

The AI Gap: How Socioeconomic Status Affects Language Technology Interactions



Tuca Vieira, 2004



Elisa Bassignana*
IT University of Copenhagen



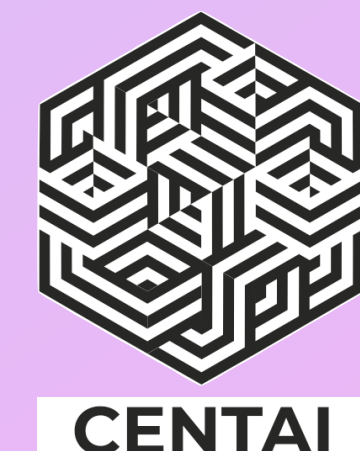
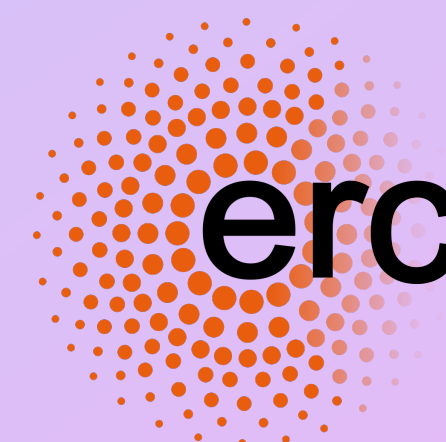
Amanda Cercas Curry*
CENTAI Institute



Dirk Hovy
Bocconi University

(*) equal contribution

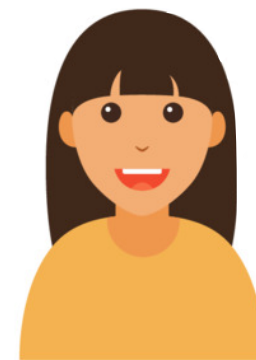
VILLUM FONDEN



Different SES ask AI different things

Topic: Food 🍎

Hello...what is the best way for me to reheat my Buffalo Wild Wings?



What are some creative uses for left-over rice?



What wine with red tuna belly



Different SES ask AI different things

Topic: Dealing with financial struggles 💰

What is the best way to reduce my debt?



Different SES ask AI different things

Topic: Travel ✈️

I am planning to travel for a vacation in Japan, do you have any recommendations?



How does SES affect the
usage of language
technologies?

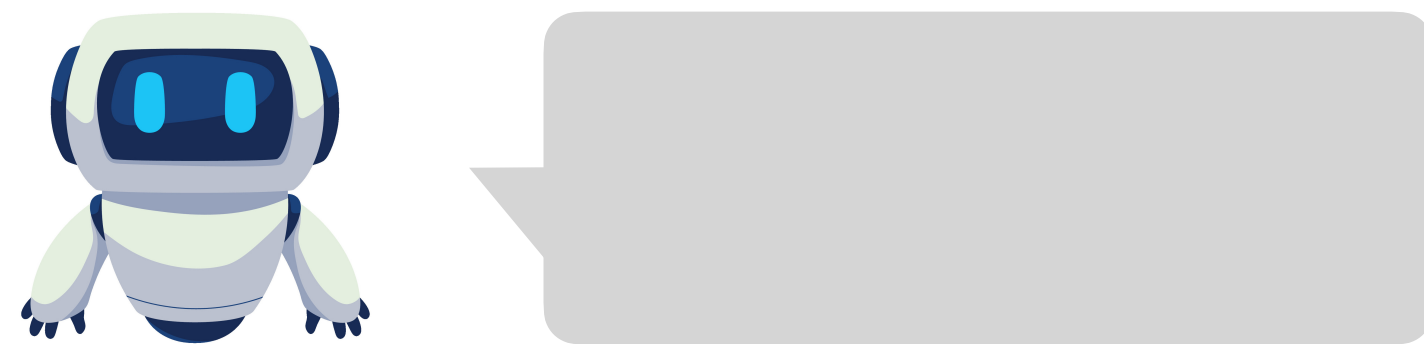
Contributions

We explore the **interplay**
between SES and LLM usage.



Contributions

We explore the **interplay between SES and LLM usage**.



We introduce and analyze a **dataset** of how people use LLMs, including **real prompts** from previous conversation with LLMs.

We **survey 1,000 individuals** on Profilic



We **survey 1,000 individuals** on Profilic

Demographics

- **SES:** 29% low, 58% middle, 13% upper
- **Location:** 41% UK and 59% US
- **Other:** age, gender, education, occupation, hobbies, parents' SES...



We **survey 1,000 individuals** on Profilic

Demographics

- **SES:** 29% low, 58% middle, 13% upper
- **Location:** 41% UK and 59% US
- **Other:** age, gender, education, occupation, hobbies, parents' SES...



Usage of language technologies

- *"How often do you use AI chatbots like ChatGPT?"*
- *"In which of the following context have you ever used AI chatbots like ChatGPT?"*
-

We **survey 1,000 individuals** on Profilic

Demographics

- **SES:** 29% low, 58% middle, 13% upper
- **Location:** 41% UK and 59% US
- **Other:** age, gender, education, occupation, hobbies, parents' SES...



