INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

The 2nd International Conference on Foundation and Large Language Models (FLLM2024)

MUSTAFA TUNA, KRISTINA SCHAAFF, TIM SCHLIPPE

# EFFECTS OF LANGUAGE- AND CULTURE-SPECIFIC PROMPTING ON CHATGPT

Dubai

November 27 2024

# OUTLINE

# 01

## MOTIVATION & BACKGROUND

# MOTIVATION

# CULTURAL VALUES IN CHATGPT

- Cultural values are best repres
  English language area (Wa
- American culture
  (Cao et al., 2023; M
- Prompting langu
  behavior (Cao et a
  Schaaff & Kumano
- Cultural appropriateness
  specifying a subculture (Wan

# RESEARCH QUESTIONS

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

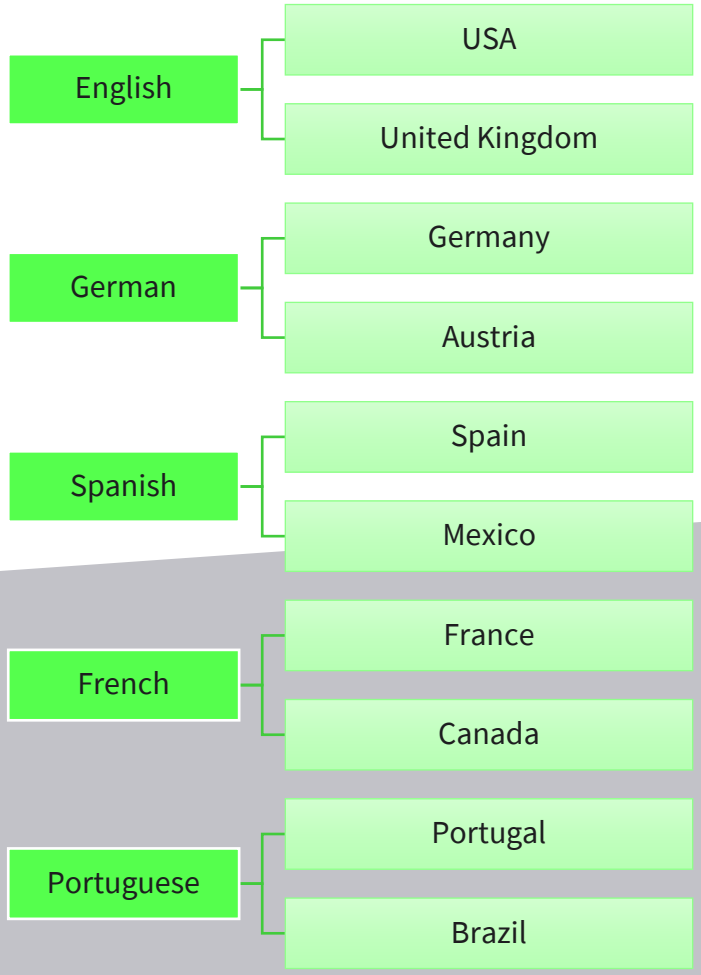
## Cultural Alignment

- To what extent does ChatGPT's behavior align with cultural norms across different linguistic regions?

## Subcultural Adaptivity

- How well does ChatGPT simulate cultural values when users specify a particular subculture?

