, we reimagined our **AI & Equality framework for the region**. Together with CENIA's "Aequitas" team, we developed deep cultural and linguistic adaptations—translating not just language but legal and social contexts. The result: a bespoke version hosted on UAbierta, the University of Chile's free education platform that reaches across Latin America, far beyond Chile's borders.

In December 2025, the **UAbierta course** launched with 1,500 registrations, with 130 practitioners completing the course for professional credit within the first week. The course analyzes regional case studies, from harmful deepfakes deployed against Chilean minors to the ethical development of LATAM GPT, proving that human rights frameworks must be grounded in local realities.

The course enabled CENIA to secure funding for **Travesías CTCI**, an experimental program using role-playing to teach high school students how to navigate AI risks in daily life. From university-level professional development to grassroots youth education, the collaboration demonstrates the innovation possible when regional institutions co-create rather than import solutions.

Translation of our technical white paper to Spanish.

Our course in UAbierta.



Results from the Travesias CTCI workshop with students.

Technical Auditing:
The Human Rights Benchmark

AI benchmarks test logical reasoning and technical speed. None test whether models respect human rights when making decisions about housing, healthcare, or legal processes.

The HumRights-Bench conceived by our Machine Learning and Society Team led by Dr Savannah Thais with Dr Wm Matt Kennedy is the first technical stress test that audits LLMs against International Human Rights Law, not statistical performance. We simulate the real work of human rights monitors using a specialized IRAC framework (Issue, Rule, Application, Conclusion). Models face complex scenarios validated by at least three legal experts, testing whether LLMs understand the obligations to Respect, Protect, and Fulfill human rights.

The findings are sobering. Leading models score 50-60% on foundational rights like access to water; a failing grade. They struggle to identify legal obligations, defaulting to implementation logistics rather than recognizing rights violations. Worse, performance is stochastic, it is wildly variable across runs, proving they lack stable legal reasoning.

As the benchmark expands into Due Process and other rights through our 2026 University Consortium,we are creating scalable evaluation frameworks that maintain expert-level rigor. Before a model is deployed in high-stakes environments, it must be proven safe not just by its code, but by the law.

Resources

HumRights-Bench Blog and Pub-Talks Series with Savannah Thais:
https://humrightsbench.com/about/blog/

Public Interest AI: Procurement as Democratic Infrastructure

AI systems are being deployed across every level of government—from welfare benefits to criminal justice to healthcare triage. Yet public agencies lack the tools to demand human rights compliance from the vendors they're purchasing from. Procurement is the most strategic lever for AI accountability we have, and it's been ignored.

Our foundational research on this gap was accepted to **CHI 2026**, the leading international conference on Human-Computer Interaction. Led by our head of AI & Equality R&D, Emma Kallina, as part of her post-doctoral research, the paper investigates how public sector agencies can leverage their purchasing power to align AI systems with public interests. Working with procurement experts across the EU and UK, we identified concrete mechanisms to control what private sector AI enters the public sector.

The cornerstone is our **Procurement Framework**, launched at UNESCO in June 2025. It provides cities and public institutions with technical and legal scaffolding to scrutinize everything from data provenance to long-term community impact, ensuring they don't inadvertently purchase bias with public funds.

Our new **Stakeholder Module** goes further, advocating that "stakeholder involvement" must be more than checkbox consultation. It shifts power to affected communities during the acquisition phase, ensuring those most impacted by automated systems have a seat at the creation table.

Learn more

The framework is gaining international traction. We moderated the **UN Business & Human Rights Forum's Safeguarding Human Rights in the Age of Artificial Intelligence**, are conducting workshops and consulting on mechanisms supporting open-source, municipal buyer groups to increase AI supplier leverage, and mandates for multi-disciplinary participation in procurement decisions.

Increasingly, we're also working with **corporations who recognize that public sector contracts require demonstrable human rights compliance**: transforming procurement standards from bureaucratic hurdle into competitive advantage. By making a human rights-based approach the cost of market entry, we're reshaping what gets built in the first place.

The AI & Equality Community:
750+ Practitioners Rewriting the Rules

Traditional AI governance happens in closed rooms. We built a digital commons where **750+ practitioners across 57 countries dissolve expertise hierarchies**. Master's students peer-review senior researchers. Policy experts find technical grounding for advocacy. Global South innovators redefine the vocabulary of technology itself.

Led by Amina Alaoui Soulimani, our AI & Equality Director, the engine is **Open Studios and Pub-Talks**, our community's town square. Between 2024 and 2025, these live sessions moved from Savannah Thais's **Human Rights Benchmark methodology** to Lilian Wanzare's **AI application for Kenyan Sign Language**, from Kristy Claassen's **philosophical disruption of Ubuntu** to Ariane Bar's deconstruction of **data-extractive politics in the French far right**. Whether debating the "New Empire of AI" with Rachel Adams or rethinking health infrastructure through "response-ability" with Amina Soulimani, the community keeps asking: for whom does AI work?

