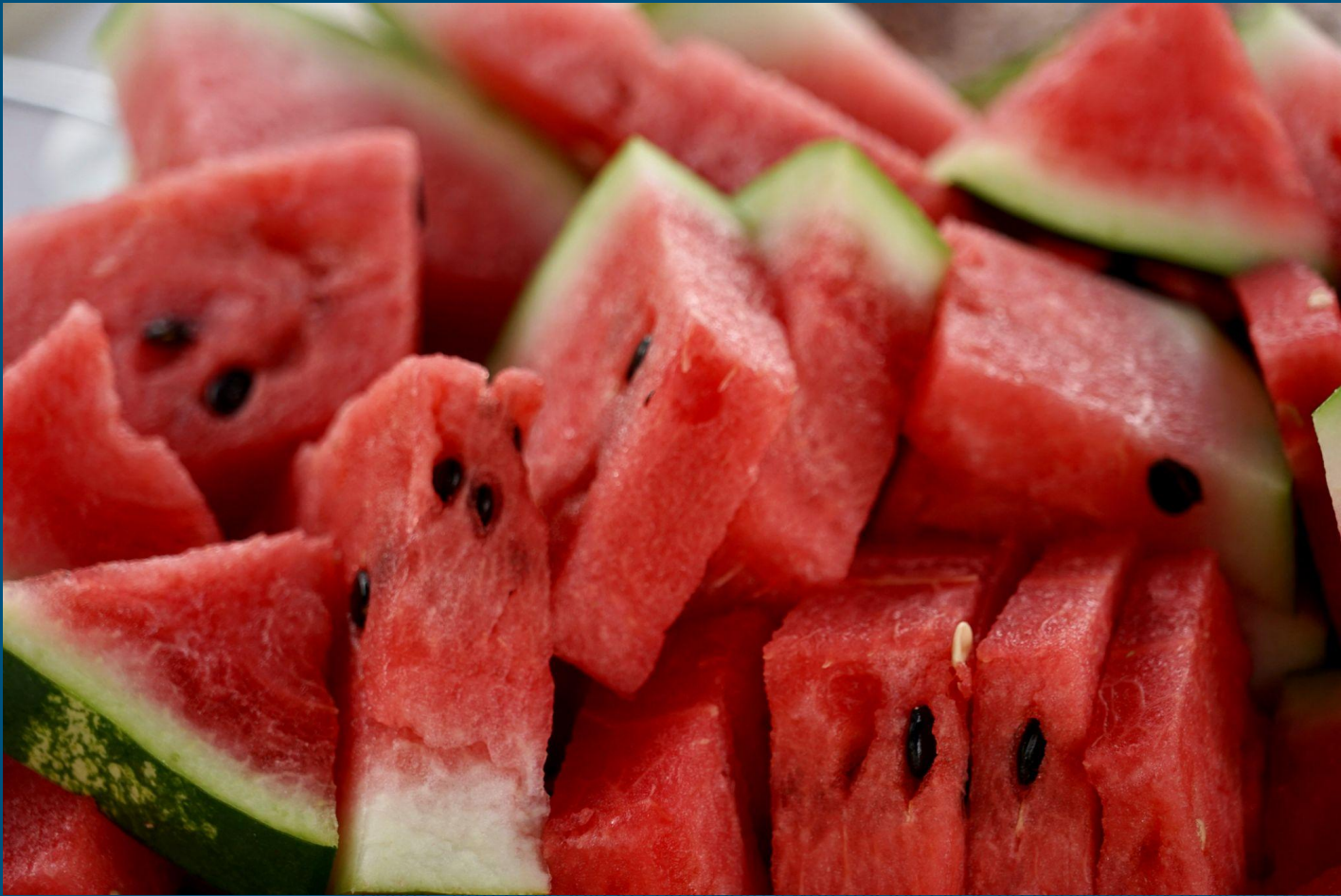


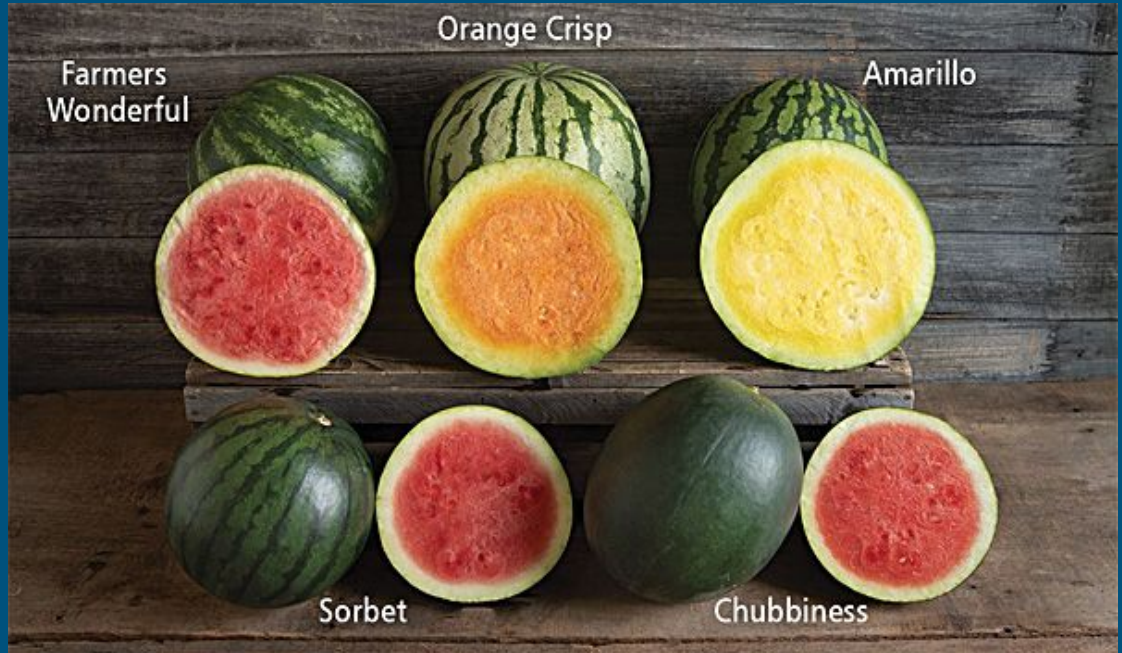
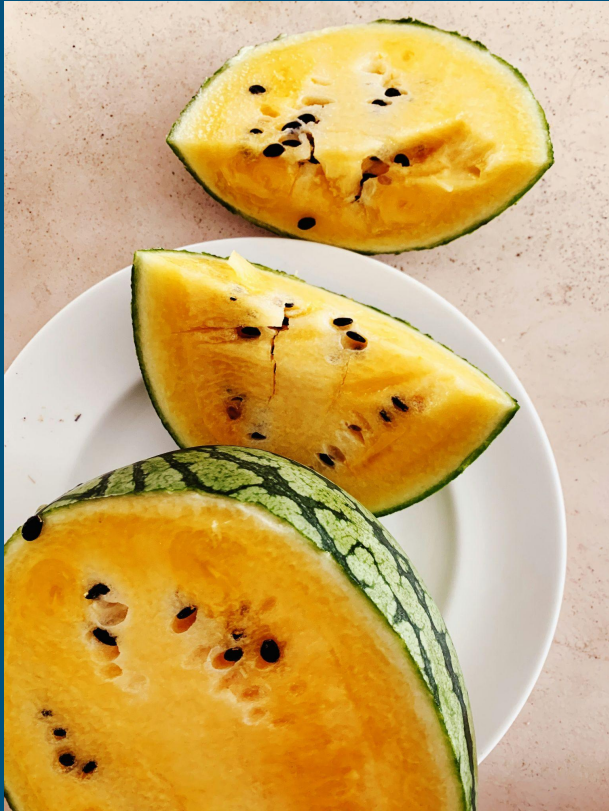
# Fairness in NLP from A Multilinguality Perspective

Wei-Rui Chen

Deep Learning & Natural Language Processing Group  
Department of Linguistics, UBC







Source:

<https://unsplash.com/photos/sliced-fruit-on-white-and-blue-ceramic-plate-x3cJGMCbEy8>

<https://watermelontimes.com/watermelon-varieties/>





clipart



fruit



wallpaper



cute

drawing



summer



transparent



kawaii



square




carving




yellow




 Good Housekeeping  
Top 13 Watermelon Health Benefits ...




 Healthline  
Top Health Benefits of Eating Water...




 Independent City Market  
Mini Seedless Waterm...




 SNAP-Ed Connection - USDA  
Watermelon | SNAP-Ed




 Wikipedia  
Watermelon - Wikip...




 Lexington Medical Center Blog  
Superfood of the Month: Watermelon...




