How People Use ChatGPT: Conversation-Level Evidence from India, Nigeria, Brazil and Pakistan

SHREYASI ROY CHOWDHURY, Independent researcher

KIRAN GARIMELLA, Rutgers University

Recent reports from OpenAI and Anthropic have begun to characterize how hundreds of millions of people use conversational AI, but these analyses rely on aggregated, privacy-preserving indicators with fixed taxonomies and inferred demographics that cannot be re-analyzed by outside researchers. We provide a complementary, conversation-level view: complete ChatGPT exports comprising 202,590 conversations from 1,252 users across India, Nigeria, Brazil, and Pakistan, paired with self-reported age and gender and spanning over three years of use (December 2022–February 2026). Using the same 24-category taxonomy as Chatterji et al. [9], we find that our sample over-indexes on information-seeking and writing and under-indexes on technical help relative to their reported global averages, with substantial heterogeneity across the four countries. Unsupervised topic discovery reveals use cases that a fixed taxonomy cannot see: health and wellness emerges as the single most prevalent theme in two of our four countries, and finance and online-earning strategies account for 10–17% of conversations, alongside translation needs, religious queries, and emotional self-reflection that vary sharply by country. Classifying conversations by task purpose (work, coursework, or personal) shows that the majority of usage is personal (55–64%), with health advice, translation, and everyday problem-solving far outweighing workplace applications. Usage patterns vary markedly by gender and age: for instance, women in India, Pakistan, and Nigeria use ChatGPT for coursework at substantially higher rates than men (up to 33% vs. 20%), and younger users concentrate on education and programming while older users shift toward finance, career development, and civic topics. Over time, users have shifted from primarily seeking information to increasingly delegating tasks to the model, and reflective, companion-like conversations are growing slowly but consistently across all four countries. These findings suggest that evaluating AI's impact through workplace productivity alone misses the dominant mode of use in these markets: ChatGPT functions less as a professional tool than as everyday infrastructure for health information, language access, education, and economic navigation—services whose value is real but largely invisible to standard productivity metrics.

ACM Reference Format:

Shreyasi Roy Chowdhury and Kiran Garimella. 2018. How People Use ChatGPT: Conversation-Level Evidence from India, Nigeria, Brazil and Pakistan. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX).* ACM, New York, NY, USA, 43 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

ChatGPT is now one of the most rapidly adopted technologies in history, with over 700 million monthly active users as of early 2025 [9]. As conversational AI becomes embedded in everyday life at this scale, a fundamental empirical question has moved to the centre of AI research, economics, and public policy: what are all these people actually doing with it?

Authors' Contact Information: Shreyasi Roy Chowdhury, shreyasi.rc12@gmail.com, Independent researcher; Kiran Garimella, kiran.garimella@rutgers.edu, Rutgers University.

Manuscript submitted to ACM

The answer so far comes from a small number of sources, each with significant limitations. OpenAI's own analysis of roughly 700 million weekly ChatGPT conversations [9] and its India-specific Signals report [21] are drawn from server-side logs at enormous scale; Anthropic's Economic Index and its geographic follow-ups [3, 4] do the same for Claude. These platform reports are invaluable for establishing broad trends, but they are, by design, aggregated and privacy-preserving: they classify conversations into a fixed taxonomy chosen in advance, infer demographics from user names or IP addresses, and release only summary statistics. An outside researcher cannot re-run a different classifier, look at the raw conversations, extract tool-use metadata, or test a hypothesis the platform operators did not report on. The WildChat corpus [30] is conversation-level and publicly available, but its users self-selected into a free research proxy, which strongly skews the sample toward technically sophisticated users; our own cross-dataset comparison (Appendix B) shows that WildChat overestimates programming by 4–7× relative to a broader sample. Recent data-donation studies [14, 19] have begun to collect naturalistic ChatGPT conversation histories at the individual level, demonstrating the feasibility of this approach, but focus on European and U.S. populations and do not provide cross-country comparisons or demographically grounded samples from the countries where ChatGPT's user base is growing fastest.

What this paper contributes. We present a conversation-level analysis of ChatGPT usage across four countries (India, Nigeria, Brazil, and Pakistan) based on complete conversation exports from 1,252 recruited users, paired with self-reported age and gender and spanning over three years of use (December 2022 through February 2026). To our knowledge, this is the first study to combine full, multi-turn ChatGPT conversation logs with verified demographic metadata at this scale for any of these countries, and the first to enable direct cross-country comparison on a common analytical pipeline.

We deliberately apply multiple classification methods to the same corpus, because each reveals something different. First, we replicate the 24-category taxonomy of Chatterji et al. [9] at the conversation level, which makes our results directly comparable to OpenAI's global averages and lets us measure how these four countries deviate from worldwide patterns. Second, we run unsupervised topic discovery (BERTopic [16]), which helps us find fine-grained topics in the data rather than projecting onto categories designed elsewhere. This turns out to matter: the unsupervised pipeline surfaces culturally specific use cases, such as health and wellness, religious queries, translation needs, digital-entrepreneurship strategies, that the predefined taxonomy absorbs into generic buckets like "Seeking Information." Third, we replicate two classifiers from Anthropic's Economic Index [3]: one that distinguishes whether users are seeking information, delegating tasks, or engaging in reflective and emotional conversation; and one that labels each conversation as work-related, coursework-related, or personal. Together, these three lenses answer not only *what* users talk about, but *how* they engage with the model and *why* (for work, school, or personal life). Finally, we extract raw metadata directly from the conversation exports, such as model versions, tool invocations, prompt and response lengths, conversation depth, subscription tier, none of which is available in any published platform report.

Main findings. We organize the results around five analytical lenses, each applied to the same corpus:

(1) **Topic distribution (Section 5.1).** Under the OpenAI taxonomy, users in our four-country sample over-index on *Seeking Information* (+4.0pp) and *Writing* (+2.5pp) and under-index on *Technical Help* (−5.5pp) relative to global averages [9], with substantial per-country heterogeneity. Gender-conditional patterns (men toward technical help and information seeking, women toward writing, practical guidance, and self-expression) mirror, at the individual-conversation level, the name-inferred patterns reported by OpenAI for India [21].

(2) **Unsupervised topic discovery (Section 5.2).** The unsupervised pipeline surfaces use cases that the predefined taxonomy cannot see: *health and wellness* as the single most prevalent theme in India (8.0%) and Brazil (9.2%); *religious queries* in Nigeria and Pakistan; *Urdu–English translation* as the top cluster in Pakistan; *self-reflection and emotional support* as a top-5 cluster in Brazil; and *online earning* strategies in India and Pakistan. These culturally specific clusters can only be found with conversation-level access to the data and an open-vocabulary method.

(3) **Task purpose (Section 5.3).** Personal use dominates in every country (55–64% of conversations), and coursework is comparable in prevalence to work (19.6% vs. 20.8% pooled). Writing and technical help make up 45–50% of work conversations, but the majority of all usage is non-work: health, education, translation, everyday problem-solving. Female users engage with coursework at substantially higher rates than male users in India (26% vs. 18%), Pakistan (33% vs. 20%), and Nigeria (24% vs. 20%).

(4) **Intent evolution (Section 5.4).** Users have shifted from primarily seeking information to increasingly delegating tasks to the model: task-delegation conversations rose from below 20% in early 2023 to 30–35% by late 2025. Meanwhile, reflective, emotional, and companion-like conversations, though still a small share, drift slowly but consistently upward across all four countries, a trend worth tracking as it carries implications for parasocial attachment and the role of AI in emotional support.

(5) **Raw metadata (Section 5.5).** Model adoption follows platform defaults rather than active user selection; advanced features (web search, code interpreter, image generation) are used in fewer than 16% of conversations; and ChatGPT Plus subscribers are concentrated almost entirely in India, with near-zero adoption in Nigeria and Pakistan, where the $20/month subscription represents 2–5% of median income.

We additionally validate these findings against the WildChat corpus [30] in Appendix B, showing that proxy-based collection systematically over-represents programming (by 4–7×) and under-represents the information-seeking and practical-guidance queries that dominate in our broader sample.

Taken together, these results paint a picture of ChatGPT usage that challenges the dominant framing of conversational AI as primarily a workplace productivity tool. The majority of usage is personal, and the most prevalent topics are health, education, translation, and everyday practical guidance. At the same time, usage is far from uniform: what people do with ChatGPT varies substantially by country, gender, and age, in ways that reflect local labour markets, linguistic needs, educational systems, and cultural context. Women's disproportionate use of ChatGPT for coursework in India, Pakistan and Nigeria, the concentration of religious queries in Pakistan, and the prevalence of translation in multilingual societies are signals that the value and risks of conversational AI will be distributed unevenly across populations. Understanding these patterns, at the conversation level and with demographic grounding, is a prerequisite for informed policy. This study is a first step. Data donation, in which users export and share their conversation histories with informed consent, is a scalable and replicable methodology for building the empirical base that policymakers, researchers, and platform designers need. We develop these implications in Section 6.

Paper structure. Section 3 describes the dataset and the WildChat corpus used for cross-validation. Section 4 details the classification pipelines. Section 5 presents results organized around the five analytical lenses. Section 6 discusses implications for economics, policy, and product design. All code, classifier prompts, and aggregate tables are released for replication.

2 Related Work

Our work sits at the intersection of four research threads: large-scale studies of LLM usage, AI adoption in the Global South, gender and age disparities in technology adoption, and methodological work on classifying user–AI interactions.

2.1 Large-Scale Studies of LLM Usage

The earliest efforts to characterize real-world LLM usage relied on research-proxy corpora. WildChat [30] released approximately one million ChatGPT conversations (expanded to 4.8M) collected through a free Gradio frontend; LMSYS-Chat-1M [31] did the same through a model-comparison platform. Both have the advantage of being conversation-level and open, but their collection methods skew the sample toward technically sophisticated users comfortable with non-standard interfaces.

The second wave of evidence has come from the platforms themselves. Chatterji et al. [9] analyzed approximately 700 million weekly ChatGPT conversations over a 15-month period, classifying each into a 24-category taxonomy over seven coarse domains (Practical Guidance, Seeking Information, Writing, Technical Help, Self-Expression, Multimedia, Other) and a three-way *Asking*/*Doing*/*Expressing* intent classification. They reported that only 27% of conversations are work-related and that self-expression is the fastest-growing category. OpenAI's Signals initiative [21] has since released a country-specific companion report for India, providing aggregated age, gender (inferred from names), and topic breakdowns. Anthropic's Economic Index [3] and its geographic follow-up [4] perform analogous analyses for Claude.ai, introducing additional measures: task purpose, multitasking, educational requirements, and the *Asking/Doing/Expressing* frame.

A third, emerging line of work collects conversation-level data directly from users through data-export or data-donation mechanisms. Karnam et al. [19] gathered 825K ChatGPT interactions from 300 users via GDPR data-export rights and documented a shift from functional to socially framed usage, including substantial growth in health and mental-health queries and rising anthropomorphization of the model. Fang et al. [14] collected 48,495 conversations from 82 U.S.-based adults through a privacy-preserving "wrapped"-style pipeline, finding a mix of instrumental and reflective use and noting that heavier users engage in proportionally more reflective exchanges. Both studies demonstrate the feasibility and value of user-donated conversation data, though neither covers the Global South countries or the cross-country comparative framing that motivate our work.

These platform reports are invaluable for benchmarking, but three limitations motivate complementary work. First, they are aggregated, privacy-preserving summaries; outside researchers cannot re-run a classifier, extract tool-use metadata, or test a hypothesis the companies did not report on. Second, demographic information is inferred from names or IP addresses rather than self-reported. Third, the topic taxonomy is fixed in advance and cannot be re-opened to discover culturally specific usage patterns. Our study directly addresses these limitations by working with complete conversation exports and self-reported demographics from a recruited cohort across four Global South countries, while using the *same* classifiers the platform reports use so that the two bodies of evidence are directly comparable.

2.2 AI Adoption in the Global South

A large literature documents persistent digital divides in the Global South along dimensions of infrastructure, literacy, language, and economic access [1, 26, 27]. Mobile-first access patterns [12], linguistic under-representation [18], and data-quality challenges [24] have all been shown to shape how users in these markets engage with digital technologies. For AI specifically, Arora and Rathi [6] examined how AI tools are perceived in Indian workplaces, and Chinchure et al. [10] analyzed early demographic patterns in AI chatbot usage. The platform reports [4, 21] are now beginning to characterize AI usage in specific Global South markets, but they do so only at the aggregate level for individual countries. Ours is, to our knowledge, the first conversation-level cross-country comparison on a common classification pipeline for India, Nigeria, Brazil, and Pakistan.

2.3 Gender, Age, and Technology Adoption

Gender disparities in technology access and usage have been extensively documented [2, 5, 17]. In the Global South, gender gaps in internet and smartphone access remain substantial, shaped by cultural norms, economic constraints, and educational disparities [15]. For LLMs specifically, Chatterji et al. [9] estimated, using name-based inference, that women's share of ChatGPT users has grown from 35% to 48% globally, and OpenAI [21] provides a similar decomposition for India. Because name-based inference is known to misclassify gender-ambiguous names (an especially relevant concern for Indian names), our self-reported gender labels provide a useful validation check. We emphasize, however, that we interpret gender-conditional *usage patterns* rather than gender *composition* as findings, since the composition of our recruited cohort reflects the Clickworker sampling frame.

2.4 AI in Education and Work

AI chatbots have attracted significant attention in educational settings [7, 11, 20, 29], including as a partial substitute for tutoring or reference material where these are scarce [28]. On the work side, Anthropic [3] report that coding, writing, and analysis dominate professional Claude usage, and Eloundou et al. [13] estimated that 80% of the U.S. workforce could see at least 10% of their tasks affected by LLMs. The geographic Claude analysis [4] finds that software and engineering tasks account for the largest share of work-related Claude usage in India, reflecting India's IT services workforce, and OpenAI [21] finds that work-related ChatGPT use in India is concentrated in writing and technical help. Our cross-country view on a common pipeline complements these India-focused aggregates with conversation-level evidence from Brazil, Nigeria, and Pakistan.

2.5 Classifying User–AI Interactions

Methods for classifying the topics and intents of user–AI interactions range from rule-based taxonomies [23] to unsupervised clustering [31] to LLM-based classification [9]. BERTopic [16] has emerged as a widely adopted method for discovering topic structure in conversational data, combining transformer embeddings with hierarchical clustering. We apply three classifiers in parallel (the OpenAI 24-category taxonomy, BERTopic over Gemini embeddings, and Anthropic's task-purpose and *Asking/Doing/Expressing* classifiers) so that our findings are both directly comparable to the platform reports and capable of surfacing patterns that a fixed taxonomy cannot see.

3 Dataset

3.1 Data Collection

We recruited participants from four Global South countries (India, Nigeria, Brazil, and Pakistan) through Clickworker, a global crowdsourcing platform. Participants were required to be active ChatGPT users and were asked to export their complete ChatGPT conversation history using OpenAI's built-in data export feature, which generates a comprehensive archive including all conversations, model metadata, and account information. Participants also provided basic demographic information (age and gender) during registration. We created a simple interface for the participants to export and upload their ChatGPT histories and before the data is exported, our tool runs a light weight client side script to remove any personal information associated with the users (like their names, emails etc). Participants were compensated for their participation. The data collection happened between December 2025–Feb 2026. Our approach parallels concurrent work by Karnam et al. [19], who collected ChatGPT exports via GDPR data-export rights, and by Fang et al. [14], who designed a privacy-preserving "wrapped"-style pipeline for the same purpose; the key difference is our focus on recruited, demographically profiled users across four Global South countries rather than convenience samples in Europe or the United States.

