The AI Gap:
How Socioeconomic Status Affects Language Technology Interactions

Elisa Bassignana(*)
IT University of Copenhagen
Pioneer Center for AI
elba@itu.dk

Amanda Cercas Curry(*)
CENTAI Institute
amanda.cercas@centai.eu

Dirk Hovy
Bocconi University
dirk.hovy@unibocconi.it

Abstract

Socioeconomic status (SES) fundamentally influences how people interact with each other and more recently, with digital technologies like Large Language Models (LLMs). While previous research has highlighted the interaction between SES and language technology, it was limited by reliance on proxy metrics and synthetic data. We survey 1,000 individuals from *diverse socioeconomic backgrounds* about their use of language technologies and generative AI, and collect 6,482 prompts from their previous interactions with LLMs. We find systematic differences across SES groups in language technology usage (i.e., frequency, performed tasks), interaction styles, and topics. Higher SES entails a higher level of abstraction, convey requests more concisely, and topics like 'inclusivity' and 'travel'. Lower SES correlates with higher anthropomorphization of LLMs (using "hello" and "thank you") and more concrete language. Our findings suggest that while generative language technologies are becoming more accessible to everyone, socioeconomic linguistic differences still stratify their use to exacerbate the digital divide. These differences underscore the importance of considering SES in developing language technologies to accommodate varying linguistic needs rooted in socioeconomic factors and limit the AI Gap across SES groups.

1 Introduction

The development of Large Language Models (LLMs), in particular "AI chatbots" like ChatGPT (OpenAI, 2023) and DeepSeek (DeepSeek-AI, 2024), are rapidly transforming how we interact with technology. However, despite widespread accessibility, how (and how frequently) people use them varies significantly between groups. Despite their enthusiastic adoption, people from various

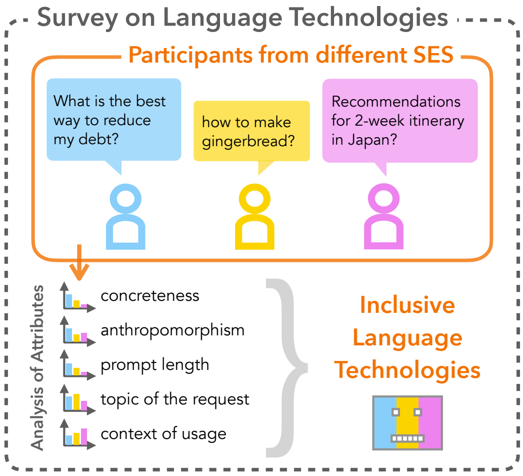


Figure 1: We survey the usage of language technologies across individual with different SES, and collect prompts from past interactions with LLMs. We find significant differences in habits and prompting strategies (see example attributes). Future language technologies should address these disparities to reduce the AI Gap.

socioeconomic backgrounds use these tools very differently. The Economist[1] reports non-white families have adopted LLMs for education more readily than their white counterparts, and UNESCO reports women are less likely to use them at work than men (UNESCO, 2023). These differences in adoption rates (the "AI Gap") have raised growing concerns over their effect on exacerbating existing inequalities (Capraro et al., 2024). This disparity is more than just a curiosity; it is an urgent signal of a growing digital divide, similar to how educational attainment, income levels, and access to digital resources have historically driven disparities.

According to the Technology Acceptance Model (TAM; Davis, 1989), adoption of new technologies is influenced by their *perceived usefulness* and their *perceived ease of use* – that is, users who see no practical benefits in using a technology, or who have previously had bad experiences, are less likely

(*) Equal contribution.

[1] The Economist. Accessed 12th February 2025

Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 18647–18664
July 27 - August 1, 2025 ©2025 Association for Computational Linguistics

to adopt it. Technology adoption rates are further shaped by factors such as access, digital literacy, and ethical/privacy concerns. Poverty and socioeconomic status (SES) are among the main drivers of the digital divide, owing to issues of access and digital literacy (Mubarak et al., 2020). However, the "digital divide" is about more than just having access to devices and the internet; it is also about using technology effectively to improve one's life.

As LLM use grows, a digital divide along SES lines could exacerbate inequalities in several ways. LLMs trained on data from higher SES usage patterns may be less effective or biased against lower SES language styles and content interests, perpetuating social biases. As a result, lower SES perspectives may be underrepresented or misrepresented, resulting in skewed narratives that do not accurately reflect society. People who use simpler language and more concrete terms may receive less effective results if models perform better with abstract or sophisticated prompts. The assistance quality could suffer, and this may not be reflected in evaluation benchmarks (Liao and Xiao, 2023). Reduced user satisfaction could further discourage adoption among lower SES groups, widening the AI Gap.

Previous work has shown that the *performance* of NLP systems is affected by sociodemographic linguistic differences, including native language (Reusens et al., 2024), race (Blodgett and Talat, 2024), and social class (Cercas Curry et al., 2024a). However, none of those works have examined how SES influences the use cases of language technologies or whether there are any systematic linguistic differences in the interactions with LLMs. Our research addresses this gap by investigating how different SES groups take advantage of current language technologies and how they interact with LLMs differently. To the best of our knowledge, we are the first to compare the use of language technologies in general and LLMs in particular across socioeconomic groups. We discovered statistically significant differences across SES groups regarding language technology adoption and specific uses.

Contributions. Our contributions are:

- A survey of language technology use across sociodemographic groups;
- The first dataset of real prompts annotated with fine-grained sociodemographic information, including SES;[2]
- A quantitative and qualitative analysis of the differences between SES groups with respect to the use of language technologies.

Our findings expand current research on the AI Gap, and on understanding LLM usage among the general public, contributing to the development of more inclusive language technologies.

2 Related Work

Interest in AI and public attitudes towards it has skyrocketed in recent years, prompting several surveys (e.g., Scantamburlo et al., 2024) investigating how people use and perceive AI and LLMs. Digital divides across sociodemographic groups have been areas of concern for some time now, inspiring several studies about the desiderata of different groups (e.g., Blaschke et al., 2024; Lent et al., 2022). Cercas Curry et al. (2024a) show that NLP systems' performance may be affected by the socioeconomic status of speakers using film and TV shows. Daepp and Counts (2024) analyze the most common intents towards chatGPT in different regions of the US and find evidence of an emerging AI Gap between regions with higher and lower incomes. Recently, several large-scale collections of prompts and interactions have been collected to understand the broad applications of LLMs, such as Kirk et al. (2024), Trippas et al. (2024), Zheng et al. (2023), Huggingface's ShareGPT datasets,[3] and Zhao et al. (2024). However, none of these have have collected information on SES. We fill the gap by surveying the usage of language technologies in general and of LLMs, collect prompts from previous interaction of the participants with AI chatbots, and analyze differences across SES groups. We highlight the importance of exploring differences across SES groups to limit the AI Gap.

3 Socioeconomic Status

Socioeconomic status (SES) refers to an individual or group's social and economic position. A person's SES[4] is a function of their economic, social, and cultural capital – factors such as income, education, occupation, and wealth typically influence the SES of an individual. Still, they are often insufficient to determine it (Bourdieu, 1987).

[2] https://huggingface.co/datasets/MilaNLProc/survey-language-technologies

[3] ShareGPT Datasets

[4] There are different social stratification systems depending on the culture of reference (e.g., the Indian caste system, Indigenous American clans, or tribes). Our study focuses on U.S. and U.K. speakers, so we refer to the Western European class model. For a broader discussion on the intricacies of social stratification systems, see (Savage and Mouncey, 2016).

SES influences almost all aspects of an individual's life: hobbies, social circle, access to experiences, and even language (Labov, 2006; Savage and Mouncey, 2016). People's perception of where they stand regarding socioeconomic status has important psychological effects, supporting the idea that subjective class is an important measure of socioeconomic status (Cercas Curry et al., 2024b). To assess SES, the typical setup is to ask participants to place themselves on the socioeconomic ladder ranging from one to ten following the Macarthur scale (Adler et al., 2000) where higher levels represent those who are more privileged.

SES and the Digital Divide: SES impacts the digital divide by shaping access to technology, digital skills, and the ability to leverage digital tools for economic and social mobility (Mubarak et al., 2020). Higher SES affords access to better devices and paid AI services. Cultural capital and *habitus* influence how people engage with technology – higher social classes may develop advanced digital literacy, using AI for learning and professional growth. These disparities may reinforce existing inequalities, as those with greater digital access and skills gain further advantages in education, employment, and social influence. With this respect, Capraro et al. (2024) posit that generative AI will widen the already existing digital divide.

