

AI BY & FOR ALL

FAIR Forward
impact report



As a federally owned enterprise, GIZ supports the German Government in achieving its objectives in the field of international cooperation for sustainable development.

PUBLISHED BY

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

REGISTERED OFFICES

Bonn and Eschborn, Germany

GLOBAL PROGRAMME DIGITAL TRANSFORMATION

FAIR Forward – Artificial Intelligence for All
Friedrich-Ebert-Allee 32 + 36
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66
E fairforward@giz.de
I www.giz.de

The programme Digital Transformation and the initiative FAIR Forward are implemented by GIZ on behalf of the **German Federal Ministry for Economic Cooperation and Development (BMZ)**. <https://www.bmz-digital.global/en/overview-of-initiatives/fair-forward/>

RESPONSIBLE

Balthas Seibold, Christian Merz, Kathleen Ziemann, GIZ FAIR Forward – Artificial Intelligence for All

EDITORS AND AUTHORS

FAIR Forward:

Abeera Dubey, Aishwarya Salvi, Balthas Seibold, Bhavika Nanawati, Christian Merz, Daniel Brumund, Deshni Govender, Elikplim Sabblah, Emmanuel Emurugat, Florence Kinyua, Florian Sabary, Golo Rademacher, Isaac Manzi, Ivan Mukibi, Jocelyn Mukamisha, Jonas Gramse, Jonas Nothnagel, Jutta Schueren, Karlina Octaviano, Kathleen Ziemann, Luisa Olaya, Mariana Carrillo, Mark Irura, Mary Afram, Nadine Dammaschk, Philipp Olbrich, Ruth Schmidt

PRODUCTION

Designed, edited and produced by Jaw Design (Pty) Ltd (South Africa).

PHOTO CREDITS/SOURCES

All pictures have been used with permission of the photographers/ image creators.



SPECIAL ACKNOWLEDGMENTS

Sincere gratitude goes to all the FAIR Forward team members across the project period 2019 – 2026, who contributed to the success of the project, promoting 'AI by and for All':

Abeera Dubey
Aishwarya Salvi
Balthas Seibold
Bhavika Nanawati
Blanca Mintz
Christian Merz
Christian Resch
Daniel Brumund
Daniel Glatte
Deshni Govender
Elikplim Sabblah
Emmanuel Emurugat
Florence Kinyua
Florian Bruetting
Florian Sabary
Florian Stratmann
Gaurav Sharma

Golo Rademacher
Gustavo Fonseca Ribeiro
Isaac Manzi
Ivan Mukibi
Jan Nehring
Jiayu Yang
Jocelyne Mukamisha
Jonas Gramse
Jonas Nothnagel
Jutta Schueren
Karlina Octaviano
Kathleen Ziemann
Katja Lueth
Kelia Mungezi
Kevin Muvunyi
Kim Willett
Kirti Pandey

Kuuku Sam
Lea Gimpel
Lisa Werkmeister
Luisa Olaya
Mariana Carrillo
Mark Irura
Mary Seiwah Afram
Milena Stolz
Nadine Dammaschk
Namritha Murali
Philipp Olbrich
Raphael Leuner
Ruth Schmidt
Sheila Kibughi
Siddharth Nandan
Vikram Kumar
Vishwas Chitale

The FAIR Forward team additionally expresses gratitude to the leadership of the GIZ Global Programme Digital Transformation – Björn Richter and subsequently Eric Heinen-Konschak – and its finance team: Jutta Schüren, Katja Lüth and Dieter Schmidt across the project period 2019 – 2026.

Above all, democratising AI worldwide, to create open, inclusive ecosystems of AI by and for all, would not have been possible without the following partners:

- Agricultural Research Council of South Africa
- Ahmadu Bello University (Nigeria)
- Alliance of Bioversity International and CIAT
- Alsisar Impact
- Artpark@IIsc (India)
- Ashesi University
- Association Maidi
- Audiopedia
- Badili Innovations
- Balipara Foundation
- Bappenas – Ministry of National Development Planning Indonesia
- Bayero University
- Biodiversity International
- Blossom Academy
- Bonn Consulting
- BRS – Business Registration Service (Kenya)
- Carnegie Mellon University Africa
- Centre for Remote Sensing and Geographic Information Services (CERSGIS)
- Centre for the Fourth Industrial Revolution Rwanda (C4IR)
- Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
- Civic Data Lab (India)
- Civil Society Budget Advocacy Group (CSBAG) (Uganda)
- Clear Global / Translators without Borders
- Climate Change AI
- Common Room Networks Foundation
- Core23Lab
- Council for Scientific and Industrial Research (South Africa)
- Data Protection Commission (DPC) (Ghana)
- Data Science Nigeria
- Data354
- Dedan Kimathi University of Technology
- Department of Communications and Digital Technologies (DCDT) (South Africa)
- Department of Science, Technology, and Innovation (South Africa)
- DFKI, German Research Centre for Artificial Intelligence
- Digifarm
- Digital Futures Lab (India)
- Digital Green
- Digital Public Goods Alliance
- Digital Umuganda
- DynAg
- Earth Analytics India Pvt Ltd
- Earth System Science Center
- Ecociencia
- ESPOL University
- ETH Zurich
- eticas.ai
- Foreign, Commonwealth & Development Office (FCDO), UK
- FOSS United (India)
- Fundación Despacio
- Future Society
- Gates Foundation
- GEO Indigenous Alliance
- GFA Consulting
- Ghana NLP
- Global Center on AI Governance (GCG)
- Global Partnership for Sustainable Development Data (GPSDD), Kenya
- Government of Telangana
- GramVaani
- Harapura Impact (Indonesia)
- HausaNLP
- Heritors Lab (Ghana)
- Hesotech GmbH
- High Carbon Stock Approach Foundation
- IIT Delhi (India)
- IIT Madras (India)
- Indian Institute of Science (IISc)
- Information Technology, Electronics, and Communications Department, Government of Telangana, India
- Intel
- International Center for Tropical Agriculture (CIAT)
- International Center of Insect Physiology
- International Development Research Centre (IDRC), Canada
- International Telecommunication Union (ITU)
- Jaringan Kerja Pemetaan Partisipatif (Seknas JKPP)
- Jomo Kenyatta University Of Agriculture And Technology
- Kabarak University
- Kaduna State University
- KaraAgro AI
- Karya Inc.
- Kigali Natural Language Processing (KiNLP) (Rwanda)
- Kumasi Technical University (Ghana)
- Kwame Nkrumah University of Science and Technology (KNUST), Responsible AI Lab
- Lahore University
- Limina Education Services
- Local Development Research Institute (LDRI), Kenya
- M-Omulimisa
- Made by People / Tanasuk Technologies Africa, Kenya
- Makerere University, Uganda
- Marine and Coastal Research Institute “José Benito Vives de Andrés” (INVEMAR), Colombia
- Masakhane Research Foundation / Masakhane African Languages Hub
- Maseno University, Kenya
- Mbaza NLP community (Rwanda)

- Meridian Institute
- Meteorological, Climatological, and Geophysical Agency (BMKG), Indonesia
- Microfinance Institution Network (MFIN)
- Ministry of Agriculture and Animal Resources (MINAGRI) (Rwanda)
- Ministry of Communication, Digital Technology and Innovations (Ghana)
- Ministry of Emergency Management Rwanda
- Ministry of Energy and Mineral Development (Uganda)
- Ministry of ICT and Innovation, Rwanda
- Ministry of ICT and National Guidance, Uganda
- Ministry of Information, Communications and the Digital Economy, Kenya
- Move Beyond Consulting
- Mozilla Common Voice Initiative
- Mozilla Foundation
- MTN Ghana
- MUSAP
- Namunyak Conservancy
- NASA SERVIR Global Collaborative
- National Association of Software and Service Companies (NASSCOM), India
- National Coffee Research Institute
- National Crops Resources Research Institute - Uganda
- National Electronic Media Institute of South Africa (NEMISA), South Africa
- National Forestry Authority (NFA) of Uganda
- NITI Aayog, India
- Nokwary Technologies
- Office of the Data Protection Commissioner (Kenya)
- Open for Good Alliance
- Open Knowledge Foundation
- Opportunity International
- Plant Protection and Regulatory Services Directorate (Ghana)
- Producers Direct
- Prosa.ai
- Qhala
- Radiant Earth Foundation
- RECOFTC
- Reiner Lemoine Institut GmbH
- Rochester Institute of Technology
- Rwanda Space Agency (RSA)
- Rwanda Utilities Regulatory Authority (RURA)
- Rwanda Information Society Authority (RISA)
- SADiLaR (South African Centre for Digital Language Resources)
- Smart Africa
- South African Institute of Chartered Accountants (SAICA)
- South African National Space Agency
- Space4Innovation
- Stifterverband / KI-Campus
- Strathmore University (iLabAfrica and Centre for Intellectual Property and Information Technology Law (CIPIT))
- Sunbird.ai
- Tech for Her
- Tech Innovators Network Kenya (THiNK)
- Tecnicafé
- The Uganda National Forestry Authority
- UNDP India
- UNESCO
- UN Global Pulse
- United States International University Africa (USIU) - Kenya
- Université d'Abomey-Calavi - Laboratoire de Biomathématiques et d'Estimations Forestières (Benin)
- University of Energy and Natural Resources (Ghana)
- University of Johannesburg (South Africa)
- University of South Africa (UNISA)
- University of Stellenbosch (South Africa)
- University of Twente
- University of the Western Cape (South Africa)
- University of the Witwatersrand (South Africa)
- Verify.earth GmbH
- Viamo
- Villgro Africa
- Wadhvani AI
- Western Cape Department of Agriculture
- World Resources Institute
- WRMS - Weather Risk Management Services
- ZINDI





GIZ FAIR Forward Impact Report & Replication Kit:

AI by and for All © 2026 by **GIZ FAIR Forward** is licensed and available in Open Access under the **Attribution-ShareAlike 4.0 International** licence CC BY-SA 4.0 DEED.

ATTRIBUTION

You must give **appropriate credit**, provide a link to the licence, and **indicate if changes were made**.

SHARE-ALIKE

If you remix, transform, or build upon the material, you must distribute your contributions under the **same licence** as the original. Any derivative work should include the following visible disclaimer: "The present work is not an official GIZ publication and shall not be considered as such."

CITATION

Please cite as follows: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) – FAIR Forward (2026). FAIR Forward Impact Report and Replication Kit: AI by and for All. Licensed under CC BY-SA 4.0 DEED.

LOGO USE ON DERIVATIVE WORKS

Use of the logos of GIZ and FAIR Forward or the imprint on the back matter of the publication is not permitted on derivative works. GIZ is not liable for any alteration of the original content as used in the derivative work. For any derivative work, please notify us at fairforward@giz.de.

DISCLAIMER

The data in the publication has been collected, analysed, and compiled with due care and prepared in good faith based on information available at the date of publication, without any independent verification. However, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH does not guarantee the accuracy, reliability, completeness, or currency of the information in this publication. GIZ shall not be held liable for any loss, damage, cost, or expense incurred or arising by reason of any person using or relying on information in this publication. The views presented in this publication are those of the authors and do not necessarily represent the views of GIZ or BMZ.

EXTERNAL LINKS

This publication contains links to external websites. Responsibility for the content of the listed external sites always lies with their respective publishers. GIZ has made concerted efforts to check the third-party content (links to external sites) to establish civil or criminal liability that may arise. However, constant review of third party links cannot reasonably be expected without a concrete indication of a violation of rights. If GIZ itself becomes aware and/or is notified by a third party that links to external sites give rise to civil or criminal liability, it will remove the link to this site immediately from this report. GIZ expressly dissociates itself from such content.

MAPS

The maps are for information purposes only and do not constitute recognition under international law of boundaries and territories. GIZ does not guarantee in any way the current status, accuracy, or completeness of the maps. All liability for any loss or damage arising directly or indirectly out of their use is excluded.

Bonn, Germany, May 2026



Africa-Asia AI Policymaker Network, Kenya 2024

TABLE OF CONTENTS



01

PREFACE

9

02

LIST OF ABBREVIATIONS

11

03

EXECUTIVE SUMMARY

13

04

ARTIFICIAL INTELLIGENCE BY AND FOR ALL – INTRODUCTION

15

A Vision for Inclusive AI – What, Why and How

16

Key Approaches and Lessons

18

Cooperative AI Sovereignty

19

Replication Kit: Join the Movement

19

Theory of Change

20

FAIR Forward's Global presence and impact

21

05

THE WHY – FAIR FORWARD APPROACH TO AI

22



6





06

THE WHAT - IMPACT

26

"What" - global impact on Democratising AI	27
Impact in numbers - from a global perspective	27

07

GLOBAL FOOTPRINT: LANGUAGE AI, CLIMATE PROTECTION, AGRICULTURE, KNOWLEDGE, AI GOVERNANCE

28

Language AI by and for all - Making AI speak your language	29
Climate change AI - Climate protection and resilience makes more possible	31
Agri AI - better harvest, food security, farmer advice and less crop diseases	33
Democratising AI Knowledge and closing the gender gap	35
State meets AI - Designing effective AI policies	37

08

39

IMPACT BY COUNTRY AND REGION

ASIA

> India	40
> Indonesia	43

AFRICA

> Ghana	45
> Kenya	48
> Rwanda	51
> South Africa	54
> Uganda	57



09

THE HOW (PART I)

Key approaches to democratise AI by and for all	60
Solving the data-conundrum globally – empowering local communities to solve real-world challenges in agriculture, climate, and language with open-source AI datasets	61
How to build your own open-source AI with reusable building blocks, datasets and AI models	62

10

OPEN-SOURCE AI IN PRACTICE I

Key benefits for society, companies, organisations, and development actors	65
Open-source AI: Distinctive benefits for society	66
Open-source AI: Distinctive benefits for companies, organisations and communities	67
Open-Source AI: Distinctive good practices of durable development cooperation	69

11

OPEN-SOURCE AI IN PRACTICE II

The challenges and solutions	78
Caveat 1 – Avoiding free-riding: Recommendations to communities, governments and large tech companies	79
Caveat 2: Mitigating risks in the area of privacy, protection of marginalised groups and potential for abuse	80

12

THE HOW (PART II)

Challenges, Lessons, Recommendations	86
Lessons learnt: AI foundations for local innovation	87
Lessons learnt: Strengthen local technical know-how in AI	89
Lessons learnt: Policy frameworks for ethical AI	90

13

OUTLOOK: AI SOVEREIGNTY: BUILD TOGETHER. A GLOBAL AGENDA

Outlook	93
Conclusion	94



01

PREFACE

When the project **FAIR Forward – Artificial Intelligence for All** was launched in 2019, it was one of the first, pioneering global initiatives to address the democratic and equitable development of Artificial Intelligence (AI) with global partners. As a flagship initiative of the German Federal Ministry for Economic Cooperation and Development (BMZ), FAIR Forward has been an integral part of Germany's AI strategy. From the outset, we at BMZ have promoted international cooperation as a necessary foundation for responsible, inclusive and open AI innovation for sustainable development.

BMZ has further deepened its commitment to fair AI through the **Hamburg Declaration on Responsible AI for the Sustainable Development Goals (SDGs)**. FAIR Forward has been an essential component of the Hamburg Declaration: by democratising AI through open-source technologies, local capacity building, and inclusive governance frameworks across seven partner countries, the initiative has transformed commitments into impact – from 36 open-source language datasets serving over 1 billion speakers in their native language to AI-powered climate and agriculture solutions reaching millions.

The project's work also represents a paramount step in preparing for sectoral AI applications with our partners in Africa, Asia and beyond. As BMZ Minister Reem Alabali Radovan said: "AI must become a force for good, a driver of inclusion, equality, and sustainable development". FAIR Forward embodies this

mission by working with partners to build, maintain and strengthen open, locally anchored AI ecosystems.

This work has empowered communities closest to real-world challenges to devise sustainable solutions to pressing issues such as food security, sustainable agriculture, climate protection, women's empowerment, the provision of public services in native languages, and many more.

BMZ would like to extend its sincere gratitude to all government partners of the Republic of Ghana, the Republic of India, the Republic of Indonesia, the Republic of Kenya, the Republic of Rwanda, the Republic of South Africa, and the Republic of Uganda. We would also like to express our gratitude to every organisation, expert, and community member who has contributed to this journey. Their commitment has been instrumental in achieving this impact.

This report serves to highlight the impact of the project over the period 2019 to 2026 but also serves as an invitation. All stakeholders are invited to extensively utilise the Digital Public Goods that FAIR Forward has co-created with its partners: the open-source datasets, tools, models, and guidelines with the lessons learned. BMZ remains committed to future action, policy-shaping and joint learning. Let us continue working together to make AI a digital public good by and for all – for inclusion, equality, and sustainable development.

Katja Vaerst
Head

Division G22 – Data Lab, AI

Federal Ministry for Economic Cooperation and Development (BMZ), Germany



02

LIST OF ABBREVIATIONS

AI

Artificial Intelligence

BMZ

German Federal Ministry for Economic Cooperation and Development

DPG

Digital Public Goods

DPGA

Digital Public Goods Alliance

GIZ

Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH

MOOC

Massive Open Online Course

NLP

Natural Language Processing

SDG

Sustainable Development Goals



03

EXECUTIVE SUMMARY

When **FAIR Forward** began in 2019, the increasing use of AI was beginning to reshape economies and societies, yet the global majority of economies were (and still are largely) written into the story as a spectator, not a co-author. Despite promises for AI-powered innovation to accelerate progress on the SDGs, its benefits risked deepening existing divides: where localised datasets, technical capacity, and enabling policies were scarce. This report outlines how the FAIR Forward initiative changed that narrative through intentional, curated and inclusion-by-design actions: co-creating inclusive AI ecosystems grounded in local languages, open-source AI, open knowledge, and equitable partnerships, making AI by and for all.

FAIR Forward – AI for All initiative, funded by the German Federal Ministry for Economic Cooperation and Development (BMZ) and implemented by GIZ, is premised on a simple conviction – that AI can and should be democratised and open-sourced.

Over the period 2019 – 2026, the project set out to put this conviction into action through carefully curated initiatives and programmes designed around open-source AI, local innovation, and inclusive governance with seven partner countries: Ghana, India, Indonesia, Kenya, Rwanda, South Africa and Uganda. The impact was local and global.



This report showcases FAIR Forward's approach, impact, and replicable solutions for democratising AI, highlighting the initiative's three areas of action:

- 1 Access to open-source training data and AI technologies for local innovation.
- 2 Strengthening local technical expertise and capacities in AI.
- 3 Developing policy frameworks for responsible AI governance.

The report is structured on a similar three-pillar concept – detailing the FAIR Forward journey from concept to impact:

- **The Why** – FAIR Forward approach and values
- **The What** – Global and regional impact
- **The How** – Key approaches, tools, lessons learnt and replicable products

This is further elaborated through a journey into the thought process behind the theory of change, the actions needed to get from concept to collaboration, from talk to tendering and finally policy to practice. The initiative's vision for inclusive AI is supplemented with key approaches and actions under the three pillars, and lessons with success factors for navigating complex topics and development terrains. Relatability and adaptability remain the benchmarks on which AI ecosystem development is founded; thus, the Replication Kit and resource materials outlined in this report contain practical, implementable guidance that can be downloaded, adapted, and reused to ensure that work can be sustained by others. Democratising AI is possible and necessary for a future where technology empowers everyone, everywhere. By embracing open-source collaboration and inclusion, it is possible to create AI that is responsible, safe, secure, and trustworthy and holistically transformative, so that it benefits all.



04 INTRODUCTION: Artificial Intelligence by and For All

A Vision for Inclusive AI - What, Why and How	16
Key Approaches and Lessons	18
Cooperative AI Sovereignty	19
Replication Kit: Join the Movement	19
Theory of Change	20
FAIR Forward's Global presence and impact	21

A VISION FOR INCLUSIVE AI – WHAT, WHY AND HOW

FAIR Forward—Artificial Intelligence for All is a project within the “Digital Transformation for Sustainable Development” global programme implemented by GIZ on behalf of BMZ, and is part of the **AI Strategy of the German Federal Government**.

FAIR Forward has pursued an open, inclusive and sustainable approach to AI on an international level. The project has worked on three main goals:



1

Expanding access to open-source training data and AI technologies for local innovation

2

Strengthening local technical know-how in AI

3

Developing policy frameworks for responsible, safe, secure, and trustworthy AI, and fostering dialogue on data protection and privacy

The project has been working with seven partner countries: Ghana, India, Indonesia, Kenya, Rwanda, South Africa and Uganda. In addition, activities with global partners in North-South and South-South development networks were undertaken.

Across Southeast Asia and Africa, FAIR Forward has demonstrated how locally grounded, inclusive AI ecosystems can help transform complex development challenges into scalable, community-led solutions.

By co-designing AI with the people it serves, the initiative advances a simple but transformative premise: equitable AI must be shaped by local contexts and open-source access to key components of AI. This is the ethos of democratising AI. Systems built on incomplete or non-localised data inevitably overlook the realities of marginalised communities, risking exclusion even if as an unintended consequence. FAIR Forward addresses this gap through a three-

pillar approach: expanding access to open-source AI datasets and models, strengthening local technical capacity, and partnering with governments to establish value-driven AI governance. The resources co-created between 2019 and 2026, including datasets, AI models and tools, are designed to be freely accessible, adaptable, and scalable, ensuring AI serves diverse communities rather than reinforcing inequalities.

Aligned with the UN 2030 Agenda, the Global Digital Compact and the Digital Public Goods Alliance, FAIR Forward treats open, responsible, safe, secure, and trustworthy AI as critical infrastructure for equitable development. Through multi-stakeholder partnerships, the initiative ensures that AI resources from language assistants in Kiswahili to satellite data for protecting Indonesian rainforests are freely adaptable, locally owned, and sustainably integrated into regional and national development strategies. The result is a replicable model for AI that serves diverse communities rather than reinforcing existing divides.

FAIR Forward's work spans language AI, climate resilience, agriculture, and governance by building the foundation for digitally enabled participation in society for **over 1 billion people in their native language** across Africa and Asia.



KEY ACHIEVEMENTS INCLUDE:

1 LANGUAGE AI

Building

OPEN-SOURCE AI FOUNDATIONS

for societal participation for over

1 BILLION PEOPLE



in their native language across Africa and Asia



2 CLIMATE ACTION

Helping protect

THE CLIMATE AND ENSURE FOOD SECURITY

through

91 LOCAL OPEN-SOURCE DATASETS

and **50+** open-source local AI models ready for replication



3 AGRICULTURE



Created AI models to detect crop diseases in India, reaching

116,000 FARMERS



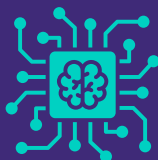
increasing yields and reducing pesticide use. Open datasets for staple crops like maize and cassava support food security.

4 POLICY AND GOVERNANCE

Supported

NATIONAL AI STRATEGIES

in Ghana, Kenya, and Rwanda, and developed **ethical guidelines** for AI procurement, data protection, and algorithmic accountability.



Over

165 MILLION PEOPLE

benefit from improved governance frameworks.



