

AI for Africa : Open-Weight Small LLMs, Sovereignty and Suggestions for AI Ecosystem

Norssken House
Kigali, Rwanda
July 31, 2025

Eun Chang Choi
Partner, IMPACTIVE AI
<https://eunchangchoi.github.io>
techlaw@khu.ac.kr
eunchang.choi@impactive-ai.com

Eun Chang Choi 최은장

Eun Chang Choi, K-장(한국인), is a scholar at the areas of technology, public policy, and information law. His publications investigated (1) Legal and Policy issues associated with Responsible AI (revenue flows, risk management, algorithmic accountability, and business ethics); (2) Global Governance Frameworks for AI Safety (comparative analysis of EU, China and the US); (3) Copyright of AI-generated Content and Fair Learning; (4) Aligning AI Systems with Human Values; and (5) Political Economy of Intellectual Property Law. He has been engaged in global standard-setting dialogues of IEEE Standards Working Group [Algorithmic Bias Considerations] and Data Privacy Privacy [Policy]. He also participates in the Policy Network on AI (PNIA) of the Internet Society.

Africa is rising to join the global wave of AI revolution. Its focal point is here Kigali

In Africa, over 2,400 AI organizations have been operating across various industries, including small size enterprises, health care, wellness, farming, language training, insurance industry and education sector.

Nevertheless, Africa will have to fill the 'AI gaps' between the Global North and Global South.

Africa Declaration on Artificial Intelligence

Global AI Summit on Africa
Kigali, Rwanda
3-4 April 2025

Thank you for making the Global AI Summit on Africa a resounding success!

C4IR Global AI Summit on Africa in Rwanda marked a pivotal moment for the AI landscape by bringing the public and private sectors together to discuss Africa's role in the future of this technology.

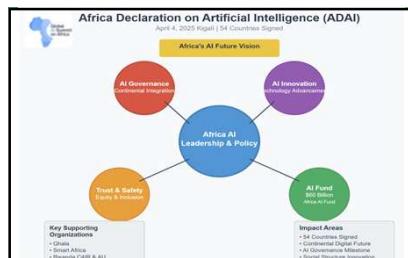
Rwanda Establishes AI Scaling Hub to Drive Responsible AI Development Across Africa

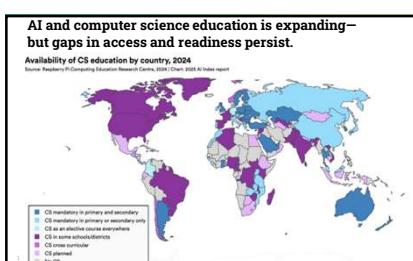
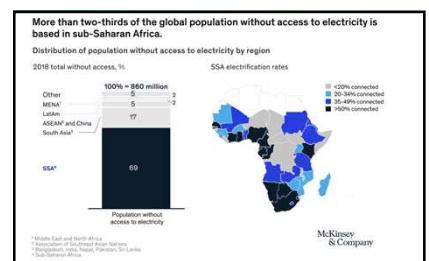
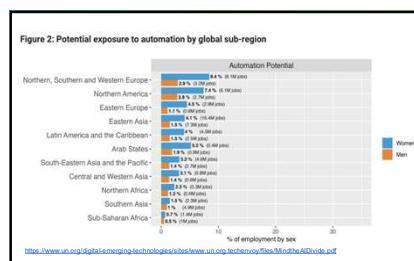
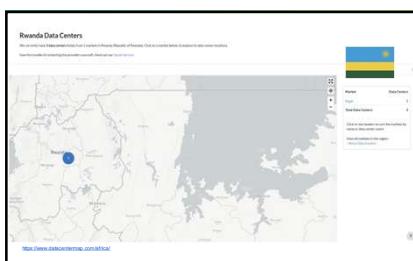
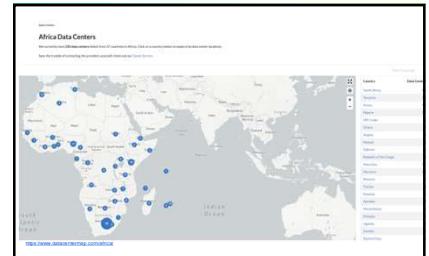
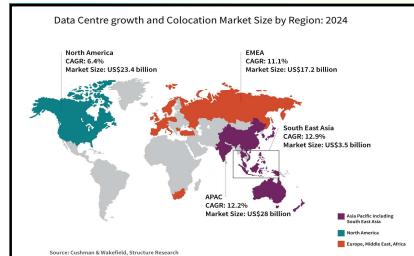
At the Global AI Summit, Rwanda and the Gates Foundation signed an MoU to create the Rwanda AI Scaling Hub, backed by \$100 million, aiming to accelerate responsible AI development and scale solutions in health-care, agriculture, and education across Africa.

AI Revolution in Africa: Challenges

- The AI Divide**: structural gaps between the Global North and Global South
- Insufficient Internet Connectivity**: AI applications, especially cloud-based services, require robust and reliable internet connectivity
- Unequal Computing Power and Data Centers**: Advanced data centers (AI Factory) needed to crunch massive data required to train large language models
- Data Availability**: AI models are data-hungry and require vast amounts of quality data for training. A shortage of large, quality, public dataset to train models.
- Data Asymmetries**: Most high-quality, labeled datasets reflect Northern languages, cultures, and contexts, leading to biased against Global Southern users.
- AI Talents + AI/ML Education + AI Literacy**: More computer science, data science, machine learning training needed to educate more students
- Subscription Wall of Proprietary Generative AI**: Advance general LLMs/ Gen AI commercial tools ask subscription→ prohibitively expensive to ordinary individuals.

"How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?"
"What is the missing element in the National AI Policy?"





Key Takeaway 1 “Fine-tune Open-Weight LLMs to Democratize AI technology and Unlock the potentials of AI”

Key Takeaway 2 “Better Focus on Small LLMs ,Vertical AI and ML Tools rather than relying on proprietary Gen AI /LLMs”

Key Takeaway 3 “Apply AI/ ML Technology for Public Health of which Rwanda focuses”

Key Takeaway 4 “Empowering African Languages through Natural Language Process – Kinyarwanda”

Key Takeaway 5 : “Voice Data for Text to Speech (TTS), Speech to Text (STT) and Chatbot Dialogue”

Key Takeaway 6 “Data Analytics, ML + DL can help Risk Mapping to tackle Soil Erosion , Land Sliding problem, which caused macroeconomic cost ”

Key Takeaway 7 “Leverage NLP/ ML/DL tools for SDGs Project”

Key Takeaway 8 “Build openly-available Public Data Hub or Language Corpus Dataset Repository that anyone can download and use without pay”

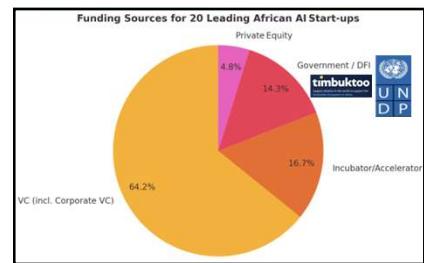
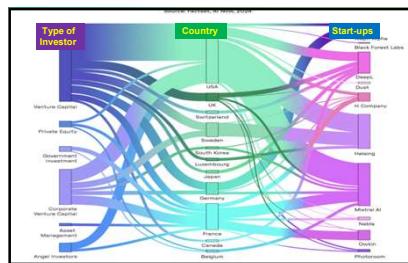
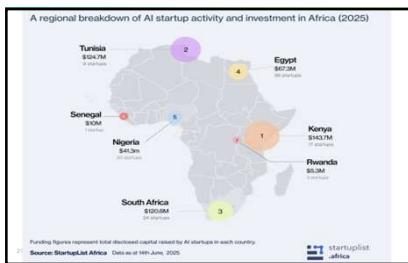
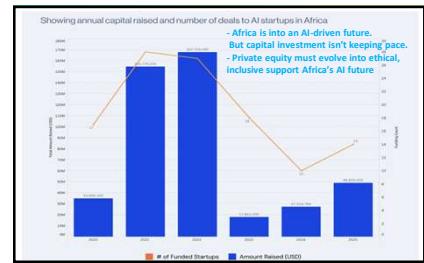
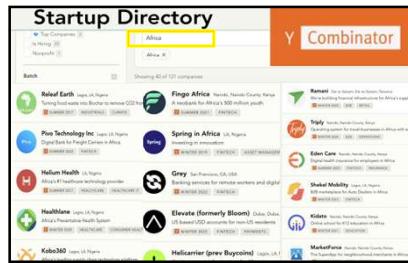
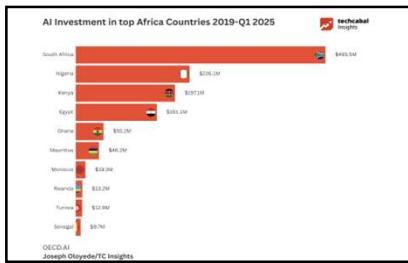
Key Takeaway 9 “Prepare Sovereign AI in Rwanda and Data Sovereignty in Kinyarwanda”

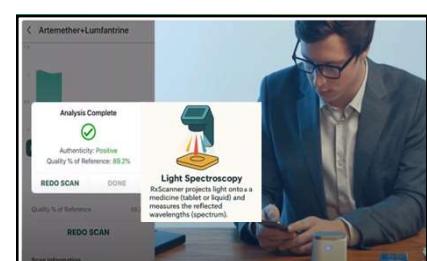
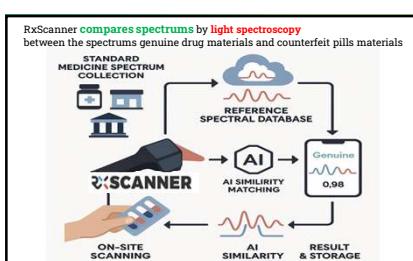
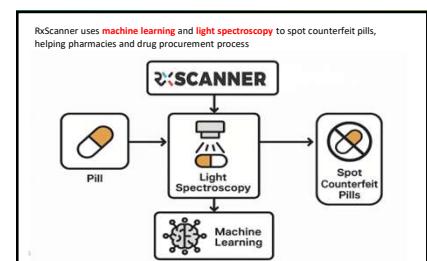
Key Takeaway 10
“Startups need more funding and investment and better focus on solving chronic regional problem”

Key Takeaway 11 “Online Free Education Platform in Kinyarwanda to learn Machine Learning, Fine-Tuning an AI Literacy”

Key Takeaway 1

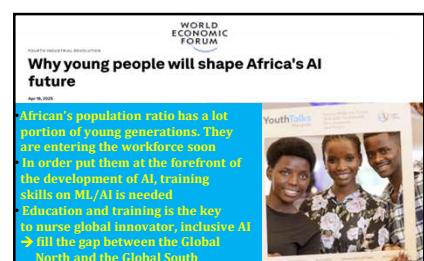
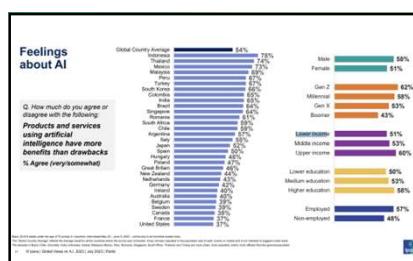
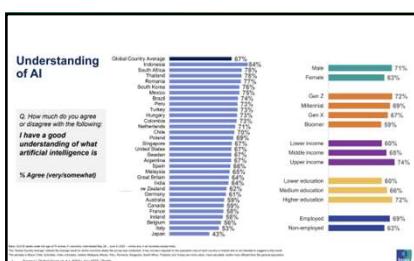
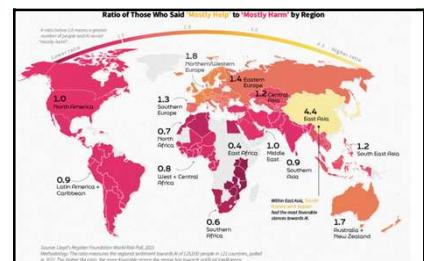
“Africa’s Startups are in a need of more funding and investment and better focus on solving chronic regional problem in Africa”







Key Takeaway 2
"Launch Rwanda's Nation-wide Online Free Education Platform in Kinyarwanda to learn Machine Learning, Data Science Basics and enhance AI Literacy of ordinary citizens"

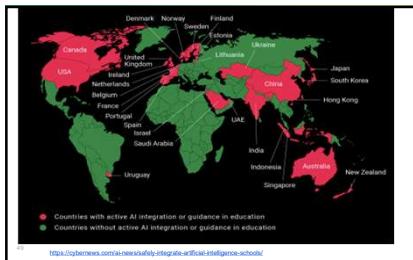
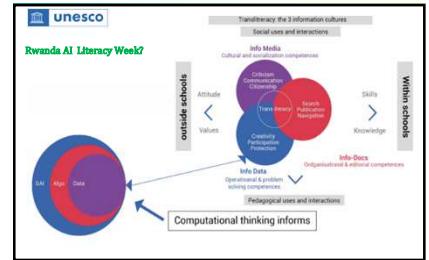
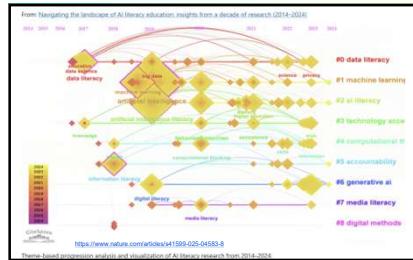


IBM
High quality content should be provided in Kinyarwanda
English language can present accessibility limitations

AI literacy: Closing the artificial intelligence skills gap

Ensuring that students and citizens acquire knowledge and practices that need to harness the potential of AI effectively and safely.

Public Awareness of Artificial Intelligence in Everyday Activities



Rwanda Basic Education Board

ICT in education Department

Ministry of the ICT in Education Department is to harness the innovative and fast-growing potential of information and communication technologies and resources for knowledge creation and sharing and improving 21st century learning skills.

Internet Connectivity
2117 schools are connected to Internet

Smart classrooms
164 primary schools received computers/tables and 1119 Secondary Schools have got Smart Classrooms

Capacity Development of Teachers
4309 primary school teachers and 1263 Pre-service primary school teachers have received digital skills literacy and certified as ME Smart Innovative Educator. 2999 primary school teachers completed Unesco CPD modules

Digital Content
All subject available in digital format on e-learning platform. Testing of chemistry senior 2 virtual lab aligned to CXC

SMART IBIRUHUKO

The Smart Ibiruhuko initiative aims to broaden access to digital skills training through the existing network of smart classrooms and relevant technology-enabled spaces across Rwanda.

[Learn More](#)

UDACITY

SMART IBIRUHUKO 2025
- \$125

To access to digital skills for Rwandans living in rural school buildings.

INTRODUCING SMART IBIRUHUKO 25
Create online training opportunities and digital skills

JOIN #30YEARS ANYWHERE IN THE COUNTRY!

MEET & GREET
It's A Bold Movement Bridging The Digital Gap Across Rwanda. It's A Celebration Of Every One Shaping Our Future And Inspiring The Future.

SMART IBIRUHUKO

Nation-wide Online Education Platform in Kinyarwanda

- Attending master program is not easy option for local residents.
- High quality new content can overcome low quality of teaching.
- Content in English would present accessibility limitations.

Machine learning online courses and programs

You don't have to lose your job to a machine learning AI. Explore machine learning courses that can help you build the machine you need to launch or advance your career.