4 Survey Setup

In our survey, we include three types of question: Sociodemographic information, inquiries about the usage of language technologies, and prompt collections (see Figure 1).[5] The first section includes 17 questions that aim to collect basic demographics (such as age and gender) as well as information about the socioeconomic background. We ask about participants' perception of their SES with respect to the Macarthur scale, as well as other individual factors such as level of education, parents' occupation and hobbies. Subjects could opt out of supplying this information. All information was treated in compliance with GDPR, in that subjects are fully anonymized (we have no way of connecting information to subjects, and the combination of features could not identify individuals). See also Section Ethical Considerations. While we cannot individually verify the sociodemographic information, we can match the information provided in our survey with the demographic profile they provide in Prolific. We find that less than 2.5% of the participants provide conflicting information (regarding gender, ethnicity or age).

For the second part of the survey, we are inspired by Lent et al. (2022) and give a broad definition of "language technologies" before asking participants about their experience these technologies (e.g., Which of the following language technologies have you used?).[6] Additionally, the second part includes questions which are more specific to the usage of LLMs, defined as "AI chatbots like ChatGPT or other similar chatbots". We ask about the frequency of usage, the applications (e.g., coding, brainstorming, writing) and the contexts of usage (e.g., working, learning, personal). For all the questions in the first and second sections, participants could select "Other" or "Prefer not to say".

In the last section, we focus on the LLMs' usage and ask participants to provide the last ten prompts used in their interaction with AI chatbots (participants are free to look at their chat log).

4.1 Pilot Studies

To refine our survey, we conduct three pilot studies with 20, 20, and 79 participants, respectively. These pilots served two main purposes: first, to test the technical robustness of our self-implemented `streamlit` app[7] hosting the survey, and second, to evaluate the clarity and constraints of our questions. Based on insights from the pilots, we refined some aspects of the survey. For instance, we introduce a requirement for participants who report using AI chatbots "every day" or "nearly every day" to provide at least five prompts. Additionally, we adjusted the wording of "please provide us with the last ten *prompts* you used for your chosen AI chatbot" to "please provide us with the last ten *questions or requests* you used for your chosen AI chatbot" to improve clarity. After the third pilot ran smoothly, we proceeded with the large-scale study.

4.2 Coverage

We distributed our survey using Prolific, a crowdsourcing platform with a wide and diverse partici-

[5] The whole survey can be found in Appendix A.

[6] *Language technology refers to any piece of software that is intended to assist humans with language specific tasks in a technological setting (i.e., on a mobile phone, tablet, computer, the internet, smart devices). Some examples of language technologies include: Spell checkers in e-mail helps people to write more professional e-mails; Google Translate helps people to translate text from one language to another; internet search engines (e.g. Google, Bing, Yahoo) help people to find websites relevant to a given query.* (Lent et al., 2022)

[7] https://streamlit.io/

pant pool from all around the globe.[8] The platform allows the selection of participants based on an extensive range of fine-grained demographic criteria. Crowdsourcing platforms are limited in terms of the population reached, but we still observe significant differences within the socioeconomic spectrum of individuals on Prolific (see Sections 5.1 and 5.2). We expect these differences to be more pronounced in a real-world population distribution.

We selected participants to be English native speakers. Except for one of our pilot studies, where we only constrained participants based on their first language, all other pilots and the large-scale study additionally required participants to reside in the United Kingdom (UK) or the United States (US). While we acknowledge that opening the participation to more countries and languages would give a broader perspective on the usage of language technologies, we decide to restrict our study to participants located in the UK and in the US. Linguistic differences between groups are not uniform across all English-speaking populations, nor are available models and resources. The complexity and the cultural dependency that would need to be addressed in a wider setup. Given our available resources, our sample would have not being representative.

We conducted the large-scale study in two phases. The first phase included 501 participants, while the second phase focused on a targeted sample of 380 participants from the low and upper social strata (see definition in Section 3). This targeted approach was necessary due to the high representation of middle-class individuals in the initial round.

5 Results

We collect a total of 1,000 answers to our survey. In Figure 2 we report the distribution of our participants over the Macarthur scale. For this study we refer to the self-reported SES and map the Macarthur scale's values to the Western hierarchical class system (see Section 3) by binning 1-3 into lower, 4-7 middle, and 8-10 into upper. We also collect data about other socioeconomic factors that have been previously used in NLP research as proxy information for estimating SES, such as education (Cercas Curry et al., 2024b). We hope this will encourage future research in NLP to consider socioeconomic status, either self-reported or via proxy factors, depending on the use case. We report

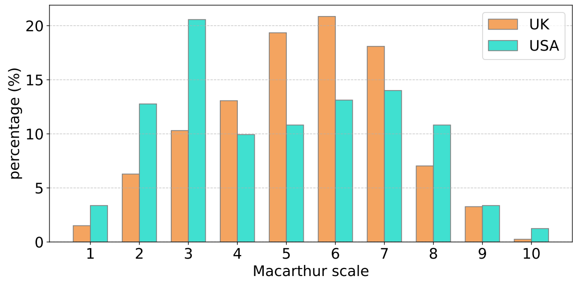


Figure 2: Distribution of participants according to the Macarthur scale (Adler et al., 2000).

in Appendix B the statistics about the level of education of the participant and of their parents, their current employment, the occupation of the participant and of the parents according to the European Skills, Competences, Qualifications and Occupations taxonomy (ESCO; le Vrang et al., 2014), the housing status (i.e., whether they own or rent) and their hobbies (following the list provided by Great British Class Survey; Savage et al., 2013). Additionally, in Appendix B we report more detailed demographics of our participants including their gender, age, nationality, the ethnicity, the marital status and religion.

5.1 Usage of Language Technologies

First, we analyze the daily access to digital devices (i.e., smartphone, tablet, laptop, smartwatch) by individuals from low, middle and upper social classes, to check if there are significant differences in the way people may access AI chatbots. Figure 3 shows that the percentage of daily access to smartphones is similar for all social classes. Differences appear instead in the daily access to tablets, laptops and smartwatches, with an usage increase for higher classes (middle and upper). The differences in daily access to digital devices are statistically significant across the three social classes, as indicated by a chi-square test of independence, χ^2 (df = 8, N = 2739) = 55.11, $p < 0.001$.[9] This suggests a strong association between socioeconomic status and daily access of digital devices.

In Appendix C we report the broad statistics relative to the type of language technologies mostly used by individuals from different SES (e.g., spell checkers, dialogue systems, speech-to-text etc.). Below instead we focus on the usage of AI generative chatbots, like ChatGPT. We identify a reverse trend in the frequency of usage of AI chatbots

[8] https://www.prolific.com/

[9] Note that this is a multiple-choice question.

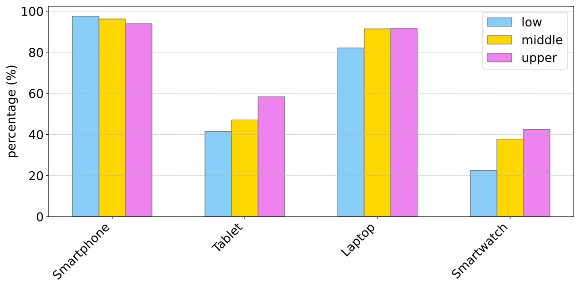


Figure 3: Daily access to digital devices from low, middle, upper classes.

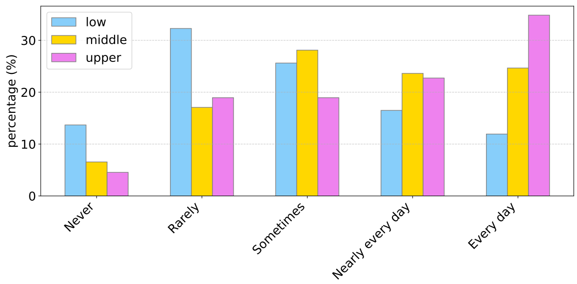


Figure 4: Frequency of usage of AI chatbots by individual from low, middle, upper classes.

across people from low and upper social classes (see Figure 4). Individuals from low socioeconomic backgrounds (values from 1 to 3 on the Macarthur scale) tend to use the chatbots less often– see decreasing values from 'Rarely' to 'Every day' in Figure 4. On the other side, individuals which placed themselves from 8 to 10 on the Macarthur scale have increasing values from 'Never' to 'Every day'. These differences are statistically significant, as indicated by a chi-square test of independence, χ^2 (df = 8, N = 997) = 67.79, $p < 0.001$, suggesting a strong association between SES and frequency of usage of AI chatbots.

In our analysis about the context of usage of AI chatbots (see Figure 6) we identify a distinct higher usage in the contexts of 'Work', 'School/University' and 'Learning' by individuals from the middle and upper classes. Additionally, the upper class uses AI chatbots more often in the 'Personal' and in the 'Technical' context, where a qualitative analysis revealed a consistent chunck of questions about machine learning. On the other side, individuals from the low class tend to interact more with AI chatbots in the contexts of 'Entertainment'. These differences are statistically significant, as indicated by a chi-square test of independence, χ^2 (df = 14, N = 2717) = 46.65, $p < 0.001$,[10] suggesting a significant association between SES and the context of AI chatbot usage.

