

Beyond English

Decolonising LLMs for Global
Justice

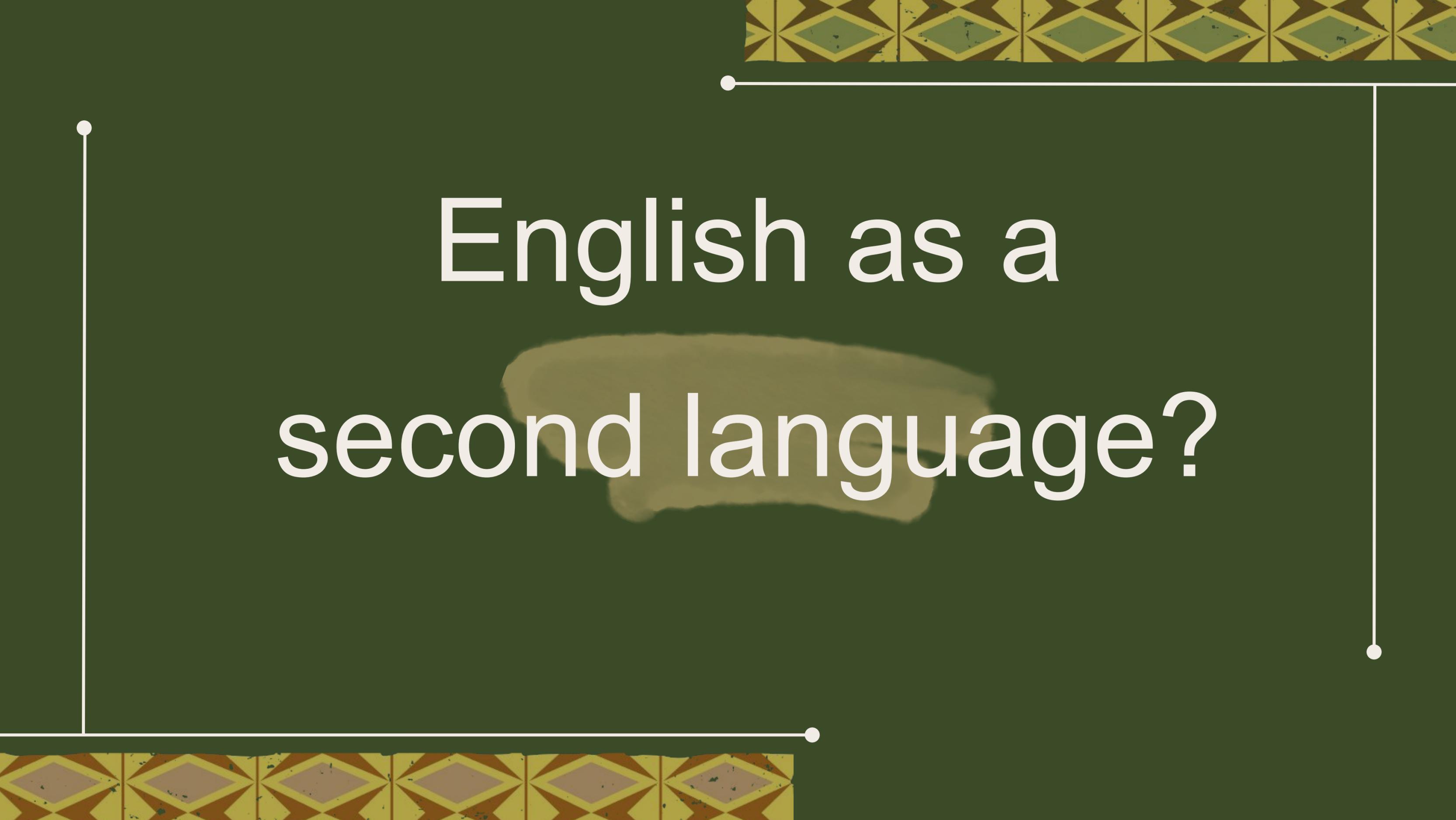
Hannah Claus

Introduction

- PhD student, Language Technology Lab
- DeepMind & Gates Cambridge Scholar
- Background in AI, Robotics, NLP, Ethics
- Worked with the German Aerospace Centre and the Ada Lovelace Institute
- Mission: build AI/Robots that work for everyone

