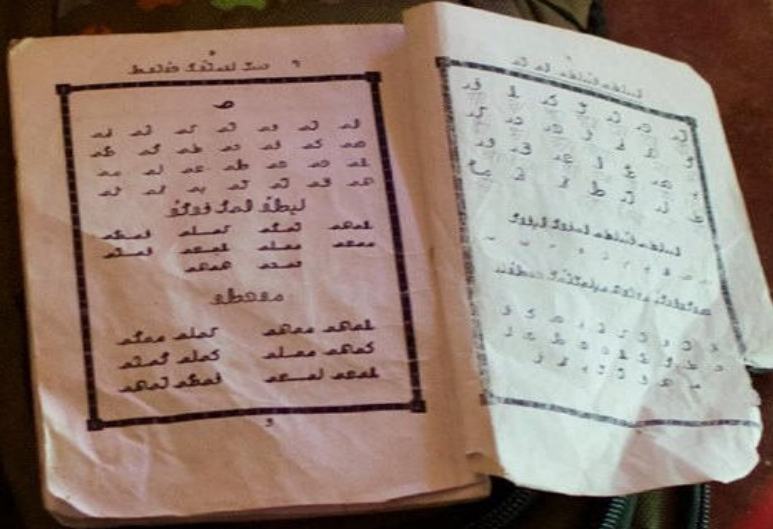




# TRANSLATORS WITHOUT BORDERS



## **About Us**

**Translators without Borders (TWB) is a non-profit organization offering language and translation technology and services for humanitarian and development agencies, and other non-profit organizations, on a global scale.**

## **Our Vision**

A world where knowledge knows no language barriers.

## **Our Mission**

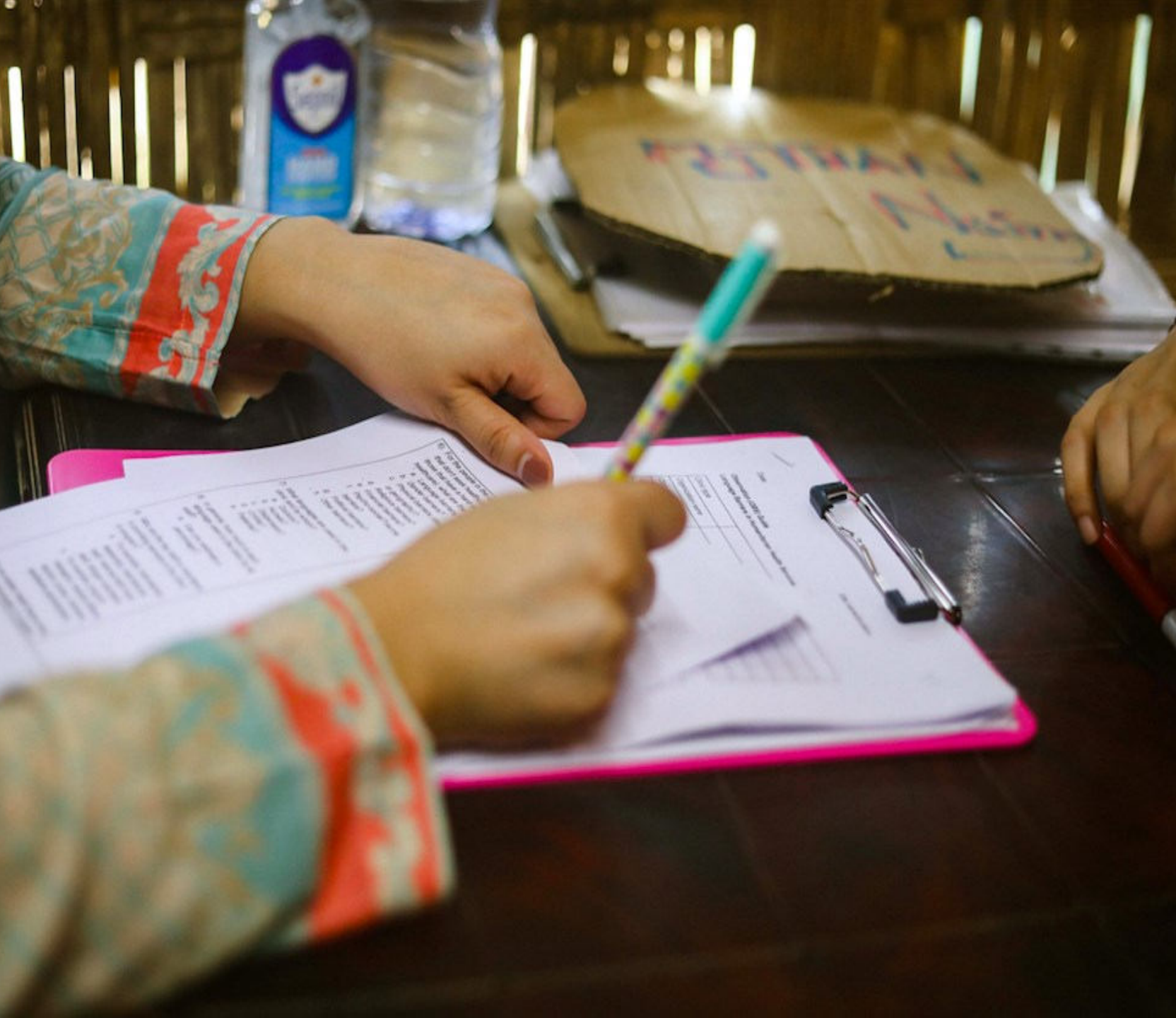
Helping people get vital information they need, and be heard, whatever language they speak.



