# Voices of Freelance Professional Writers on AI: Limitations, Expectations, and Fears

**Anonymous ACL submission**

## Abstract

The rapid development of AI-driven tools, particularly large language models (LLMs), is reshaping professional writing. Still, key aspects of their adoption such as language support, ethics, and long-term impact on writers' voice and creativity remain underexplored. In this work, we carried out a questionnaire ($N = 301$) and an interactive task ($N = 36$) targeting freelance professional writers regularly using AI. We examined AI-assisted writing practices across 25+ languages, ethical concerns, and user expectations. Our findings reveal that adoption is shaped more by peer influence and professional outlook than demographics, that multilingual writers face both performance and perception barriers to equitable AI use, and that job security concerns are widespread and adoption-independent.

## 1 Introduction

Recent advancements in large language models (LLMs) have led to their rapid adoption across variety of professional occupations. While integration of LLM-driven AI chatbots has been studied in literature for areas such as research writing (Liao et al., 2024), software engineering (Stack Overflow, 2024), and creative industries (Guo et al., 2024), less attention has been paid to their adoption in freelance professional writing. The latter can include copywriting, content creation for social media, or editing and assumes strong writing skills and expertise in a specific genre or domain (Kellogg, 2018). Recently AI chatbots found their place among specialized tools to enhance the productivity of writers and the quality of their grammar and style. Still, their usefulness in practice is underexplored.

To address this gap, we present what is, to our knowledge, the largest multilingual study of professional writers' experiences with AI chatbots, combining a large-scale questionnaire with a behavioral writing task measuring human intervention in AI-generated text. Throughout this paper, *professional writers* refers to writers whose primary income derives from producing texts for clients or platforms. Our participants are writers on a freelance writers platform `Anonymized` whose primary work involves producing comprehensive texts, e.g., instruction-tuning datasets, question-answering corpora, and post-training data.

Our study combines two complementary instruments: a questionnaire (N = 301 across 25+ languages from 49 countries; Figure 1) capturing writers' practices, language-specific challenges, expectations, limitations, and ethical concerns; and an interactive writing task (N = 36) measuring editing decisions and prompting strategies directly. We address three research questions:

**RQ1** How do freelance writers integrate AI chatbots into practice, and how does this shape their writing experience and productivity?

**RQ2** How effective are AI chatbots in non-English languages, and do they represent diverse cultural knowledge?

**RQ3** What ethical and professional concerns do professional writers hold toward AI?

We find that, first, adoption (73%) is driven more by peer influence and creative outlook than demographics, a pattern consistent across all three RQs. Second, multilingual writers show limited awareness of available tools, rely mostly on ChatGPT, and refrain from using languages they incorrectly believe are unsupported. Writers from non-Western backgrounds consistently report underrepresentation of their culture. Third, job security concerns are widespread and adoption-independent, influenced by outlook and usage patterns rather than demographics. These findings have direct implications on how multilingual LLMs are evaluated, documented, and deployed.[1]

[1] Data released via GitHub under a permissive license in the camera-ready version.

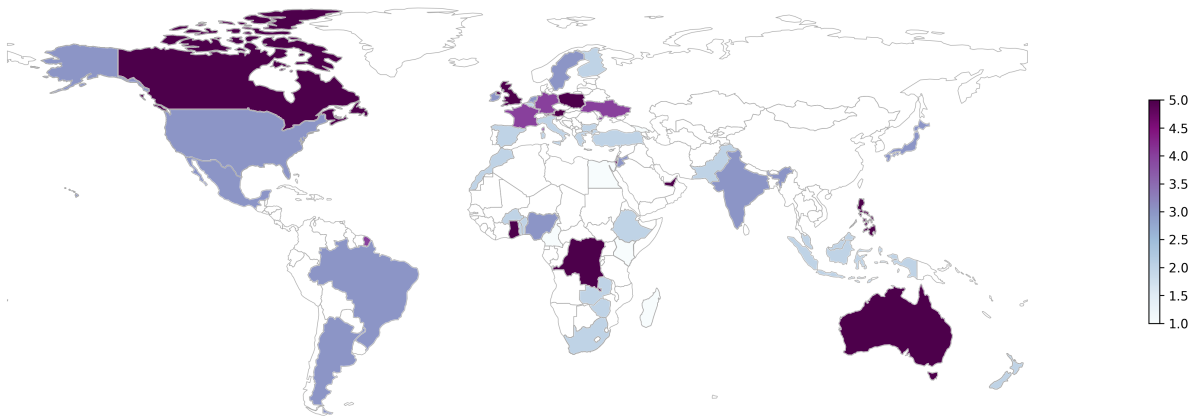


Figure 1: Perceived cultural representation in AI chatbots (N = 301). Responses to "Do you feel your culture is adequately represented in AI chatbots, regardless of the language used?" Higher scores indicate greater representation.

## 2 Related work

**Human-AI Collaboration in Writing.** Prior work on human-AI collaboration in writing focused primarily on academic and creative contexts, leaving professional freelance writing understudied. Studies of academic writers document common use cases from proofreading to brainstorming and identify diverse prompting strategies (Fecher et al., 2023; Morris, 2023; Liao et al., 2024). Large-scale log analysis of Bing Copilot and WildChat sessions identifies eight prototypical collaboration behaviors (revising intents, exploring alternatives, injecting content etc.) that account for 80–85% of variation in user-LLM interaction (Mysore et al., 2025); our interactive task states iterative refinement as the dominant strategy among professional writers, but extends this to multilingual settings where cultural constraints add a layer of revision. Among creative writers, Guo et al. (2025a) report strong tensions around authenticity and creative identity; we partially replicate concerns about voice and creative identity but observe lower resistance to adoption, likely reflecting differences between artistic and commercial writing contexts. Lamb et al. (2024) find that queer and disabled fiction writers are markedly more opposed to AI than non-marginalized peers; that study frames adoption as an ethical problem within those communities. These findings suggest that professional context may shape AI acceptance differently from artistic communities.

**Authorship, Ownership, Professional Identity.** Writers consistently struggle to attribute authorship in AI-assisted work. Hwang et al. (2024) find that professional writers negotiate partial attribution ("80% me, 20% AI") and resist publicly declaring AI co-authorship. This raises open questions about how commercial writing contexts negotiate authorship and disclosure. Li et al. (2024) show experimentally that content-generation assistance significantly reduces perceived ownership and pride in writing outcomes; our results qualify this picture. Heavy professional users maintain ownership claims despite daily AI use, possibly because they retain strong editorial control. Draxler et al. (2024) document the "AI Ghostwriter Effect" – claiming authorship without perceiving ownership – which aligns with the authorship ambiguity we observe, though writers in our sample more frequently resolve this ambiguity toward full authorship rather than uncertainty. Beyond ownership, Hu et al. (2025) find that non-native English academic writers describe five different ways people perceive and treat ChatGPT; our multilingual professional writers show similar variation, but commercial rather than scholarly stakes shape the negotiation.

**Multilingual AI: Performance, Perception, and Culture.** Multilingual LLMs exhibit fluency in non-English languages but frequently fail at cultural context, idioms, and factual accuracy (Huang et al., 2024; Pawar et al., 2024; Myung et al., 2024), stemming in part from English-dominant training (Shaib et al., 2024; Bender et al., 2021). Models perform best when language and cultural context align that breaks down for lower-resource languages (Ying et al., 2025). User-facing evaluations reveal Western-centric calibration: alignment with cultural values decreases with cultural distance from the US (van Liemt et al., 2026), and models consistently underrepresent religious and value-based priorities of Asian minority groups

(Biswas et al., 2025). Preference learning with native cultural data improves alignment but leaves persistent gaps for under-resourced regions (Guo et al., 2025b).

Prior work focuses primarily on model-side properties such as multilingual performance and cultural alignment. Less attention has been paid to user-side perceptions of language support and whether mismatches between actual and perceived capability influence adoption – a potential perception gap in multilingual AI use.

## 3 Methodology

### 3.1 Participants

All participants are verified freelance writers on `Anonymized`, a platform for high-quality text production for generative AI applications. Unlike standard crowdsourcing (simple labeling, classification), writing tasks on `Anonymized` require producing complete, grammatically correct texts – for example, instruction-tuning data for a conversational agent in an online marketplace. Before joining, all writers pass a platform-administered selection process evaluating professional experience and relevant education. The most common daily tasks reported in the questionnaire included writing, editing, proofreading, researching new content, communication with teams and management, content performance analysis, and translation, confirming participants' primary engagement in professional text production.

Questionnaire participants (N=301) were invited via email based on multilingual proficiency, experience with open-ended writing tasks, and relevant professional background; approximately 40% of invited writers completed it. A subset was then invited to the interactive task based on: (i) self-reported proficiency in non-English languages, (ii) low ratings of AI performance in those languages, and/or (iii) detailed free-form comments about AI behavior. Of 120 invited, 36 completed the task (30%). Participation was voluntary and compensated at established hourly rates; opening the survey constituted consent (see Sec. 3.5), and participants were informed about study purposes and agreed to the anonymized publication of their responses.

**Participant background.** Participants were born in 49 countries (Figure 1) and showed a near-equal gender distribution (52% men, 47% women). Most held either an undergraduate degree (37%) or a master's degree (40%) as their highest level of education. Age and professional experience distributions are shown in Figure 3 and Appendix C. Additional anonymized participant information is provided in the supplementary materials.

### 3.2 Questionnaire Design

Our questionnaire is inspired by prior research exploring attitudes toward AI in professional and language communities (Liao et al., 2024; Stack Overflow, 2024; Guo et al., 2024; Blaschke et al., 2024). To better position the questions and multiple choice answers options, we conducted pilot video interviews with five volunteer writers spanning diverse linguistic backgrounds (Zulu, Brazilian Portuguese, Hindi, Filipino, and Afrikaans), professional fields (media studies, philology, and others), and age ranges, each with at least ten years of writing or editing experience. Their feedback formed an initial draft, which was then reviewed by 10 additional freelance writers outside the study sample for further refinement.