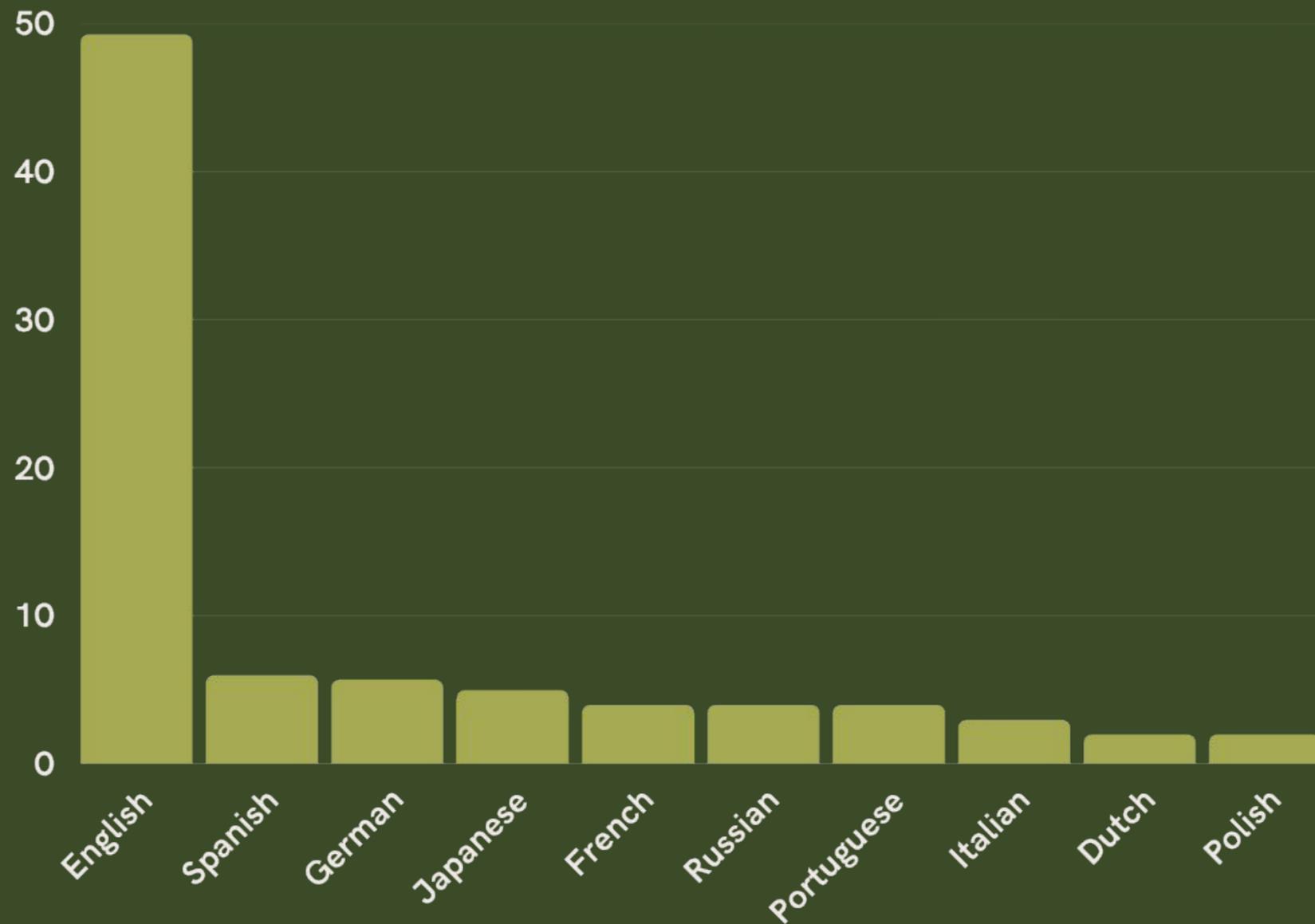


Hannah Claus



English as a
second language?