5 CAPACITY BUILDING

Trained

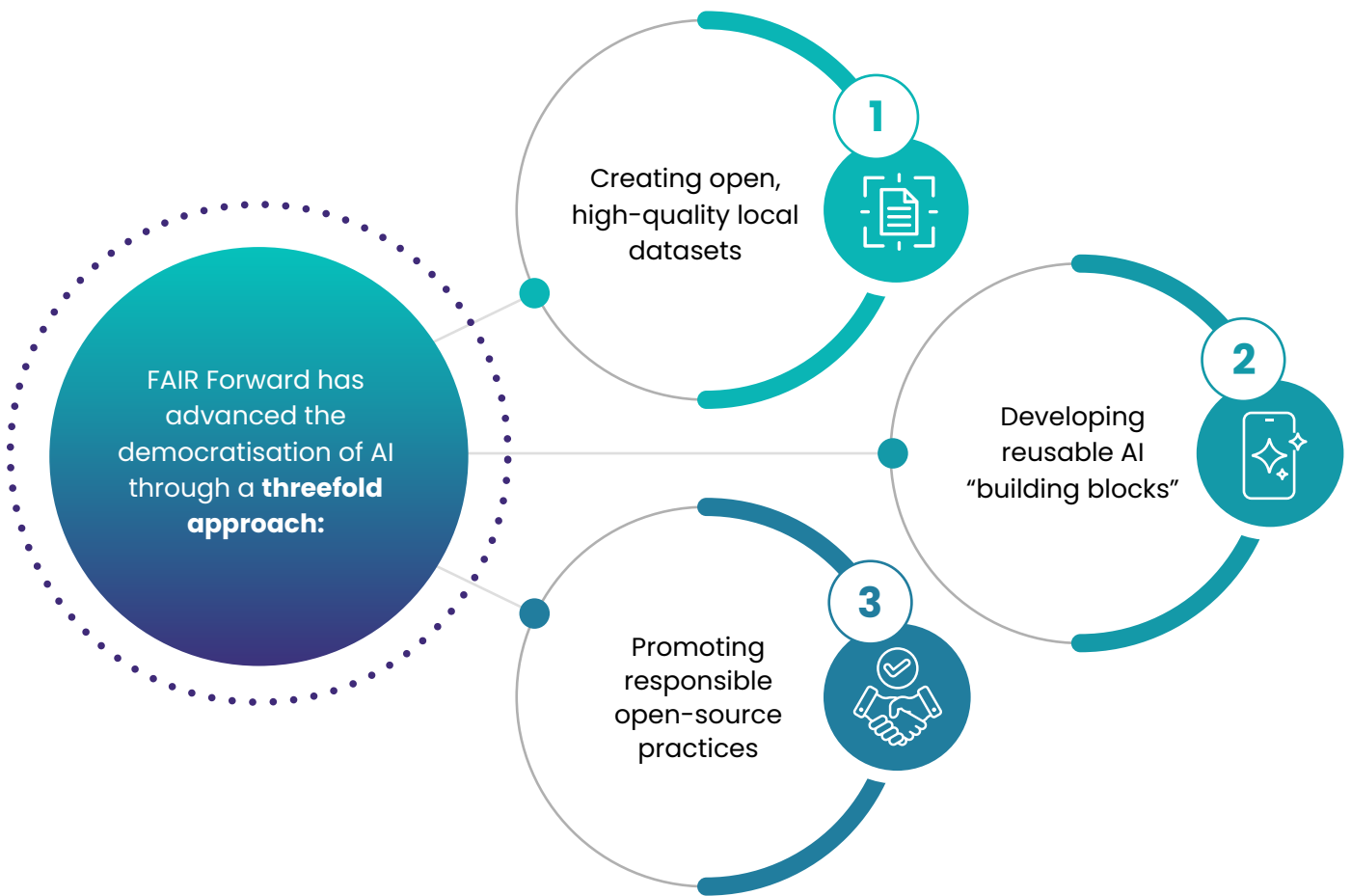
256,000+ PROFESSIONALS

in AI, with a focus on

WOMEN AND MARGINALISED GROUPS

through initiatives like Africa's first AI and Data Science Bootcamp for Women.





Central to this effort has been the Lacuna Fund, which addressed a critical barrier in the global majority countries: the lack of accessible, locally relevant data.

The Fund developed over 75 datasets that enable practical AI solutions in areas such as agriculture, climate, health, and language inclusion. Building on this foundation, FAIR Forward and its partners have developed open-source models, datasets, and code for more than 56 use cases, allowing others to replicate and adapt solutions as digital public goods rather than starting from scratch.

The open-source approach fosters fairer markets, shared digital sovereignty, faster capacity building, and efficient use of public

funds, benefiting societies, companies, and development actors alike. Simultaneously, challenges such as free-riding, privacy risks, and potential misuse pose a real risk to society. FAIR Forward has been addressing them through tailored licensing, responsible AI assessments, and safeguards for marginalised groups.

Lessons learnt highlight the importance of investing in shared, reusable assets, strengthening local ecosystems, and adopting a systemic perspective that connects data, skills, infrastructure, and governance. Overall, FAIR Forward demonstrates that open, locally anchored, and responsibly governed AI ecosystems enable communities closest to real-world challenges to develop sustainable and scalable solutions.

COOPERATIVE AI SOVEREIGNTY

As AI governance grows increasingly fragmented, FAIR Forward advances a model of cooperative AI sovereignty. This enables nations to retain strategic control over their digital infrastructure and data content whilst actively participating in global knowledge networks to address global challenges such as climate change and inequality. **By leveraging digital public goods, countries can:**

- > Reduce dependencies on proprietary systems.
- > Strengthen autonomy through transparent, adaptable technologies.
- > Foster inclusive economic growth by enabling local innovation.

FAIR Forward's work demonstrates that open-source AI is not just a technical approach but a strategic governance tool, one that prioritises equity, collaboration, and public benefit. As AI continues to evolve, open and cooperative frameworks will be essential to ensure technology serves all of humanity, not just a privileged few.

REPLICATION KIT: JOIN THE MOVEMENT

Open-source AI on its own does not guarantee impact and remains theoretical. Implementation experience with partners has clarified what works in practice and what to avoid, offering a clear roadmap for practitioners, policymakers, and development partners. The methodologies listed in this report are packaged as modular **replication kits** to accelerate uptake in new contexts. These include open-source datasets, use cases, and e-learning courses to enable faster, gender-transformative and cost-effective AI development, alongside tools for responsible implementation and risk mitigation of AI systems.

For institutions and governments, it outlines key service offerings, including chatbot solutions for accessible public services, AI policy design support, climate-focused AI applications, and targeted training programmes such as bootcamps for women and marginalised groups.

By combining technical tools, policy guidance, and capacity building, the Replication Kit enables stakeholders to develop context-specific, responsible, and scalable AI solutions whilst fostering inclusion, local ownership, and long-term sustainability.



[Access the Online Replication Kit Here](#)

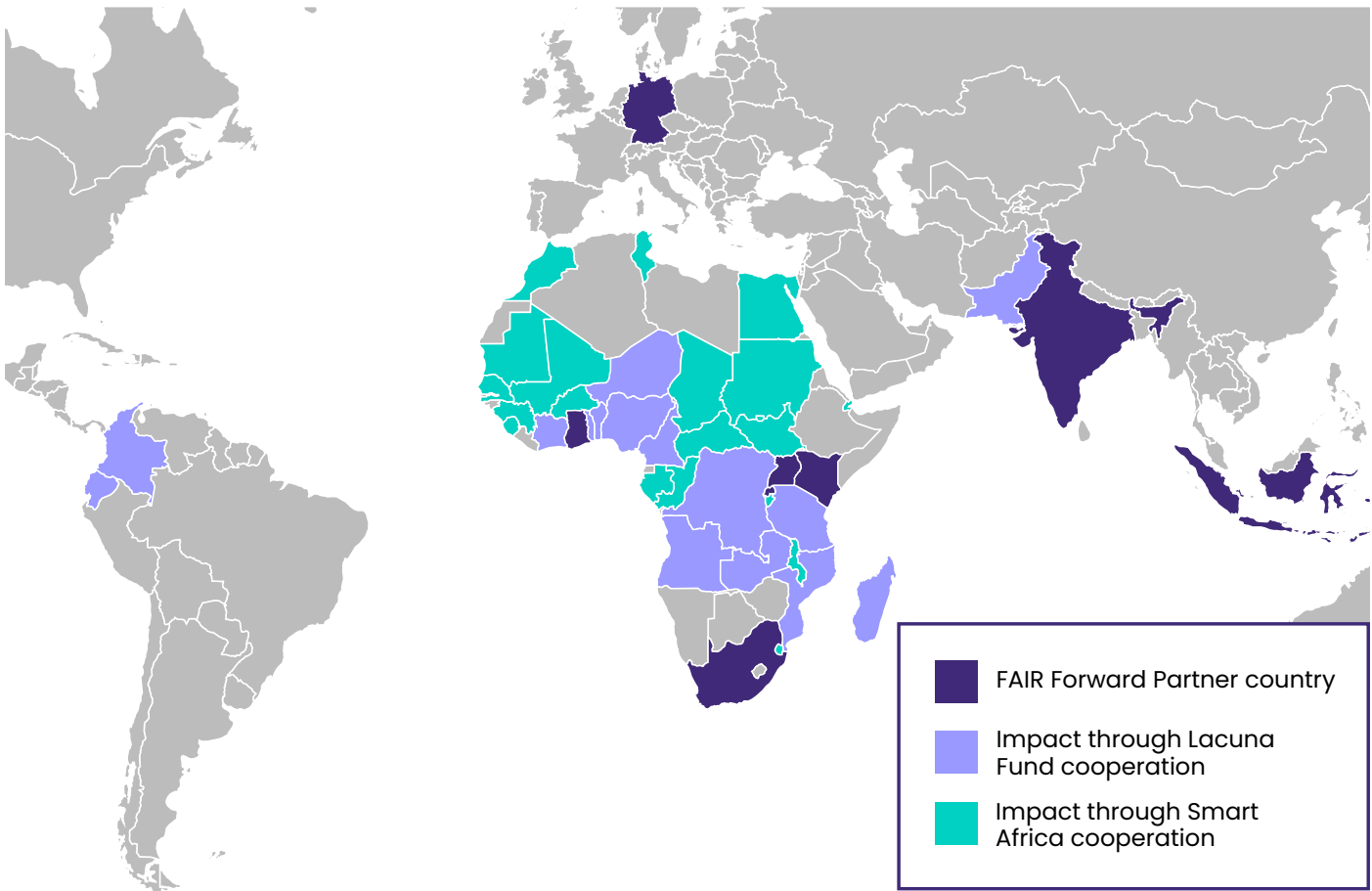
THEORY OF CHANGE

The following illustration makes transparent the adopted change theory of the initiative. It shows how FAIR Forward undertook the change process and tested against the benchmarks for success (via success indicators and outcome indicators).

FAIR Forward's results matrix: how we strive to contribute to change and how we measure this contribution



FAIR FORWARD'S GLOBAL IMPACT:





05

THE WHY- FAIR FORWARD APPROACH TO AI

“Why” – FAIR Forward approach to AI

Too often, AI systems fail the people in the global majority of countries because openly available, unbiased, and localised AI resources (training datasets, AI models, etc.) are missing. A coffee farmer in Uganda faces pest outbreaks that could be mitigated with timely, localised agronomic advice. Yet AI-powered advisory tools trained predominantly on data from Latin American or Southeast Asian coffee plantations misdiagnose regional disease patterns or recommend unsuitable interventions.

The resources that can help create appropriate products and services, fit for purpose, do not yet exist in a manner that can be easily accessed, utilised and adapted for a digital context. Combined with a shortage of local technical skills and political frameworks that are slow to materialise, these gaps represent a major barrier to making AI work in Africa, Asia, and elsewhere in the global majority.

What does “democratising AI worldwide” truly mean? It means fostering the growth and enhancement of open-source AI resources for local development challenges.

These resources can be freely used, shared, remixed and innovated on across projects, companies, organisations, and the globe. These open-source, reusable resources are called digital public goods. Digital public goods include open-source software, open standards,

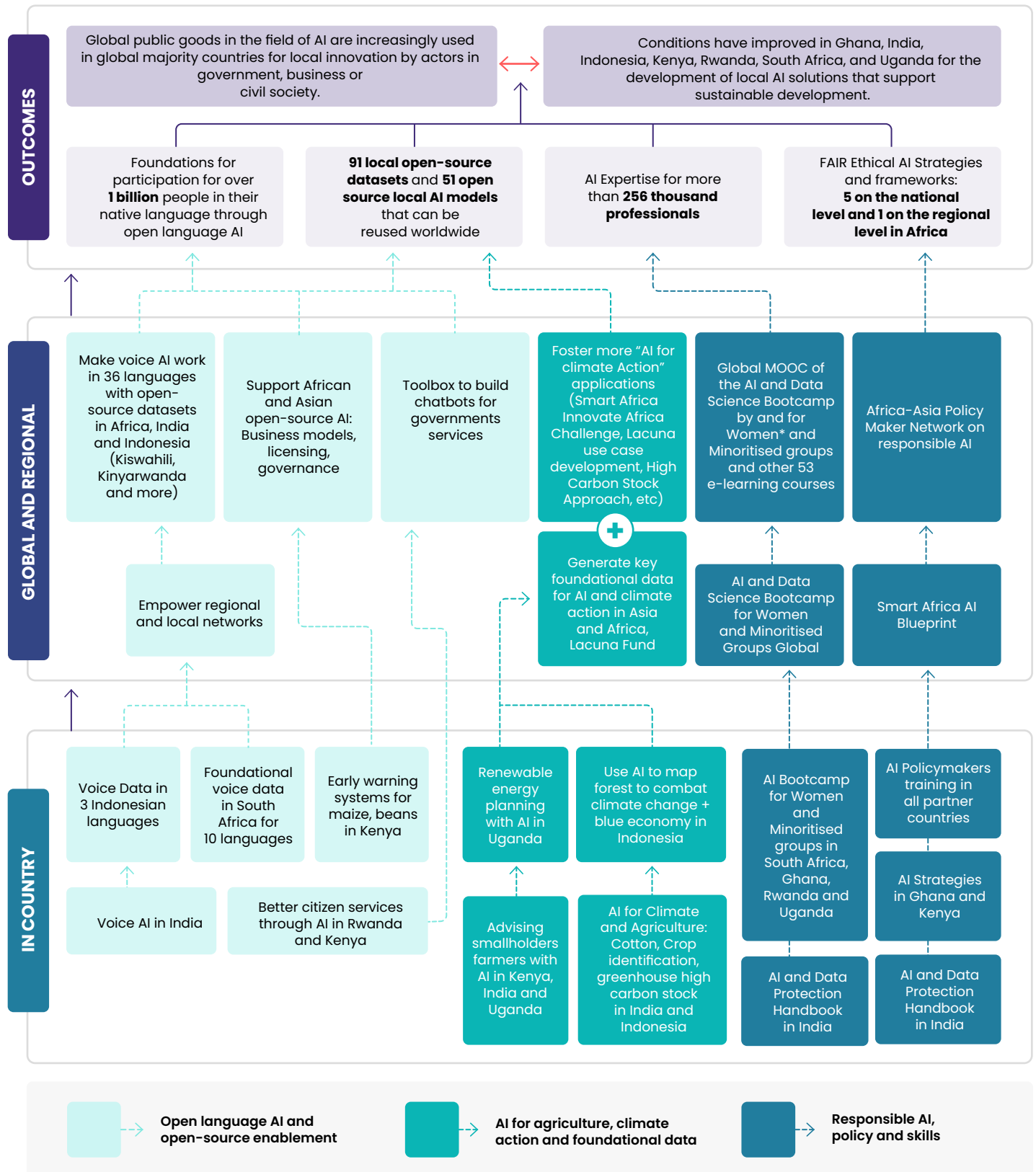
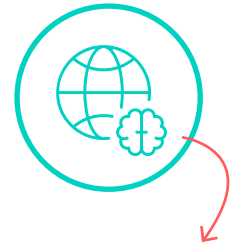
open data, open AI systems, and open content collections that adhere to privacy and other applicable best practices, do no harm, and contribute meaningfully to the achievement of the United Nations 2030 SDGs. FAIR Forward’s approach aligns with this definition from the Digital Public Goods Alliance by developing and promoting reusable, open, and rights-respecting digital solutions with AI for public benefit. Project partners in government, business and civil society have benefited from the development of digital public goods for local innovation. This has helped to create local market ecosystems, improve government services, strengthen digital justice and cooperation, and address shared development challenges.

The AI resources produced take the form of open, modular “AI building blocks” built for cross-context adaptation, and scaling hinges on strategic partnerships. Through targeted multi-stakeholder, multi-disciplinary collaborations, the initiative successfully advanced two interconnected objectives: combating inequalities to leave no one behind and boosting local innovation through open-source AI.



Democratising AI worldwide - the approach in one chart

Fostering local innovation by open-sourcing AI and fostering the principle of leaving no one behind to combat inequalities



1

CHALLENGE:

Benefits of AI and digital transformation are concentrated amongst those already privileged.

How can AI development and deployment reach marginalised communities?

This can be achieved by creating open, locally relevant training datasets, supporting speech technologies for underrepresented languages, and building AI solutions for climate resilience and agriculture. Furthermore, it requires reducing barriers and improving access and awareness, so AI directly serves marginalised groups rather than reinforcing inequalities.



2

CHALLENGE:

AI is most effective when localised to context and adapted to socio-economic needs.

What resources and frameworks can assist in achieving this?

Open-source AI technologies, freely accessible datasets, large-scale capacity-building initiatives, and collaboration with local research institutes and innovators equip regional developers to build AI systems aligned with local languages, development priorities, and economic realities.



3

CHALLENGE:

Regulatory frameworks can be restrictive or encouraging.

How can the same framework protect citizen rights but also encourage innovation through sovereign digital transformation?

The approach involves grounding AI governance in human rights principles whilst supporting countries in shaping their own AI governance and ecosystems. Using tools like Responsible AI Assessments helps to foster privacy, transparency, and accountability without limiting national autonomy or innovation.



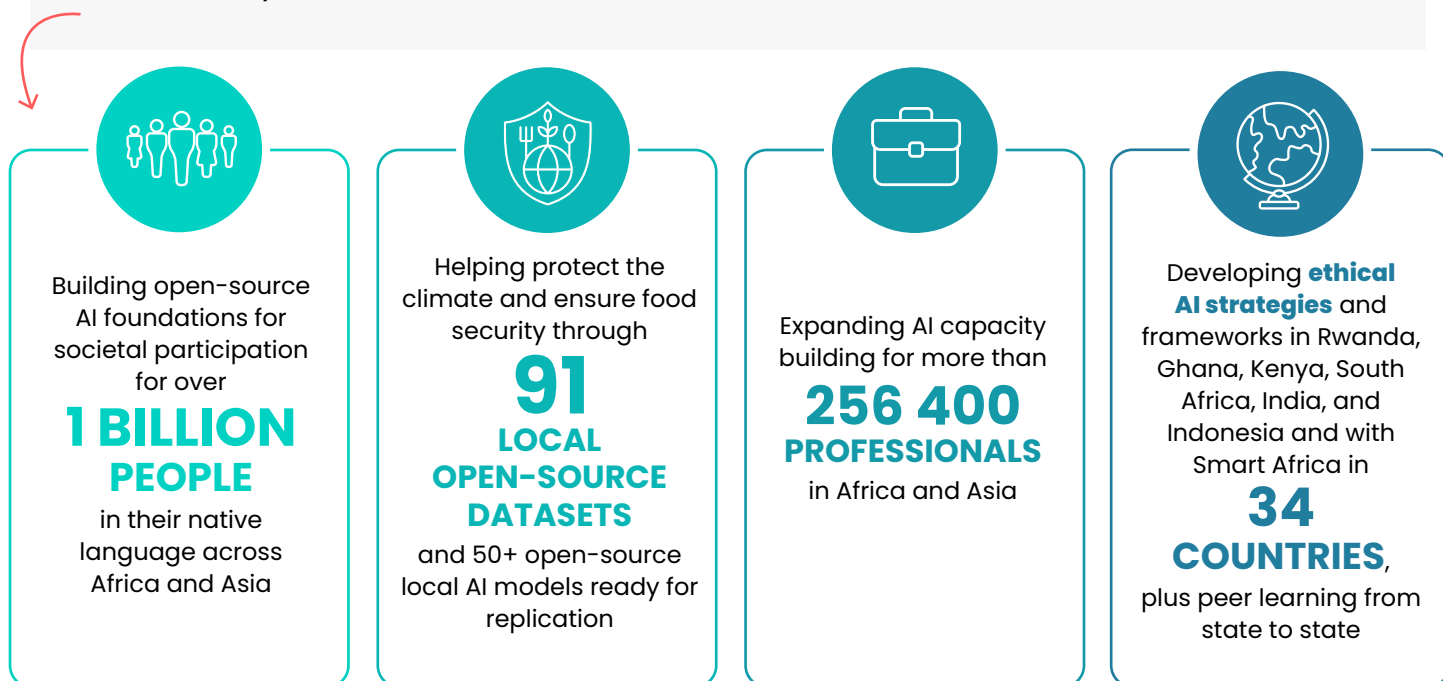


06 THE WHAT-IMPACT

"What" - global impact on Democratising AI	27
Impact in numbers - from a global perspective	27

IMPACT IN NUMBERS – FROM A GLOBAL PERSPECTIVE

Summary of results in the three action areas overall





GLOBAL FOOTPRINT:

Language AI, climate protection, agriculture, knowledge, AI governance

Language AI by and for all - Making AI speak your language	29
Climate change AI - Climate protection and resilience makes more possible	31
Agri AI - better harvest, food security, farmer advice and less crop diseases	33
Democratising AI Knowledge and closing the gender gap	35
State meets AI - Designing effective AI policies	37

FAIR Forward has created globally reusable 'AI building blocks' for 20+ AI application areas, including sustainable agriculture, climate action, renewable energy and public service delivery. Together with its partners in Africa and Asia, the initiative developed 91 open-source datasets to address local challenges.

LANGUAGE AI BY AND FOR ALL – MAKING AI SPEAK YOUR LANGUAGE

The lack of accessible AI training data in local languages means that billions of people cannot use AI-powered applications in their preferred language, dialect, or native tongue.

Therefore, many AI-powered applications that could improve lives will never reach underprivileged communities, simply because they do not speak English or the mainstream languages prioritised in digital communication. Since its inception, FAIR Forward has identified this critical gap and carefully structured actions to address this shortage and inaccuracy of AI training data in the global majority of countries.

Economically, it enables inclusive digital trade, allowing entrepreneurs and professionals to create access to tools, markets, and financial services in the local main language, boosting local innovation and broader participation in the global digital economy. Socially, it is foundational for equitable development in sectors such as education, health and agriculture. This is not merely a technical convenience but a matter of urgency and preservation of culture and identity. It ensures the digital future is not a space of linguistic exclusion but one where their local knowledge, culture, and voice can shape shared progress.

The creation of open-source language datasets with trusted expert partners has led to **91 open-source language datasets**. This has created the foundation for digitally enabled participation in societies for over **1 billion people in their native language across Africa and Asia**, with usable AI building blocks for local start-ups, government and other organisations, to develop products and services with locally relevant language technologies.

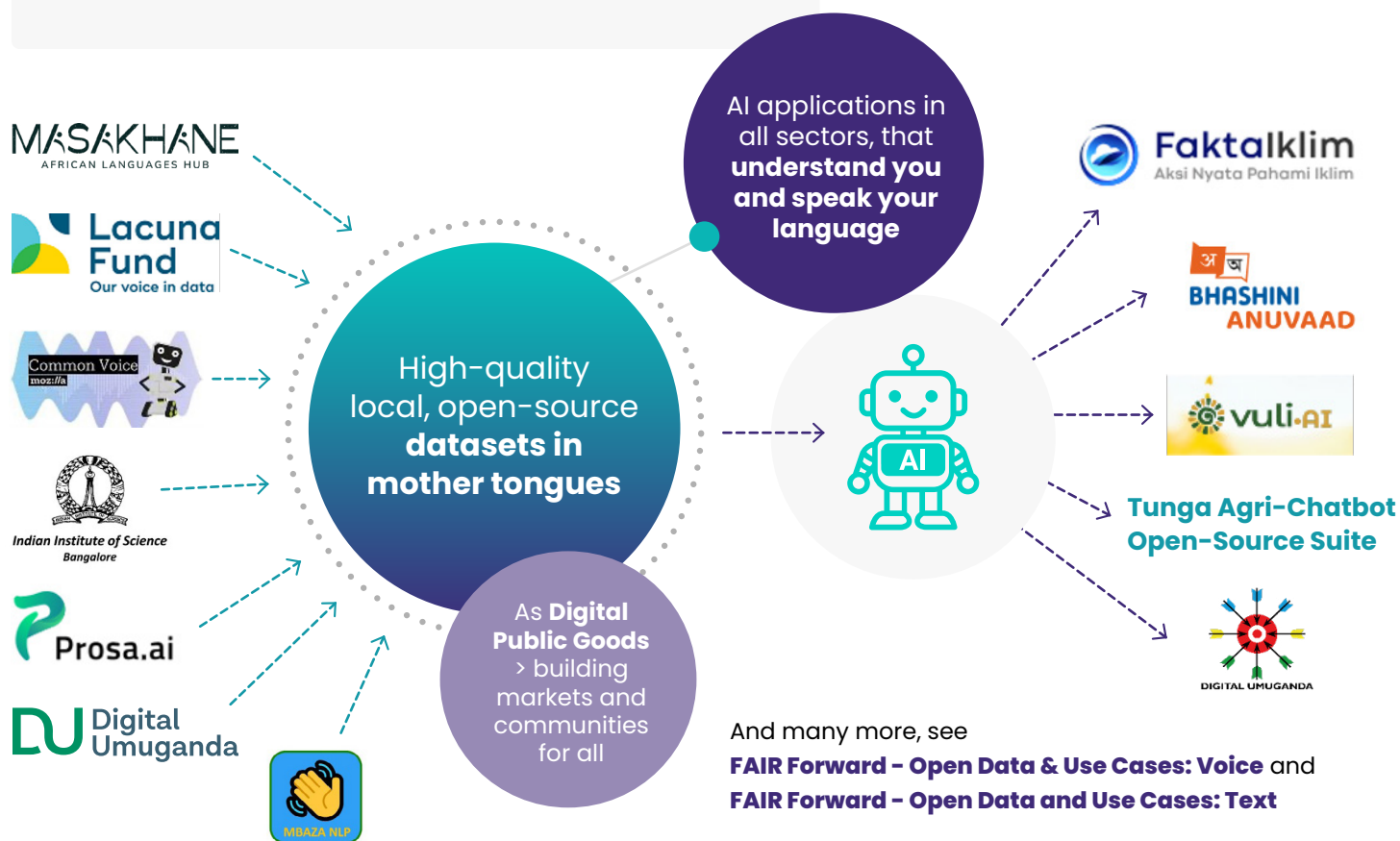
As of May 2026, these open AI language datasets have been downloaded more than 250 000 times for research, experimentation, and further use in applications. Below is a selection of use cases where cross-collaborative, multi-stakeholder partnerships led to tangible impact for citizens, through the creation of open-source language AI.