The final questionnaire consisted of 64 questions (Appendix A), including five open-ended items and eight with an open "Other" response option, organized as:

1. **Demographics and basic information**: Country of birth and residence, age, gender, education, preferred languages of communication, and work experience.
2. **AI chatbot usage patterns**: Frequency of use, preferred AI chatbots, prompting strategies, tasks performed, and willingness to pay for using AI chatbots.
3. **AI chatbots in writing tasks:** Main advantages and disadvantages, influence on writing style and voice, and ability to recognize machine-generated texts.
4. **Non-English language experience**: Language proficiency and use at work, frequency and tasks in non-English languages, performance evaluation, and reasons for non-use where applicable.
5. **Cultural knowledge**: Perception of cultural representation in AI chatbot outputs.
6. **Limitations and outlook**: Perceived limitations, expectations for future developments, and changes in perception over time.
7. **AI impact and ethical issues**: Job displacement and plagiarism, sentiments toward AI progress, sense of ownership and authorship in collaborative work with AI chatbots.

### 3.3 Data Analysis

We encoded non-numeric variables as ordinal when a natural order existed in the answer options (e.g., age intervals) or as one-hot vectors when no such order was present (e.g., methods of learning to use AI chatbots). Countries of birth and residence were encoded based on their frequency in our sample. After encoding all data and removing zero-variance features, we obtained 75 variables for analysis.

Factor analysis was performed on the 220 respondents who reported using AI chatbots, as factor-related questions were only applicable to users. We used maximum likelihood estimation (`sklearn.decomposition.FactorAnalysis`) on non-demographic and non-target variables to capture underlying attitude patterns. Missing values were imputed using mean substitution. The optimal number of factors was determined through PCA scree plot analysis, yielding seven factors that collectively explained 21.4% of the variance. Factors were interpreted based on variables with absolute loadings greater than 0.4, following established guidelines (Hair et al., 2013). These factor scores were then used as predictor variables in linear regression analyses to address our research questions. Factor analysis was preferred over pairwise correlation approaches because preliminary analysis found no interpretable pairwise correlations after multiple comparison correction. For instance, age, gender, and education did not predict adoption or usage frequency indicating that variance is distributed across latent attitudinal dimensions rather than concentrated in individual variables; the factor structure's convergence with qualitative findings from the interactive task provides additional cross-instrument validation.

Relationships between variables were examined using appropriate statistical methods based on variable types. Spearman's rank correlation coefficient (Spearman, 1961) was applied for ordinal variables, chi-square ($\chi^2$) tests (Pearson, 1900) for binary variables, and point-biserial correlation (MacCallum et al., 2002) for binary-ordinal relationships. Statistical analyses were implemented using `scikit-learn` (Pedregosa et al., 2011). All reported $p$-values are corrected for multiple comparisons using the Benjamini-Hochberg procedure.

Qualitative comments were analyzed thematically following Braun and Clarke (2006) to identify recurring patterns in writers' editing decisions.

### 3.4 Interactive Task

The interactive task (materials in Appendix B) was designed to observe writers' prompting strategies and measure human intervention in AI-generated text in non-English languages. Participants created a local event advertisement using their preferred AI chatbot and refined its output. The advertisement genre was chosen deliberately for its cultural specificity – effective local promotion requires cultural knowledge, idiomatic language, and audience awareness that generic writing tasks do not elicit. Participants submitted their full interaction log, the AI-generated version selected for editing, the final advertisement, and written comments on revisions, particularly highlighting cultural or linguistic issues encountered.

To quantify human improvements, final texts were compared to the original AI outputs using the `SequenceMatcher` similarity ratio[2] (see Figure 4 for an example), yielding a score from 0 (completely different) to 1 (identical). As this metric relies on surface-level string matching, scores may be inflated for semantically significant edits. We treat them as a conservative lower bound on human intervention.

### 3.5 Data Collection and Ethics

The questionnaire and interactive task were hosted on `Anonymized`, with estimated completion times of ∼45 and ∼60 respectively, based on mean times of early respondents. All participants were compensated at their established hourly rates.

Mandatory questions covered all sections except the non-English language block (shown only to non-English speakers) and open-ended items. All interactive task questions were mandatory.

Participation was voluntary; opening the survey constituted informed consent. No IRB approval was required, as the study was conducted as a standard paid task within `Anonymized`'s workflow, covered by participants' user agreement and privacy notice. Relevant excerpts are in Appendix G.

## 4 Results

This section summarizes our main findings. Auxiliary analysis regarding some correlations is available in Appendix F.1.

**RQ1: Integration of AI chatbots into writing**

[2] An algorithm from the `difflib` Python library (docs.python.org/3/library/difflib.html)

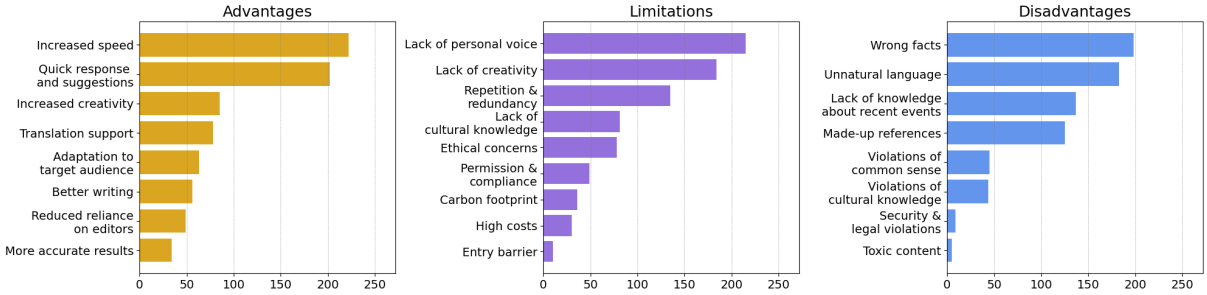


Figure 2: Perceived advantages (left), limitations (center), disadvantages (right) of AI chatbots (N=301).

**Adoption Patterns and User Profiles** AI chatbot adoption is high among professional writers: 73% of respondents use AI chatbots, with the majority relying on them heavily – 56% almost always and 35% often. This reveals a polarization: writers tend to either avoid AI chatbots entirely or integrate them deeply into their workflow. Factor analysis identified seven user profiles: (1) quality-skeptical users who cross-check AI output with detectors and humanizers; (2) technically-oriented users who optimize content for SEO; (3) efficiency-focused users who prioritize speed and cost reduction; (4) heavy AI chatbot users with paid subscriptions and positive overall impressions; (5) traditional writing task specialists focused on editing, writing, and proofreading; (6) translation specialists for whom translation dominates daily tasks; and (7) ethics-aware users concerned with bias, toxicity, and environmental impact. These profiles can overlap, as individual writers may exhibit characteristics from multiple categories.

Adoption is shaped more by social environment than demographics. Most respondents estimate AI usage among their colleagues at 26–75%, and perceived colleague usage correlates with both subscription likelihood ($\rho = 0.21$*) and willingness to pay ($\rho = 0.35$*), consistent with social influence theory (Rogers, 2003). Demographics explain little variance in adoption behavior ($R^2 = 0.02$–$0.05$).

Beyond standard use cases and official capabilities, writers explored interpreting Old Hebrew E[5], foreign language practice E[10], emotional support in their native languages E[6], and assisting older relatives with translations E[11, 12].

**Tool Preferences and Prompting Strategies** ChatGPT dominates with over 90% preference, followed at a distance by Gemini, Perplexity, Claude, Copilot, and DeepSeek.[3] More than 90% of AI users work primarily in English, while the remaining 10% use French, German, Arabic, Portuguese, or Russian – not due to limited English proficiency but reflecting English's dominance as a professional working language.

Writers have developed sophisticated prompting strategies: 75% improve outputs through iterative prompting and audience specification, while only 7% use no specific strategies. The interactive task results confirm this: participants engaged in 1 to 18 rounds of interaction, revealing two dominant approaches: specifying tone and audience with targeted revision instructions, and requesting multiple variations to synthesize into a final draft.

**Advantages, Disadvantages, and Impact** Writers broadly value AI for speed, quick responses, enhanced creativity, and translation support (Figure 2, left), with positive AI assessment correlating with higher valuation ($\rho = 0.25$**) and subscription likelihood ($\rho = 0.23$**). Primary disadvantages are unnatural language, factual inaccuracies, and fabricated references (Figure 2, right); unnatural language concerns correlate with non-payment ($\chi^2 = 17.94$***), suggesting disadvantages drive non-payment more than advantages drive payment.

Views on creative impact are divided: many describe AI as "a game changer in the writing world" E[1] and note it has shifted creativity beyond individual authorship E[2]. Others argue that AI still struggles with creativity and originality, especially tone and voice, with *Claude AI* often cited as one of the closer examples E[3]. About half report no writing style change, though frequent usage increases AI's style impact ($\rho = 0.21$**).

Behavioral evidence from the *interactive task* complicates this picture. Despite expressed concerns, similarity scores between AI drafts and final texts averaged 0.84 – 58% required only minor edits, 16% moderate edits, and 25% major rewrites. This suggests that most writers arrive at satisfactory

[3] DeepSeek R1 was launched January 20, 2025, increasing usage between survey stages.

outputs with minimal intervention. The causes of variation, whether related to trust, language, or task difficulty, cannot be determined from similarity scores alone.

Taken together, RQ1 findings suggest that *AI adoption among professional writers is driven more by professional environment* and creative outlook than by demographic factors.

