



Masakhane

“We build together”

<https://masakhane.io>

Dataset Creation

Language Technology Model Building

ARTIFACTS

Original Content

Annotated Content

Datasets

AI model

feedback

AGENTS

content creator

annotator

curator

language technologist

evaluator

REQUIREMENTS

keyboard,
online platform

dictionaries,
keyboard, CAT
tools

crawlers,
lang ID

AI toolkits,
preprocessors,
compute

automatic & human
metrics
linguistic knowledge

NLP Participatory Research

Participatory Research as a means to ensure everyone who should be in the room is in the room (e.g: civil society, linguists, developers, etc.)

Masakhane Community

2,115 members
150+ African languages
70+ Publications
3+ Awards
200+ Datasets
12 workshops & socials



Artwork by Keneilwe Mokoena