Every session lives in our **comprehensive digital librar**y, a permanent, searchable archive of community-generated knowledge.

Policy and publications

Beyond dialogue, the community is a policy forge.

Through targeted Consultation Pieces, the network intervenes at the highest levels of international governance, ensuring "human-rights-based AI" becomes standard and regulatory requirement, not aspiration. We're not just critiquing the rules, we're writing them.

Community Commentary:
Artificial Intelligence, Cultural Rights, and the Right to Development
AI & Equality Community Commentary in response to OHCHR's Call for input for EMRTD study "Artificial Intelligence, Cultural Rights, and the Right to Development"

Community Publication:
Re-Visions of Now and Future 4
A new volume of essays written by participants of the AI & Equality Human Rights online course, held during the 2025 June/July Summer School.

Community Commentary:
How can AI advance or hinder SDGs?
Input for the Hamburg Sustainability Conference & UNDP on How AI can advance or hinder SDGs. Our comment focuses on the three Ps of People, Planet, and Partnership.

Community Publication:
Re-Visions of Now and Future 3
This volume brings together nineteen essays written by participants of the AI & Equality Human Rights online course, held during the 2025 January Winter J-Term.

Community Commentary:
Independent International Scientific Panel on AI and Global Dialogue on AI
Input for the UN Digital Compact on defining the mandate and framework for the Independent AI Panel and Global AI Governance Dialogue.

Community Commentary:
AI and the UN Guiding Principles on Business and Human Rights
Comment on OHCHR Input call on The use of Artificial Intelligence and the UN Guiding Principles on business and Human Rights

Community Publication:
Re-visions of Now and Future 2
Essays from diverse disciplines and backgrounds to show new perspectives and advocate for what we want our future to look like.

Community Commentary:
UNESCO Publication on AI Regulation
Comment on UNESCO Publication on AI Regulation: Emerging Approaches Across the World

Strategic Partnerships: **From Research to Reality**

Our work is only as strong as our network. We partner with leading institutions to ensure every methodology is academically rigorous, regionally grounded, and strategically deployed.

This architecture, balancing academic rigor with regional expertise and policy access, ensures we're not just writing papers. We're writing the rules.

Academic and technical rigor comes from SCAI (Sorbonne Center for Artificial Intelligence), Cambridge University, Chile's National Centre for Artificial Intelligence (CENIA) and the Alan Turing Institute. These partnerships ensure our frameworks can withstand scholarly scrutiny.

Regional grounding comes from the African Centre for Technology Studies (ACTS) and the University of Chile, ensuring technical standards aren't Eurocentric exports but frameworks refined through local expertise and infrastructure realities.

Implementation reach comes from Code for Africa and AI4D, and CENIA putting our tools in the hands of innovators across both the African continent and Latin America.

Policy influence comes from our collaborations within the International Gender Champions, Action Coalition on Technology & Innovation fo Gender Equality, the Gender in Digital Coalition and others allowing us to shape international human rights standards and diplomatic protocols directly.

2026 and Beyond: **Co-Creating the Future**

We're building the civic infrastructure layer of Digital Public Infrastructure—the accountability mechanisms that strengthen democracy rather than erode it.

When AI systems respect human rights by design, they provide stability rather than volatility, legitimacy rather than backlash. This is the missing layer governments are building DPI without: the equality-centered foundation that makes digital transformation sustainable.

Procurement: Changing development from the ground up

Our CHI 2026 paper establishes what we've long argued: procurement isn't about auditing what vendors built, it's about changing what they choose to build. We're running workshops with VNG (the Netherlands Municipalities Association), the UK's Local Government Association, and Sweden's Demand Acceleration Community to begin 2026. When municipalities demand human rights compliance as baseline, and when corporations recognize this opens markets rather than restricts them, development priorities fundamentally shift. We're creating market conditions where rights-based design isn't ethics theater, it's a competitive necessity. The future we're building: AI systems conceived with human rights as foundation, not scrambled in as afterthought.

HumRights-Bench University Consortium: The standard

We anticipate that ten universities across continents and languages will join us in a 2026 Consortium. Each will develop one new right in depth—Due Process, Housing, Health and beyond, and then test their chosen model or language across all rights the consortium has collectively built.

But this isn't just academic research. We're building the leaderboard that LLMs will be measured against, the benchmark that developers will optimize toward, and the standard that governments will require in procurement. When human rights performance becomes as visible and comparable as reasoning speed, the entire development landscape shifts. The Consortium trains the next generation of researchers while establishing the infrastructure that makes rights-based evaluation inevitable, not optional.

Community: From Critique to Creation

Our 750+ practitioner network continues expanding through Open Studios and targeted policy interventions shaping international governance. Our inaugural Festival of Ideas in 2026 will bring the global cohort together not just to critique existing systems but to prototype the alternatives. We're nurturing a global community that doesn't wait for permission to reimagine what's possible.

The future of AI isn't predetermined.
We have the agency to shape it.

AI & EQUALITY

Join us in creating AI that works for all of us.

aiequalitytoolbox.com