BUILDING HIGH-QUALITY AI LANGUAGE TRAINING DATASETS IN LOCAL LANGUAGES IS A CRITICAL INVESTMENT WITH PROFOUND ECONOMIC, SOCIAL, AND CULTURAL IMPACT



**Better government services, sustainable agriculture,
climate action through open-source language AI**



Building inclusive language AI for over a billion people: **THE MASAKHANE AFRICAN LANGUAGES HUB**



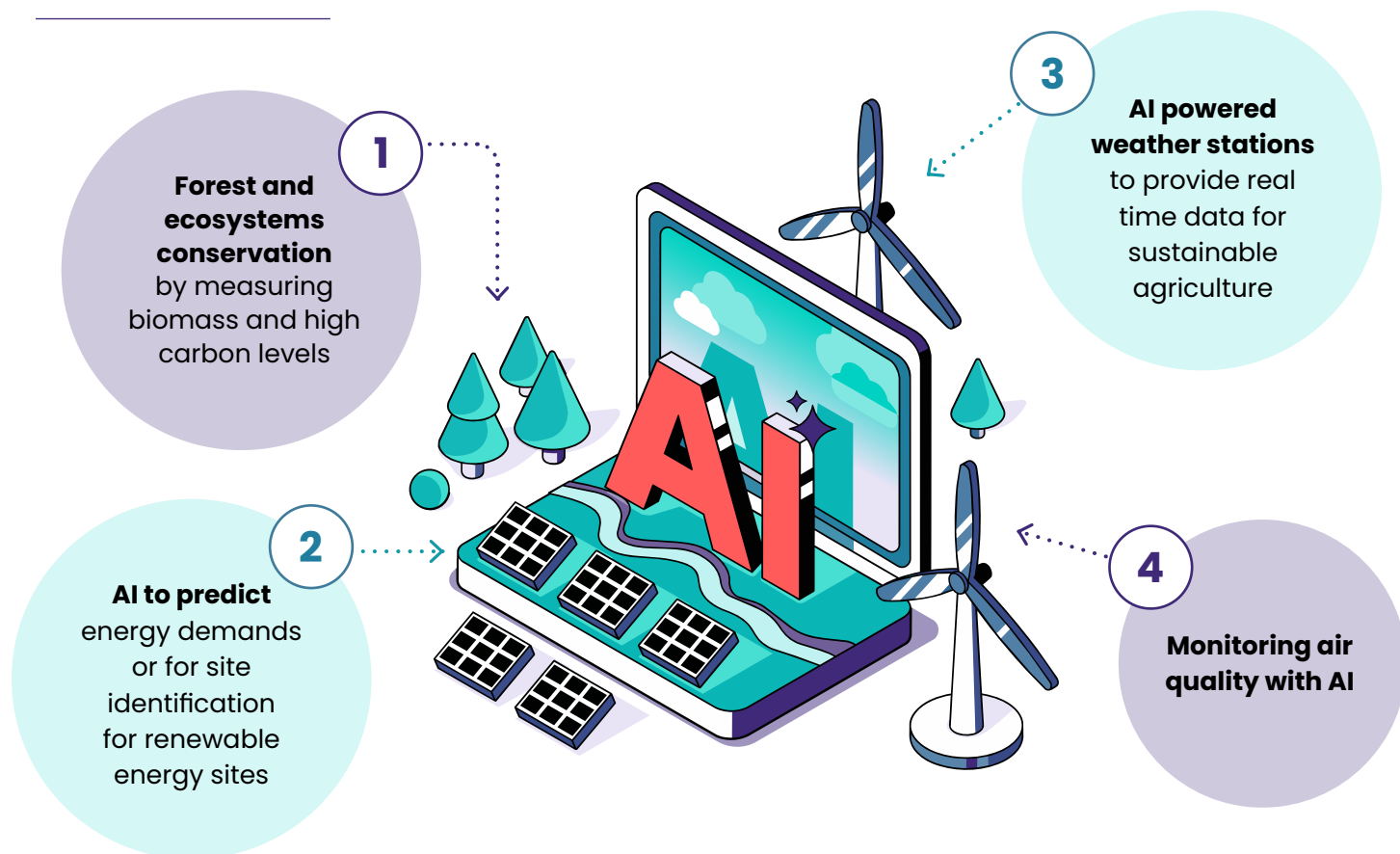
Masakhane (which means “let’s build together”) is a grassroots community of over 2 000 African NLP researchers at the forefront of building AI by, with, and for Africans. In April 2025, the Masakhane Hub for African Languages was launched to address the critical underrepresentation of African languages in AI. FAIR Forward has supported the creation and structuring of the Hub through technical consultation rounds, multi-stakeholder design workshops in Ghana and Kenya, and with inclusive data governance models.

Supporting the Hub creation has led to the collective action by the

to foster the African language AI ecosystem. By developing open-source datasets, AI models, and scalable use cases tailored to local contexts, the Hub is working towards democratising language AI technology. Anchored in community-driven approaches, the Hub aspires to empower over one billion Africans by 2029 by enhancing equitable access to AI technologies.

SEE: [Masakhane – Masakhane African Languages Hub](#)

OUR APPROACHES:



Deforestation is a major contributor to climate change, particularly affecting tropical countries. To counteract this, FAIR Forward has implemented an AI-enabled ‘High Carbon Stock’ approach to promote forest conservation and ecosystem protection.

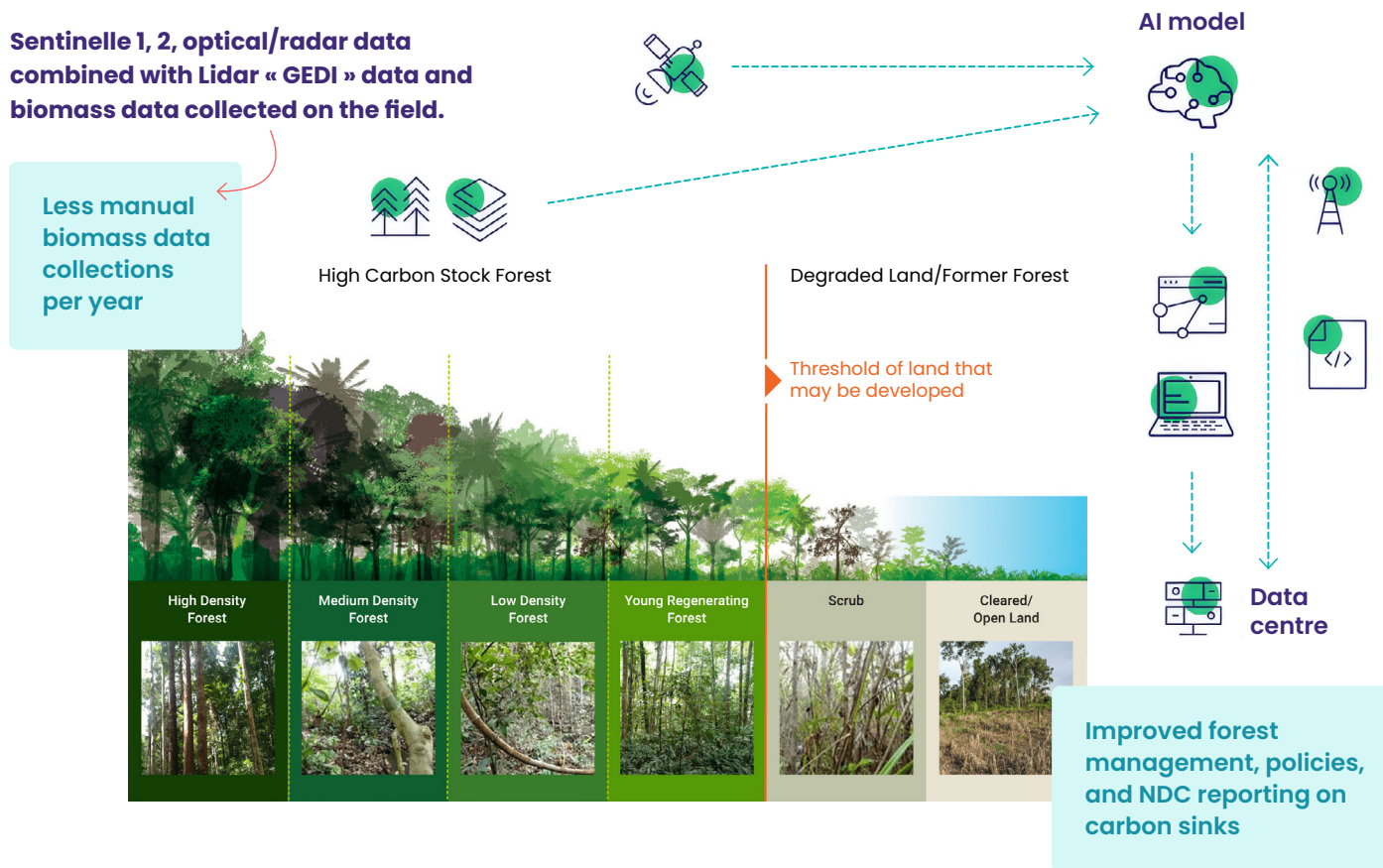
This approach enables the identification, management, and monitoring of forests with high carbon levels (carbon sinks). In Indonesia, AI was used to create large-scale maps of 1.5 million hectares of its tropical forests, combining open-source satellite data, forest field data, and input from local communities. Through these efforts, it is now possible to quickly identify

the location of carbon-rich and endangered rainforests that need to be preserved. Such forests can subsequently be considered carbon reservoirs in national climate protection plans. Leveraging its global network, FAIR Forward expanded and scaled this community-centred approach to forest conservation, fire-risk mapping and carbon sink/ NDC tracking from Indonesia to India and Senegal.

FAIR Forward partnered with Sunbird AI in Uganda to develop an AI-based tool which uses satellite images and AI-powered analysis to detect and sort types of buildings to predict energy demand. This tool creates location recommendations for renewable energy sources and optimises energy supply in rural areas.

In the cities of Gurugram and Patna in India, FAIR Forward supported over 150 citizen scientists using more than 100 low-cost sensors to gather more than 1 million air quality data points. The goal was to generate real-time, hyperlocal data for better air quality management.

This hyperlocal data helped identify specific neighbourhoods most affected by poor air quality, guiding local authorities to develop targeted interventions, such as optimising traffic routes, monitoring construction dust, and urban planning.



Start-ups combating climate change through AI – the Smart Africa Innovate Africa Challenge

Together with Smart Africa, FAIR Forward created the Innovative Africa Challenge- AI for Climate Action, which supported seven start-ups in four countries, providing technical and business expertise to help them develop AI solutions that address climate change through both mitigation and adaptation strategies.

The winner, IMARIKA from Strathmore University's iLabAfrica – chosen for its potential to advance climate adaptation for smallholder farmer communities – builds low-cost automatic

weather station networks to provide access to accurate, local weather information in rural Kenya. This granular data allows advisory services for smallholder farmers to shift from using generic information to location-specific guidance, delivered through farmer organisations and digital platforms. IMARIKA also works on translating weather data into actionable recommendations, addressing farmers' difficulties in interpreting weather information.

Building on the first iteration, Smart Africa and partners are launching further challenges, including on civil drones and AI for climate-resilient agriculture.

Per the Food and Agriculture Organisation (FAO) of the United Nations, smallholder farmers are estimated to produce around a third of the world's food but are severely affected by climate change.

Cashew nuts, cocoa, cotton, and coffee are vital income sources for smallholder farmers across Africa and Asia. Coffee production alone sustains an estimated 125 million people globally, whilst cotton supports around 100 million rural households globally. Corn, cassava, bananas, beans and tomatoes secure the food supply. Cashew nuts, cocoa, cotton and coffee provide important sources of income. For the livelihood of these citizens and the downstream economic consequences, FAIR Forward, with partners, has worked to develop AI-powered plant disease detection tools. These work simply by uploading a mobile phone photo to an app. The mobile app Cotton Ace, developed by the Indian NGO

Wadhwani AI, helps Indian cotton farmers to recognise plant diseases with just one click. Cotton Ace assisted 116 000 farmers in eleven Indian states to protect their harvest. The app was supported by Google.org, and their impact report confirms that since the release of the Cotton Ace app, farmer profits have increased by 20% and pesticide usage has decreased by 25%. The AI models are now being integrated into India's National Pest Surveillance System (NPSS) and extended to ten more staple crops, including rice, wheat, and corn.

Localised datasets alone are not enough. For AI-powered innovation to meaningfully support smallholder farmers, other critical data gaps must be closed. Through the Lacuna Fund initiative and FAIR Forward, over 180 researchers from renowned African and Asian universities developed ground-truth training datasets for crops. This effort also helped establish a market for open-source based, AI-powered agricultural innovations.

**From waiting weeks to instant answers:
How AI is transforming farming advice**



For many years, smallholder farmers have struggled to access reliable advice for their farms. Local extension officers, who are meant to advise farmers, are often stretched thin. In many regions, one extension worker serves up to 5 000 farmers, leaving countless questions unanswered. Visits are rare, and when they do happen, the advice can come too late.

Human advisors also come with high costs of up to \$35 per visit. Millions of smallholder farmers in India, Kenya, and beyond face the same

challenge: farming without timely, trustworthy information.

That's where the Agriculture Information Exchange Platform (AIEP) stepped in. AIEP, funded by the Gates Foundation and coordinated by the BMZ-Initiative FAIR Forward, set out to answer a simple but urgent question:

What if every farmer, no matter how remote, could get personalised, real-time advice in their own language **at little or no cost?**

Since 2024, AIEP has grown into a digital lifeline for more than 40 000 farmers in Kenya and Bihar, India. Instead of waiting weeks for answers, farmers can now open an app, call an interactive voice line, or chat on WhatsApp or Telegram to instantly receive guidance tailored to their crops, soil in local languages.

AIEP leverages Generative AI to tackle persistent challenges in public extension systems, such as limited access to high-quality advisory services, exclusion of women and marginalised groups or inefficiencies in scaling localised extension.

“

Answers are instant. The answer that they provide is always about the question that I have asked, and so far, by using their advice, I have seen improvement in my crop production.

THIS SHOWS THAT THEIR INFORMATION IS RELIABLE.

Female farmer, 49
About AIEP



The following five cohorts developed **minimum viable products**. Early surveys with more than 800 users show high satisfaction, especially amongst women farmers.

OMNICHANNEL DIGITAL ASSISTANT

(Viamo, Producers Direct, HarvestPlus, and Sahaj):

A voice-based advisory solution for beans and wheat farmers in Bihar and Kenya, available in English, Hindi, and Swahili.

1

TECH FOR HER

(Dalberg and DeHaat):

The solution from DeHaat and Dalberg is an open-source architecture enabling customised, inclusive and interactive information exchange by last-mile women farmers & extension agents, interoperable with omnichannel interfaces.

2

MSHAURI

(Opportunity International, DigiFarm, Gooley.ai):

A WhatsApp Chatbot integrated with an AI workflow platform by Gooley.ai. It's an iterated solution initially custom built and first tested in Malawi.

3

FARMER.CHAT

(Digital Green, Karya, Gooley.ai):

An AI-powered agricultural assistant developed for smallholder farmers in rural, low-literacy, resource-constrained environments. It supports seven languages (Swahili, Amharic, Hausa, Hindi, Odiya, Telugu, English) via text, voice, and image and offers personalised advice across 40+ crops based on local practices and needs.

4

DYNAG

(International Rice Research Institute (IRRI), CIMMYT, Gramhal, H3I, IFFCO Kisan, Dexian and Sumarth):

Offers advice through the app ai.sakhi and a voice hotline, covering crops like mushrooms, paddy, and onions in local languages including Hindi, Magahi, Bhojpuri, and Maithili.

5

FAIR Forward has supported significant capacity development initiatives addressing the design, deployment and use of local AI.

In partnership with local institutions across the seven partner countries and the GIZ project atinzi, **over 256 000 people** have been trained in the field of AI. Topics range from AI basics, AI for policymakers, Python coding, to creating and sustaining open-source datasets for applications like natural language processing and computer vision with satellite data.

Recognising that over 75% of AI professionals globally are currently men or male identifying persons, this gender imbalance is reflected by AI systems that do not equally serve all genders

and identities. FAIR Forward created the **first Africa-wide AI and Data Science Bootcamp for Women and Minoritised groups**. This bootcamp was designed to actively address the visible and invisible barriers that women and minoritised groups face in entering and staying in AI. It empowers them with advanced technical and personal development skills to succeed in the AI field.

E-LEARNING COURSES FOR ALL!

FAIR Forward has developed, adapted and/or published

53 E-LEARNING COURSES

on atinzi, from which

254 127 LEARNERS

have benefited.



AI and Data Science Bootcamp for Women and Minoritised Groups

The Bootcamp was piloted in South Africa, then adapted for and scaled to Ghana, Rwanda, and Uganda. Close to 8 000 applications were received for only 180 seats across the Bootcamp locations. This initiative has concretely made a difference in the lives and livelihoods of 136 women and two non-binary persons.

The training was intentionally designed to address barriers that women and minoritised groups face in the tech ecosystem and empower them with advanced AI and personal development skills to succeed in a digitally driven economy. With adaptations like childcare spaces, hybrid schedules that consider participants' work commitments and caregiving duties, safe spaces

to access computers, as well as invisible social and cultural responsibilities placed on women/minority groups - the design was reflected in the results.



The programme modules included, inter alia, technical coding skills (Python), data and statistical analysis, data visualisation, machine learning fundamentals, and AI and data ethics.

To enhance confidence and address the unconscious biases and systemic psychological impact of a gender-imbalanced industry, modules such as communication and presentation skills, product design, gender bias awareness, data and AI feminism, and entrepreneurship were also included.

Intel Corporation supported the initiative by providing its 'AI for Current Workforce Programme', which was incorporated into the overall Bootcamp curriculum.

“

As women, it can be challenging; we have this notion that 'Oh, tech and coding is for guys, so we cannot find our foot in that space.'

HOWEVER, THANKS TO GIZ, IT'S CHANGING.

And it's not just about gender; it's about the value you bring to the table.

Grazia, Ghana

Interested in replicating this Bootcamp? Check the [replication kit](#) with many tools to assist!



AI is often described as a double-edged sword. On the one hand, AI offers new opportunities for countries to achieve SDGs and presents promising avenues for economic and social development and environmental protection. On the other hand, it can create substantial harm if left untethered.

“

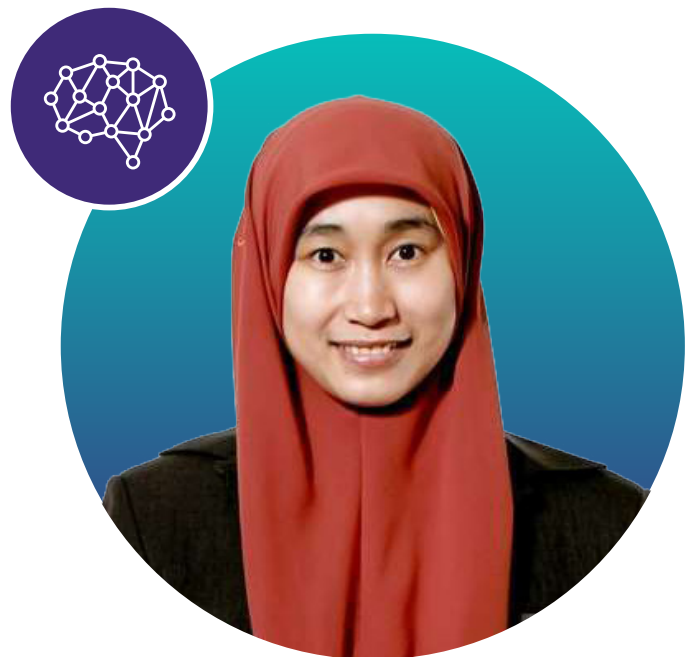
For too long, the conversation around AI has been dominated by concerns of efficiency and innovation, but as we move forward, we must remember that technology is not an end in itself –

**IT IS A MEANS TO
IMPROVE THE LIVES OF ALL
INDONESIANS.**

Prof. Ayu Purwarianti

Head of the Artificial Intelligence Center
at Bandung Institute of Technology

Simultaneously, economies might risk being left behind in meaningfully participating in the intensifying AI competition, especially when they face historical and systemic barriers, including a lack of policy frameworks to guide the responsible use of data and AI development, whilst still trying to be economically competitive.



Together with local partners (academia, government, civil society and industry), with the support of FAIR Forward, local partners successfully concluded locally led initiatives to develop national AI strategies in Africa in the Republic of Ghana, Republic of Kenya and Republic of Rwanda.

Additionally, this work has supported inter-government peer learning across partner countries which accompanied the creation of six ethical and regulatory guidelines in Asia and Africa covering areas such as AI procurement, data protection, and AI maturity assessments (See section on South Africa for the **AI Maturity Assessment Framework (AI MAF)** aimed at assessing South Africa's readiness for AI

integration and on India for regulatory work in India). Presently, **over 165 million citizens of the countries of Ghana, Kenya, South Africa, and Rwanda** would see a benefit from enhanced government frameworks for ethical, inclusive, and responsible AI.

Moreover, FAIR Forward, in collaboration with Smart Africa and the Republic of South Africa, helped champion the development of the **first regional AI policy blueprint** for the continent and the 42 member States of the Smart Africa Alliance in 2021.

Local Answers to Global AI Challenges: The AI Policy Playbook

AI may hold extraordinary promise, but for whom, and by whom? Africa and Asia comprise more than 75% of the world's population, yet they contribute a fraction of global AI research output. This disparity reveals a troubling reality: the risks, rules, and rewards of AI are being crafted without the global majority. Change requires inclusive and responsible AI development, use and governance. The **AI Policy Playbook** emerges as a practice-oriented guide for crafting inclusive, responsible, and context-aware AI governance – through the eyes and experiences of policymakers across Africa and Asia.

Co-created by the **Africa-Asia AI Policymaker Network** and FAIR Forward, the playbook charts a grounded journey through the real-world complexities of AI policymaking in the global majority. It draws from lived experience across seven partner countries – **Ghana, India, Indonesia, Kenya, Rwanda, South Africa, and Uganda** – to offer more than just principles but practical tools, insights, and strategies to shape policy that serves local needs. The playbook is a robust, practice-oriented guide for crafting inclusive, responsible, and context-aware AI governance through the eyes and experiences of policymakers and local partners on the ground, across seven countries in Africa and Asia.