Table 1 summarizes the key characteristics of our dataset. In total, we collected conversation exports from 1,252 users encompassing 202,590 unique conversations, spanning from December 2022 (shortly after ChatGPT's launch on November 30, 2022) through February 2026, a period of over three years covering the introduction of GPT-3.5, GPT-4o, and the GPT-5 family.

Table 1. Dataset overview by country. Convos/User reports conversations per user (mean with interquartile range).

Country	Users	Conversations	Mean	Median	P25–P75	Date Range
India	557	88,958	159.7	51	11–197	Dec 2022 – Feb 2026
Nigeria	243	44,114	181.5	91	20–225	Dec 2022 – Dec 2025
Brazil	246	40,067	162.9	62	26–159	Dec 2022 – Dec 2025
Pakistan	206	29,451	143.0	55	13–153	Dec 2022 – Dec 2025
Total	**1,252**	**202,590**	**161.8**	**60**	**15–193**	**Dec 2022 – Feb 2026**

India contributes the largest share of users (44.5%) and conversations (43.9%), followed by Nigeria (19.4% of users), Brazil (19.6%), and Pakistan (16.5%). The high variance in conversations per user (IQR of 15–193 overall) reflects a mix of casual and power users across all countries, with Nigeria having the highest median engagement (91 conversations per user). Monthly active users and conversation volume by country are shown in Figure 34.

Table 2 presents the demographic composition of our sample. Overall, 65.9% of users are male and 34.1% are female. However, this gender distribution varies substantially across countries: India (77.2% male) and Pakistan (75.7% male) show the most skewed ratios, while Brazil is the only country where female users outnumber males (55.3% female). Nigeria shows near-parity with a slight male majority (53.1%).

Table 2. Demographic composition by country, gender, and age group (user counts).

	Gender		Age Group			
Country	Male	Female	18–25	26–35	36+	Mean Age
India	430 (77.2%)	127 (22.8%)	287 (51.5%)	187 (33.6%)	83 (14.9%)	27.8
Nigeria	129 (53.1%)	114 (46.9%)	56 (23.0%)	133 (54.7%)	54 (22.2%)	31.0
Brazil	110 (44.7%)	136 (55.3%)	45 (18.3%)	110 (44.7%)	91 (37.0%)	33.3
Pakistan	156 (75.7%)	50 (24.3%)	118 (57.3%)	55 (26.7%)	33 (16.0%)	27.1
Total	**825 (65.9%)**	**427 (34.1%)**	**506 (40.4%)**	**485 (38.7%)**	**261 (20.8%)**	**29.6**

Age distributions also differ markedly: India and Pakistan skew young (median age 25 and 24, respectively, with over 50% in the 18–25 bracket), while Brazil is notably older (median 32, with 37% over 36). Nigeria falls in between (median 30). These demographic differences likely reflect both the age profiles of Clickworker participants in each country and broader patterns of technology adoption. Figure 1 visualizes the demographic composition.

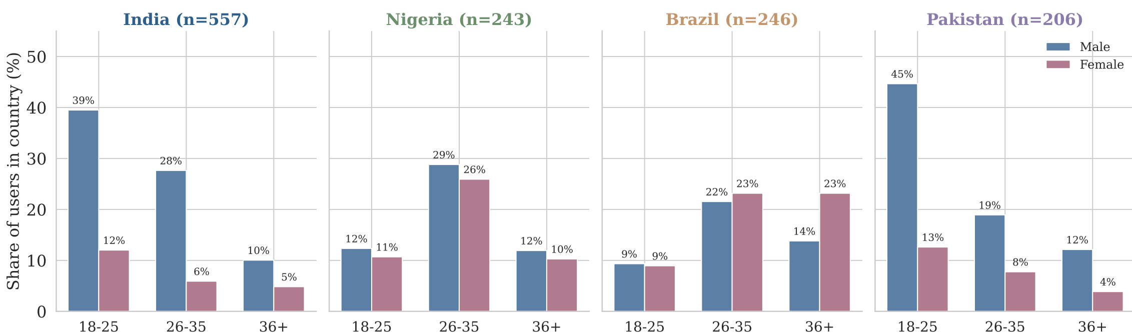


Fig. 1. Demographic composition by country, gender, and age group. Bars show the share of each country's participants falling in each age×gender cell (so panels are directly comparable despite different per-country sample sizes); labels on bars are within-country percentages.

Nomenclature and fields. Each participant's archive contains a complete set of their ChatGPT sessions, and their account-level metadata. We adopt the nomenclature implicit in that structure: a *conversation* is a single ChatGPT session (one page in the ChatGPT web or mobile UI) and contains a back-and-forth sequence of *messages*. Each message carries the sender role (user, assistant, or tool), the text content, and per-message metadata including fields identifying the underlying model version and records of any tool invocations (web search, code interpreter, image generation, memory updates, file uploads). From `user.json` we obtain each participant's subscription tier (Free, Plus, or Pro) and self-reported age and gender.

Note on sample bias. The conversations analyzed in this paper are a small fraction of the ChatGPT user base in each country, and the sample is *not* designed to be demographically representative: it is a convenience sample of Clickworker participants willing to export and share their complete ChatGPT history for compensation, and no stratified effort was made to match the broader national distributions of age, gender, education, or urban/rural residence. Results should therefore be interpreted as patterns *conditional on* being an active, digitally literate, Clickworker-accessible ChatGPT user, not as estimates of country-wide prevalence. Despite these limits, complete, user-level ChatGPT exports paired with self-reported demographics are, to our knowledge, not publicly available at this scale for any of these four countries, and the dataset offers a first conversation-level view of real-world usage that complements the aggregated server-side indicators that OpenAI and Anthropic have released.

3.2 WildChat Dataset

Ours is not the first publicly analyzable corpus of in-the-wild ChatGPT conversations. The best-known alternative is WildChat [30], which we use in Appendix B as a cross-dataset validation source. We briefly describe it here for completeness.

WildChat was constructed by offering free access to ChatGPT via a Gradio-based web interface in exchange for users' affirmative, consensual opt-in to share their conversation transcripts anonymously for research.

The original WildChat release contains over one million timestamped conversations comprising more than 2.5 million interaction turns, collected between April 2023 and May 2024 via APIs backed by GPT-3.5-Turbo and GPT-4 [30]. An expanded version, WildChat-4.8M, extends the corpus to approximately 4.8 million conversations. Roughly 24-26% of conversations in the original release used the GPT-4 API, with the remainder using GPT-3.5-Turbo. Each conversation record contains the full conversation history, timestamp, the model version used, toxicity flags from the OpenAI Moderation API, hashed IP addresses, and coarse geographic metadata (state and country). WildChat covers more than 60 languages, offering a linguistically diverse outlook of global ChatGPT usage. A key methodological limitation of WildChat is that, users who interact with a dedicated research platform are likely to be limited by the observer effect, participants who know their data is being collected may alter their behaviour. Additionally, its collection method, although consent-based, tends to skew toward users seeking free access to ChatGPT.

Because WildChat is freely available, it has become a de facto reference corpus for empirical studies of ChatGPT usage, and findings drawn from it are sometimes treated as generalizable. Our recruited cohort gives us a natural point of comparison: we apply the same topic classifier to both corpora and, in Appendix B, quantify how the proxy-based sampling frame systematically over-represents technical and programming-oriented use at the expense of the information-seeking and practical-guidance queries that dominate broader user populations.

4 Methods

This section describes the classification pipelines applied to the conversation corpus and a few conventions that are shared across all subsequent analyses. At a high level, every conversation is passed through three independent classifiers: (i) OpenAI's published 24-category topic taxonomy, (ii) unsupervised topic discovery via BERTopic with Gemini embeddings, and (iii) a replication of two of the Anthropic Economic Index's classifiers (*Asking/Doing/Expressing* intent and *work/coursework/personal* task purpose). The first two answer "what is this conversation about?" from two complementary angles (a fixed, directly comparable taxonomy and an open-vocabulary clustering); the third answers "what kind of task is the user doing?" Raw metadata (model version, tool invocations, subscription tier, conversation depth, prompt/response lengths) is extracted directly from the JSON export described in Section 3. All timestamps are stored as UTC Unix timestamps and converted to the primary timezone of each country for local time-of-day analyses: IST (UTC+5:30) for India, WAT (UTC+1) for Nigeria, BRT (UTC−3) for Brazil, PKT (UTC+5) for Pakistan; trend analyses aggregate at monthly granularity and exclude months with fewer than 20 conversations.

4.1 Topic Classification

We classify conversation topics using two complementary methods.

4.1.1 OpenAI Taxonomy. To enable direct comparison with Chatterji et al. [9], we apply their published 24-category fine-grained taxonomy, organized into 7 coarse domains: Practical Guidance, Seeking Information, Writing, Technical Help, Self-Expression, Multimedia, and Other/Unknown. We use the GPT-4o API with the same classification prompt as Chatterji et al., localized only where necessary (e.g., category name consistency).

The main procedural difference between our setup and theirs is that we label each *conversation* with a single topic rather than labelling every individual message. Chatterji et al. label at the message level because they have server-side access to the entire message stream; for an external replication, per-message labelling would multiply API cost by the median conversation depth (~6 turns), which is infeasible at the 202K-conversation scale of our corpus. Because conversations in our data are typically on a single topic (a user begins a new ChatGPT session when switching task), a single conversation-level label recovers most of the signal at a fraction of the cost. To stay as close as possible to Chatterji et al.'s classification prompt (which takes as input a "current" user message together with up to ten preceding messages as context), we construct the prompt input as follows. For conversations with more than ten messages, we use the 11th message as the "current" message and the preceding ten as context; for shorter conversations, we use the final message as the "current" message and whatever precedes it as context. Each message is truncated to 5,000 characters to avoid the classification instability that long prompts introduced during pilot runs, and messages are serialized in the familiar `[User]: ...\n[Assistant]: ...` format before being sent to the classifier.

4.1.2 BERTopic. The OpenAI taxonomy is designed to be stable across hundreds of millions of conversations and, for that reason, is deliberately coarse: a bucket like *Seeking Information* covers anything from symptom interpretation to celebrity trivia to legal clarifications. A privacy-preserving, server-side methodology cannot easily look inside these buckets, but a conversation-level corpus can. We therefore complement the supervised classifier with an unsupervised pipeline that is free to discover whatever structure exists in the data.

Concretely, we embed each conversation using Google's `gemini-embedding-001` model (3,072-dimensional vectors computed over the first ten user–assistant turns), run MiniBatch K-means with $k = 500$ to obtain semantically narrow micro-clusters, and then merge the 500 centroids via agglomerative hierarchical clustering with cosine distance until roughly 50 interpretable top-level topics remain per country. Cluster labels are drafted by `gpt-4o-mini`, adjudicated by Claude Sonnet 4.6 acting as a judge (low-confidence or non-homogeneous clusters are flagged for relabelling or reassignment), and manually reviewed by the authors. The final pipeline yields 50, 45, 36, and 53 topics for India, Nigeria, Brazil, and Pakistan respectively, covering 91–95% of conversations per country; remaining low-confidence conversations are reassigned to existing clusters by embedding-based cosine similarity with a per-cluster z-score threshold of 1.5. Complete hyperparameters, dendrograms, and the full judge prompts are documented in Appendix A.

4.2 Intent and Task-Purpose Classifiers

Anthropic [3] introduced a set of conversation-level classifiers that measure *how* a user is using the model, orthogonally to the topic of the conversation. We replicate two of these on our corpus using the GPT-4o API with the prompts released with the Anthropic Economic Index, applied independently to each conversation. Concretely:

- **Asking/Doing/Expressing.** The three-way intent classification introduced by Chatterji et al. [9] and also used by Anthropic: *Asking* conversations seek information or decision support, *Doing* conversations delegate a task (drafting, transforming, coding, executing) to the model, and *Expressing* conversations use the model for reflection or emotional communication.
- **Work / coursework / personal.** Whether the conversation is work-related, coursework-related, or personal.

5 Results

We organize our results around five analytical lenses applied to the same 1,252-user, 202,590-conversation corpus: (1) the OpenAI 24-category taxonomy [9] applied at the conversation level; (2) unsupervised topic discovery, which surfaces culturally specific themes that a predefined taxonomy absorbs into generic buckets; (3) the *Asking/Doing/Expressing* intent classifier of Chatterji et al. [9]; (4) the Anthropic-style work/coursework/personal task-purpose classifier [3]; and (5) raw-metadata descriptives on model adoption, tool use, conversation depth, prompt/response lengths, and subscription tier that are simply not available in any published platform report. Each lens is examined at the aggregate, per-country, gender-conditional, age-conditional, and temporal levels. A cross-dataset comparison against the WildChat corpus [30] is reported in Appendix B.

5.1 OpenAI Taxonomy: Topic Distribution and Demographic Conditioning

We apply the OpenAI 24-category fine-grained taxonomy (organized into seven coarse domains) from Chatterji et al. [9] at the conversation level using GPT-4o; details of the classifier are in Section 4. An aggregate treemap of the topic distribution appears in Figure 26.

5.1.1 Cross-country variation. Country-level breakdowns (Figures 2–3) reveal that the pooled average hides substantial per-country heterogeneity. India leads in *computer programming* (7.6%), *mathematical calculation* (3.1%), and *tutoring/teaching* (9.3%), consistent with India's large IT sector and the OpenAI Signals finding that Indian users are roughly three times above the global median for coding queries [21]. Nigeria dominates in *personal writing and communication* (12.7%) and shows the strongest engagement with text editing and critique. Brazil has the highest *health/fitness/beauty/self-care* share (7.7%, nearly double India's) and the lowest programming share (2.4%). Pakistan leads in *translation* tasks and matches Nigeria in *personal writing or communication* (12.7%).

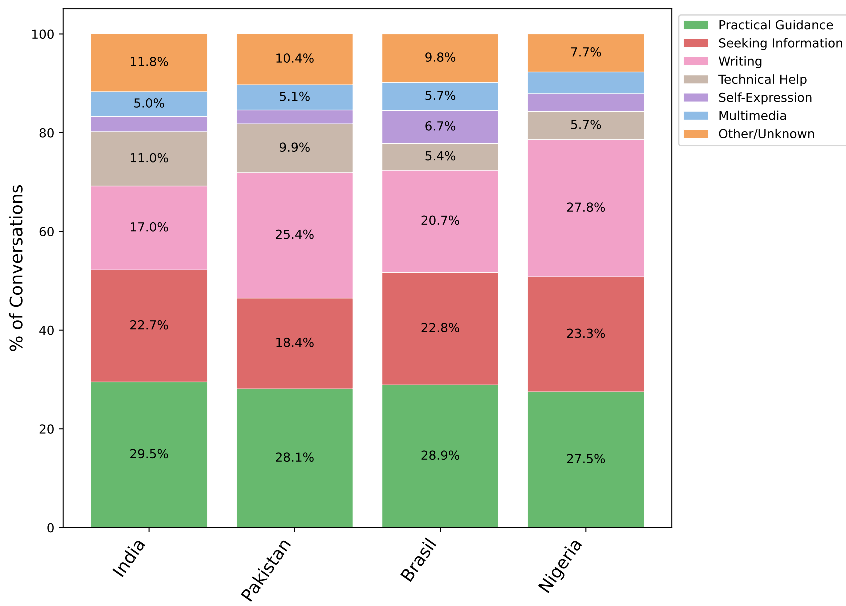


Fig. 2. Coarse topic distribution by country. Practical Guidance, Seeking Information, and Writing dominate across all four countries, collectively accounting for over 65% of conversations.