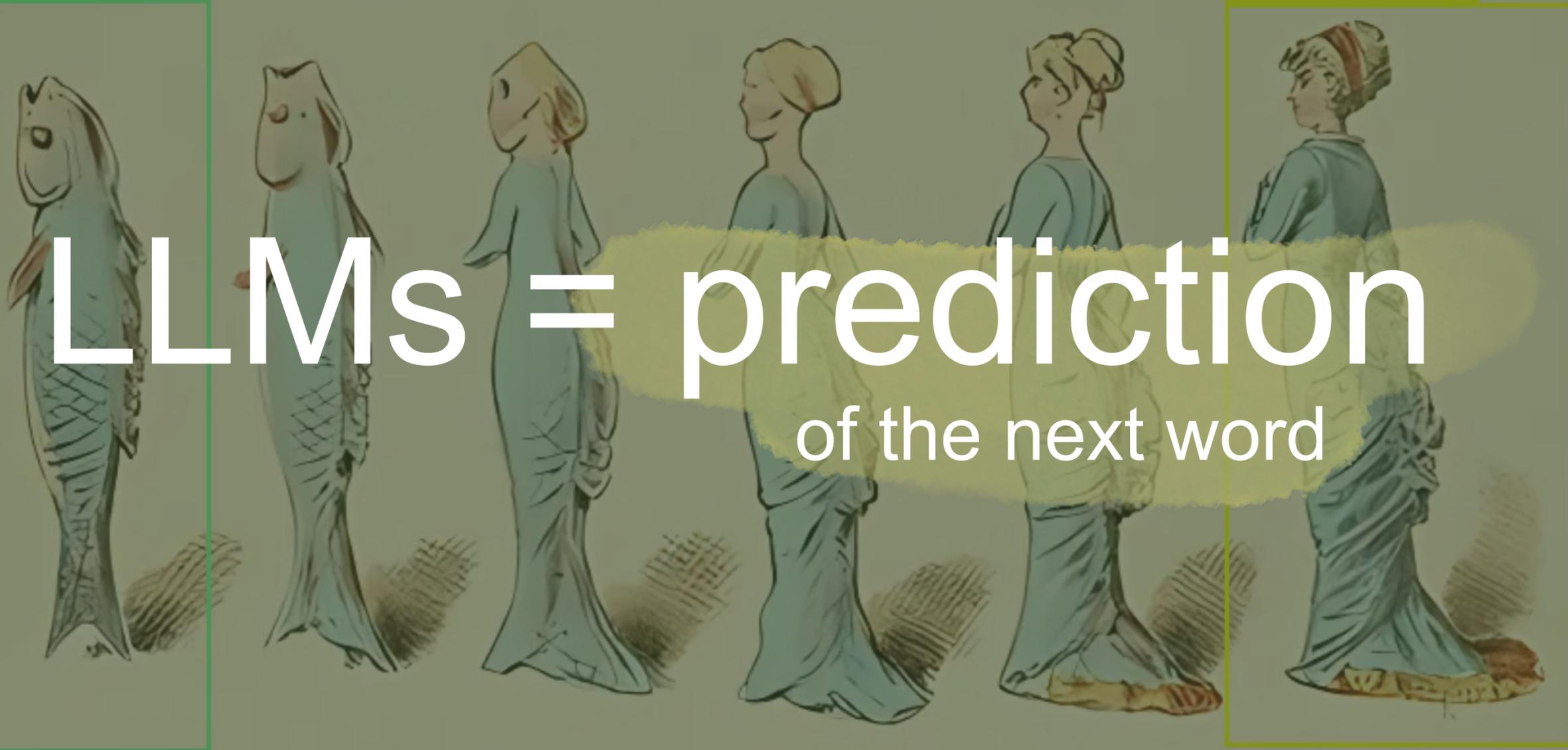
Data



- There are **7,000+** languages worldwide
- Half of the internet's content is in **English**
- But English is only spoken by **1.5 billion** people
- With only **26%** being native speakers