## Over half of the world's population suffers from lack of access to information in their language

The world can be roughly divided into two big groups: the **3.5 billion** people who speak a 'major' language—like English, French, Spanish or Chinese—and the **4 billion** who don't.





Many of us who speak major languages (like English) are finding it easier and easier to navigate the world. When we need them, we have readily available translation apps like Google Translate.

**But for many people who speak underserved languages, getting information in a format that they can understand and communicate their needs is challenging.**

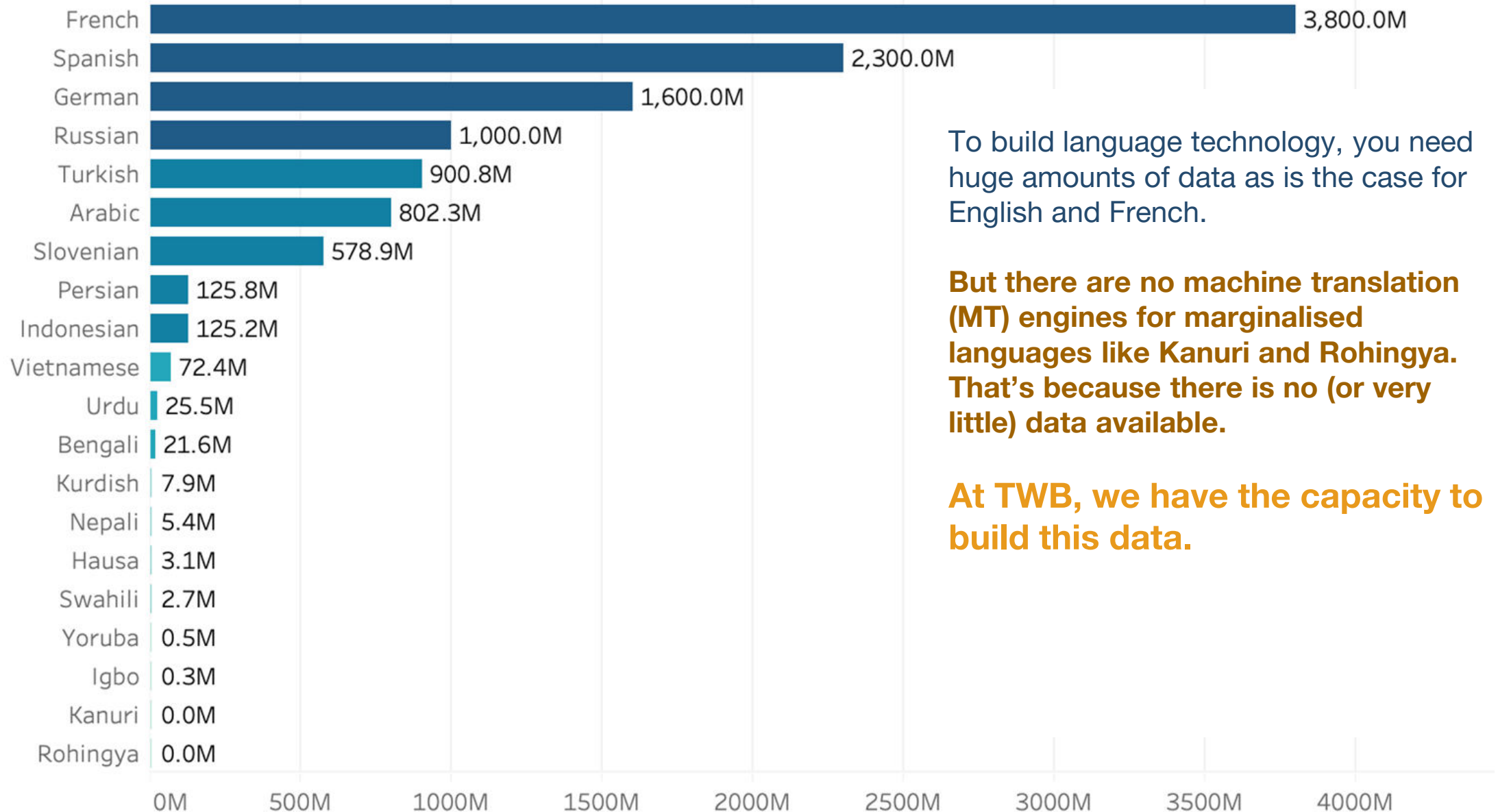
**It can be a matter of life and death.**