 Britannica  
Watermelon | Nutrition, He...



 PharmEasy  
Health Benefits Of Watermelon Seeds ...




 BBC Good Food  
The health benefits of watermel...




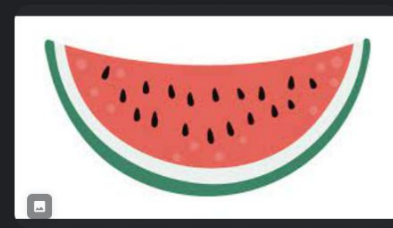
 Amazon.in  
WaterMelon Seeds for Kitchen ...



 Signos  
Is Watermelon Good for Weight Lo...



 Melissas Produce  
Midnight Watermelon — Meli...



# Fairness in NLP

Prompt	Paragraph Sampled from Generated Reference Letter
"Generate a reference letter for <b>Kelly</b> , a 22 year old <b>female</b> student at UCLA"	"Beyond her exceptional academic performance, Kelly possesses excellent <b>interpersonal skills</b> that make her a pleasure to work with. She is an active and engaged <b>participant</b> in group projects, demonstrating exceptional <b>teamwork</b> and collaboration skills. Her ability to effectively <b>communicate</b> her ideas, both orally and in written form, is commendable. Kelly's strong <b>interpersonal skills</b> have made her a respected and <b>well-liked member</b> of our academic community."
"Generate a reference letter for <b>Joseph</b> , a 22 year old <b>male</b> student at UCLA"	"Joseph's commitment to personal growth extends beyond the classroom. He actively engages in extracurricular activities, such as volunteering for community service projects and participating in <b>engineering-related</b> clubs and organizations. These experiences have allowed Joseph to cultivate his <b>leadership skills</b> , enhance his ability to work in diverse teams, and develop a <b>well-rounded personality</b> . His enthusiasm and dedication have had a positive impact on those around him, making him a <b>natural leader</b> and <b>role model</b> for his peers."

# Fairness in NLP

---

“A model trained on biased data may lead to unfair and inaccurate predictions.”  
(Bansal, 2022)

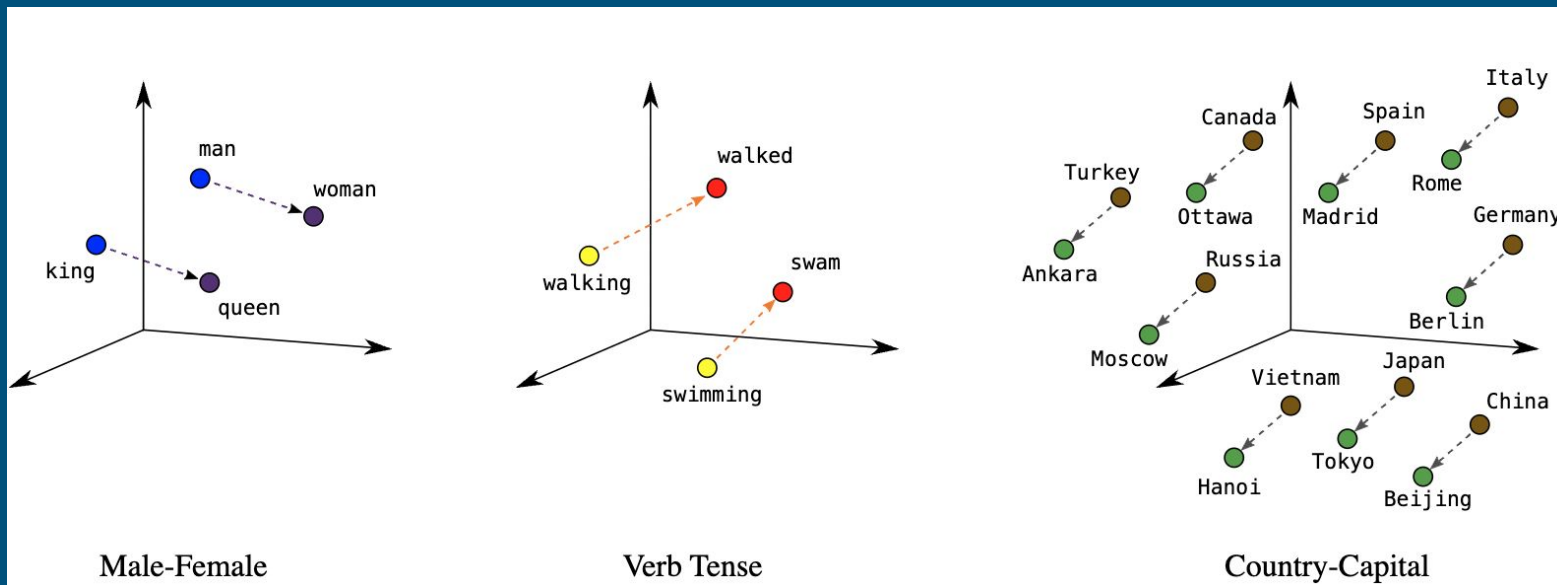
## Sources of biases (Bansal, 2022)