5.1.2 Comparison with OpenAI global averages. Table 3 compares the pooled coarse-topic shares across our four countries with those reported by Chatterji et al. [9]. Practical Guidance is the single largest domain (28.8%), closely matching the global share of 29.0%. The most informative comparisons are the deviations: our sample over-indexes on *Seeking Information* (+4.0pp) and *Writing* (+2.5pp) and under-indexes on *Technical Help* (−5.5pp) and *Multimedia*

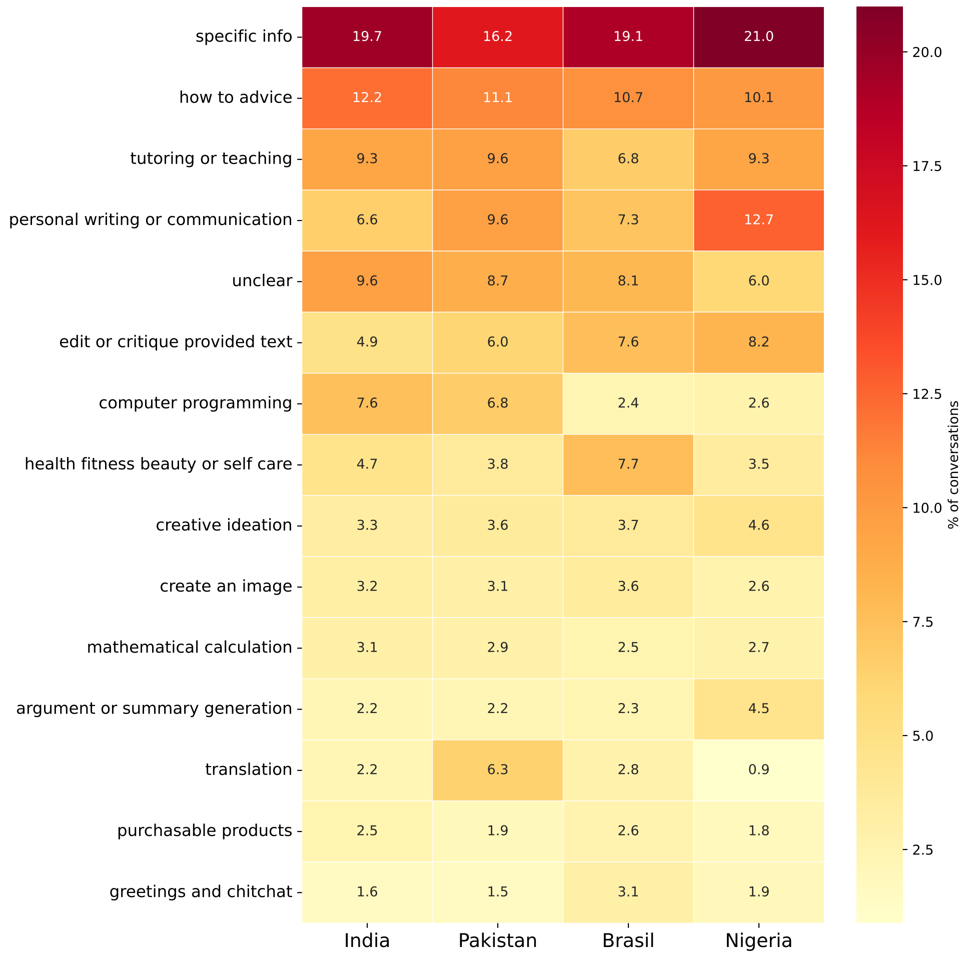


Fig. 3. Top 15 fine-grained topics by country. Colour intensity indicates the share of conversations devoted to each topic in each country.

(−2.0pp). At the fine-grained level, the five most common categories are *specific information* (19.3%), *how-to advice* (11.3%), *tutoring or teaching* (8.9%), *personal writing or communication* (8.5%), and *edit or critique provided text* (6.3%).

5.1.3 Gender and age conditioning. We emphasize up front that the gender and age composition of our sample reflects the convenience sample on Clickworker recruitment channel rather than the underlying ChatGPT user base; we therefore interpret gender- and age-*conditional* usage patterns, not gender or age composition, as findings.

Figure 4 shows that male users over-index on *Technical Help* (driven by computer programming) in all four countries, while female users over-index on *Practical Guidance* and *Self-Expression*, with the largest female skews in cooking, health/self-care, and tutoring. The raw gender shares per topic (Figure 27) confirm that these patterns hold after controlling for the overall gender composition of each country's sample. This conditional pattern is consistent across

Table 3. Coarse topic distribution: our four-country sample vs. OpenAI's reported global averages [9].

Topic	OpenAI global (%)	Ours (%)	Diff (pp)
Practical Guidance	29.0	28.8	−0.2
Seeking Information	18.0	22.0	+4.0
Writing	19.0	21.5	+2.5
Technical Help	14.0	8.5	−5.5
Self-Expression	5.0	3.9	−1.1
Multimedia	7.0	5.0	−2.0
Other/Unknown	8.0	10.2	+2.2

all four countries and directly matches, at the individual-conversation level, the name-inferred pattern reported by OpenAI for India [21].

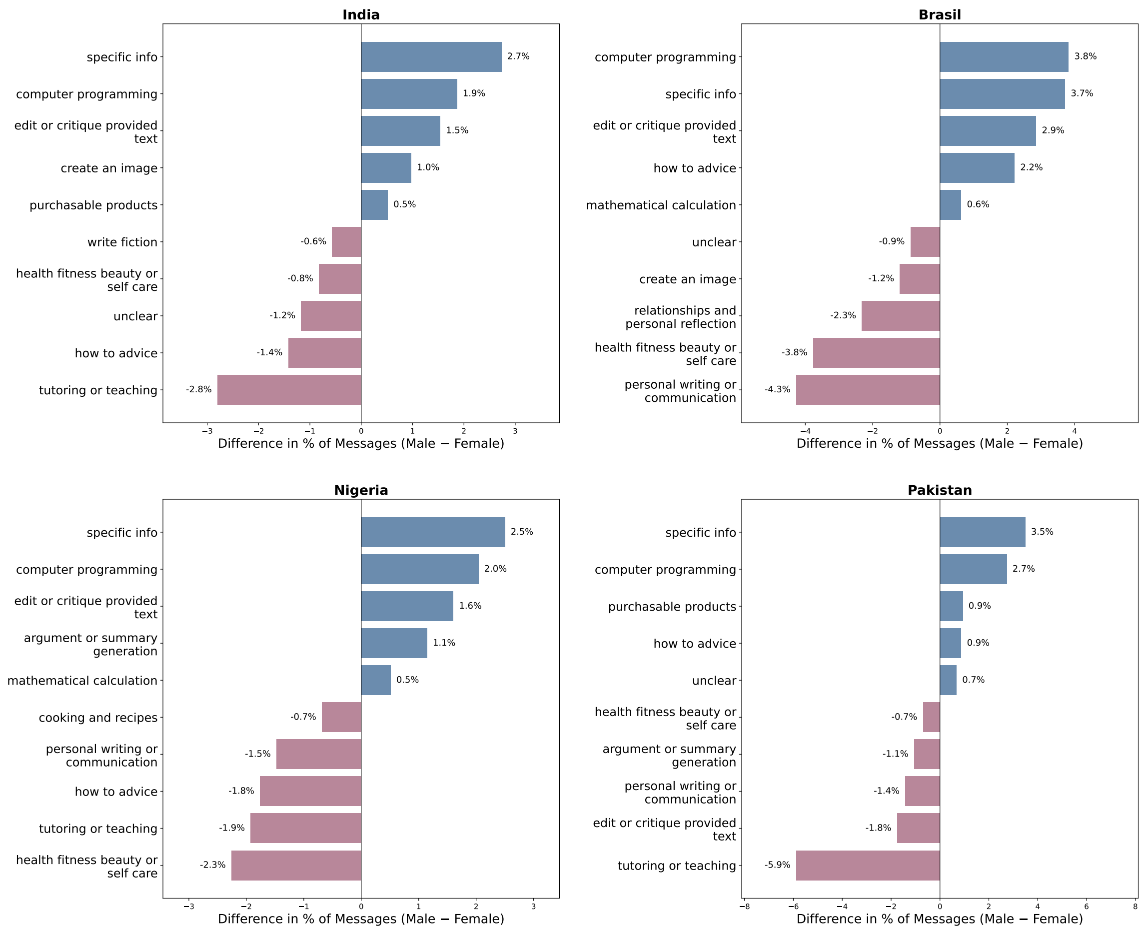


Fig. 4. Gender topic divergence: difference in topic share (Male − Female) per country. Blue indicates male over-representation; pink indicates female over-representation.

On age (Figure 28), the dominant cohort varies by country: 18–25 year olds contribute the largest share of messages in India and Pakistan (with particularly pronounced representation in *Technical Help*), while the 26–35 cohort is largest in Nigeria and Brazil. The 36+ cohort is consistently the smallest in India and Pakistan but substantial in Brazil (where older users are well represented across all topics). In India, this youth dominance is consistent with OpenAI [21], which reports that 18–24 year olds account for nearly half of Indian ChatGPT messages.

5.1.4 Temporal evolution. Figure 5 shows the share of each coarse topic over time in our four countries. The early period of ChatGPT adoption (2023) was dominated by *Writing* and *Technical Help* queries, which decline sharply (most dramatically in Brazil, where *Writing* briefly exceeded 40% before collapsing) while *Practical Guidance* and *Seeking Information* grow steadily and become the dominant categories by 2025. This convergence to guidance and information-seeking matches the late-2025 snapshot reported in OpenAI Signals for India [21]. A stacked-area version of the same data, which additionally overlays monthly conversation volume, is deferred to Appendix C (Figure 35).

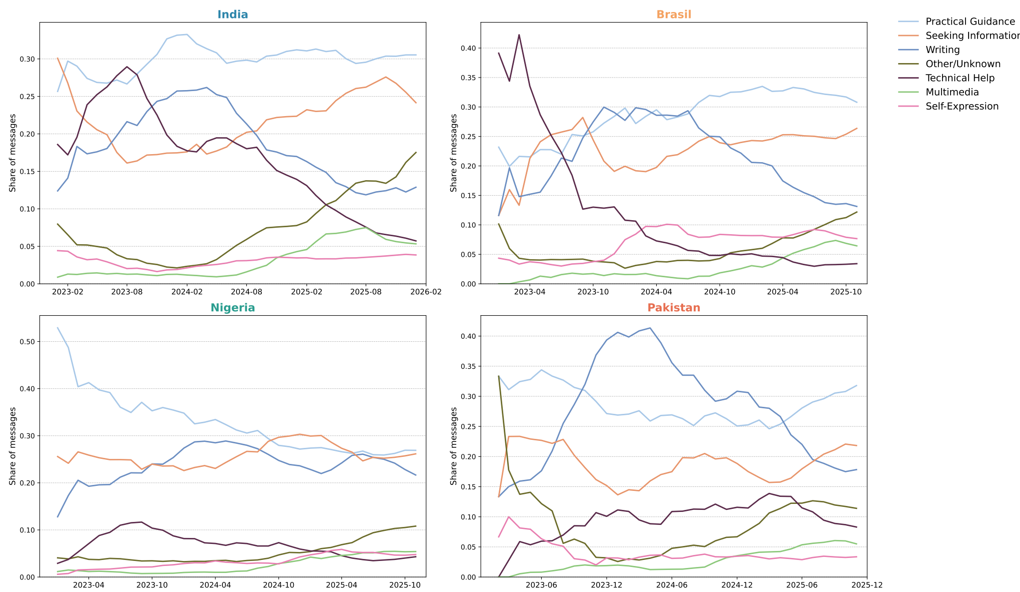


Fig. 5. Share of coarse topics over time, non-work messages only.

5.2 Unsupervised Topic Discovery: Culturally Specific Themes

A predefined 24-category taxonomy cannot, by construction, surface topics that the taxonomy designers did not anticipate. We therefore complement the supervised classifier with an unsupervised pipeline (BERTopic with Gemini embeddings) described in Section 4. To avoid confusion with the OpenAI "fine-grained" 24-category level, we refer to the per-country outputs of this pipeline as *topic clusters* (approximately 50 per country) and to the cross-country rollup as *themes* (ten categories).

5.2.1 Country-specific topic clusters. Figure 6 shows the top 15 topic clusters per country. Three observations stand out.

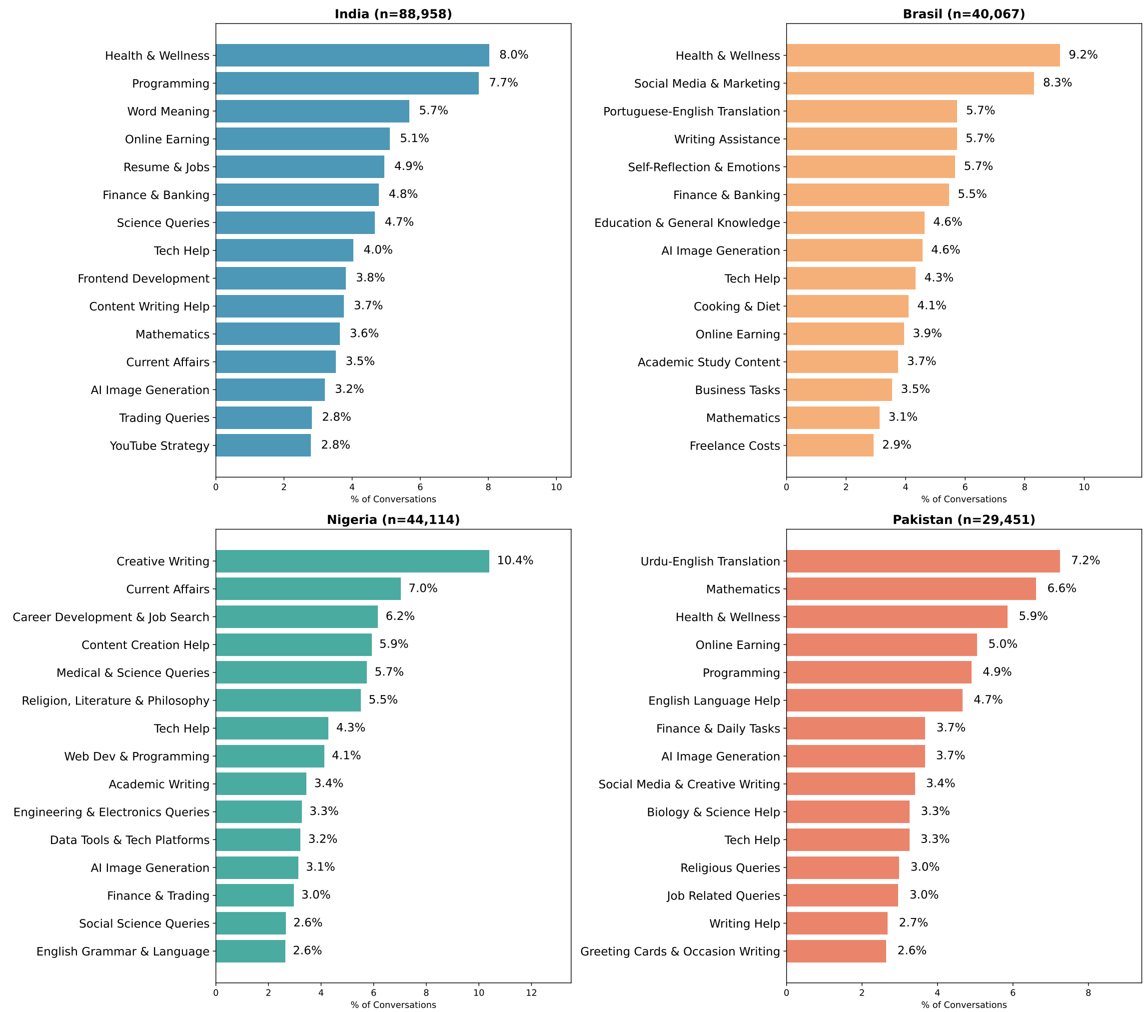


Fig. 6. Top 15 unsupervised topic clusters per country. The pipeline surfaces detailed information-seeking behaviour (e.g. the prevalence of Health and Wellness) and culturally specific clusters (e.g. Nigeria's religious topics, Brazil's self-reflection conversations, Pakistan's Urdu–English translation needs, India's digital-entrepreneurship themes) that the 24-category OpenAI taxonomy absorbs into generic buckets.