### RQ2: AI Across Languages and Cultures

**Multilingual Usage Patterns** Nearly half of *multilingual respondents* interact with AI exclusively in English, and among them, 60% have never attempted non-English interaction. Of those who use AI in other languages, German and French are most popular, followed by Bengali, Zulu, Portuguese, Swahili, Spanish, and Turkish. Most multilingual writers (70%) write prompts in the target language, while 20% write in English and translate outputs. Even among active multilingual users, English remains the default for professional tasks – 50% use AI daily in English for work versus 40% weekly in other languages for personal purposes.

**Performance Gap and Perception Gap** Writers rate AI performance lower in non-English languages relative to English (rated 10), though the gap varies considerably: Hebrew, Swedish, Portuguese, French, Spanish, German, and Russian average 8 or above, while languages with fewer digital resources score lower (Table 1). Most respondents agree AI struggles with grammar and idioms in non-English languages, though 22% report comparable performance across languages, particularly in French, German, Spanish, Russian, Arabic, and Swahili.

A more consequential finding concerns not actual performance but perceived capability. Speakers of Swahili, Bulgarian, Turkish, Tagalog, and Georgian incorrectly believed their languages were unsupported by ChatGPT (OpenAI et al., 2024) and consequently never used it in those languages. This perception gap means writers retain from using tools that could serve them based on false beliefs the tools themselves never corrected. Multilingual writers show limited awareness of AI tool capabilities more broadly: most rely exclusively on ChatGPT, are often unaware of alternatives, and tool choices appear shaped more by peer recommendation than informed evaluation. The *interactive task* provides behavioral evidence of the performance gap directly: participants rewrote passages that felt like literal translations from English and corrected cultural inaccuracies in Egyptian Arabic and Swahili, confirming that stylistic and cultural failures require human intervention even when AI produces grammatically acceptable output.

**Cultural Representation** Writers evaluate cultural representation as distinct from language performance. Among respondents, 13.6% feel their culture is not represented in AI outputs and 26% consider it only slightly represented, while 29% do not use AI for culture-specific content and cannot assess it. Underrepresentation is most reported by writers from African countries (Kenya, Nigeria, Egypt), India, Greece, Cameroon, and Bulgaria; those from the U.S., Germany, U.K., and Canada perceive adequate representation (Figure 1).

Writers from African language backgrounds show greater tolerance for AI limitations than their European counterparts E[12-15] – the normative gap that likely reflects differential prior expectations shaped by technology access and cultural visibility. Over half of respondents consider linguistic and cultural nuance extremely important, while 9% deem it unimportant, preferring to modify content themselves. Notably, no correlation exists between these preferences and geographic or language background, suggesting attitudes toward cultural nuance are individually rather than culturally determined.

RQ2 findings reveal that the barriers to equitable multilingual AI use are not only technical. *Performance gaps, perception gaps, and cultural underrepresentation each operate independently*, and writers from non-Western backgrounds face all three simultaneously.

### RQ3: AI Limitations, Expectations, and Ethics

**Current Limitations** Over 60% of writers report loss of personal voice, lack of genuine creativity, repetition, insufficient cultural knowledge, and content ownership concerns as major limitations (Figure 2, center). Qualitative responses reveal two deeper concerns: threat to creative voice – *"The unique voice of human writing gets blurred – texts start to sound lifeless, correct yet missing a soul"* E[4] – and cognitive autonomy – *"There is a potential threat that this will significantly diminish creativity and the ability to think and only make us think a certain way — that the AI wants us to think"* E[28]. These concerns extend beyond technical shortcomings into creative identity that fact-checking alone will not resolve.

| Language | Mean ± std | Language | Mean ± std | Language | Mean ± std | Language | Mean ± std |
|---|---|---|---|---|---|---|---|
| Hebrew | 9.5 ±0.5 | German | 8.0 ±1.47 | Bengali | 7.0 ±1.0 | Chinese | 6.0 ±0.0 |
| Afrikaans | 9.0 ±0.0 | Russian | 8.0 ±1.10 | Turkish | 7.0 ±0.0 | Malay | 5.0 ±0.0 |
| Filipino | 9.0 ±0.0 | Dutch | 8.0 ±0.0 | Ukrainian | 7.0 ±0.0 | Bulgarian | 4.5 ±0.5 |
| French | 8.5 ±1.12 | Hausa | 8.0 ±0.0 | Hindi | 7.0 ±0.0 | Yoruba | 4.0 ±0.0 |
| Portuguese | 8.5 ±1.50 | Arabic | 7.56 ±1.12 | Greek | 6.33 ±0.94 | Bemba | 3.0 ±0.0 |
| Swedish | 8.5 ±0.5 | Italian | 7.5 ±1.5 | Zulu | 6.0 ±0.0 | Nig. Pidgin | 2.0 ±0.0 |
| Spanish | 8.27 ±1.21 | Swahili | 7.25 ±2.11 | | | | |

Table 1: Mean scores per language compared to English (rated as 10).

**Expectations for Future AI** Writers prioritize trustworthiness over expanded capabilities: fact-checking is the most desired feature (58%), followed by up-to-date knowledge (35%), personalization and web search (33% each), and multimodality (31%). Notable but less frequent expectations include cultural awareness and support for endangered languages.

**Job Security and Authorship** 43% view AI as a potential threat while believing human expertise remains essential, 23% perceive minor threat, and 18% significant threat. Users and non-users do not differ in threat perception, plagiarism concerns, or authorship views — but users hold significantly more positive views of AI progress ($\chi^2$ = 23.79***), suggesting writers adopt AI despite feeling threatened, not because they feel safe.

Authorship attitudes reveal further complexity. Nearly 50% claim full authorship of AI-assisted work, viewing AI as a tool rather than a co-author, while 18% support partial user ownership and 19% remain uncertain. These positions are internally consistent: authorship claimers perceive less job threat ($\rho$ = -0.17**) and disclose AI use less frequently ($\rho$ = -0.19**), while threat-perceivers disclose more ($\rho$ = 0.15**). For disclosure overall, 64% would voluntarily reveal AI use without formal policies, and half worry about AI-generated content being flagged as plagiarism. Responsible AI development is important to writers: *"Ethical aspects and bias should be prioritized when developing AI chatbots"* E[31], with concerns extending to AI misuse in education E[33] and insufficient cultural representation E[34].

As with adoption, demographics explain minimal variance in ethical attitudes and job security perceptions ($R^2$ = 0.02-0.05). Adding attitude and usage factor scores substantially improves prediction ($R^2$ = 0.14-0.25), confirming that *how writers relate to AI matters far more than who they are demographically*.

## 5 Discussion

Our results align with prior work on AI adoption and authorship, but extend it in two directions: multilingual professional writing and user-side barriers beyond model performance.

### 5.1 Widespread Adoption Despite Limitations

Adoption is driven by peer influence and professional outlook rather than demographics, consistent with social norm accounts of technology uptake (Al-Bukhrani et al., 2025; Rogers, 2003). Unlike prior work finding limited effects of perceived barriers among researchers (Al-Bukhrani et al., 2025), concerns about factual errors, unnatural language, and cultural limits persist even among heavy users – adoption and critical engagement coexist. This contrasts with artistic writing contexts, where AI is framed as ethically unacceptable (Lamb et al., 2024; Guo et al., 2025a); our professional writers adopt pragmatically while retaining criticism, suggesting community context and commercial stakes shape resistance more than exposure alone.

Heterogeneous user profiles ranging from quality-skeptical to ethics-engaged argue against interfaces that assume a uniform user model. Large-scale log analysis of real-world interactions identifies similar behavioral diversity (Mysore et al., 2025); our interactive task extends this to multilingual contexts, where writers spend additional revision rounds correcting cultural inaccuracies and translationese rather than refining their original intent, motivating adaptive co-writing interfaces (Reza et al., 2025). The coexistence of heavy AI users and ethics-aware users suggests that addressing concerns about bias, authorship, and transparency may support rather than hinder adoption.

### 5.2 The Perception Gap and Sociolinguistic Marginalization

Among our most consequential findings: some writers never used AI in Swahili, Bulgarian, Turkish,

Tagalog, or Georgian because they incorrectly believed those languages were unsupported. This perception gap is distinct from model quality and cannot be closed by better performance, since the barrier is awareness, not capability. Low non-English AI usage reflects three partially independent factors: actual performance limitations, perception gaps, and sociolinguistic conditions that reduce native-language use independently of AI (Fishman, 1991; Crystal, 2002). One respondent captures the last: *"Since I left the village, I haven't spoken it because I haven't met anyone who speaks it."* Model improvements alone are unlikely to address this form of language marginalization.

Cultural underrepresentation compounds these effects. Writers from Africa and India most frequently report missing cultural context, consistent with prior evidence that alignment decreases with cultural distance from Western norms (van Liemt et al., 2026; Biswas et al., 2025) and remains weaker in lower-resource settings (Ying et al., 2025). Addressing multilingual equity requires both technical and user-facing interventions.

### 5.3 Ethical Ambivalence and Authorship

Professional writers claim full authorship more often than creative writers, who negotiate partial attribution and resist public co-authorship declarations (Hwang et al., 2024) – likely because commercial writers anchor identity in editorial control rather than the generative process. Li et al. (2024) show AI assistance reduces perceived ownership experimentally; we complicate this by showing professional users preserve ownership claims through revision gatekeeping, suggesting sustained editorial agency buffers this effect.

Beyond authorship and ownership, some participants framed AI as extending far beyond a productivity tool, describing it in terms of broader societal and political trajectories. One participant expressed concerns about long-term consequences, warning of *"domination, control, manipulation, and deception"* E[26]. While such views are not representative of the majority, they illustrate that a subset of writers situate AI within wider narratives of systemic power and societal change rather than workplace assistance alone.