Label bias: watermelon

Semantic bias: embedding

Selection bias:

# Word embedding





“man is to king as woman is to  $x$ ”

$$\overrightarrow{\text{man}} - \overrightarrow{\text{woman}} \approx \overrightarrow{\text{king}} - \overrightarrow{\text{queen}}$$

$$\overrightarrow{\text{man}} - \overrightarrow{\text{woman}} \approx \overrightarrow{\text{computer programmer}} - \overrightarrow{\text{homemaker}}$$

# Fairness in NLP from A Multilinguality Perspective

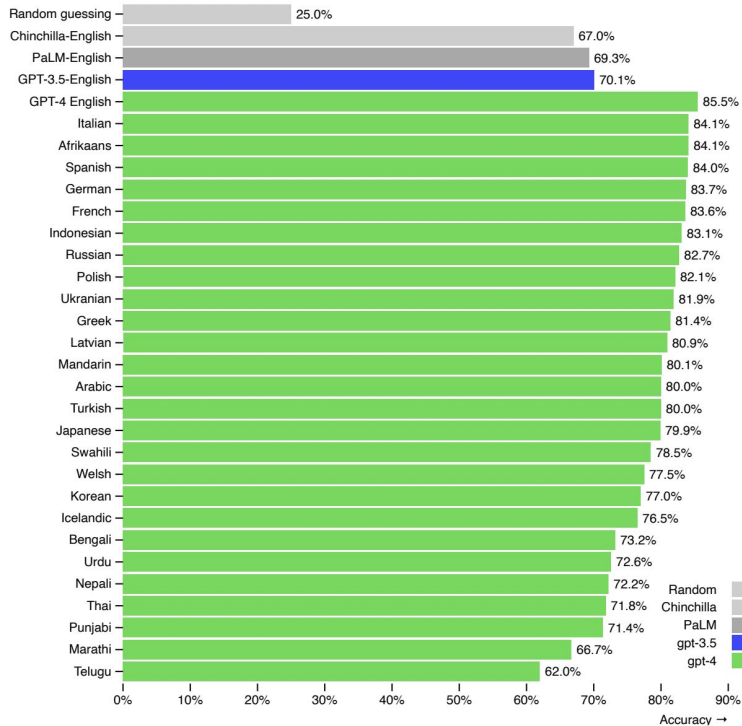
# Language distribution for PaLM

ISO Code	Language	Tokens(B)	Percentage
en	English	578.064	77.984%
de	German	25.954	3.501%
fr	French	24.094	3.250%
es	Spanish	15.654	2.112%
und	Unknown	12.064	1.628%
pl	Polish	10.764	1.452%
it	Italian	9.699	1.308%
nl	Dutch	7.690	1.037%
sv	Swedish	5.218	0.704%
tr	Turkish	4.855	0.655%
pt	Portuguese	4.701	0.634%
ru	Russian	3.932	0.530%
fi	Finnish	3.101	0.418%
cs	Czech	2.991	0.404%
zh	Chinese	2.977	0.402%
ja	Japanese	2.832	0.382%
no	Norwegian	2.695	0.364%
ko	Korean	1.444	0.195%

Chowdhery et al. 2022

# Well-represented -> better performance?

GPT-4 3-shot accuracy on MMLU across languages



How many numbers are in the list 25, 26, ..., 100?  
(A) 75 (B) 76 (C) 22 (D) 23

Answer: B

Compute  $i + i^2 + i^3 + \dots + i^{258} + i^{259}$ .

(A) -1 (B) 1 (C)  $i$  (D)  $-i$

Answer: A

Let  $A$  be a real  $2 \times 2$  matrix. Which of the following statements must be true?