Last, we ask more specifically about the tasks performed with the AI chatbots, and report the statistics in Figure 5. We identify that people from higher social classes (i.e., middle and upper) use the LLMs more frequently for writing and writing related tasks (i.e., paraphrasing, proofreading/editing, summarizing), and for more technical tasks (i.e., coding, solving mathematical and logical problems, analyzing data). On the other hand, individuals from the low class tend to use the AI chatbots for more generic tasks like brainstorming, generic chatbot conversations and answering questions about general knowledge. We perform a chi-squared statistical test, χ^2 (df = 36, N = 4809) = 88.40, $p < 0.001$[11] revealing a strong association between SES and tasks performed with AI chatbots.

5.2 Linguistic Analysis of the Prompts

We collect a total of 6,482 real prompts used by our participants in previous interactions with LLMs. We perform a linguistic analysis to investigate key characteristics, with a focus on how prompting strategies differ among individuals with varying socioeconomic backgrounds.

Prompt length. We find differences in the average length of prompts written by people from different SES. Specifically, individuals from higher social classes tend to write shorter and more concise prompts. The average length in terms of number of words is 27.0 for low, 22.3 for middle and 18.4 for upper class. A bootstrap resampling significance testing indicates a statistically significant difference between the low and upper classes ($p < 0.05$).[12] We speculate this divergence to be a consequence of higher class individuals having a wider vocabulary, which allows them to express themselves using less words.[13]

Concreteness. Bernstein (1960) posits that people from higher class families are more encouraged to use language for abstract thinking in contrast to people from lower class families, which are "limited" to more concrete concepts. To asses the level

[10] Note that this is a multiple-choice question.

[11] Note that this is a multiple-choice question.

[12] We use the implementation by Ulmer et al. (2022).

[13] We test several metrics for computing the vocabulary diversity across social strata (including TTR and entropy), but do not find statistically significant differences. We attribute this to the restricted nature of the data itself, which consists of prompts, that inherently limit the range of vocabulary used.

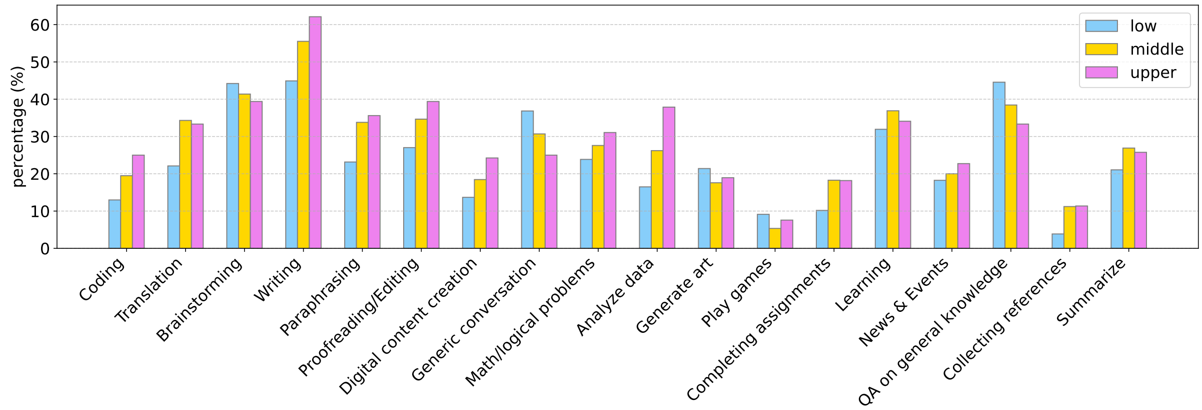


Figure 5: Tasks performed by individual from low, middle, upper classes with AI chatbots.

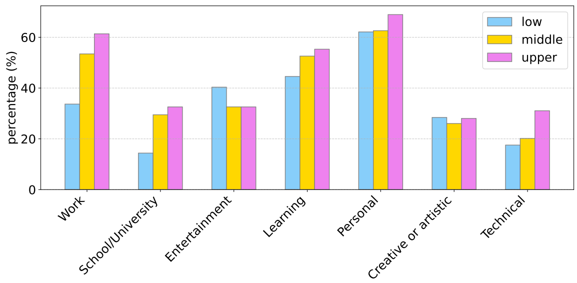


Figure 6: Context of usage of AI chatbots by individual from low, middle, upper classes.

of concreteness and abstraction of the collected prompts, we use the list of 40,000 common English words proposed by Brysbaert et al. (2014). In this collection, each word is evaluated on a scale from one (abstract) to five (concrete) by at least 25 participants. The average concreteness score of the prompts written by people from the low social class is 2.66, while it decreases (the lower the more abstract) to 2.63 and 2.57 respectively for people from the middle and upper class. While these differences are relatively small, the upper class differs significantly from the others based on bootstrap resampling ($p < 0.05$). Bernstein (1960)'s theory on the different level of concreteness of the language used by people from different socioeconomic background is reflected in the way people interact with LLMs: People from the upper social class interact with a more abstract language.

Logistic Regression Classifier. From the analysis above, we suspect that the prompts written by individual from different socioeconomic background are significantly different and easily distinguishable. To confirm our hypothesis, we train a simplistic bag-of-word classifier, which achieves a Macro-F1 score of 39.25, compared to a majority baseline of 25.02 Macro-F1. Since we split our data into train and evaluation sets using random seeds, we apply the Almost Stochastic Order test (Del Barrio et al., 2018; Dror et al., 2019) as implemented by Ulmer et al. (2022) with a confidence level $\alpha = 0.05$ to assess that the simplistic bag-of-word classifier is stochastically dominant over the majority baseline ($\epsilon_{\text{min}} = 0$).

5.3 Clustering of the Prompts

Methodology. To analyze the topics addressed in the collection of prompts, we perform topic modeling. Specifically, we (1) embed the prompts using SentenceTransformer (Reimers and Gurevych, 2019)[14] and M3-Embedding (Chen et al., 2024),[15] (2) cluster them using UMAP (McInnes et al., 2018) and HDBSCAN (McInnes et al., 2017), (3) assign them a short and distinct description using GPT-4 (OpenAI et al., 2024)[16] and (4) manually evaluate the clusters and the descriptions.

Clustering results. We find commonalities in the topics identified across the three social classes. These include for example *translation* (e.g., "Translate Good morning into Japanese.", "translate to dutch 'I am fine thank you'"), *mental health* (e.g., "How do i best overcome social anxiety ?", "Give me advice on living with depression"), *medical advice* (e.g., "what are the dangers of etopic pregnancy", "Why does my 1 year old have a running nose all the time"), *writing* (e.g., "write an essay of about the dangers of using AI to generate text", "Hi I'm in need of some ideas for what to write in my wife's birthday card.") and *text edit-*

[14]We use `all-MiniLM-L6-v2`.
[15]We use `bge-large-en-v1.5`
[16]We use GPT-4o.

	Low	Middle	Upper
Finance	What is the cheapest place to live in the US?	write a business plan for an agricultural based business	What are the current crypto market trend
	Do I need to be a member of a credit union to apply for their loans	here are my monthly earnings and spends , show me how to save money	when should I buy a house
	What is the best way to reduce my debt	How do I start a successful small business?	how long should I invest for?
Job	What are some WFH jobs that require no experience or degrees.	Please can you write an application letter to a school for work experience.	Create a cover letter for a new role as a communication manager
	Please write an email to get some experience and/or paid employment	Please write a covering letter to an employer which summarises my skills.	Can you suggest some effective leadership strategies for managing a team
	Write me a covering letter for a restauraunt job	Write a cover letter for assistant accountatn to match the below job description.	How can I improve my email writing to sound more professional?
Food	Hello...what is the best way for me to reheat my Buffalo Wild Wings?	What are some creative uses for leftover rice?	Give me some healthy meal ideas I could cook for a family of three.
	give me new recipes	what can I cook with leftover pork	What wine with red tuna belly
	I don't have peas and carrots	how to make gingerbread	What's the best filtered water pitcher? Preferably one that eliminates any of the following: PFAS, PFOAS, microplastics and/or nanoplastics

Table 1: Example of prompts written by individuals from different social classes, divided by macro-topics.

ing (e.g., "Rewrite this text in formal English.", "Please rephrase the below email to sound more professional and authoritive").

We also find some topics to be distinctive for specific social classes. This is especially valid for the upper class, where we identify a cluster related to *travel destinations* (e.g., "I am planning to travel for a vacation in Japan, do you have any recommendations?", "4 week itineray for seniors, travelling in Vietnam") and several clusters related to abstract concepts like *inclusivity* (e.g., "How can we create a more inclusive environment for all genders?", "What are some specific actions, practices, or class features that make you feel most supported and valued as an LGBTQ+ student in a college course environment?") and *good communication* (e.g., "How can we improve internal communication across different departments within our organization?", "Barriers of good communication").