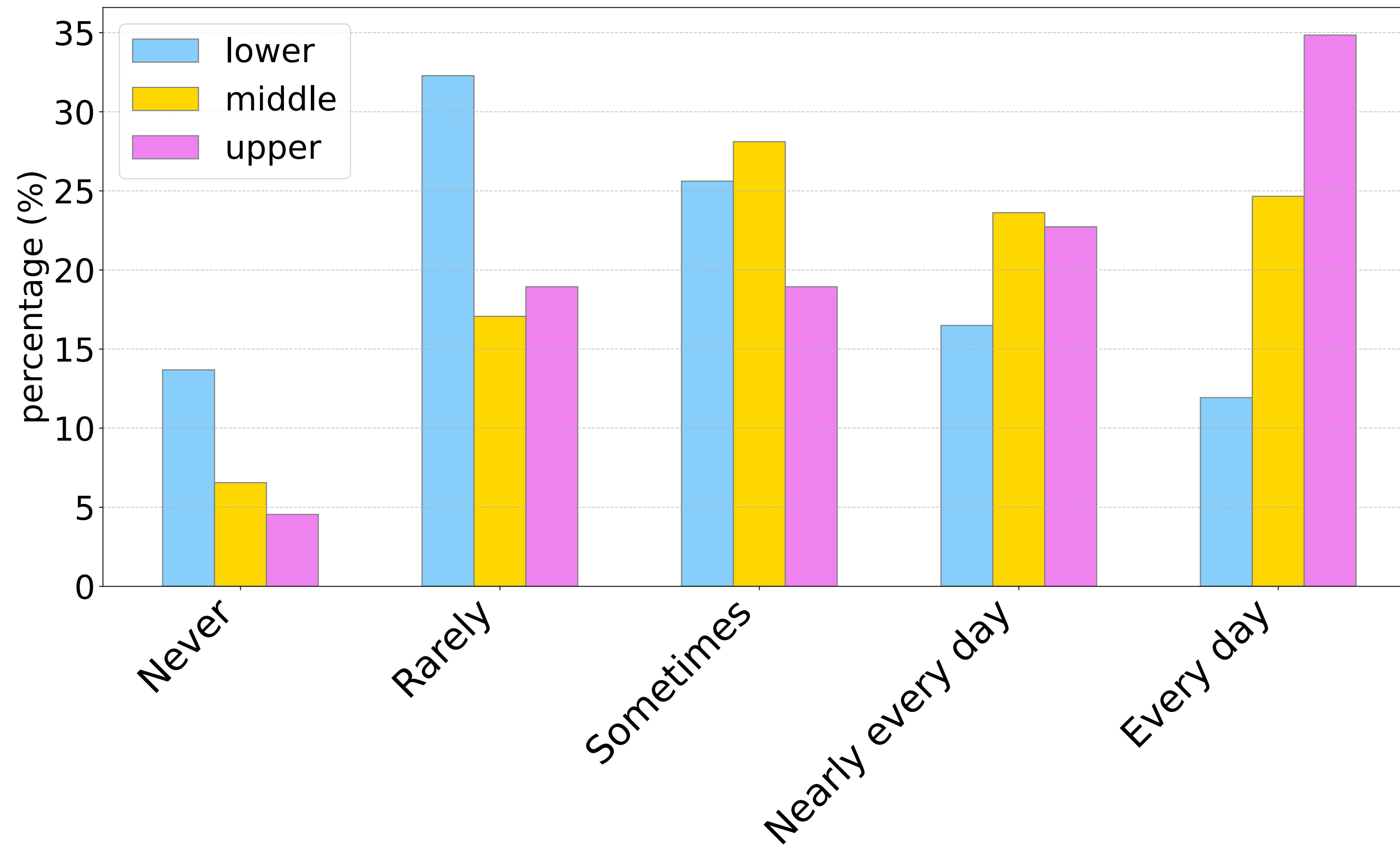
Usage of language technologies

- *"How often do you use AI chatbots like ChatGPT?"*
- *"In which of the following context have you ever used AI chatbots like ChatGPT?"*
-

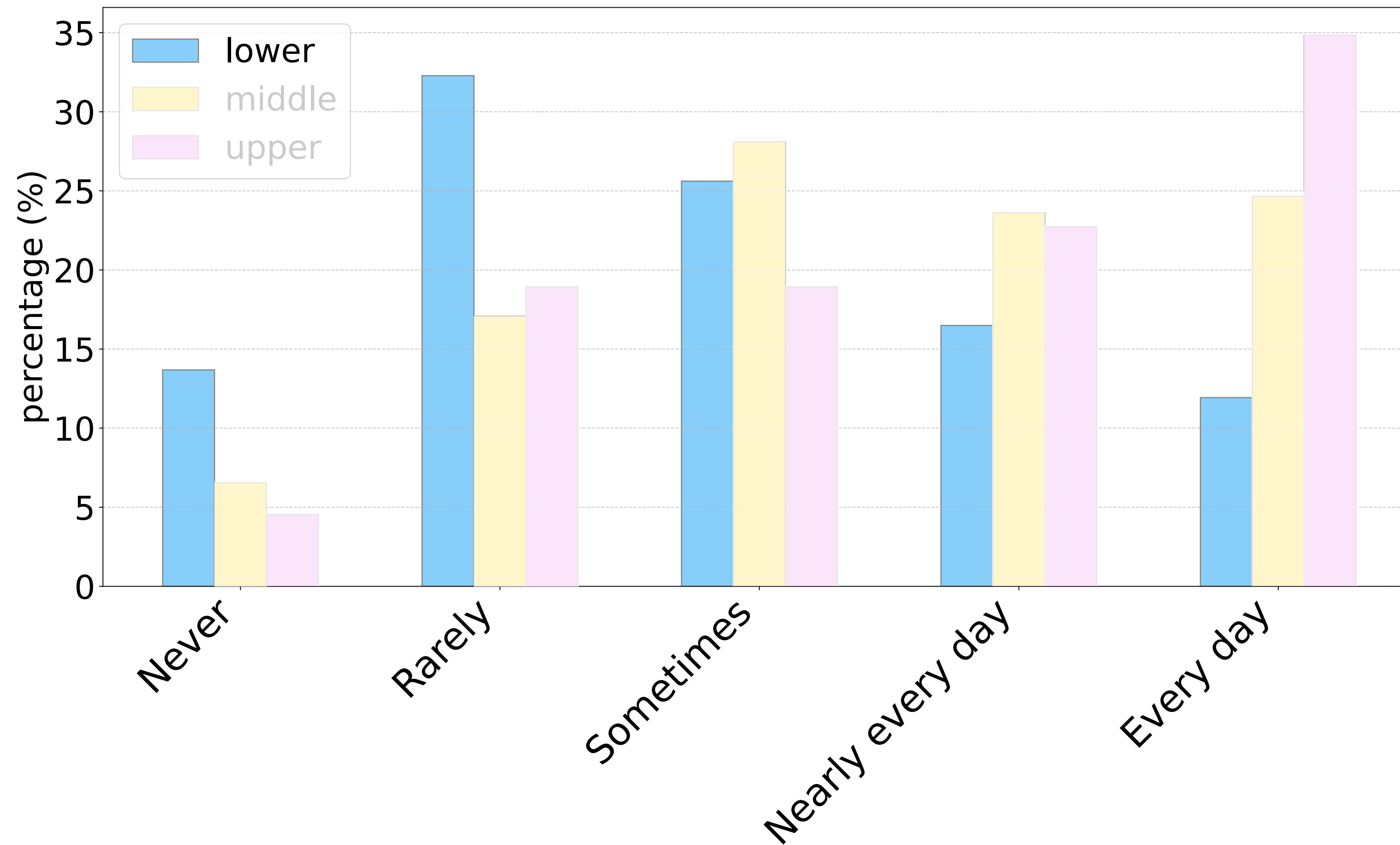
 **Last 10 prompts** submitted to an AI chatbot

- obtained 6,500 prompts

How often do individuals use AI chatbots?