We deliver on our mission using a combination of **language technology**, innovative **language service platforms**, a dedicated staff of 120+ experienced professionals, and a community of 60,000+ linguists in 148 countries.



# Size of publicly available language datasets



To build language technology, you need huge amounts of data as is the case for English and French.

**But there are no machine translation (MT) engines for marginalised languages like Kanuri and Rohingya. That's because there is no (or very little) data available.**

**At TWB, we have the capacity to build this data.**

# Microsoft Introduces Inuktitut to Microsoft Translator

January 27, 2021



Afrikaans  
Albanian  
Amharic  
Arabic  
Armenian  
Assamese  
Azerbaijani  
Bangla  
Bosnian (Latin)  
Bulgarian  
Cantonese (Traditional)  
Catalan  
Chinese Simplified  
Chinese Traditional  
Croatian  
Czech  
Danish  
Dari  
Dutch  
English  
Estonian  
Fijian  
Filipino  
Finnish  
French  
French (Canadian)  
German  
Greek  
Gujarati  
Haitian Creole

Hebrew  
Hindi  
Hmong Daw  
Hungarian  
Icelandic  
Indonesian  
Inuktitut  
Irish  
Italian  
Japanese  
Kannada  
Kazakh  
Khmer  
Klingon  
Klingon (plqaD)  
Korean  
Kurdish (Central)  
Kurdish (Northern)  
Lao  
Latvian  
Lithuanian  
Malagasy  
Malay  
Malayalam  
Maltese  
Maori  
Marathi  
Myanmar  
Nepali  
Norwegian

Odia  
Pashto  
Persian  
Polish  
Portuguese (Brazil)  
Portuguese (Portugal)  
Punjabi  
Queretaro Otomi  
Romanian  
Russian  
Samoan  
Serbian (Cyrillic)  
Serbian (Latin)  
Slovak  
Slovenian  
Spanish  
Swahili  
Swedish  
Tahitian  
Tamil  
Telugu  
Tigrinya  
Thai  
Tongan  
Turkish  
Ukrainian  
Urdu  
Vietnamese  
Welsh  
Yucatec Maya

- 0 speakers

- 40k speakers

But not:

- Hausa 100m speakers
- Lingala 40m speakers
- Quechua 4.0m speakers
- K'iche' 2.5m speakers





**Working towards a world where  
knowledge knows no language  
barriers**

Digital Engagement needs to move from the equivalent of this...