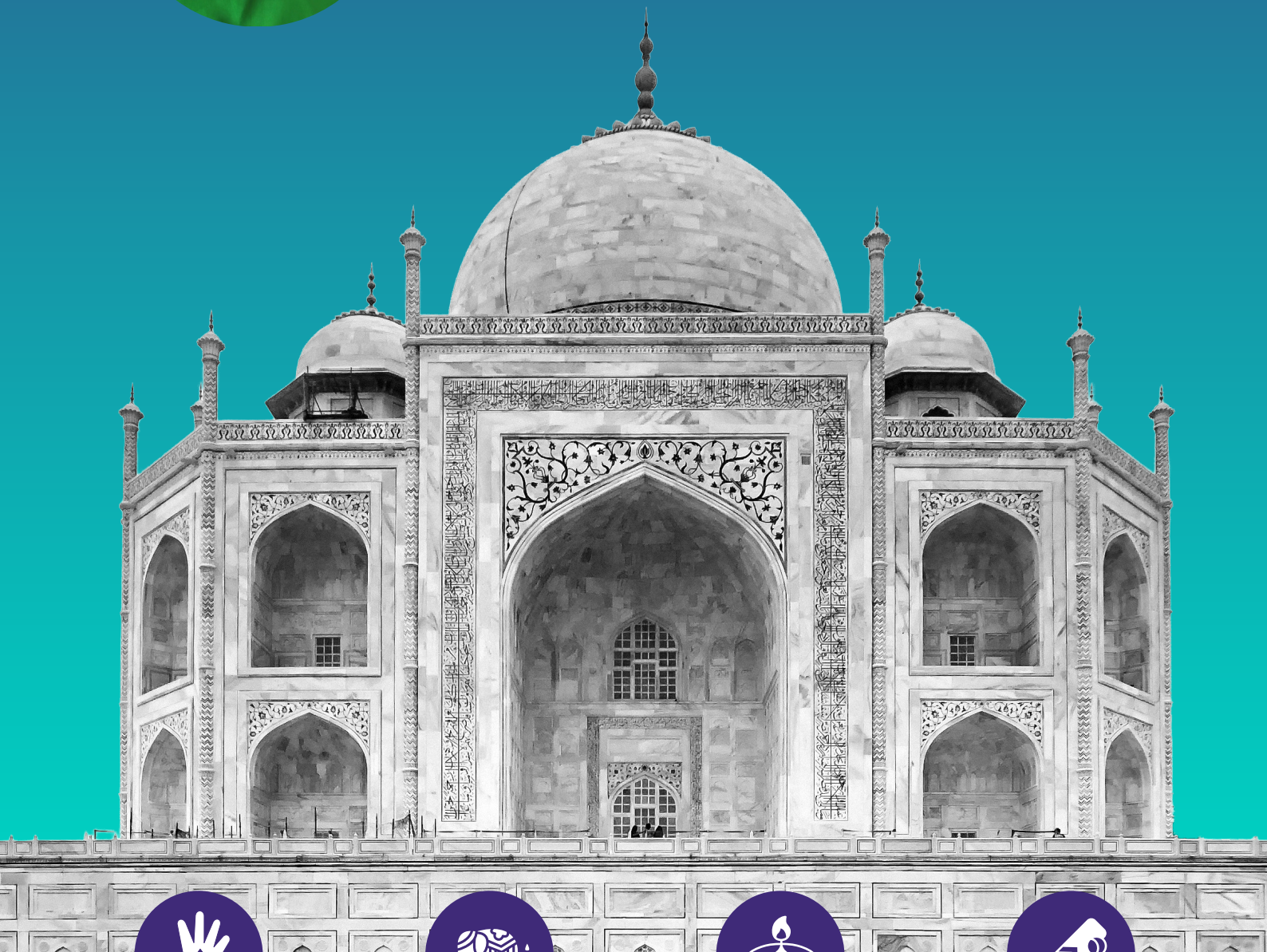
08 IMPACT BY COUNTRY AND REGION

ASIA	
> India	40
> Indonesia	43
AFRICA	
> Ghana	45
> Kenya	48
> Rwanda	51
> South Africa	54
> Uganda	57

ASIA



REPUBLIC OF INDIA



INDIA FACES A MASSIVE CHALLENGE

– bridging the digital divide in a country with more than 22 official languages and even more dialects, and a large population dependent on agriculture.



To address this, FAIR Forward partnered with the Ministry of Electronics and Information Technology (MeitY) and leading research institutions like the Indian Institute of Science (IISc) to develop AI solutions that work for everyone – not just English or Hindi speakers.

As voice technologies are essential for inclusion, the project has focused on building the foundations for AI-based digital services in local languages that can be accessible to millions. One example of this is the Synthesising Speech in Indian Languages (SYSPIN) project, which created 720 hours of voice data across nine Indian languages. This free, open data helps developers create voice assistants and other accessible digital solutions and advisory services. These tools have the potential to unlock services for over 500 million speakers who can interact with technology in their native language.

Moving from language to climate, through AI-powered innovations, FAIR Forward helps farmers adapt to changing climate conditions. With Digital Green, the initiative brought data-driven advice directly to the fields of 300 000+ farmers. Furthermore, by sponsoring challenges for developers, FAIR Forward crowdsourced new solutions for tracking crop health and estimating yields.

Moreover, to ensure India's AI journey is responsible and equitable, FAIR Forward forged key strategic partnerships with:

- 1 BHASHINI – the National Language Translation Mission and the Indian government's flagship initiative to digitally connect citizens in their native languages.
- 2 National Association of Software and Service Companies (NASSCOM).
- 3 Digital Futures Lab.

Together, they developed policy frameworks, best practices and toolkits for developers to build and scale open-source responsible AI. This approach helped to effectively bridge the gap between regulation and development.

“Open, inclusive voice tech in Indian languages is not a luxury–
IT'S A NECESSITY

Mr. Amitabh Nag
CEO of the Digital India
Bhashini Division

Translate complex laws into practical steps for AI developers

The Handbook on Data Protection and Privacy for Developers of AI in India, developed jointly with NASSCOM, the Data Security Council of India (DSCI) and IKIGAI Law, helps bridge the divide between legal requirements and the coding reality. It provides clear explanations of how India's new Digital Personal Data Protection Act, 2023 applies specifically to AI, and provides checklists and case studies for AI developers. By incorporating a thorough review of best practices for embedding privacy, it empowers not only Indian developers but also the global AI community to build AI solutions that are not only innovative but also safe and trustworthy.

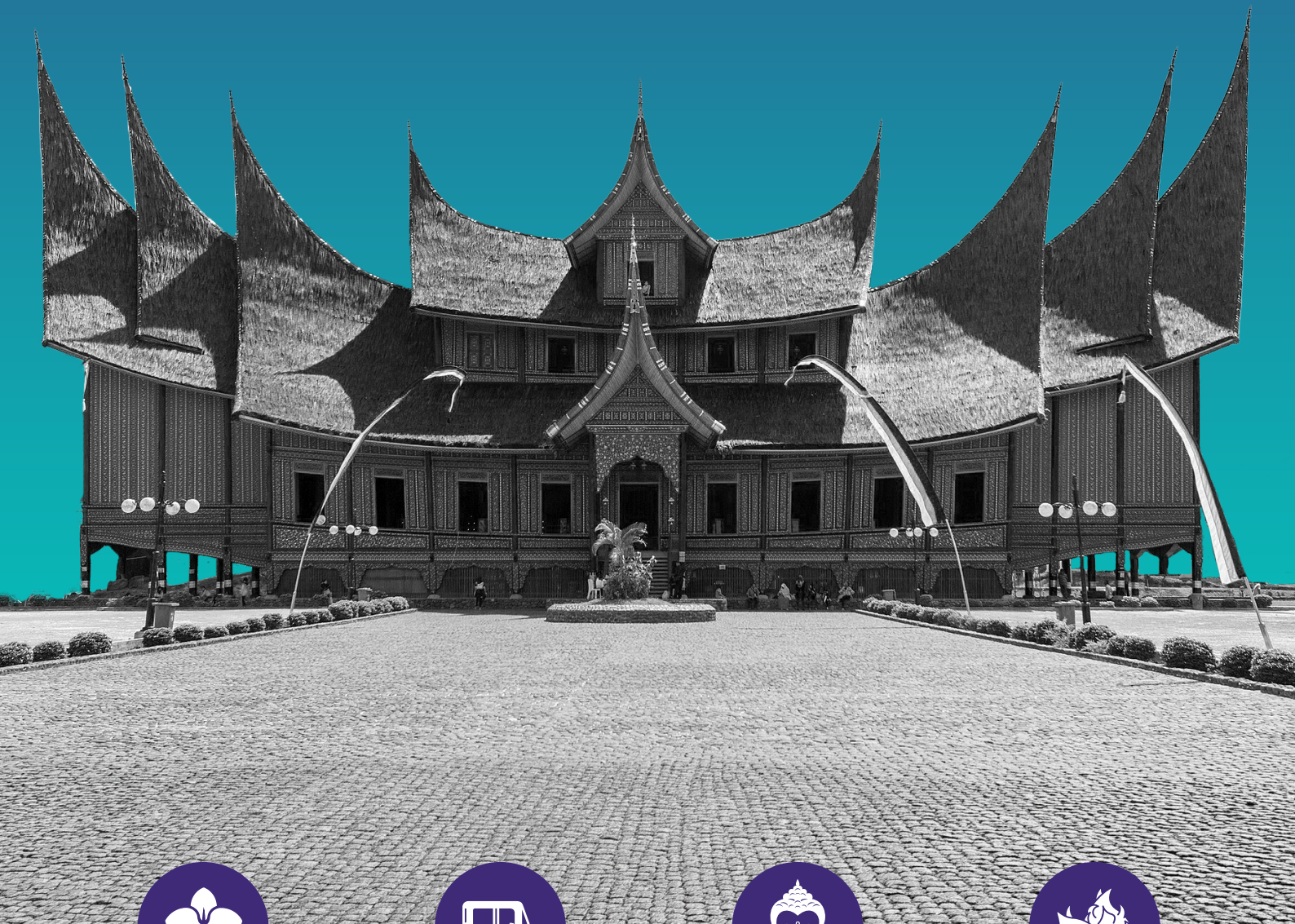




ASIA



REPUBLIC OF INDONESIA



43



In **Pulo Aceh**, one of the hardest hit regions of the 2004 tsunami, coastal communities are now facing new challenges: rising sea levels, unpredictable weather, and dwindling fish stocks. However, history will not repeat itself here:

THE RESPONSE IS PROACTIVE AND POWERED BY TECHNOLOGY.

Supported by FAIR Forward in partnership with the Ministry of National Development Planning (BAPPENAS), the Community-based Innovation Lab for Climate Resilience (Co_LABS), and implemented by **Common Room Networks Foundation**, local women in these remote villages are building digital confidence and becoming drivers of climate resilience in their communities. However, the work does not stop at skills training. Women and communities take an active role in co-creating local solutions to pressing climate challenges, connecting their traditional knowledge with modern technology. Consequently, this initiative created the first community-based weather stations and distributed 53 DIY environmental sensors across 10 provinces, demonstrating how locally led innovation can drive climate adaptation and environmental stewardship.

Indonesia is, however, not only one of the most biodiverse countries but also the second most linguistically diverse country in the world, with over 700 living languages. In collaboration with the woman-led start-up Prosa.ai, the project supported the collection of over ten million 'word' datasets in each of the Minangkabau, Balinese and Bugis languages. By offering data annotation roles that allowed flexible, remote participation, women and rural communities were empowered to contribute to Indonesian datasets. These datasets in 3 local languages were used for the development of Faktalklim, Indonesia's



Deployed two IoT sensors in Pulo Aceh

first AI-powered chatbot that addresses climate misinformation in local languages. The aim is to empower users to verify climate claims using accessible, evidence-based information. It is a powerful example of how AI can support public awareness – by speaking the language of the people.



AI should be able to help journalist work to shorten the time to produce credible and trusted news through Faktalklim. **WE HOPE IT CAN BE A TOOL FOR JOURNALIST PERFORMANCE.**

Fira Abdurachmand
Secretary General of Indonesian
Environmental Journalists.

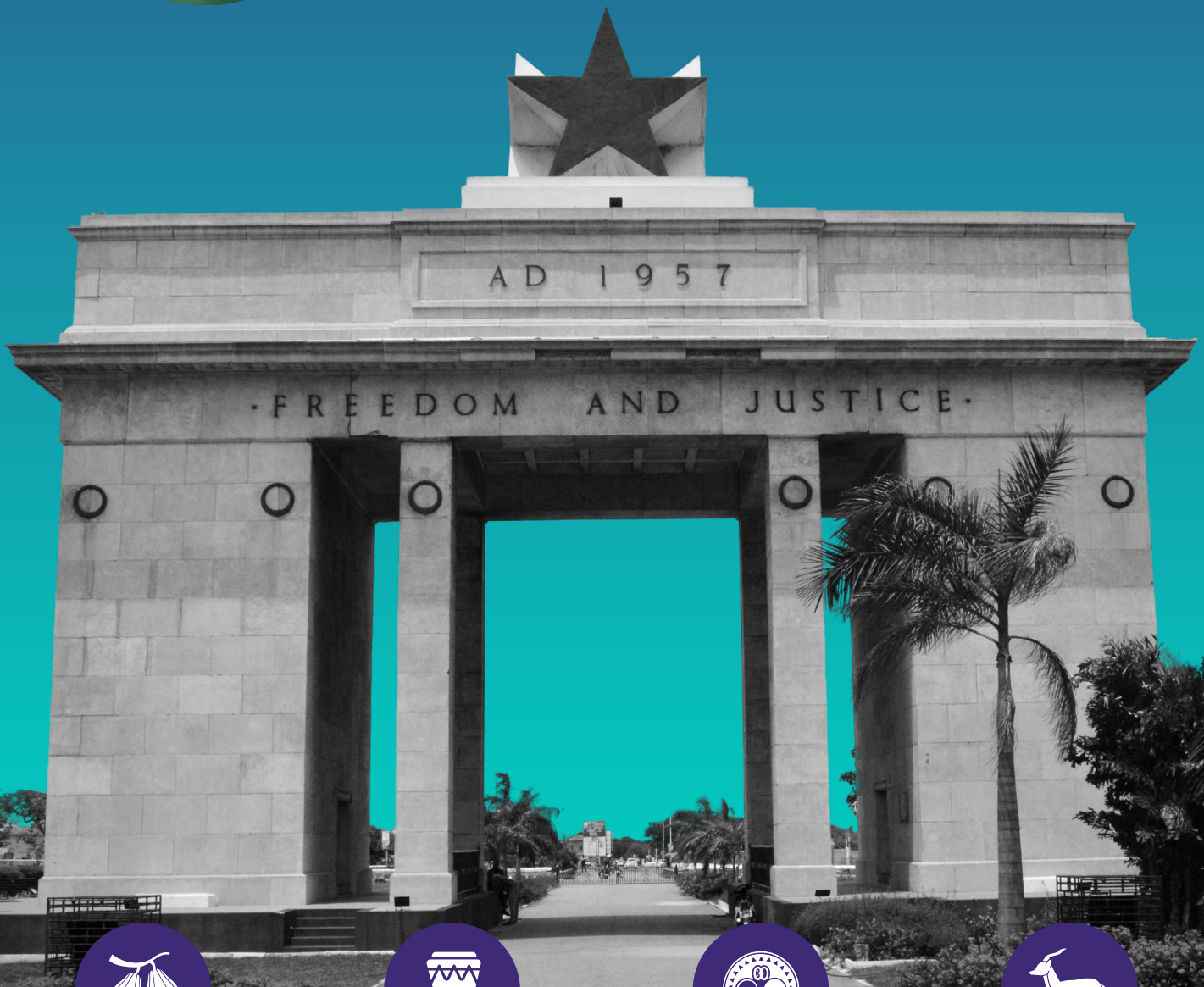
Following inclusion methods used for the development of AI technologies in Indonesia, FAIR Forward, with Harapura and Aptaworks delivered a three-month AI governance training programme for 23 policymakers, encompassing AI ethics, sector-specific policy design, and peer learning from international counterparts – building their capacity to develop ethical policy instruments. This resulted in the foundation of an AI Taskforce, which has already contributed to the drafting of inclusive, human-centred AI policies.



AFRICA



REPUBLIC OF GHANA



Ghana's AI journey in the past few years has been marked by the genuine desire of various actors to leverage AI for sustainable development. The drive has been remarkable, yet it was characterised in the beginning by little coordination of stakeholders and fragmented initiatives.



Also, AI development in the country has been hampered by the lack of openly accessible resources, a reliable expert talent pool and a national strategy to bind it all together. FAIR Forward has contributed to addressing these gaps through targeted interventions in policy, research, openly accessible resources, and capacity building.

By supporting the Ministry of Communication, Digital Technology and Innovations in drafting the Ghana National AI Strategy, the project helped catalyse engagement from national and international partners such as the Foreign, Commonwealth & Development Office (FCDO), strengthening the review and adoption process. This effort has also enabled the Responsible AI Lab (RAIL) at Kwame Nkrumah University of Science and Technology (KNUST) to take on a coordinating role, linking government, academia, and development partners around AI governance and implementation.

With support from FAIR Forward and implemented by Heritors Labs, the **Ghana AI Practitioners' Guide** helped developers, policymakers, and organisations implementing AI systems to translate high-level policy principles into applied AI practice across sectors. The Guide provides actionable guidance on responsible AI development, including ethics, data governance, transparency, and risk management. On the application side, the project has supported the development of AI solutions and datasets

created by Ghanaian innovators specifically for the Ghanaian context. The Cashew Disease Identification (CADI) project demonstrates that AI and drone imagery can detect crop diseases earlier and at a lower cost compared to traditional methods. This has practical implications for smallholder farmers, particularly in improving early intervention and reducing crop losses in an industry that contributes to Ghana's agricultural exports and rural livelihoods.

The programme has further facilitated the creation of open, locally relevant datasets for staple crops such as maize, pepper, and tomatoes. These datasets are publicly accessible and are already being used by local innovators and researchers to develop AI-driven tools, including potential applications for farmers. This contributes to building a foundation for AI and open data-driven agricultural innovation and improved food security outcomes.

Overall, FAIR Forward's interventions have contributed to strengthening coordination, building human capacities, enabling research infrastructure, and supporting practical AI applications in Ghana, whilst maintaining a focus on responsible and inclusive AI development.

Ghana trains 1 300+ new AI talents with a special focus on women

FAIR Forward, in collaboration with Blossom Academy and industry partners, has trained over 1 300 individuals in AI, equipping them with skills to harness their potential within their communities and within Ghana's industry, with a strong emphasis on women's participation.

As part of this initiative, the AI Fellowship for women targeted individuals with little to no coding experience. Over 50 fellows were shortlisted out of 5 000+ applicants, showing a strong demand and thirst for knowledge in

AI. Over a three-year period, FAIR Forward and its partners executed three AI fellowships, each providing a six-month structured programme consisting of three months of technical training followed by a three-month industry placement.

A special focus lay on linking technical knowledge with practical skills needed in Ghana's industry through mentorship exposure. Beyond individual outcomes, the fellowship has influenced Ghana's training ecosystem, which has adopted some key elements of the curriculum, particularly on Responsible and Ethical AI.



The Responsible AI Lab at KNUST: a new regional hub for ethical AI and real-world impact projects

At the institutional level, FAIR Forward co-supported the establishment of the Responsible AI Lab (RAIL). This was done in partnership with and catalysed by the International Development Research Centre (IDRC) of Canada. The lab has become a key platform for applied AI research, knowledge development, and skills training.

As of January 2026, RAIL has trained over 5 000 researchers and 10 000 students in AI and data science-related fields, contributing to the growth of a local talent pipeline. It also supports postgraduate research and the development of open-source AI resources. Through initiatives such as the Responsible AI Network (RAIN Africa)

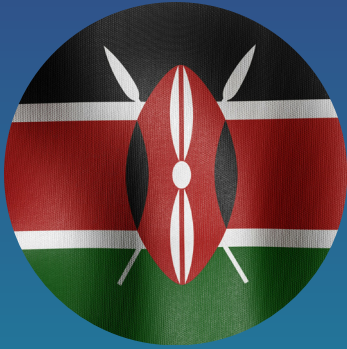
and in collaboration with the Technical University of Munich, RAIL has contributed to advancing research on responsible AI.

This work has produced outputs including an edited volume on AI ethics in higher education **The Need for AI Ethics in Higher Education**, helping to shape discourse on responsible AI in African contexts.

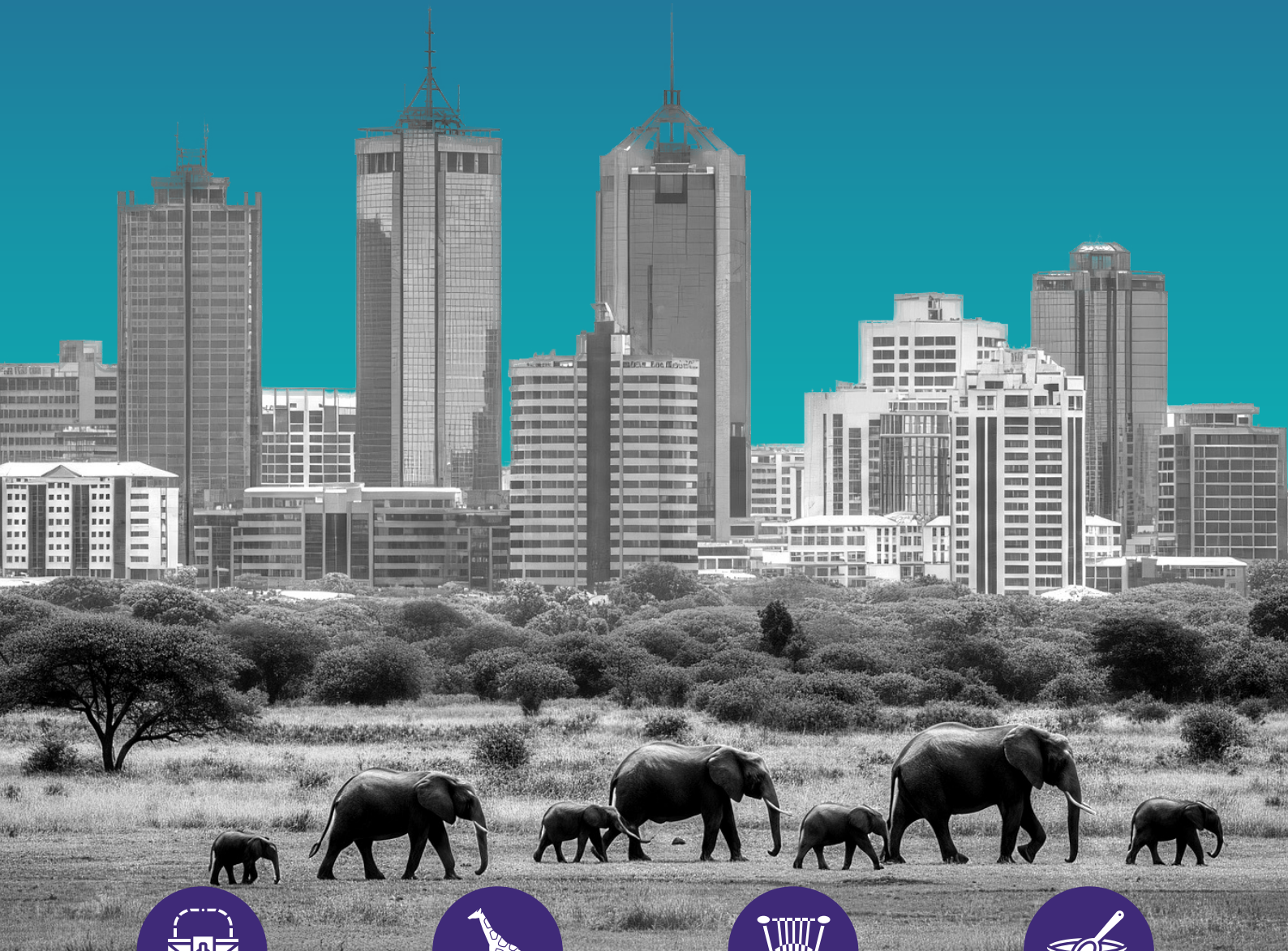




AFRICA



REPUBLIC OF KENYA



KENYA IS WIDELY RECOGNISED AS ONE OF AFRICA'S LEADING DIGITAL INNOVATION HUBS



However, despite rapid growth in digital services and mobile technology, the design, deployment, and adoption of AI were constrained by foundational gaps like limited public sector capacity, restricted availability of local language datasets, and the absence of coordinated national policy frameworks to guide responsible AI development.

Strengthening public sector leadership on AI

was the foundational work needed to establish an engaged and informed ecosystem of decision-makers across finance, agriculture, energy, the judiciary, and others. This enabled public institutions to take informed strategic action around AI adoption and governance on Kenya's national AI governance landscape. With the Digital Transformation Centre in Kenya, a series of multi-stakeholder consultations led to the development of the AI Practitioners Guide in 2023 – supporting responsible and practical AI implementation – and consequently, **Kenya's National AI Strategy**, launched in April 2025.

FAIR Forward worked with Kenyan universities and the Mozilla Foundation to create new datasets accompanied by data governance mechanisms in **Kiswahili and Dholuo**, two widely spoken local languages. These include the use of equitable AI data licensing frameworks, such as Nwulite Obodo Open Data License (NOODL) that aims to protect the agency of source communities over their language data, not only in Africa but across the Global South more broadly.

Citizen-centred, practical AI use cases were a priority-testing an early chatbot deployment and conversational AI for Business Registration Services (BRS) and the Office of the Data Protection Commissioner. Public service chatbots in Kiswahili, English, and Sheng were developed to help citizens access government information in everyday local conversation. However, citizens' struggles persisted with different government departments for e-services. The "GovBot" solution was created – an AI Chatbot interface that connects citizens to services across multiple institutions through a single interaction point on the eCitizen platform.