Finally, some topics are addressed by individual from the whole socioeconomic spectrum, but within varying framings (see examples in Table 1). Among these, the most prominent is *finance*. While within the low class we find advices for money saving, in the corresponding upper class cluster there are requests for investments advice. We also identify a common *job* cluster across all three social classes. Within this, the requests from the middle class mostly involve job applications for specific positions, the ones from the low class suggestions for more generic low-skilled jobs, and the ones from the upper class often imply an high-level job positions of the user. Last, we identify a cluster of prompts related to *food*. Here, the upper class is the most distinct with requests specifically targeting healthy and expensive dishes.

5.4 User Perceptions

User perceptions of a system, whether they perceive it as a tool or something else, affect how they interact with and use it (Delcker et al., 2024). Cues suggesting humanlikeness (such as the use of natural language) trigger social scripts and a mental model of a system with humanlike qualities (Reeves and Nass, 1996). At a meta-level, human metaphors (such as deep *learning*) are common when discussing AI (Ye and Li, 2024). Metaphorical language plays an important role in understanding complex systems (Lakoff and Johnson, 2008) but the use of human metaphor can convey more humanlikeness than intended (Epley et al., 2007). There are growing concerns about anthropomorphism in systems, its implications for digital literacy and how these may lead to overreliance (Abercrombie et al., 2023; Akbulut et al., 2024), potentially differently between social groups.

Anthropomorphism. Although some metrics have been proposed to measure anthropomorphism in the models' outputs (e.g. Cheng et al., 2024), to the best of our knowledge, no such metrics ex-

	Jargon	Metaphor	Phatic	Verbs
Low	3.32	25.07	6.34	63.04
Middle	4.16	23.72	5.08	64.98
High	4.94	23.49	4.29	63.42

Table 2: Mean percentage of prompts that contain jargon, metaphors, phatic expressions, and verbs.

ist for user prompts. Instead, we study linguistic markers typical of human-human dialogue and metaphorical language which may be indicative of the user's mental model of the system. We measure (1) the use of politeness markers, e.g., *thank you*, and phatic expressions such as *hi*; (2) the use of metaphorical verbs and jargon (e.g., *write* vs *generate*), and 3) the use of complete sentences (e.g., "weather in Rome" vs. "What's the weather like in Rome"), which convey a naturalness not always necessary. We use keyword spotting for (1) and (2) using manually compiled lists (more details in Appendix E). For (3) we use SpaCy to find whether the prompt contains a verb. The results are shown in Table 2. We find general trends in the data: jargon is more common for upper and middle class participants, while metaphorical language and phatic expressions are more common among lower SES participants. However, we do not find these differences to be statistically significant.

Search engine questions. We investigate the extent to which individuals are replacing the use of search engines with LLMs. As a proxy, we find that a notable proportion of 46.6%, 43.5% and 45.4% of the prompts written by people with low, middle and upper socioeconomic background respectively contains at least one of the question words: "who", "what", "when", "where", "why", "how". In this case, the trend is that people from the low and upper classes tend to make slightly more usage of the LLMs with questions like 'Which country celebrates new years first?' or 'What is the difference between Espresso and regular whole bean coffee?'.

6 Discussion and Future Directions

We show that the adoption rates and use of language technologies vary significantly based on the SES of the users. In terms of contexts where the interactions take place, we find that mid- and higher-SES participants use LLMs more commonly for work and education. These differences may be explained by matters of access, digital literacy or *habitus*, but they may exacerbate existing inequalities. In a recent report, UNESCO (2023) has brought attention to this gap, and Capraro et al. (2024) posit that the AI Gap will lead to further inequality, as certain communities benefit more from the advantages of language technologies, while already marginalised communities are increasingly left behind. For example, the benefits of generative AI in the workplace are centred around middle-class jobs, as shown in our results and recently reported in a recent report by Anthropic on the economic tasks of LLMs (Handa et al., 2025).

We also find potential for concern in terms of the robustness of current evaluation benchmarks. The applications associated with higher SES participants (such as paraphrasing, summarizing, and mathematical problems) are generally suited for ground-truth evaluations (e.g. Hendrycks et al.), while the tasks more often reported by lower SES uses rely more heavily on human preference evaluation. Our results also support the notion that there are significant linguistic differences between groups of different SES. Although in recent years there has been growing interest in human-centred evaluation (e.g. Xiao et al., 2024; Blodgett et al., 2024; Ibrahim et al., 2024), participatory design (such as Caselli et al., 2021) and perspectivism (Frenda et al., 2024), currently no benchmarks exist to quantify how SES differences affect NLP systems or what their real-world potential impact may be. Future work should focus on benchmarking model performance in realistic scenarios that represent the full socioeconomic spectrum, aiming to create resources and systems that address and mitigate the digital divide in NLP technologies.

7 Conclusion

We survey the usage of language technologies among individuals with different SES. We find statistically significant differences both in the adoption of language technologies and in the specific uses people give them. In particular, we find that upper class individuals have access to a wider variety of digital devices, use AI chatbots more frequently and with the goal to improve their work through more technical tasks like coding, data analysis or writing. We collect 6,482 prompts from previous interactions of our participants with LLMs, where we find statistically significant differences in the length and concreteness level across SES groups. From a qualitative analysis, we find further differences in the topics and framings of the prompts, and in the user perceptions of the systems (i.e., anthropomorphism). Our work calls for the development of inclusive NLP technologies to accommodate different SES needs and habitus and mitigate the existing AI Gap.

Limitations

Our study is limited to U.S.- and U.K.-based crowdworkers on the Prolific platform, and may not be representative of the broader population. In terms of socioeconomic status, we expect the Prolific population to be skewed towards the middle to low social class. Furthermore, given crowdworkers' familiarity with technology, they may be more likely to use language technologies than the general population. Crowdworkers may use LLMs to complete the survey itself, e.g., by generating ten example prompts rather than providing their own, despite platform policies against the use of LLMs. We use the Macarthur Scale for measuring SES: Subjective metrics are prone to ambiguity and bias – with most people comparing themselves to their peers and thinking of themselves as being somewhere in the middle. However, we also note that one's perception of their SES plays an important role in behavior and attitudes and for this reason we chose to use it in this survey.

Ethical Considerations

The survey was approved by the ethics board of the IT University of Copenhagen. Crowdworkers were compensated for their time in accordance to the platform's recommendation of £9/hour. They may withdraw from the study by contacting the researchers. Subjects were fully anonymized in compliance with GDPR, and could opt out of supplying sensitive sociodemographic information.

Acknowledgments

We thank the MilaNLP group in Bocconi University and the NLPnorth group at ITU for feedback on earlier version of this draft, as well as the reviewers for their helpful comments. A special thanks to Mike Zhang for help with the design of Figure 1. Elisa Bassignana is supported by a research grant (VIL59826) from VILLUM FONDEN. Dirk Hovy was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 949944, INTEGRATOR). He is a member the Data and Marketing Insights Unit of the Bocconi Institute for Data Science and Analysis (BIDSA).

References

Gavin Abercrombie, Amanda Cercas Curry, Tanvi Dinkar, Verena Rieser, and Zeerak Talat. 2023. Mirages. on anthropomorphism in dialogue systems. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 4776–4790, Singapore. Association for Computational Linguistics.

Nancy E. Adler, Elissa S. Epel, Grace Castellazzo, and Jeannette R. Ickovics. 2000. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 19 6:586–92.

Canfer Akbulut, Laura Weidinger, Arianna Manzini, Iason Gabriel, and Verena Rieser. 2024. All too human? mapping and mitigating the risk from anthropomorphic ai. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 7(1):13–26.

Basil Bernstein. 1960. Language and social class. *The British Journal of Sociology*, 11(3):271–276.

Verena Blaschke, Christoph Purschke, Hinrich Schuetze, and Barbara Plank. 2024. What do dialect speakers want? a survey of attitudes towards language technology for German dialects. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 823–841, Bangkok, Thailand. Association for Computational Linguistics.

Su Lin Blodgett, Jackie Chi Kit Cheung, Vera Liao, and Ziang Xiao. 2024. Human-centered evaluation of language technologies. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Tutorial Abstracts*, pages 39–43, Miami, Florida, USA. Association for Computational Linguistics.

Su Lin Blodgett and Zeerak Talat. 2024. Llms produce racist output when prompted in african american english.

Pierre Bourdieu. 1987. What makes a social class? on the theoretical and practical existence of groups. *Berkeley journal of sociology*, 32:1–17.

Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior Research Methods*, 46:904–911.

Valerio Capraro, Austin Lentsch, Daron Acemoglu, Selin Akgun, Aisel Akhmedova, Ennio Bilancini, Jean-François Bonnefon, Pablo Brañas-Garza, Luigi Butera, Karen M Douglas, et al. 2024. The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS nexus*, 3(6).