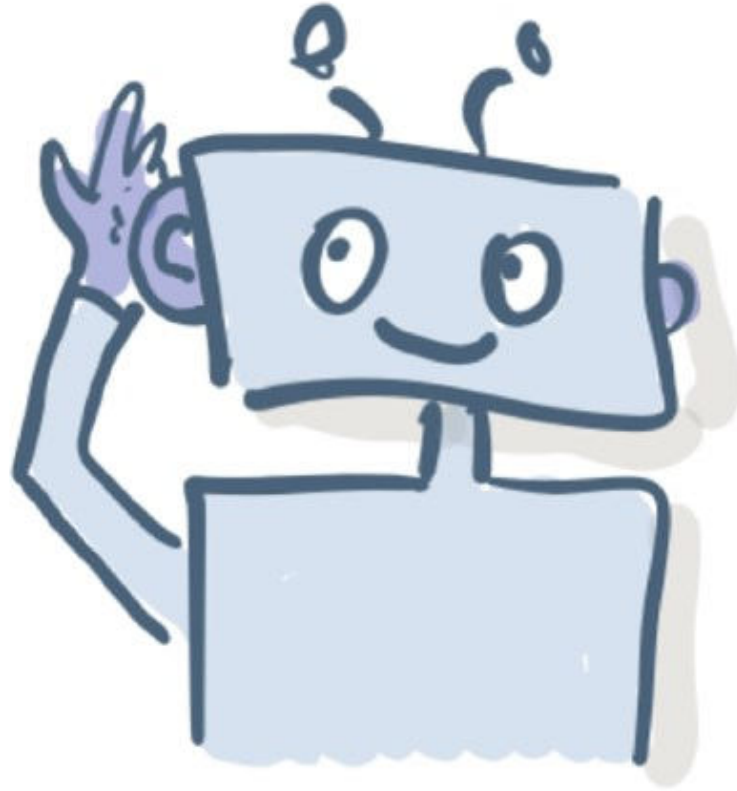


... to something more like this.





“I’m listening”



# “Fattening the tail”

An Incremental Approach to Low-Resource  
Language Support

# “Fattening the tail”

An Incremental Approach to Low-Resource  
Language Support

VOICE INPUT (ASR)

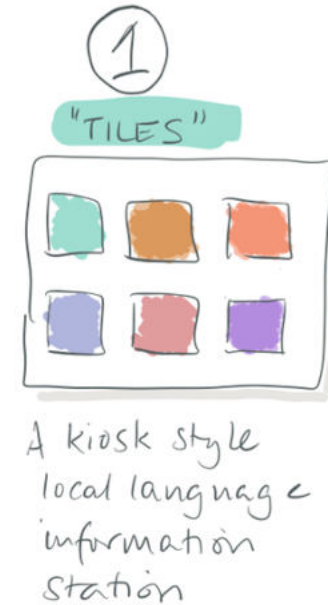
Diverse speaker  
language models  
for keyword spotting