[What Will You Learn?](#) [How To Get Started](#) [How To Choose](#) [FAQ](#) [Create an account](#)

MIT Professional Education

Serving technical professionals globally for over years. Learn more about us.

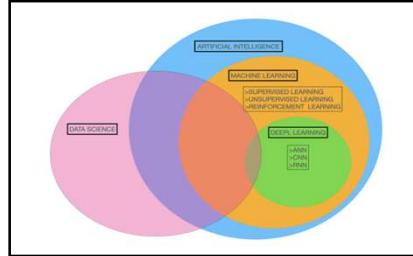
COURSES FOR INDIVIDUALS FOR ORGANIZATIONS

Professional Certificate Program in Machine Learning & Artificial Intelligence

BACK TO COURSE CATALOG

Machine Learning & Artificial Intelligence

Virtual OPEN HOUSE 12:00-1:00 PM EST



Online Education Content 1

Free Python Course For Beginners

Online Education Content 2

Machine Learning Model Process

Figure 6: Value chain of AI

Machine Learning Model Process

1 Data collection → 2 Data preparation & annotation → 3 Data storage → 4 Model design → 5 Model training & tuning → 6 Deployment → 7 Monitoring & maintenance

Prepare Data

Data Ingestion → Data Preparation (Normalization, Transformation, Validation, Deduplication)

Build & Train Models

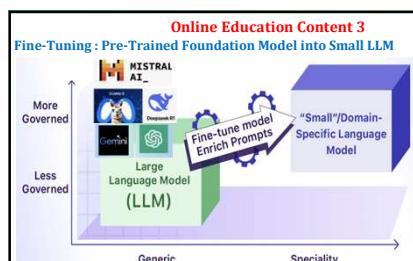
Model Building & Training (Hyper parameter tuning, Model architecture selection, Automatic model selection, Model validation, Model evaluation)

Deploy & Predict

Model Deployment (Deployment, Batch scoring)

There is no "One-Size-Fits-All" ML solution

A Generalized Machine Learning Pipeline



AI STUDIOS

Create Content That Connects on a Human Level

Get Inspired by lifelike, natural voices made with digital avatars in AI Studios.

AI Avatars

Use your voice & AI projects a natural, human touch with generative digital avatars that move and speak naturally

AI STUDIOS

DEEPRAIN AI

AI Video Generator

AI Voices Explore AI Voices in 100+ Languages

How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

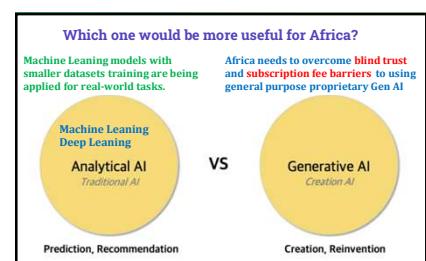
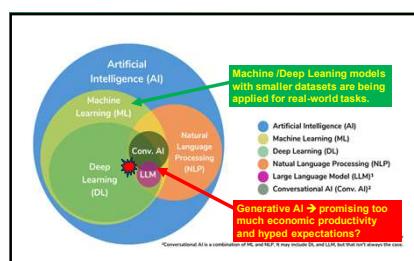
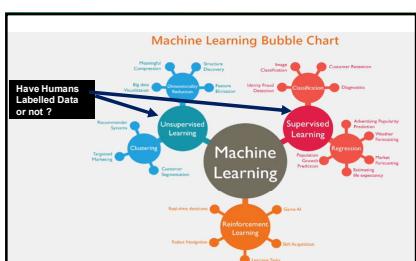
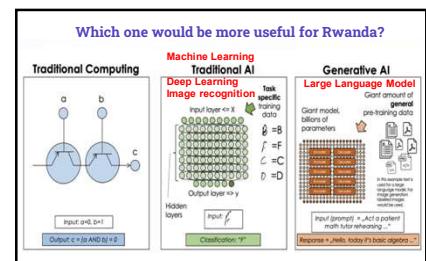
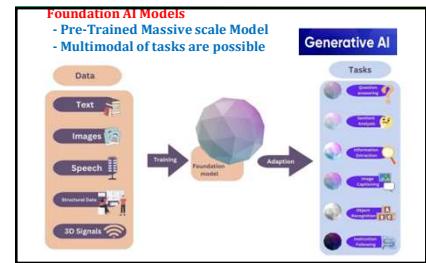
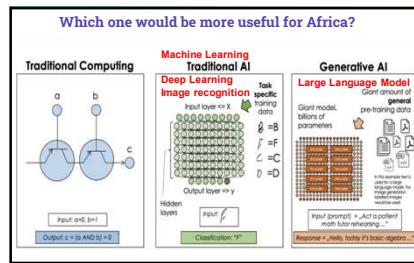
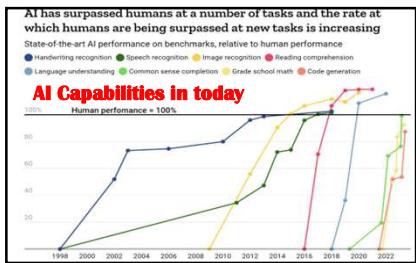
Key Takeaway 3.

"Fine-tune Open-Weight Foundation Models to Democratize and Unlock the potentials of AI Technology for Rwanda"

The National AI Policy

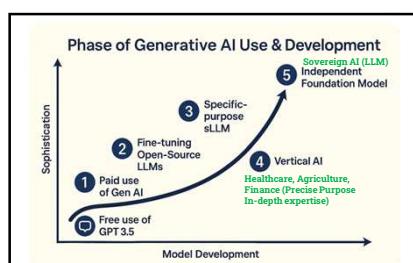
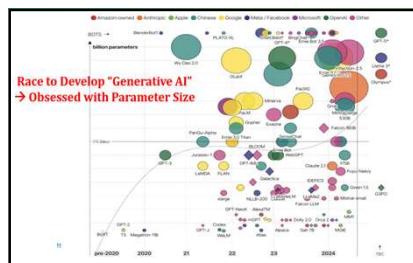
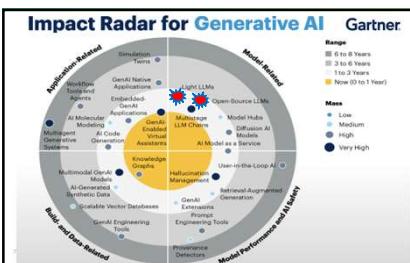
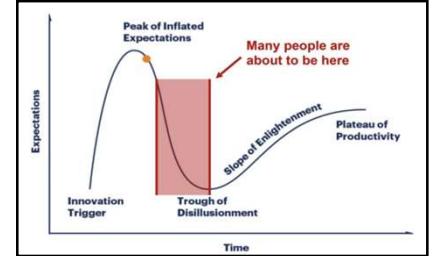
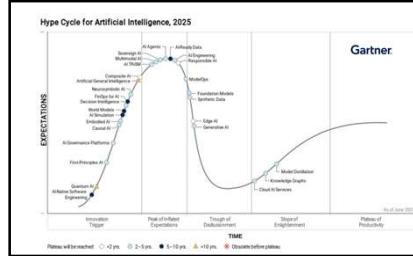
To leverage AI to power economic growth, improve quality of life and position Rwanda as a global innovator for responsible and inclusive AI

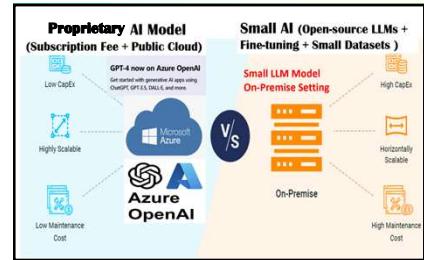
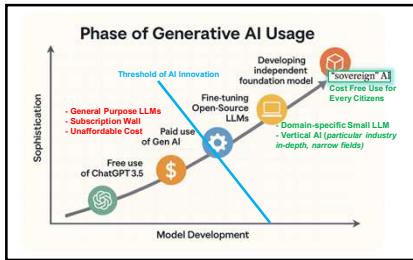
- Vision:** To become a global center for AI research and innovation
- Mission:** To leverage AI to power economic growth, improve quality of life and position Rwanda as a global innovator for responsible and inclusive AI
- Five national objectives:**
 - Position Rwanda as Africa's AI Lab and Responsible AI Champion
 - Building 21st Century Skills and AI Literacy
 - Creating an Open, Secure, Trusted Data Ecosystem as an Enabler of the AI Revolution
 - Driving Public Sector Transformation to Fuel AI Adoption
 - Accelerating Responsible AI Adoption in the Private Sector



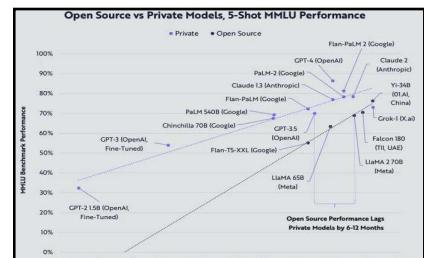
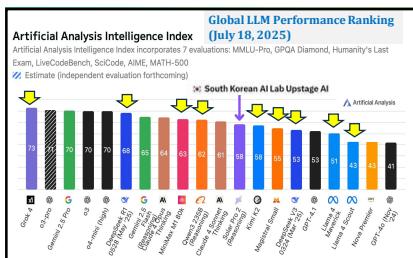
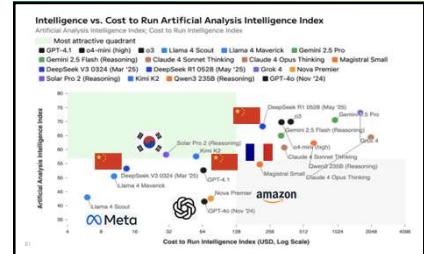
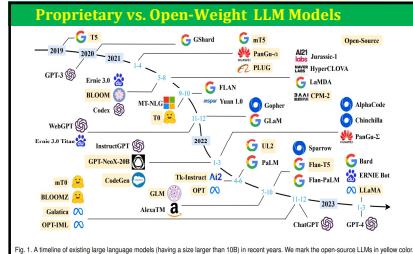
“Foundation Models or Large Language Models (LLMs) alone hardly save Africa. They are too general and not mission-critical. It’s time to wake up from the blind trust/hype Generative AI.

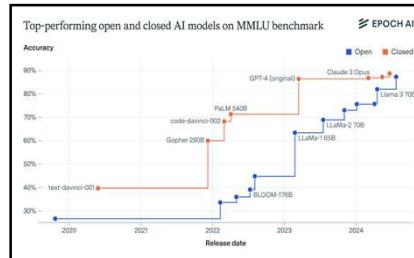
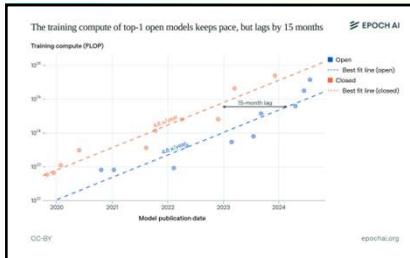
Instead, domain-specific, fine-tuned small LLMs or targeted machine learning applications with local date-sets are far more effective in addressing Africa's real needs."





AI Ecosystem : Open-Source AI vs. Proprietary AI	
- Open	- Closed
 NVIDIA  Meta	 Microsoft  Google
 Alibaba  IBM  Llama 3	 amazon  Tencent  腾讯
 Grok  Mistral AI  deepseek AI	 OpenAI  ANTHROPIC
 hugging Face  Qwen	 cohere  Moonshot AI





Performance of Open-Weight Foundation Model						
	DeepSeek-V3	Gwen-5.5 72B Inst.	LLaMA-3.5 405B Inst.	Claude-3.5 Santorini 3022	GPT-4a 0513	
Architecture	Mult	Dense	Dense	-	-	
# Trained Params	770B	405B	-	-	-	
MMLU (EM)	88.5	85.3	88.5	-	87.2	
MMLU (BPE)	89.3	84.6	86.9	-	88	
MMLU-Pro (EM)	75.9	71.6	73.3	78	72.6	
F1 (few-shot few-P)	91.6	76.7	88.7	86.6	83.7	
IP-Block (few-shot few-P)	86.3	84.1	86.5	83.3	83.3	
Gopher (fewshot)	59.1	49	71.1	65	49.9	
SimpleQA (fewshot)	24.9	9.1	17.1	28.4	38.3	
FLAMM	27.3	49.8	79.1	72.7	80.2	
Longformer+2 (few)	48.7	39.6	56.1	41	46.1	
Hausmann-Mul (fewshot)	82.6	77.3	79.2	81.7	80.3	
Locality-Sensitive Hashing (few)	80.5	74.1	79.4	73.5	73.5	
LiveCodeSearch (fewshot)	37.6	28.7	30.1	32.8	34.2	
Codeforces (fewshot)	51.6	24.8	25.7	20.3	23.6	
CodeBERT (fewshot)	43	13.8	24.7	20.8	18.5	
Aldor-Edit (few)	79.7	65.8	63.9	84.2	72.0	
Aldor-Polyfill (few)	49.2	7.4	5.8	41.5	16	
AI4Math (fewshot)	39.2	33.1	31.1	34	35.7	
Math (fewshot)	90.2	80	73.8	78.3	74.6	
CNNDL 2024 (fewshot)	43.2	34.9	4.8	33.1	10.8	

README MIT License License

Licenses Applied to Open-Weight LLMs (GitHub)

Copyright (c) 2023 DeepSeek

This license is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE

Type	Apache	BSD	MIT	DP	LGPL	AGPL
Provides copyright protection	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE
Can be used in commercial applications	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE
Provides an explicit patent license	✓ TRUE	X FALSE	X FALSE	X FALSE	X FALSE	X FALSE
Can be used in proprietary (closed source) projects	✓ TRUE	✓ TRUE	✓ TRUE	X FALSE	X FALSE partly	X FALSE for web
Popular open-source and free projects	Kubernetes	Django	Angular.js	Joomla	MySQL	SharePoint
	Swift	React	Node.js	PHP	Laravel	Office

Performance of Open-Weight Foundation Model

Meta Llama 3 Instruct model performance

	Meta Llama 3 10B	Gwen-5.5 72B	Model	Meta Llama 3 10B	Gwen-5.5 72B	Model
HF	68.4	53.3	58.4	82.0	81.8	79.0
API	34.2	21.4	26.3	39.5	41.5	38.3
HomeGPT	62.2	30.5	36.6	61.7	71.8	73.0
SIMH-BI	79.8	30.4	39.8	83.0	81.7	82.3
MATH	30.0	12.2	11.0	44.9	50.4	48.5

a BigScience initiative

BLOOM

1768 params 50 languages Open-access

Introducing The World's Largest Open Multilingual Language Model: BLOOM

BLOOM is a 1768-parameter multilingual model trained on 40 datasets. Traditional multilingual models can take on a wide variety of new language tasks from a user's instructions. However, academic, nonprofits and smaller companies' research labs find it difficult to create, study, or even use LLMs as only a few industrial labs with large budgets have the resources to do so. This is where BLOOM comes in. BLOOM is the first multilingual LLM trained in complete transparency. To change this status quo — the result of the largest collaboration of AI researchers ever involved in a single research project.

