

University
of Essex



THE UNIVERSITY
of EDINBURGH

Language Technologies as if People Mattered: Centering Language Communities in Speech Technology development

Nina Markl

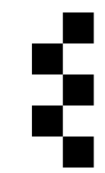
University of Essex

Catherine Lai

Lauren Hall-Lew

University of Edinburgh

LREC-Coling, 2024



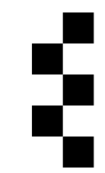
Attending to big questions

normative questions can be addressed through an interdisciplinary lens

sociolinguistic theory

“AI ethics” and science and technology studies

input from language communities (the most important part)



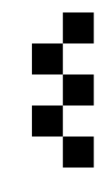
Attending to big questions

big questions

how should we develop language technologies that work for everyone?

who should develop language technologies?

at what costs should we develop language technologies?



Attending to big questions

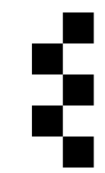
ethical, social and legal questions in AI development have no easy answer

trade-offs between privacy and accessibility

trade-offs between accuracy and surveillance

trade-offs between convenience and data sovereignty

trade-offs between varieties, speech styles and communities



Language Technologies for All?

LTs tend to work better for (high status) varieties of high-status speakers

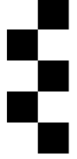
global inequity across languages

local inequity within languages

LTs are expanding to new social, cultural and linguistic contexts

high-stakes contexts (hiring, policing)

“under-resourced” varieties



Language Technologies by Whom?

DONAVYN COFFEY

SECURITY 28.04.2021 06:00 AM

Māori are trying to save their language from Big Tech

Te Hiku Media gathered huge swathes of Māori language data. Corporates are now trying to get the rights to it

The Masakhane project wants machine translation and AI to transform Africa

Language as...

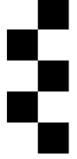
Data?

Capital?

Commodity?

Heritage?

Patrimony?



Language Technologies at what cost?

Artificial intelligence (AI)

'It's destroyed me completely': Kenyan moderators decry toll of training of AI models

Employees describe the psychological trauma of reading and viewing graphic content, low pay and abrupt dismissals

Niamh Rowe

Wed 2 Aug 2023 16:00 BST

A new front in the water wars: Your internet use

In the American West, data centers are clashing with local communities that want to preserve water amid drought

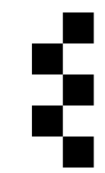


By [Shannon Osaka](#)

April 25, 2023 at 6:30 a.m. EDT

Extraction of...

- Labour?
- Energy?
- Minerals?
- Water?



Attention, not quick solutions

Big questions cannot be solved definitively or alone

taking an ethical stance

working together

shifting research cultures