Tommaso Caselli, Roberto Cibin, Costanza Conforti, Enrique Encinas, and Maurizio Teli. 2021. Guiding principles for participatory design-inspired natural language processing. In *Proceedings of the 1st Workshop on NLP for Positive Impact*, pages 27–35, Online. Association for Computational Linguistics.

Amanda Cercas Curry, Giuseppe Attanasio, Zeerak Talat, and Dirk Hovy. 2024a. Classist tools: Social class correlates with performance in NLP. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12643–12655, Bangkok, Thailand. Association for Computational Linguistics.

Amanda Cercas Curry, Zeerak Talat, and Dirk Hovy. 2024b. Impoverished language technology: The lack of (social) class in NLP. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8675–8682, Torino, Italia. ELRA and ICCL.

Jianlyu Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, and Zheng Liu. 2024. M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 2318–2335, Bangkok, Thailand. Association for Computational Linguistics.

Myra Cheng, Kristina Gligoric, Tiziano Piccardi, and Dan Jurafsky. 2024. AnthroScore: A computational linguistic measure of anthropomorphism. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 807–825, St. Julian's, Malta. Association for Computational Linguistics.

Madeleine IG Daepp and Scott Counts. 2024. The emerging ai divide in the united states. *arXiv preprint arXiv:2404.11988*.

Fred D Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340.

DeepSeek-AI. 2024. Deepseek-v3 technical report. *Preprint*, arXiv:2412.19437.

Eustasio Del Barrio, Juan A Cuesta-Albertos, and Carlos Matrán. 2018. An optimal transportation approach for assessing almost stochastic order. In *The Mathematics of the Uncertain*, pages 33–44. Springer.

Jan Delcker, Joana Heil, Dirk Ifenthaler, Sabine Seufert, and Lukas Spirgi. 2024. First-year students ai-competence as a predictor for intended and de facto use of ai-tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1):18.

Rotem Dror, Segev Shlomov, and Roi Reichart. 2019. Deep dominance - how to properly compare deep neural models. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2773–2785, Florence, Italy. Association for Computational Linguistics.

Nicholas Epley, Adam Waytz, and John T Cacioppo. 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114(4):864.

Simona Frenda, Gavin Abercrombie, Valerio Basile, Alessandro Pedrani, Raffaella Panizzon, Alessandra Teresa Cignarella, Cristina Marco, and Davide Bernardi. 2024. Perspectivist approaches to natural language processing: A survey. *Language Resources and Evaluation*.

Kunal Handa, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, Sarah Heck, Jared Mueller, Jerry Hong, Stuart Ritchie, Tim Belonax, et al. 2025. Which economic tasks are performed with ai? evidence from millions of claude conversations.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding. In *International Conference on Learning Representations*.

Dirk Hovy, Shiri Melumad, and J Jeffrey Inman. 2021. Wordify: A Tool for Discovering and Differentiating Consumer Vocabularies. *Journal of Consumer Research*, 48(3):394–414.

Lujain Ibrahim, Saffron Huang, Lama Ahmad, and Markus Anderljung. 2024. Beyond static ai evaluations: advancing human interaction evaluations for llm harms and risks. *arXiv preprint arXiv:2405.10632*.

Hannah Rose Kirk, Alexander Whitefield, Paul Röttger, Andrew Michael Bean, Katerina Margatina, Rafael Mosquera, Juan Manuel Ciro, Max Bartolo, Adina Williams, He He, et al. 2024. The prism alignment dataset: What participatory, representative and individualised human feedback reveals about the subjective and multicultural alignment of large language models. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.

William Labov. 2006. *The Social Stratification of English in New York City*, 2 edition. Cambridge University Press.

George Lakoff and Mark Johnson. 2008. *Metaphors we live by*. University of Chicago press.

Martin le Vrang, Agis Papantoniou, Erika Pauwels, Pieter Fannes, Dominique Vandensteen, and Johan De Smedt. 2014. Esco: Boosting job matching in europe with semantic interoperability. *Computer*, 47(10):57–64.

Heather Lent, Kelechi Ogueji, Miryam de Lhoneux, Orevaoghene Ahia, and Anders Søgaard. 2022. What a creole wants, what a creole needs. In *Proceedings of* the Thirteenth Language Resources and Evaluation Conference*, pages 6439–6449, Marseille, France. European Language Resources Association.

Q Vera Liao and Ziang Xiao. 2023. Rethinking model evaluation as narrowing the socio-technical gap. *arXiv preprint arXiv:2306.03100*.

Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205.

Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. 2018. Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29):861.

Nicolai Meinshausen and Peter Bühlmann. 2010. Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(4):417–473.

Farooq Mubarak, Reima Suomi, and Satu-Päivi Kantola. 2020. Confirming the links between socio-economic variables and digitalization worldwide: the unsettled debate on digital divide. *Journal of Information, Communication and Ethics in Society*, 18(3):415–430.

OpenAI. 2023. Chatgpt: Large language model. `https://chat.openai.com`.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong, Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder, Mario Saltarelli, Ted Sanders, Shibani Santurkar, Girish Sastry, Heather Schmidt, David Schnurr, John Schulman, Daniel Selsam, Kyla Sheppard, Toki Sherbakov, Jessica Shieh, Sarah Shoker, Pranav Shyam, Szymon Sidor, Eric Sigler, Maddie Simens, Jordan Sitkin, Katarina Slama, Ian Sohl, Benjamin Sokolowsky, Yang Song, Natalie Staudacher, Felipe Petroski Such, Natalie Summers, Ilya Sutskever, Jie Tang, Nikolas Tezak, Madeleine B. Thompson, Phil Tillet, Amin Tootoonchian, Elizabeth Tseng, Preston Tuggle, Nick Turley, Jerry Tworek, Juan Felipe Cerón Uribe, Andrea Vallone, Arun Vijayvergiya, Chelsea Voss, Carroll Wainwright, Justin Jay Wang, Alvin Wang, Ben Wang, Jonathan Ward, Jason Wei, CJ Weinmann, Akila Welihinda, Peter Welinder, Jiayi Weng, Lilian Weng, Matt Wiethoff, Dave Willner, Clemens Winter, Samuel Wolrich, Hannah Wong, Lauren Workman, Sherwin Wu, Jeff Wu, Michael Wu, Kai Xiao, Tao Xu, Sarah Yoo, Kevin Yu, Qiming Yuan, Wojciech Zaremba, Rowan Zellers, Chong Zhang, Marvin Zhang, Shengjia Zhao, Tianhao Zheng, Juntang Zhuang, William Zhuk, and Barret Zoph. 2024. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

Byron Reeves and Clifford Nass. 1996. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK*, 10(10):19–36.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing* *and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.

Manon Reusens, Philipp Borchert, Jochen De Weerdt, and Bart Baesens. 2024. Native design bias: Studying the impact of english nativeness on language model performance. *arXiv preprint arXiv:2406.17385*.

Mike Savage, Fiona Devine, Niall Cunningham, Mark Taylor, Yaojun Li, Johs Hjellbrekke, Brigitte Le Roux, Sam Friedman, and Andrew Miles. 2013. A new model of social class? findings from the bbc's great british class survey experiment. *Sociology*, 47(2):219–250.

Mike Savage and Peter Mouncey. 2016. Social class in the 21st century.

Teresa Scantamburlo, Atia Cortés, Francesca Foffano, Cristian Barrué, Veronica Distefano, Long Pham, and Alessandro Fabris. 2024. Artificial intelligence across europe: A study on awareness, attitude and trust. *IEEE Transactions on Artificial Intelligence*.

Johanne R Trippas, Sara Fahad Dawood Al Lawati, Joel Mackenzie, and Luke Gallagher. 2024. What do users really ask large language models? an initial log analysis of google bard interactions in the wild. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2703–2707.

Dennis Ulmer, Christian Hardmeier, and Jes Frellsen. 2022. deep-significance: Easy and meaningful signifcance testing in the age of neural networks. In *ML Evaluation Standards Workshop at the Tenth International Conference on Learning Representations*.

UNESCO. 2023. Ai literacy and the new digital divide: A global call to action.

Ziang Xiao, Wesley Hanwen Deng, Michelle S. Lam, Motahhare Eslami, Juho Kim, Mina Lee, and Q. Vera Liao. 2024. Human-centered evaluation and auditing of language models. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA. Association for Computing Machinery.

Zhanglei Ye and Jian Li. 2024. Artificial intelligence through the lens of metaphor: Analyzing the eu aia. *International Journal of Digital Law and Governance*, 1(2):361–381.

Wenting Zhao, Xiang Ren, Jack Hessel, Claire Cardie, Yejin Choi, and Yuntian Deng. 2024. Wildchat: 1m chatgpt interaction logs in the wild. In *The Twelfth International Conference on Learning Representations*.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Tianle Li, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zhuohan Li, Zi Lin, Eric. P Xing, Joseph E. Gonzalez, Ion Stoica, and Hao Zhang. 2023. Lmsys-chat-1m: A large-scale real-world llm conversation dataset. *Preprint*, arXiv:2309.11998.