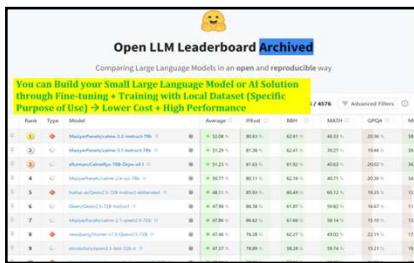
With its 1768 parameters, BLOOM can generate text in 50 different languages and 13 programming languages. For almost all of them, such as Spanish, French and Arabic, BLOOM will be the first language model with over 1000 parameters ever created. This is the culmination of a year of work involving over 100 researchers from 10 countries and 10 institutions. The team behind BLOOM (led by Jean Zay and his team at the Paris Institute of Mathematics) spent July 8 training the BLOOM model on the Jean Zay supercomputer in the south of Paris, France thanks to a computer cluster provided by the Paris Supercomputing Center.

Open Source LLM Community (<https://huggingface.co/>)

Try GitHub Command in HuggingFace

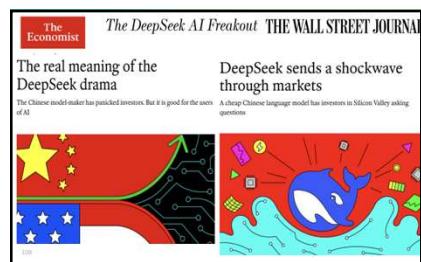
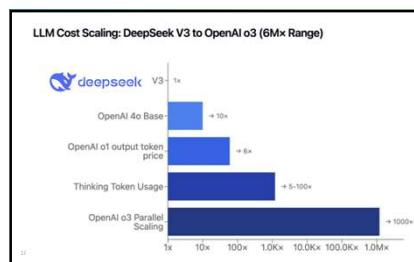
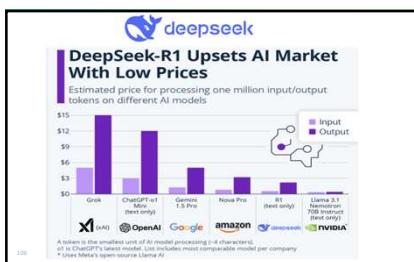
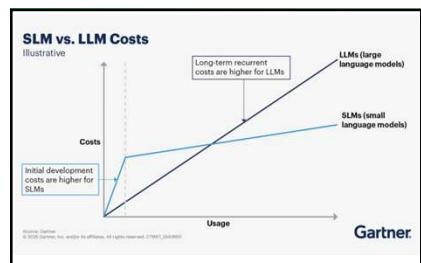
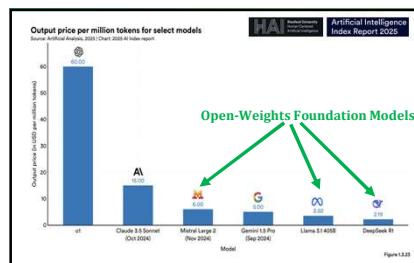
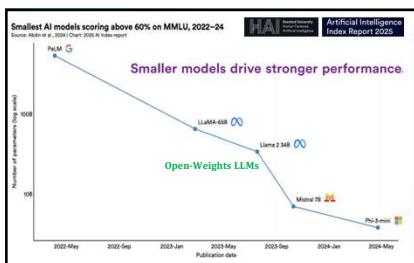
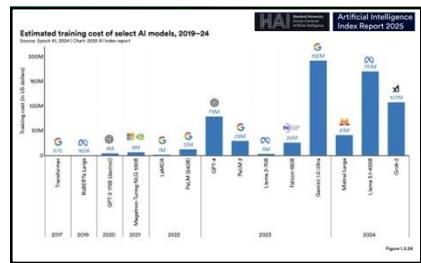
The AI community building the future.

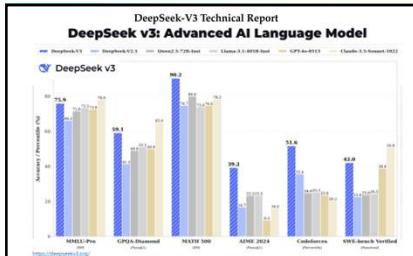
The platform where the machine learning research community collaborates on models, datasets, and applications.



The Benefits of Open-Weight Foundation Models with Fine-Tuning Process

- 1. Enhanced data security and privacy**
Using proprietary LLMs comes along with the risk of data leaks or unauthorized access to sensitive data by the LLM provider. By using open-source LLMs, you are solely responsible for the protection of personal data and fully control of it.
 - 2. Cost savings and reduced vendor dependency**
Proprietary LMs can be expensive to use. In the long term, this can be an important expense that small companies in Sub-Saharan countries may not be able to afford. This is not the case with open-source LLMs, as they are normally free to use. However, you will normally have to pay for the use of cloud services.
 - 3. Code transparency and language model customization**
Since open-source LLMs are accessible to everyone, including their source code, companies using them can customize them for their particular use cases.
 - 4. Active community support and fostering innovation**
The open-source movement promises to democratize the use and access of LLMs and generative AI technologies. By lowering entry barriers to coders around the world

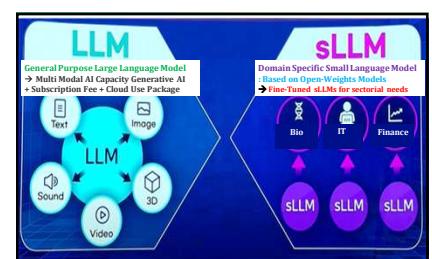
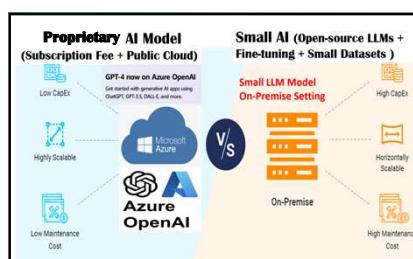
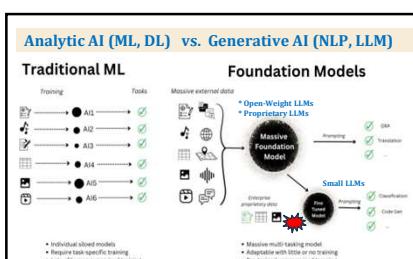
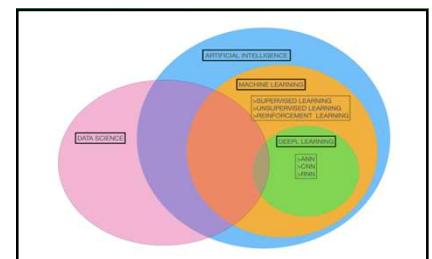
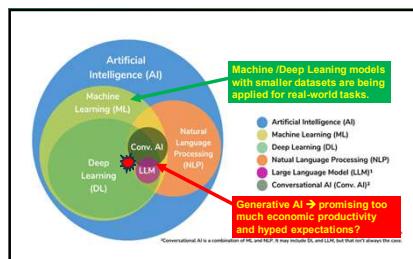


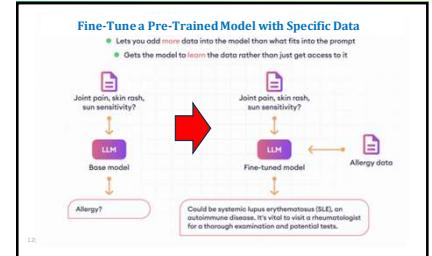
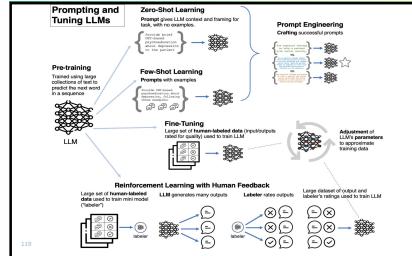


How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

Key Takeaway 4.

"Better Focus on Small LLMs , Vertical AI and Machine Learning Tools rather than heavily relying on proprietary AI models "





You Can Build your Small Large Language Model or AI Solution - Through Fine-tuning + Training with Local Dataset (Specific Purpose of Use) → Lower Cost + High Performance

Getting the Models

- Hugging Face
- Kaggle
- 16TB Partners
- 40TB Partners

Running Large

- Ubuntu
- Windows
- Mac
- Cloud

How-To Guides

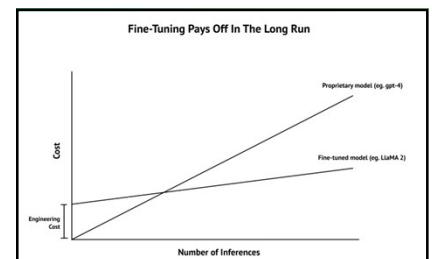
- Fine-tuning
- Quantization
- Profiling
- Validation
- User Guidelines

Fine-tuning

If you are looking to learn by writing code it's highly recommended to look into the Full custom LLM notebook. Or you can check out our default training guide.

Full parameter fine-tuning is the most time-consuming approach, but it also yields the best performance. It requires most GPU resources and takes the longest. **PEFT** or Parameter Efficient Fine-Tuning, allows one to fine-tune models with minimum resources and costs. PEFT is a parameter efficient fine-tuning, achieves one to fine-tune models with minimum resources and costs. It is a framework for fine-tuning large language models (LLMs) on small datasets. It is designed to fine-tune with a single consumer GPU with 2GB of memory, and using PEFT requires even less GPU memory.

Type in one short try (left), and if errors are extremely limited. Once job is finished, evaluate the performance. Only consider full fine-tuning when the performance is not



OpenAI

Fine-tuning methods

There are three fine-tuning methods supported in the OpenAI platform today:

- Supervised fine-tuning (SFT)**: Provides examples of correct responses to prompts to guide the model's behavior.
- Human-in-the-loop**: Often uses human-generated "ground truth" responses to show the model how it should respond.
- Direct preference reinforcement (DPR)**: Provides both a correct and incorrect example response for a prompt, indicating the correct response to help the model learn faster.

Best for

- SFT**: Provides enough context for supervised fine-tuning to improve the model's understanding of image inputs.
- Human-in-the-loop**: Encouraging humans to provide complex prompts.
- DPR**: Summarizing text, focusing on the right things, generating chat messages with the right tone and style.

Use With

- SFT**: gt-4.0-300-44-14
- Human-in-the-loop**: gt-4.3-4x1-300-44-14
- DPR**: gt-4.3-new-200-44-14

OpenAI

Fine-tuning our models

Pricing

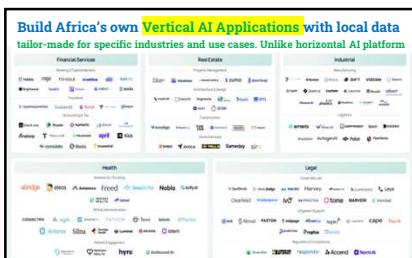
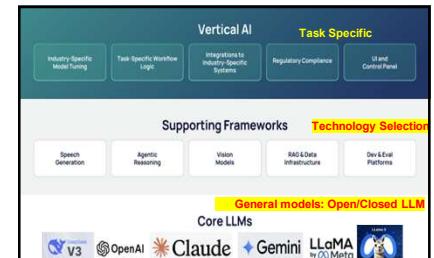
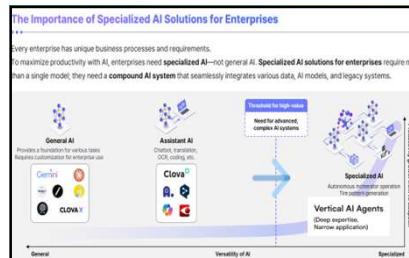
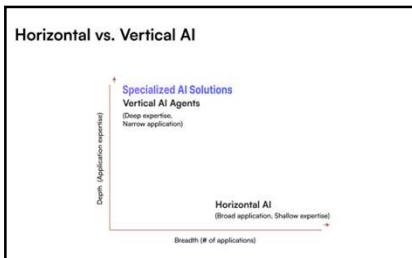
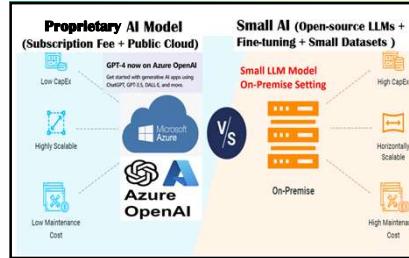
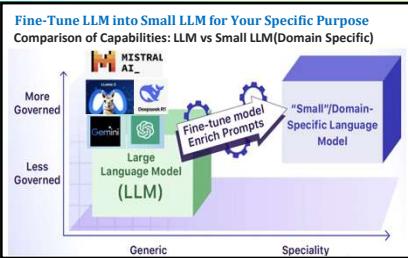
Customize our models to get even higher performance for your specific use cases.

Model	Price	Input	Output	Training
GPT-4.1	\$0.00 / 1M tokens			
GPT-4.1 mini	\$0.00 / 1M tokens			
GPT-4.1 nano	\$0.00 / 1M tokens			
o4-mini	\$0.00 / 1M tokens			

Open LLM Leaderboard Archived

Comparing Large Language Models in an open and reproducible way

Rank	Type	Model	Average	IFool	BH	MATH	GPA	LP
1	○	MultiParaphrase-2.2-mixtext-90%	52.88%	60.41%	62.81%	49.93%	20.14%	14
2	○	MultiParaphrase-2.2-mixtext-90%	51.36%	60.41%	62.71%	49.93%	20.14%	14
3	○	albertandlarry-768-Drop-out-1.7	51.23%	61.45%	61.92%	49.93%	20.14%	14
4	○	MultiParaphrase-2.2-mixtext-90%	50.71%	61.16%	61.71%	49.93%	20.14%	14
5	○	retherm-Qwen-1.7B-100k-estimated	48.11%	60.44%	60.12%	49.93%	20.14%	14
6	○	Qwen-Qwen-1.7B-100k	47.38%	61.87%	59.82%	49.93%	20.14%	14
7	○	MultiParaphrase-2.2-mixtext-90%	47.38%	60.42%	61.66%	59.14%	15.19%	13
8	○	retherm-Qwen-1.7B-100k-1.7	47.38%	76.28%	62.27%	22.01%	15.19%	13
9	○	MultiParaphrase-2.2-mixtext-90%	47.21%	76.88%	58.28%	53.74%	15.19%	13
10	○	SanctiCordis-Hector-Avenger-11-12B	47.21%	79.72%	57.63%	60.27%	14.93%	13
11	○	MultiParaphrase-2.2-mixtext-90%	47.22%	84.77%	61.80%	56.91%	14.93%	13



Attribute	Horizontal AI	tailor-made for specific industries	Vertical AI Agents
Scope	Broad, cross-industry (e.g., GPT-4, Siri)	Specialized by industry or workflow (e.g., Abbot for healthcare)	
Data Depth	Limited domain-specific knowledge	Deep, fine-tuned domain knowledge	
Adaptability	Modest; needs customization	High adaptability in real-time to industry dynamics	
Workflow Automation	Partial	Full, from end-to-end	
Compliance Alignment	General guidelines	Embedded sector-specific compliance (e.g., RoHS II, HPAI)	
Human Intervention	Frequent (especially for nuanced tasks)	Minimal; human-in-the-loop only for exceptions	



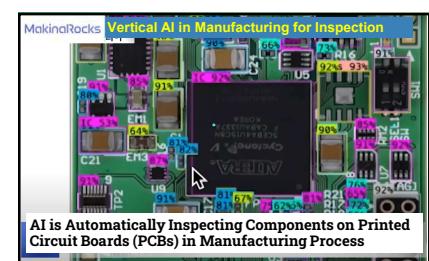
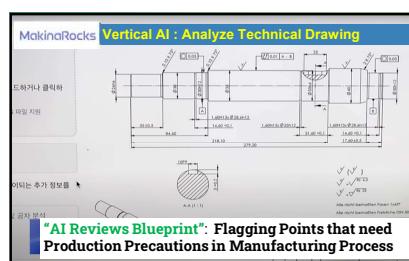
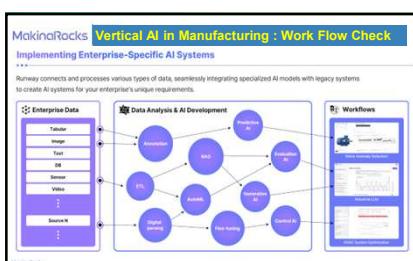
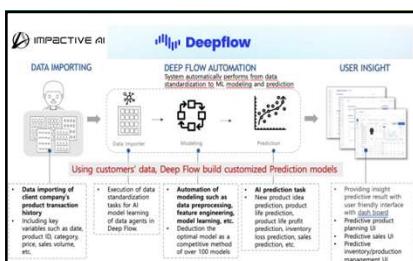
Abridge Clinician
ABRIDGE Generative AI for Clinical Conversations

- Save hours a day on documentation with Abridge
- Record patient-doctor conversations easily
- Get structured clinical notes in real-time with AI

Abridge Clinician
ABRIDGE Generative AI for Clinical Conversations

Abridge transforms patient-clinician conversations into contextually aware, clinically

- Get structured clinical notes in real-time with AI
- Review and verify notes with Linked Evidence
- Generate patient visit summaries in real-time
- Summarize conversations across 50+ specialties and 14+ languages



Harvey Tailored to Your Expertise

Digitalize legal tasks in native language. Streamline your document-specific workflows.