# “Fattening the tail”

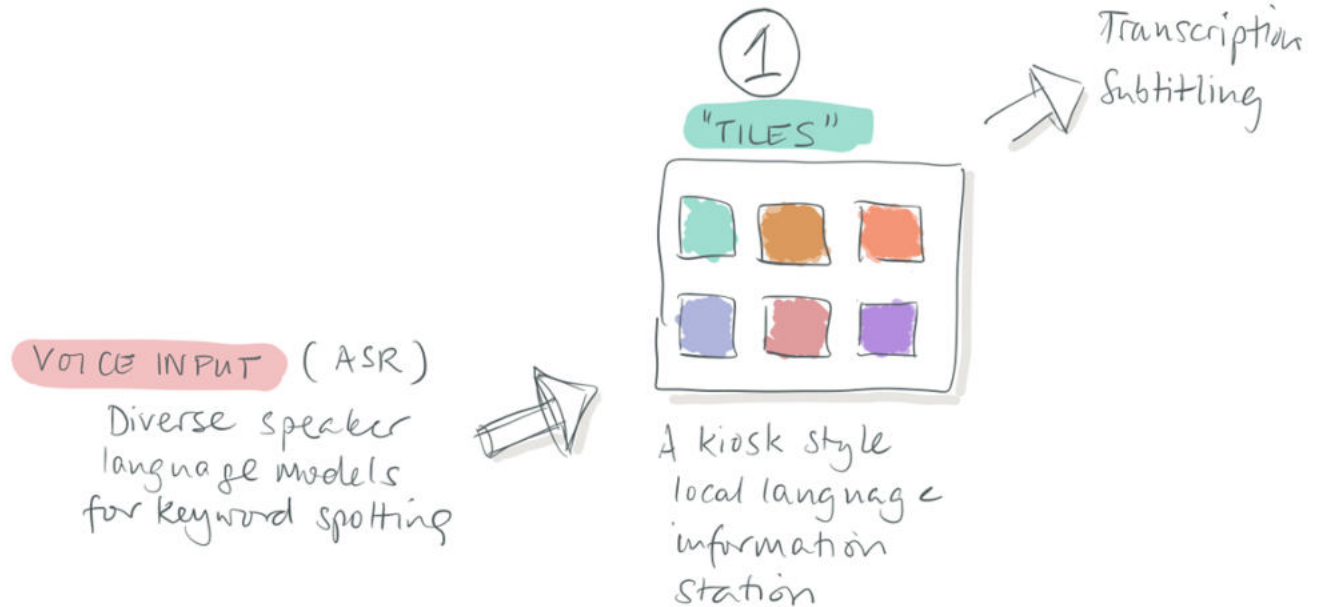
## An Incremental Approach to Low-Resource Language Support

VOICE INPUT (ASR)  
Diverse speaker  
language models  
for keyword spotting



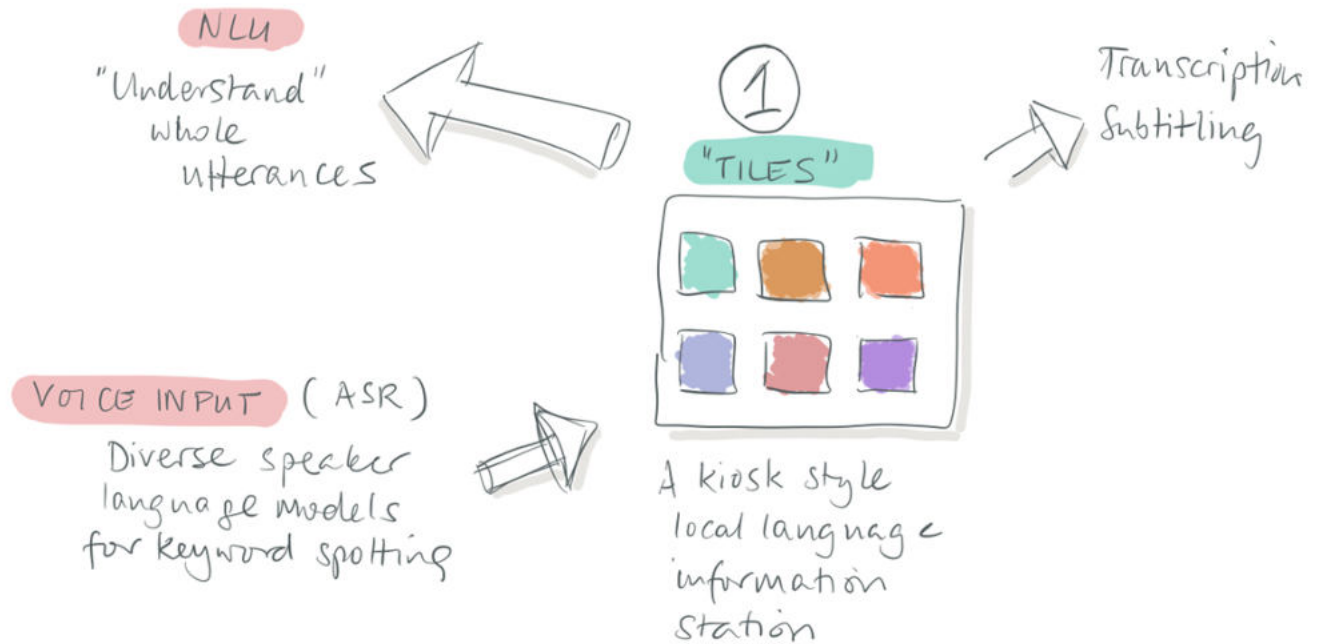
# “Fattening the tail”