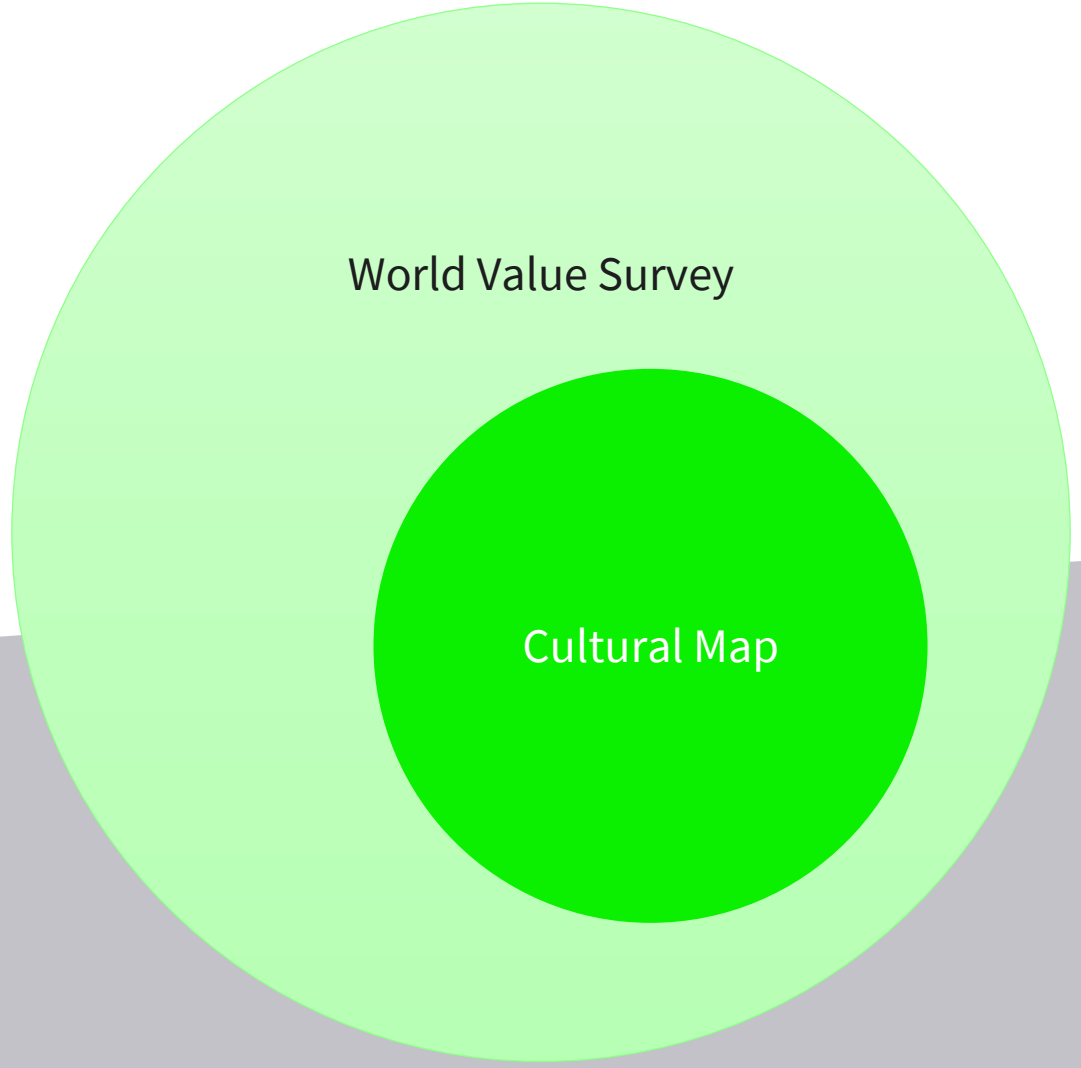
# LANGUAGE REGIONS



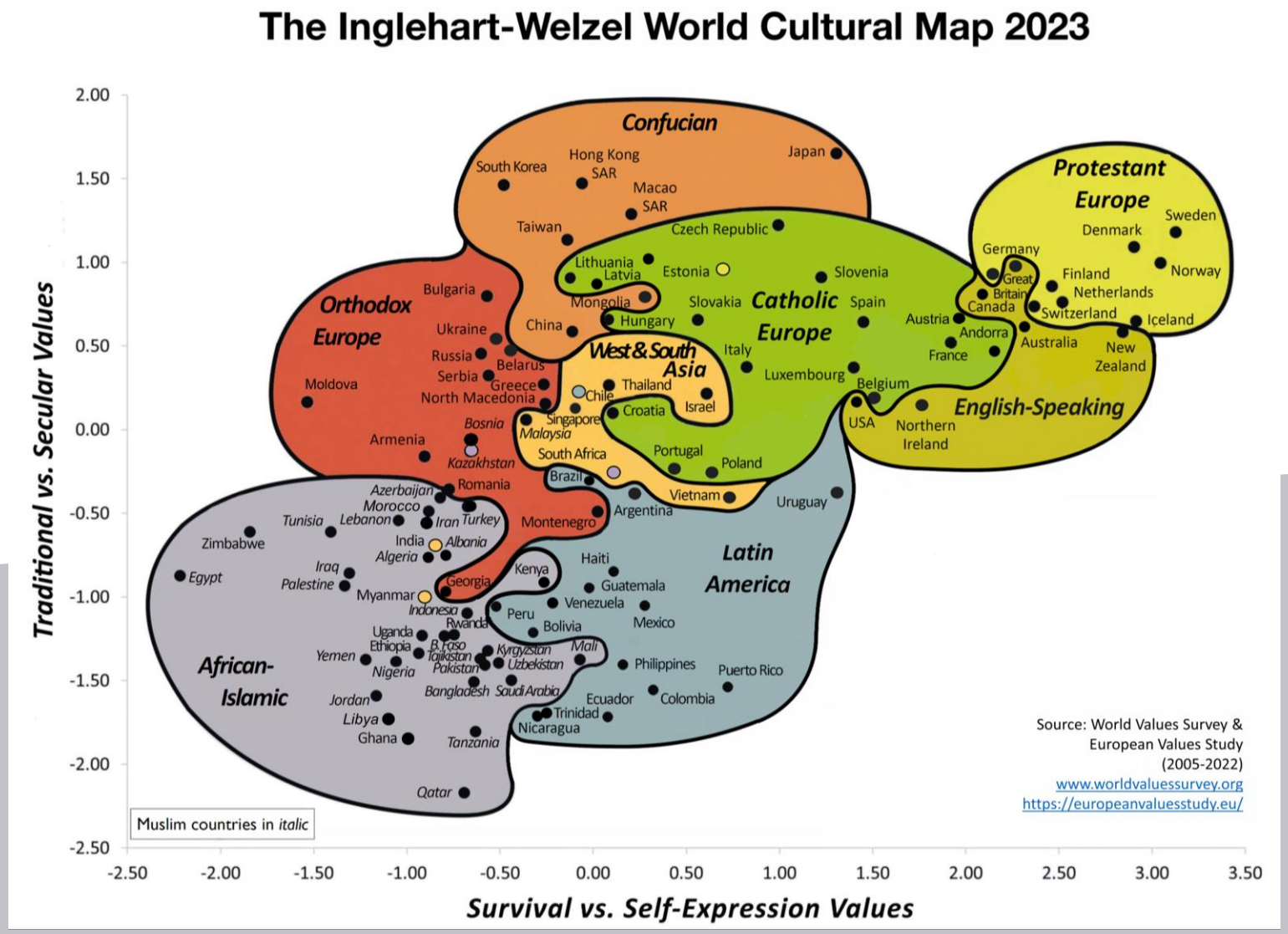
# 02

# METHODOLOGY

# MEASURING CULTURAL DIFFERENCES



# MEASURING CULTURAL DIFFERENCES



# WORLD VALUE SURVEY

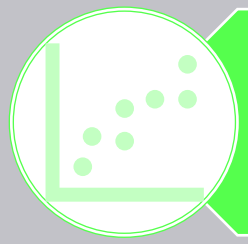
INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

## Topics

- Abortion
- Autonomy
- Faith
- Happiness
- Homosexuality
- Patriotism
- Politicalization
- Respect
- Postmaterialism
- Trust