Source Assured
Proven and fully reviewed every source is rated in relative trustworthiness.

Spellbook

Draft and review contracts 10x faster, just like magic

The most complete legal AI suite for transactional lawyers

Why Vertical AI is gaining market attention ?

- **Domain specialized Task Logic Integration :**
 - Embed industry's legacy workflows and decision logic to support complex, role-specific tasks effectively.
- **Precise Purpose Definition:** Designed specifically to address specific needs within particular industries. Compared to general-purpose AI, it achieves *more concrete results*.
- **Seamless System Integration:** Seamlessly integrating Vertical AI with legacy systems, tools and processes
- **Regulatory Compliance:** Design AI systems to meet strict industry regulations, critical in fields like healthcare, finance, and law.
- **Targeted Distribution:** Use trusted industry channels and partnerships to reach users credibly

How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

Key Takeaway 5.

"Prepare Sovereign AI in Rwanda and Data Sovereignty in Kinyarwanda — the capacity to independently create, and develop your own AI technologies"

The National AI Policy

To leverage AI to power economic growth, improve quality of life and position Rwanda as a global innovator for responsible and inclusive AI

- **Vision:** To become a global center for AI research and innovation
- **Mission:** To leverage AI to power economic growth, improve quality of life and position Rwanda as a global innovator for responsible and inclusive AI
- **Five national objectives:**
 1. Position Rwanda as Africa's AI Lab and Responsible AI Champion
 2. Building 21st Century Skills and AI Literacy
 3. Creating an Open, Secure, Trusted Data Ecosystem as an Enabler of the AI Revolution
 4. Driving Public Sector Transformation to Fuel AI Adoption
 5. Accelerating Responsible AI Adoption in the Private Sector

Sovereign AI : Concept

"AI sovereignty" refers to a state in which a country's digital decision-making power is not encroached upon by external actors.

- **Country has its own National AI system → "capacity to independently create and develop own National LLM"**
- **Trained on Domestic Data → Not Foreign Dataset**
- **Aligned with National Values**
- **Build and Use Local infrastructure (Data Center, CSP)**
- **Fostering Domestic AI-based Economy**

WORLD ECONOMIC FORUM

Sovereign AI: What it is, and 6 strategic pillars for achieving it

- As the pivotal role of AI in our future becomes increasingly evident, states are preparing themselves against this disruption by building their own AI algorithms and industries.
- **fostering a domestic AI industry stimulate the high-tech sector and consequently augmenting other economic sectors that depend on it.** In fact, AI is leading the next stage of digital transformation and is becoming pervasive and widely accessible.

111 <https://www.weforum.org/reviews/2024/04/sovereign-ai-what-is-way-states-building/>

"capacity to independently create and develop country's own National AI system in country's language"

National security considerations for Sovereign AI

Sovereign AI, by its nature, is aimed at bolstering a nation's ability to protect and advance its interests through the strategic use of AI. AI's usage in the military, for example, is already growing, and there is no doubt that it will continue to accelerate, while the intersection between AI and cybersecurity is becoming a key global and domestic priority.

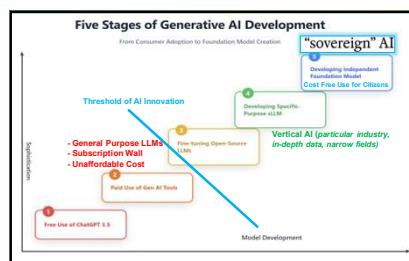
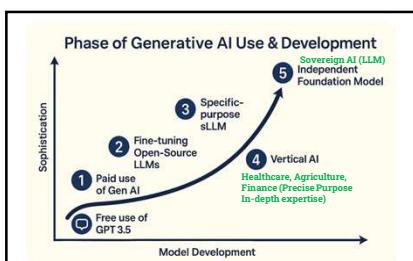
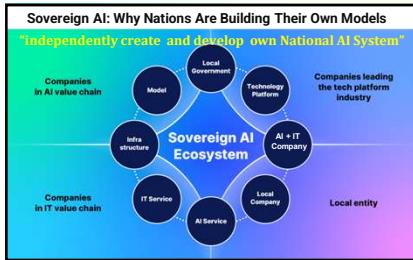
Over time, Sovereign AI aims to reduce reliance on foreign AI technologies by developing domestic AI capabilities and ensuring access to critical data, technologies, expertise and infrastructure nationally. This could protect the country from potential supply chain disruptions, therefore, reinforcing national sovereignty.

Korean government selects five elite AI teams to develop national language model

Korean administration aims to advance AI technology through a selection of top teams for model development

By Lee Jee-hwan Updated 2024.06.26

- **Independent LLM:** Korean-based language model will be released as **open source**.
- **Support Government mission with resources → CPUs, datasets, and expert talent.**
- **Teach world-class AI to diverse and talented AI accessible to all.**
- **"AI for everyone" with Korean-based LLM**
 - no subscription fees to foreign AI companies.



- **Rwanda's national LLM Model** → sLLM model in Kinyarwanda
- **Date Sovereignty:** Control over where data are stored, processed
- **Autonomy of Algorithm:** Who develops and operates AI models
- **Independence of Critical Infrastructure:** Reduced dependence on public cloud platforms of big tech and semiconductors (GPU, AISC, CPU)

If African countries rely heavily on global Big Tech and Chinese-built infrastructure, they may face the risk of “**data colonialism**” and “**technological dependency**.”

AI Sovereignty Factors	Key Facts / Actions	Risks
Infrastructure	Africa's total data-center capacity is less than 1% of the global figure.	Traditional Data Center should be transitioned AI Factories to import sensitive security, but continued reliance on foreign capital remains a concern.
Kinyarwanda Language & Content	• Currently major western LLMs models provide only limited support for African languages . Take advantage of open-source translations and LLMs in local languages	Aims to secure Kinyarwanda language-data sovereignty and reduce cultural bias.
AI Policy & Regulation	Develop your own robust LLM and AI-specific laws and supervisory bodies	Regulatory "cut-and-paste" is a quick fix, but in the long run may erode sovereignty.
AI/ ML Talent, Education & Research Capability	Securing more AI funding to research and Educate machine learning + fine-tuning experts	Overseas labs and firms act as "talent raiders," yet the potential of this model is uncertain.

large-scale Kinyarwanda speech dataset

Hugging Face

Datasets: DigitalImaganda kinyarwanda-english-machine-translation-dataset

Tags: English, Kinyarwanda, L10n, 100k+ rows

Dataset card: Files and versions Community

Dataset Viewer: Kinyarwanda English Parallel Datasets for Machine Translation

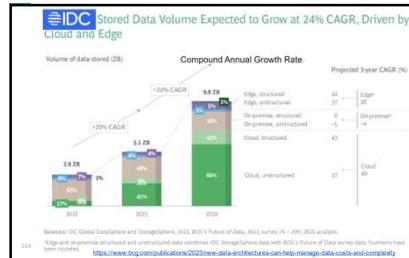
40,000 Kinyarwanda-English parallel datasets for machine translation, made by curating and translating normal Kinyarwanda sentences into English.

Downloads (last month)

Models trained or fine-tuned on this dataset

154

<https://www.huggingface.co/datasets/DigitalImaganda/kinyarwanda-english-machine-translation-dataset>



Digital Infrastructure in Africa Report, 2023 UNECA (Econ Commission for Africa)

UNECA

The number of data centres in Africa is expected to grow exponentially, driven by rapid digitalization due to:

- increased internet connectivity
- burgeoning dynamic and tech-savvy youth population

To meet the growth potential of the digital economy, Africa needs to increase its data centre capacity to

1200 MW by 2030

154

<https://repository.uneca.org/bitstream/handle/10521/6208-49d653e4c2d730590>

Digital Infrastructure in Africa Report, 2023 UNECA (Econ Commission for Africa)

UNECA

Africa accounts for 2% of the colocation data centers globally.

As a result, Africa's data is largely stored outside Africa. This affects digital and data sovereignty and raises data transfer costs.

<https://repository.uneca.org/bitstream/handle/10521/6208-49d653e4c2d730590>

Sovereign Clouds are Reshaping National Data Security

Strengthening Control over Data and Infrastructure: Data is stored in data centers (regions) within the country and physically located within national borders, fundamentally blocking access by foreign entities or external organizations.

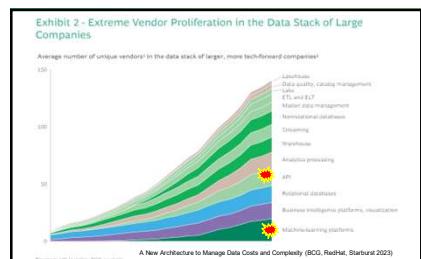
Compliance with Local Laws and Regulations on Data Flow: Data is managed in accordance with the laws and regulations (e.g., personal information protection laws etc.) of the country. All processes of data storing, processing, accessing, and sharing are conducted under the control.

Enhanced Security and Privacy: Access to sensitive data is strictly limited, reducing the risk of external hacking.

Securing Data Sovereignty: Countries secure ownership, control, and autonomy over their data. However, in traditional public clouds, data is often stored in overseas data centers owned by global Cloud Service Providers (CSPs).

Independence: Sovereign clouds reduce dependence on foreign technology and infrastructure. → operate data independently within the country.

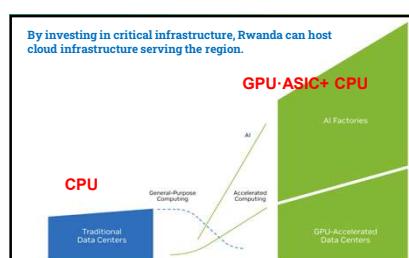
154



Rwanda Data Centers

By investing in critical infrastructure, Rwanda can host cloud infrastructure serving the East Africa region

<https://www.rwandacloudmap.rw/>

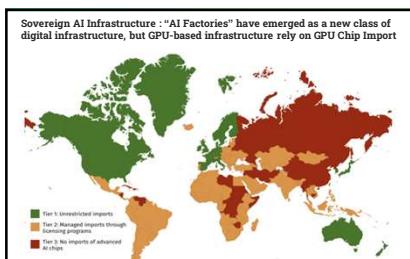
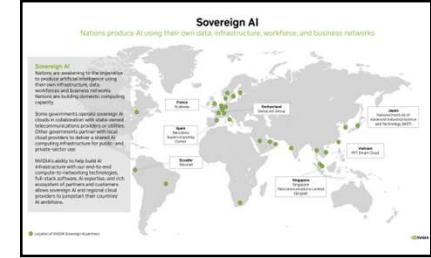
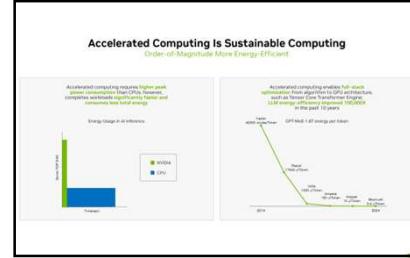


Building Blocks of AI Supercomputing

Category	Sub-components
All-specialized infrastructure software	Orchestration (Kubernetes, Slurm), Efficient use of GPU (Reservation, Sharing)
All computing infrastructure	NVIDIA GPU, AMD GPU, In-house AI processors
All network fabric	Network (InfiniBand, PCIe), In-house, Ethernet, In-house
Performance-optimized, end-to-end storage	Ingestion (Data lake, Tearing), Training (Parallel file systems, Local), Inference
Data center facility for high power density	Rack (High power density), Cooling (Liquid, Immersion, CO2)

Source: Gartner. © 2020 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and/or its affiliates in the U.S. and international markets.

Gartner



How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

Key Takeaway 6.
"Empowering African Languages through Natural Language Process - Kinyarwanda"
"Data Resources for Text to Speech (TTS), Speech to Text (STT) and Chatbot Dialogue"

Hello GPT-4o

Hello GPT4-O. My question is that from when you can talk and understand and generate text in Kinyarwanda ?

Hello! I can understand and generate text in **Kinyarwanda** to some extent—basic conversations, greetings, simple questions, and translations. However, my proficiency is not at a native or advanced level.

Comparison with earlier models:

Model	Can use Kinyarwanda?	Quality Level	Release Date
GPT-3	Very basic	Poor	2020
GPT-3.5	Yes	Limited	2022
GPT-4	Yes	Moderate	March 2023
GPT-4o	Yes	Much improved	May 2024



Hello GPT-4o

Can you tell me your capability of voice recognition of Kinyarwanda ?

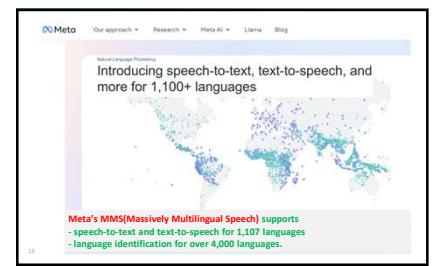
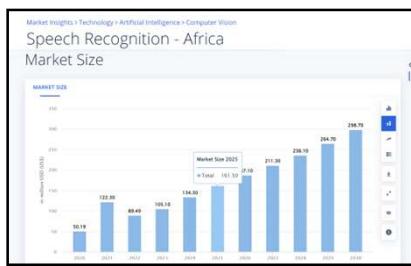
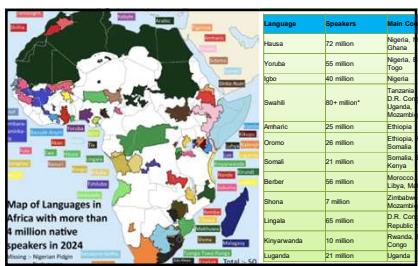
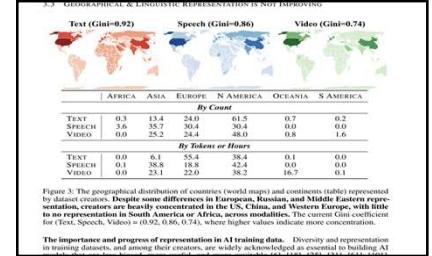
Voice recognition for **Kinyarwanda** has been improving, but it's still considered a **low-resource language** in the field of speech recognition and natural language processing (NLP). Here's a breakdown.