For many smallholder farmers in Kenya, access to timely information about weather and crop diseases is limited, particularly for farmers in rural areas who rely on basic mobile feature phones.

The Agricultural Information Exchange Platform (AIEP) initiative enables data-backed support to smallholder farmers with crop disease identification and informed crop management choices, using tools combining image recognition, African NLP, and AI-powered advisory services delivered through low-tech channels. For a maize farmer in western Kenya like Achieng, this means receiving timely alerts and practical, actionable advice through digital channels that work with basic mobile devices, removing the wait time for extension officers and taking the guesswork out of farming.



AI-powered Early Warning System: Advisory services on climate-smart farming for Kenyan smallholder farmers

The Early Warning System (EWS) created by the Local Development Research Institute (LDRI) of Kenya is an AI-powered platform that monitors farming activities and supports climate-smart precision agriculture for smallholder farms.

It uses geo-referenced data collected from farms through mobile phones and high-resolution satellite imagery. The system provides crop monitoring, yield prediction, pest/ disease detection, and localised agricultural advisory services in multiple languages.

The platform is currently accessed through both WhatsApp by farmers and a self-serve platform/ API for institutional users.

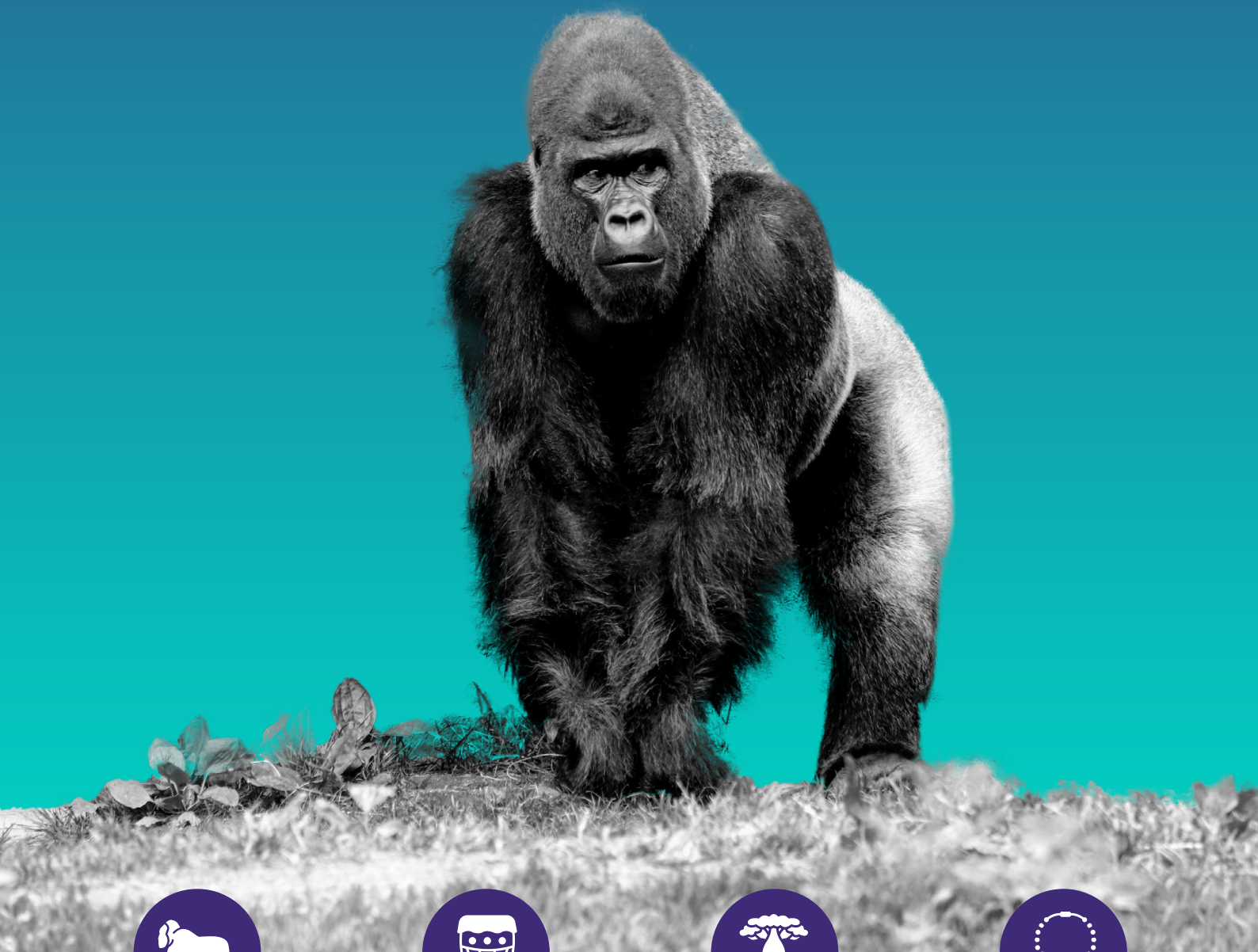
It covers multiple crop types, including maize, beans, sweet potatoes, cassava, kales, Irish potatoes, cowpeas, soya beans, and indigenous vegetables, as well as the following languages: KiEmbu, Kikuyu, Kiswahili and Luhya.

Data was collected over multiple growing seasons from 2022 to 2024 by LDRI in partnership with champion farmers and agronomists.





REPUBLIC OF RWANDA



RWANDA HAS SET ITSELF THE GOAL OF BECOMING ONE OF AFRICA'S **LEADING INNOVATION HUBS**



Known for its strong public sector coordination, future-oriented policymaking, and openness to digital transformation, the country offers fertile ground for exploring how AI can be applied for the public good. At the same time, Rwanda's development realities clearly shaped FAIR Forward's work: a large share of the population depends on agriculture, climate-related risks directly affect livelihoods, and digital tools can only be truly inclusive if they are accessible in the local language and usable through simple technologies.

To strengthen the Rwandan AI ecosystem as a whole, FAIR Forward therefore followed a collaborative approach that convened different actors and capacitated them.

An AI Innovation Lab is a space where the AI community, start-ups, and public institutions could explore new ideas and jointly develop solutions. Alongside this, inclusive and tailored training programmes like the AI Fellowship, the AI for Women Bootcamp, courses for Machine Learning for Earth Observation, and AI training for the public sector helped participants understand AI as a technology, how to use it responsibly, and apply it to real challenges in their professional context. This combined focus on collaboration and skills helped ensure that Rwanda's progress in AI is driven by local talent, practical application and shared knowledge rather than by isolated pilot projects.

Building strong local foundations for AI also meant investing in training data and language technology. FAIR Forward supported the collection of local language data and contributed to the development of the **first text-to-speech model for Kinyarwanda**.

This was a critical step towards making AI systems more accessible, familiar, and relevant to Rwandan users. In a country where digital inclusion depends on connectivity as well as language and usability, these building blocks are essential for ensuring that technological progress can reach a wider population.



These foundations enabled the development of practical AI solutions directly relevant to people's everyday lives. Together with local start-ups, FAIR Forward supported **chatbots** such as Mbaza, which provided **reliable information to millions** during the COVID-19 pandemic, and Tunga, which offers smallholder farmers practical agricultural advice through simple mobile phones.

In Rwanda, where 62.9% of the working-age population is engaged in agriculture, such solutions can make a tangible difference by helping people access information quickly, in their own language, and through channels they already use.

Towards safe, inclusive, and accountable AI innovation – Rwanda's National AI Policy and Rwanda's Responsible AI Office

Rwanda's National AI Policy was developed to advance the country's vision of becoming a leader in responsible and inclusive AI-driven innovation. The policy provides a strategic framework for strengthening national AI skills, improving data ecosystems, and enabling ethical adoption across priority sectors.

In its development, FAIR Forward provided technical expertise and supported multi-stakeholder consultations that shaped and refined the policy. FAIR Forward also contributed to the establishment and operationalisation of the Rwanda Responsible AI Office, a key institution for coordinating and overseeing implementation. Through this support, the policy is positioned to transform Rwanda's AI ambitions into sustainable, achievable outcomes that deliver broad socio-economic benefits.

FAIR Forward also supported the development of AI solutions to strengthen monitoring and early warning capacities in two highly relevant areas for Rwanda. The first was agriculture, where improved data and forecasting can help reduce climate change-induced risks for farmers and improve government services. The second was disaster risk management, especially in relation to landslides, which are the deadliest disaster in Rwanda.

In this way, FAIR Forward in Rwanda showed how AI can be translated into concrete public value: by strengthening local innovation, expanding access to information, and helping address some of the country's most pressing social, economic, and environmental challenges.

Advancing digital public services with Kinyarwanda-language AI – Tunga chatbot

FAIR Forward supports Rwanda in making digital public services more inclusive by enabling AI solutions in Kinyarwanda that address concrete local needs. Tunga is a 24/7 AI voice assistant in Kinyarwanda that supports the National Agriculture Call Centre at the Ministry of Agriculture and Animal Resources (MINICT) by providing instant advice on good farming practices, pest and disease control and, agricultural subsidy programmes, potentially reaching millions of smallholder farmers across Rwanda.

Built on earlier efforts in language AI, such as the Mbaza COVID-19 Chatbot, Tunga proves that reusing open-source AI building blocks can save time and costs, reduce technical risks and make it easier to develop and roll out new, locally relevant AI services at scale.



AFRICA



REPUBLIC OF SOUTH AFRICA



South Africa,
the 'rainbow nation',
is a country whose rich
diversity and resilience
cultivated an ecosystem
that is both complex
and innovative.



FAIR Forward in South Africa builds on this legacy, with initiatives that enhance, celebrate, and reflect inclusion, illuminating and enabling the distinct economic, linguistic, and social realities – making diversity a doorway to opportunity, not an impediment. The goal was to move from fragmented pilots, disjointed events and siloed research into a more coordinated, participatory national conversation.

Take the people to AI or take AI to the people?

An inclusive and contextual AI-powered application comes from a multi-stakeholder and multi-disciplinary initiative. A first-of-its-kind initiative, the **ML4EO-2022** (Machine Learning for Earth Observation) programme was a joint initiative of the Department of Science and Innovation, South African National Space Agency, the Council for Scientific and Industrial Research, Wits University, and Agricultural Research Council of South Africa, supported by FAIR Forward.

The programme was designed for a local audience but received 234 applications globally, with applicants requesting international iterations. A rigorous selection process saw 20 professionals (9 female, 11 male) with backgrounds ranging from remote sensing and ML applications to smart agriculture and actuarial science, alongside PhD and post-doctoral students. Designed to be practical and interactive, the curriculum included topics such as introduction to remote sensing in agriculture, classification-based remote sensing

applications, and machine learning modelling of important agricultural crop parameters, and necessitated a project presentation and research report of the class and field work, applied to a case study.



Expectation on Day 1:

Want to get more knowledge and in-depth in terms of GIS.

Outcome:

This was quite an exciting journey for me. I did not have any background in RS or GIS or EO. My background is agriculture. It was fascinating to see available opportunities – merging the industry in terms of the agriculture industry, RS, and EO – and putting it all together.

Pre and post programme survey compiled with anonymous feedback of participants.



Fieldwork in Limpopo, South Africa



Given the advanced industry sector, coupled with the public sector making strides to catch up and a thriving academic ecosystem, AI policy actions have served a dual purpose: governance as well as cohesion.

The **AI Maturity Assessment Framework (AI MAF)** aimed to assess South Africa's readiness for AI integration, a first of its kind in the country and across the continent. It serves as a foundational tool to understand, measure, and advance national readiness and maturity of AI, with a data-driven approach that identifies current capabilities, opportunities for growth, and gaps in AI integration across sectors.

Unlike global benchmarking tools that prioritise cross-country comparisons, this open-source framework is specifically designed to assess South Africa's unique socio-economic, cultural, and technological landscape. Using the ISO 37004 standardised maturity measurement scale makes the framework globally applicable, whilst keeping it locally relevant. The launch was hailed as 'eye-opening' by ministry partners, who viewed the framework and national assessment outcome as a baseline to benchmark the country's readiness. FAIR Forward's governance support further extended to participating in the development of the National Recommendation towards AI Strategy, with both activities providing a foundation for the draft national AI Policy, soon to be presented to Parliament for review and gazetting.

Despite progressive gender laws, only 25% of all professionals in the field of AI in South Africa are women. FAIR Forward carefully conceptualised and curated the **AI & Data Science Bootcamp 4 Women**, a pilot aiming to dismantle harmful stereotypes of women and minority groups in the field of AI. The unanimous sentiments expressed – "Flexible, accommodating, community" –

embodied in the project, enabled active mothers to fully and meaningfully participate, truly leaving no one behind. This Bootcamp also laid the foundation for the free e-learning MOOC now available on the webpage of the National Electronic Media of South Africa (NEMISA). 60 participants started, with 38 graduating (36 females and 2 non-binary individuals).

Another first for FAIR Forward in South Africa was having the country's local languages represented on Mozilla's Common Voice Platform, through the Voices of Mzansi project – part of UNESCO's 'Decade of Indigenous Languages' project. This facilitated text datasets in the 10 local vernacular languages collected, collated and uploaded.

South Africans are now able to join the international common voice community in collecting and sharing speech data, joining initiatives such as Masakhane, Lelapa AI and IndabaX to advance African language AI for local innovation.





REPUBLIC OF UGANDA



Uganda is a country whose abundant natural resources, youthful population, and cultural diversity provide a strong foundation for sustainable development, including through AI.

FAIR Forward's work in Uganda has focused on advancing inclusion and addressing the "AI part" in the country's critical development priorities – particularly in agriculture and renewable energy, under the rural electrification strategy.

With rising demand for clean power, improved agricultural productivity, and climate-resilient livelihoods, Uganda is at a pivotal moment. In this context, FAIR Forward has concentrated on developing replicable AI use cases with the potential to enhance the quality of life across the country.



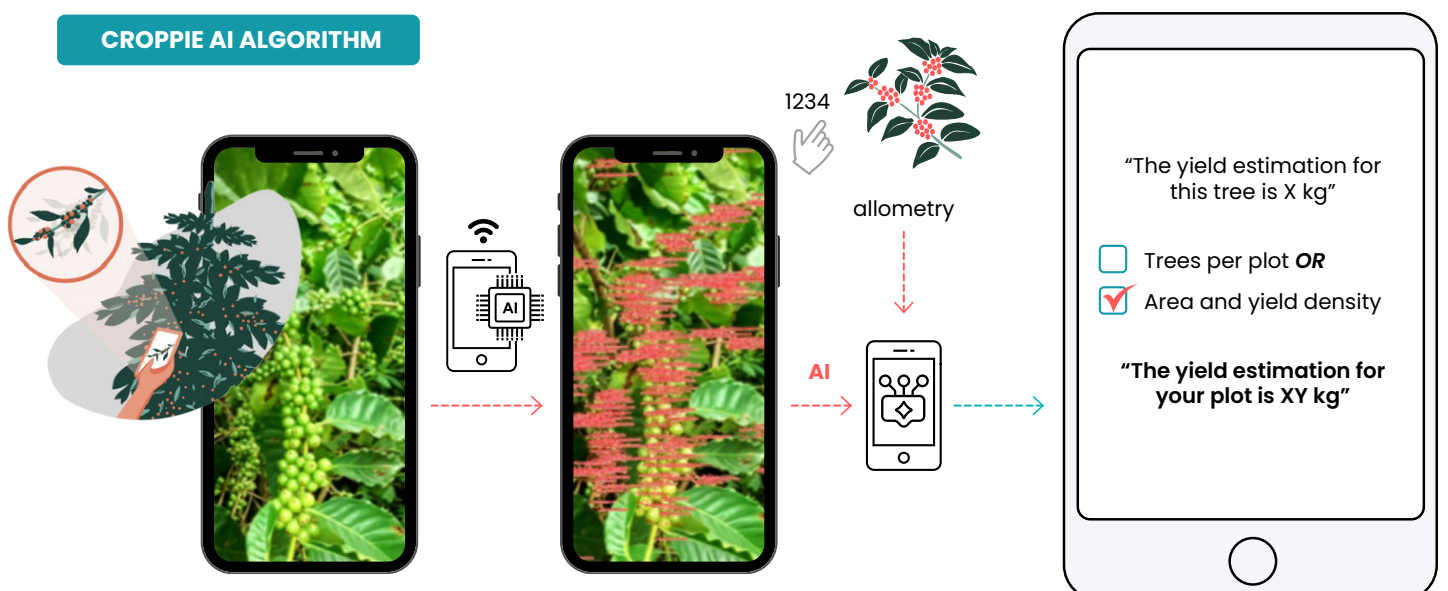
AI in Coffee – Croppie

Contributing almost a third of foreign export earnings, coffee is one of the main cash crops in Uganda. Presently, many smallholder farmers, who rely on subsistence coffee farming for their livelihoods, face challenges such as unpredictable weather and volatile market prices. Traditionally, they have relied on manual yield estimates, but these are time-intensive to perform and not always reliable, leaving many farmers unable to safely plan their harvests or manage their resources effectively.

Supported through **FAIR Forward**, **Producers Direct**, an NGO run by smallholder farmers, the **International Centre for Tropical Agriculture (CIAT)**, in alliance with **Bioversity International** and **M-Omulimisa**

developed '**Crop Yield Predictions for smallholders**', shortened to Croppie, a tool powered by AI that helps coffee farmers to predict their coffee yields. A total of 359 farmers contributed more than 1 121 images, helping to build a representative dataset for the Ugandan coffee sector.

The tool harnesses AI technology to give farmers real-time insights, offering personalised yield predictions and agronomic advice. In that way, Croppie helps farmers to tackle climate challenges and improve their practices. Beyond agronomy, Croppie is also exploring how AI-generated data can help farmers access financial services – such as loans and insurance – by using yield data as evidence of future productivity.



How open data and AI help to find the best energy sources for villages in Uganda

Only 27% of Uganda's population has access to any kind of electricity; the number drops to 10% in rural areas. One promising remedy lies in making better use of Uganda's existing but often untapped renewable energy sources such as solar, wind, and biomass. To pave the way for a just energy transition, FAIR Forward partnered up with Sunbird AI and the Ugandan Ministry of Energy and Mineral Development to support the implementation of the National Electrification Strategy (NES) that aims at connecting 10 million households to the most optimal energy source by 2030.

This partnership produced the Lamwo Electrification App, a site identification tool that uses satellite imagery and various other data sources to provide recommendations for renewable energy sites in Lamwo district. This App helps public officials to better plan and

implement the optimal energy supply in rural areas so that people can receive the reliable electricity access they need for lighting their homes and running their businesses.

The solar-powered mini-grids have so far been installed in 25 villages across Lamwo. These mini grids produce between 20 – 80 kWh worth of energy, serving between 50 – 200 households according to a village's needs.

“

I am able to continue serving my customers till as late as 10 pm.
THIS EXTENSION HAS DOUBLED MY DAILY INCOME!

Nancy Amito

A small restaurant owner in Palabekogilli Trading Centre, a largely refugee-populated village in Lamwo district in Northern Uganda

Moreover, to foster a self-reinforcing cycle of innovation and the ethical development and use of AI, FAIR Forward prioritised equipping all parts of the ecosystem, from frontline practitioners and historically marginalised communities to policymakers with strong, responsible AI capabilities.

By the time this report was written, 181 practitioners and multipliers in Uganda had been trained.

In particular, the country's youthful population benefited from two programmes: an AI Skills Development programme open to all participants, and an AI Bootcamp for Women, which focused on building AI skills amongst women.





09 THE HOW (PART 1)

Key approaches to democratise AI by and for all	61
Solving the data-conundrum globally - empowering local communities to solve real-world challenges in agriculture, climate, and language with open-source AI datasets	62
How to build your own open-source AI with reusable building blocks, datasets and AI models	64

Democratising AI requires understanding the ecosystem, actively closing critical data gaps, and creating open AI 'building blocks' for reuse. Open-sourcing knowledge and content is not without its challenges – transparency is key – and “The how” discusses key challenges, caveats, and proposed solutions – why to go for open-source and how to do it responsibly.

The Lacuna Fund, as one key partner, has addressed one of AI's biggest data bottlenecks in the global majority of countries: lack of high-quality, openly available datasets built by and for local communities. Many datasets have been curated by a broad coalition to enable concrete solutions in agriculture, climate, health, and low resource languages. Building on that data foundation, FAIR Forward and partners have created globally reusable AI components (models, datasets, and code) for high-value use cases in climate, agriculture, and language.

This chapter contains the open repositories needed to replicate, adapt, or extend proven components as digital public goods, instead of starting from scratch.

Implementation lessons from work across 99 communities across Africa and Asia elaborate on

why open-source AI is an effective path. This path leads to fair markets, shared digital sovereignty, faster capacity building, and better stewardship of public funds, and shows how this translates into benefits for societies, companies, and communities. It also surfaces real-world caveats and solutions: avoiding free-riding through fit for purpose licensing and capacity support, protecting the privacy of marginalised groups, and mitigating abuse with responsible AI assessments.

Short case studies (e.g. India's Bhashini, Rwanda's Mbaza community, and the open-source business model mentorship programme) show how open innovation, resources, local ownership, and cross-sector coalitions convert shared assets into value-adding and sustainable services.

OVERALL, THE “HOW” APPROACH OUTLINES:

- 1 Fund and catalyse open-source high-quality local data for AI;
- 2 Package and publish reusable AI building blocks; and
- 3 Adopt open, responsible practices that anchor innovation in partner ecosystems so that communities closest to the challenges can develop and scale the solutions.

SOLVING THE DATA-CONUNDRUM GLOBALLY – EMPOWERING LOCAL COMMUNITIES TO SOLVE REAL WORLD CHALLENGES IN AGRICULTURE, CLIMATE, AND LANGUAGE WITH OPEN-SOURCE AI DATASETS

The Lacuna Fund is the first global initiative dedicated to closing one of AI's most persistent blind spots: the lack of high-quality, openly available datasets created by and for people in Africa, Asia, and Latin America.

FAIR Forward has supported significant dataset development, along with the Lacuna Fund, enabling researchers, cooperatives, and public institutions to build machine-learning models that reflect their realities.

Farmers now forecast crop diseases with locally trained models. Health workers detect preventable illnesses earlier because data finally mirror the conditions on the ground. Speakers of underrepresented languages can access digital tools in their own language. Climate researchers can track the pressure on forests or energy systems with unprecedented precision. For a full account of key results to key problems, see the chart below.

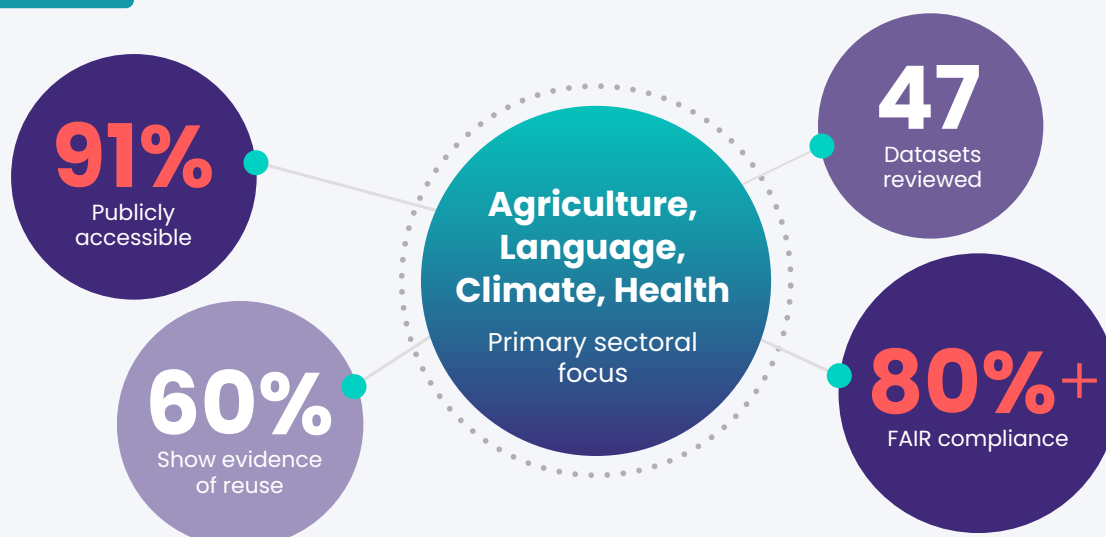
Germany's Federal Ministry of Economic Cooperation and Development (BMZ) played a crucial role as a co-founder of Lacuna and supported **27 datasets in 19 countries** across three rounds of open calls for funding to communities.