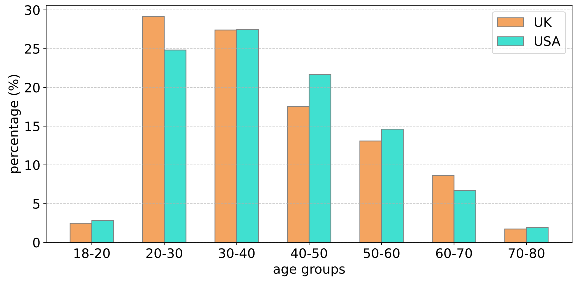


Figure 7: Distribution of participants according to the age group and divided by country of residence.

A Survey Interface

In Figure 27 we report the interface of our survey, including all the questions.

B Demographic Statistics

We report below all the information regarding the demographics and the socioeconomic factors of our participants. The nationality is split as 55.3% US, 37.8% UK, 6.9% 'Other' including double nationality cases (mostly Canada and Nigeria). Of the total of our 1,000 participants, 52.5% identify as men, 45.6% women, 1.6% non-binary and 0.3% prefer not to say. We plot the other statistics, i.e., age (Figure 7), ethnicity (Figure 8), marital status (Figure 9), religion (Figure 10), level of education (Figure 11), mother's level of education (Figure 12), father's level of education (Figure 13), employment status (Figure 14), occupation according to the European Skills, Competences, Qualifications and Occupations taxonomy (ESCO; le Vrang et al., 2014) (Figure 15), mother's occupation (Figure 16), father's occupation (Figure 17), housing situation (Figure 18) and hobbies, following the list provided by Great British Class Survey (Savage et al., 2013) (Figure 19).

C Language Technologies

We report in Figure 20 the statistics related to the question 'Which of the following language technologies have you heard about?'. The distributions across the low, middle and upper classes are statistically significant, as indicated by a chi-square test of independence, χ^2 (df = 24, N = 6449) = 166.44, $p < 0.001$.

In Figure 21 we report the statistics relative to the question 'Which of the following language technologies have you used?'. The distributions across the low, middle and upper classes are statistically significant, as indicated by a chi-square test of independence, χ^2 (df = 24, N = 4889) = 166.34, $p <$

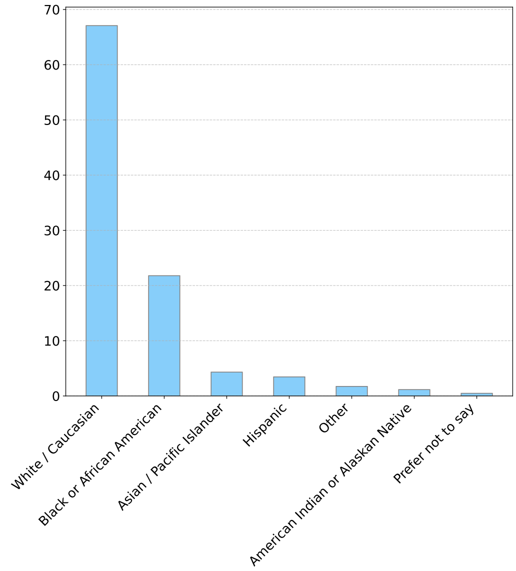


Figure 8: Distribution of participants' ethnicity.

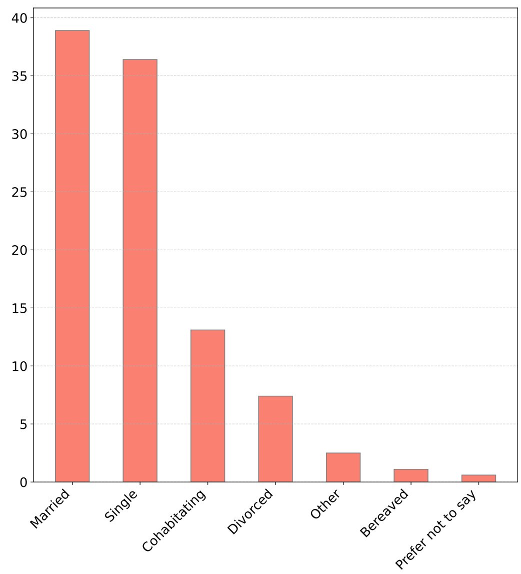


Figure 9: Distribution of participants' marital status.

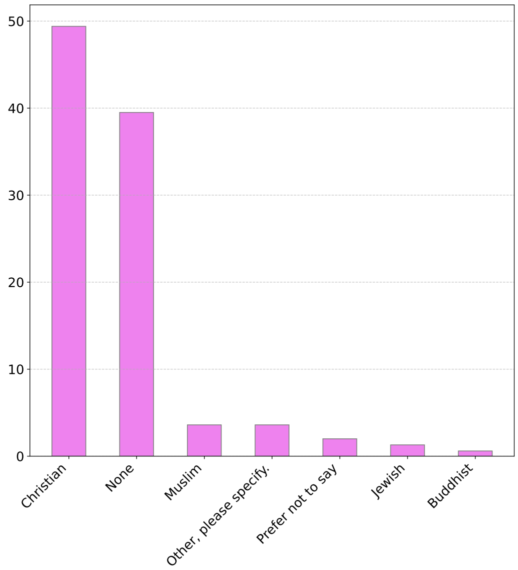


Figure 10: Distribution of participants' religion.

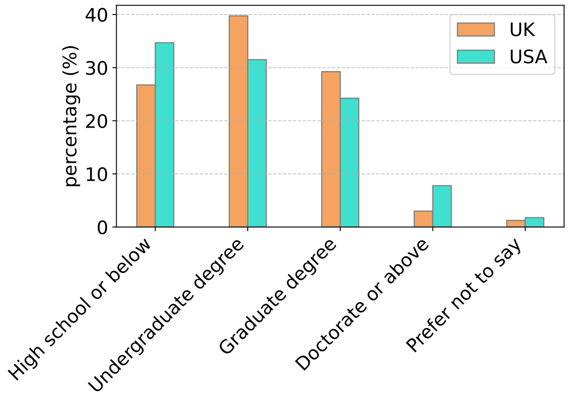


Figure 11: Distribution of participants according to the level of education and divided by country of residence.

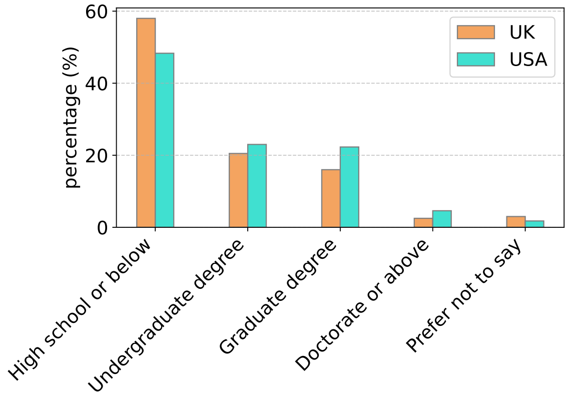


Figure 12: Distribution of participants according to the mother's level of education and divided by country of residence.

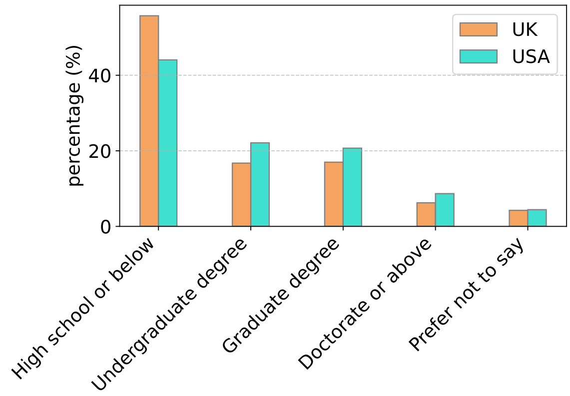


Figure 13: Distribution of participants according to the father's level of education and divided by country of residence.

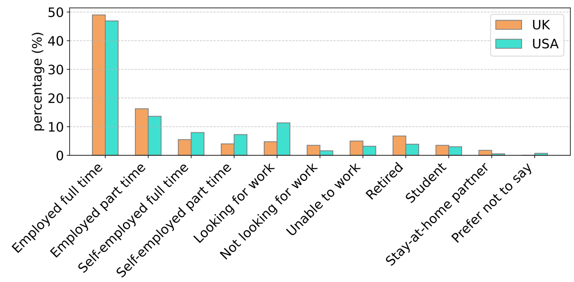


Figure 14: Distribution of participants according to the employment status and divided by country of residence.