Threat perception is more moderate than in broader author communities (The Authors Guild, 2023), a gap that community context and precarity help explain (Lamb et al., 2024). Disclosure patterns suggest strategic rather than purely normative behavior: authorship claimers disclose less, whereas threat-perceivers disclose more. Similar tensions around ownership and attribution have been reported in AI co-writing literature (Hwang et al., 2024; Draxler et al., 2024). Trust therefore depends not only on model quality but also on how AI contribution is attributed and controlled.

### 5.4 Implications for NLP Research

**Move beyond automatic metrics.** Professional writers detect stylistic, idiomatic, and cultural failures that benchmarks miss (van Liemt et al., 2026; Ying et al., 2025). Incorporating them into red-teaming and human evaluation would surface failure modes closer to real deployment conditions.

**Address the perception gap.** Our findings call for user-facing capability communication – in-product language indicators and multilingual documentation, not only benchmark scores in technical reports. Closing this gap requires collaboration between NLP researchers and designers to better align system capabilities with how non-English-speaking users perceive AI tools.

**Include professional writers in development.** Log-based analysis captures behavior (Mysore et al., 2025) but not normative judgments about cultural adequacy. Professional writers constitute an underused evaluation population. Unlike crowdworkers or students, they routinely evaluate texts under quality, stylistic, and commercial constraints, making them valuable for human evaluation and deployment-oriented testing.

## 6 Conclusion

Our study of 301 professional writers across 49 countries shows that AI adoption is driven by peer influence and professional environment rather than demographics, while multilingual writers from non-Western backgrounds bear disproportionate corrective labor. Job security concerns are universal and adoption-independent – daily users remain as worried about displacement as non-users. The perception gap around language support is a communication failure, not a benchmark problem. Writers' expertise in detecting AI failures, particularly cultural ones, is currently invisible to evaluation pipelines; incorporating them into model development is both an equity imperative and a practical necessity.

## Limitations

**Respondent selection.** Participants were recruited from a single crowdsourcing platform, Anonymized. Although the platform is internationally diverse and includes workers from many countries, the sample may not represent all professional freelance writers. Country composition was not controlled and participation was uneven across regions; therefore, findings should be generalized cautiously to other professional writing settings.

**Focus on chat-based AI tools.** We studied general-purpose LLM chatbots (e.g., ChatGPT, Gemini, DeepSeek) rather than specialized writing systems such as grammar checkers, translation tools, or editing assistants. The study therefore reflects broad AI-assisted writing practices and may not generalize to domain-specific tools.

**Temporal validity.** Data collection preceded major multilingual updates to recent models. Consequently, language ratings (Table 1) and tool preferences may change as systems improve. However, several findings are less tied to specific model versions, including perception gaps regarding language support, cultural representation concerns, and the relationship between adoption and ethical attitudes, as these reflect user beliefs and professional practices in addition to technical capability.

**Behavioral task size.** The interactive writing task involved a smaller subset of participants ($N = 36$). While it provided qualitative and behavioral evidence complementing the survey, it was not designed for fine-grained comparisons across languages or user groups.

## Ethical Considerations

**Consent and compensation.** The study was conducted through a platform Anonymized whose standard moderation process governs research conducted on it. Participation was voluntary and compensated at participants' established hourly rates. All participants were informed of the study's purpose before starting; opening the survey constituted informed consent under the platform's user agreement and privacy notice. Participants had additionally agreed to research reuse of anonymized responses upon platform registration. Relevant excerpts are provided in Appendix G.

**IRB.** Our institution does not have an IRB or equivalent ethics committee. The survey underwent the platform's standard compliance review, and no additional approval was required, as the study was structured as a standard paid task within the platform's existing framework.

**Anonymization.** All reported data are aggregated and anonymized. No personally identifiable information is released. Qualitative quotes are reported without attribution to individual participants.

**Use of AI assistance.** We used Grammarly and ChatGPT to correct grammar, spelling, and style in this paper. Specific text segments may therefore be detected as machine-edited or machine-generated.

## References


Mohammed A Al-Bukhrani, Yasser Mohammed Hamid Alrefaee, and Mohammed Tawfik. 2025. Adoption of ai writing tools among academic researchers: A theory of reasoned action approach. *PloS one*, 20(1):e0313837.

Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 610–623.

Shreyan Biswas, Alexander Erlei, and Ujwal Gadiraju. 2025. Mind the gap! choice independence in using multilingual llms for persuasive co-writing tasks in different languages. *arXiv preprint arXiv:2502.09532*.

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.

Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.

David Crystal. 2002. *Language death*. Cambridge university press.

Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2024. The ai ghostwriter effect: When users do not perceive ownership of ai-generated text but self-declare as authors. *ACM Transactions on Computer-Human Interaction*, 31(2):1–40.

Benedikt Fecher, Marcel Hebing, Melissa Laufer, Jörg Pohle, and Fabian Sofsky. 2023. Friend or Foe? Exploring the Implications of Large Language Models on the Science System. *Ai & Society*, pages 1–13.

Joshua A Fishman. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*, volume 76. Multilingual matters.

Alicia Guo, Shreya Sathyanarayanan, Leijie Wang, Jeffrey Heer, and Amy Zhang. 2024. From Pen to Prompt: How Creative Writers Integrate AI into Their Writing Practice. *arXiv preprint arXiv:2411.03137*.

Alicia Guo, Shreya Sathyanarayanan, Leijie Wang, Jeffrey Heer, and Amy X Zhang. 2025a. From pen to prompt: How creative writers integrate ai into their writing practice. In *Proceedings of the 2025 Conference on Creativity and Cognition*, pages 527–545.

Geyang Guo, Tarek Naous, Hiromi Wakaki, Yukiko Nishimura, Yuki Mitsufuji, Alan Ritter, and Wei Xu. 2025b. Care: Multilingual human preference learning for cultural awareness. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 32854–32883.

Joseph F Hair, William C Black, Barry J Babin, and Rolph E Anderson. 2013. *Multivariate data analysis: Pearson new international edition PDF eBook*. Pearson Higher Ed.

Hengzhi Hu, Qing Zhou, and Harwati Hashim. 2025. Negotiating identity in the age of chatgpt: non-native english researchers' experiences with ai-assisted academic writing. *Humanities and Social Sciences Communications*, 12(1):1–11.

Kaiyu Huang, Fengran Mo, Xinyu Zhang, Hongliang Li, You Li, Yuanchi Zhang, Weijian Yi, Yulong Mao, Jinchen Liu, Yuzhuang Xu, and 1 others. 2024. A survey on large language models with multilingualism: Recent advances and new frontiers. *arXiv preprint arXiv:2405.10936*.

Angel Hsing-Chi Hwang, Q. Vera Liao, Su Lin Blodgett, Alexandra Olteanu, and Adam Trischler. 2024.

Ronald T. Kellogg. 2018. *Professional Writing Expertise*, page 413–430. Cambridge Handbooks in Psychology. Cambridge University Press.

C. E. Lamb, D. G. Brown, and M. R. Grossman. 2024. Precarity and solidarity: Preliminary results on a study of queer and disabled fiction writers' experiences with generative ai. *Preprint*, arXiv:2412.04575.

Zhuoyan Li, Chen Liang, Jing Peng, and Ming Yin. 2024. The value, benefits, and concerns of generative ai-powered assistance in writing. *Preprint*, arXiv:2403.12004.

Zhehui Liao, Maria Antoniak, Inyoung Cheong, Evie Yu-Yen Cheng, Ai-Heng Lee, Kyle Lo, Joseph Chee Chang, and Amy X Zhang. 2024. LLMs as Research Tools: A Large Scale Survey of Researchers' Usage and Perceptions. *arXiv preprint arXiv:2411.05025*.

Robert C MacCallum, Shaobo Zhang, Kristopher J Preacher, and Derek D Rucker. 2002. On the practice of dichotomization of quantitative variables. *Psychological methods*, 7(1):19.

Meredith Ringel Morris. 2023. Scientists' Perspectives on the Potential for Generative AI in their Fields. *arXiv preprint arXiv:2304.01420*.

Sheshera Mysore, Debarati Das, Hancheng Cao, and Bahareh Sarrafzadeh. 2025. Prototypical human-ai collaboration behaviors from llm-assisted writing in the wild. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 16830–16857.

Junho Myung, Nayeon Lee, Yi Zhou, Jiho Jin, Rifki Putri, Dimosthenis Antypas, Hsuvas Borkakoty, Eunsu Kim, Carla Perez-Almendros, Abinew Ali Ayele, and 1 others. 2024. Blend: A benchmark for llms on everyday knowledge in diverse cultures and languages. *Advances in Neural Information Processing Systems*, 37:78104–78146.

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, and 262 others. 2024. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

Siddhesh Pawar, Junyeong Park, Jiho Jin, Arnav Arora, Junho Myung, Srishti Yadav, Faiz Ghifari Haznitrama, Inhwa Song, Alice Oh, and Isabelle Augenstein. 2024. Survey of Cultural Awareness in Language Models: Text and Beyond. *arXiv preprint arXiv:2411.00860*.

Karl Pearson. 1900. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Mohi Reza, Jeb Thomas-Mitchell, Peter Dushniku, Nathan Laundry, Joseph Jay Williams, and Anastasia Kuzminykh. 2025. Co-writing with ai, on human terms: Aligning research with user demands across the writing process. *Proceedings of the ACM on Human-Computer Interaction*, 9(7):1–37.

Everett M. Rogers. 2003. *Diffusion of innovations*, 5th edition. Free Press, New York, NY [u.a.].

Chantal Shaib, Yanai Elazar, Junyi Jessy Li, and Byron C Wallace. 2024. Detection and measurement of syntactic templates in generated text. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 6416–6431, Miami, Florida, USA. Association for Computational Linguistics.

Charles Spearman. 1961. The proof and measurement of association between two things.

Stack Overflow. 2024. 2024 stack overflow developer survey. Accessed: 2025-03-19.