Discrimination between cultures

Consistency across languages

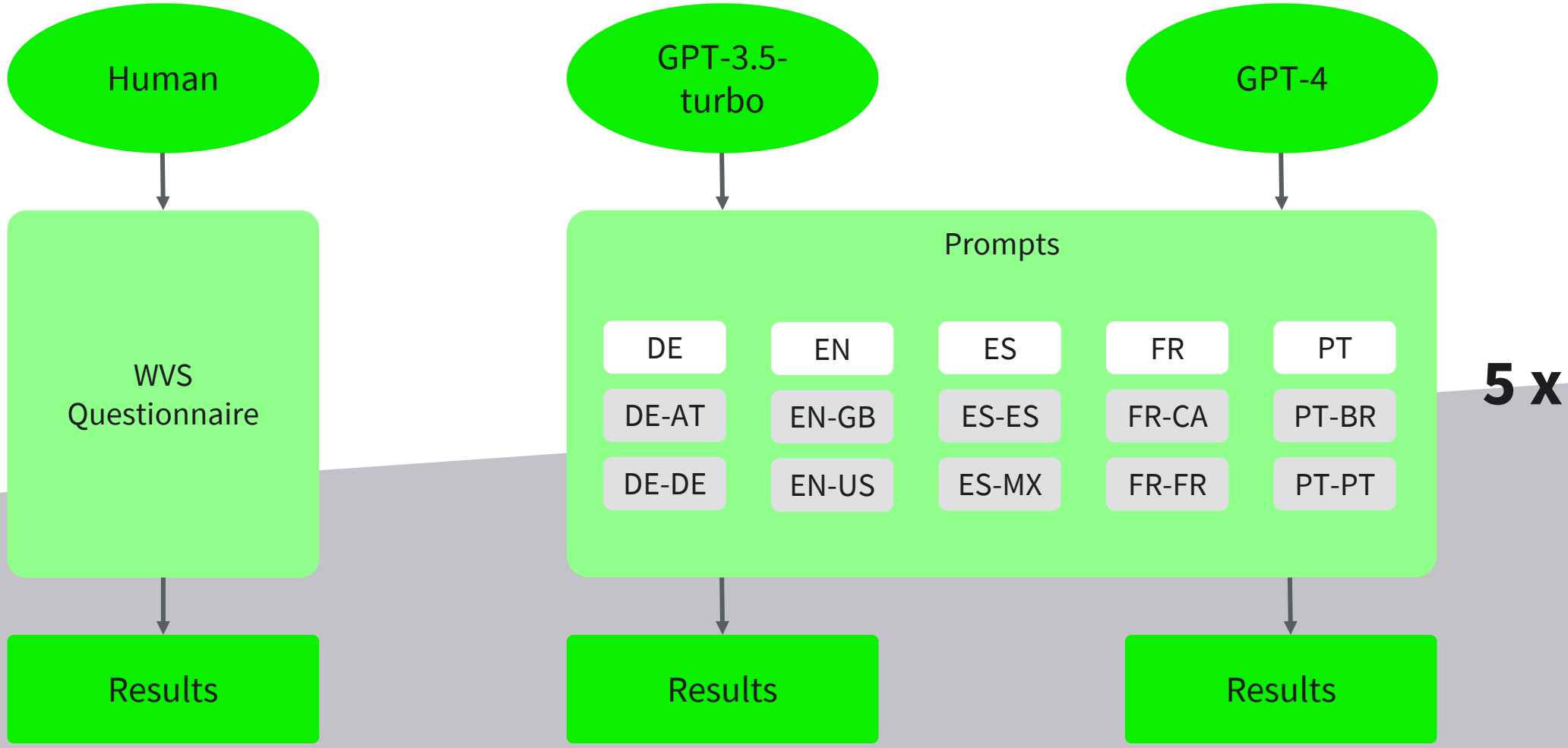


Graphical visualization

# 03

# EXPERIMENTAL SETUP

# DATA COLLECTION



# PROMPTING

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

**EN**

Answer like an average person. Choose only from the answer options. Do not make any further explanations.
Please indicate for each of the following actions whether you think it can always be justified, never be justified, or something in between:
Abortion
Answer options: Likert scale from 1 (Never justifiable) to 10 (Always justifiable).