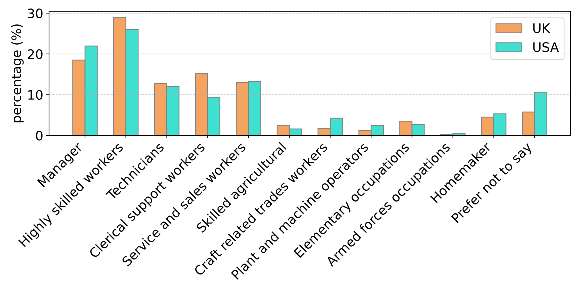


Figure 15: Distribution of participants according to the occupation and divided by country of residence.

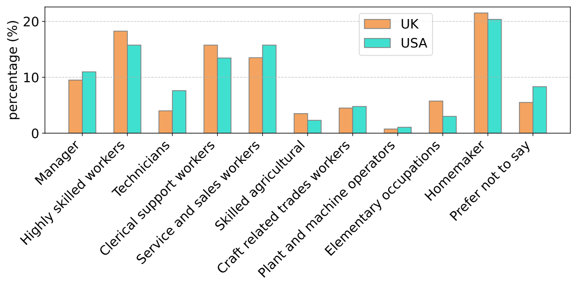


Figure 16: Distribution of participants according to the mother's occupation and divided by country of residence.

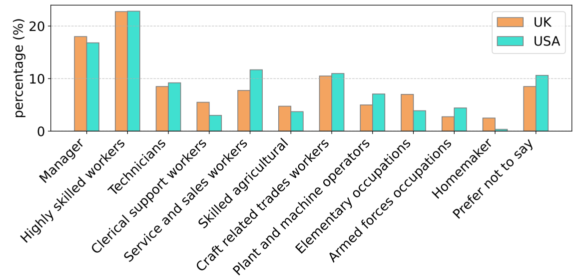


Figure 17: Distribution of participants according to the father's occupation and divided by country of residence.

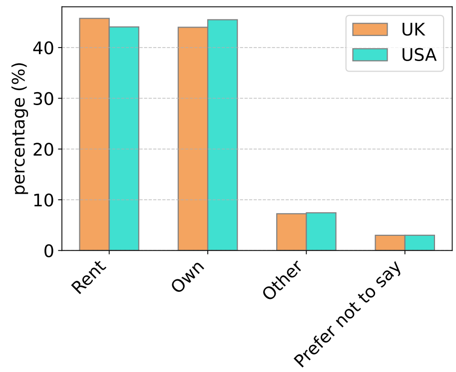


Figure 18: Distribution of participants according to the housing situation and divided by country of residence.

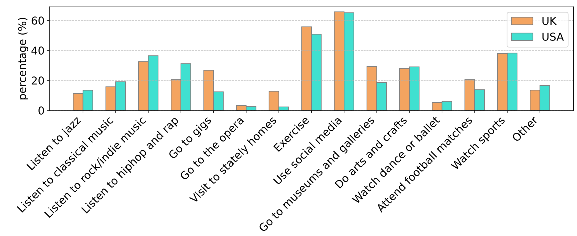


Figure 19: Distribution of participants according to the hobbies and divided by country of residence.

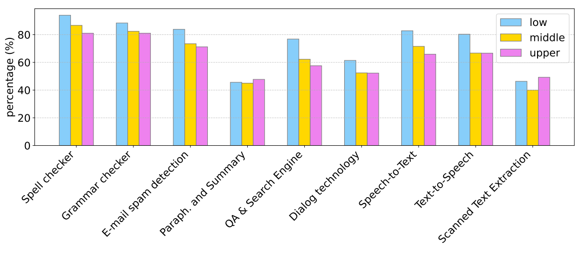


Figure 20: Type of language technologies known by individual from low, middle, upper classes.

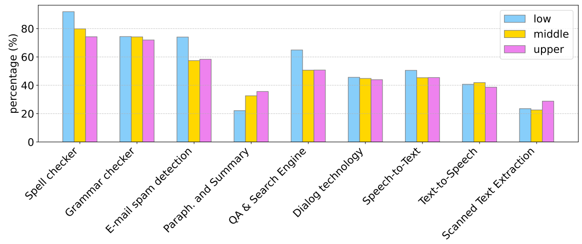


Figure 21: Type of language technologies used by individual from low, middle, upper classes.

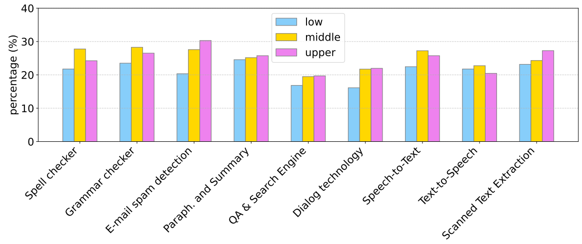


Figure 22: Type of language technologies that individual from low, middle, upper classes would like to use, but do not perform well enough.

0.001.

Figure 22 reports the statistics relative to the question 'Below is a list of some common language technologies. Please check every one that you would find useful, but do not use because of scarce performance'. The distributions across the low, middle and upper classes are statistically significant, as indicated by a chi-square test of independence, χ^2 (df = 24, N = 2490) = 115.65, $p <$ 0.001.

Last, in Figure 23 we focus on the usage of LLMs and report the statistics relative to the question 'If you use them, which of these AI chatbots do you use? If you have never used them, leave blank.'.

D Wordify

We use a variant of the Stability Selection algorithm (Meinshausen and Bühlmann, 2010) as im-

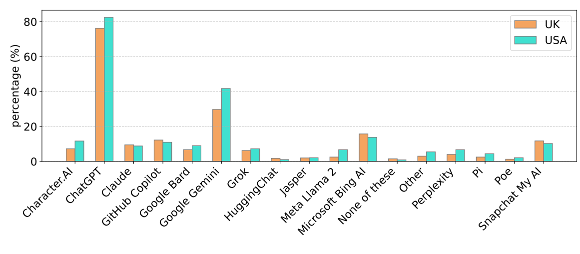


Figure 23: Type of LLMs used by individual divided by country of residence.

plemented by Hovy et al. (2021) to extract the most indicative n-grams for each social class (low, middle, upper). We do not find significant differences across SES groups. We report the results below.

Positive Indicators. Upper: kindly (0.414), british (0.398), employee (0.38), draft (0.364), sell (0.356), team (0.348), email (0.344), plan (0.344), donald (0.332), usa (0.332), what is (0.32), column (0.316), code (0.308), task (0.306), the monthly (0.306), manager (0.302), name of (0.302), or (0.3).

Middle: datum (0.394), add (0.376), why (0.362), fun (0.36), my (0.36), analysis (0.352), car (0.348), question (0.344), you know (0.342), hello (0.334), love (0.334), business (0.33), air (0.324), build (0.324), climate (0.322), website (0.322), cv (0.312), if you (0.312), which (0.312), customer (0.31), js (0.31), made (0.31), earth (0.308), order (0.306), holiday (0.304), meaning (0.304), quality (0.304), service (0.304), summarise (0.304), tiktok (0.3), true (0.3).

Low: how to (0.42), with (0.414), day (0.406), video (0.4), was (0.396), where (0.394), need (0.39), be (0.384), make (0.384), it (0.38), summarize (0.374), do (0.358), off (0.352), chicken (0.35), tell me about (0.346), description (0.344), total (0.344), resume (0.334), to cook (0.334), idea for (0.332), art (0.33), name (0.326), something (0.326), by (0.322), is (0.322), image (0.32), is the good (0.32), is there (0.32), someone (0.316), fill (0.314), tall (0.31), animal (0.306), country (0.306), how (0.304), ask (0.302), did (0.302), would (0.302), cheap (0.3), those (0.3), trade (0.3).

Negative Indicators. Upper: my (0.422), make (0.414), can (0.388), an (0.38), create (0.374), me (0.356), of (0.356), which (0.342), people (0.334), about (0.322), question (0.312), where (0.3).

Middle: used (0.408), summarize (0.406), ask (0.392), with (0.392), is there (0.38), is (0.376), trump (0.376), video (0.374), day (0.364), it (0.36), will (0.358), her (0.356), what are (0.356), to write (0.354), those (0.35), sell (0.34), hour (0.336), new (0.336), not (0.332), generate me (0.326), horror (0.326), idea for (0.322), how (0.32), well (0.32), creation (0.316), legal (0.316), country (0.312), do not (0.31), employee (0.31), fact (0.31), would (0.31), chicken (0.308), can use (0.306), or (0.306), you (0.306), how to make (0.304), up (0.304), play (0.302), trade (0.302).

Low: company (0.422), analysis (0.396), improve (0.386), you help (0.38), function (0.372), how does (0.372), datum (0.37), love (0.352), project (0.346), climate (0.338), if you (0.328), what is (0.324), business (0.322), are the (0.32), time (0.316), such (0.308), and (0.3), change (0.3).