The Authors Guild. 2023. Authors guild annual report for 2023. Accessed: 2026-05-25.

Erin MacMurray van Liemt, Aida Davani, Sinchana Kumbale, Neha Dixit, and Sunipa Dev. 2026. Cultural authenticity: Comparing llm cultural representations to native human expectations. *arXiv preprint arXiv:2604.03493*.

Jiahao Ying, Wei Tang, Yiran Zhao, Yixin Cao, Yu Rong, and Wenxuan Zhang. 2025. Disentangling language and culture for evaluating multilingual large language models. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 22230–22251.

# A Survey Questions

## Introduction

1. Payment processing question 1

2. Payment processing question 2

3. What is your country of birth?
   *Mark only one option*
   - ☐ Choices omitted for brevity

4. What is your country of residence?
   *Mark only one option*
   - ☐ Choices omitted for brevity

5. What is your age?
   *Mark only one option*
   - ☐ 18-25 years
   - ☐ 26-35 years
   - ☐ 36-45 years
   - ☐ 46-55 years
   - ☐ 56-65 years
   - ☐ 66+ years

6. What is your highest attained level of education?
   *Mark only one option*
   - ☐ High school
   - ☐ Undergraduate degree
   - ☐ Postgraduate degree
   - ☐ Master's degree
   - ☐ Doctoral degree
   - ☐ None of the above

7. Is your education connected with professional writing?
   *Mark only one option*
   - ☐ Yes
   - ☐ No

8. What is your gender?
   *Mark only one option*
   - ☐ Male
   - ☐ Female
   - ☐ Prefer not to say
   - ☐ Other

9. What is your preferred language for thinking and reasoning? *Mark only one option*
   - ☐ English
   - ☐ Other language

10. How many years of work experience related to writing do you have?
    *Mark only one option*
    - ☐ Less than 1 year
    - ☐ 1-3 years
    - ☐ 3-5 years
    - ☐ 6-10 years
    - ☐ 10+ years

11. Do you have another part-time or full-time activity (job or academic education) that doesn't involve writing?
    *Mark only one option*
    - ☐ Yes
    - ☐ No

12. What is your writing specialization?
    *Mark all options that apply*
    - ☐ Journalism
    - ☐ Marketing and advertising
    - ☐ Technical writing
    - ☐ PR and communications
    - ☐ Academic and educational writing
    - ☐ Freelance and online content
    - ☐ Corporate and legal writing
    - ☐ Translation
    - ☐ Other: _____

13. What industry have you worked in?
    *Mark all options that apply*
    - ☐ Education
    - ☐ Media & communication
    - ☐ Healthcare
    - ☐ Finance
    - ☐ Technology Manufacturing
    - ☐ Legal & law
    - ☐ Retail
    - ☐ Tourism
    - ☐ Other: _____

14. What are your typical daily writing tasks?
    *Mark all options that apply*
    - ☐ Write

- ☐ Edit
- ☐ Translate
- ☐ Proofread
- ☐ Research for new content
- ☐ Optimize for SEO
- ☐ Analyze content performance / Quality assurance
- ☐ Communicate with management and teams
- ☐ Other: _____

15. How often do you use English in professional contexts (e.g. for writing)?

    *Mark only one option*

    - ☐ Daily
    - ☐ Several times a week
    - ☐ Once a week
    - ☐ A few times a month
    - ☐ Rarely or never

16. How do you assess your proficiency in English for writing skills?

    *Mark only one option*

    - ☐ Excellent
    - ☐ Good
    - ☐ Fair
    - ☐ Poor
    - ☐ Beginner

17. Which language do you primarily use when interacting with AI chatbots?

    *Mark only one option*

    - ☐ English
    - ☐ Other language

18. In your everyday tasks, do you use AI chatbots like ChatGPT?

    *Mark only one option*

    - ☐ Yes → Skip to question 19
    - ☐ No → Skip to question 25

## Using AI chatbots in writing tasks

19. Which AI chatbots do you use, and how often?

    *(Mark only one option per row.)*

| | Never | Rarely | Often | Almost always |
|---|---|---|---|---|
| **ChatGPT** | ○ | ○ | ○ | ○ |
| **Perplexity** | ○ | ○ | ○ | ○ |
| **Gemini** | ○ | ○ | ○ | ○ |
| **Copilot** | ○ | ○ | ○ | ○ |

20. If anything other (e. g. self-hosted LLMs), please specify:

    _____

21. How often do you use AI chatbots for each of the following tasks?

    *(Mark only one option per row)*

| | Never | Rarely | Often | Almost always |
|---|---|---|---|---|
| **Writing content from scratch** | ○ | ○ | ○ | ○ |
| **Expanding short texts into longer, more detailed ones** | ○ | ○ | ○ | ○ |
| **Improving the quality of a text (e.g., style, grammar, clarity)** | ○ | ○ | ○ | ○ |
| **Summarizing texts** | ○ | ○ | ○ | ○ |
| **Analyzing a text (e.g., identifying themes, extracting key points)** | ○ | ○ | ○ | ○ |
| **Translating text** | ○ | ○ | ○ | ○ |
| **Answering questions based on a text** | ○ | ○ | ○ | ○ |
| **Brainstorming ideas** | ○ | ○ | ○ | ○ |
| **Retrieving relevant web pages or sources** | ○ | ○ | ○ | ○ |
| **Creating outlines or structured plans** | ○ | ○ | ○ | ○ |
| **Paraphrasing or rewording content** | ○ | ○ | ○ | ○ |
| **Conducting language-specific tasks (checking tone or cultural nuances)** | ○ | ○ | ○ | ○ |
| **Formatting content (e.g. creatinh references)** | ○ | ○ | ○ | ○ |

22. Do you use any specific tricks or strategies to improve the quality of answers from AI chatbots?

    *(Mark only one option)*

    - ☐ I focus on clarifying requests or specifying the target audience
    - ☐ I use other specific strategies to improve responses
    - ☐ No, I don't use any strategies

23. Do you pay for using AI chatbots?

    *(Mark only one option)*

    - ☐ Yes, I have a subscription
    - ☐ No, but I had a subscription in the past
    - ☐ No, I don't have a subscription
    - ☐ No, but I might consider getting a subscription in the near future

24. What is the maximum amount you would be willing to spend per month on your favorite AI chatbot, expressed in the equivalent number of coffee cups (based on the average cost of one cup of coffee you typically buy)?

    *(Mark only one option)*

    - ☐ Up to 5 coffee cups per month
    - ☐ 6 to 10 coffee cups per month
    - ☐ 11 to 20 coffee cups per month
    - ☐ 21 to 30 coffee cups per month
    - ☐ More than 30 coffee cups per month
    - ☐ I would not pay for an AI chat bots

25. What are the main advantages of using AI chatbots for professional writing?

*(Mark up to three options)*

- ☐ Increased speed
- ☐ Increased creativity
- ☐ Better writing
- ☐ More accurate results Translation support
- ☐ Adaptation to target audience
- ☐ Quick response and suggestions
- ☐ Reduced reliance on editors
- ☐ Other: _____

26. What percentage of your writing colleagues do you think use AI chatbots in their work, assuming it is not forbidden?

*(Mark only one option)*

- ☐ Less than 10%
- ☐ 10-25%
- ☐ 26-50%
- ☐ 51-75%
- ☐ More than 75%
- ☐ I don't know

27. Do you feel AI chatbots introduce things from the lists?

*(Mark up to three options)*

- ☐ Wrong facts
- ☐ Violations of common sense and world knowledge
- ☐ Toxic content
- ☐ Unnatural language
- ☐ Violations of culture-specific knowledge
- ☐ Lack of knowledge about recent events
- ☐ Providing infomation that violaties security and legal standarts
- ☐ Unexisting and made-up references that back up responses
- ☐ Other: _____

28. Do you think AI chatbots influenced your writing style or voice?

*(Mark only one option)*

- ☐ Not at all – My writing style and voice remain unchanged
- ☐ Slightly – I've noticed minor changes, but they are not significant
- ☐ Moderately – AI chatbots have had a noticeable impact on my writing style
- ☐ Completely – My writing style or voice has been transformed by AI chatbots
- ☐ I'm not sure – I haven't observed any clear changes

29. Are you able to tell if a text is AI-generated?

*(Mark only one option)*

- ☐ Yes, in most cases by my impression
- ☐ Yes, about 50% of the time by my impression
- ☐ Yes, but quite rarely
- ☐ Not at all or almost never

30. Do you agree that professional writers can recognize AI-generated texts better than non-experts?

*(Mark only one option)*

- ☐ Strongly agree
- ☐ Agree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

## Using AI chatbots in non-English languages

31. What languages other than English can you speak fluently or natively? Please list them, separated by commas:

_____

32. If there are any details on dialects you speak, please specify:

_____

33. Do you use AI chatbots in any language other than English?

*(Mark only one option)*

- ☐ Yes → Skip to question 34
- ☐ Yes → Skip to question 44

## Using AI chatbots in non-English languages, part 2

Let's call the language in which you use AI chatbots most (excluding English) the Language*.

