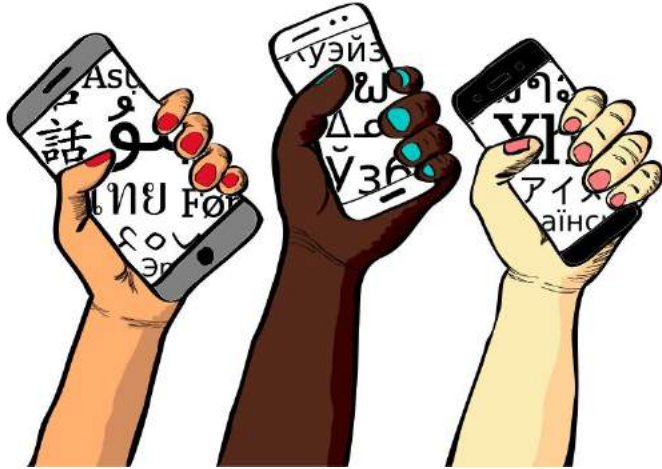


DECOLONIZING the Internet's Languages



...and questions of
epistemic (in)justice

Anasuya Sengupta

June 2020



What do we care about?

- 130 million books have been published in only about 480 languages (of the 7000+ languages that exist).
- 7% of the world's languages are captured in published material, and an even smaller fraction digitised.



What do we care about?

- **59% of the world is online**
- **75% of those online are from the Global South**
- **45% of all women are online**
- **10 of the 7000+ world languages make up 84.3 percent of people using the Internet (English and Chinese dominate)**



Why do we care?

The internet does not look or sound like most of us in the world.

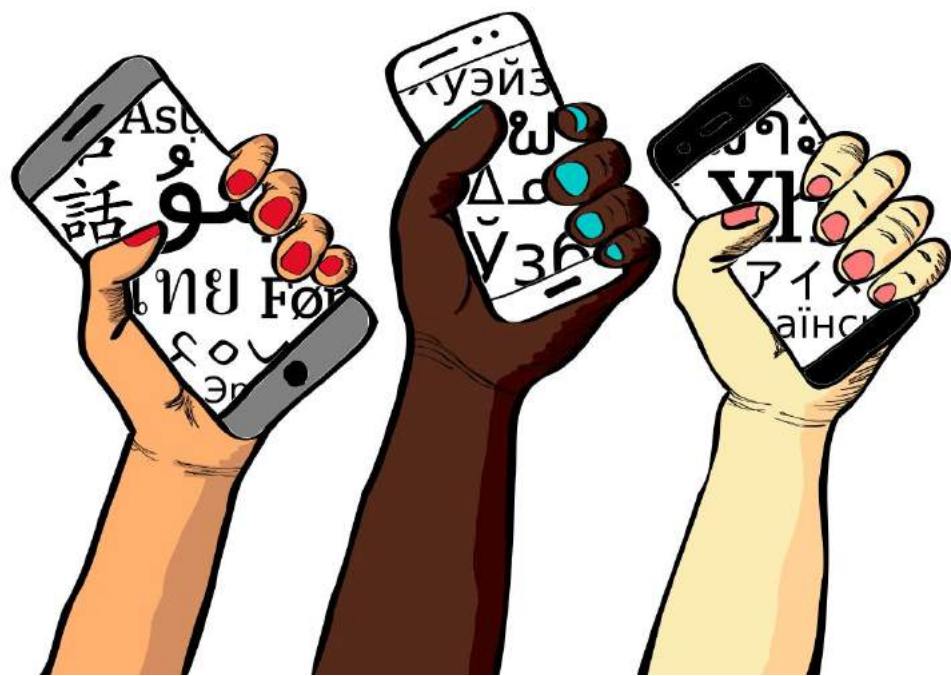


DEcolonize
theINTERNET





What does decolonizing the internet's languages mean?



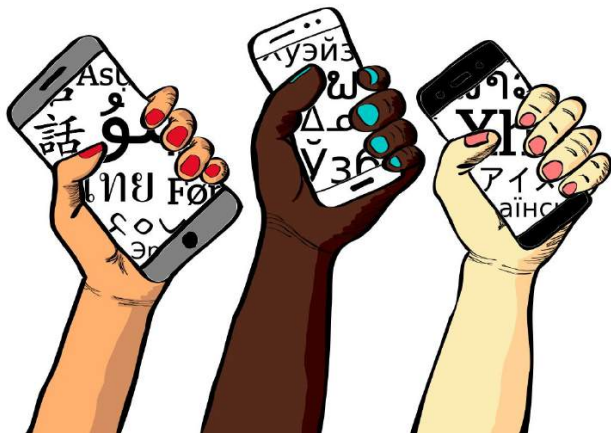
- Seeing language as a proxy for knowledge
- Creating a truly multilingual internet
- Embodying multiple forms of knowledge



The State of the Internet's Languages report (later this year!)