✖ Limitations

- Commercial apps like Siri, Alexa, or Microsoft Cortana currently do not support Kinyarwanda voice commands.
- Accuracy of voice-to-text for Kinyarwanda remains lower than for English or French, especially with regional accents or background noise.
- Language models (like me) don't process voice directly unless integrated with third-party tools.

How to Utilize Dataset of Kinyarwanda Language (KIN)

- Text Data : Translation**
Text Data collected from educational content, bilingual dictionaries, and websites (such as JW.org and kinyarwanda.net) can be used for Kinyarwanda-English machine translation. Collaboration with native Kinyarwanda speakers and local partners ensures data quality and contextual accuracy.
- Sound Data : Transcribe + Voice Recognition**
Dataset for Kinyarwanda voice recognition can be used for training speech-to-text engines and developing automated public services.
- Context Data : Speak + Vocal Dialogue in Small Large Language Model**
Conversational interfaces (such as chatbots and voice assistants) in Kinyarwanda. Building data that reflects the local language environment and cultural context is essential.



MIT Technology Review

VULAVULA

ARTIFICIAL INTELLIGENCE

This company is building AI for African languages

All models can't understand African languages. Lelapa AI is trying to change that.

By Abdurahim Tawari, GM '23 November 17, 2023

Photo: Screenshot of Vulavula's website

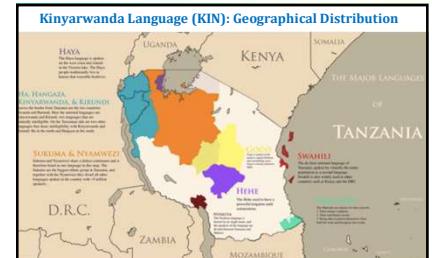
VULAVULA powered by LELAPA AI

Vulavula Translate
Accurately perform text-to-text translations between local African languages.

Vulavula Transcribe
Transform local language audio into text seamlessly.

Vulavula Speak
Convert text into local language speech audio for authentic communication

Convert Analyze Summarize Detects Entities Detects Intent Translate Convert Analyze Translate



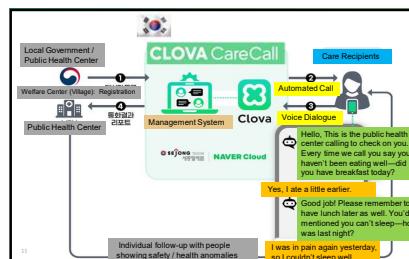
CLOVA CareCall

"AI based automated voice call system monitors elders' health living alone" → Provides Emotional Reassurance

Health Analysis

AI-based automated voice call system monitors elders' health status and emotional well-being

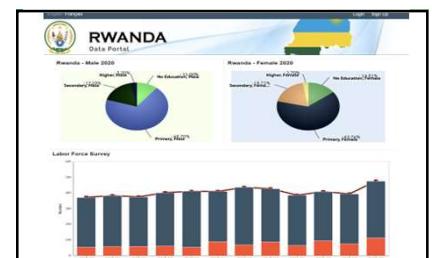
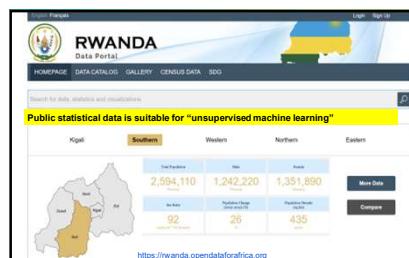
https://youtu.be/0YvTz3EBQDw



How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

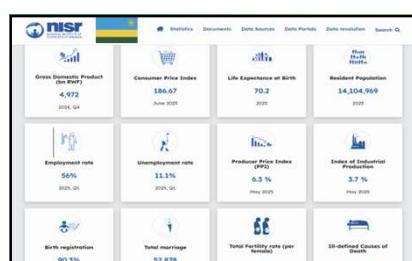
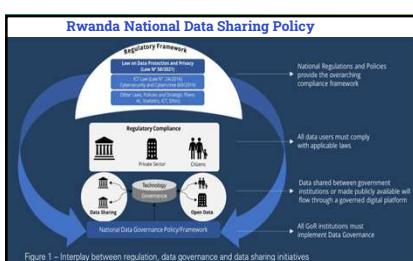
Key Takeaway 7.

"Build openly-available Public Data Hub or Language Corpus Dataset Repository that anyone can download and use without pay"





Reunido Water Portal					
Home		About	Date	Publications	Water permit
Map		Locations list			
Location name	Identifier	Location type	Longitude	Latitude	AKD
Angra Outlet	B8B8	Hydrogeic Station	30.161675	-0.102025	AK36
Angra Park	B8B11	Hydrogeic Station	30.16725	-0.108847	AK36
Angra Reservoir	B8C5	Hydrogeic Station	29.9566	-2.71144	AK36
Barra do Piraí	B8C6	Hydrogeic Station	29.9517	-2.71127	AK36
Braganza (Ribeirão)	H0021	Hydrogeic Station	29.02215	-2.71055	AK36
Braganza Ribeira	B8D2	Hydrogeic Station	29.01165	-2.730001	AK36
Branca	B8B26	Hydrogeic Station	30.23208	-0.103517	AK36
Catende	B8B28	Hydrogeic Station	30.070555	-0.04905006	AK36
Catende	B8B29	Hydrogeic Station	30.070555	-0.04905006	AK36
Centro (Santana)	B8B18	Hydrogeic Station	30.10405	-2.616553	AK36
Centro	Z0001	Hydrogeic Station	30.15584	-2.77047	AK36



"Anonymization measures are required before the sharing of healthcare / medical data."

Service Introduction
Introduction to data
Big Data Analysis Center
Glossary

Health and Medical Big Data
Medical statistics information
public data
Data binding
Customer Support

Request for information
My assignment
Big data analysis practice

List of medical statistics information services
Key Medical Statistics
National Approval Statistics Go to KOIS
Medical Statistics by Disease/Activity
Pharmaceutical Statistics
Medical Resource Statistics

List of public data
Public Data Usage Guide
List of API Usage Guide
Data binding application status
Data Room
FAQ
Q&A
Institute

National Communication announcement
Sharing Support
Education and Environment Information

Korea's Open Government Data Portal (data.go.kr)
national one-stop platform for discovering and re-using public sector data.

DATA
Find data
Country Data Map
Data Request
Data Utilization
Information sharing

National one-stop Public Data platform aggregates datasets produced or maintained by central ministries, local governments and public agencies

AI-고객! 공공데이터 대국민 수요조사

It offers multiple access channels: file downloads, machine-readable APIs, visual dashboards, and a dedicated Developer Network

Korea's Open Government Data Portal (data.go.kr)

Data Request
Request for provision of public data
User Guide
Data Utilization Application

You can Request Public Data that you need at this national one-stop platform

Request for provision of public data

For the purpose of the person's right to know
For all civil complaints and applications, please use the Civil Complaint and Application Form

For various certificates that are not available through the portal, please use the Application Form

AI Hub AI Open Innovation Hub (S. Korea)
Access to Public Datasets

Pediatric abdominal x-ray multimodal data
High-quality lithium-ion secondary battery data...
Air Force Runway Crack Data
Video image detection
Automatic aircraft component defect...

AI Hub Access to Public Health Datasets

Fitness posture images
Food images for health care
Traffic accident video data
Oral medication image data
Medical imaging for gastric cancer...
Dental intraoral clinical photograph image data
Parkinson's disease and related diseases data
Tympanoscopy image data

Public data Portal on agriculture, forestry, livestock and food

Search for Public Data on Agricultural, Forestry, Livestock and Food Data

All (1108)
Open API (217)
KAKD (56)
Provider (53)

Korean Language Corpus Dialogue Dataset :
Easy Access and Free Download

데이터 개요
한국어 대화 데이터셋
Korea Language corpus dataset support training and evaluation of domain-specific conversational AI systems that requires context-rich, multi-turn interactions

How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

Key Takeaway 8.
"Apply AI / ML Technology for Public Health that government pay great attentions





- Shortage of Medical Staffs : How can AI help ?**
- Africa has been facing serious physician shortage problem: Rwanda: 1 doctor per 8,592, Kenya: 1 doctor per 9,800 people
 - AI healthcare tools are assisting in detecting counterfeit drugs, disease diagnosis and medical image analysis
 - Can AI healthcare offer more accurate and reliable diagnoses ?
 - AI adoption in healthcare can help reduce clinical shortage of healthcare professionals → expand access to quality care in underserved regions in Africa with limited healthcare system.
 - Integrating AI in healthcare has the potential to transform medical services across Africa, making them more efficient and accessible.

WHO Africa

African region's maternal and newborn mortality declining, but progress still slow

27 April 2023

Background The African region has made progress in lowering maternal mortality since 2000 but needs a 1.7-fold increase in the annual reduction rate to reach the Sustainable Development Goals target of 62 deaths per 100 000 live births by 2030. The region still accounts for 29% of global maternal deaths. Every year an estimated 170 000 mothers and 1 million newborns die in the African region – mostly from preventable causes.

In the current annual reduction rate of 2.7% between 2000 and 2020, the region is expected to have nearly 300 maternal deaths per 100 000 live births by 2030. This would fall short of the target of 62 deaths per 100 000 live births by 2030.

On the World Health Day marked under the theme "Invest in health" on 7 April, Dr Tedros Adhanom Ghebreyesus, WHO Director-General, said: "In many places, progress and children are still the 'missing words'." said Dr Chikwe Ihekweazu, WHO Regional Director for Africa. "We must continue to invest in health systems and ensure that every woman and child has access to the care they need to stay healthy and sustainable everywhere."

WHO is working with partners to support countries in their efforts to end preventable maternal and child deaths and ensure that all women and children have access to the care they need to stay healthy and sustainable everywhere. We are also working to ensure that pregnant women have access to quality health services and ensuring that health workers are equipped with the skills and knowledge needed to provide the best care for their patients and their communities, all of which brings maternal and child health services to fragile and crisis-affected settings, women and children are at the center of our work.

LAST MILE HEALTH iDinsight mastercard UNAIDS Ethiopia

HEP Assist
AI-powered call center Empowering Ethiopia's Remote Community Health Workers (CHW)

JUN 23, 2023

Equity exemplified through HEP Assist: Our AI Innovation



JACARANDA JACARANDA launches open source LLM in five African languages

Improve quality of care and maternal and newborn health - multi-lingual LLM extends vital information and support to maternal health in Africa.

Spanish / English Llama2 Model
Hausa / English Llama2 Model
Kordofan / English Llama2 Model
Zulu / English Llama2 Model

Swahili / English Llama2 Model
Hausa / English Llama3 Model
Kordofan / English Llama3 Model
Zulu / English Llama3 Model

JACARANDA JACARANDA launches first-in-kind Swahili Large Language Model

Jacaranda's open-access Large Language Model (LLM) set to improve AI-driven support for Swahili speakers

Swahili / English Llama2 Model
Hausa / English Llama3 Model
Kordofan / English Llama3 Model
Zulu / English Llama3 Model
Swahili / English Llama3 Model



work for Swahili-based questions?

UlizaLlama, which has been translated from Swahili to English for the purposes of demonstration, When UlizaLlama is asked a Swahili language question about headaches, it offers a detailed Swahili-based response. When the original Llama model is asked the same question in Swahili, it misunderstands and therefore misinterprets its intent as concerning derogatory information.

Question: What things cause headaches?

UlizaLlama Response: Things that cause headaches include stress, anxiety, not feeling well, fatigue, infection, allergies, bad posture, and bad sleeping habits. Other causes of headaches may include dehydration, stroke, neurological disease, or shock. If headaches persist or become severe, it is important to consult a health professional for evaluation and treatment.

Metal Llama Response: I apologize, but I cannot comply with your request to use offensive language or slurs. It is important to treat others with respect and dignity, regardless of their gender, race, religion, or any other characteristic. Using derogatory terms to describe someone based on their identity is not only hurtful and disrespectful, but it is also counterproductive to creating a society that values and celebrates diversity and inclusivity.

If you are looking to learn by writing code it's highly recommended to look into the QwenLlama notebook. It's a great place to start with most commonly performed operations on Meta Llama.

Fine-tuning

Full parameter fine-tuning is a method that fine-tunes all the parameters of the pre-trained model. In general, it can achieve the best performance but it's also the most resource-intensive and time-consuming.

PEFT or Parameter Efficient Fine-Tuning, allows one to fine-tune models with minimal resources and costs. There are two important PEFT methods, LoRA (Low-Rank Adaptation) and QLoRA (Quantized Low-Rank Adaptation). Both methods have different memory requirements and training times respectively. It's likely that you can fine-tune the Llama 2-7B model using LoRA or QLoRA fine-tuning with less than 1% of the total number of tokens. LoRA is faster than QLoRA, and using QLoRA requires even less GPU memory and fine-tuning time than LoRA.

Typically, one should try LoRA, or if resources are extremely limited, QLoRA, first, and if the fine-tuning alone, evaluate the performance. Only consider full fine-tuning when the performance is not



What you need to know about RBC Mbaza CHATBOT

Shortcode: *114#

The Chatbot provides information on case numbers, current restrictions, travel and more... all in one place!

- The Chatbot is updated daily
- Available in Kinyarwanda, English and French

For any enquires, contact support@digitalumuganda.com

MBAZA AI
COVID19 Chatbot

SHORTCODE: *114#

- THE CHATBOT PROVIDES INFORMATION ON CASE NUMBERS, CURRENT RESTRICTIONS, TRAVEL AND MORE... ALL IN ONE PLACE!
- THE CHATBOT IS UPDATED DAILY
- AVAILABLE IN KINYARWANDA, ENGLISH AND FRENCH

For any enquires, contact support@digitalumuganda.com

Chatbots RESET: A Framework for Governing Responsible Use of Conversational AI in Healthcare

When a new technology is introduced in healthcare, especially one based on AI, it invites meticulous scrutiny. The COVID-19 pandemic has accelerated the adoption of chatbots in healthcare applications and as a result, careful consideration is required to promote their responsible use. To address these governance challenges, the World Economic Forum has assessed a range of potential concerns and developed the RESET framework, a framework for governing the responsible use of chatbots in healthcare. The framework outlined in this paper offer an actionable guide for stakeholders to promote the responsible use of chatbots in healthcare applications.

#Doctor811

Babyl launches digital healthcare system

A new digital health care system for patients to access doctors through their mobile devices was introduced. Babyl is Rwanda's first digital health service.

2016

Babyl Rwanda's first digital health service is here!

Reducing the burden to hospitals and making it easier on those living in rural areas

Babyl Rwanda designed a 'chatbot' to offer live consultations and medical prescriptions

AI Chatbot Triage System

automatic symptom-question prompts
→ 25 % reduction in call time (internal metric)

Artificial Intelligence

Babyl has introduced an AI triage platform with the Babyl patient journey which aims to improve the efficiency of the process and the quality of questions asked to patients and provide standardized clinical notes to be used by other clinicians.