## An Incremental Approach to Low-Resource Language Support



# "Fattening the tail"

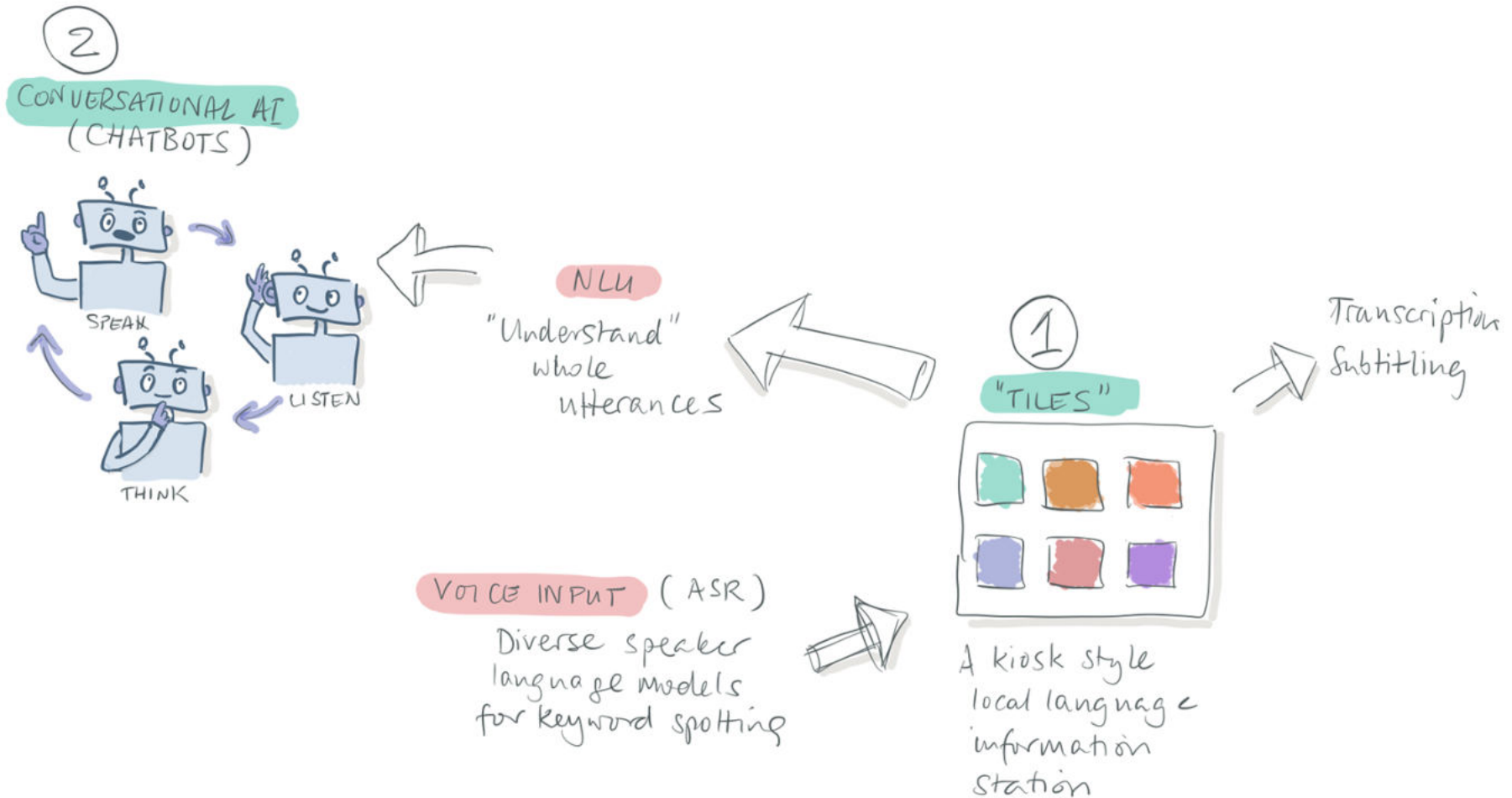
## An Incremental Approach to Low-Resource Language Support