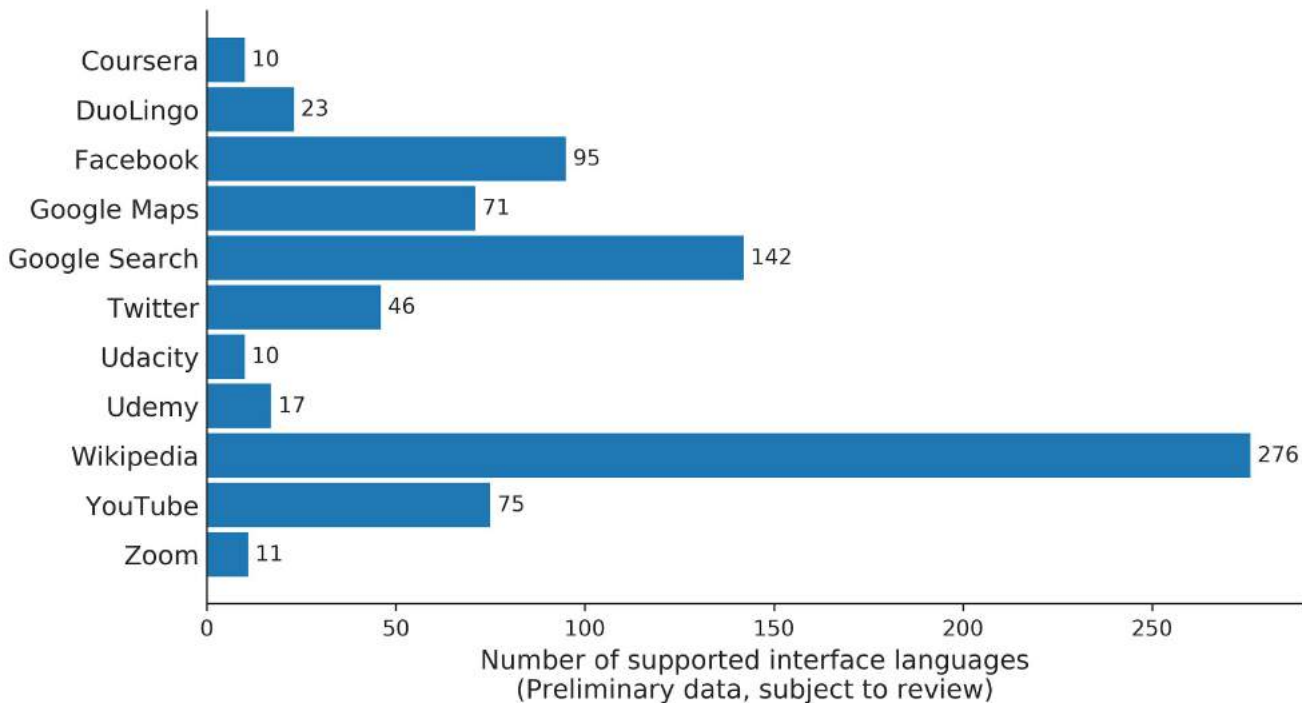


- Showing numbers and experience
- Mapping challenges and opportunities
- Suggesting a practical agenda(s) for action