Health and wellness as a dominant theme. The OpenAI taxonomy subsumes these conversations under the smaller "health/fitness/beauty/self-care" fine category (7.7% in Brazil, 4.4% in India); the unsupervised clustering shows that once related wellness sub-topics—symptom interpretation, diet and fitness advice, mental-health queries—are aggregated semantically, health is the modal use case in two of our four countries. This growth in health-related queries is consistent with the trajectory documented by Karnam et al. [19], who find substantial increases in health and mental-health usage over time in their GDPR-donated ChatGPT corpus, and with ChatGPT serving as an accessible health-information substitute where formal consultation is expensive or uneven.

Culturally specific clusters that the OpenAI taxonomy does not have a name for. Religious clusters appear only in Nigeria (*Religion, Literature & Philosophy*, 5.5%) and Pakistan (*Religious Queries*, 3.0%), and are entirely absent from India's and Brazil's top-15 lists. *Urdu–English translation* is the single most prevalent cluster in Pakistan (7.2%); *Portuguese–English translation* is a top-5 cluster in Brazil (5.7%). *Self-reflection and emotional conversations* (5.7%) appears as a top-5 Brazilian cluster, unique among our four countries. *Online earning* and *YouTube monetization strategies* are top clusters in India (5.1% and 2.8%) and also surface prominently in Pakistan (5.0%). The prevalence of these clusters is likely amplified by the Clickworker recruitment channel (participants are, by definition, active gig-economy workers), but the fact that the same digital-entrepreneurship signature does *not* appear in Brazil or Nigeria (who were recruited through the same channel) suggests this reflects genuine cross-country differences in how ChatGPT is used as a resource for online income generation, rather than a pure sampling artifact.

Cross-country convergence on digital-entrepreneurship themes. Grouping the topic clusters into ten broad themes that span all four countries (Figure 7) shows that *Finance/Earning* is strikingly concentrated in India (17.7%) and Brazil (11.6%) but nearly absent in Nigeria (3.3%). *Programming/Tech* is most prominent in India (21.0%) and least in Brazil (8.3%). *Writing/Creative* is largest in Nigeria (11.6%) and Pakistan (8.6%). *Religion* is exclusively a Nigerian/Pakistani theme. *Translation/Language* is highest in Pakistan (11.2%).

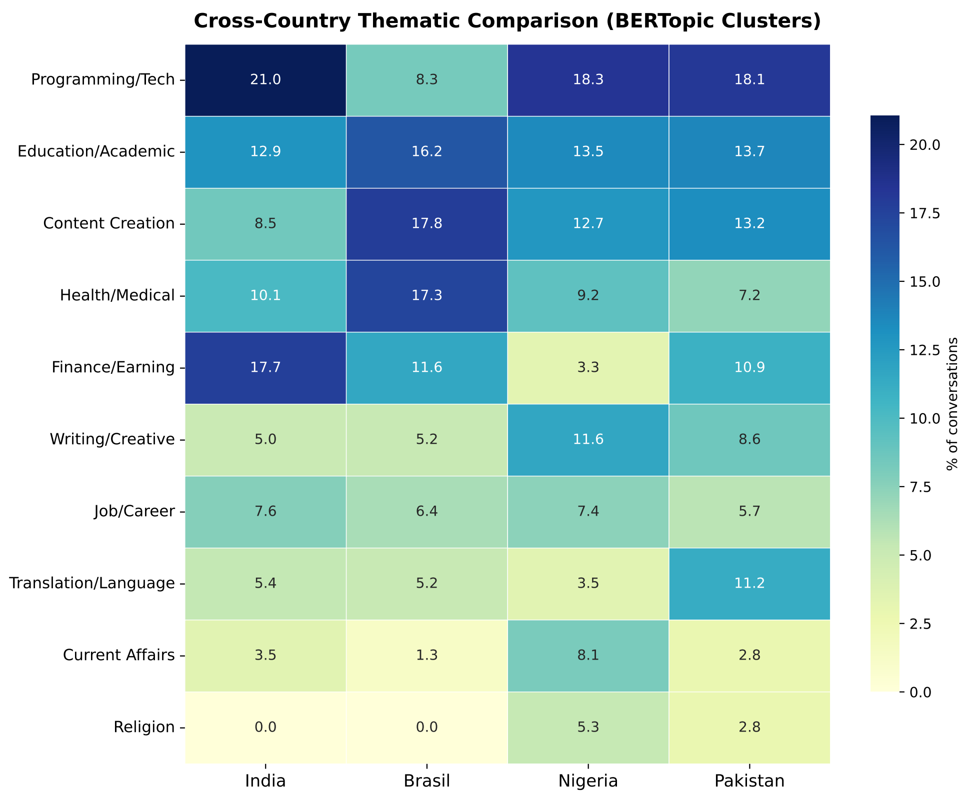


Fig. 7. Heatmap of ten broad themes (cross-country rollup of the unsupervised clusters) across countries. Colour intensity indicates the share of conversations per country.

5.2.2 Gender conditioning (themes). Figure 29 shows that gender-conditional patterns under the unsupervised lens are *sharper* than under the OpenAI taxonomy, because the clusters are thematically tighter. Male users over-index on *Programming/Tech* in every country and on *Finance/Earning* in India and Pakistan; *Religion* is a distinctly male theme in Nigeria and Pakistan. Female users over-index on *Health/Medical* in every country, and on *Content Creation* in Nigeria and Brazil. *Education/Academic* is the most female-skewed theme in Nigeria and Pakistan, suggesting that women in these countries disproportionately use ChatGPT as an academic support tool.

5.2.3 Age conditioning (themes). Figure 30 shows a clear generational gradient. The 18–25 cohort over-indexes on *Education/Academic* and *Programming/Tech* in India, Nigeria, and Pakistan, and on *Content Creation* and *Health/Medical* in Brazil, a student-oriented profile. The 26–35 cohort distributes most uniformly across themes. The 36+ cohort has the most distinctive profile: in India it over-indexes strongly on *Finance/Earning* and *Content Creation*; in Nigeria and Pakistan on *Religion* and *Current Affairs*; in Brazil on *Job/Career* and *Writing/Creative*. Taken together, these patterns describe a gradient from education and skill-building among younger users toward financial, professional, and civic topics among older users, visible across otherwise very different national contexts.

5.2.4 Temporal evolution (themes). Figure 31 shows the monthly evolution of the ten themes. The broad trends mirror the OpenAI-taxonomy view: early dominance of programming and writing themes gives way to a more diversified distribution where health, education, and finance grow as shares of monthly volume.

5.3 Work, Coursework, and Personal Use

Following the task-purpose framework in Anthropic's Economic Index [3], we label each conversation as work-related, coursework-related, or personal. The classifier was applied identically to all four countries. Personal use dominates everywhere: 61.5% of conversations in India, 63.7% in Brazil, 55.0% in Nigeria, and 55.4% in Pakistan. Work and coursework are of roughly comparable prevalence in the pooled sample (20.8% work vs. 19.6% coursework), though Brazil is an outlier with a substantially lower coursework share (13.6%) consistent with its older user base, and Pakistan has the highest coursework share (23.1%) consistent with its very young cohort. The 18.5% work share in our Indian subsample is noticeably lower than the 27% reported for global ChatGPT usage by Chatterji et al. [9], again consistent with the younger demographic composition of our sample and its larger share of students.

Figure 8 shows two conditional patterns worth flagging. First, as expected, the share of coursework drops monotonically with age in every country (from ~24–36% among 18–25 year olds to ~9–16% among 36+), while the share of work correspondingly rises (from ~10–18% to ~23–29%); this gradient is visible in all four countries despite their different age profiles. Second, and more striking, female users use ChatGPT for coursework at substantially higher rates than male users in three of four countries: 26.1% vs. 18.2% in India, 32.8% vs. 19.7% in Pakistan, and 23.5% vs. 20.2% in Nigeria. Brazil is the exception (12.4% female vs. 14.9% male), in line with its distinct older, more female-skewed, more work-oriented profile. The India/Pakistan/Nigeria pattern is consistent with ChatGPT serving as an academic support resource that women in these countries lean on disproportionately, complementing the female over-representation on the *Education/Academic* theme we reported under the unsupervised lens.

5.3.1 Topic distribution conditioned on task purpose. Figure 9 quantifies the topic composition of work vs. non-work conversations under the OpenAI taxonomy. The first striking feature is that *Writing* and *Technical Help* combined account for roughly 45–50% of work-related conversations in every country (44.6% in India, 48.3% in Nigeria, 47.3% in Brazil, 50.0% in Pakistan); within that pair, *Writing* alone is consistently the single largest work bucket, accounting for

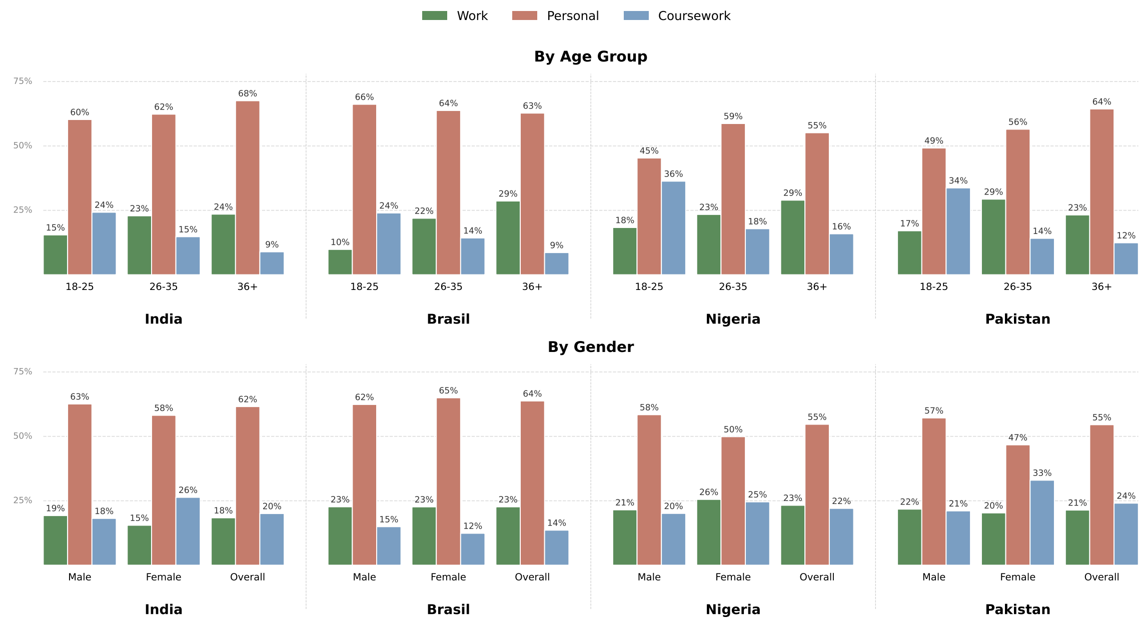


Fig. 8. Distribution of task purpose (work, coursework, personal) by age group (top) and gender (bottom), by country.

28–42% of work messages. Even at the end of 2025, in other words, ChatGPT as a workplace tool is still predominantly a writing-and-editing tool for our four countries; the coding assistant framing that dominates popular discourse is secondary. The second feature is the mirror image on the non-work side: *Practical Guidance* is consistently the largest non-work bucket (28–31% across countries), followed by *Seeking Information* (20–26%), together covering more than half of non-work conversations. *Self-Expression* is nearly three times more prevalent in non-work than in work (3–8% vs. 1–3%), and only shows up meaningfully in the personal setting. The cross-country stability of this pattern (writing-heavy work, guidance-and-information-heavy personal use) mirrors the OpenAI Signals finding for India [21].

Conditioning the unsupervised topic clusters on task purpose (Figures 10–11 and 32) makes the content of these work/personal conversations concrete. The writing-heavy work profile surfaced above resolves, under the unsupervised lens, into a small number of recurring work themes: *CV and cover-letter drafting*, *job-application and interview preparation*, *professional email and correspondence*, and *content creation for social and short-video platforms*. These are the specific tasks behind the "Writing at 28–42% of work" finding. *Online earning* appears as a top-5 work cluster in India, Pakistan, and Brazil, suggesting that across these economies users frequently turn to ChatGPT to navigate precarious or entrepreneurial labour-market conditions. Resume building and job applications dominate work-related use in India, while career development and content creation lead in Nigeria. Programming and technical tasks appear consistently but are never the largest work cluster in any single country, reinforcing that the technical-help slice of work is narrower and more targeted than the writing slice.