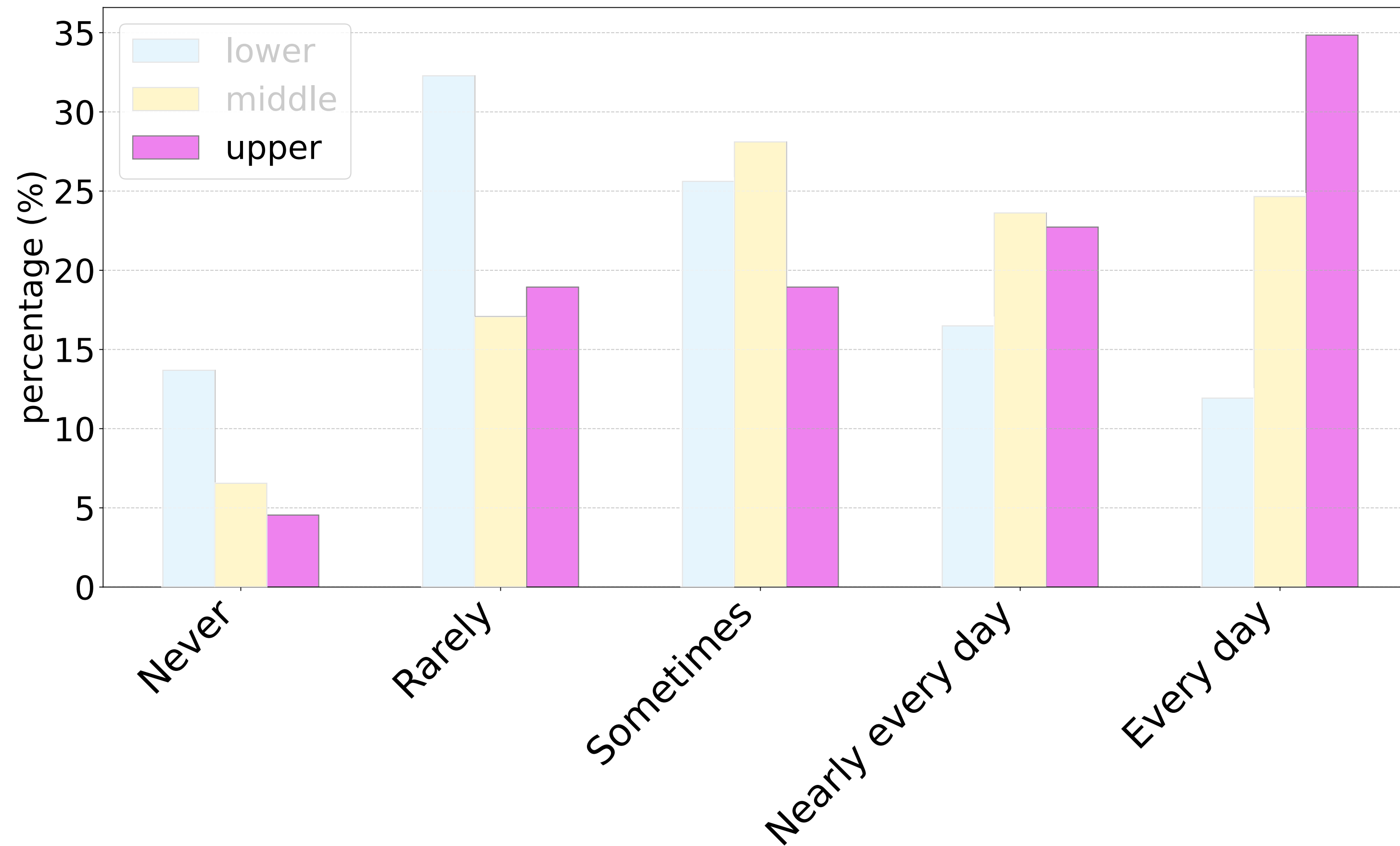


How often do individuals use AI chatbots?



Lower SES **rarely** uses LLMs.

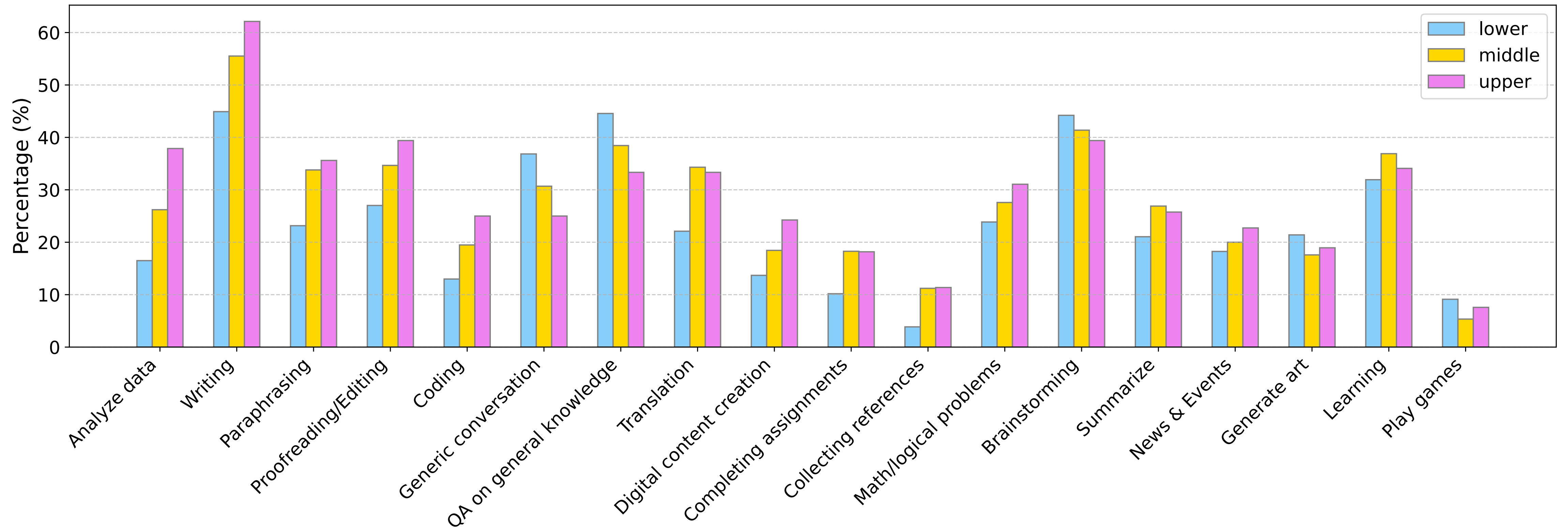
How often do individuals use AI chatbots?