This work, implemented by FAIR Forward, focused on three areas where data scarcity was holding back innovation:

- 1 Understanding how **climate change affects forests**, enabling early warning systems and more sustainable land-use planning Improving
- 2 **energy systems and infrastructure**, supporting adaptation and mitigation strategies rooted in local data
- 3 Digital and societal inclusion, particularly through AI for **low-resource languages**, ensuring people can access information and digital services in the languages they speak every day

What emerges from these efforts is a growing ecosystem of locally led AI innovation. The Lacuna Fund shows what becomes possible when data belongs to everyone, and when people closest to the challenges are equipped to shape the solutions. Lacuna-Fund has been supported by a multitude of partners, including the Rockefeller Foundation, Canada's International Development Research Centre, BMZ through FAIR Forward (also as part of Lacuna's steering committee), Google.org, the Wellcome Trust, the Patrick J McGovern Foundation, Robert Wood Johnson Foundation, and the Gordon and Betty Moore Foundation.

RESULTS:



Key results of the Lacuna Fund related to the key problems, and links to key outputs:



LACK OF DATA

OVER 75 new machine learning datasets in agriculture, climate, health, and low-resource languages

89 GRANTS made and administered

SEE:
Catalog of Lacuna-Fund Datasets and Use Cases

GRANTS MADE:

Agriculture	11
NLP	30
Health	18
Climate	30
TOTAL:	89



SCATTERED APPROACHES OF GLOBAL FUNDERS

OVER 20 USD Mio pooled funding

8 FUNDERS:
mix of governments, foundations and private sector/tech

Google.org

GORDON AND BETTY
MOORE
FOUNDATION

IDRC · CRDI

Patrick J McGovern
FOUNDATION

The Rockefeller
Foundation

FAIR FORWARD
Artificial Intelligence for all.

W
wellcome

rwjf robert wood johnson
foundation



NO AGREED QUALITY-STANDARDS

Dataset
Hosting and
Documentation
Guidance

Guidance on
Environment,
Conflict,
and Gender

Intellectual
Property (IP)
Policy

ML Dataset
Sustainability
Toolkit (prepared
by LDRI)

Community
Engagement
Framework for
Machine Learning
Datasets



NO COOPERATION OF MACHINE LEARNING "DATA-PRODUCERS" ACROSS COUNTRIES

Learning networks in key areas such as agric AI, languages AI, health AI etc. (grantee meetings, Online-network via slack, work in the consortia)



BENEFITS OF RESPONSIBLE LOCAL AI OUT OF REACH TO MANY COMMUNITIES

Lacuna Fund Dataset Impact
Assessment Report

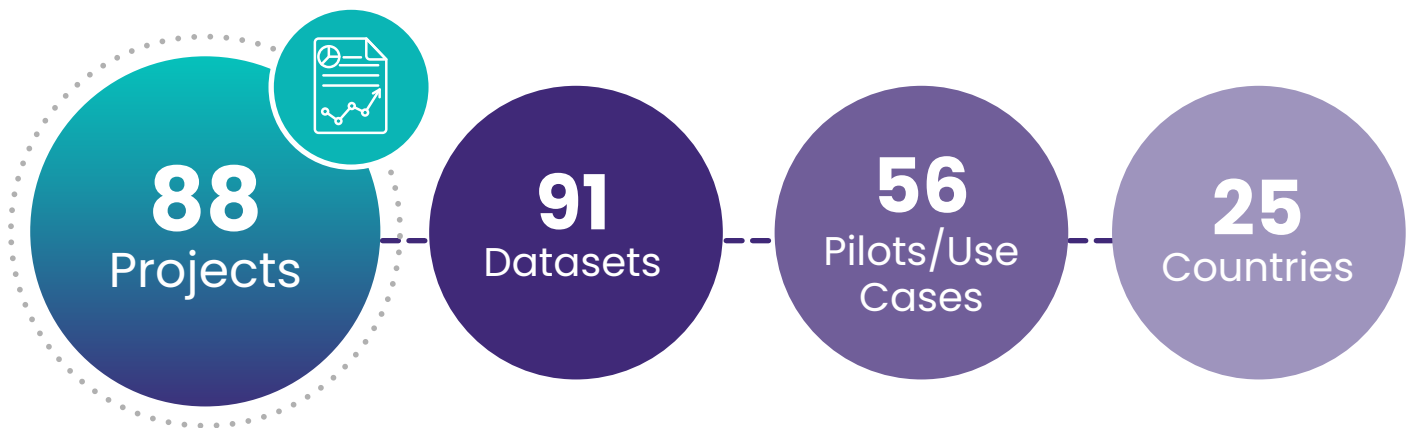
HOW TO BUILD YOUR OWN OPEN-SOURCE AI WITH REUSABLE BUILDING BLOCKS, DATASETS AND AI MODELS

FAIR Forward, with and through its partners, has created globally reusable 'AI building blocks' for the most important AI application areas for climate and agriculture, and language AI.

These datasets, open-source AI models and corresponding documentation work locally and can be shared, adapted and further developed across regions as a digital public good.

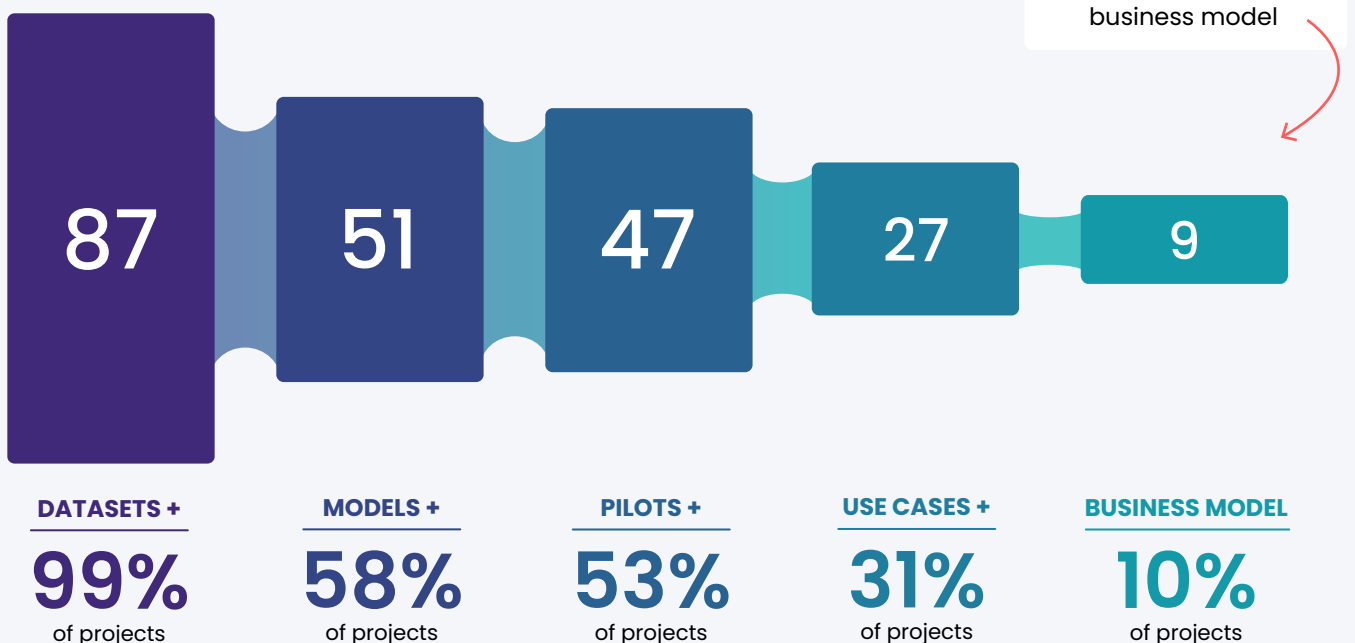
See overview of FAIR Forward work and learn more about the **91 local open-source datasets** and **51 open-source local AI models** ready for replication and building new AI on these foundations

FAIR Forward Catalog of Open AI Data & Use Cases:
<https://fair-forward.github.io/datasets/>



PROJECT MATURITY PIPELINE

How many projects have reached each stage





OPEN-SOURCE AI IN PRACTICE I

Key benefits for society, companies, organisations, and development actors	66
Open-source AI: Distinctive benefits for society	67
Open-source AI: Distinctive benefits for companies, organisations and communities	69
Open-Source AI: Distinctive good practices of durable development cooperation	74

OPEN-SOURCE AI IN PRACTICE I:

Key benefits for society, companies, organisations, and development actors

10

AI is fuelled by data. Inaccurate, incomplete, missing or unavailable high-quality localised training data remains one of the main obstacles for AI innovation in Africa and Asia.

The initiative supports open-sourcing and sharing local training data for more local AI innovation. FAIR Forward has worked with **99 organisations and communities** of researchers, start-ups and civil society organisations from across the seven partner countries who have meticulously collected, verified, and shared data needed for AI to work in local contexts.

Why open-source? At its core, open-source AI is about **promoting fair and responsible sharing of AI that can challenge AI monopolies and systemically enable local innovation and build open AI markets**. Building and open-sourcing AI components should strengthen the common

good and be done in close coordination with local stakeholders, including AI communities, business associations and governments.

Start-ups and companies have a profit motive; freelance AI developers have a curiosity to learn and share; universities have a research interest and public service mandate; civil society has a critical perspective on power distribution and the fundamental rights of citizens; and finally, society as a whole has its right to economic, social and environmental development through sustainable technology – this is why open-source is critical. FAIR Forward's theory of change around open-source innovation and all learnings annexed detail how to make open-source AI work in local contexts.

Open-source AI is good for society, companies and communities, and good development cooperation practice. The following section details the real-world benefit for the three groups.



OPEN-SOURCE AI: DISTINCTIVE BENEFITS FOR SOCIETY

1

MARKETS INSTEAD OF MONOPOLIES

Open-source AI creates open local markets and provides economic opportunities for a multitude of local builders of AI instead of reinforcing AI monopolies

2

SHARED DIGITAL SOVEREIGNTY INSTEAD OF DEPENDENCIES

helps retain more control over AI systems and reduce reliance on closed-source technologies. At the same time, it opens access to shared global public goods and saves resources by collaboratively building on the work of others

3

MORE FAIRNESS INSTEAD OF RISING INEQUALITY

Open-source AI is the route to identify and start tackling inequalities and tweak systems in a way that no one is left behind, e.g. through adding local language capabilities

4

EQUAL OPPORTUNITIES INSTEAD OF PREFERENTIAL TREATMENT

Open-source AI allows for equal access from all companies and communities in a marketplace, instead of subsidies for some

5

RAPID TECHNOLOGY EDUCATION

with open resources makes spreading technical know-how on AI easier

6

AI INNOVATION FOR PUBLIC INTEREST

Open-source AI simplifies steering priorities and ensures that AI innovation will benefit the public interest by better controlling the AI stack in its software and data dimensions

7

BETTER USE OF PUBLIC MONEY

through funding openly accessible AI, which remains available even if the original companies building it cease business operations

8

MORE SECURE AND TRUSTED AI

Open-source AI makes it possible to inspect all parts of an AI system, and collaboratively identify and mitigate security risks

9

MORE TECH ALLIANCES WITH OTHER COUNTRIES INSTEAD OF MORE ISOLATION OR DEPENDENCIES

Open-source AI helps boost international technological cooperation on shared datasets, models and systems ("cooperative" or "shared" AI sovereignty)

Creating markets: Open-source AI for Inclusive Voice Technologies in India

A case of Open-source AI as a distinct benefit for society.

India is a country steeped in rich cultural history, with over **19 000 dialects** of the 121 languages spoken, only 22 of which are official languages.

Trying to access e-government services, healthcare, or find legal information when the internet doesn't speak a person's native language is simply against humanitarian rights.

Enter **Bhashini** – launched by the Indian government as an open-source voice technology initiative that enables digital sovereignty and promotes equality. By making voice technology relatable and accessible to citizens, Bhashini provides a platform for local voice tech solutions in Indian languages and promotes public-interest AI. Researchers, start-ups, and communities can contribute language datasets and AI models as digital public goods, aligning with open-source principles and values on which Bhashini is built.

In return, the contributors can access a shared, local infrastructure enabling speech recognition, translation, and speech generation across all 22 scheduled languages in India. Democratising voice tech is a win for all.

FAIR Forward has contributed to this work, collaborating with the Indian Institute of Science (IISc) in Bangalore, which hosts nine language datasets and AI models on Bhashini. The project focuses on text-to-speech (TTS) AI systems where contributors can generate high-quality voice files with simple text input.

Conversely, speech-to-text data was collected for the same datasets, supported by the Gates Foundation, leading to over 720 hours of high-quality STT data for nine Indian languages. Consequently, Bhashini is able to host 18 language models, offering a male and female option for all. As digital public goods, developers can independently build authentic and culturally appropriate solutions without vendor lock-in.



OPEN-SOURCE AI: DISTINCTIVE BENEFITS FOR COMPANIES, ORGANISATIONS AND COMMUNITIES

FAIR Forward's cooperation with the private sector and communities in Ghana, India, Indonesia, Kenya, Rwanda, South Africa, and Uganda has shown that open-sourcing training data carries significant benefits for companies, institutions, and communities, enabling them to build, train, and fine-tune their own AI models – **echoing broader open-source software paradigms:**

1

LOWER COSTS & SUSTAINABLE PROFIT BY STANDING ON THE SHOULDERS OF GIANTS

by building on a pool of high-quality and high-relevance open-source training data to allow developers to build quality services in AI faster and more cost-efficiently

2

GENERATE REVENUE STREAMS THROUGH OPEN-SOURCE-STYLE BUSINESS MODELS

and/or financial sustainability due to lower costs for data collection (and partly compute)

3

CONTINUOUS IMPROVEMENT AND SHARING OF COSTS

through collaboration in collecting AI training datasets that leads to AI commons, as individuals and institutions across countries contribute to a shared pool of open-source training data and improve it over time – everyone benefits

(SEE: Page 60)

4

DISCOVER AND MITIGATE BIASES AND OTHER RISKS

best achieved through easier identification of biases and malfunctions (e.g. through algorithmic auditing) with open-source, paving the way for adjustments before and after deployment. Open-source AI systems promote transparency, accountability – in contrast to opaque proprietary AI where biases go undetected

5

REDUCED LEGAL RISKS

through open training data which minimises legal liabilities associated with opaque data collection approaches such as web-scraping or unlawful discrimination

6

IMPROVED CYBERSECURITY

when using open AI training data and AI models, as it enables enhanced systematic and crowd-based detection of malicious action and prevention of tampering

Open-source AI sparks business and impact: Mentorship programme for open-source AI entrepreneurs

A case of Open-source AI as a distinct benefit for organisations.

From Kenya to India, entrepreneurs are using AI-powered applications to provide agricultural advice, empower women to enforce rights, and improve public services, making data and models openly available.

How can companies drive local innovation and create social impact whilst remaining economically viable? A **mentorship programme by Villgro Africa and 'FAIR Forward - Artificial Intelligence for All'** has supported them in refining sustainable business models and growing a global network of open-source AI innovators.



Defining business metrics and working with our mentors gave us a clearer picture of the impact of our work on various stakeholders in the cotton value chain, and helped us strengthen our strategy.

J.P. Tripathi

Director of Agriculture Programmes,
Wadhvani AI



Our trustful relationship with farmers and their benefits are close to our hearts. That is why we are carefully considering how to offer premium services that allow us to generate revenue whilst, keeping our services free to access for them. Anything we do, we co-create with them as our partners. That is what responsible AI looks like for us.

Leonida Mutuku

Head of AI development,

Local Development Research Institute (LDRI)

The following resources are available for replication:

- 1 OPEN MENTORSHIP TOOLKIT:**
Practical guide to build a mentorship programme for open-source AI entrepreneurs (incl. links to all training materials). It is designed to aid in creating and running a successful mentorship programme that will support and empower open-source AI entrepreneurs to develop sustainable businesses and funding models for their projects.
- 2 WEBINAR:**
Lessons from the mentorship programme and insights into how open-source AI is transforming the business landscape and driving innovation across industries.



Digital Umuganda and the Mbaza Community – leveraging the power of open-source AI

A case of Open-source AI as a distinct benefit for companies.

The start-up “Digital Umuganda” in Rwanda built training datasets for Kinyarwanda using the Mozilla Common Voice Platform.

Based on this, Digital Umuganda built speech recognition and other AI models. These were further refined by the Rwandan “Mbaza NLP community”, which grew around the initial work of Digital Umuganda and other AI actors in the local ecosystem. Jointly, the community contributed to improving the Kinyarwanda models – basing their work on foundational models and fine-tuning them for enhanced performance in local languages and dialects.

Masakhane and the Masakhane African Languages Hub – Making AI work for African communities and research, by Africans in Africa

A case of Open-source AI as a distinct benefit for communities.

Starting as a grassroots community – focused on African NLP research – **Masakhane** – through their mission of ‘building together’, a common value underpinning African communal ethics – empowers African communities to shape and own the technological advances of AI. Bootstrapping work on African AI and African Natural Language Processing (NLP) provided the foundations, which led to pioneering cross-African cooperation, building and preserving cultural concepts in AI. This ultimately resulted in the founding of the **Masakhane African Languages Hub**. The hub was born through a one-year strategy process led by a community of pan-African NLP experts. This was catalysed by FAIR Forward and by the language group of the AI for **Development Funders collaborative** (Including Gates Foundation, International Development Research Centre, Foreign, Commonwealth and Development Office, and others). The Hub is now the major Africa-led platform focused on addressing the lack of representation of African languages in AI research, data, models, and applications, with Masakhane as a key continental catalyst and community.

Here is how Masakhane’s year-long, Africa-wide collaboration on curating local AI training datasets in NLP has evolved into new research and a new shared pool of open-source training data that is accessible and available to all:

- 1 Collaborative tools such as MasakhaNER have become the de facto benchmark for evaluating Large Language Models (LLMs) on African languages.
- 2 Training data collections such as MakerereNLP in have been critical for speech and text apps in East Africa by allowing the building of speech-to-text AI in languages spoken in East Africa, e.g. for agricultural advice or healthcare.
- 3 The Hub has launched two cycles of Africa-wide grants to advance inclusive, ethical, and open AI for more African languages; and invested in gender-responsive use cases in health and livelihoods.
- 4 Project Lingua Africa was launched: to focus on developing a solution-based approach to dataset curation. Masakhane has established a new approach to strengthening review integrity in African language AI funding through a transparent Technical Reviewer roster.

FAIR Forward has been a long-time collaborator and supporter of Masakhane in its institutional development by supporting the Masakhane team to establish and register the Masakhane Research Foundation. Through this entity, the Foundation has established a track record in grant funding and grant implementation from GIZ and other entities. Through collaboration in collecting AI training datasets, a more robust AI commons develops, as individuals and institutions across countries contribute to a shared pool of open-source training data and improve it over time – everyone can benefit from all collective works.

1 DIGITAL UMUGANDA | RWANDA

Digital Umuganda was founded in 2018 with a focus on building inclusive language technology in African languages. It has been working with FAIR Forward since then on open-source language AI - including datasets and models in Kinyarwanda. They also co-developed the Mbaza Chatbot for COVID-19.

FOR MORE INFORMATION:

<https://digitalumuganda.com/> (and the sub-chapter “Rwanda” above)



GHANANLP | GHANA 2

GhanaNLP builds translation and speech recognition tools for Ghanaian languages, such as Twi and contributes to open-source language tools. GhanaNLP has worked with FAIR Forward on NLP-related work and has contributed to works of the Open For Good Alliance, e.g. article “Open-Source AI Data Sharing: yes! Data Colonialism: no! | by Open for Good Alliance”.

FOR MORE INFORMATION:

<https://ghananlp.org/>

3 KARAAGRO AI | GHANA

KaraAgro is a Ghanaian start-up that leverages AI for supporting food security in Ghana. Their goal is to help farmers increase their crop yields more generally, help countries (with a focus on LMICs) achieve food security through AI-supported precision agriculture solutions. FAIR Forward has worked with KaraAgro on several actions, which are outlined in the sub-chapter “Ghana” above.

FOR MORE INFORMATION:

<https://www.karaagro.com/>



LESAN.AI | ETHIOPIA 4

Lesan works on state-of-the-art Ethiopian Language AI with a focus on translation across the languages Amharic, Tigrinya, Oromo, Somali and English. Lesan contributed to open corpora and benchmarking datasets through [Lacuna-Fund's NLP work](#).

FOR MORE INFORMATION:

[Lesan AI](#).

African and Asian companies that build open-source AI

5

PROSA.AI | INDONESIA

Prosa.ai is an Indonesian start-up that specialises in NLP solutions for a wide range of clients, especially for Indonesian languages. With FAIR Forward, Prosa.ai collected open data and trained models for Balinese, Bugis and Minangkabau.

FOR MORE INFORMATION:

"Discover Indonesia's linguistic diversity: FAIR Forward and Prosa.ai on the road to inclusive AI technology | BMZ Digital.Global" and Prosa.ai.



TECH INNOVATORS NETWORK KENYA (THiNK) | KENYA

6

THiNK is fostering an ecosystem approach to building tech, and tapping into available software developers in Kenya. Their approach promotes an alternative way to showcase local tech talent and promote the use of local solutions. FAIR Forward has worked with THiNK on several NLP and chatbot actions, which are outlined in the sub-chapter "Kenya" and the info-box "AI product Chatbot" in the chapter.

FOR MORE INFORMATION:

<https://think.ke/>

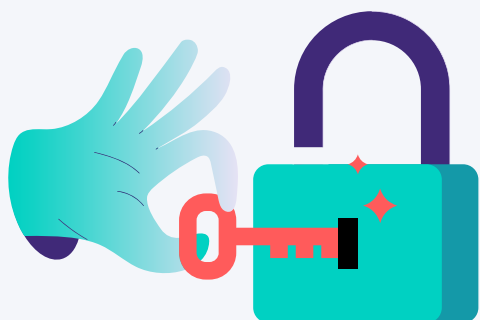
7

KARYA | INDIA

Karya is an India-based company focusing on responsible AI data. FAIR Forward has worked with Karya to create a dataset corpus for a text-to-speech (TTS) system specifically for **Mundari, a low-resource Indian language.**

FOR MORE INFORMATION:

<https://www.karya.in/>



OPEN-SOURCE AI: DISTINCTIVE GOOD PRACTICES OF DURABLE DEVELOPMENT COOPERATION

Better harvests for smallholder farmers, climate justice or better citizen services, i.e. development policy goals, can be most effectively achieved if AI technology is developed and sustained locally by a diverse group of actors. Highlighted below is how the open-source approach could operationalise a key practical challenge of digital development cooperation: sustainable local anchoring and self-empowered continuation after the project's cooperation. Concrete examples of funder initiatives that implement similar good practices are provided.

For more overall good practices, also check the section FAIR Forward **"Principles for implementation of AI in projects"**.

1 CREATE LOCAL AI MARKETS FOR BROAD-BASED ECONOMIC GROWTH, PROSPERITY AND MORE EQUALITY:

For inclusive, broad-based economic development through AI to happen in the global majority, the promotion of start-ups locally needs to be combined with local market creation and enabling systems for AI applications. For this, AI should be a global and regional DPG. DPGs allow developers to collaboratively build on the work of others, enabling small and under-resourced countries to advance on their predecessors' work. To this end, there must be equity for local players to develop AI systems. Non-commercial actors – research, public administration, and NGOs – should have access to AI data, models and infrastructure. With open-source, everyone can build a common public good (or use global DPGs) for commercial or non-profit intents, preventing monopolies. Such ecosystems are self-reinforcing in which open-source AI grows local AI applications and creates added value for societies and broad-based economic development.

2 FROM CONSUMPTION TO CREATION:

Broad-based economic development in the global majority is marred by barriers to accessing data, compute, and tech power hegemonies, which hamper the scaling of localised solutions. Open-source changes the concept of AI from being a technology for consumption to a digital public good asset. Open-sourcing of AI systems democratises access to, inter alia, local start-ups, researchers, and government to previously guarded assets, necessary for developing foundational systems and citizen-centred localised solutions.

3 MEMO TO GLOBAL NORTH:

Be a fair and trusted partner for mutually beneficial AI cooperation and shared sovereignty: Disingenuous actors try to cast AI as a 'race to the top' that can only be won by focusing on proprietary, closed solutions. Foreign AI systems are perceived as potentially creating dependency or an entry door for foreign monopoly actors. Through open-source AI cooperation, a mutually beneficial way of building AI is possible, and leads to affordable AI sovereignty in a win-win team effort instead of a zero-sum game. Open-source and sovereignty are not mutually exclusive concepts. Open-source AI enables the self-determined use and deployment of AI systems without vendor lock-ins or opaque AI models.