Personal conversations (Figure 11) are anchored by *Health and Wellness* in India and Brazil (where it is the top personal theme) and Pakistan (second). Beyond health, personal use is markedly more culturally differentiated: Nigeria stands out for creative writing and current affairs; Pakistan for Urdu–English translation and religious queries; Brazil

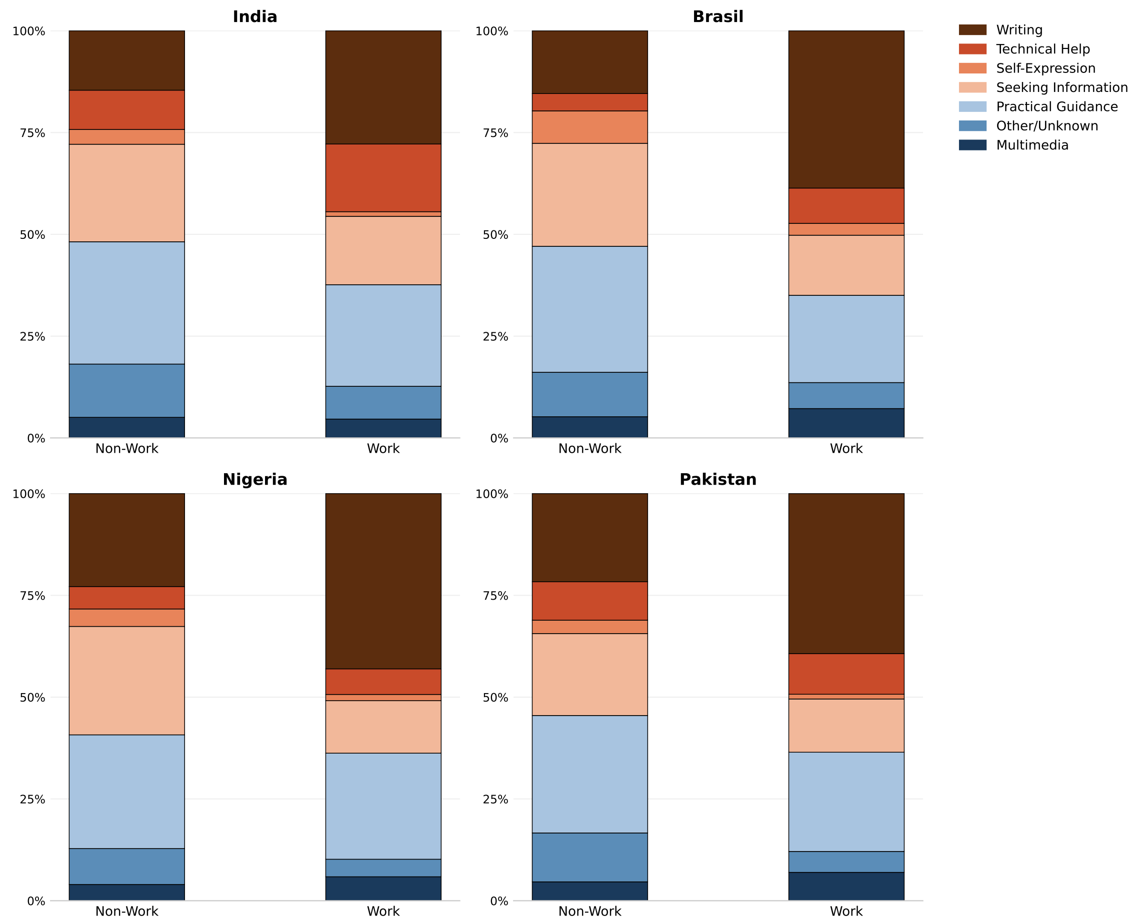


Fig. 9. Share of coarse topics for work-related vs. non-work ChatGPT messages, by country.

for self-reflection and emotional support, suggesting a more expressive, interpersonal mode of engagement that is largely absent from the other three countries.

Coursework conversations (Figure 32) are dominated by STEM subjects in every country. Programming and mathematics lead in India and Pakistan; Nigeria concentrates on academic writing, engineering, and data tools; Brazil is distinctive in its large share of general-knowledge queries and Portuguese–English translation, pointing to language support as a significant driver of educational use.

5.3.2 Temporal evolution of task purpose. Figures 12–13 show that India, Nigeria, and Pakistan exhibit a remarkably stable work/non-work split throughout the observed period, with non-work messages consistently accounting for 75–85% of all queries. Brazil is the notable exception, beginning in early 2023 with an almost even split between work and non-work usage before converging to the ~80% non-work share seen elsewhere by mid-2024. The stability of the Indian split is consistent with OpenAI [21], which reports that the shift toward non-work dominance in India occurred before mid-2024 and has remained steady since.

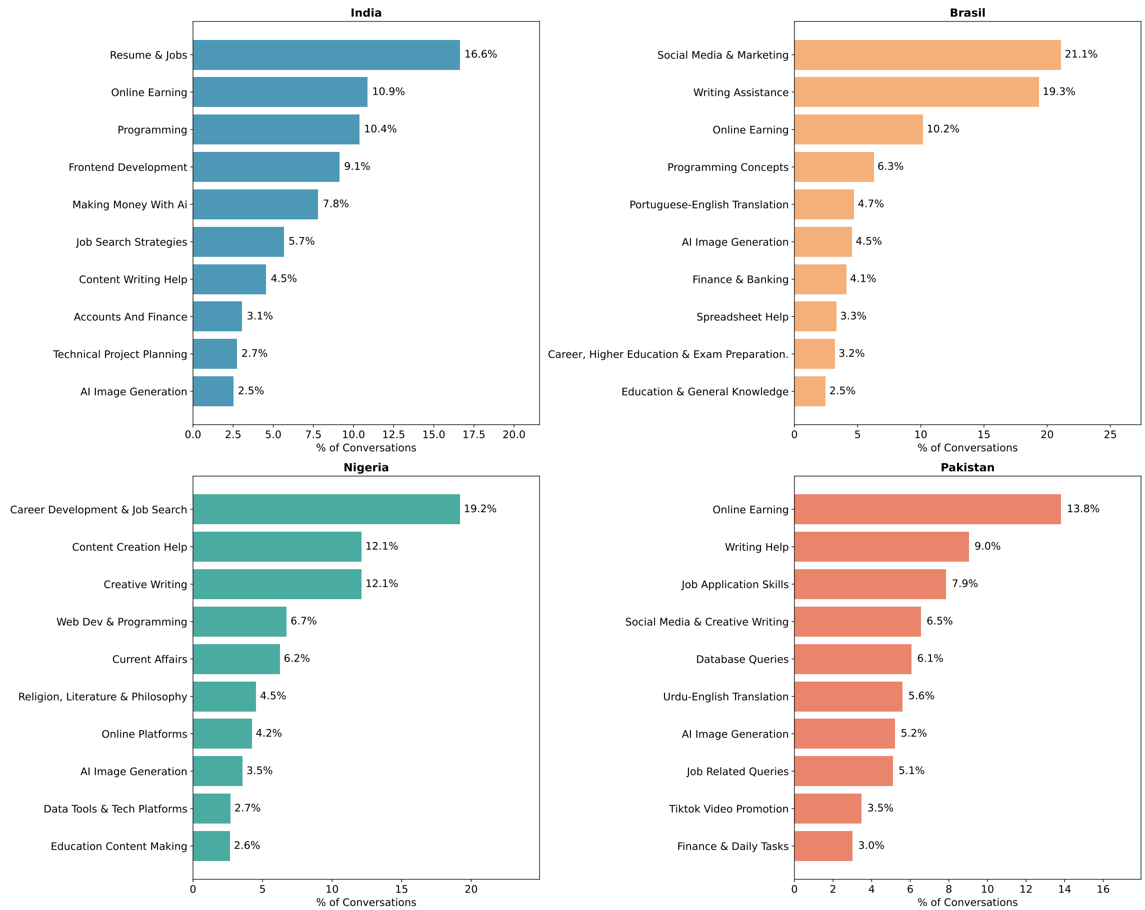


Fig. 10. Top 10 unsupervised topic clusters for work-related conversations, by country.

5.4 Asking, Doing, Expressing

We apply the three-way *Asking/Doing/Expressing* intent classifier introduced by Chatterji et al. [9] (*Asking* = seeking information or decision support; *Doing* = requesting the model to execute a task; *Expressing* = using the model for reflection or emotional communication). Figure 33 shows the distribution conditioned on task purpose (work vs. non-work), and Figure 14 shows the monthly trend.

The two comparisons tell a consistent story. Work-related conversations are disproportionately *Doing* (drafting, transforming, or executing tasks) while non-work conversations are disproportionately *Asking* and *Expressing*. The shift is most pronounced in Brazil, Nigeria, and Pakistan, where *Doing* becomes the plurality or outright majority of work conversations (50.2%, 52.5%, and 49.1% respectively), while *Asking* drops by 13–18 percentage points from its non-work level. India shows a directionally consistent but smaller shift: *Asking* remains the plurality even in work (42.2%), with *Doing* at 38.7%. Across all four countries, the *Expressing* share declines from non-work to work, consistent with the OpenAI Signals finding for India that work conversations are more task-oriented [21].

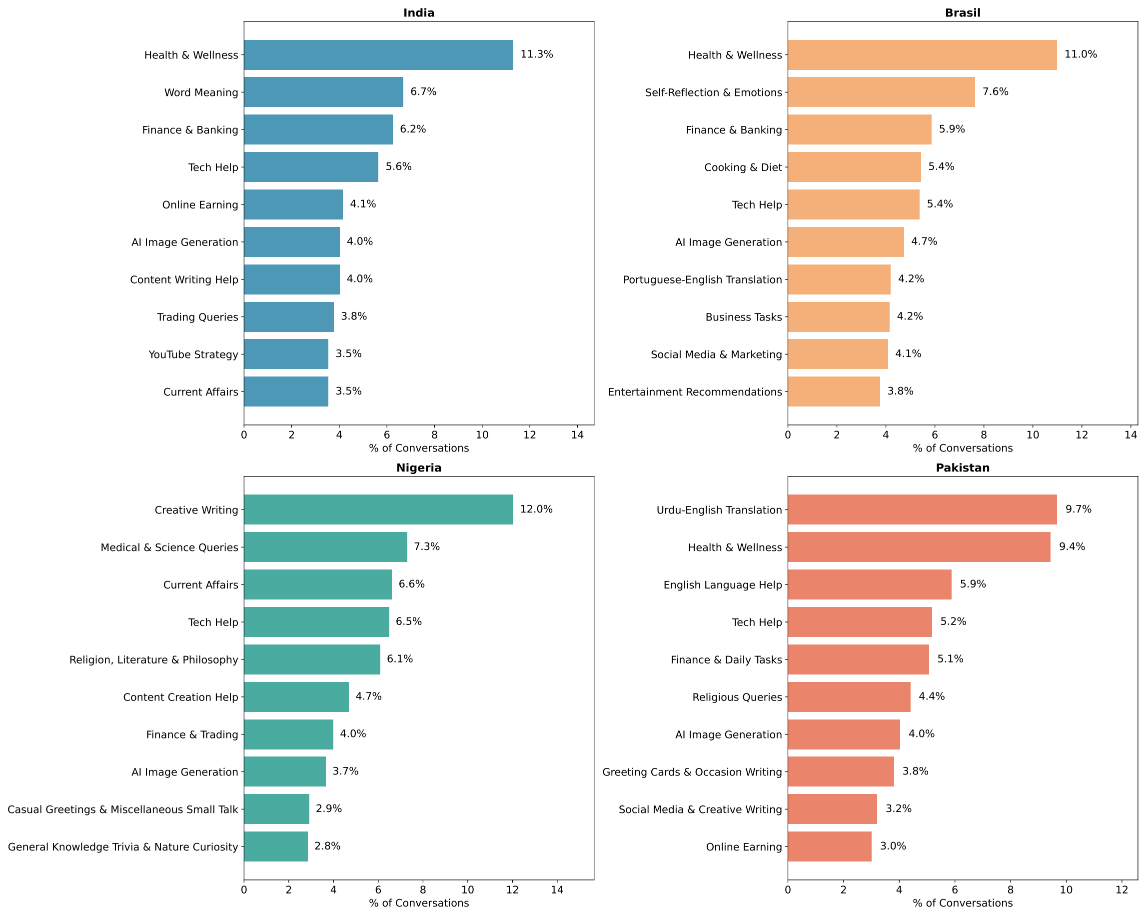


Fig. 11. Top 10 unsupervised topic clusters for personal conversations, by country.

Two points are worth emphasizing across these temporal trends. First, although *Asking* is the dominant mode in every country throughout the observation period, its share has gradually declined as *Doing* has grown, most visibly in India and Nigeria, where *Doing* rises from below 20% in early 2023 to 30–35% by late 2025. Chatterji et al. [9] report a similar global gradient, with *Doing* reaching roughly 40% of messages by 2025; our four-country corpus reproduces this trajectory, suggesting that as users mature on the platform they progressively delegate more autonomous task execution to the model. Second, there is meaningful cross-country heterogeneity even though the relative ordering (*Asking* $\gg$ *Doing* $\gg$ *Expressing*) is consistent: India and Nigeria show the most pronounced shift toward *Doing*, Brazil maintains the highest share of *Expressing* throughout, and Pakistan's *Doing* share remains smaller and more volatile. Perhaps most interestingly, the *Expressing* share (though small in absolute terms) exhibits a gentle upward drift over time across all four countries. This is consistent with two independent lines of evidence: Karnam et al. [19] document rising social framing and companion-like treatment of ChatGPT over time, and Fang et al. [14] find that heavier users engage in proportionally more reflective (as opposed to purely transactional) exchanges. We flag this as an underexplored trend worth following in future work (Section 6): if users are using ChatGPT for reflective, affective, or companion-like

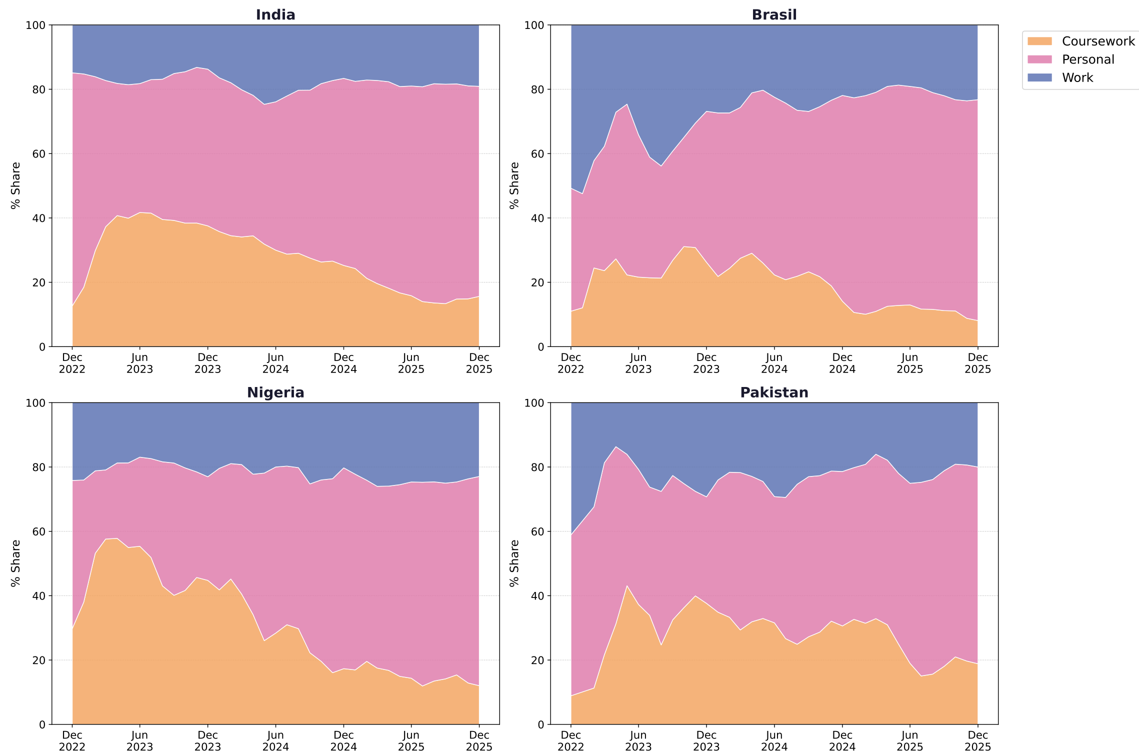


Fig. 12. Temporal share of work, coursework, and personal conversations (3-month rolling average).

conversation at a slowly growing rate, the downstream implications for well-being, emotional-support infrastructure, and product design are substantial.