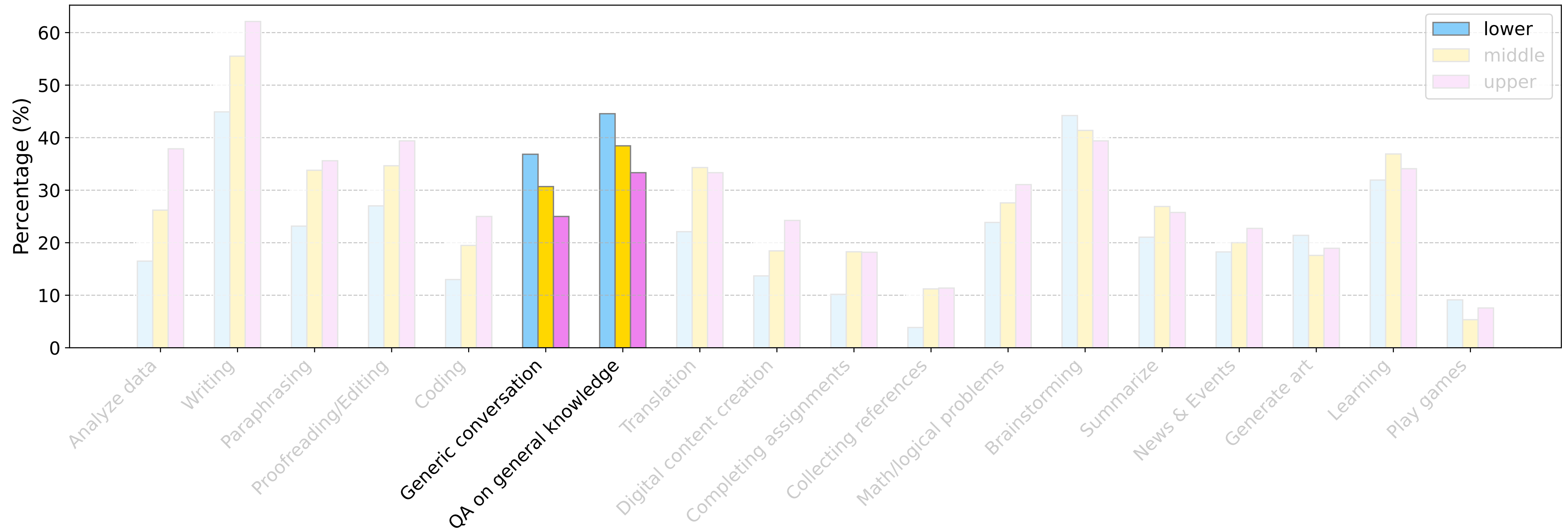
Lower SES **rarely** uses LLMs.

Upper SES mostly uses LLMs
every day.

What do users use LLMs for?

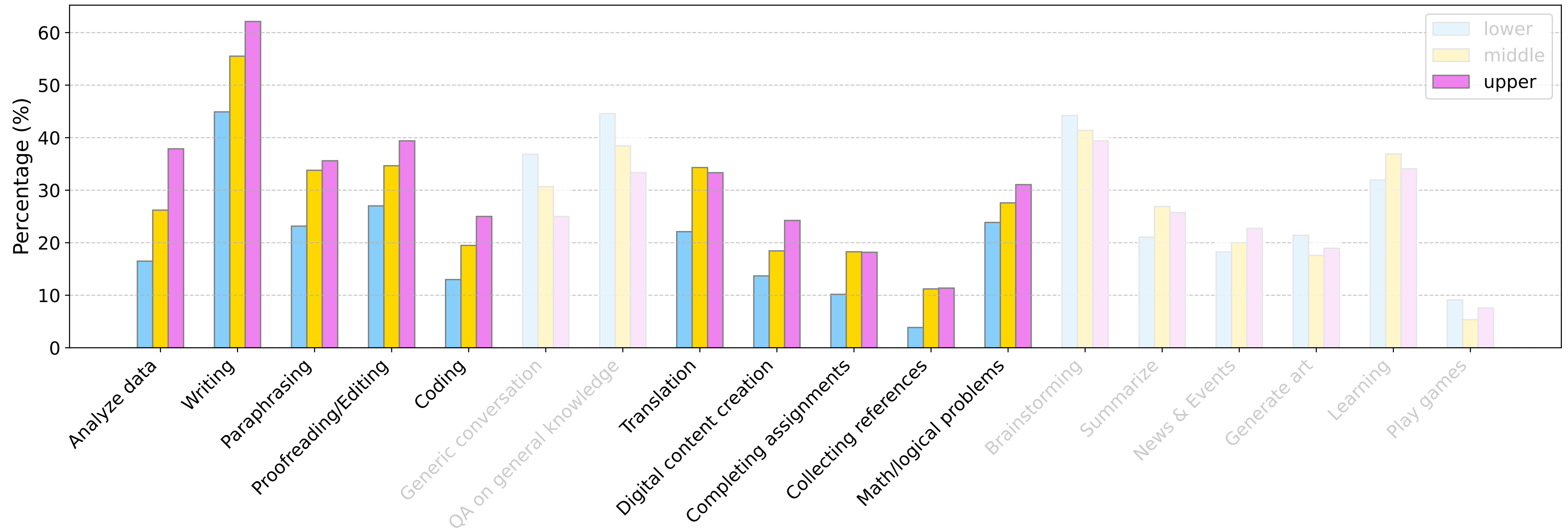


What do users use LLMs for?



Lower SES uses LLMs
more for **entertainment**.

What do users use LLMs for?



Lower SES uses LLMs more for **entertainment**.

Upper SES uses LLMs more for **work** and **studying**.