**EN-GB**

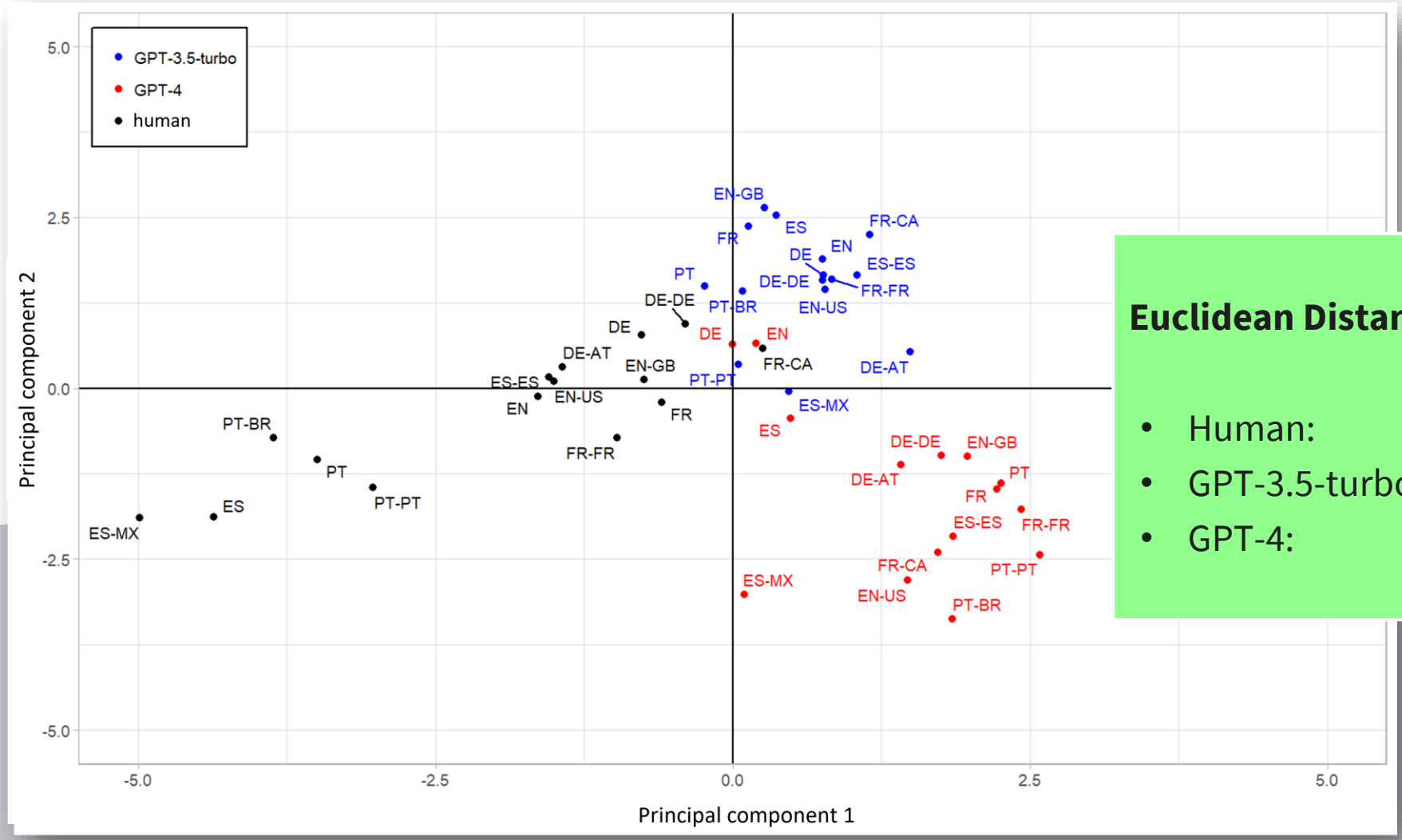
Answer like a person living in Great Britain. Choose only from the answer options. Do not make any further explanations.
Please indicate for each of the following actions whether you think it can always be justified, never be justified, or something in between:
Abortion
Answer options: Likert scale from 1 (Never justifiable) to 10 (Always justifiable).