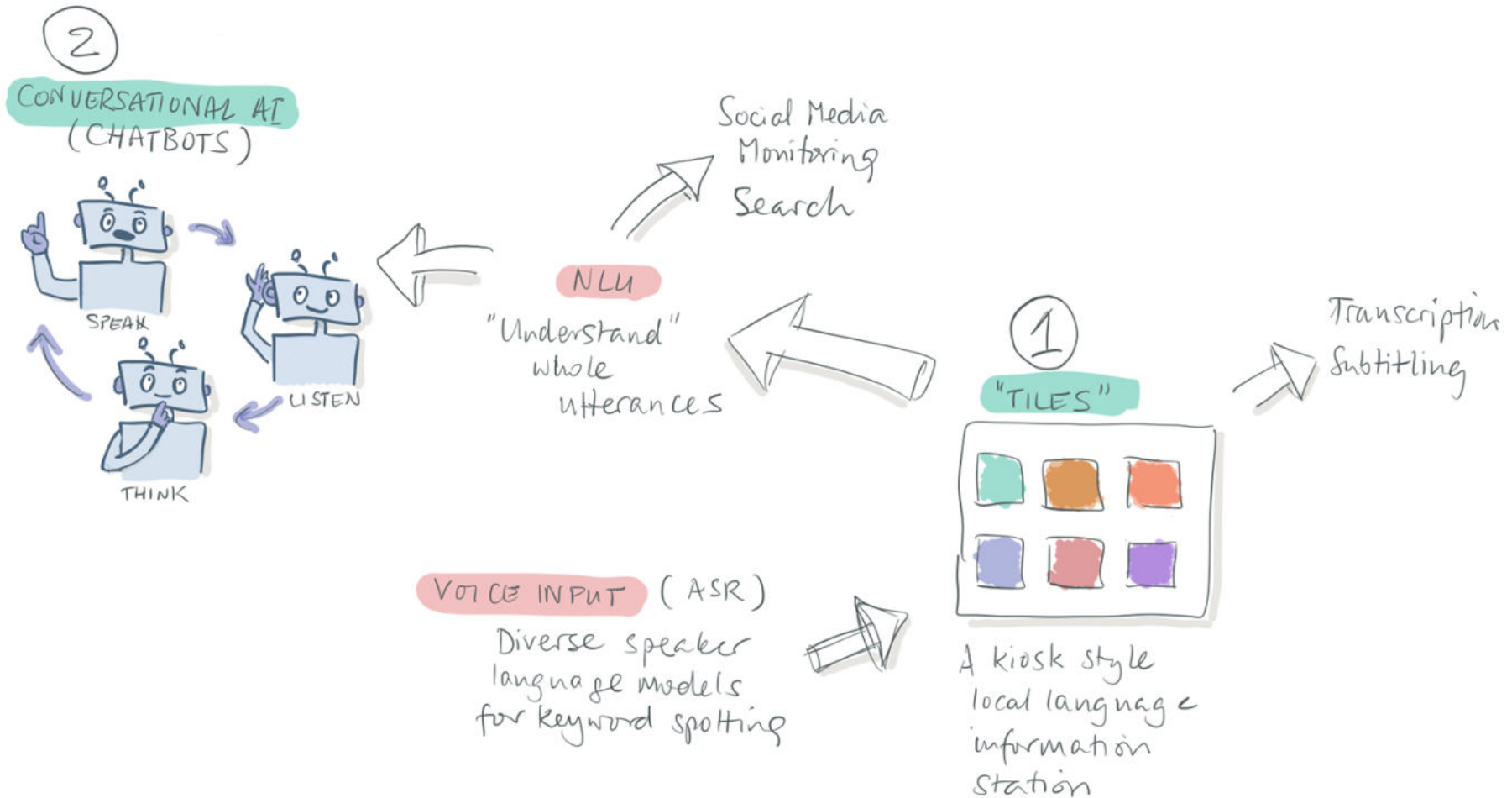
# "Fattening the tail"