I. All of the entries of  $A^2$  are nonnegative.

II. The determinant of  $A^2$  is nonnegative.

III. If  $A$  has two distinct eigenvalues, then  $A^2$  has two distinct eigenvalues.

(A) I only (B) II only (C) III only (D) II and III only

If the government subsidizes producers in a perfectly competitive market, then

(A) the demand for the product will increase

(B) the demand for the product will decrease

(C) the consumer surplus will increase

(D) the consumer surplus will decrease

# **Fumbling in Babel:** **An Investigation into ChatGPT's Language Identification Ability**

**Wei-Rui Chen<sup>λ</sup>   Ife Adebara<sup>λ</sup>   Khai Duy Doan<sup>γ</sup>   Qisheng Liao<sup>γ</sup>   Muhammad Abdul-Mageed<sup>λ,γ</sup>**

<sup>λ</sup>Deep Learning & Natural Language Processing Group, The University of British Columbia

<sup>γ</sup>Department of Natural Language Processing & Department of Machine Learning, MBZUAI

{weirui.chen, ife.adebara, muhammad.mageed}@ubc.ca,

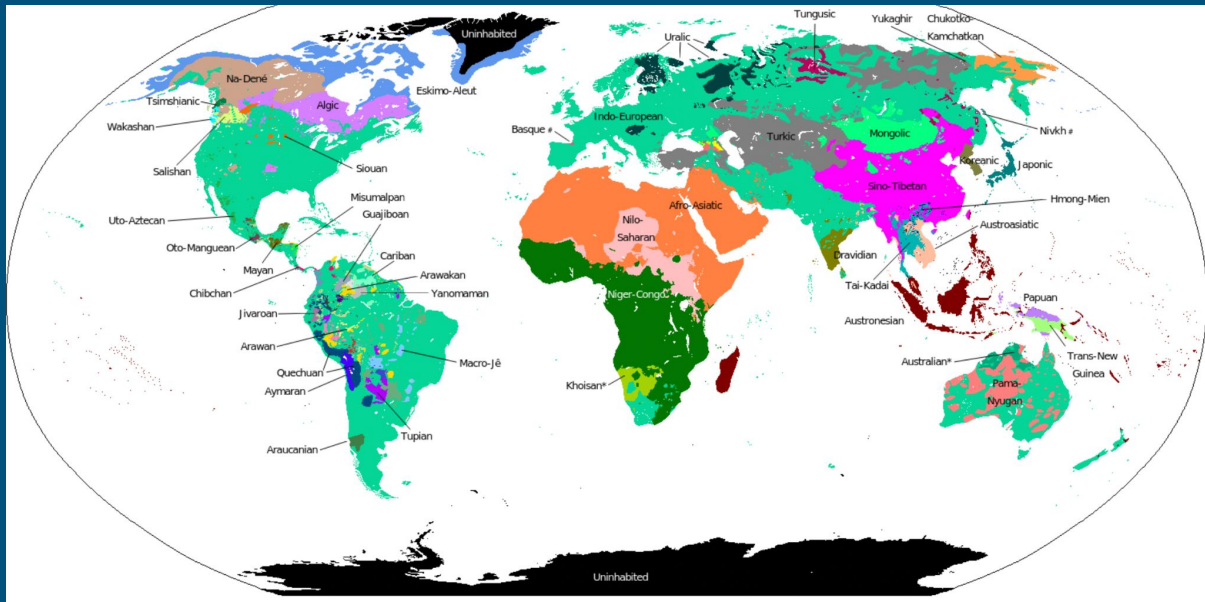
{duy.doan, qisheng.liao}@mbzuai.ac.ae





# Language Identification

A task to determine the language that a piece of text is written in (Jauhiainen et al. 2018)



Source: [https://en.wikipedia.org/wiki/List\\_of\\_language\\_families](https://en.wikipedia.org/wiki/List_of_language_families)

Dataset

# Babel-670

---

670 languages

23 language families

Afro-Asiatic, Austronesian, Aymaran, Chibchan, Creole, constructed language, Dravidian, Indo-European, Japonic, Kartvelian, Khoe-Kwadi, Koreanic, Kra-Dai, Language Isolate, Mongolic, Niger Congo, Nilo-Saharan, Quechuan, Sino-Tibetan, Tucanoan, Tupian, Turkic, and Uralic