Prompt analysis



Prompt analysis



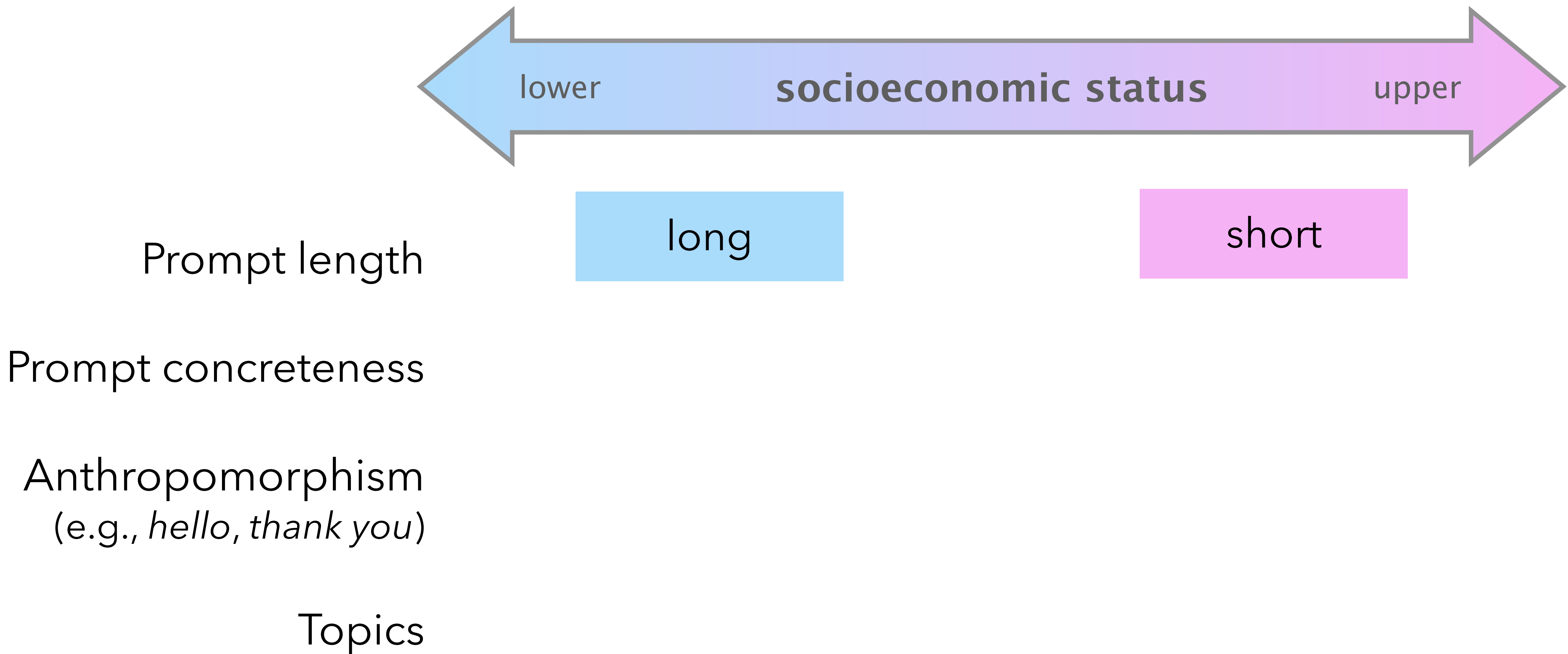
Prompt length

Prompt concreteness

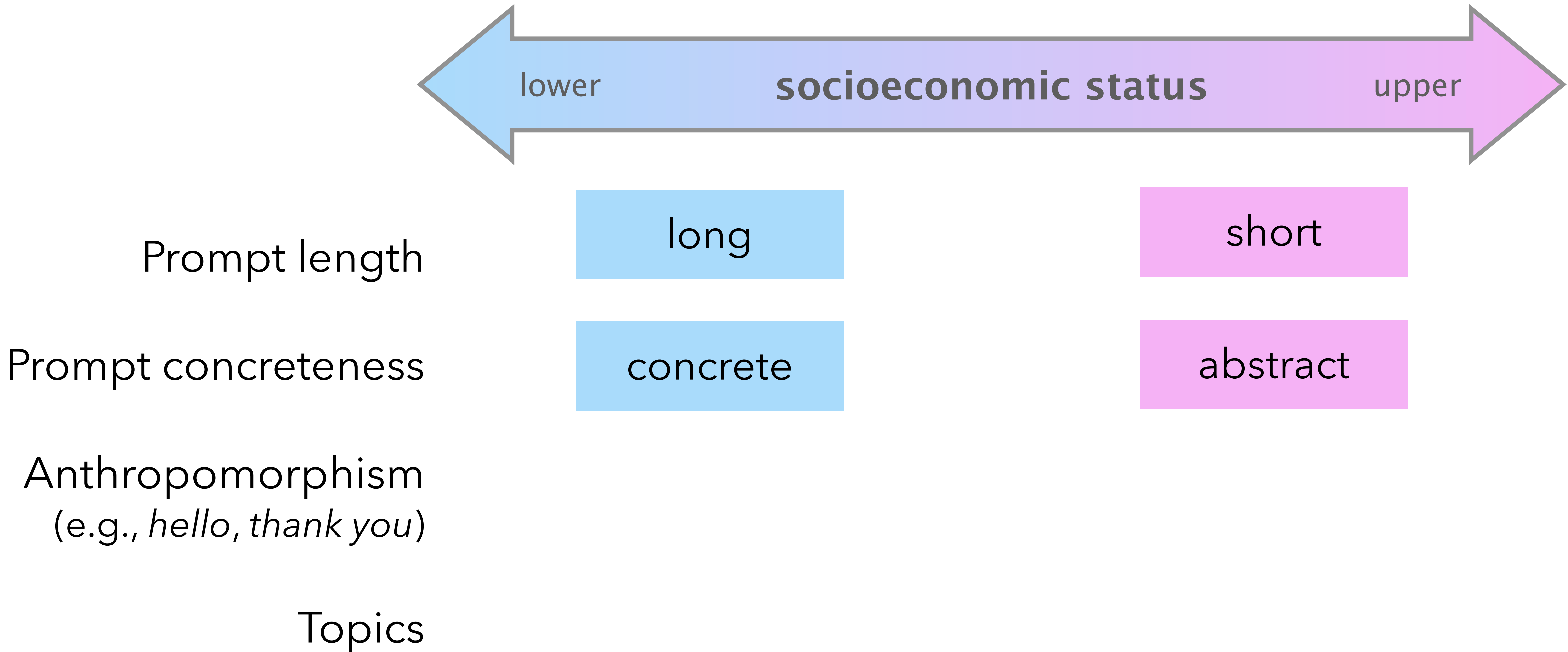
Anthropomorphism
(e.g., *hello, thank you*)