fish

woman



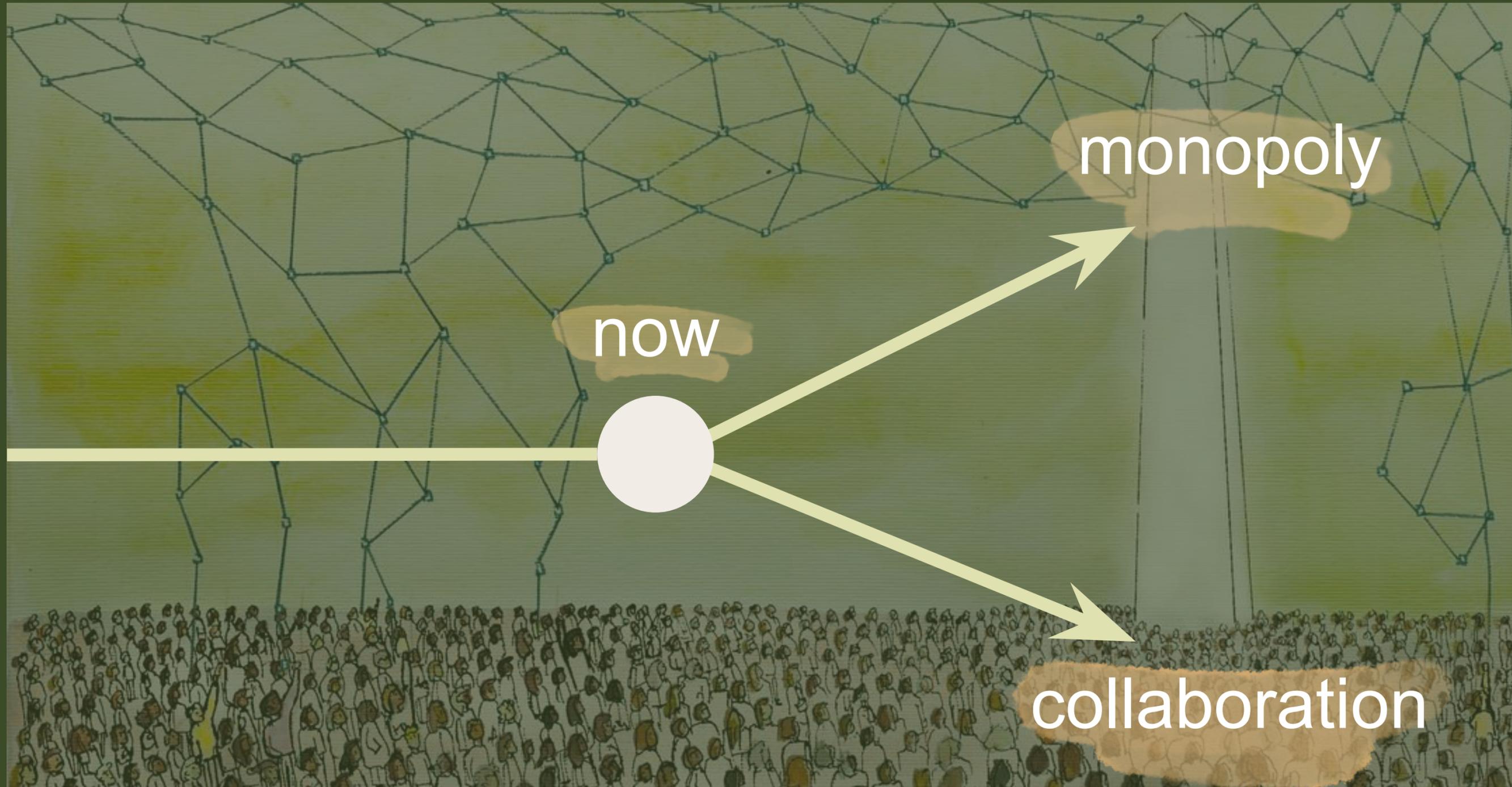
LLMs = prediction
of the next word