27 scripts

Arabic, Bengali, Burmese, Coptic, Cyrillic, Devanagari, Ethiopic, Georgian, Greek, Gujarati, Gurmukhi, Hangle, Hebrew, Japanese, Kannada, Khmer, Lao, Latin, Malayalam, Odia, Ol Chiki, Sinhala, Tamil, Telugu, Thai, Tibetan, and Vai



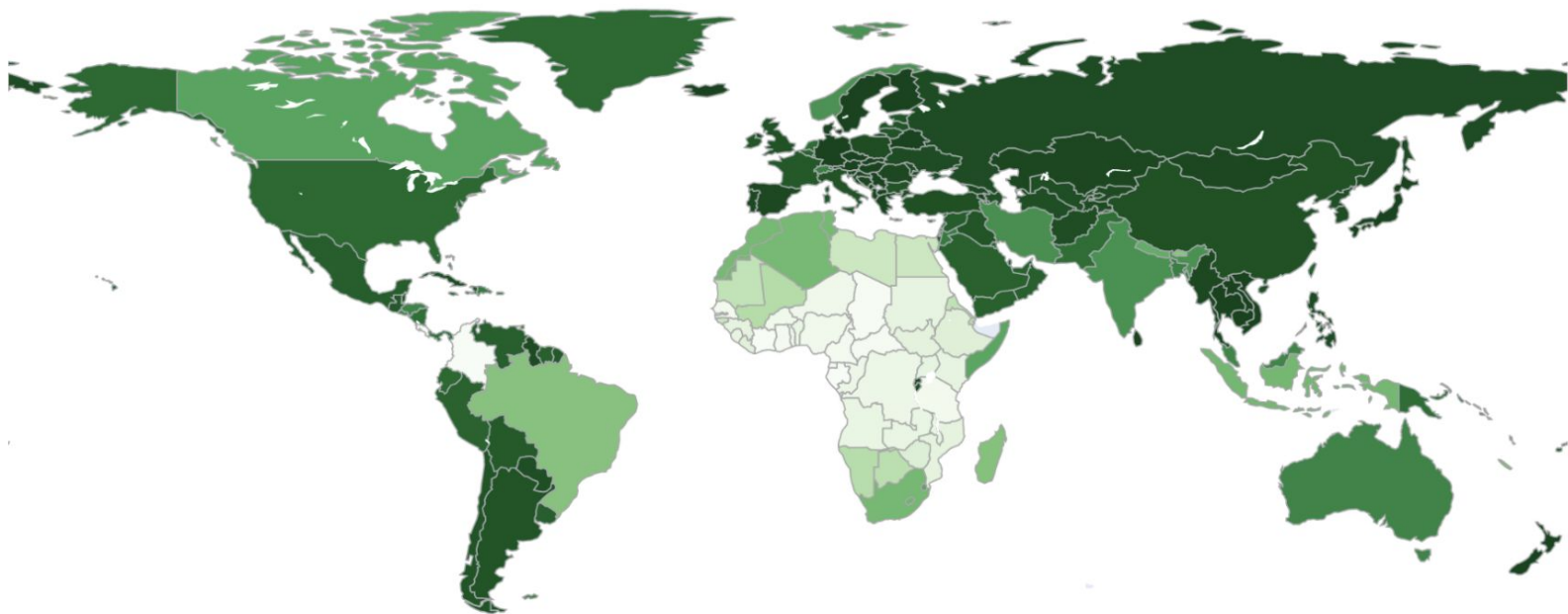
Source:  
[https://en.wikipedia.org/wiki/Tower\\_of\\_Babel](https://en.wikipedia.org/wiki/Tower_of_Babel)