Topics

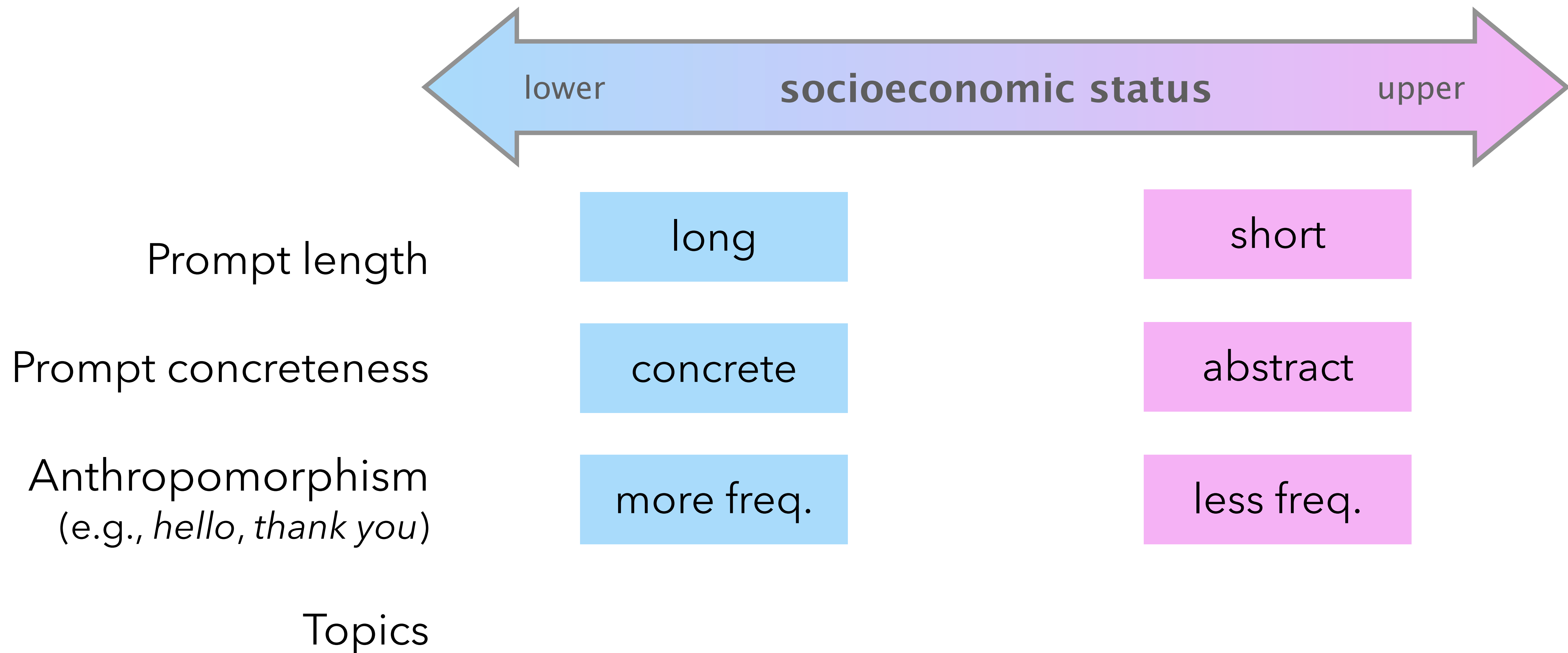
Prompt analysis



Prompt analysis



Prompt analysis



Prompt analysis



Prompt length

long

short

Prompt concreteness

concrete

abstract

Anthropomorphism
(e.g., *hello, thank you*)

more freq.

less freq.

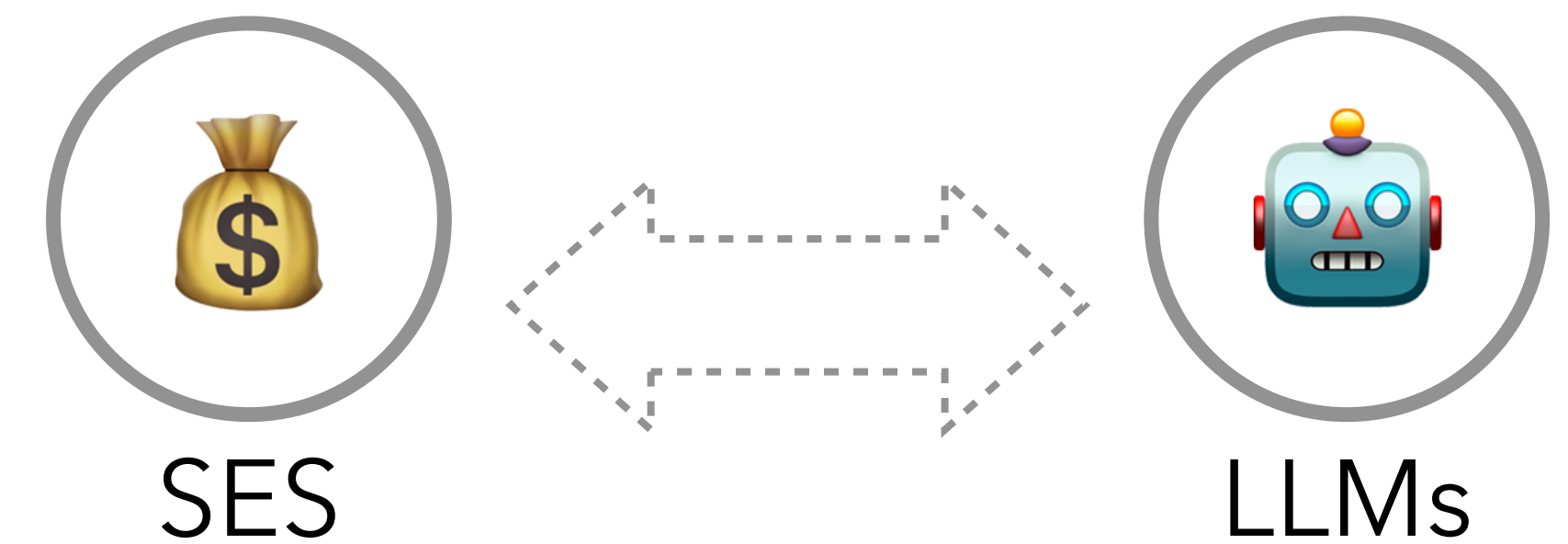
Topics

financial
struggle, ...

investments,
travel, ...