5.5 Raw-Metadata Descriptives: Model Adoption, Tools, and Engagement

Unlike platform reports, our conversation exports expose raw client-side metadata: the model version used for each turn, tool invocations (web search, code interpreter, image generation), prompt and response lengths, conversation depth (turn counts), and Free vs. Plus subscription status. We summarize these below.

Model version adoption. Figure 15 shows three distinct eras. The initial Text-Davinci model family (GPT-3.5) was universal through early 2024. GPT-4o then displaced it within roughly two months of its mid-2024 launch, dominating at over 95% of conversations through mid-2025. The GPT-5 family (GPT-5, GPT-5.1, GPT-5.2) began displacing GPT-4o from mid-2025 onward, with GPT-5.2 accounting for the majority of conversations by early 2026. The rapidity of each transition suggests that most users in our sample adopt whatever default model the platform offers rather than actively selecting advanced variants.

Tool and feature usage. Figure 16 shows feature adoption rates by country, recomputed from the raw JSON conversation data (authoritative user list only, n = 1,252). Web search is the most-used tool, appearing in 12.3–15.9% of conversations, with India (15.9%) and Brazil (15.6%) leading. Code interpreter usage is relatively uniform (3.4–4.1%)

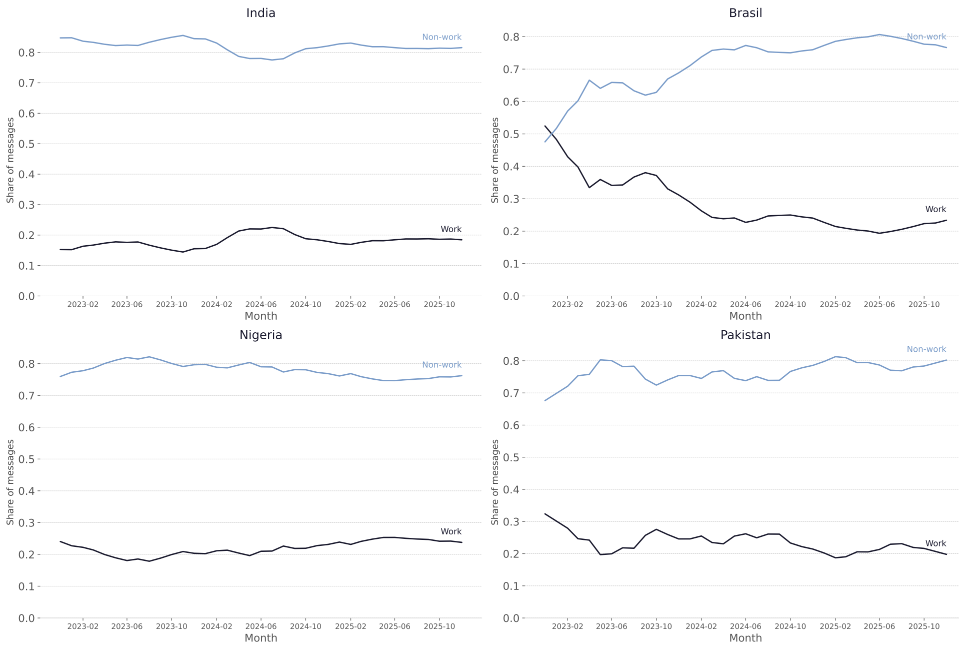


Fig. 13. Share of work-related vs. non-work messages over time, by country.

and image generation usage remains below 1% everywhere. These rates are modest overall, pointing to either limited feature awareness, limited relevance to user tasks, or free-tier access restrictions.

Free vs. Plus. Only 10.7% of users in our cohort (134 of 1,252) are ChatGPT Plus subscribers, and the distribution is strongly skewed toward India (116 of 134); Nigeria has only 1 Plus user and Pakistan only 3 (Table 4). Plus users run deeper conversations, write longer prompts, and receive longer responses (Figure 17; all comparisons shown with bootstrap 95% CIs). The causal direction is ambiguous: users who invest in a paid tier may naturally use it more intensively, or the tier itself unlocks deeper interaction.

Table 4. Free vs. Plus user counts by country (authoritative 1,252-user cohort).

Country	Free	Plus
India	441	116
Nigeria	242	1
Brazil	232	14
Pakistan	203	3
Total	**1,118**	**134**

Conversation depth and prompt/response lengths. Figure 18a shows the mean number of turns per conversation by country with bootstrap 95% confidence intervals. Mean conversation depth ranges from 11.6 turns in Nigeria to 14.6 in Pakistan; all CIs are tight, so the differences are statistically reliable despite being moderate in magnitude. The

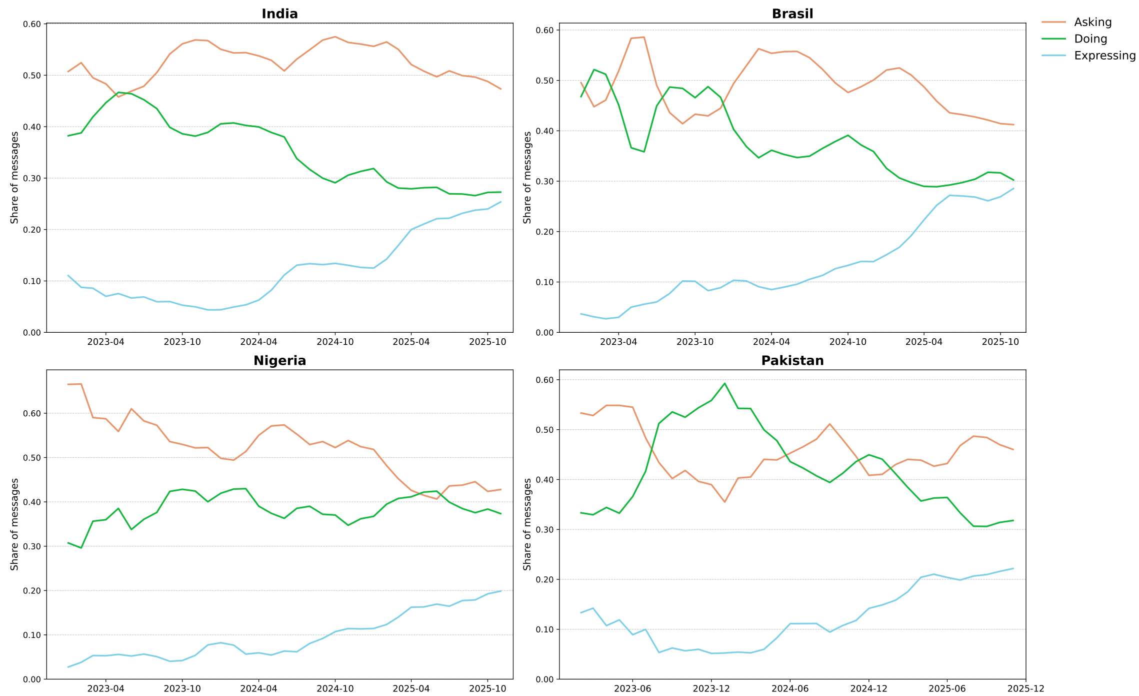


Fig. 14. Temporal shift in *Asking*/*Doing*/*Expressing* messages (3-month rolling average) by country.

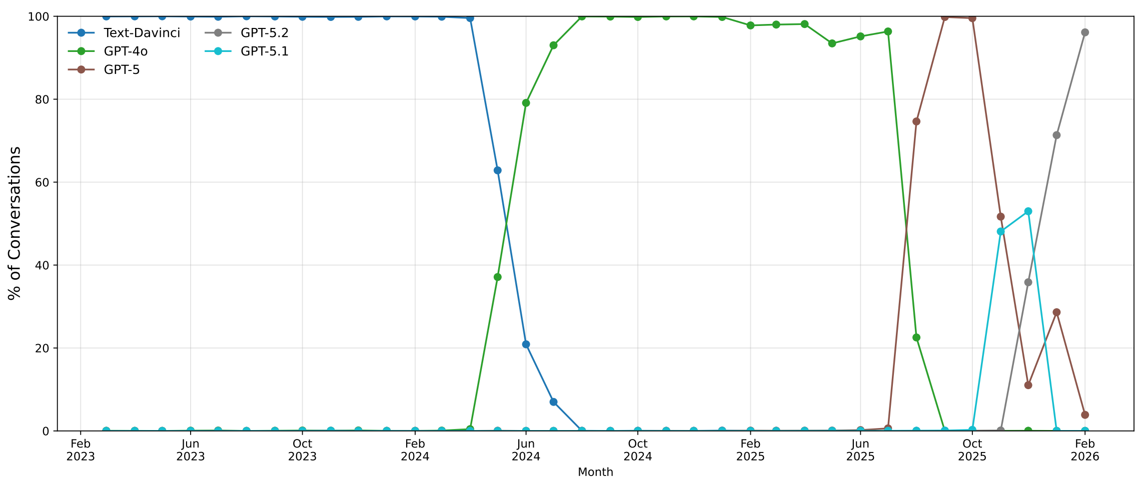


Fig. 15. Model version adoption over time. Three successive waves are visible: Text-Davinci (GPT-3.5), GPT-4o, and the GPT-5 family.

distributions are heavily right-skewed, so the mean reflects both the typical short exchange and the long-tail power users. Figure 18b shows mean prompt and response lengths per conversation with bootstrap 95% CIs. Pakistani users

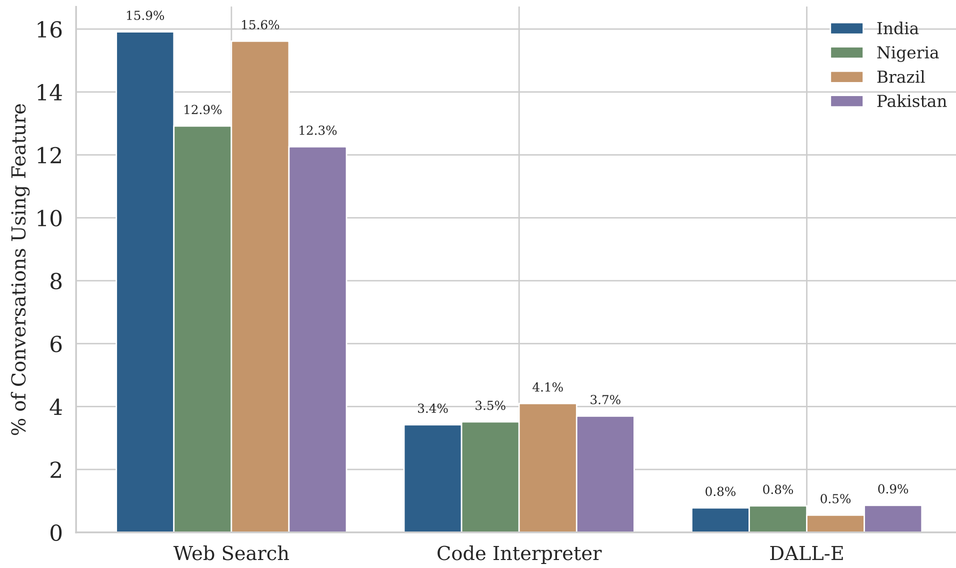


Fig. 16. Advanced feature adoption rates by country, computed over the raw conversation JSON for the authoritative 1,252-user cohort.

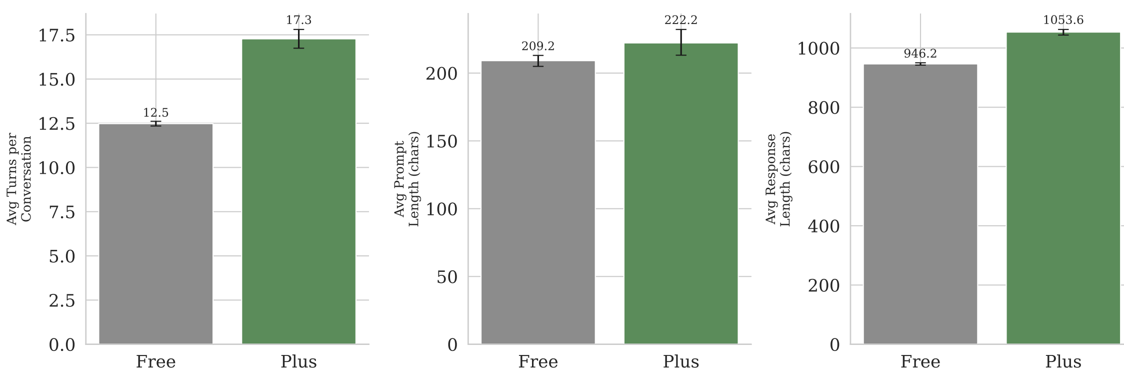
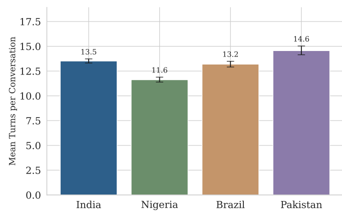


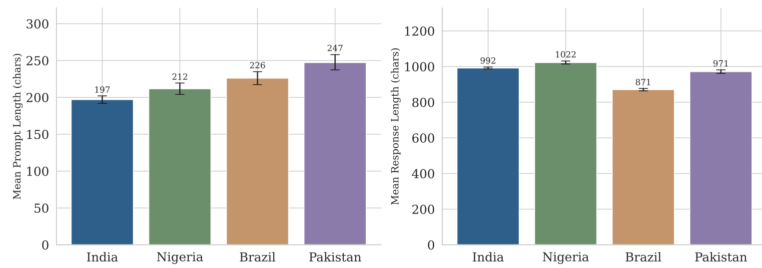
Fig. 17. Behavioural comparison between Free and Plus users (bootstrap 95% CIs).

write the longest prompts on average, consistent with their higher share of technical and educational queries; response lengths are more uniform and consistently several times longer than prompts.

Adoption curves and demographic share of activity. Figure 34 shows monthly active users and conversation volume by country; India has the steepest trajectory. Figures 19–20 show gender and age shares of weekly active users, estimated via bootstrapping with equal samples per demographic group. All four countries show a slight increase in the female share of conversations over time, and in India the growth accelerates from mid-2025, matching the trend reported by OpenAI [21]. We do not treat the cross-country level differences in gender or age shares as substantive findings (they reflect the sampling bias) but the *within-country trends over time* are consistent with platform-level reports.