Ecosystem partners can meaningfully cooperate for fair and mutually beneficial systems predicated on trust and transparency, whilst still open-sourcing AI.

4 CREATE LASTING LOCAL AI SYSTEMS INSTEAD OF “PILOTS” AND MAKE HAND-OVER POSSIBLE:

Imported or closed-source AI systems are expensive, less locally anchored, and therefore less sustainable. They are often not tailored to the local realities where they are deployed. Open-source, by contrast, fosters local AI innovation and collaboration. It enables four groups to work together:

1. **Local research**
2. **Local companies that can adapt and improve AI**
3. **Civil society and**
4. **State that orchestrates such an “innovation system”**

Open-source AI and open documentation ensure that all parts of an AI system can be easily transferred to partners and used freely without licensing costs or accruing third-party rights. See FAIR Forward’s replication-oriented by **FAIR Forward Catalog of Open AI Data & Use Cases** (see “How to build open-source AI with reusable building blocks, datasets and AI models”).

5 CREATE AI BUILDING BLOCKS ONCE AND THEN RE-USE = EFFICIENT USE OF SCARCE PUBLIC FUNDS:

A growing number of local open-source AI assets are available as open-source building blocks – software, AI training data, and models to use without having to undertake development again. This ensures the rapid and less costly adaptation and uptake of AI systems and data anywhere in the world for the development of local AI applications. In an era of natural resource scarcity and greater scrutiny over the use of public funds, this argues for efficiency and impact.

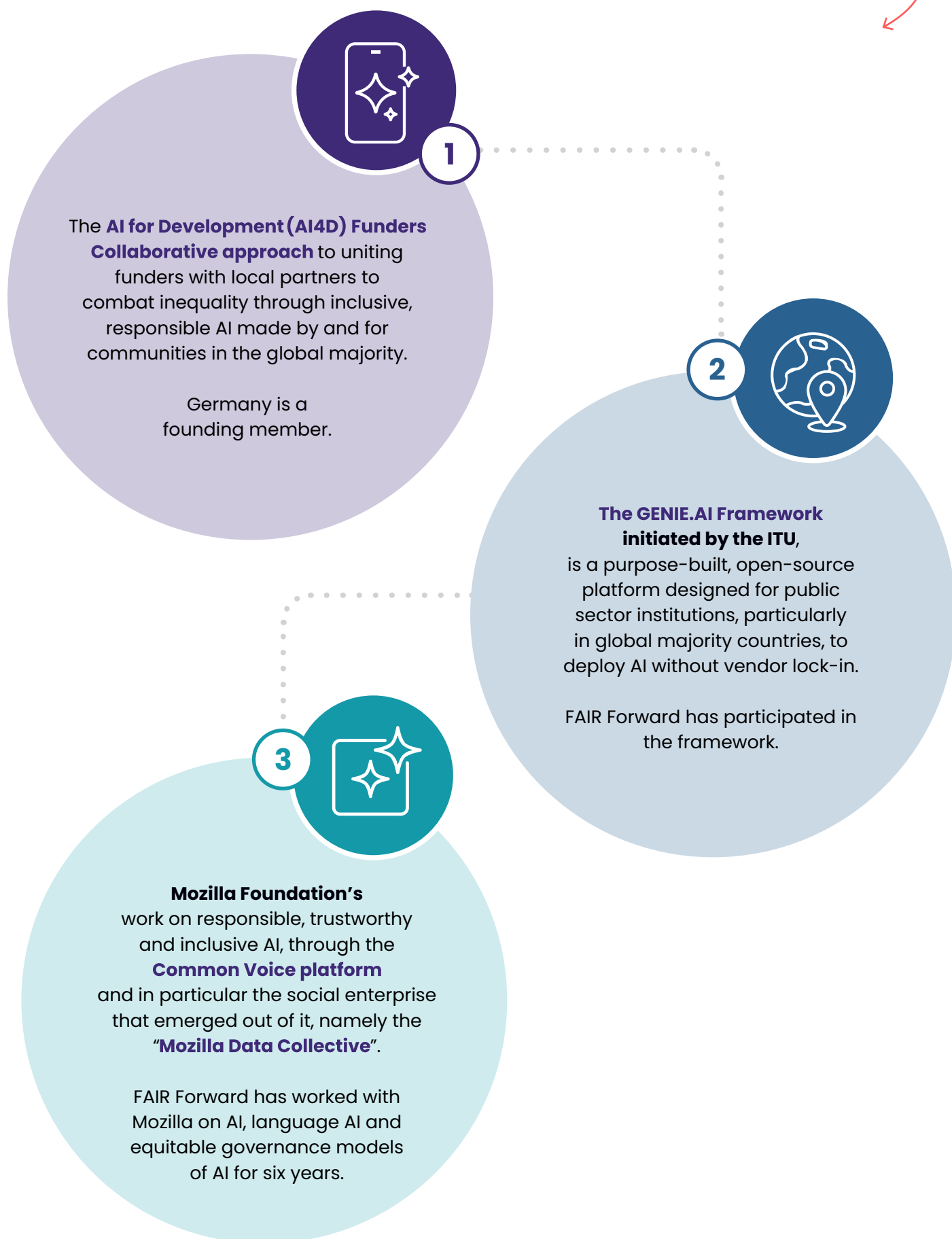
6 AVOID THE EMERGENCE OF UNJUST, DISCRIMINATORY OR EXCLUSIONARY AI:

Open-source ensures transparency and avoids the deployment of ‘black box’ AI systems that discriminate or exclude minoritised groups. Transparency and consequently trusted systems alone are insufficient for responsible and inclusive AI systems. A community approach is needed to guarantee that the voices affected by AI systems are included.

Moreover, accessibility in local languages, and by extension, contextually localised applications are essential for human rights-based and responsible AI systems.



What does this look like in practice? An overview of three concrete approaches of pioneering development cooperation actors that embody inclusive practices around open-source AI.





AI4D Funders Collaborative

ITU's GENIE.AI Framework

Mozilla's work on AI and Data Collective

1

Create local AI markets and ecosystems through open-source

Strong focus on **locally led AI ecosystems** in Africa, Asia, and Latin America; promotes skills, **representative data** and sustainable compute as a **digital public good**, supports research hubs, local start-ups, policy actors, and civil society.

Promotes multi-stakeholder ecosystems, interoperability, and access through standards; **supports participation of SMEs**, governments, and academia.

Funds and convenes data stewards, researchers, and communities to **co-create open datasets** and tools that reflect regional contexts and needs.

2

Fair and trusted cooperation & shared sovereignty through open-source

Built explicitly as a **multi-donor, partnership-based initiative**; emphasises **local leadership**, equity, and **global majority ownership of AI agendas and data**.

Strong focus on **global cooperation**, neutrality, trust, and inclusive governance, enabling **shared AI sovereignty**.

Mozilla Data Collective promotes **open community-led data governance models** that prioritise consent, agency, and shared decision-making, enabling more equitable participation in AI development.

3

Create lasting local systems and enable handover through open-source

Invests in **AI research institutes, training programmes, and policy capacity**; supports **long-term institutional development of open-source communities**, e.g. Masakhane African Languages hub, and not just pilots.

Emphasises **capacity building**, institutional development, and long-term enablement beyond pilots.

Invests in **long-term capacity** by embedding skills and governance frameworks into data collection standards and other tools.

4

Reuse AI building blocks through open-source (efficiency & scaling)

Encourages **knowledge sharing, regional collaboration, and reuse of research outputs**; supports **open data and collaborative research models**

Strong emphasis on **standards, interoperability, modular systems**, and knowledge sharing

Provides a **global platform for reuse and sharing** through its **Mozilla Data Collective** Platform. Supports open datasets, open tooling, and interoperable standards – for reuse and adaptation, reduces duplication and accelerates responsible innovation.

5

Avoid harmful, biased, exclusionary AI through open-source

Strong focus on **responsible AI, gender equality, inclusion, and local language AI**; supports **community-driven approaches and context-aware solutions**.

Strong alignment with **ethical AI**, human rights, inclusivity, and local relevance.

Actively funds and advocates for datasets and practices that address bias, represent marginalised communities, and embed ethical safeguards into data collection.



OPEN-SOURCE AI IN PRACTICE II

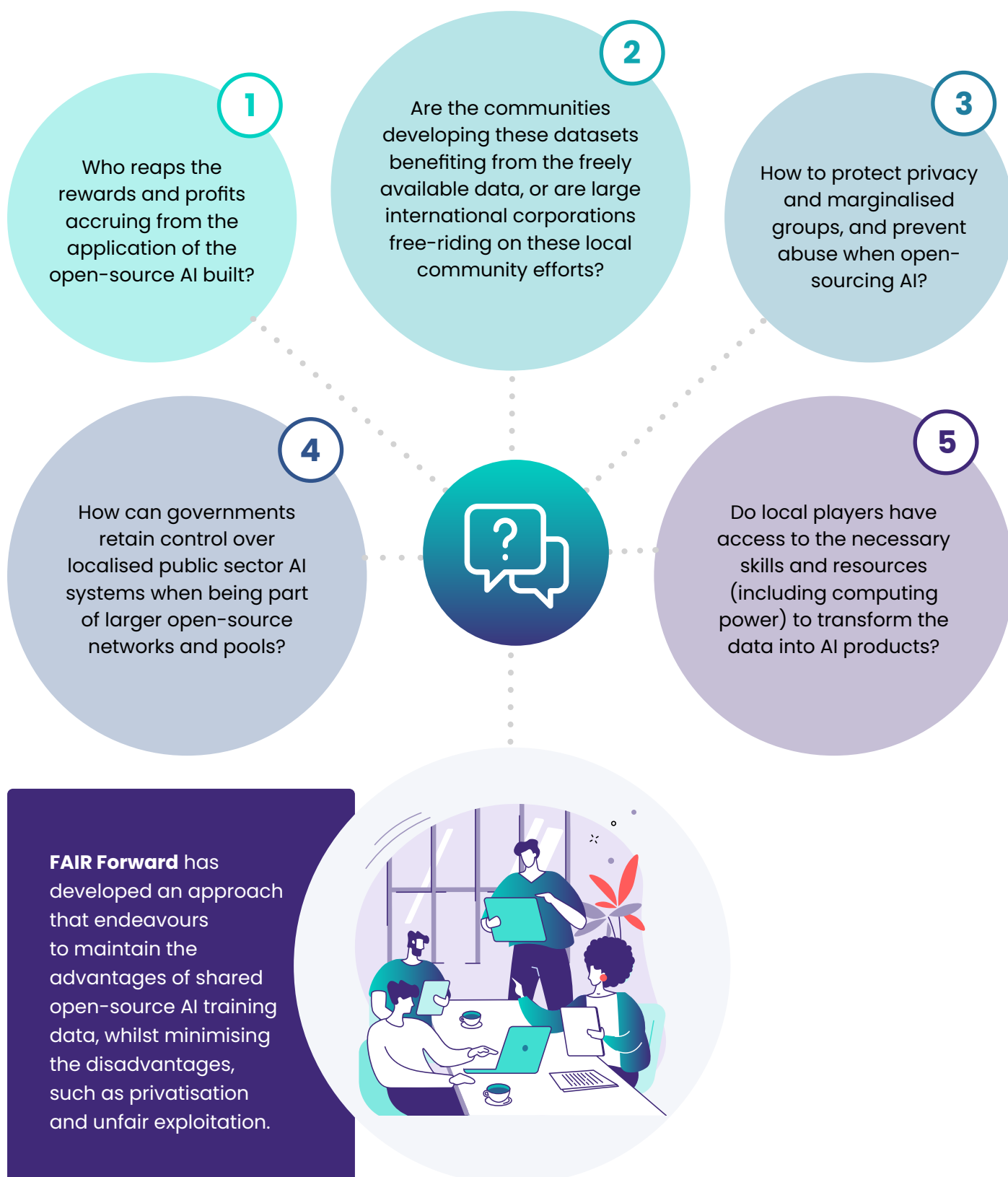
The challenges and solutions	79
Caveat 1 – Avoiding free-riding: Recommendations to communities, governments and large tech companies	80
Caveat 2: Mitigating risks in the area of privacy, protection of marginalised groups and potential for abuse	84

OPEN-SOURCE AI IN PRACTICE II:

The challenges and solutions

11

During FAIR Forward's journey with 99 organisations and communities in building local and global open-source AI, these communities and partner governments raised some thorny questions around digital sovereignty and free-riding that come with going open-source.



CAVEAT 1 – AVOIDING FREE-RIDING:

RECOMMENDATIONS TO COMMUNITIES, GOVERNMENTS AND LARGE TECH COMPANIES

Unfortunately, the fear of exploitation, free-riding of commons, and unfair competition discourages some communities and industry players from adopting open-source AI. The following are some of the strategies FAIR Forward has found to keep African and Asian training datasets open-sourced whilst protecting them against privatisation and exploitation.

Below are six groups that need to feel that a fair sharing of any (monetary or general) value is happening:



Given the diverse interests of these groups, there is no silver bullet to satisfy all. Instead, multi-stakeholder dialogues are needed to get everyone heard and agree on fair sharing principles and appropriate safeguards.

One tested method is to ensure that anyone who uses any open-source AI resource is also obliged to give back any improved or augmented version of the AI resource to the local and/or global community. Another novel option is licensing frameworks that limit free access and usage to certain categories of entities.

Licensing alone might not be a robust enough response. It fails to acknowledge power differentials and structural resource gaps: the step of transforming data to AI models and products is expensive, both in terms of resources and skills.

This creates a gap where communities create the data, but financially resourceful actors reap the benefits. Further, the multiplication in value through the data training process reverts to these resourceful actors because the communities are not nearly as well-equipped (in capacity or resources) to take part in this data-to-model-to-system transformation step. This common structural barrier creates an ecosystem imbalance, which has adverse downstream consequences.

Copyleft licensing needs to be complemented by frameworks for capacity building within the communities and a pathway for them to access the same resources (compute, skills) that others accessing the data have. The recommendations in the box **“Avoiding free-riding”** were built as part of the **Open for Good Alliance** with more than 15 institutions from across FAIR Forward’s partner countries in Africa and Asia, and beyond.



SPOTLIGHT:

Open-source, data governance and licensing – An African conversation

Open innovation is not without complexities and consequences. Making AI training data openly available for open-source uses and development illuminates the boundaries and considerations of sustainability of the data pool, using representative datasets, legalities of mixing differently licensed data and using synthetic data, amongst others. FAIR Forward therefore joined forces with like-minded partners in the **Open for Good** Alliance, a global group of institutions that are focused on improving access to localised training data and AI technologies to spur local open innovation.

In 2023, FAIR Forward, through the Open for Good Alliance hosted a two-part, online workshop by and for African experts on Open Data Licensing aimed to tackle issues facing practitioners creating or using open-source AI. The intent was to investigate what practical solutions are available that foster open data use and development, improve access to open data, showcase available licensing options and highlight the risks and advantages of open-source from combined perspectives.

The process and results of the workshop are **presented in a report**, along with quick access **posters** that were developed to better explain open-source benefits to communities and start-ups.

Through the work of dataset development (text and voice) in local languages, such as the **Voices of Mzansi** project (localising South African languages for Mozilla Common Voice) as well as other initiatives of **Lacuna Fund**, it became clear that open means open for all – this includes free-riders.

Engagements with active community participants such as Masakhane Research Foundation and Lelapa AI, have yielded that communities need protection. Through our work with Mozilla Foundation and Maseno University, the limitations of Open Access or Creative Commons Zero (CC0) licensing models were interrogated, specifically the assumption that CC0 licensing is universally beneficial. Empowering the Dhuluo language community in Kenya through curated capacity development, returned agency to the community through governance of their language datasets. The community selected a data governance model which illuminates a new approach to citizen-centred governance and, consequently, human-centred policy making.



The **Nwulite Obodo Open Data Licence** (NOODL), developed by African researchers – M. Omino and C. Okorie in 2024 – was most adaptable to the community's needs. It provides a balance to open access to work, with the right to share future returns emanating from the data use. Years of advocacy together with the Dhuluo project contributed to the development of the Mozilla Data Collective – a data sharing platform that allows retention of ownership and control over use. The **Dholuo Dhonam dataset**, along with many others, is housed on the platform, together with the community-created licensing terms of use. A recent report on the NOODL licence commissioned by Mozilla and written by Open Future provides further insights: **“How Open Licensing is Changing with AI: The NOODL Licence”**.

For more recommendations on how to tackle the tension between openness and free-riding, see this sub-chapter and the opinion piece of the

Open For Good Alliance: **“Open source AI Data Sharing: yes! Data Colonialism: no!”**

The real work starts at the concept phase – how to correctly and respectfully collect language data. Lanfrica and Masakhane collaborated on a grant issued by FAIR Forward to share knowledge on data collection for African NLP.

They produced a practitioner's guide with learnings and recommendations from the language data collection community, as well as a handy guide on IP, FAIR Use and Copyright in Africa and ideal management of data governance for collected datasets.

Avoiding free-riding: Recommendations to communities and governments that plan to provide open-source AI through a community:

- 1** Balance the interests of all groups for a fair sharing of any (monetary or general) value. This can be done via multi-stakeholder dialogue and a “free, prior, and informed consent” process with affected communities. Consider existing data cooperative initiatives, e.g. **Karya** in India, that are focused on returning the proceeds of data collection to the communities.
- 2 If the community or government is worried that their AI commons might be unfairly exploited:**
 - I. Consider “share-alike” approaches to keep the power of community-driven innovation whilst avoiding privatisation of open-source AI resources
 - II. Collect and prioritise any additional needs

and demands that are needed from anyone (re)using developed open-source and/or share-alike data, such as

- Sustainable access to compute to enable developers to fully build an AI model and AI application
- Research and operational support in very specialised domains such as rooting out gender and race bias in datasets, conducting error analysis, etc
- Capacity building and skills development to get from AI data to a running AI system (e.g. AI operations)
- Quantised models that fit the computing or memory constraints of low-resource contexts, or free credits to pre-trained embeddings
- Skills development in business models and sustainability plans that couple open-source (copyleft) AI with revenue streams
- Other tools that might be needed to provide a level playing field

- 3 On licensing, be aware of potential downsides of a “share-alike” approach or of tailored community-centred licences vis-à-vis a fully liberal open-source approach: e.g. less uptake in large-scale solutions that are operated by large companies, challenges to enforce share-alike requirements, limitations to some business models.
- 4 Meaningfully discuss such a governance framework for open-source AI and the economics of open data in AI with those who collect AI data, those who aggregate, curate and vet such data, and small, local institutions that nurture and use such data and the data subjects.
- 5 Converse with communities about the balance between openness, fairness to all parties, principles of trustworthy and responsible AI, free-riding, and any other concerns or fears of local communities. Not everything open is good content, and not everything should be open.

Consider more nuanced approaches to data sharing:

- > Copyleft commons coupled with demands for capacity building and compute support for the community.
 - > Licences with restrictions, e.g. based on geographies or other restrictions, that may balance open access to created datasets with a right to future returns.
 - > Open-source licensing, but vetting of access to the data, e.g. through a data custodian.
- 6 Discuss with communities potential downsides of closing down data / restricting access based on geographies or other criteria: e.g. less future uptake and contributions by others of created data and AI models, less value to society, more legal liabilities or limited identification of biases and malfunctions.

Recommendations to (big) tech and other institutions planning to use open-source AI, which has been developed by a community or under-resourced governments:

To entities utilising open-source AI data, models or other resources which were collected, aggregated, curated and vetted by local communities or under-resourced governments, contribute to a fair sharing of any (monetary or general) value.

Concretely, this means:

- 1 Acknowledge and attribute any use of such data (mandated by almost all open-source licences out there, but should be done in all cases).
- 2 Share back any improved version to the community or government, thereby jointly building a better AI for all parties (mandated by all “share-alike” licences but highly encouraged in all cases).
- 3 Contact and converse with communities to see how a fairer sharing of value can be organised jointly. There might be several options to balance needs and values and benefits (such needs might include access to compute to enable communities to fine-tune models for local contexts as a way to counterbalance resource imbalances in the field or capacity building).
- 4 Discuss a pathway for communities / under-resourced governments to access the same resources (compute, skills) available, as well as discuss frameworks for capacity building or other needs with the community, to create a level playing field and allies for joint work on an AI commons.

CAVEAT 2 – MITIGATING RISKS IN THE AREA OF PRIVACY, PROTECTION OF MARGINALISED GROUPS AND POTENTIAL FOR ABUSE

There are a small number of cases in which the advantages and disadvantages of open-source must be weighed against other concerns, and potentially adverse consequences of open-source have to be mitigated.

The following optimal approaches and recommendations result from the work advanced by Lacuna Fund:

1 PROTECTION OF PRIVACY

AI training datasets or AI systems (often in the health sector) should not be made openly available if this allows conclusions to be drawn about individuals.

SOLUTIONS AND MITIGATION:

Pseudonymisation before publication, grouping people or “interspersing misinformation to protect people”, conducting responsible AI assessments before publication, federated learning approaches, or simply choosing not to open highly sensitive data.



2 INDIGENOUS DATA SOVEREIGNTY

The need of groups to maintain control over their identity, knowledge, or economic value of data. An example of this is the decision of Maori communities in Northern New Zealand to protect their language data with a [customised license reflecting their cultural values](#).

SOLUTIONS AND MITIGATION:

More closed collections of commons that are only shared within specific groups as well as open-source approaches that leave the data sovereignty and economic value of data with the respective data subjects (such as Digital Green’s [FarmStack system](#)), conducting responsible AI assessments before publication; licensing arrangements that create a reciprocal relationship between data provider communities and data users (such as NOODL).

3 POTENTIAL FOR ABUSE

AI training datasets or AI systems (e.g. in urban planning) should not be made openly available if they can be used to locate marginalised individuals or groups, which can be misused for persecution.

SOLUTIONS AND MITIGATION:

Summarising data sets or “interspersing misinformation to protect against localisation” before publication, conducting “responsible AI assessments” before publication.



How “Responsible AI Assessments” were used to mitigate risks and protect marginalised groups

The **Responsible AI Assessments** offer a structured method to identify, assess, and mitigate risks in AI systems, before, during and after deployment – with a focus on use cases in Sub-Saharan Africa and Asia. The concept and guides were co-created by FAIR Forward, Eticas, and a diverse community of AI inclusion experts from Sub-Saharan Africa and Asia-Pacific region. Overall, FAIR Forward has worked with and supported thirteen start-ups and organisations across Sub-Saharan Africa and Asia on applying these principles to AI systems.

Here are examples of how the assessments have proven to become a relatable, implementable way to conduct a risk assessment for responsible AI development, serving evaluators, auditors and developers alike.

1 GHANA

An NLP chatbot was developed to provide an accessible chat interface to help Ghanaian entrepreneurs, small business owners, and other users locate relevant policy information easily. Applying the assessments, questions posed during the scoping included: what sources make up the policy corpus, and how to verify authenticity and versioning? Looking at transparency and explainability, aspects of answer provenance (sources and citations) and tracking confabulations and misinformation were checked. This highlighted risks of outdated or incomplete policy information, uneven performance across languages, and the potential for fabricated or ungrounded responses. Applying the reliability checks, attention was paid to how the system signals its limitations to users and how responsibility is assigned for content updates and error correction.

2 KENYA

GIZ supported the development of a chatbot that responds to questions on data protection and privacy. Applying the assessment enhanced accountability of how users are informed about the system’s role, how errors or misdirection are handled, and how responsibility is shared between technical providers and public institutions.

3 UGANDA

The Croppie AI system is designed to support smallholder coffee farmers in Uganda by estimating crop yields from smartphone images of coffee plants and providing feedback to improve farm management and sustainability. Applying the assessment illuminated the need to expand and diversify training data beyond a limited set of regions to reduce geographic and practice-related bias, including access to smartphones.