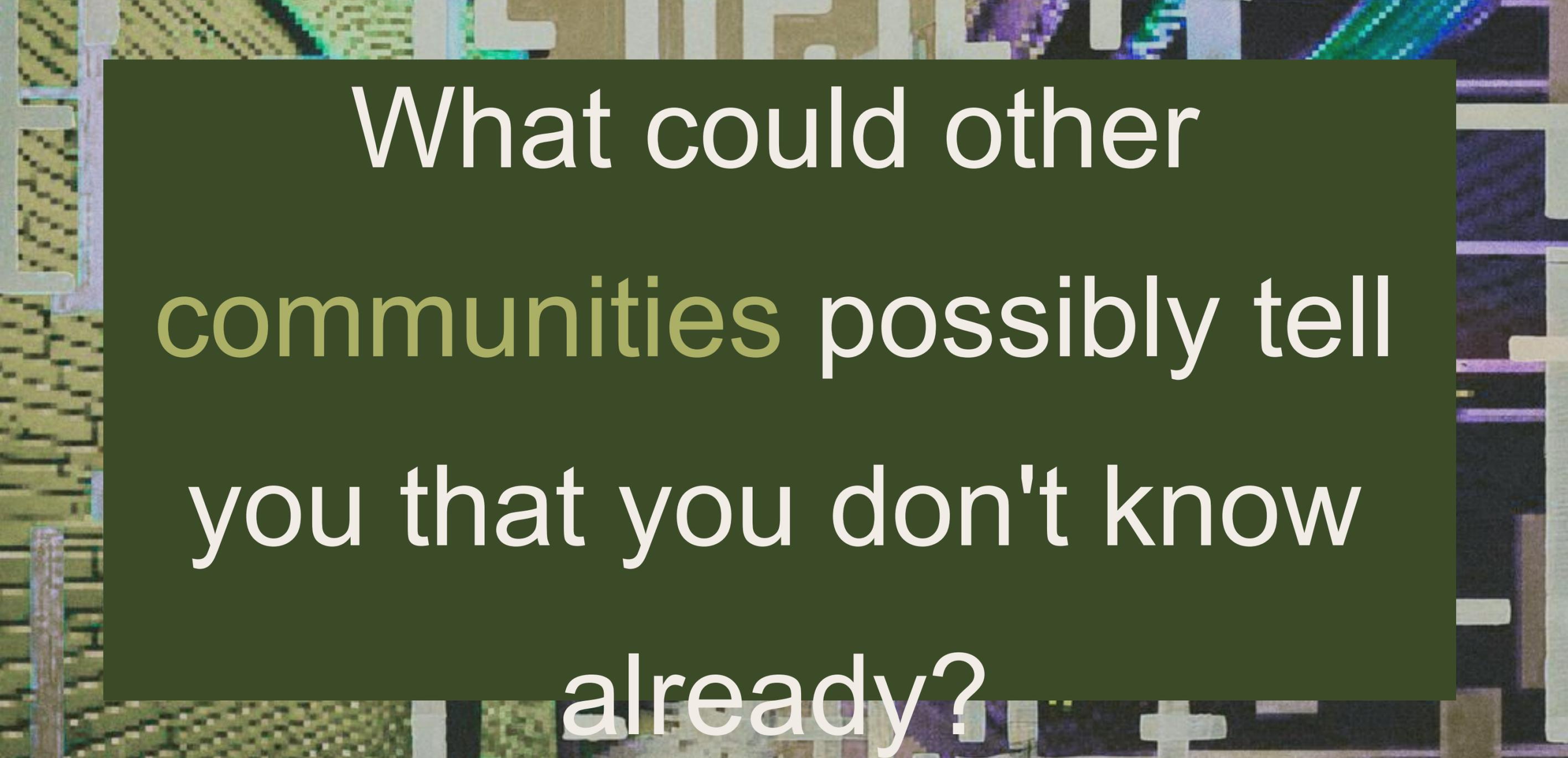
Basque vs Swahili



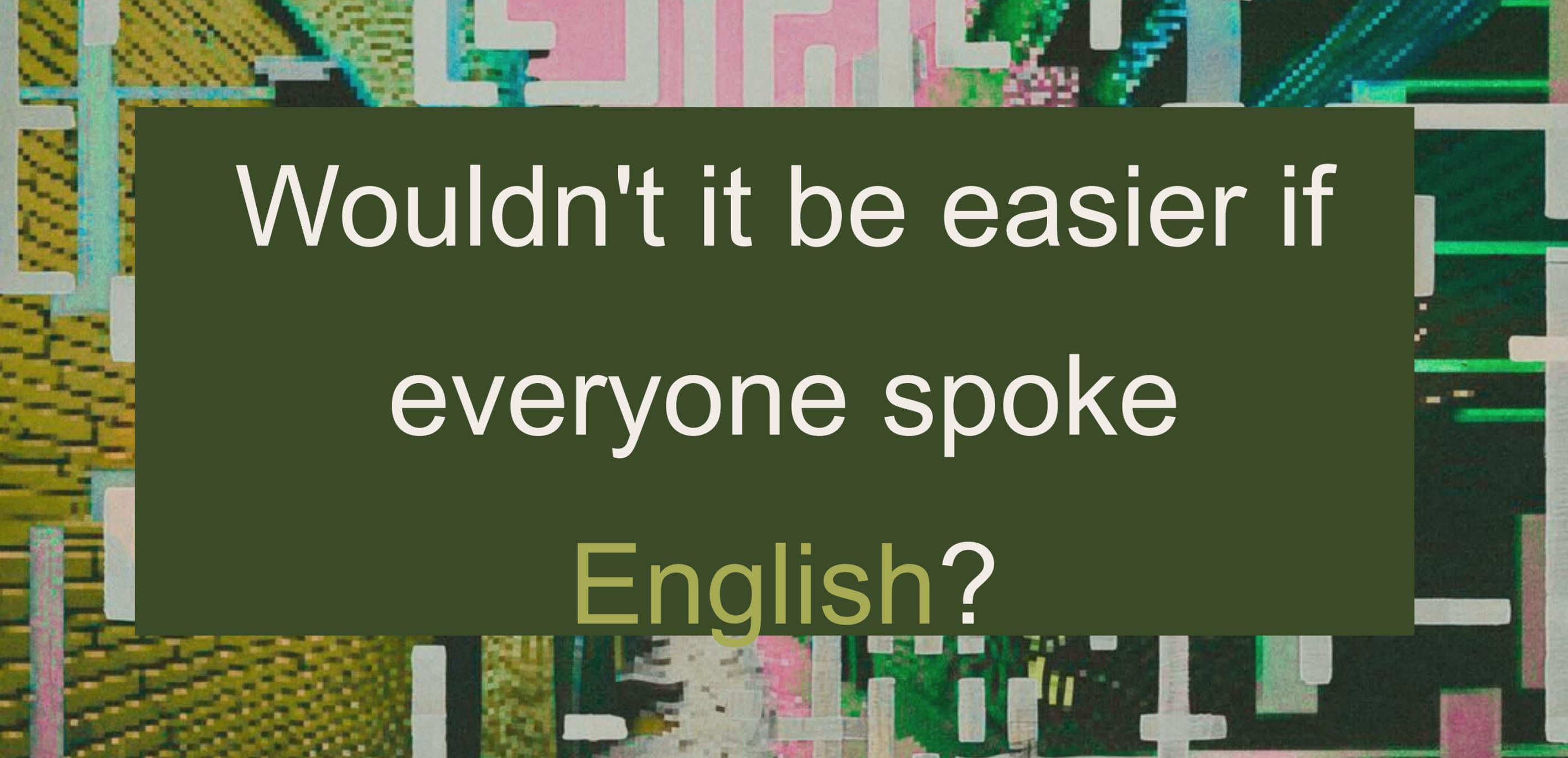