(a) Conversation depth



(b) Prompt and response lengths

Fig. 18. (a) Mean turns per conversation by country; (b) mean prompt length (left) and response length (right) per conversation by country. All panels show bootstrap 95% CIs.

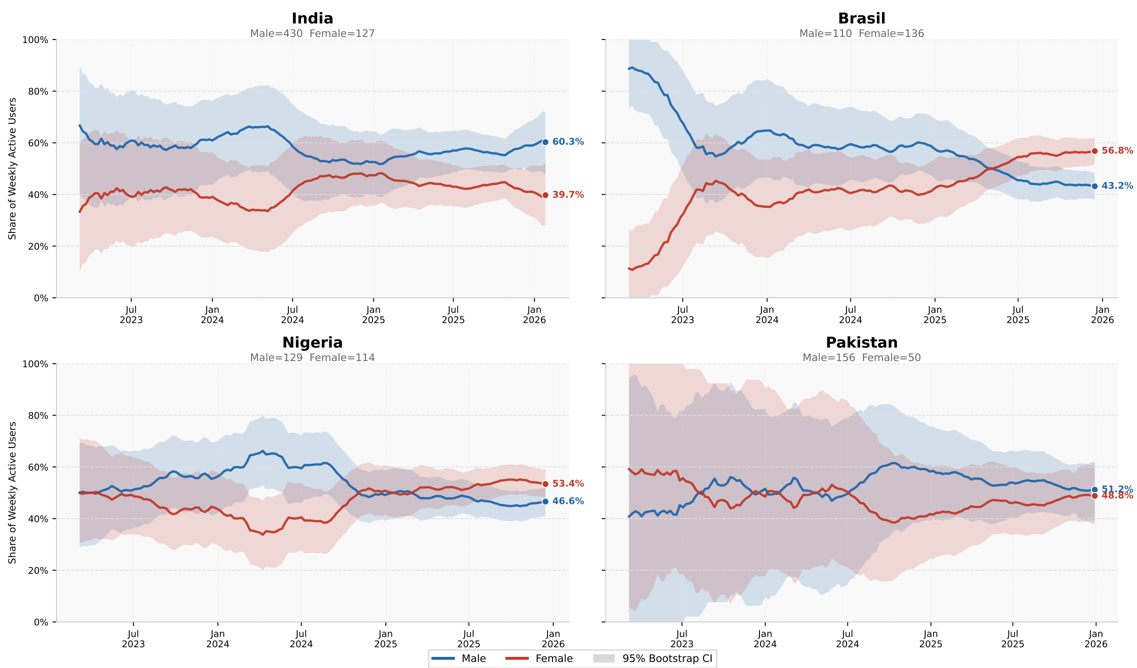


Fig. 19. Weekly-active-user share by gender and country (bootstrap CIs, 500 iterations).

Figure 21 shows usage intensity by hour of day and day of week (all timestamps converted to each country's local timezone). Indian and Pakistani users are most active in the evening and late night; Nigerian users peak during business hours; Brazilian users are comparatively uniform throughout the day. Weekend activity is slightly lower than weekday activity across all countries. We report these as descriptive context rather than as substantive findings.

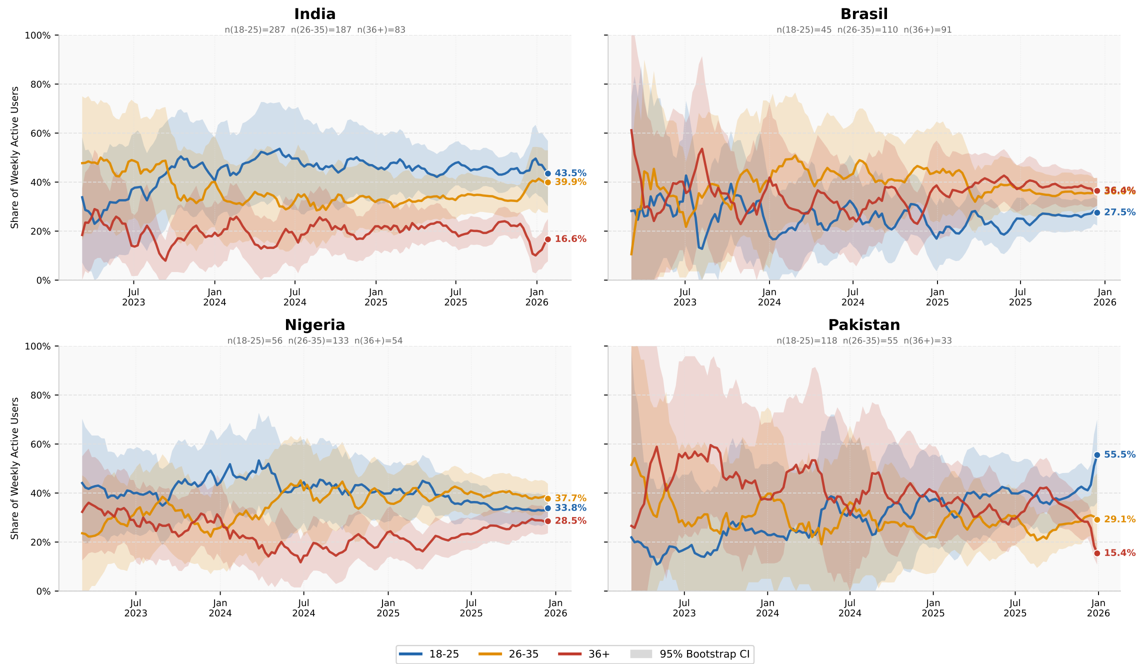


Fig. 20. Weekly-active-user share by age group and country (bootstrap CIs, 500 iterations).

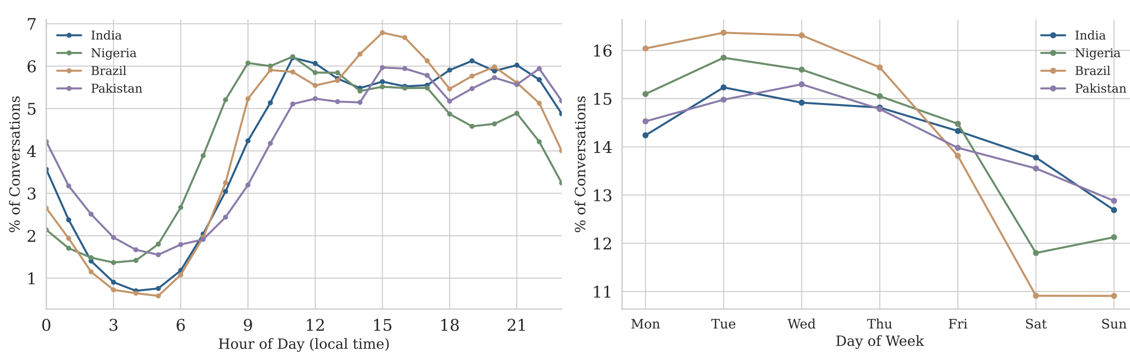


Fig. 21. Usage intensity by hour of day (left) and day of week (right) in local time, one line per country.

6 Discussion

Our results paint a picture of ChatGPT usage in the Global South that diverges in important ways from the productivity-focused narrative that dominates public and policy discourse around large language models. We organize this discussion around the broader implications of our findings, what they suggest for future research, and what they cannot tell us.

6.1 ChatGPT as Invisible Infrastructure

The single most striking pattern across our results (visible under every analytical lens we applied) is that ChatGPT in these four countries is overwhelmingly a *personal* tool. Personal conversations account for 55–64% of usage in every country; *Practical Guidance* and *Seeking Information* are the dominant coarse categories; and the unsupervised pipeline surfaces health and wellness, translation, and everyday problem-solving as the largest topic clusters. Even within work-related conversations, the dominant tasks are not the high-skill professional applications that receive the most attention (software engineering, data analysis, legal reasoning) but rather job-search support, CV and cover-letter drafting, and online-earning strategies, tasks at the boundary of formal employment and personal economic survival.

This matters for how we think about the economic value of conversational AI. The mainstream framing, reflected in Eloundou et al. [13] and Anthropic [3], centres on workplace productivity: how much faster can a programmer, writer, or analyst work with an LLM? That framing is important but structurally incomplete. A user who asks ChatGPT to interpret medical symptoms before deciding whether to pay for a doctor visit, or to diagnose why a household appliance is malfunctioning, or to draft a message to a landlord in a second language, is extracting real economic value, but it is value that does not appear in any standard measure of labour productivity. Economists recognize this as household production [8]: the non-market work that sustains daily life but is invisible to GDP accounting. Our data suggest that the primary mode of ChatGPT usage in the Global South falls squarely in this category. If so, productivity-oriented evaluations of AI impact will systematically undercount the technology's value in precisely the markets where it may matter most: markets where formal consultation (medical, legal, financial) is expensive relative to income, and where an always-available, zero-marginal-cost advisor fills a genuine gap.

The *content* of personal use reinforces the point: it is not frivolous. Health and wellness is the single most prevalent unsupervised theme in India (8.0%) and Brazil (9.2%), and a top-5 theme in Pakistan, in countries where the ratio of physicians to population is a fraction of that in the OECD. Translation is the top cluster in Pakistan and a top-5 cluster in Brazil, bridging the gap between a user's dominant language and the language in which information exists. A student in Pakistan who uses ChatGPT to understand a difficult concept in English (because no widely available Urdu explanation exists) is receiving a service with direct educational value. A gig worker in India who uses it to draft a competitive freelancing proposal is investing in income generation. These returns are real but invisible to the question "what percentage of knowledge workers use AI at their desks?" Any aggregate cost-benefit analysis that focuses exclusively on the workplace will miss most of the picture in these markets.

We cannot assess the quality of ChatGPT's health advice from our data, and we make no claims that it is a substitute for professional care. But the sheer volume of health-related usage suggests that millions of users are already treating it as a first-line health-information resource. Understanding whether this substitution improves or harms health outcomes is an urgent empirical question that our data can motivate but not answer.

6.2 What Fixed Taxonomies Cannot See

A methodological lesson runs through our results: a taxonomy designed in one context will structurally fail to see the things that matter in another. OpenAI's 24-category classification is well-designed for its purpose (stable, reproducible, directly comparable across billions of conversations) but it has no category for Islamic jurisprudence queries, Urdu–English code-switching, YouTube monetization strategy, or emotional self-reflection in Portuguese. These are not edge cases in our data; they are top-5 clusters in their respective countries. The unsupervised pipeline finds them because it is free to discover whatever structure the data contains, rather than projecting onto a fixed vocabulary.

This observation generalizes beyond our specific context. Any study that classifies AI usage through a fixed taxonomy will inherit the blind spots of whoever designed that taxonomy. In a field where the most influential usage statistics come from platform operators in a small number of countries, this creates a systematic risk: the categories through which we understand AI usage worldwide are calibrated to the usage patterns of early adopters in high-income, English-speaking markets. Conversation-level access to the underlying data (the kind of access our study provides) is the only corrective.

6.3 From Asking to Doing to ... Expressing

The temporal trajectory of *Asking*, *Doing*, and *Expressing* carries implications beyond a simple adoption curve. The rise of *Doing* (from below 20% in 2023 to 30–35% by late 2025 across all four countries) reflects growing user trust and model capability: users who once asked ChatGPT for information increasingly delegate the execution of tasks to it. This matches the global trend reported by Chatterji et al. [9] and is consistent with a maturing user base that has learned what the model can reliably do.

The more provocative signal is the slow but consistent upward drift in *Expressing*. It remains the smallest of the three modes throughout our observation window, so we are cautious about over-interpreting it, but the direction is visible in all four countries. Brazil, which has the most distinctive *Expressing* profile (anchored by a large self-reflection and emotional-conversation cluster), may be a leading indicator. This trajectory aligns with independent evidence from Karnam et al. [19], who document that ChatGPT interactions are becoming increasingly socially framed, with users more frequently treating the system as a companion, and from Fang et al. [14], who observe that heavy users gravitate toward reflective exchanges. Several forces could drive this trend: improvements in model conversational warmth and emotional register; the normalization of talking to AI as a cultural practice; or, more soberly, the absence of affordable alternatives for emotional support in contexts where mental-health services are scarce and stigmatized. The downstream implications depend heavily on which of these forces dominates. If users are turning to ChatGPT for companionship or emotional processing at scale, questions of parasocial attachment [25], displacement of human social ties, and the quality of AI-mediated emotional "support" become urgent, and they are questions that the current generation of platform reports, which classify at the topic level rather than the relational level, are not designed to answer.

6.4 Demographic Patterns: What We See and What We Cannot Claim

The gender-conditional patterns in our data are consistent and cross-nationally stable: men over-index on programming, finance, and religion; women over-index on health, education, and content creation. These reproduce the name-inferred patterns reported by OpenAI [21] for India, providing independent validation at the individual-conversation level.

The harder question is what these patterns *mean*. One reading is that gendered AI usage mirrors existing occupational and social roles: men in South Asia are over-represented in IT and finance; women carry a disproportionate share of health-care and educational responsibilities within families. Under this reading, AI usage reflects societal structure rather than reshaping it. An alternative reading, more hopeful, is that women's disproportionate use of ChatGPT for education and coursework (particularly the striking female coursework skew we observe in India (26.1% vs. 18.2%), Pakistan (32.8% vs. 19.7%), and Nigeria (23.5% vs. 20.2%)) represents AI serving as an equalizer: a private, judgement-free educational resource that partially compensates for barriers to formal education. Our data cannot distinguish between these readings. Otis et al. [22] argue that generative AI's impact on global inequality will depend on whether it amplifies existing advantages or compensates for existing disadvantages; the gendered usage patterns we observe are a concrete, measurable instantiation of that question, but resolving it requires larger samples and, ideally, panel data that track individual users over time.

The age gradient is cleaner to interpret: younger users concentrate on education and skill-building, older users on finance, professional tasks, and civic topics. This is consistent across all four countries despite their different age profiles, and it is exactly what one would expect from a life-cycle model of technology use.

6.5 A Digital Sophistication Gradient

The raw-metadata descriptives expose an access-and-sophistication gradient that compounds existing inequalities. Plus subscribers are concentrated almost entirely in India (116 of 134 in our sample); Nigeria has only 1 Plus user and Pakistan only 3. The behavioural gap between Free and Plus users (deeper conversations, longer prompts, more feature use) suggests that this access difference translates into qualitatively different AI experiences. At $20/month, ChatGPT Plus costs roughly 2–5% of median monthly income in Nigeria and Pakistan; the near-zero adoption rate is unsurprising but worth stating plainly, because it means that the most capable version of the most widely used AI tool is effectively inaccessible in two of our four countries.

Advanced-feature adoption is modest even among users who could access them: web search in 12–16% of conversations, code interpreter in 3–4%, image generation below 1%. Whether this reflects limited awareness, limited perceived relevance, or interface friction is an open question with direct product-design implications.