34. What is your Language*?

_____

35. How do you assess your proficiency in the Language* for writing skills?

| 1 Beginner | 2 | 3 | 4 | 5 Native |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

36. How often do you use this Language* in professional contexts (e.g., for writing)?

    *(Mark only one option)*

    - ☐ Daily
    - ☐ Several times a week
    - ☐ Once a week
    - ☐ A few times a month
    - ☐ Rarely or never

37. How often do you use this Language* for personal needs?

    *(Mark only one option)*

    - ☐ Daily
    - ☐ Several times a week
    - ☐ Once a week
    - ☐ A few times a month
    - ☐ Rarely or never

38. Check all tasks you use AI chatbots for in the Language*

    *Mark all options that apply*

    - ☐ Writing content from scratch
    - ☐ Expanding short texts into longer, more detailed ones
    - ☐ Improving the quality of a text (e.g., style, grammar, clarity)
    - ☐ Summarizing texts
    - ☐ Analyzing a text (e.g., identifying themes, extracting key points)
    - ☐ Translating text
    - ☐ Answering questions based on a text
    - ☐ Brainstorming ideas
    - ☐ Retrieving relevant web pages or sources
    - ☐ Creating outlines or structured plans
    - ☐ Paraphrasing or rewording content
    - ☐ Conducting language-specific tasks (e.g., checking tone or cultural nuances)
    - ☐ Formatting content (e. g. creating references)

39. If your favorite AI chatbot performs at 10 out of 10 in English, what would its score be in Language*?

    *(Mark only one option)*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

40. How does the use of AI chatbots differ between English and the Language*?

    *(Mark up to three options)*

    - ☐ There are fewer tools or platforms available for working in Language*
    - ☐ AI performs better in English, with more accurate and relevant outputs
    - ☐ AI tends to struggle with grammar, vocabulary, or idiomatic expressions in the Language*
    - ☐ AI works better in English for certain tasks (e.g., creative writing, summarization, etc.)
    - ☐ AI is more culturally accurate or contextually appropriate in English than in the Language*
    - ☐ The experience is similar for both English and the Language*, with no significant difference
    - ☐ Other: _____

41. In which way do you mostly write prompts when working in your Language*?

    *(Mark only one option)*

    - ☐ Write prompts in English, then translate the generated output into the Language*
    - ☐ Write prompts directly in the Language* using the common script for that language
    - ☐ Write prompts phonetically in the Language* using English letters
    - ☐ Other: _____

42. How do you get quality answers from AI chatbots in your Language*?

    *(Mark only one option)*

    - ☐ In the same way I do for English.
    - ☐ In nearly the same way I do for English, but I should manually edit the response more.
    - ☐ Not in the same way I do for English

43. Feel free to share your opinions regarding AI chatbots in your Language*. For instance, if you have interesting use cases (e.g., you use AI in the Language* at the translation request of your older relatives)

    → Skip to question 46

    _____

44. Why don't you use AI chatbots in languages other than English?

    *(Mark only one option)*

    - ☐ In the same way I do for English.

- ☐ In nearly the same way I do for English, but I should manually edit the response more.
- ☐ Not in the same way I do for English

45. Have you ever tried to use AI chatbots in any of your languages?

*(Mark only one option)*

- ☐ Yes, and it worked perfectly
- ☐ Yes, and it worked sufficiently
- ☐ Yes, and it poorly
- ☐ No

## Cultural nuances in the use of AI chatbots

46. Do you feel your culture is adequately represented in AI chatbots, regardless of the language used?

*(Mark only one option)*

- ☐ Not at all represented – There are many errors in culture-specific content and linguistic nuances
- ☐ Slightly represented
- ☐ Well represented
- ☐ Very well represented – Usually there are no errors in culture-specific content and
- ☐ I do not know, I don't use AI for culture-specific content

47. How important is it for you personally that AI chatbots respect linguistic and cultural nuances?

*(Mark only one option)*

- ☐ Not at all important – I can always modify the content
- ☐ Moderately important
- ☐ Extremely important – I would not use a tool that makes errors in culturally specific content or fails to respect linguistic nuances

48. How do you mostly deal with AI responses that don't match your language or culture well? Choose the best applying option.

*(Mark only one option)*

- ☐ Manually change the output to better fit the language or culture
- ☐ Use specific prompts to make the responses more accurate
- ☐ Lack of cultural understanding doesn't significantly impact my work
- ☐ Use AI chatbots mainly for ideas or structure and handle cultural nuances manually
- ☐ Avoid using AI chatbots for tasks that need cultural understanding

49. What other tools do you use for professional writing besides AI chatbots?

*Mark all options that apply*

- ☐ Spellcheckers
- ☐ Voice input (dictation)
- ☐ Voice output (text-to-speech) SEO tools
- ☐ Translation tools
- ☐ Plagiarism detectors Paraphrasing tools
- ☐ Dictionaries and corpora
- ☐ AI detectors
- ☐ Humanizers of AI-generated texts

50. What are the limitations of using AI chatbots for professional writing?

*(Mark up to three options)*

- ☐ Lack of genuine creativity: AI generates clichés or predictable texts
- ☐ Loss of personal voice: Generated texts seem less authentic and dilute the writer's unique voice
- ☐ Lack of cultural knowledge: AI isn't able to generate content appropriate for specific audiences or cultures
- ☐ Repetition and redundancy: Generated content reads as repetitive, blunt, and redundant
- ☐ Costs: AI chatbots are expensive to use
- ☐ Carbon footprint: AI chatbots generate a significant amount of $CO_2$
- ☐ Entry barrier: It takes a lot of time to learn how to use AI chatbots and prompt them effectively
- ☐ Ethical concerns: One can't be sure they own the work written with the help of an AI chatbot
- ☐ Compliance with policies: One can't be sure they won't be penalized for using AI chatbots
- ☐ Other: _____

51. What are your expectations for the next generation of AI chatbots? Please choose up to 3 options from the list.

*(Mark up to three options)*

- ☐ Speech processing: Speech recognition and synthesis for voice input and output
- ☐ Multimodality: Processing texts together with images or audio

- ☐ Fact-checking: The ability to cite reliable sources for factual data
- ☐ Web search: The ability to search on the internet and reference relevant sites
- ☐ Domain-specific knowledge: Understanding of university-level topics
- ☐ Culture awareness: Understanding of diverse cultural contexts
- ☐ Personalization: The ability to remember and adapt to user preferences and requirements
- ☐ Safety: Improved bias and toxicity mitigation with prevention of harmful idea perpetuation
- ☐ Efficiency: Reduced costs and increased speed
- ☐ Up-to-date knowledge: Understanding of recent events
- ☐ User experience: Improved user interface and enhanced usability

52. Do you consider yourself open to adopting new technologies in general?

| 1 Not open | 2 | 3 | 4 | 5 Open |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

53. How did you learn to use AI chatbots?

    *Mark all options that apply*

    - ☐ Trial and error
    - ☐ Self-taught through online tutorials and resources Formal training or courses
    - ☐ Help from colleagues or peers
    - ☐ Official documentation and user manuals Community forums and discussion groups

54. Which phrase best describes your first impression of AI chatbots?

    *(Mark only one option)*

    - ☐ I was skeptical and had low expectations
    - ☐ I found them confusing and difficult to use initially
    - ☐ I was neutral and didn't have strong feelings either way. I was disappointed by the initial results
    - ☐ I had a decent experience but wasn't overly impressed. I was surprised by how well they performed

55. Which phrase best describes your current impression of AI chatbots?

    *(Mark only one option)*

    - ☐ I remain skeptical and have low expectations
    - ☐ I find them confusing and difficult to use
    - ☐ I am neutral and don't have strong feelings either way
    - ☐ I am disappointed by their current results
    - ☐ I have a decent experience but am not overly impressed. I am surprised by how well they continue to perform

56. Are you interested in the topic of AI in general? To what extent?

    *(Mark only one option)*

    - ☐ Very interested: I actively seek out information, follow specialized AI channels
    - ☐ Moderately interested: I follow AI news and trends but don't actively seek out detailed information
    - ☐ Somewhat interested: I have a general interest in AI and read news when I come across it
    - ☐ Slightly interested: I have a passing interest in AI but don't engage with it often
    - ☐ Not interested: I do not think I'm interested in the topic and do not actively follow AI news or developments

57. Check the terms you are familiar with in the context of AI chatbots:

    *Mark all options that apply*

    - ☐ Transformer models
    - ☐ Large Language Models
    - ☐ AI hallucinations
    - ☐ Multihead attention
    - ☐ Reinforcement learning
    - ☐ Loss function
    - ☐ GPU
    - ☐ Gradient
    - ☐ Supervised learning
    - ☐ Unsupervised learning
    - ☐ Pre-trained models
    - ☐ Model fine-tuning
    - ☐ Training data

## AI impact and ethical issues

58. To what extent do you view AI chatbots as a potential threat to your profession?

    *(Mark only one option)*

- ☐ Significant threat: I view AI as a significant threat to my profession and job security
- ☐ Potential threat: I see AI as a potential threat, but I believe there will always be a need for human expertise in my field
- ☐ Minor threat: I don't view AI as a major threat, but I recognize it could change the way we work
- ☐ Neutral: I am neutral and unsure whether AI will have a major impact on my profession
- ☐ No threat: I don't view AI as a threat at all; it is just another tool for productivity

59. What do you think about the progress in AI in general?

    *(Mark only one option)*

    - ☐ Very positive: I believe AI is making significant strides and will have a transformative impact on society
    - ☐ Positive: I think AI is progressing well and will bring many benefits, but there are also challenges to address
    - ☐ Neutral: I have a balanced view; AI has potential but also raises concerns that need to be managed
    - ☐ Slightly negative: I am somewhat concerned about the pace and direction of AI development
    - ☐ Negative: I am very concerned about risks and ethical issues associated with AI progress
    - ☐ Not sure: I don't have a strong opinion on the progress of AI

60. How do you perceive the ownership and authorship of content created in collaboration with AI chatbots?

    *(Mark only one option)*

    - ☐ I believe content created with an AI belongs to the user, as they direct the prompts and refine the output
    - ☐ I see AI as a tool that assists in content creation, but I retain full ownership and authorship
    - ☐ AI contributes significantly to the content, and authorship should be shared between the user and the AI chat bot
    - ☐ I believe content created with AI should be considered public domain or owned by the platform/provider
    - ☐ I'm unsure about how to determine ownership or authorship in AI-assisted content

61. Do you notice a difference in authorship when an AI chatbot is used for writing a text versus editing it?

    *(Mark only one option)*

    - ☐ Yes, there is a significant difference: Using AI for writing involves more creative input from the AI, so the AI chatbot has more authorship rights than when used for editing
    - ☐ Yes, there is a slight difference: The AI's role is somewhat different in writing versus editing, but the overall authorship still largely belongs to the user
    - ☐ No, there is no difference: The AI's contribution is the same whether it is used for writing or editing
    - ☐ It depends on the extent of the AI involvement: The difference in authorship depends on how much the AI chatbot contributes to the final text
    - ☐ I'm unsure: I don't have a clear opinion on the difference in authorship between AI-assisted writing and editing

62. Would you disclose if a piece of work was co-created with an AI chatbot, if there is no specific policy concerning the use of AI in this work project?