language = identity

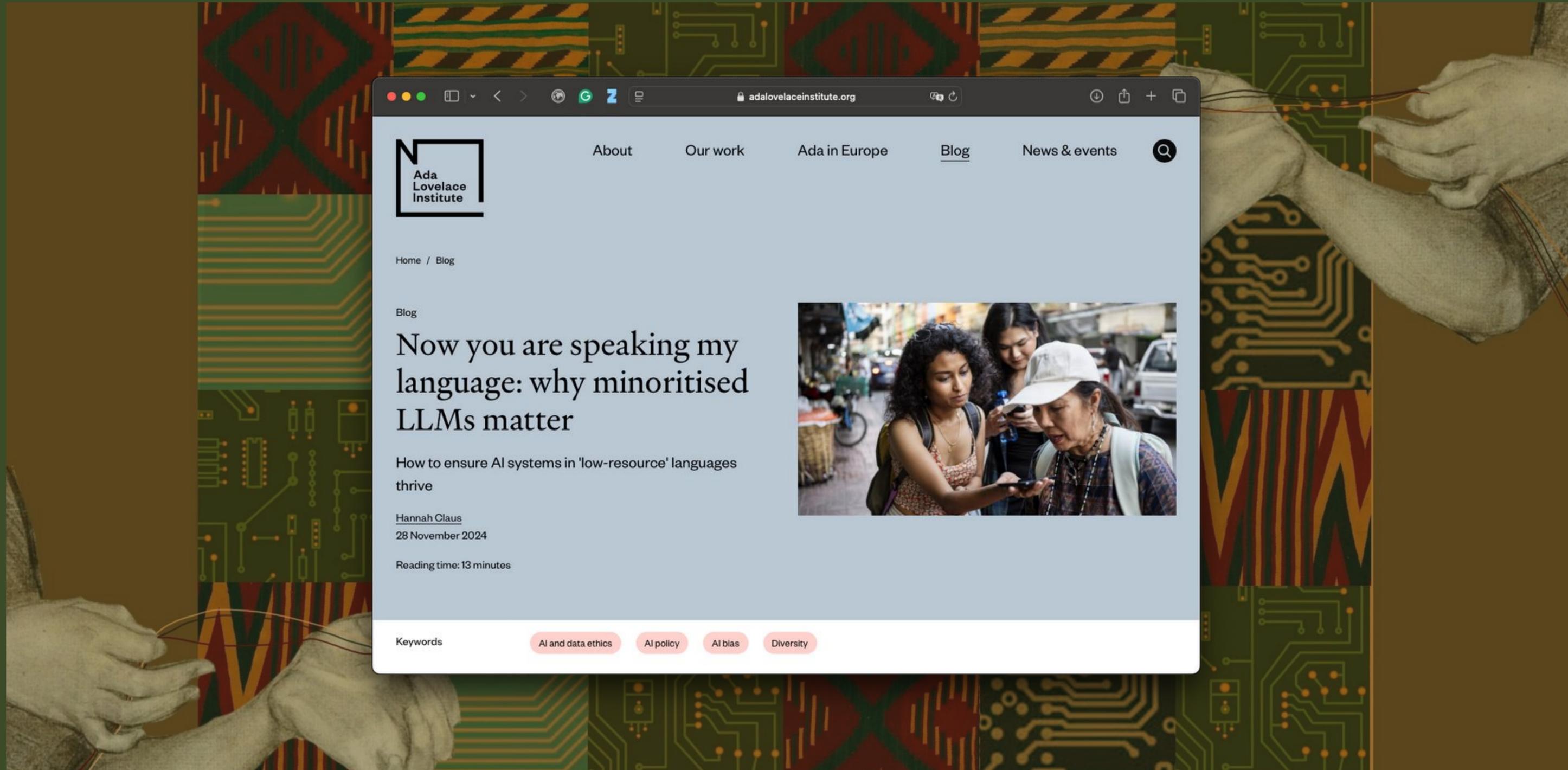




What could other
communities possibly tell
you that you don't know
already?



Wouldn't it be easier if
everyone spoke
English?