Conclusions

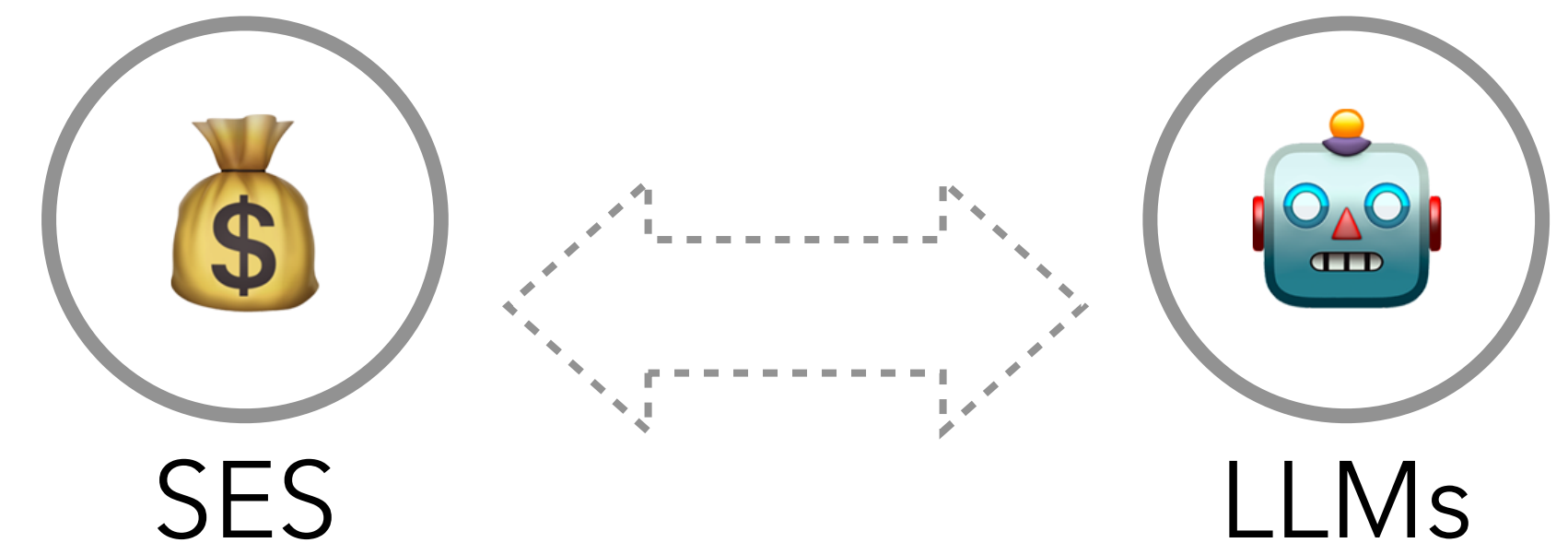
We **survey** the usage of LLMs among SES and **share the data**.



Conclusions

We **survey** the usage of LLMs among SES and **share the data**.

We show that the **adoption rates and context of use of LLMs vary significantly** based on the SES of the user.

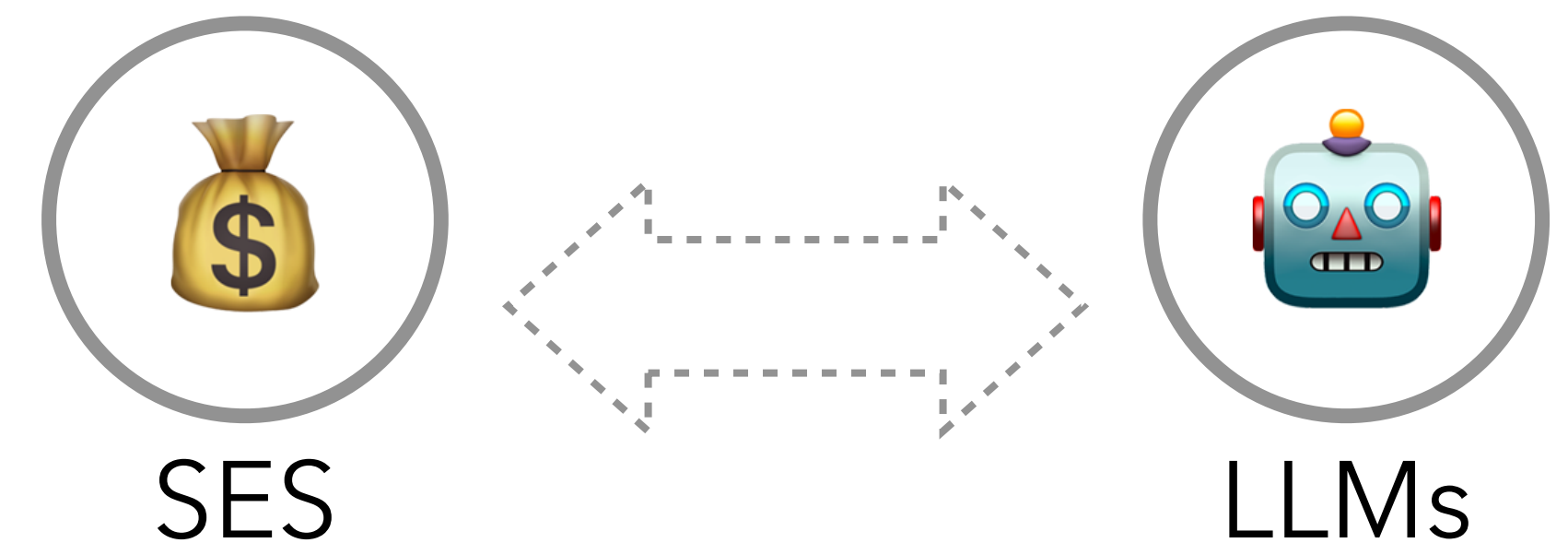


Conclusions

We **survey** the usage of LLMs among SES and **share the data**.

We show that the **adoption rates and context of use of LLMs vary significantly** based on the SES of the user.

We find **significant linguistic differences** (concreteness, length etc.) in the formulation of the prompts.