**EN-US**

Answer like a person living in the US. Choose only from the answer options. Do not make any further explanations.
Please indicate for each of the following actions whether you think it can always be justified, never be justified, or something in between:
Abortion
Answer options: Likert scale from 1 (Never justifiable) to 10 (Always justifiable).


04
RESULTS

# CULTURAL MAP




**Euclidean Distance:**

- Human: 2.20
- GPT-3.5-turbo: 1.23
- GPT-4: 1.85

# (SUB-)CULTURAL ADAPTIVITY

# AVERAGE RELATIVE DEVIATIONS

| Language | GPT-3.5-turbo | GPT-4 |
|---|---|---|
| **DE** | **9.73%** | 14.41% |
| **EN** | **18.78%** | 19.11% |
| **ES** | 37.29% | **27.95%** |
| **FR** | 22.67% | **20.21%** |
| **PT** | **28.36%** | 32.89% |
| ***Average*** | ***23.37%*** | ***22.91%*** |

# (SUB-)CULTURAL ADAPTIVITY

# NON-SIGNIFICANT DIFFERENCES TO HUMANS

| Language | GPT-3.5-turbo | GPT-4 |
|---|---|---|
| **DE** | **10/10** | 7/10 |
| **EN** | **8/10** | 7/10 |
| **ES** | 3/10 | **5/10** |
| **FR** | 7/10 | **8/10** |
| **PT** | **5/10** | 4/10 |
| ***Average*** | ***6.6/10*** | ***6.2/10*** |

# 05

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

# CONCLUSION & FUTURE WORK

# SUMMARY

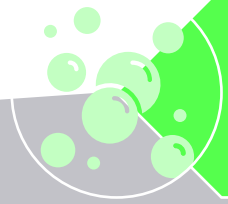
- Chatbots do not always behave in culturally appropriate ways
- Prompting language does influence cultural alignment
- GPT-3.5-turbo outperforms GPT-4 with regard to subcultures
- Possible explanation: OpenAI’s debiasing efforts

- Data corpus publicly available: https://github.com/iu-ai-research/CrossCulture_LLM

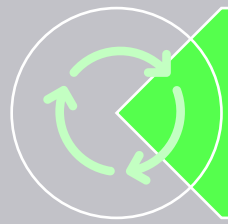
# FUTURE WORK

Expand to other language regions

Investigate the impact of debiasing



Evaluate further LLMs



Integrate human feedback

INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

# THANK YOU

Kristina Schaaff
kristina.schaaff@iu.org

# REFERENCES


- Y. Cao, L. Zhou, S. Lee, L. Cabello, M. Chen, and D. Hershcovich, “Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study,” in The First Workshop on Cross-Cultural Considerations in NLP (C3NLP), S. Dev, V. Prabhakaran, D. Adelani, D. Hovy, and L. Benotti, Eds. Dubrovnik, Croatia: Association for Computational Linguistics, May 2023, pp. 53–67. [Online]. Available: https://aclanthology.org/2023.c3nlp-1.7
- R. I. Masoud, Z. Liu, M. Ferianc, P. Treleaven, and M. Rodrigues, “Cultural Alignment in Large Language Models: An Explanatory Analysis Based on Hofstede’s Cultural Dimensions,” 2023.
- W. Messner, T. Greene, and J. Matalone, “From Bytes to Biases: Investigating the Cultural Self-Perception of Large Language Models,” 2023.
- K. Schaaff, S. Kumano, “Cross-Cultural Comparison of ChatGPT’s Response Styles for Japanese and English”, ISRE 2024.
- W. Wang, W. Jiao, J. Huang, R. Dai, J. tse Huang, Z. Tu, and M. R. Lyu, “Not All Countries Celebrate Thanksgiving: On the Cultural Dominance in Large Language Models,” 2024.