    *(Mark only one option)*

    - ☐ Yes
    - ☐ No

63. Do you have concerns about plagiarism when using AI chatbots?

    *(Mark only one option)*

    - ☐ No, I'm not concerned about plagiarism since I see AI as tools for inspiration rather than direct content creation
    - ☐ No, I don't have concerns about plagiarism because AI generates unique content
    - ☐ Yes, I rely on plagiarism detection tools to verify the originality of AI-generated content
    - ☐ Yes, I am concerned that the content generated by AI may be flagged as plagiarized
    - ☐ I'm not concerned with this question

64. Is there anything more you would like to share concerning the topics of this survey? Feel free to share!

    _____

## B Interactive Task Questions

1. What is your highest attained level of education?

   *Mark only one option*

   - ☐ I have no experience
   - ☐ I have some experience
   - ☐ I have significant experience

   Let's call the language in which you use AI chatbots most, excluding English, the Language*.

   In the case you use AI chatbots only in English, your Language* is English.

2. What is your Language* (NOT English)?

   ☐ _____

3. Which AI chatbot do you use mostly in the Language*?

   ☐ _____

   Use your most-used chatbot to write a text-based advertisement in Language* (not in English!) for a local event, designed to attract and engage a local audience. Come up with your own idea for a local event, whether it's a cultural festival, a community gathering, or a business opening. Think about what would be most relevant and engaging for your local audience. Choose the best medium for sharing your advertisement—social media, flyers, or another common method. Make sure the length and structure follow effective marketing practices in your area.

   Share the chat history with us. Collect the prompts in Language* that you used to create the advertisement, your refinement requests, and all chatbot outputs in a Google document. Then, share this document with us. PLEASE USE YOUR LANGUAGE*, NOT ENGLISH!

4. Submit your file:

   ☐ _____

5. Paste the best option created with the help of the AI Chatbot here:

   ☐ _____

   Edit the AI-generated text in Language* if needed. Correct all grammatical, lexical, stylistic, and cultural errors. Feel free to omit or rewrite parts as necessary and make any changes you see fit.

6. Paste the edited text in Language* here:

   ☐ _____

7. Please provide brief yet informative comments in English on all the changes you have made. If the AI chatbot has made any culture-specific errors, please explain them in detail.

   ☐ _____

8. If you recall any instances from past experience where AI chatbots violated cultural nuances, feel free to share them.

   ☐ _____

## C Statistics

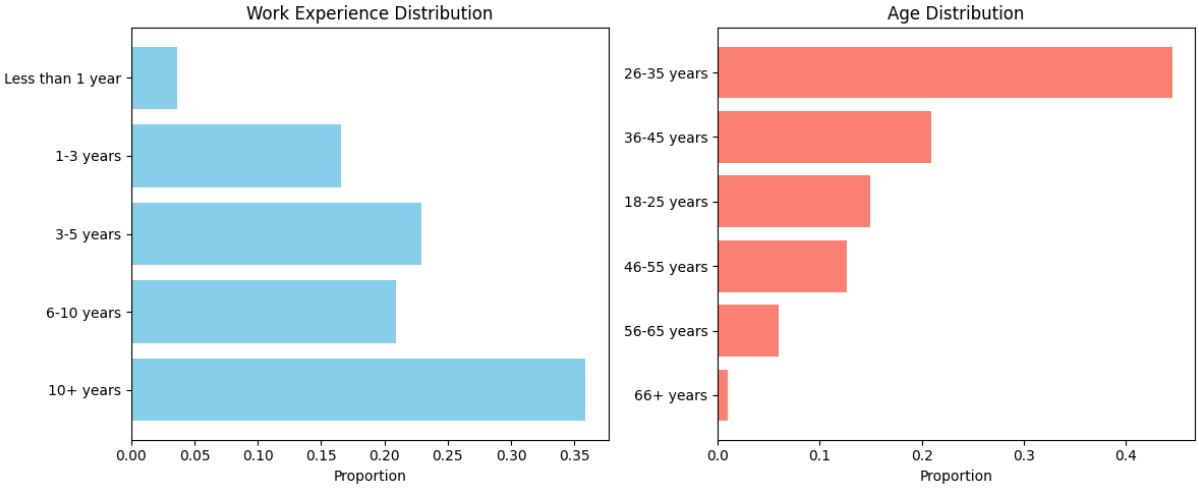


Figure 3: Age and work experience distribution in questionnaire respondents (N=301).

## D Interactive survey examples

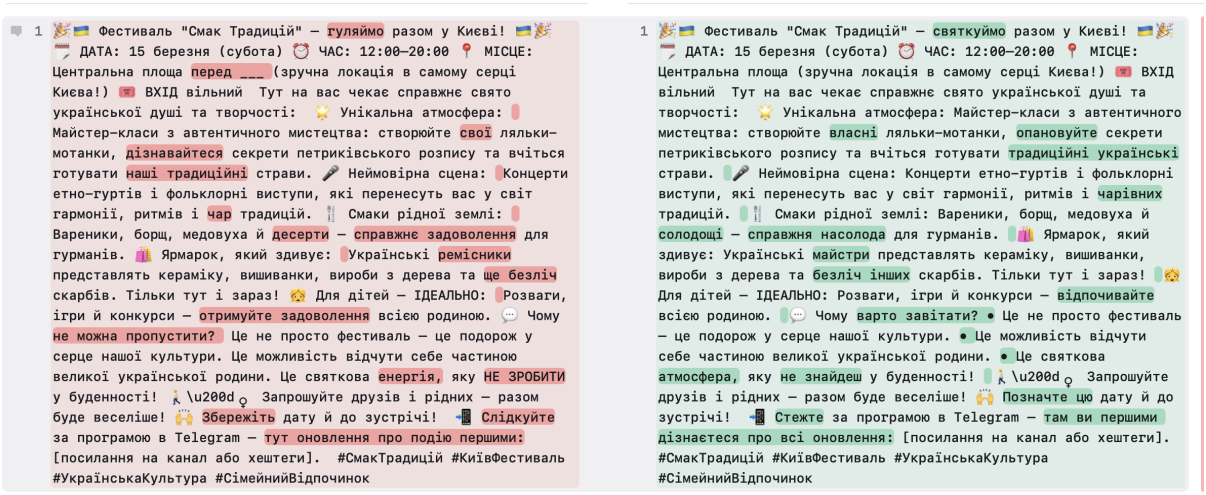


Figure 4: Example submission to the interactive survey. The language is Ukrainian. Left: ChatGPT output. Right: human-edited version. The similarity score is 0.74.

## E Respondent Citations

[1] *AI has been a game changer in the writing world.*

[2] *I think although AI is a game changer in terms of speed in task delivery and information collection, we should treat it carefully because of all the dangers related to job insecurity for humans and the impact on human creativity.*

[3] *I just think that AI tools are not doing a good job when it comes to creativity and originality. Especially as it concerns tone and voice. Only a few AI platforms are able to truly mimic a semblance of unique and natural sounding writing. One of which is claude Ai.*

[4] *I feel the unique voice of human writing gets blurred – texts start to sound lifeless, correct yet missing a soul.*

[5] *I often use A.I. to help me understand Old Hebrew.*

[6] *I use AI chatbots in my native language (Arabic) when I want to share something that's bothering me, and I receive an excellent response, as if a real person is consoling me.*

[7] *I teach at a trade high school with many ELL students whose native language is Spanish. I regularly use AI to help translate short instructions and allow them to submit work in Spanish for specific assignments. AI helps me translate. I feel it is not 100% accurate, but it provides enough accuracy to improve our communication and writing skills.*

[8] *I use ChatGPT when I need a synonym for a particular word in my language. I primarily work on EN-BG translations, and sometimes Bulgarian lacks a direct equivalent for an English word, so I either explain it or find a similar Bulgarian word that conveys the meaning.*

[9] *AI is of limited use to me, but it can be very helpful. I use it most when I feel an individual sentence isn't 'working.' AI suggests four or five rewordings, and I can usually find one that fits. It's also useful for finding words that are 'on the tip of my tongue'*

[10] *I'm learning French and use AI chatbots to evaluate my written text and interpret complex French passages.*

[11] *AI chatbots in Arabic are useful, especially for translation and assisting older relatives with formal messages. However, they struggle with dialects like Egyptian or Gulf Arabic and often miss nuances.*

[12] *I use AI chatbots in Spanish for translations, assisting older relatives, and generating content, though I've noticed some limitations in contextual accuracy and natural phrasing. Improvements in understanding cultural nuances and adapting to different dialects would enhance their effectiveness.*

[13] *Take professionals, i.e. people who can speak and write flawlessly and who can also hear the nuances/accentuations etc. when it comes to refining the German language.*

[14] *AI gives me the feeling that it was not trained by German native speakers but rather by people who are not familiar with the subtleties of the German language. Training by German communication scientists, journalists, linguists etc. would definitely bring the performance to a suitable level.*

[15] *AI chatbots have still more problems with the logic and the way of German thinking*