Masakhane

A grassroots NLP community for Africa, by Africans



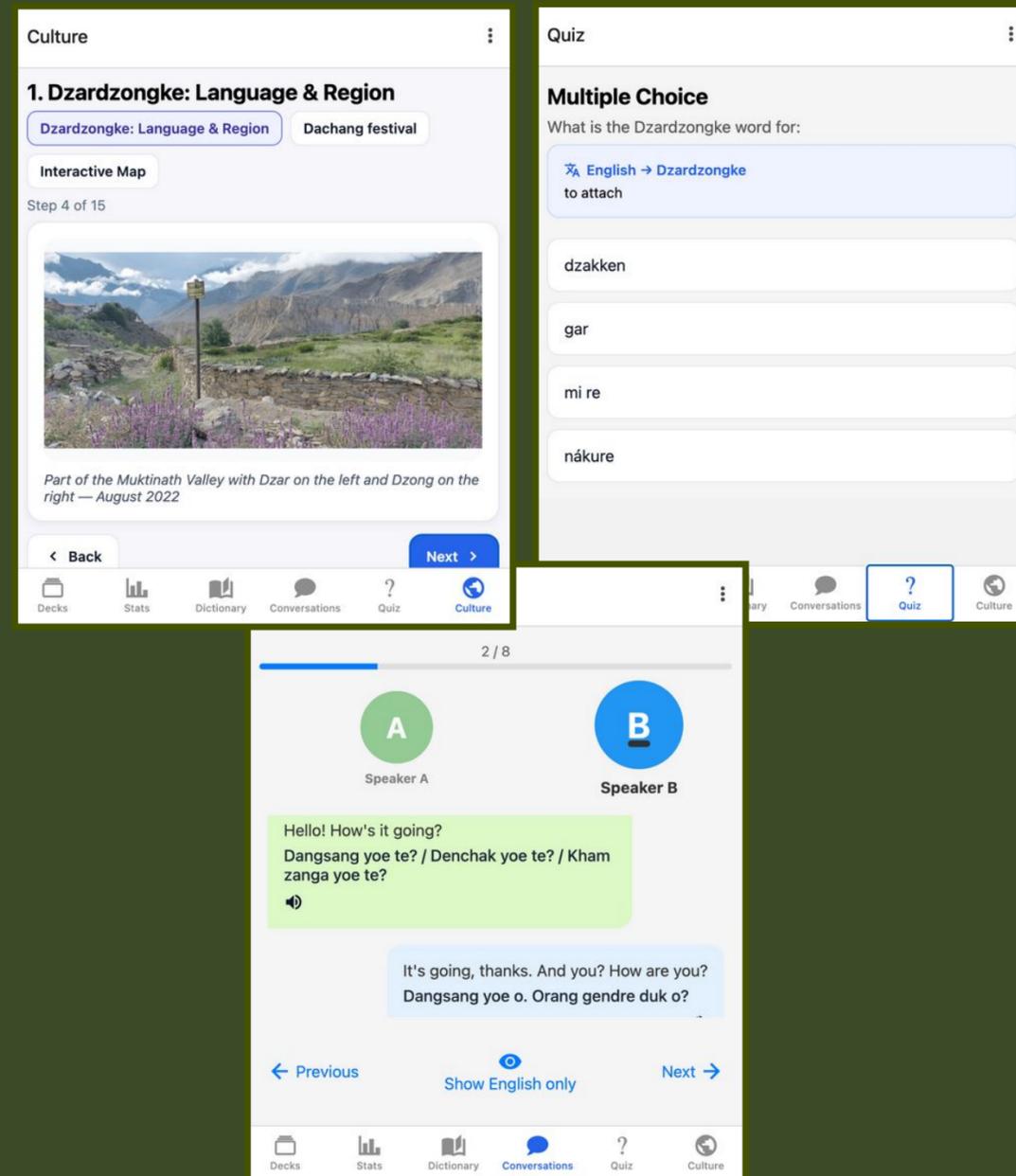
LELAPA AI



DEEP LEARNING INDABA

Case Study

- **Dzardzongke: an oral language** with ~3,000 speakers
- Spoken in a village in **Nepal**
- **No written learning resources** or standardised writing system
- Community interest in **preserving language** and developing literacy tools
- Limited digital resources and **low AI readiness** in local context



- Create a **standardised writing system**
- Co-design a **language learning app** with the community
- Support Dzardzongke speakers in learning to **read and write** their language
- Develop **language models** in Dzardzongke
- Build **NLP applications** benefiting the community

Q&A



Hannah Claus

Let's connect
on LinkedIn