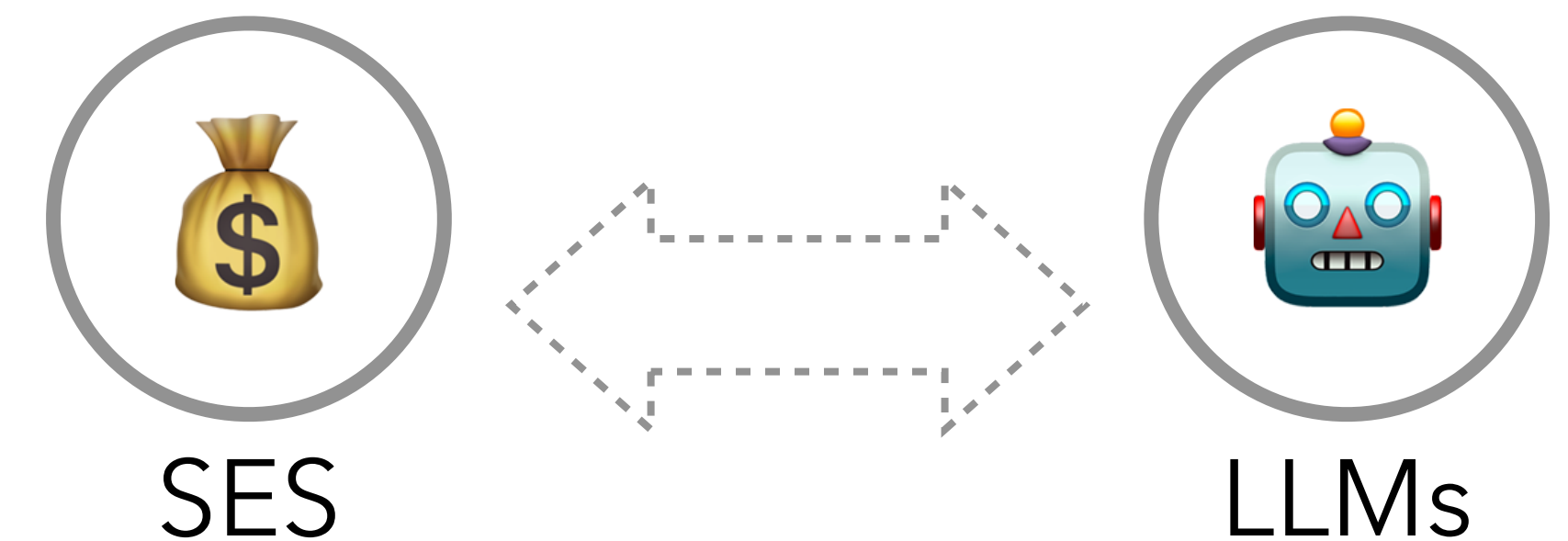


Conclusions

We **survey** the usage of LLMs among SES and **share the data**.

We show that the **adoption rates and context of use of LLMs vary significantly** based on the SES of the user.

We find **significant linguistic differences** (concreteness, length etc.) in the formulation of the prompts.



Implication for **benchmarking, alignment, fairness!**

How can we develop
technologies that take into
account **different SES**?

Thank you!



dataset 😊



Tuca Vieira, 2004



Elisa Bassignana*



Amanda Cercas Curry*



Dirk Hovy

References

- William Labov. 2006. The Social Stratification of English in New York City, 2 edition. Cambridge University Press.
- Basil Bernstein. 1960. Language and social class. The British Journal of Sociology, 11(3):271-276.
- Eve V Clark and Marisa Casillas. 2015. First language acquisition. In The Routledge handbook of linguistics, pages 311-328. Routledge.
- Kraus, M. W., Park, J. W., & Tan, J. J. X. (2017). Signs of Social Class: The Experience of Economic Inequality in Everyday Life. *Perspectives on Psychological Science*, 12(3)

Appendix

Social class

Grouping people according to socioeconomic status (SES) based on:

- ❖ Wealth
- ❖ Education
- ❖ Social capital
- ❖ Occupation
- ❖ ...

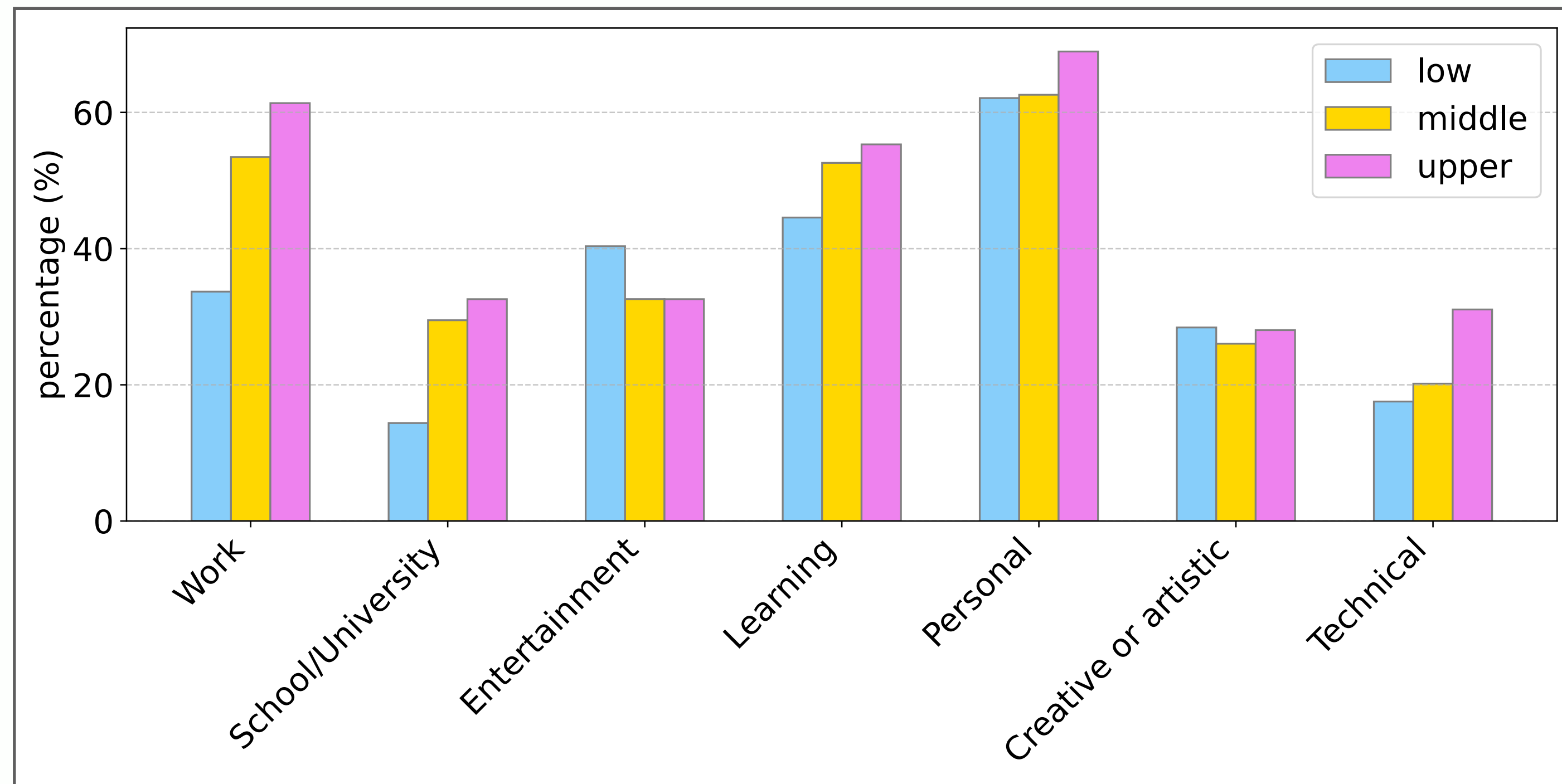
There are different classification systems based on cultures.

We focus on the European class system.



In which context do individuals use AI chatbots?

Context of usage



The Macarthur scale

(Adler et al., 2000)

Based on the picture on the side, where would you place yourself in the socioeconomic ladder in terms of wealth?

