

EXPOSÉ • GRAND PUBLIC

L'IA au secours des langues africaines

L'intelligence artificielle peut-elle apprendre, traduire, et même enseigner les langues d'Afrique ?

Présentation • 2026



Trois questions, un même enjeu



Apprendre

Une IA peut-elle « assimiler » une langue africaine à partir de textes et de voix ?



Traduire

Peut-elle traduire ces langues de façon fiable, dans les deux sens ?



Enseigner

Peut-elle aider à transmettre et à apprendre ces langues ?

L'enjeu : préserver un patrimoine linguistique unique et le rendre vivant à l'ère numérique.

LE POINT DE DÉPART

Un patrimoine unique... et menacé

≈ 2 000

langues parlées en Afrique :
près d'un tiers des langues
du monde

428

langues africaines
menacées, selon l'Atlas de
l'UNESCO

10 %

pourraient disparaître d'ici
la fin du siècle

Chaque langue qui s'éteint emporte une culture, une mémoire et une façon de voir le monde.

Sources : UNESCO ; Atlas UNESCO des langues en danger (via Wikipedia, « Language endangerment in Africa »).

Le fossé numérique linguistique



Pourquoi l'IA « ignore » la plupart des langues africaines

- Les IA apprennent à partir de données ; or ces langues ont peu de textes et d'enregistrements numérisés.
- Avant 2022, moins de 25 langues africaines étaient correctement prises en charge par les grands traducteurs (Meta).
- La majorité des langues sont quasi absentes du web, des claviers et des assistants vocaux.
- Résultat : un cercle vicieux : peu de données → peu d'outils → encore moins de contenus créés.

Source : Meta AI, « No Language Left Behind (NLLB-200) », 2022.

Comment une IA « apprend » une langue



1. Rassembler

Des données : textes, traductions, voix enregistrées.



2. Entraîner

Le modèle repère seul les régularités de la langue.



3. Utiliser

Il traduit, transcrit ou génère du texte.



Le « zero-shot » : certaines IA (ex. Google) apprennent une langue à partir de textes non traduits, sans exemples bilingues. Puissant... mais « ce n'est pas parfait » : la qualité reste inégale.

Source : Google, « Google Translate learns 24 new languages » (Zero-Shot Machine Translation), 2022.

LA RÉPONSE : OUI

Traduire : déjà réel, et en plein essor



Google Translate

- 2022 : +5 langues africaines (bambara, éwé, lingala, luganda, oromo) via « zero-shot ».
- 2024 : +110 langues, la plus grande extension - près d'un quart sont africaines (wolof, kikongo, ga, luo, fon...).
- Portée : 614 millions de locuteurs, propulsé par le modèle PaLM 2.



Meta - NLLB-200

- Un seul modèle d'IA traduit 200 langues, dont 55 langues africaines.
- Avant : moins de 25 langues africaines bien prises en charge - plus du double désormais.
- Modèle et données publiés en open source (banc d'essai FLORES-200).

Sources : Google (Blog Translate, 2022 & 2024) ; Meta AI, NLLB-200, 2022.

PAR ET POUR LES AFRICAINS

L'Afrique construit sa propre IA



Masakhane

Communauté panafricaine de chercheurs : une IA du langage « pour l'Afrique, par les Africains ».



Lelapa AI (Afrique du Sud)

InkubaLM : un « petit » modèle (400 M paramètres) pour 5 langues - isiZulu, yoruba, haoussa, swahili, isiXhosa.



GhanaNLP - appli Khaya

Traduction mobile des langues du Ghana (twi, éwé, ga...), pensée pour un usage local.



Mozilla Common Voice

Des données vocales ouvertes, créées par les communautés (ex. le kinyarwanda, parmi les plus grands corpus).

Sources : masakhane.io ; lelapa.ai ; ghananlp.org ; Mozilla Foundation (Common Voice).

Enseigner : une promesse émergente



Tuteurs & chatbots

Dialoguer avec une IA pour s'exercer à converser dans sa langue.



Prononciation

La reconnaissance vocale corrige et guide l'apprenant.



Contenus pédagogiques

Générer exercices, lexiques et histoires pour l'école.



À nuancer : l'enseignement par l'IA reste naissant. Il dépend des mêmes données rares, et très peu d'applications grand public existent par langue.

QUAND ÇA CHANGE LA VIE

De l'outil à l'impact réel



Le kiswahili, langue-pont

- Plus de 200 millions de locuteurs : la langue la plus parlée d'Afrique subsaharienne.
- Langue officielle de l'Union africaine.
- Journée mondiale du kiswahili (7 juillet) : 1^{re} langue africaine honorée par l'ONU.

Les secteurs qui en profitent



Santé : diagnostic et traitement des patients



Éducation : apprentissage dans sa langue



Administration



Accès à l'information

Source : ONU, Journée mondiale de la langue kiswahili (un.org).

Limites, risques et éthique



Qualité imparfaite

Le « zero-shot » se trompe encore ; accents et diacritiques mal gérés.



Données rares & biaisées

Peu de corpus, souvent non représentatifs des usages réels.



À qui appartiennent les données ?

Consentement et bénéfices doivent revenir aux communautés.



Dépendance aux plateformes

Le risque de dépendre de quelques géants technologiques.



Uniformisation

Standardiser une langue peut effacer dialectes et variantes.



Rythme inégal

Les grandes langues avancent vite ; les plus petites restent à la traîne.

Conditions de réussite



Des données ouvertes

Constituées avec et pour les communautés locutrices.



Co-construction

Concevoir les outils avec les locuteurs, pas seulement pour eux.



Soutien public

Politiques linguistiques et financements pérennes.



Talents locaux

Former chercheurs et développeurs sur le continent.



Mesurer la qualité

Des bancs d'essai (type FLORES) pour progresser honnêtement.

À RETENIR

L'IA est un outil puissant, pas une baguette magique.

La préservation des langues reste une œuvre humaine et communautaire.



Apprendre

✓ déjà une
réalité



Traduire ✓

en plein essor



Enseigner

✓ prometteur, à
bâtir

Merci — Questions ?

POUR ALLER PLUS LOIN

Références

UNESCO — Les langues en Afrique : diversité culturelle et multilinguisme.

unesco.org/en/articles/unesco-and-promotion-languages-africa-cultural-diversity-and-multilingualism

UNESCO / Atlas — Langues africaines en danger (428 menacées).

en.wikipedia.org/wiki/Language_endangerment_and_extinction_in_Africa

Google — « Translate learns 24 new languages » (zero-shot), 2022.

blog.google/products/translate/24-new-languages

Google — « 110 new languages, biggest expansion » (PaLM 2), 2024.

blog.google/products-and-platforms/products/translate/google-translate-new-languages-2024

Meta AI — « No Language Left Behind (NLLB-200) », 200 langues.

ai.meta.com/blog/nllb-200-high-quality-machine-translation

Masakhane — Communauté NLP panafricaine.

masakhane.io

Lelapa AI — InkubaLM — petit modèle pour langues africaines.

lelapa.ai/blog/inkubalm-small-language-model

GhanaNLP — Application de traduction Khaya.

ghananlp.org

Mozilla — Common Voice — données vocales ouvertes.

mozillafoundation.org/en/blog/common-voice-african-languages-stt-model-sprint

ONU — Journée mondiale de la langue kiswahili.

un.org/en/observances/kiswahili-day