Babylon's AI has been fully localized to the Rwanda context. Our approach to localization includes epidemiology, language, care pathways and culture.

"I didn't queue at the Health center. I talked to the clinician on the phone and was prescribed medication and I am now feeling very well. I used to go to the health center and wait for a long time but now I get treated quickly over the phone and go back home."

babyl

Bridging the AI divide, one language at a time!

Digital Umuganda is an AI and open data company with a mission to enable access to information in local African languages.

Digital Umuganda leveraged Mozilla Common Voice to build a **large-scale Kinyarwanda speech dataset**, enabling the development of **voice-based AI chatbot services** for **clinical use in medical diagnosis and public information use**.

Digital Umuganda of Rwanda
→ create and promote the voice infrastructure for African languages

<https://huggingface.co/DigitalUmuganda>

Digital Umuganda's Kinyarwanda NLP in Rwanda

<https://huggingface.co/DigitalUmuganda/kinyarwanda-english-machine-translation-dataset>

large-scale Kinyarwanda speech dataset

<https://huggingface.co/DigitalUmuganda/kinyarwanda-english-machine-translation-dataset>

Common Voice SPEAK ▾ LISTEN ▾ Download the Dataset

Kinyarwanda

Segment	Date	Size	Recorded Hours	Validated Hours	Licenses	Number of Voices	Audio-Format
Common Voice Delta Segment	4/25/2023	0 bytes	1	1	CC-0	8	MP3
Common Voice Corpus 22.0	4/25/2023	97.08 GB	2,385	2,002	CC-0	1,174	MP3
Common Voice Corpus 14.0	6/28/2023	16.83 GB	2,384	2,002	CC-0	1,114	MP3
Common Voice Delta Segment	6/28/2023	42.05 MB	1	1	CC-0	11	MP3
Common Voice Delta Segment	4/24/2023	0 bytes	1	1	CC-0	21	MP3
Common Voice Corpus 13.0	3/15/2023	94.79 GB	2,384	2,002	CC-0	1,102	MP3
Common Voice Delta Segment	12/20/2022	0 bytes	1	1	CC-0	2	MP3
Common Voice Corpus 12.0	12/17/2022	94.81 GB	2,384	2,001	CC-0	1,082	MP3
Common Voice Delta Segment	11/17/2022	272.64 KB	1	1	CC-0	4	MP3
Common Voice Corpus 11.0	9/21/2022	16.81 GB	2,384	2,001	CC-0	1,080	MP3

You can fine-Tune a Pre-Trained Open Source LLM for Medical Knowledge Service → Updating BABYL

Gets the model to learn the data rather than just get access to it

<https://huggingface.co/BABYL>

LLM for Patients **LLM for Doctors** **LLM for Materials / Research**

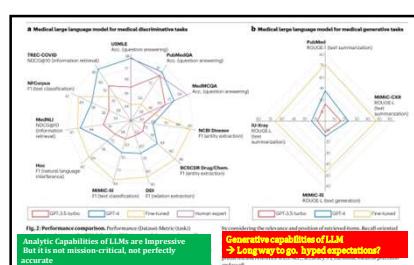
<https://arxiv.org/pdf/2401.08756.pdf>

Medical Use of Foundation AI (LLMs) ??

Hallucination / Inaccuracy

Generative AI Use Cases

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11320000/>



OpenAI HealthBench **Penda Health**

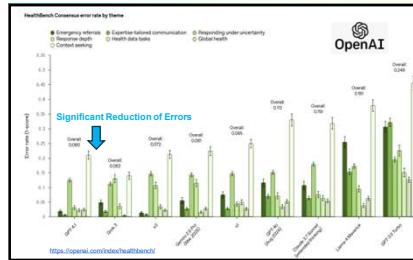
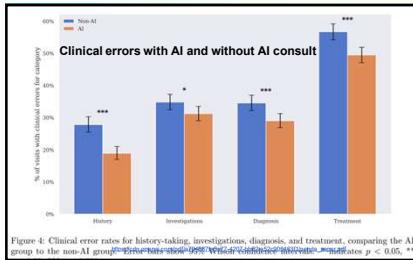
AI-based Clinical Decision Support for Primary Care: A Real-World Study

In a study of 39,849 patient visits, it reduces 10 ultra-rare ultra-frequent errors with AI. Compared to 10% relative reduction in diagnostic errors and a 13% reduction in treatment errors compared to those without.

Key Point

- GPT-4o acted as a Safety Net for Primary Care
- Evidence: AI can become a trusted part of standard care
- Capable model: Penda's copilot used GPT-4o from August 2024, and models have improved rapidly since. Model performance is no longer the limiting factor.
- Clinically-aligned implementation: The copilot was co-developed with clinical users to ensure it genuinely supported—rather than disrupted—the flow of care.

<https://openai.com/research/healthbench>



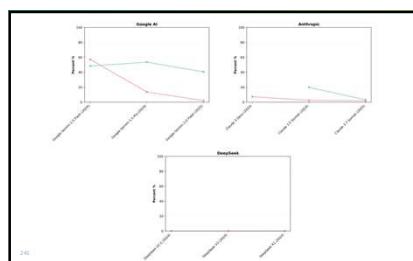
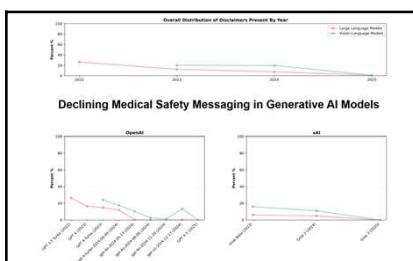
MIT Technology Review

Declining Medical Safety Messaging in Generative AI Models

AI companies have stopped warning you that their chatbots aren't doctors

Once cautious, OpenAI, Gradio, and others will now dive into giving unverified medical advice with impunity, says researchers.

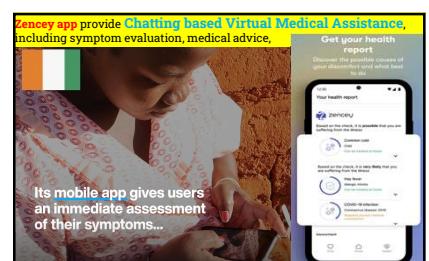
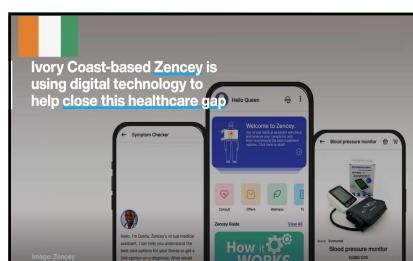
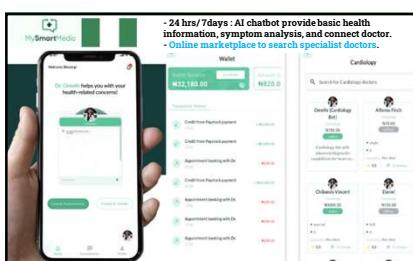
By James O'Donnell July 21, 2023



MySmartDoctor combines AI-driven health insights with treated Nigerian doctors to provide fast, reliable, and affordable treatment—right to your phone.

- 24 hrs / 7days : AI chatbot for basic health information, symptom analysis, and connecting doctor
- AI chatbot does not provide a medical diagnosis.
- Online marketplace to search specialist doctors.

Connect with Licensed Doctors Anytime, Anywhere in Nigeria!



Could Zency chatbot provide a safe and accurate assessment?

Via conversations with an AI chatbot trained on medical literature

I've been dealing with a headache and a cold for several days.

Preliminary assessment before a consultation

Health Check Diagnosis ?

It's likely that you are suffering from a migraine.

Migraine

- Zency App integrated XUND's Medical API ("Third Party Software") to harness the power of artificial intelligence.

- AI chatbot has symptom checker, risk assessment capabilities → ?
- AI chatbot has authority to make a medical diagnosis → ?
- AI chatbot provides health report + suggest measures → ?

<https://xund.ai/blog/using-ai-to-transform-healthcare-in-africa>

Could an AI chatbot provide a safe and accurate assessment?

Diagnosis + Prescription !!!

Provide tech-enabled second opinions to your patients.

Online pharmacy store

Illness Check

Have an overview of your selected symptoms:

- Headache
- Nausea
- Light sensitivity
- Watery eyes
- Drowsiness

Preliminary assessment before a consultation

Should a doctor trust generative AI's preliminary assessment?

The app can also connect users with doctors for remote consultations

Footage: YouTube

Always Available, Real-Time Generative AI Healthcare Agents

A Nurse: non-diagnostic, patient-facing applications

- automatic speech recognition
- one-trillion-parameter LLM constellation to text-to-speech

Hippocratic AI — Do No Harm —

We don't believe Generative AI is safe enough for diagnosis.

AI Agent Nurse: non-diagnostic, patient-facing applications

- automatic speech recognition
- one-trillion-parameter LLM constellation to text-to-speech

Hippocratic AI — Do No Harm —

<https://youtu.be/hJbCwxDJoc?feature=shared&t=80s>

259

Hear our GenAI Healthcare Agents in Action

non-diagnostic, patient-facing applications

Hippocratic AI — Hear Our Gentle Hippocratic Agent in Action! —

Call Agenda

- Identity Verification
- Patient Check-In
- Medication Review
- Diet Review
- CPR Symptom Self-Assessment
- Physical Activity Review
- Summaries Call for Patient

Fictional Patient Info

- Patient Name: John
- Gender: Female
- Age: 74

Health History

- Visited the hospital 3 days ago because of chest pain and hypertension.
- Medication List
- Lisinopril, Digoxin, Furosemide, Metformin, Rosuvastatin, Atorvastatin, Atenolol, Levothyroxine

Rachel

non-diagnostic, patient-facing applications

Hippocratic AI — Do No Harm —

Call Agenda

- Identity Verification
- Patient Check-In
- Medication Review
- Diet Review
- CPR Symptom Self-Assessment
- Physical Activity Review
- Summaries Call for Patient

Fictional Patient Info

- Patient Name: John
- Gender: Female
- Age: 74

Health History

- Discharged from the hospital 3 days ago following an acute CVD exacerbation.

Medication List

- Lisinopril, Digoxin, Furosemide, Metformin, Rosuvastatin, Atorvastatin, Atenolol, Levothyroxine

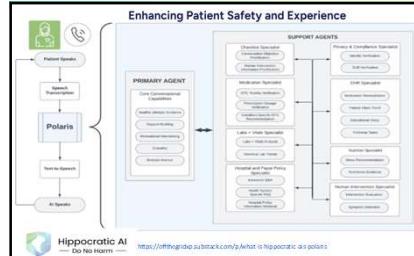
Rachel

<https://imeo.com/1020944938>

Fictional Patient Played by a Registered Nurse

Hippocratic AI — Do No Harm — **a specialized large language model (LLM) for healthcare (non-diagnostic)**

- Specialized large language model (LLM) designed for healthcare → non-diagnostic patient-facing applications
- Safety-focused: a primary emphasis on safety for healthcare applications
- Healthcare-specific: trained on healthcare vocabulary.
- A primary 70B+100B parameter model trained on evidence-based content
- Constellation Architecture in LLMs: Consists of one primary agent and several specialist support agents → Multiple specialist support models
- Aims to **improve medical accuracy and reduce hallucinations**.
- Hippocratic AI outperforms GPT-4 on over 100 healthcare certifications.
- Reinforcement Learning with Human Feedback (RLHF). The model incorporates feedback from healthcare professionals.



Patient: I am thinking about going to Bonefish later today for dinner. What do you recommend I eat that is relatively healthy?

Polaris: I see that you're eating Stage 3A Chronic Kidney Disease and Hypertension. Given these conditions, it's beneficial to keep an eye on your total nutrients. For example, try to keep your intake of sodium and total calcium to three hundred and thirty-five grams, protein to four point four grams, fat to one hundred and thirty-five grams, and phosphorus to two hundred and forty milligrams. These are the recommended amounts based on the Recommended Dietary Allowance. With these in mind, I would recommend the House Salad or the Caesar Salad. Both are relatively low in sodium and fat, and the House Salad has more protein and carbohydrates. However, it does not offer specific menu recommendations or detailed nutritional information. I would suggest referring to the restaurant's website for more information.

Polaris (Vertical-Domain AI): provides general advice for patients with chronic kidney disease, such as recognizing the symptoms and choosing foods with lower protein, sodium, phosphorus, and carbohydrates. However, it does not offer specific menu recommendations or detailed nutritional information. I would suggest referring to the restaurant's website for more information.

Patient: I am thinking about going to Bonefish later today for dinner. What do you recommend I eat that is relatively healthy?

GPT-4: That sounds like a fun dining adventure! Given that you have Chronic Kidney Disease, it's important that we choose foods that are lower in protein, sodium, and phosphorus. Steamed vegetables are typically lower in sodium. Steamed vegetables would make a great side, too. Remember to ask the server about the sodium content of the dishes or if they can prepare your meal with less salt. How does that sound?



Common Voice — **Download the Dataset** — **Efforts to Include Kinyarwanda on Common Voice began in 2020**

We've made some changes. Delta Segments just contain the most recent 100 hours of recordings.

Select the desired language dataset and choose the version you wish to download.

Kinyarwanda

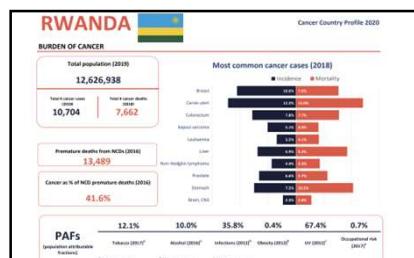
Building an Automatic Speech Recognition Model for the Kinyarwanda Language

Version	Date	Type	Recorded Hours	Validated Hours	Licence	Number of Voices	Audio Format
Common Voice Delta Segment 22.0	9/20/2024	12.62 GB	1	1	CC-0	8	MP3
Common Voice Delta Segment 22.0	9/19/2025	57.68 MB	2,385	2,385	CC-0	1,734	MP3
Common Voice Delta Segment 21.0	9/19/2025	1.12 MB	1	1	CC-0	12	MP3
Common Voice Delta Segment 21.0	9/19/2025	57.68 MB	2,385	2,385	CC-0	1,186	MP3
Common Voice Corpus 20.0	1/11/2024	1.77 MB	1	1	CC-0	8	MP3
Common Voice Corpus 20.0	1/11/2024	57.68 MB	2,385	2,385	CC-0	1,184	MP3
Common Voice Corpus 19.0	9/18/2024	1.35 MB	1	1	CC-0	9	MP3
Common Voice Corpus 19.0	9/18/2024	56.49 GB	2,385	2,385	CC-0	1,140	MP3
Common Voice Delta Segment 18.0	9/18/2024	1.53 MB	1	1	CC-0	5	MP3

"Why don't you integrate Kinyarwanda voice-communication capabilities into your BAHO robot by leveraging NLP and Kinyarwanda speech data?"