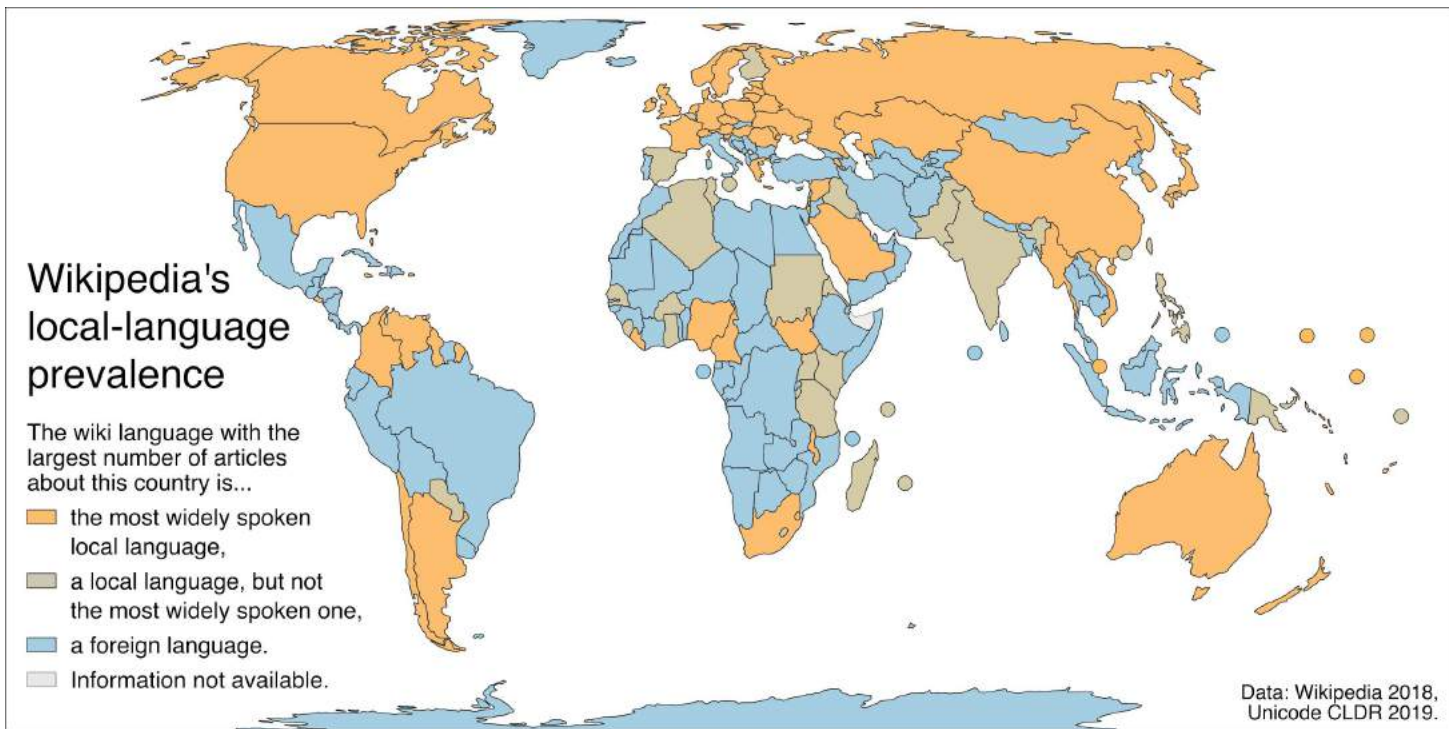


The internet's languages



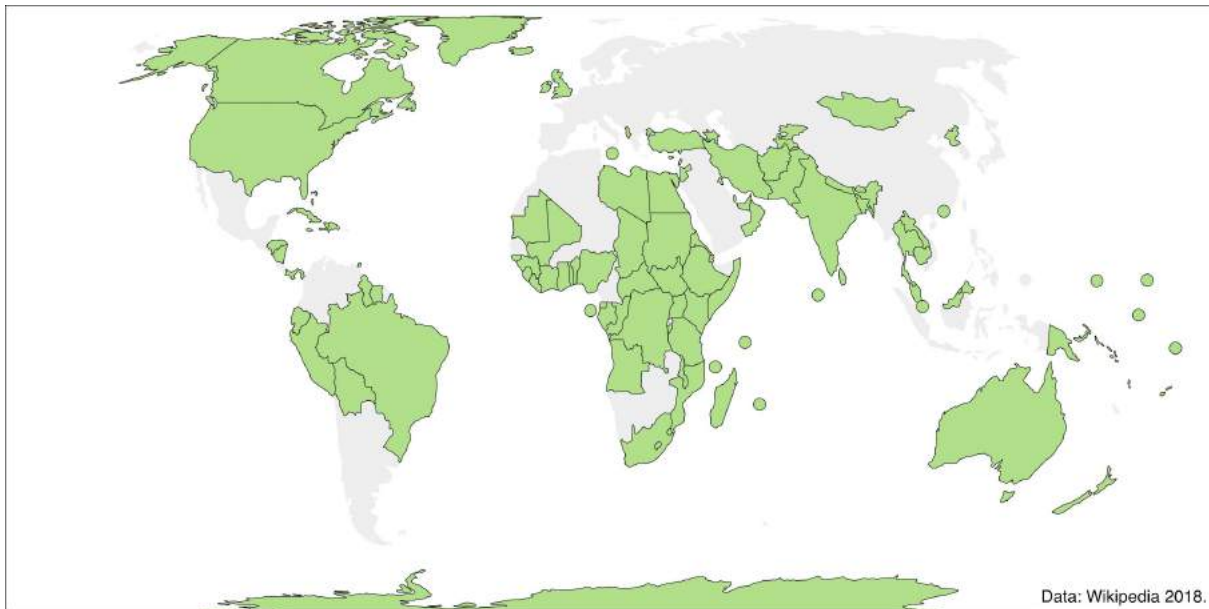


Wikipedia's languages





Wikipedia's language



Countries where most Wikipedia articles about local places are in **English**



Unicode's language

While almost all [Court] statutes state that when there is inconsistency between the different versions, the Sinhala text shall prevail, the Sinhala text is rarely available online.

Sri Lankan conversations on human rights and democracy often take place in English.

Unicode Sinhala does not allow you to insert combined letters



A Sinhala letter that is becoming non-existent as no keyboards recognise it



(dā)



**People
are
Knowledge**

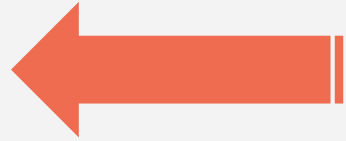
Image by Alamcsd at wts wikivoyage (Own work) [Public domain], via Wikimedia Commons

Epistemic Injustice



Testimonial

Black Lives Matter [logo](#) PD, via Wikimedia Commons



Hermeneutical

By [Marc Nozell](#) CC BY 2.0, via Wikimedia Commons



Whose language? Whose knowledge?

- Can you only be on the internet in your nearest colonial language?
- ...How much difference (and different knowledges) is being flattened out, lost, or not preserved?

whoseknowledge.org



Adele Vrana

Anasuya Sengupta

Siko Bouterse