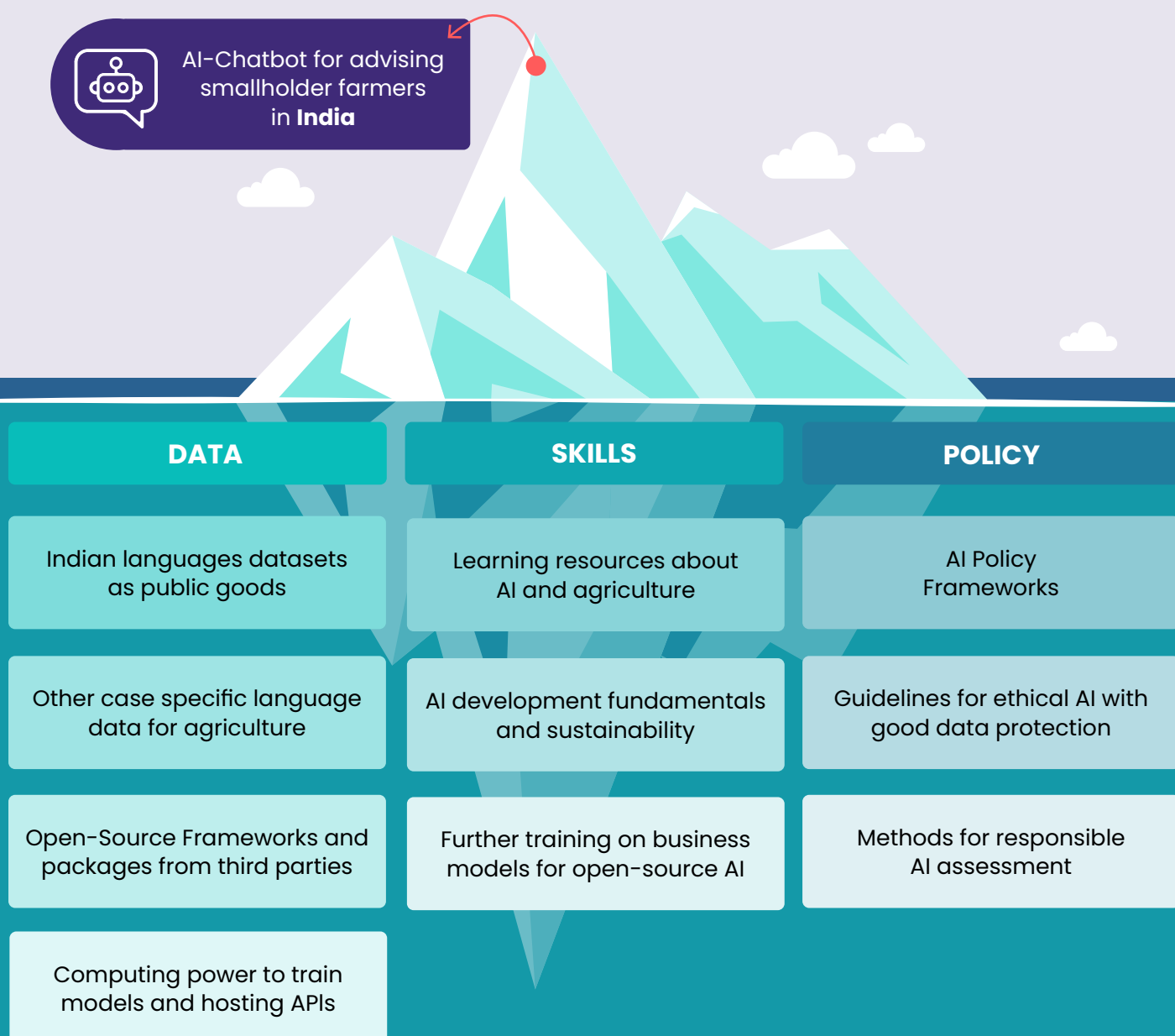
THE HOW (PART II)

Challenges, Lessons, Recommendations	87
Lessons learnt: AI foundations for local innovation	89
Lessons learnt: Strengthen local technical know-how in AI	90
Lessons learnt: Policy frameworks for ethical AI	91

Many of FAIR Forward's implementation experiences have illustrated that what looks like a single AI "use case" in practice depends on a complex set of building blocks below the surface.

Many 'hidden' elements are required, such as, inter alia, high-quality language data, skilled service providers who understand both the local context and willing sectors, AI models and sufficient computing power for training and deployment. These components are complex and costly to build from scratch for each use case, but they can be turned into a key resource if purposely designed for reuse.

ILLUSTRATION OF THE COMMON 'HIDDEN PIECES'



The project's experience has shown that resources are used most efficiently for open and accessible AI when they are directed towards shared assets and locally anchored structures, rather than one-off solutions.

Investments in open-source materials such as language datasets, models, training guides, and mentorship resources can benefit a wide range of actors instead of a single organisation. When working with individual partners, it is important to support activities that create value beyond the immediate project and remain openly available, such as Mozilla's Common Voice or Local Development Research Institute's partially open datasets. Having staff in global majority countries and working closely with local institutions helps to ensure that these efforts are relevant, sustainable,

and rooted in local ecosystems, whilst hosting datasets and models on public platforms like Hugging Face, GitHub, Mozilla Data Collective, or through DPG registries supports reuse and scaling. Simultaneously, the initiative's way of working has offered important lessons on what to keep and what to adjust to going forward. Prioritising local ownership, including diverse stakeholders in decision-making, and being open to learning together with partners has helped make activities more relevant, inclusive, and sustainable.

However, over the course of implementation, it became apparent that applying a systemic lens would have improved results. This lens would consider all the hidden elements illustrated above, intentionally linking them whilst avoiding a narrow focus on isolated stages of the AI life cycle.



Mitigating risks of AI and maximising impact? FAIR Forward's "Principles for implementation of AI in projects"

For development partners working on AI for development or a project incorporating AI into its work, the "FAIR Forward Principles for implementation of AI in projects" may be helpful to mitigate some of the apparent and imperceptible risks of AI in development cooperation.

The following guiding principles help maximise impact and minimise harm:

1. Aligning AI with the Sustainable Development Goals

2. Paying attention to clear added value through AI
3. Developing and using AI responsibly, ethically, and safely
4. Long-term equal opportunities for all
5. Ensuring the long-term use of AI applications
6. Making AI sustainable where it is applied: strengthening local AI markets and ecosystems, avoiding AI monopolies, and supporting digital sovereignty
7. Using open-source AI for innovation and fairness

For more specific recommendations on 'Using open-source AI for innovation and fairness', also see the sub-chapter on "[Open-source AI: Distinctive good practice of durable development cooperation](#)" ([page 63](#))

This document captures key lessons learned from the implementation of FAIR Forward's AI initiatives across various countries and contexts. It brings together practical insights on what worked, what did not, and how approaches

can be improved or adapted. The aim is to support future projects by providing actionable guidance, helping partners design, implement, and scale AI solutions in a more effective, responsible, and sustainable way.

LESSONS LEARNT: AI FOUNDATIONS FOR LOCAL INNOVATION

Datasets, use cases, business models



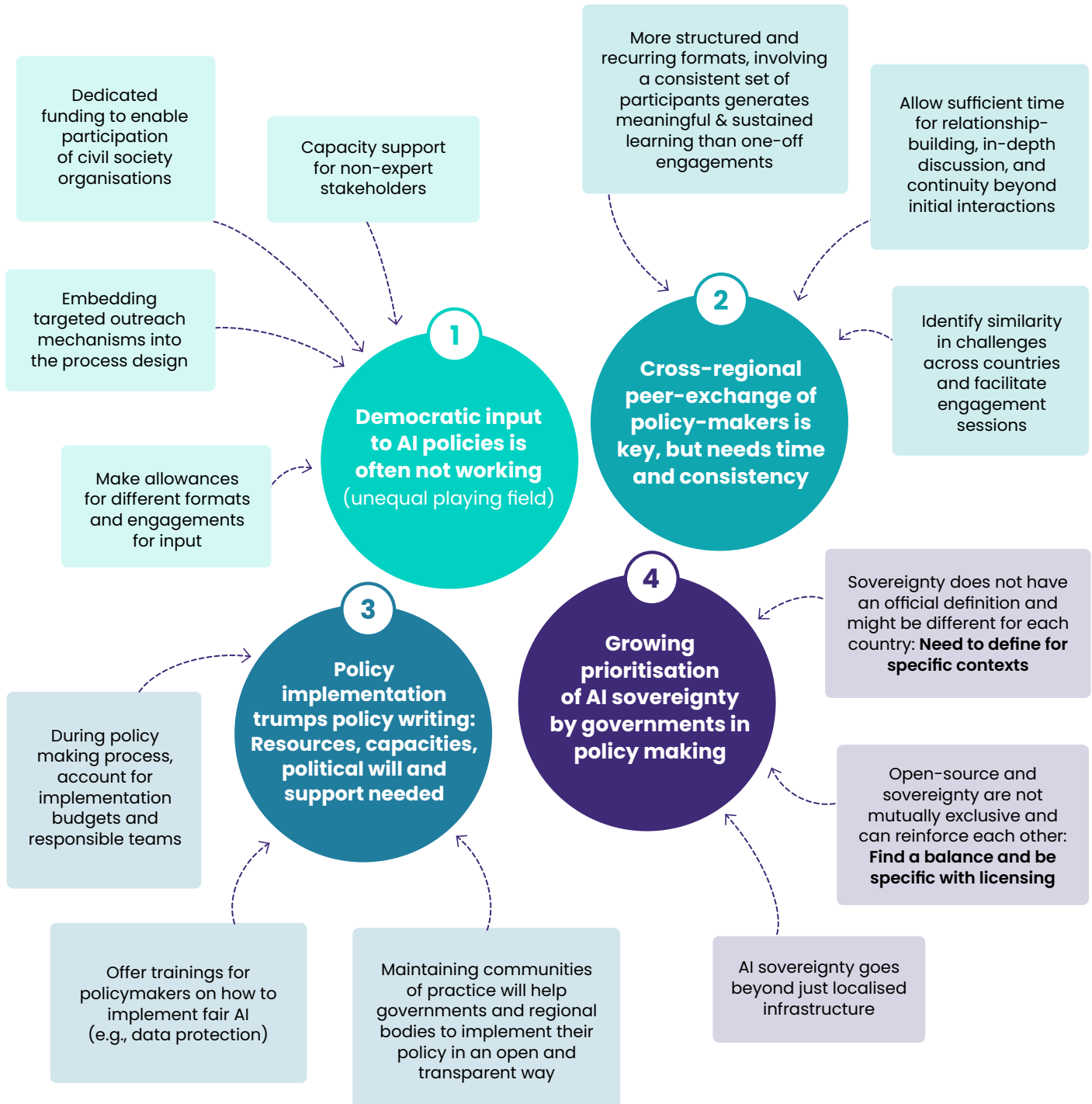
For a deeper understanding of these learnings, refer to **Replication Kit (2): Lessons Learnt**.

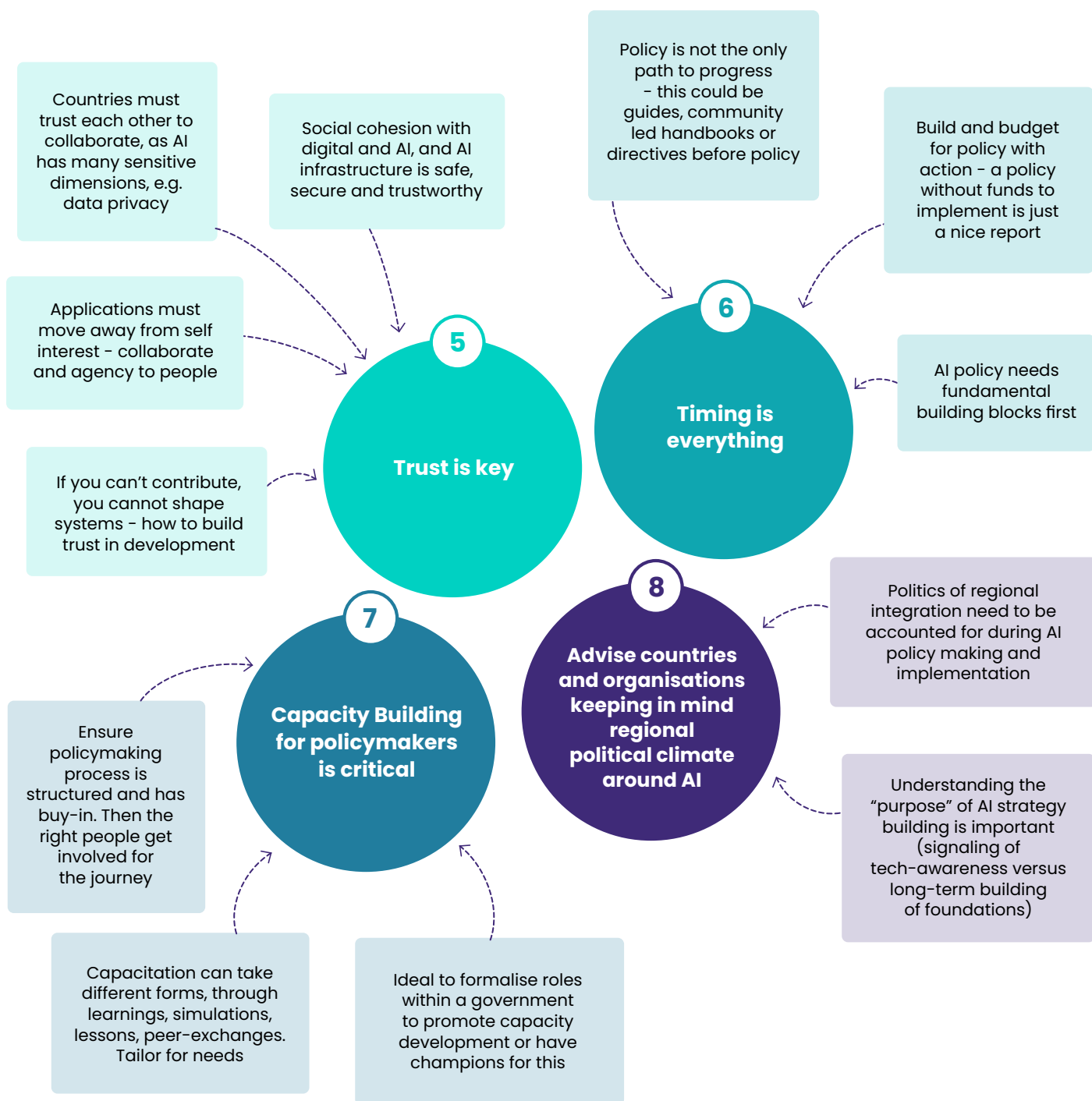
Capacity Building: democratising AI knowledge



For a deeper understanding of these learnings, refer to **Replication Kit (2): Lessons Learnt**.

Deep-dive into learnings





For a deeper understanding of these learnings, refer to **Replication Kit (2): Lessons Learnt.**



OUTLOOK: AI sovereignty: build together. A global agenda

Outlook	94
Conclusion	98

OUTLOOK:

AI sovereignty: built together. A global agenda

13

Current times represent an age of international turmoil. Emerging and frontier technologies are rapidly becoming flashpoints and symbols of geopolitical rivalries. States prioritise reducing dependencies on and increasing control over AI.

Many nations are building “national AI champions” and see themselves in a competitive race for bigger, faster and more powerful AI.

Another world exists. One in which **AI sovereignty of states and citizens is built upon more global teamwork** instead of less. A world where **more democratisation of AI through open-source cooperation** leads to effective, shared AI sovereignty instead of splintering AI silos and growing closed AI monopolies.

This conviction builds upon six years of cooperative work with over 130 Institutions around the globe towards democratising AI and sharing it as open-source resources.

It might seem futile to call for more global teamwork on AI when everyone is retreating to control and competition. **AI technology itself might promote a more cooperative approach.**

For everyone to effectively benefit from AI-powered technology, it is impossible to start from scratch and build an independent, self-functioning, isolated AI stack. Instead, it is more effective, economical and resource efficient for the vast majority of countries to build a shared AI stack and apply this stack to major societal challenges such as reducing inequalities, mitigating climate change, and solving other humanitarian problems.

Even regional associations of states would struggle to own and manage all components of AI. Hence, AI isolation and proprietary solutions will not deliver in the manner it claims, as argued by the [Digital Public Good Alliance here](#), the [German Government/BMZ in its AI strategy here](#) and by [GIZ here](#).



Three networks “for the future” that are already working on “shared AI sovereignty and democratisation of AI”, and how FAIR Forward contributed to them:

1 THE DIGITAL PUBLIC GOOD ALLIANCE

brings together countries and organisations to create a thriving global ecosystem for digital public goods. FAIR Forward has co-shaped its work on “AI as a digital public good”, amongst other things, by helping register some key [AI systems as global DPGs](#), by contributing to its [Open Data for Public Interest AI Toolkit](#), mobilising community input to its [recommendations for the DPG Standard Council on updating the DPG Standard](#) for assessing AI systems, and by co-editing the [“Core Considerations for Exploring AI Systems as Digital Public Goods”](#) of the DPGA.

THE “AI FOR DEVELOPMENT FUNDERS COLLABORATIVE”

2

has committed to “supporting multidisciplinary skills, representative data, and sustainable compute as digital public goods for local ecosystems”, amongst its 4 key priorities. Its vision is to “[unite with local partners to combat inequality](#) through inclusive, responsible AI made by and for communities in the global majority.”

FAIR Forward teamed up proudly with the collaborative as founding implementors and shared our lessons on language AI, gender and AI capacity building, amongst others, with BMZ representing Germany as funder.

3 THE “OPEN FOR GOOD ALLIANCE”

is a network of 16 organisations that work on localised open-source AI data as a commons. It produced a series of recommendations on the collection of open-source datasets and applications of AI, worked on the tension between [“Open-source AI Data Sharing: yes! Data Colonialism: no!”](#) and shared successful examples through members-only meetings and podcasts (on agriculture, voice & language tools, Radiant Earth, Makerere AI Lab etc.) and webinars (on open-source AI and data in law and justice systems, open geospatial datasets for machine learning etc). FAIR Forward has been involved as a catalysing and founding member.

As **Mozilla and Open-Source Initiative** have argued [here](#), and as the United Nations has put forward in its Digital Global Compact: Closed systems create dependency; open systems create capacity

Two international declarations “for the future” that are already calling for more “shared AI sovereignty and democratisation of AI”

A

Hamburg Declaration on Responsible AI for the Sustainable Development Goals

More than 50 global leaders jointly endorsed the [Hamburg Declaration on Responsible AI](#) in June 2025. Initiated by BMZ with the UNDP, the initiative is the first global declaration focused specifically on AI in international development. It outlines shared principles and commitments, amongst them to treat “AI systems and associated datasets (publicly available, non-personal, and anonymised) as global digital public goods” and to strengthen “partnerships and foster a globally inclusive AI ecosystem that upholds equity, innovation opportunities, and shared responsibility”. Many of the endorsers have been partners of FAIR Forward, and FAIR Forward also made a significant contribution to several BMZ deliverables.

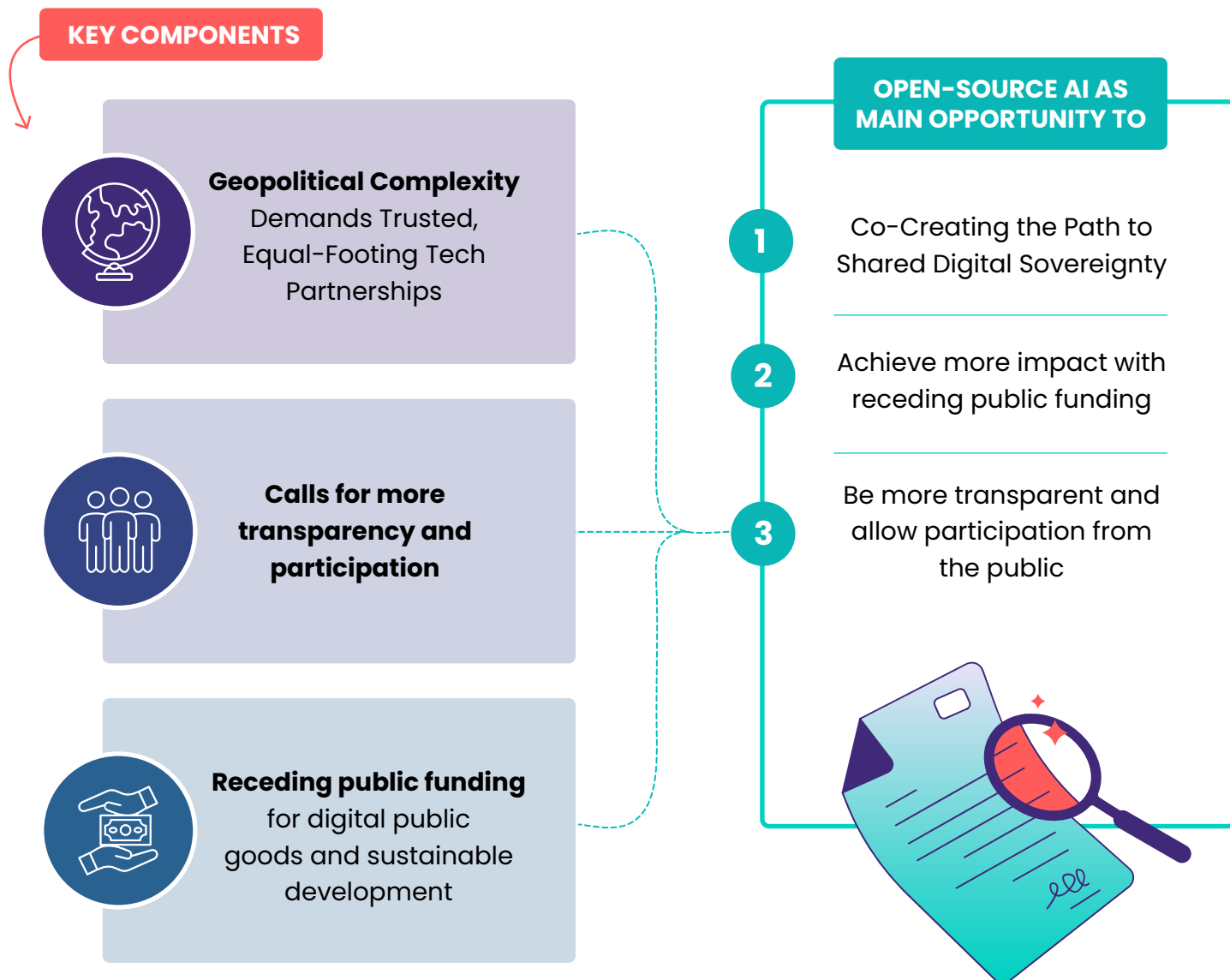
B

The United Nations Digital Global Compact

This was adopted by UN member states in September 2024. The [Global Digital Compact](#) commits everyone to “(a) Develop, disseminate and maintain, through multi-stakeholder cooperation, safe and secure open-source software, open data, open AI models and open standards that benefit society as a whole”. FAIR Forward’s work, especially its 91 local open-source datasets and 50+ open source local AI models, is a contribution to this commitment.

Jointly building open-source AI as a global team is also a necessary practice of development cooperation. In discussions with numerous policymakers and technologists in Africa, Asia, and Latin America, the project has repeatedly heard that “Europe’s main opportunity to offer something unique and valuable to the global majority is to work together on eye level towards effective, shared digital sovereignty.”





FAIR Forward's three pillars of open data, local expertise, and ethical governance were not separate workstreams. They were three angles tackling the same problem: dismantling power hegemonies, to make AI work for all, created by all.

This report has illustrated that open-source AI matters because it cannot be monopolised. Local technical capacity is key because borrowed expertise comes with predetermined priorities. Ethical policy frameworks are paramount because technology without accountability is automating existing inequalities. Together, they lead towards an AI ecosystem that is not just more inclusive, but more robust, inclusive, and human-centred for the global majority.

FAIR Forward has supported 99 communities and institutions that have built open-source AI across national and regional borders.

Based on these grounded experiences, the recent assessment of the Digital Public Good Alliance on AI sovereignty stating that “open-source technologies and digital public goods are a critical part of the puzzle in advancing ational digital sovereignty” – is grounded in reality, supported by project evidence presented in this report.

AI Digital Public Goods provide foundational, open and accessible components and solutions that nations and communities can own, adapt, and control. By doing so, they reduce dependence on proprietary systems, strengthen autonomy, and advance transparency and accountability – essential aspects to building trusted and reliable digital solutions. Moreover, open-source technologies and digital public goods are inherently designed to be built by communities, facilitating broader participation in technology development and promoting democratic, commons-based governance.



The collaborative, cross-continental teamwork approach has made such a complex implementation possible. No actor, however well-resourced, would succeed alone. This is the grounded data that founds the argument that nations, regions, institutions and people need to foster more teamwork underpinned by open-source to achieve shared AI sovereignty and manage the evolving nature of AI.

FAIR Forward closes this chapter not with a finished project, but with a shifted baseline, one in which the ecosystem has been improved and enhanced in the countries and areas of operation.

Democratising AI is an emerging global movement, not a project of any single institution – FAIR Forward considers itself privileged and honoured to have participated, shaped, and contributed to this movement.

FAIR FORWARD'S CALL TO JOINT ACTION:

Let us jointly foster strong, trusted, open-source AI by and for all – for less inequalities, for better public services and good governance, for a stronger and more inclusive economy, and to solve the SDGs whilst enhancing sovereignty and reducing dependencies.