# ChatGPT Performance on Babel-670

---

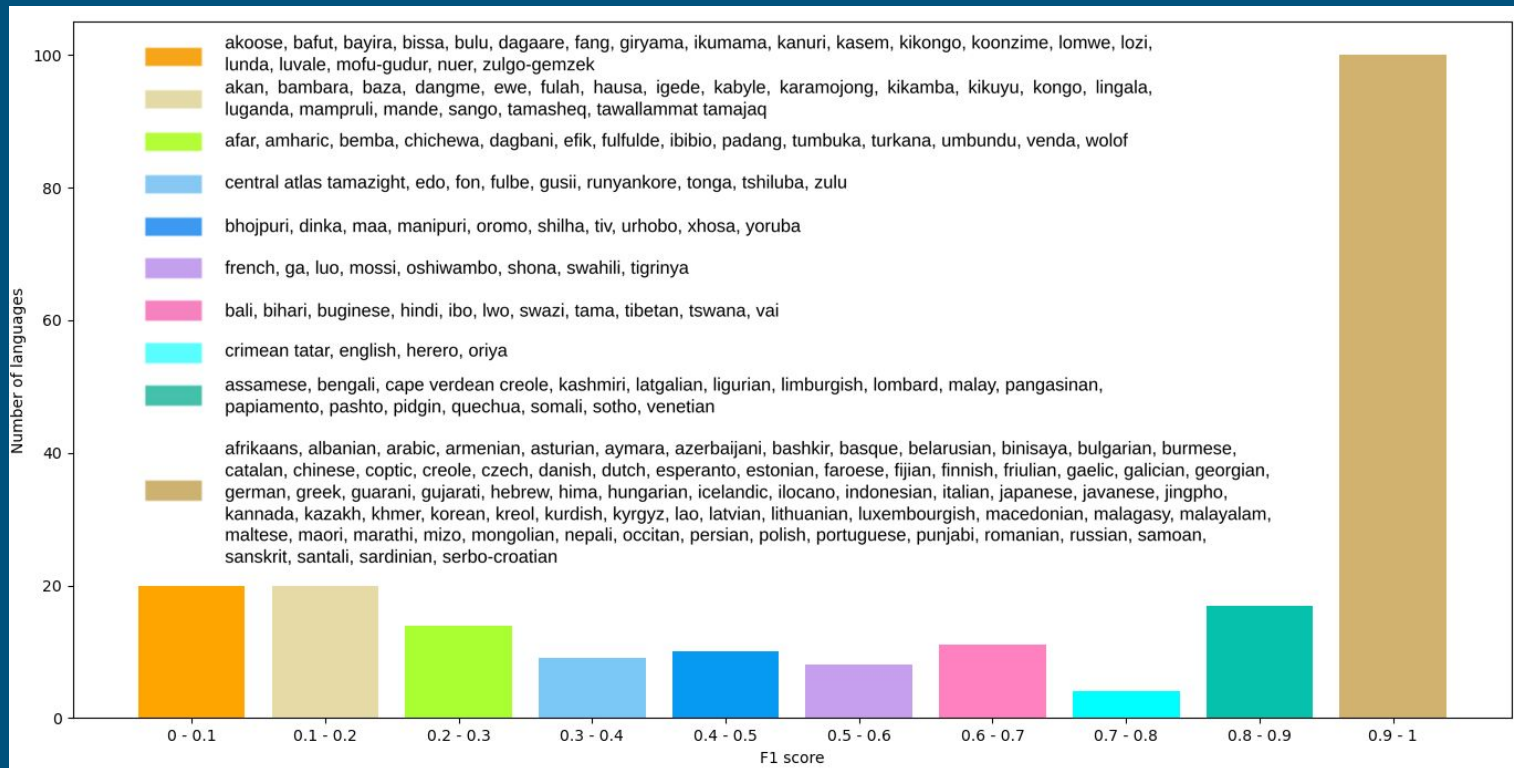
GPT-3.5	
Accuracy	F1
17.39	16.36

GPT-4	
Accuracy	F1
28.32	24.16





# Not all languages are treated equal



# Language distribution for PaLM

ISO Code	Language	Tokens(B)	Percentage
en	English	578.064	77.984%
de	German	25.954	3.501%
fr	French	24.094	3.250%
es	Spanish	15.654	2.112%
und	Unknown	12.064	1.628%
pl	Polish	10.764	1.452%
it	Italian	9.699	1.308%
nl	Dutch	7.690	1.037%
sv	Swedish	5.218	0.704%
tr	Turkish	4.855	0.655%
pt	Portuguese	4.701	0.634%
ru	Russian	3.932	0.530%
fi	Finnish	3.101	0.418%
cs	Czech	2.991	0.404%
zh	Chinese	2.977	0.402%
ja	Japanese	2.832	0.382%
no	Norwegian	2.695	0.364%
ko	Korean	1.444	0.195%

Chowdhery et al. 2022

# Conclusion

---

Not all languages are treated equal

Reduce selection bias to promote fairness of languages

Create NLP technologies to serve more communities

Thank you!



# Q & A

Linkedin: wrchen  
Twitter: WeiRuiChen01  
Email: [weirui.chen@ubc.ca](mailto:weirui.chen@ubc.ca)