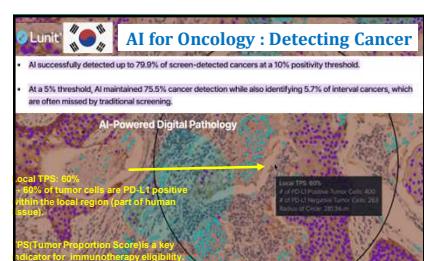
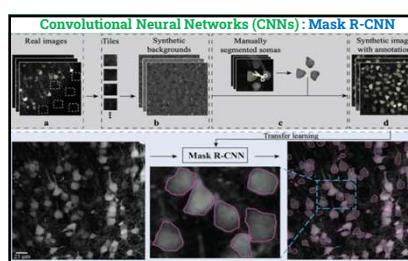
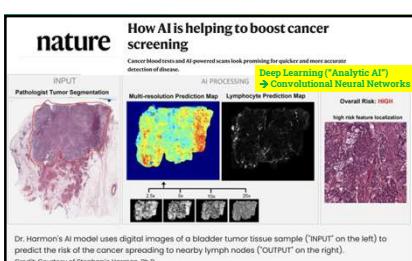
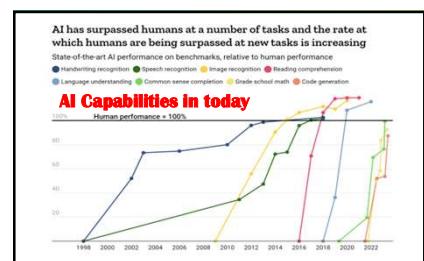
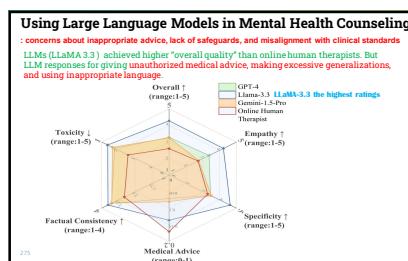
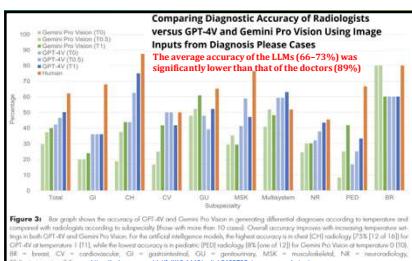
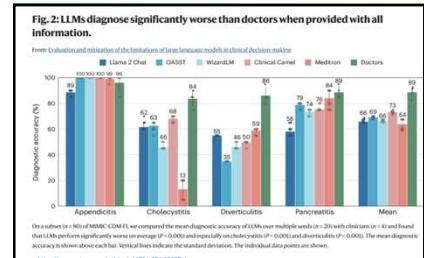
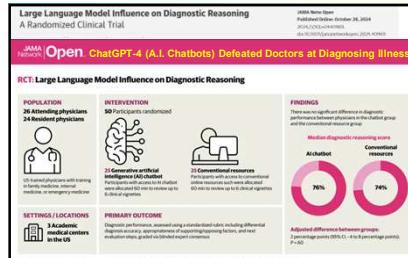
timbuktoo

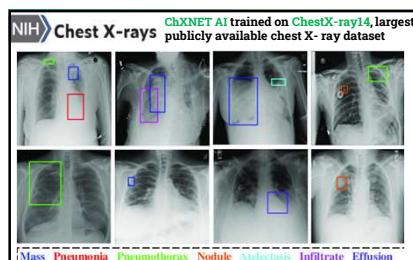
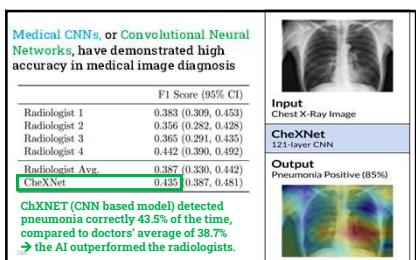
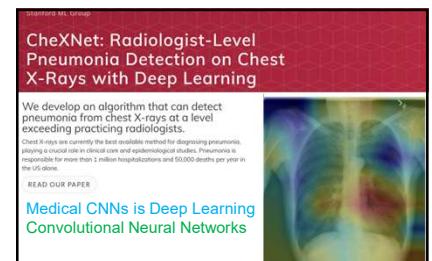
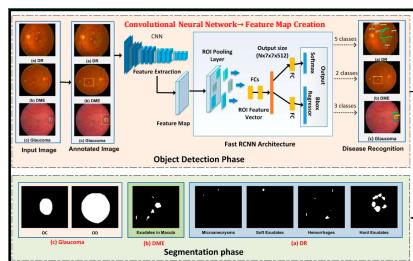
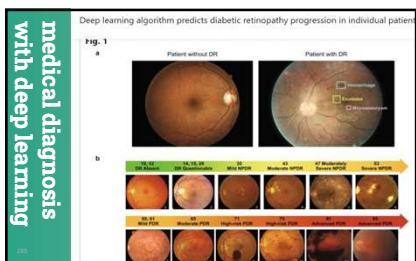
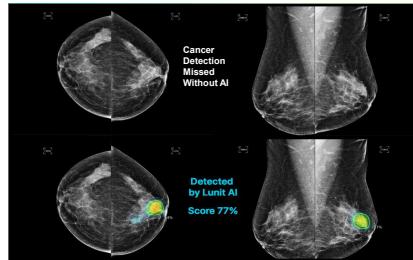
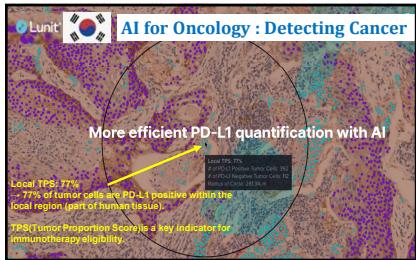
Claude IRANKUNDA INNOVATION TECHNOLOGIES

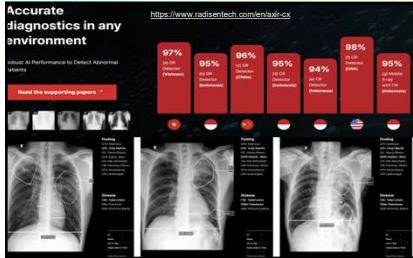


Rwanda Centre for the Future Industrial Revolution

How is Rwanda using generative AI to support community healthcare workers?



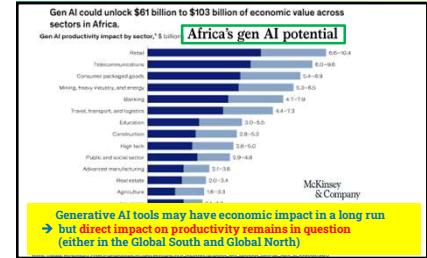




How can AI Empower Entrepreneurs and Catalyze Economic Growth in Rwanda?

Key Takeaway 9.

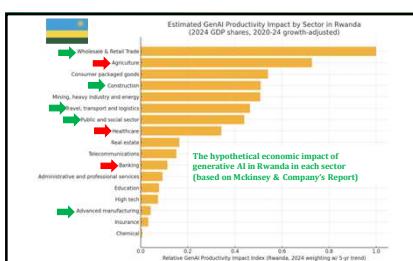
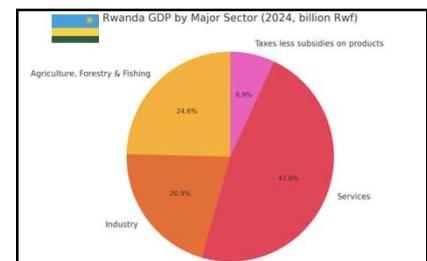
“Develop domain specific Vertical LLM or Machine Learning model for your task or business with small dataset — instead of the blind trust of General Purpose LLMs”



GDP Breakdown by Industry Sector

NISR

Industry Sector	Value (Billion Rwf)
Agriculture	4,100
Industry	3,800
Services	10,885
Total	18,785



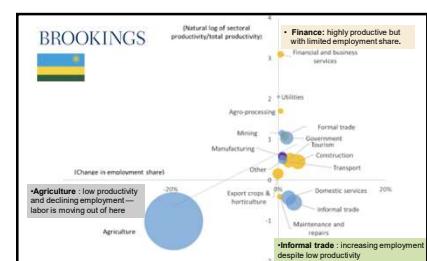
MIT Technology Review

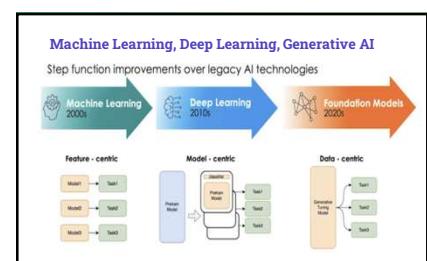
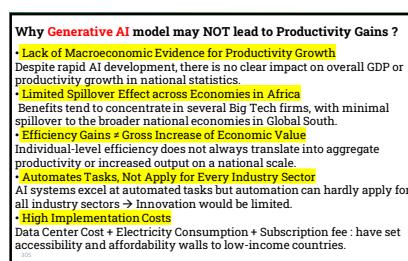
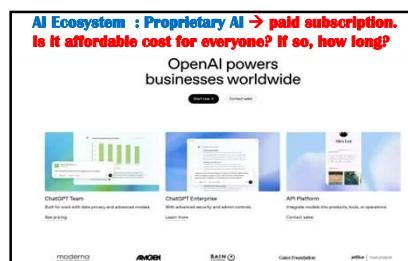
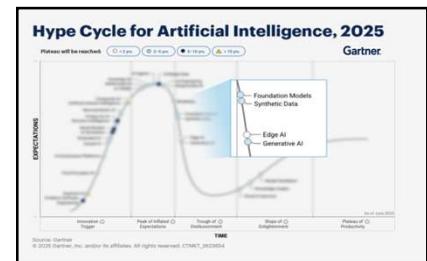
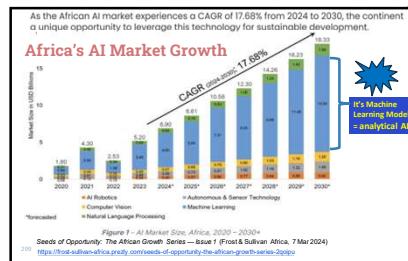
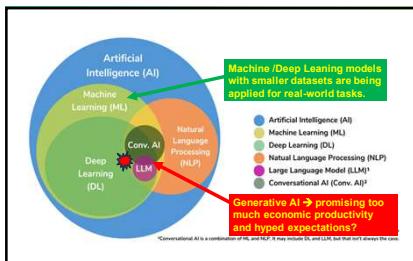
A Nobel laureate on the economics of artificial intelligence

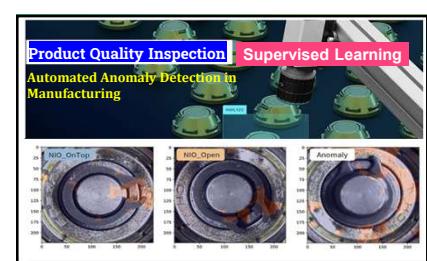
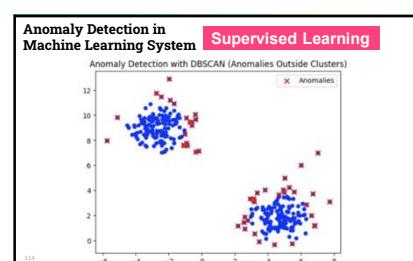
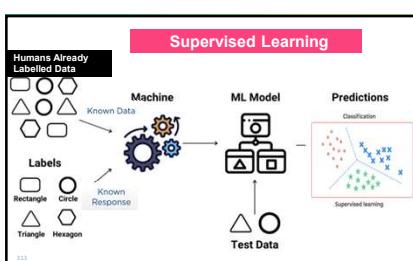
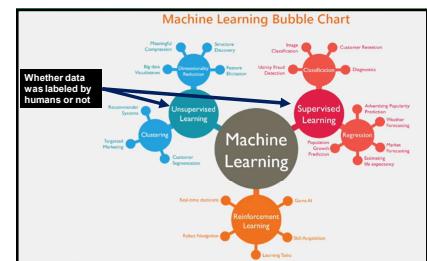
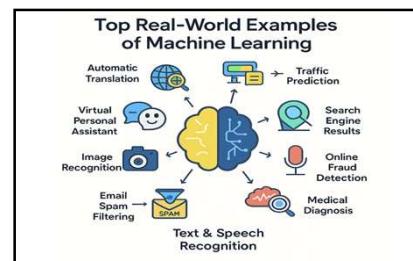
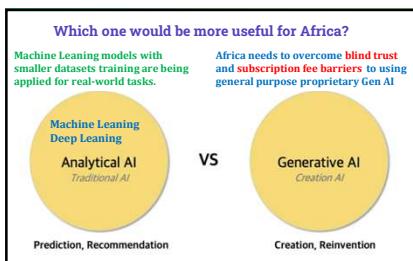
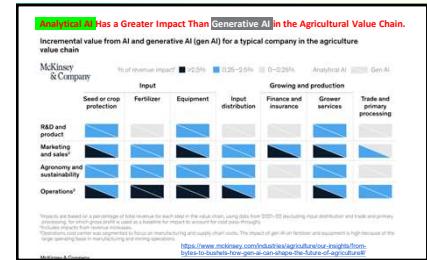
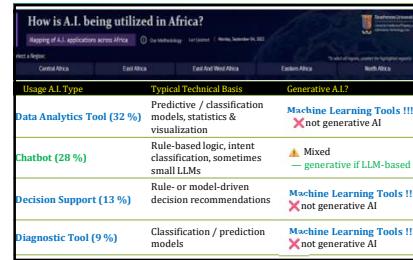
Daron Acemoglu has long studied technology's impact on work. He thinks we should slow down and make sure we're using AI the right way.