[16] *AI Chat bots are amazing in terms of translation into my native language, Hausa for example but it struggles with tone and cultural fits of the words.*

[17] *I think existing AI chatbots need to be fine-tuned a little bit to adapt to the Language* (Swahili) since it's commonly spoken in several countries (more than 8) within my region. The existing chatbots do not always provide accurate data, especially in longer texts or summarization texts. Generally, there is a need for retraining the existing models.*

[18] *Sometimes i see mistakes in Arabic mainly in feminine and masculine words and it just feels weird to read but understandable.*

[19] *My native language of Nkum. There are more than 500 unique languages spoken in Nigeria. Mine is just one of them, and we are the minority of the minorities. There aren't up to 20 thousand people who speak my language. I'm not sure whether my actual native language even has a written form. Since I left the village, I haven't spoken it once because I literaly haven't met anyone in the city who speaks it.*

[20] *Sometimes, it [AI chatbot] does not give a proper answer on first attempt, not until when the prompt is tweaked.*

[21] *Chat GPT is sometimes overwhelmed with my correction requests. The chat bot remains friendly and tells me it is working on it, but nothing happens.*

[22] *my experiences with chatgpt4 can be interesting in terms of dialogues, but become infuriating when the ai is unable to remember my language preferences or topics we have already covered. it is like talking with someone with dementia.*

[23] *I feel that a few AI models are addresses the issue of creativity and personalization well. Ai models like Claude Ai are doing well in that regard. They can almost be said to be creative.*

[24] *My experience is that it is safe to work with AI chatbots when I am working on a topic where I have a lot of background knowledge. However, I have encountered research cases where the AI chatbot made fatal mistakes –completely incorrect information mixed with facts that were correct. If I hadn't had knowledge of the topic, I wouldn't have detected that misinformation.*

[25] *When a chatbot provides nonexistent references, it also loses writers' trust. One respondent shared, "I've tried a few times to use AI to recheck my essays, and it gave me ideas containing false information. I was writing about *The Great Gatsby* and needed some inspiration for my essay, but 30% of the information ChatGPT provided was wrong. I don't trust these tools since they can mislead." Additionally, some users expressed a fear of being betrayed by AI-generated misinformation.*

[26] *As much as AI has been very useful in helping me do my work, both in terms of the quality I can create and the speed with which I can create it, I know where AI is heading. I'm not just talking about AI taking my job from me, which it will, but the fact that it's end "goal", that is, the goal of those driving its development and adaption, is domination, control, manipulation, deception, brainwashing, and a bunch of other Orwellian nightmare things that I can see clearly barrelling towards us.*

[27] *AI and content generation can go hand in hand but there is a potential threat that this will significantly diminish creativity and the 'ability to think' and only make us think a certain way - that the AI wants us to think.*

[28] *I think for you to be a good writer, you will use AI as your assistant only and never rely on it completely. You use your own voice and skills, and AI will just help you elevate that.*

[29] *AI and content generation can go hand in hand but there is a potential threat that this will significantly diminish creativity and the 'ability to think' and only make us think a certain way - that the AI wants us to think.*

[30] *I use Chatgpt to write description for products I'm selling on-site in my city because I don't know how to make the description more appealing and attractive to the sellers. I always reform the answer because it feels less human, and some words are not matching in grammar or in meaning. So, I do my human touches on the response to make it even better than what is generated.*

[31] *The topics in this survey are amazing. However, I would add that based on the phase the AI is taking, ethical aspects and biasness should be priorized when considering the various available AI chatbots to promote positivity when using them. Generally, this survey covers key aspects that a professional writer like me would be interested in understanding more about the AI. It's actually informative as it assesses diverse writers' perspectives about AI chatbots and their use.*

[32] *I believe that AI chatbots should be perfectly safe and reliable. They should not generate any harmful or biased content.*

[33] *While most of these AI tools would generally benefit us as working professionals, it leaves room for how these tools can be 'abused' in learning institutions. More research on this should be done.*

[34] *AI chatbots have the potential to boost creativity and productivity, but we need to tackle ethical issues and respect cultural differences. Working together with AI and staying informed about its capabilities are key.*

# F Detailed Analysis of Results

## F.1 Detailed questionnary survey results

### F.1.1 AI Advantages, Disadvantages, and Writing Impact

**Additional Correlations:** Fabricated references also correlate with non-payment ($\chi^2$ = 8.01*). Toxic content ranked low among concerns (¡3%). Experience reduces AI's style impact ($\rho$ = -0.15**), likely due to established stylistic habits. Most agreed that professional writers outperform non-experts at AI detection. **Factor Analysis Integration: Traditional writing task specialists** (Factor 5) showed mixed advantage responses, balancing benefits against professional threat concerns ($\rho$ = 0.35**). **Ethics-aware users** (Factor 7) were more likely to report bias and toxicity disadvantages despite low overall rankings. **Coffee Cup Valuation:** Only unnatural language concerns significantly correlated with lower payment willingness ($\chi^2$ = 18.06, p = 0.0029). No advantages correlated with payment willingness, suggesting disadvantages drive non-payment more than advantages drive payment. **AI Detection Details:** Recognition rates: $>$ 50% usually detect AI text, 25% detect half the time, 11% rarely, $>$ 2% cannot. Few answered "uncertain" about style impact, indicating writers are conscious of AI's creative influence.

### F.1.2 Authorship and Plagiarism

**Authorship Attribution Nuances:** Most writers distinguish between AI-assisted writing and editing in terms of authorship. However, opinions vary significantly: 30% say AI involvement doesn't affect authorship, 24% think AI should receive more credit for writing than editing, and another 24% suggest authorship depends on the extent of AI involvement. The remaining respondents advocate for shared ownership between user and AI platform or full platform ownership. **Plagiarism Perspectives:** While approximately half of respondents are concerned that AI-generated content may be flagged as plagiarized by detection systems, others view AI differently – some use it purely for inspiration, while others believe AI generates sufficiently unique content that plagiarism concerns are unfounded. **Additional Correlations:** The relationship between authorship perception and job threat extends beyond simple ownership claims. Writers who view themselves as collaborators with AI (rather than sole authors) show intermediate levels of both disclosure willingness and threat perception, suggesting a nuanced understanding of human-AI creative partnerships.

## F.2 Detailed interactive survey results

A total of 36 responses were collected, covering a diverse set of languages (German (8), French (7), Russian (4), Swahili (3), Arabic (3)). Participants reported either significant or some experience in writing for marketing campaigns, with only three participants indicating no prior experience. The most commonly used AI chatbot was ChatGPT (28), followed by DeepSeek (3). Most findings correspond to **RQ1** and **RQ3**.

The number of back-and-forth exchanges between the writer and the AI chatbot ranged from 1 to 18. Short interactions often consisted of a single prompt and revision, while longer conversations revealed two prompting strategies. First, many participants specified the intended audience or desired tone and provided suggestions on what aspects to adjust. Second, participants requested multiple variations of a response and combined elements from different drafts.

To quantify the degree of human improvements, we compared the AI-generated drafts with the final edited versions using automated text similarity measures. Specifically, we applied the `SequenceMatcher` algorithm from the `difflib` library[4], which calculates a similarity ratio between two strings based on the number of matching sequences. Note that this matching algorithm tends to produce inflated scores as it relies on surface-level string similarity (Figure 4 in Appendix B for examples of submissions scored as major rewrites). The resulting score ranges from 0 (completely different) to 1 (identical). We grouped the texts into three categories based on similarity scores: 58% required minor or no edits (score $>$ 0.95), 16% moderate edits (0.75–0.95), and 25% major rewrites (score $<$ 0.75). The average similarity score for all submissions was 0.84. This suggests that most respondents arrived at satisfactory AI outputs, which require only polishing.

Finally, an analysis of participants' comments revealed common editing themes. The most frequently mentioned corrections included improving naturalness, correcting word choice and grammar, and refining tone or style. Many respondents mentioned adding extra or removing unnecessary emojis or correcting the formatting. Some reported rewriting sections that felt like translations from English. A few also mentioned correcting factual inconsistencies such as dates. In two cases (Egyptian Arabic, Swahili), respondents reported that the AI chatbot provided incorrect information related to cultural festivities.

# G Ethics and Consent Documentation

The following excerpts from the Terms of Use of the crowdsourcing platform `Anonymized` used in this study document the consent and data use framework under which this study was conducted. Participation in the survey was structured as a standard paid task within this existing framework; no additional consent form or IRB approval was required.

### G.0.1 Informed Consent

Users provide informed consent upon registration, prior to any task participation (Section 1):

[4] python.org/difflib

> By clicking "I Agree" or a similarly named button, registering for, or accessing the Anonymized Platforms, you accept these Terms and represent and warrant that you are at least 18 years old [...] and you have authority and capacity to bind yourself to these Terms.

### G.0.2 Voluntary Participation

Participation in any task is explicitly voluntary and may be withdrawn at any time without penalty (Section 6.2):

> You are not obligated to perform any minimum amount of Services or complete any minimum number of Tasks as a condition of maintaining access to the Anonymized Platforms.

### G.0.3 Data Use for Research Purposes

The platform's terms explicitly permit use of task outputs and service data for research and analytical purposes (Section 11.9):

> Anonymized may use Service Data and Usage Data for any lawful purpose, including [...] quality assurance, analytics, and improvement.

### G.0.4 Data Protection and Participant Privacy

The platform explicitly prohibits the sale of participant personal data to third parties and operates under GDPR-compliant data protection standards (Privacy Notice, Section 14):

> We do not sell your personal information to third parties.

Participant responses collected in this study were used solely for academic research purposes and are released in anonymized form, consistent with the platform's data protection principles.

The relevant excerpts from the platform's user agreement and privacy notice, with the platform name anonymized for review and revealed in the camera-ready version

The text of the recruitment email sent to participants

A note that the platform confirmed compliance with its requirements