E Anthropomorphism

Below we report the list of words used for the analysis of user perceptions with respect to anthropomorphism.

Phatic expression: thank, thanks, please, hi, hello.

Metaphorical verbs: ask, assess, care, choose, create, decide, describe, discover, empathize, engage, enjoy, evaluate, explain, express, feel, find, hear, imagine, improve, invent, judge, know, learn, listen, look, observe, plan, predict, prioritize, rate, react, reason, recommend, remember, respond, search, see, solve, speak, suggest, think, translate, understand, watch, write

Jargon: activate, compute, detect, evaluate, forecast, generate, ingest, infer, input, map, monitor, optimize, output, parse, prioritize, process, query, rank, rationalize, register, render, resolve, respond, retrieve, score, select, simulate, store, synthesize, track.

Tell us a bit about you

We are conducting research about the ways in which people of all backgrounds are using AI. To understand if there are differences in the ways different people are using AI chatbots and other technologies, we are running a survey.

Any data published will be fully anonymised.

USE TAB TO GO TO THE NEXT SECTION. DO NOT PRESS ENTER TO MOVE TO THE NEXT SECTION OR THE FORM WILL SUBMIT!!!

Gender*

Choose an option

If you selected other, please specify:

Age*

○ 18-24 ○ 25-34 ○ 35-44 ○ 45-54 ○ 55-60 ○ 60+

Nationality (You may select multiple):*

Select all that apply.

What is your ethnicity? You may select more than one.*

Select all that apply.

If you selected other, please specify:

What is your marital status?*

○ Married ○ Cohabitating ○ Bereaved ○ Divorced ○ Single
○ Other (Please Specify) ○ Prefer not to say

If you selected other, please specify:

What is your first language? You may select more than one, if applicable.*

Select all that apply.

If you selected other, please specify:

Do you have a religious affiliation? If so, which one?*

Choose an option

If you selected other, please specify:

Current education level*

Choose an option

What was the highest level of education achieved by your mother?*

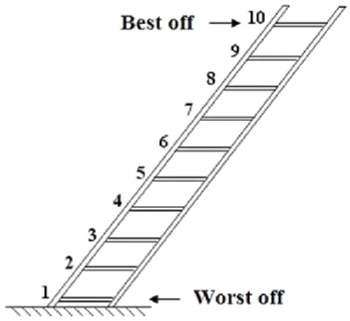
Choose an option

What was the highest level of education achieved by your father?*

Choose an option

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

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10
○ Prefer not to say



Macarthur scale

Do you own or rent your home?*

○ Rent ○ Own ○ Other (Please specify) ○ Prefer not to say

If you selected other, please specify:

What is your current employment status?*

Choose an option

What is/was your current occupation? You may select more than one.*

Select all that apply.

What is/was your mother's occupation? You may select more than one.*

Select all that apply.

What is/was your father's occupation? You may select more than one.*

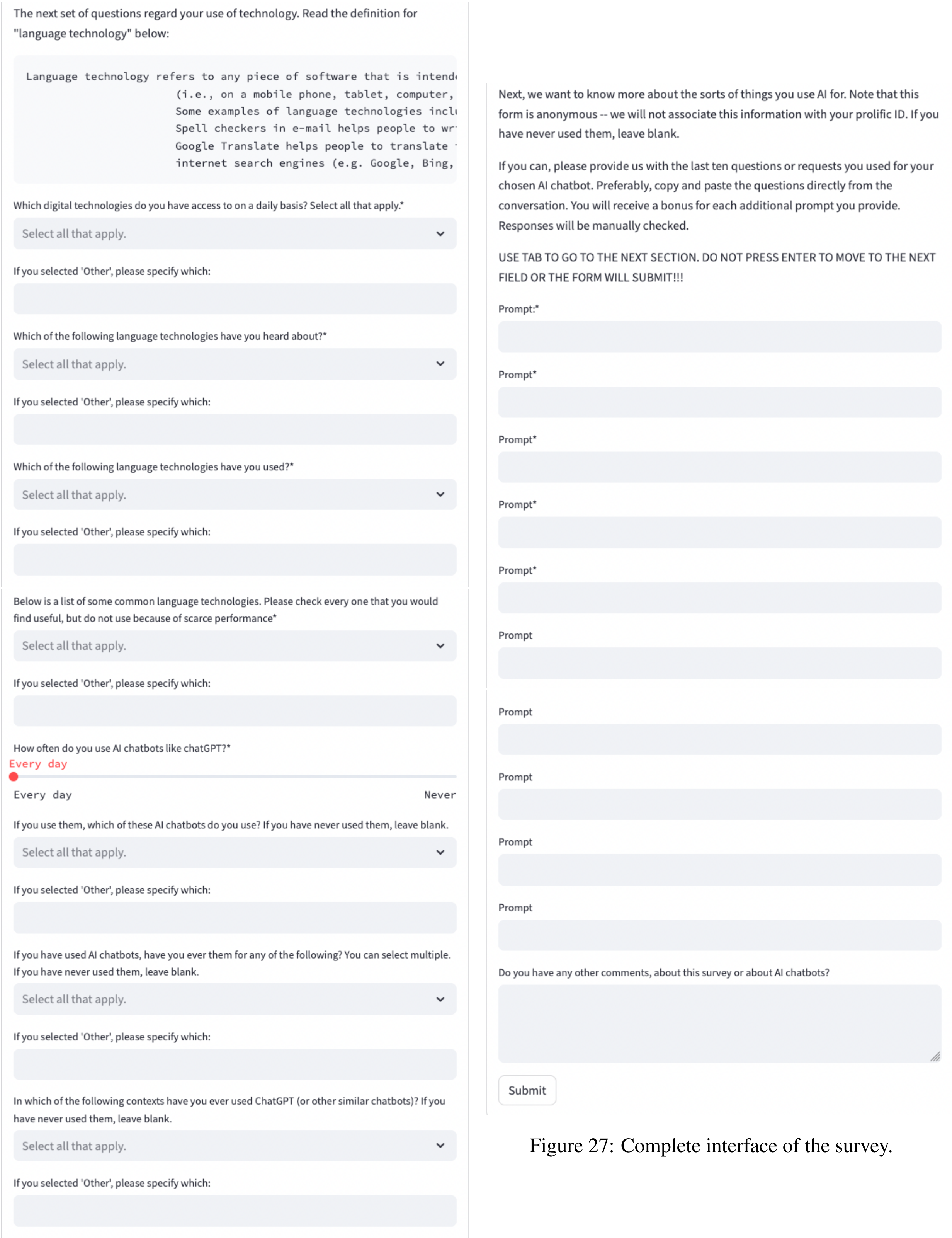
Select all that apply.

What sorts of things do you do in your free time? You may select more than one.*

Select all that apply.

If you selected other, please specify:

[continue in the next page]


The next set of questions regard your use of technology. Read the definition for "language technology" below:

```
Language technology refers to any piece of software that is intend
                         (i.e., on a mobile phone, tablet, computer,
                         Some examples of language technologies incl
                         Spell checkers in e-mail helps people to wr
                         Google Translate helps people to translate
                         internet search engines (e.g. Google, Bing,
```

Which digital technologies do you have access to on a daily basis? Select all that apply.*

Select all that apply.

If you selected 'Other', please specify which:

Which of the following language technologies have you heard about?*

Select all that apply.

If you selected 'Other', please specify which:

Which of the following language technologies have you used?*

Select all that apply.

If you selected 'Other', please specify which:

Below is a list of some common language technologies. Please check every one that you would find useful, but do not use because of scarce performance*

Select all that apply.

If you selected 'Other', please specify which:

How often do you use AI chatbots like chatGPT?*

Every day

Every day Never

If you use them, which of these AI chatbots do you use? If you have never used them, leave blank.

Select all that apply.

If you selected 'Other', please specify which:

If you have used AI chatbots, have you ever them for any of the following? You can select multiple. If you have never used them, leave blank.

Select all that apply.

If you selected 'Other', please specify which:

In which of the following contexts have you ever used ChatGPT (or other similar chatbots)? If you have never used them, leave blank.

Select all that apply.

If you selected 'Other', please specify which:

Next, we want to know more about the sorts of things you use AI for. Note that this form is anonymous -- we will not associate this information with your prolific ID. If you have never used them, leave blank.

If you can, please provide us with the last ten questions or requests you used for your chosen AI chatbot. Preferably, copy and paste the questions directly from the conversation. You will receive a bonus for each additional prompt you provide. Responses will be manually checked.

USE TAB TO GO TO THE NEXT SECTION. DO NOT PRESS ENTER TO MOVE TO THE NEXT FIELD OR THE FORM WILL SUBMIT!!!

Prompt:*

Prompt*

Prompt*

Prompt*

Prompt*

Prompt

Prompt

Prompt

Prompt

Prompt

Do you have any other comments, about this survey or about AI chatbots?

Submit

Figure 27: Complete interface of the survey.