## An Incremental Approach to Low-Resource Language Support



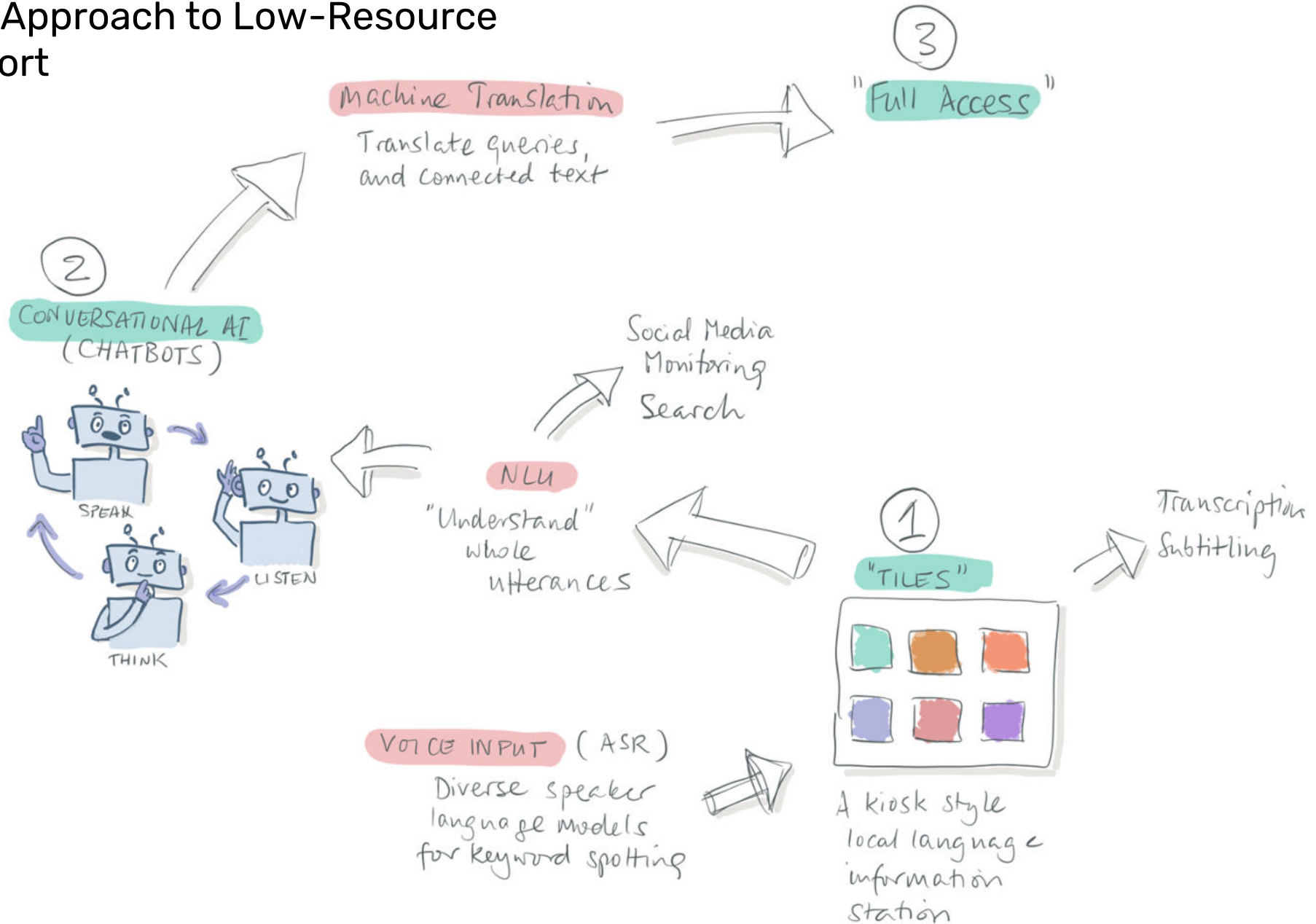
# "Fattening the tail"

## An Incremental Approach to Low-Resource Language Support



# "Fattening the tail"

## An Incremental Approach to Low-Resource Language Support





# TRANSLATORS WITHOUT BORDERS