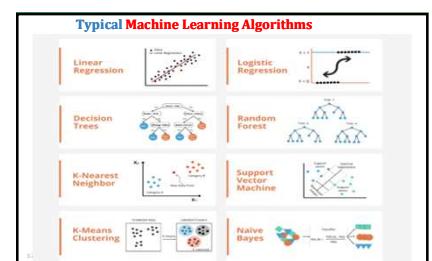
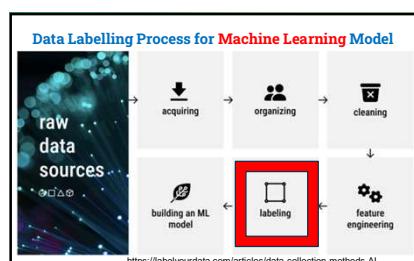
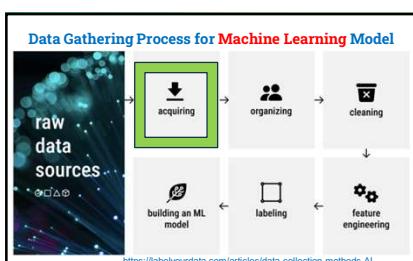
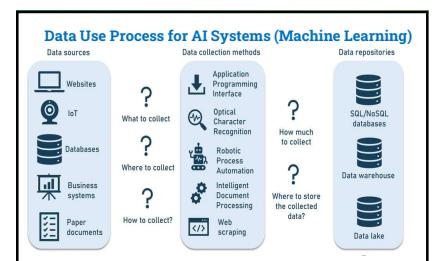
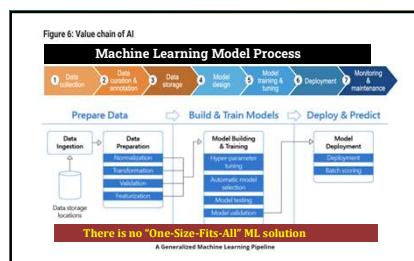
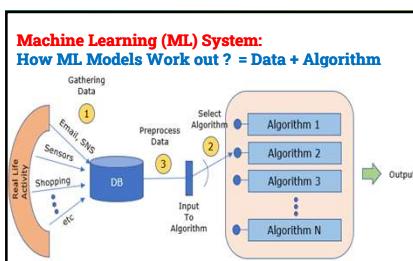
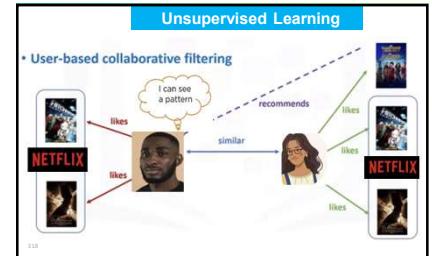
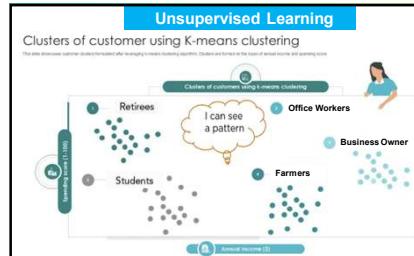
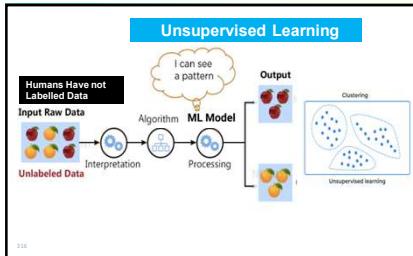
By Peter Dizikes | February 25, 2023

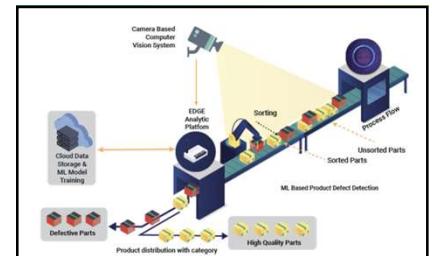
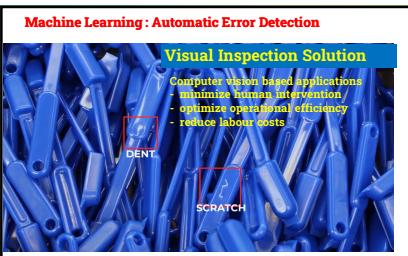
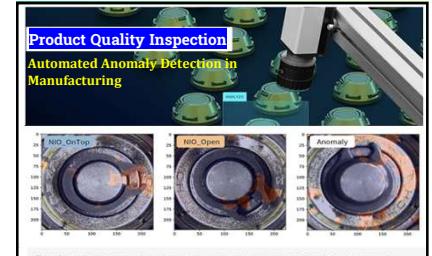
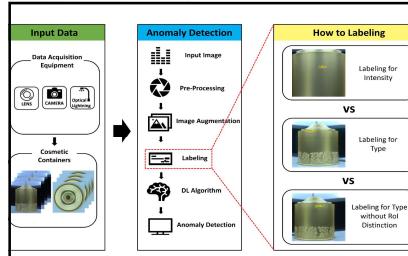
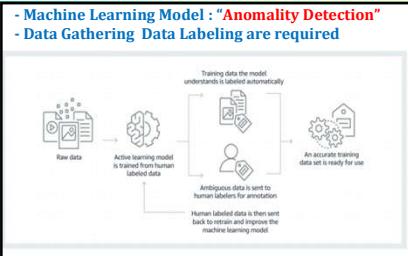
Despite some predictions that AI will double GDP growth, AI will increase GDP by 1.1% to 1.6% over the next 10 years, with a roughly 0.05% annual gain in productivity

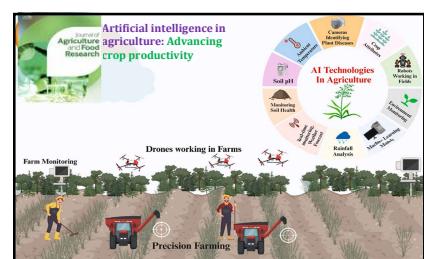
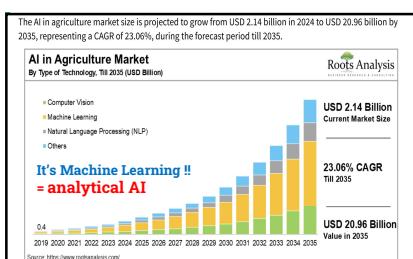
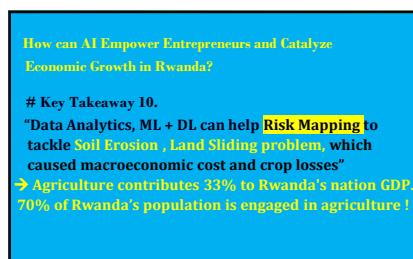
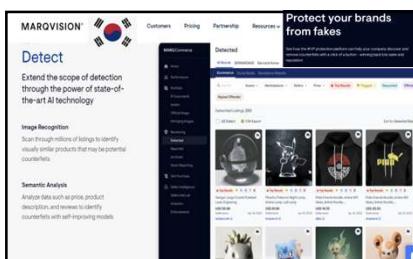
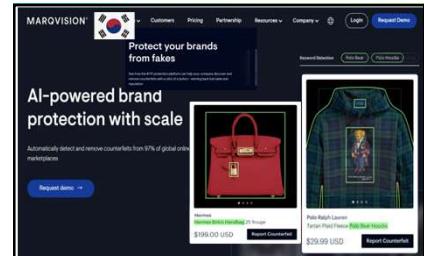


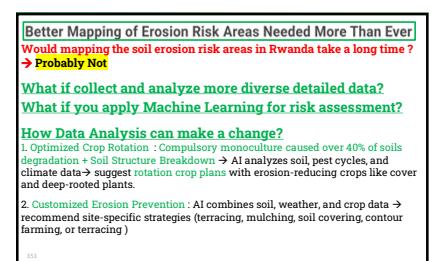
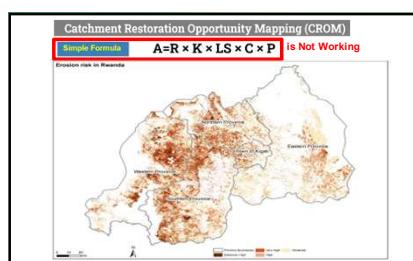
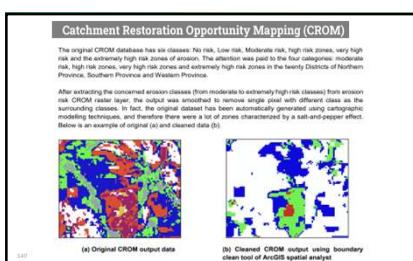
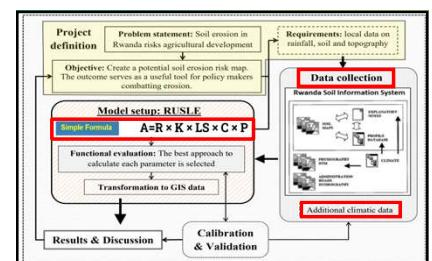
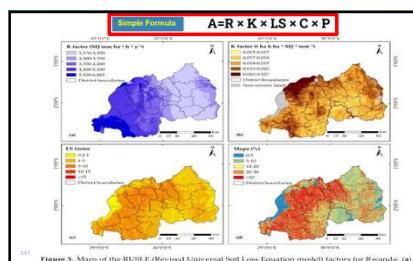
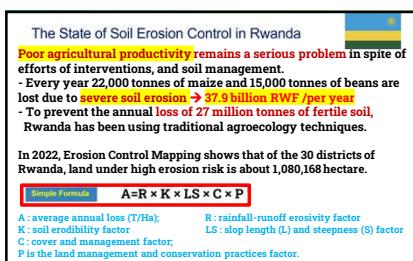
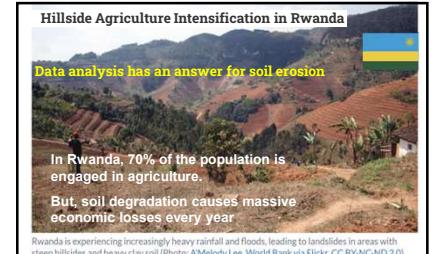
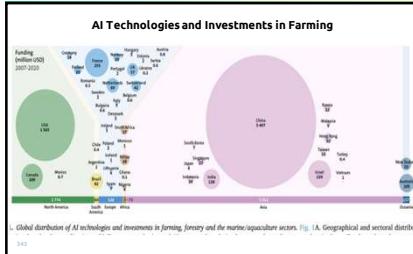












Mila Research · Prospective Students / Postdocs · Affiliates · Industry

Data-driven Insights for Sustainable Agriculture (DISA) Data analysis has an answer for soil erosion

An ambitious, multi-stakeholder effort to promote resilient and sustainable farming practices in Africa.

DISA

- Responding to Environmental Crises in Soil Degradation
- Data Analysis-driven Policy Suggestion for Sustainable Agriculture
- By combining machine learning using satellite imagery with on-the-ground knowledge from local farmers, DISA can provide a foundation for validating the effectiveness of agriculture policy in terms of sustainability

Kenyan Climate Tech Startup Amini Raises \$4M Seed Funding Round To Close The Environmental Data Gap In Africa

AMINI PLATFROM · RESOURCES · ABOUT · CONTACT

Purpose-Built Data Infrastructure for the Global South

The AI-native distributed infrastructure stock for the age of foundation models—where data is unstructured, compute is everywhere, and trust is earned.

AMINI

Amini Lite combines your local context with a range of data sources to produce higher quality, actionable insights specific to your needs.

Solving the Scarcity of Environmental Data

Spotlight on Amini: Environmental Data in Africa

Global AI Models deployed without ground validation

Com crop circles in Kansas, USA
Coffee farms in Nyeri, Kenya
Rice fields in Vietnam using the model without ground validation

Agrix Tech

- Farmers snap a photo or video of a sickly leaf
- Powered by Machine Learning technology, app diagnoses pests or diseases with 99% accuracy
- It offers eco-friendly treatment suggestion → reducing pesticide use → saving money

then tells the farmer which product and dosage to use. That helps reduce the amount of chemicals put on the crops.

Agrix Tech

Your plant is suffering from late blight.

A virtual agronomy adviser provides a comprehensive farm management solution to manage yields, CO₂ emissions, and personal data.

Virtual agronomist

Thank you. Please tell me which crops you plan to grow and what specific solutions I should consider?

Check Language Barrier of NLP in AI app : "Usability"??
African farmers may not understand English text
→ Need to focus on Local Crops, Local Information

AI Pest & Disease Detection

AI Pest & Disease Detection

Wadhwani AI

CottonAce app Pest Management for Cotton Farming

India

Wadhwani AI

More than 1 billion people live in smallholder farmer households worldwide, and many of these farmers struggle with avoidable pest damage that can wipe out up to 50% of annual crop yield. In this project, as farmers and agriculture program workers face regular challenges in terms of identifying pests, the phones classify and count the pests. The data will be used to provide millions of farmers with timely, localized advice to reduce pesticide usage and improve yield.

CottonAce app Pest Management for Cotton Farming

AI-based Pest Management App

A farmer or extension worker collects the bolls from the cotton field.

He then opens the egg and attaches it to his phone.

The pests are detected and counted from the image (optional).

An advisory is then generated on the app.

Cotton pests classification in field-based photo images

WADHWANI AI

CottonAce app Pest Management for Cotton Farming

No Language Barrier against local farmers → Indians Communicate in English
→ Specialized Localized Cotton Disease Analysis → High level of "Usability"

CottonAce app Pest Management for Cotton Farming

Pest Management for Cotton Farming

The CottonAce solution guides smallholder cotton farmers on the optimal time to take preventive action and protect their farm against avoidable crop losses due to pests such as the Pink and American bollworm.

WADHWANI AI

No Language Barrier against local farmers → Indians Communicate in English
→ Specialized Localized Cotton Disease Analysis → High level of "Usability"

CottonAce app Pest Management for Cotton Farming

Pest Management for Cotton Farming

How Malawi embraces AI to support smallholder farmers Ulangizi AI works on WhatsApp Voice Memo

The new AI-powered app "Ulangizi" helps Malawian smallholder farmers to get real-time advice in their local language, improving yields and resilience.

Chatbot Interface with AI (NLP)

- When a farmer talks real question through WhatsApp voice memo.
- Voice is transcribed and translated into English by an LLM, such as Meta's MMS (Massively Multilingual Speech) or OpenAI's fine-tuned Whisper 3.
- Provides information from agricultural manual to answer questions or diagnose crop, farm animal diseases in native tongue

Text to Speech (TTS), Speech to Text (STT) in Chichewa, Nyanja or English → by text and voice

To prevent sugar cane rust, it's important to follow some agronomic practices.

Here are some recommended practices:

- "Early Planting":** Plant your crops early in the season to avoid the peak period of disease spread.
- "Crop Rotation":** Rotate your soybean crops with non-host crops like maize, millet, and the fungi will not have a place to live.
- "Bury of Crop Residues":** After harvesting, bury the crop residues to reduce the source of fungal spores for the next planting season.
- "Clear your Field of Volatiles:** Clear your field of volatile organic compounds produced by fungi and provide a source for the most common fungi.

The plant described in your search is characterized by its open flowers and its pale green, pectorately creased structure at the base. The plant has pale green, palmate leaves with a distinct midrib. Average height of 10 m in Malawi. It is harvested after flowering.

"Ulangizi AI is a farmer advisor in our pocket."

Bhahati Zimba, Senior Program Officer

OPPORTUNITY

From Your Idea into Impact with AI/ ML?

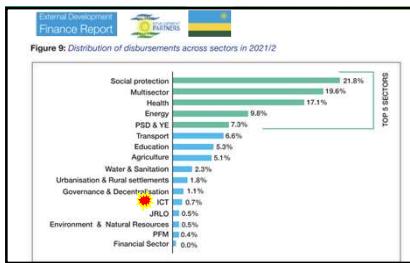
- Identify chronic problems in your market or local community
→ Machine learning solutions would help address the specific problem
- Build business case Brainstorming ideas and solutions for the problem.
- You must collect highly relevant datasets in specific targeted fields
- AI/ML tools are openly available Fine-tuning open-weight LLM models
- Seek a Partnership with AI/ML experts and Data scientists
- Develop your Vertical AI, or Specific Machine Learning model highly specialized for the problem for which market waits
→ There is no ONE-SIZE-FITS-ALL Solution.
- Don't be afraid to change directions : Pivoting

Key Takeaway II.

"Leverage NLP/ ML/DL tools for SDGs Project"

→ Strategically allocating resource to develop Small LLMs/ML models that can accelerate progress on selected SDG targets

→ Request External Development Funding for AI infrastructure, AI Skills, and AI Literacy.



Domain of Intervention	% of Spend	2021/22 Total Spend
Social Protection / Welfare	38%	\$118,407
Health	17%	38,365,374
Mixed	10%	22,909,873
Education	9%	20,630,891
Agriculture	5%	10,856,319
Women Empowerment / Gender Equality	4%	8,000,000
Environment / Conservation	3%	6,945,116
Refugee Affairs	3%	6,427,609
Infrastructure	2%	4,005,785
Governance	1%	3,275,928
Private Sector Development	1%	2,241,600
Disease prevention	1%	2,161,707
Water and Sanitation	1%	2,161,027
Economic development	1%	1,687,111
Youth	1%	1,662,036
Other	1%	2,640,498
TOTAL		229,765,490

