

# INTERACT Deliverable 8.3: Panel on Crisis MT

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement No 734211.

Attendance at the Crisis MT Workshop 2019 has been sponsored by the INTERACT project, with funding from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska Curie grant agreement No 734211.



**TRANSLATORS  
WITHOUT BORDERS**  
#LanguageMatters

A light gray world map is shown in the background, centered on the Atlantic Ocean. The map includes outlines of all major landmasses and country borders.

# 39

**Countries currently experiencing  
humanitarian emergencies.**



The background of the entire image is filled with a repeating pattern of grey human icons. These icons are arranged in a grid-like fashion, with some appearing as solid grey and others as outlines, creating a textured, population-like background.

# 175,843,646

**People in need of humanitarian assistance**

# How many languages are spoken in those countries?

# 3,136

# How many of these are 'high resource' languages with openly available MT?

# 10-20

MALAWI  
HEALTH PASSPORT  
WOMAN HEALTH PROFILE

NAME                     

Dzina / Zina

DATE OF BIRTH

Tsikulu Lobadwa

VILLAGE

Mudzi / Kuka

Please bring book

Chonde bweretsa

Chonde yeghanu

Issuing per

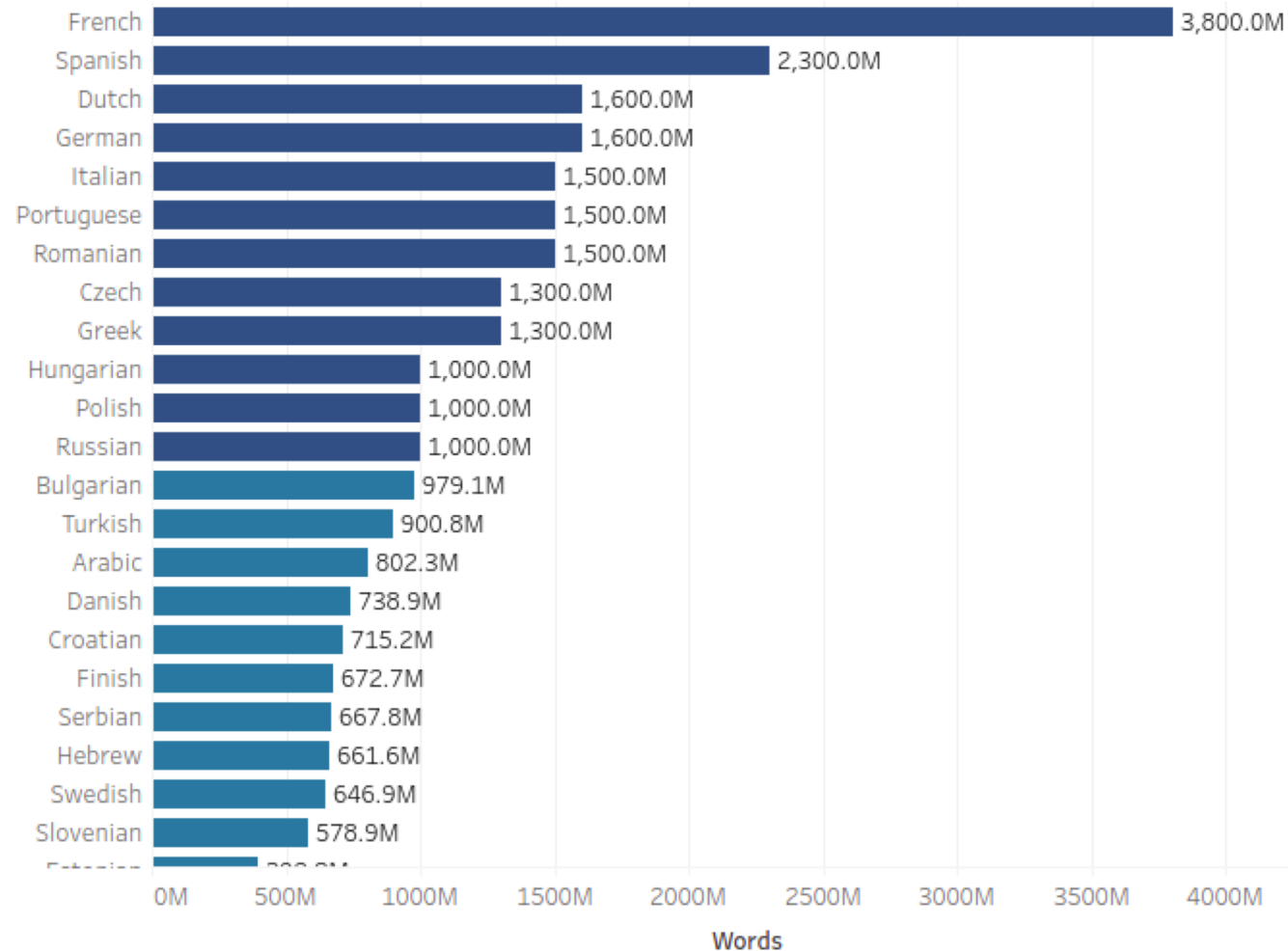
**What *is* crisis MT?**

# Initial response

- **Opportunisticly develop individual solutions**
- **Opportunistic in terms of availability of data as well as tech**

# Available data

Size of publicly available language datasets



full list at  
[https://public.tableau.com/viz/Sizeofpublicallyavailablelanguage/datasets/Dashboard1?embed=y&:display\\_count=yes&:origin=viz\\_share\\_link](https://public.tableau.com/viz/Sizeofpublicallyavailablelanguage/datasets/Dashboard1?embed=y&:display_count=yes&:origin=viz_share_link)

# At TWB, crisis MT

- **applies to marginalized languages, ie. languages spoken by people whose marginalization is at least partially due to sociolinguistic factors**
- **is not strictly limited to MT, not the least because of literacy implications and data situation**
- **is characterized by lack of data (the vicious investment circle)**

# Gamayun

## **The TWB language equality initiative**

- **Unified purpose-built data generation strategy**
- **Showcase**
- **Explore crosslingual and multilingual strategies**
- **Include voice technology from the outset**
- **Language variants, a universal challenge**

# Data collection

**Leverage and adapt our existing stack of translation tools and processes**

- **Kató**
- **Kató Speak**
- **Kató To Go**
- **TWB audio glossaries**



## Use cases

**As we're making progress, the crucial importance of actual use cases has become apparent**



**TRANSLATORS**  
WITHOUT BORDERS

**#LanguageMatters**