6.6 Limitations

Several limitations qualify our findings. First, our sample is recruited through Clickworker, which introduces selection bias toward digitally literate, English-proficient, gig-economy-adjacent individuals who are not representative of the broader ChatGPT user base in each country. The prominence of *online earning* themes, for instance, is likely amplified by this channel, though the fact that the same theme does *not* appear in Nigeria (recruited through the same channel) provides some evidence that it is not purely a sampling artifact. Second, while our dataset is large by the standards of demographically grounded studies (n = 1,252 users, >200K conversations), it is orders of magnitude smaller than server-side analyses, limiting statistical power for fine-grained subgroup comparisons. Third, some additional conversation-level labels (multitasking, human-alone, user and AI education levels) were collected only for the Indian subsample; a cross-country comparison on those dimensions awaits future work. Fourth, our topic classifiers are themselves LLM-based and may introduce systematic biases for non-English content or culturally specific topics that the classifier was not trained on. Fifth, our temporal coverage coincides with a period of rapid model improvement and feature expansion (the release of GPT-4o, reasoning models, improved tool use), making it difficult to fully disentangle user behavioural evolution from changes in system capabilities. Finally, we observe behaviour but not outcomes: we can see that users ask health questions, but we cannot assess whether the answers they receive are accurate or helpful.

6.7 Future Directions

This study is a proof of concept. The methodology (recruiting users through a crowdsourcing platform, collecting ChatGPT exports with consent, applying multiple classification pipelines on a common corpus) is straightforward to scale. A tenfold expansion of the sample, covering additional countries in Sub-Saharan Africa, Southeast Asia, and Latin America, would enable the kind of fine-grained demographic and geographic comparisons that our current sample cannot support. We see several specific directions that are feasible with this methodology but beyond the scope of the current paper:

- **Health-information quality.** Given the volume of health queries, evaluating the medical accuracy and safety of the responses users receive (perhaps by expert annotation of a stratified subsample) would address one of the most policy-relevant questions our data raises.
- **Longitudinal tracking.** Repeated exports from the same users over time would enable within-person analysis of how usage evolves, complementing the cross-sectional patterns we report here.
- **Multilingual analysis.** Our pipeline operates primarily on English text; dedicated analysis of Urdu, Portuguese, Hausa, and Hindi conversations would capture code-switching patterns and language-specific usage that our current approach likely under-represents.
- **Gender gaps and AI equity.** The gendered usage patterns we observe are a concrete instance of the broader question posed by Otis et al. [22]: does generative AI narrow or widen existing inequalities? A larger, more demographically balanced sample would allow this question to be addressed with statistical rigor rather than suggestive pattern-matching.
- **Community infrastructure.** We release classifier prompts, aggregate tables, and anonymized metadata to support independent re-analysis. We hope this paper demonstrates the value of conversation-level, demographically grounded data on AI usage, and we encourage the research community to build on this methodology: through expanded data collection, alternative classification schemes, or cross-disciplinary analyses that connect usage patterns to outcomes in health, education, and economic welfare.

7 Conclusion

We analyzed 202,590 ChatGPT conversations from 1,252 users across India, Nigeria, Brazil, and Pakistan, to our knowledge the first conversation-level, demographically grounded comparison of LLM usage across multiple Global South markets on a common classification pipeline. Four findings stand out: (1) users in our four countries over-index on information-seeking and writing relative to global averages, with substantial per-country heterogeneity; (2) unsupervised topic discovery surfaces culturally specific use cases (health, religion, translation, digital entrepreneurship, emotional self-reflection) that a fixed taxonomy absorbs into generic buckets; (3) *Doing* has risen steadily while a quiet upward drift in *Expressing* signals an evolving relationship between users and conversational AI; and (4) personal use is the majority everywhere, coursework rivals work in prevalence, and female coursework engagement is markedly higher in India, Pakistan, and Nigeria.

As we argue in Section 6, the broadest implication is that workplace productivity, the dominant lens for evaluating AI's economic impact, captures only a minority of what these users actually do with the technology. Health advice, translation, educational support, and everyday problem-solving are forms of household production whose returns are real but invisible to standard metrics. Product design calibrated to platform-wide averages risks under-serving these use cases, and pricing that places advanced features behind a $20/month paywall makes the most capable version of the tool effectively inaccessible in markets like Nigeria and Pakistan.

This study is a proof of concept: the methodology is straightforward to scale, and we release classifier prompts, aggregate tables, and anonymized metadata for replication. The questions it raises (whether AI-mediated health information helps or harms, whether gendered usage patterns reinforce or compensate for existing inequalities, what the growth of *Expressing* means for human–AI relationships) are larger than any single paper. We hope this work motivates expanded, community-driven data collection across a broader set of countries and populations.

References

[1] Jenny C. Aker and Isaac M. Mbiti. 2010. Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives* 24, 3 (2010), 207–232.

[2] Nicholas O. Alozie and Patience Akpan-Obong. 2011. The Digital Gender Divide: Confronting Obstacles to Women's Development in Africa. *Development Policy Review* 29, 2 (2011), 161–179.

[3] Anthropic. 2025. The Anthropic Economic Index. (2025). https://www.anthropic.com/research/the-anthropic-economic-index

[4] Anthropic. 2025. *The Anthropic Economic Index: Geographic Variation in Claude Usage*. Technical Report. Anthropic. https://www.anthropic.com/research/anthropic-economic-index

[5] Amy Antonio and David Tuffley. 2014. The Gender Digital Divide in Developing Countries. *Future Internet* 6, 4 (2014), 673–687.

[6] Payal Arora and Aashish Rathi. 2023. The Value of AI in the Global South: Perspectives from Indian Workplaces. *AI & Society* (2023).

[7] David Baidoo-Anu and Leticia Owusu Ansah. 2023. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI* 7, 1 (2023), 52–62.

[8] Gary S. Becker. 1965. A Theory of the Allocation of Time. *The Economic Journal* 75, 299 (1965), 493–517.

[9] Aaron Chatterji, Tom Cunningham, David Deming, Zoë Hitzig, Christopher Ong, Carl Shan, and Kevin Wadman. 2025. How People Use ChatGPT. (2025). https://cdn.openai.com/pdf/a253471f-8260-40c6-a2cc-aa93fe9f142e/economic-research-chatgpt-usage-paper.pdf

[10] Aakash Chinchure, Shail Patel, Priyanshu Jain, and Apoorv Agrawal. 2023. Demographic Patterns in AI Chatbot Usage. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

[11] Jiahong Deng and Yuanyuan Lin. 2023. The Benefits and Challenges of ChatGPT: An Overview. *Frontiers in Computing and Intelligent Systems* 2, 2 (2023), 81–83.

[12] Jonathan Donner. 2015. *After Access: Inclusion, Development, and a More Mobile Internet*. MIT Press.

[13] Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. *arXiv preprint arXiv:2303.10130* (2023).

[14] Cathy Mengying Fang, Sheer Karny, Chayapatr Archiwaranguprok, Yasith Samaradivakara, Pat Pataranutaporn, and Pattie Maes. 2026. AI-Wrapped: Participatory, Privacy-Preserving Measurement of Longitudinal LLM Use In-the-Wild. In *Proceedings of CHI*. https://arxiv.org/abs/2602.18415

[15] Alison Gillwald, Onkokame Mothobi, and Broc Rademan. 2019. The State of ICT in Africa. *Research ICT Africa* (2019).

[16] Maarten Grootendorst. 2022. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv preprint arXiv:2203.05794* (2022).

[17] Martin Hilbert. 2011. Digital Gender Divide or Technologically Empowered Women in Developing Countries? A Typical Case of Lies, Damned Lies, and Statistics. *Women's Studies International Forum* 34, 6 (2011), 479–489.

[18] Pratik Joshi, Sebastin Santy, Amar Buber, Kalika Bali, and Monojit Choudhury. 2020. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In *Proceedings of ACL*. 6282–6293.

[19] Sai Keerthana Karnam, Abhisek Dash, Krishna P. Gummadi, Animesh Mukherjee, Ingmar Weber, and Savvas Zannettou. 2026. Bowling with ChatGPT: On the Evolving User Interactions with Conversational AI Systems. In *Proceedings of The Web Conference (WWW)*. https://arxiv.org/abs/2602.01114

[20] Enkelejda Kasneci, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences* 103 (2023), 102274.

[21] OpenAI. 2026. *How India Uses ChatGPT: Early Evidence from OpenAI Signals*. Technical Report. OpenAI. https://cdn.openai.com/signals/how-india-uses-chatgpt.pdf

[22] Nicholas G Otis, Solene Delecourt, Katelyn Cranney, and Rembrand Koning. 2024. *Global evidence on gender gaps and generative AI*. Harvard Business School Boston, MA, USA.

[23] Siru Ouyang, Shuohang Zhang, Meiqi Zhao, Yelong Gao, and Jiawei Chen. 2023. Shifted Diffusion for Text-to-Image Generation. *arXiv preprint arXiv:2306.02115* (2023).

[24] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M. Aroyo. 2021. "Everyone Wants to Do the Model Work, Not the Data Work": Data Cascades in High-Stakes AI. In *Proceedings of CHI*. 1–15.

[25] Sherry Turkle. 2017. *Alone Together: Why We Expect More from Technology and Less from Each Other* (3rd ed.). Basic Books.

[26] Jan A. G. M. van Dijk. 2020. The Digital Divide. (2020).

[27] Mark Warschauer. 2004. *Technology and Social Inclusion: Rethinking the Digital Divide*. MIT Press.

[28] Mark Warschauer, Wavell Tseng, Soobin Yim, Tamara Webster, Sharin Jacob, Qian Du, and Tammy Tate. 2023. The Affordances and Contradictions of AI-Generated Text for Second Language Writers. *Journal of Second Language Writing* 62 (2023), 100894.

[29] Lixiang Yan, Lele Sha, Linxuan Zhao, Yuheng Li, Roberto Martinez-Maldonado, Guanliang Chen, Xinyu Li, Yueqiao Jin, and Dragan Gasevic. 2024. Practical and Ethical Challenges of Large Language Models in Education: A Systematic Scoping Review. *British Journal of Educational Technology* 55, 1 (2024), 90–112.

[30] Wenting Zhao, Xiang Ren, Jack Hessel, Claire Cardie, Yejin Choi, and Yuntian Deng. 2024. WildChat: 1M ChatGPT Interaction Logs in the Wild. In *Proceedings of ICLR*. https://arxiv.org/abs/2405.01470

[31] Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. 2023. Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. In *Advances in Neural Information Processing Systems (NeurIPS)*.

A BERTopic Pipeline Details

This appendix documents the full unsupervised topic-modelling pipeline summarized in Section 4.

Embedding. Each conversation is represented by a dense semantic embedding obtained from Google's `gemini-embedding-001` model, which produces 3,072-dimensional vectors. The embedding input is the first ten user–assistant turns of the conversation, serialized in the `[User]: ...\n[Assistant]: ...` format, truncated at a per-message cap of 5,000 characters. Using the first ten turns rather than the whole conversation keeps per-conversation cost bounded and, in practice, captures the topic of the conversation well (users typically establish their task in the opening exchanges).

Initial clustering. For each country we fit MiniBatch K-means with $k = 500$ on the embedding matrix. The deliberately high k produces narrow, internally coherent micro-clusters (e.g. "debugging Python code for web scraping" rather than the generic "programming"). Starting from an over-segmented solution lets the subsequent aggregation be driven by semantic similarity rather than forced early generalization.

Hierarchical agglomeration. The 500 K-means centroids are then merged with agglomerative hierarchical clustering using cosine distance. The cutoff is selected per country by visual inspection of the dendrogram for natural breakpoints, targeting approximately 50 interpretable top-level topics, a number that balances cross-country comparability with sufficient granularity to reveal local patterns.

Topic labelling and refinement. Each resulting cluster is labelled by summarization with `gpt-4o-mini`, using the top documents and most distinctive keywords in the cluster as context. To enhance label validity, we used Claude Sonnet 4.6 as a judge: clusters with ambiguous or low-confidence labels were either relabelled or flagged as non-homogeneous. Conversations in flagged clusters were then reassigned to defined topics by cosine-similarity scoring against cluster centroids, using a per-cluster z-score threshold of 1.5 as the reassignment floor. The final pipeline yields **50**, **45**, **36**, and **53** topics for India, Nigeria, Brazil, and Pakistan respectively, covering 91–95% of conversations per country; the remaining unassigned conversations (typically very short, off-topic, or multilingual/code-heavy) are excluded from BERTopic analyses but retained elsewhere.

Cross-country theme grouping. For the ten-theme cross-country view in Section 5.2, we group the per-country clusters into ten broad themes (Programming/Tech, Finance/Earning, Writing/Creative, Religion, Translation/Language, Health/Medical, Education/Academic, Job/Career, Content Creation, Current Affairs) by keyword matching on the cluster labels. This is a lossy aggregation (some cluster labels match multiple themes) but it makes the four countries comparable on a shared vocabulary while retaining the per-country cluster resolution for the cluster-level figures.

B Cross-Dataset Validation: WildChat

To assess the generalizability of our findings beyond our recruited cohort and to compare with global trends, we analyze conversation topics in WildChat-4.8M [30], an independent corpus of ChatGPT interactions collected via a Gradio-based research proxy. WildChat provides geographic metadata (derived from IP addresses) for a large, diverse user base. From the top 20 countries by volume we situate our India and Brazil findings within a broader global comparison.

B.1 Sampling and Classification

We selected the top 20 countries by total conversation volume in WildChat (Table 5). To ensure sufficient per-user coverage, we sampled users with 10 or more conversations; for the five largest (United States, Russia, China, Germany,

Table 5. WildChat sample overview. Top 20 countries by total conversation volume, before and after sampling users with 10+ conversations. For United States, Russia, China, Germany, and Hong Kong we further sampled 500 users each.

Country	Total Conversations	Sampled Users	Sampled Conversations
United States	944,046	500	36,394
Russia	262,050	500	36,770
China	253,752	500	14,296
Germany	114,837	500	34,507
United Kingdom	103,614	445	43,936
Japan	98,810	446	12,924
Vietnam	75,220	258	59,186
Hong Kong	70,225	500	14,526
India	62,804	577	17,315
France	62,309	483	19,430
South Korea	61,894	231	15,697
Brazil	57,702	427	16,033
Canada	51,265	320	18,228
Taiwan	41,691	409	20,256
Australia	37,080	259	12,707
Italy	36,827	262	9,177
Spain	33,019	240	10,178
The Netherlands	32,578	207	9,048
Singapore	30,899	375	11,567
Egypt	27,708	219	4,640

Hong Kong), we further subsampled 500 users each. This yielded 398,050 classified conversations across 20 countries. We applied the same OpenAI taxonomy classifier (GPT-4o, 24 fine-grained categories, 7 coarse domains) used for our primary dataset, enabling direct comparison.

A key difference between the two datasets is their user populations. WildChat users self-selected into a research proxy to access ChatGPT for free, creating a sample that skews toward technically sophisticated, often English-speaking users, particularly in non-Anglophone countries. Our primary dataset, by contrast, represents a broader demographic cross-section through crowdsourced recruitment.

B.2 Coarse Topic Distribution Across Countries

Figure 22 presents the coarse topic composition across all 20 WildChat countries. Several patterns are immediately apparent. *Writing* dominates in Anglophone countries, reaching 66.0% in the United Kingdom and 46.3% in the United States, driven primarily by fiction writing, which alone accounts for 40.2% and 18.5% of conversations in these countries respectively. *Multimedia* (primarily image generation) is concentrated in East Asian countries: Japan (49.8%), Hong Kong (43.2%), and China (41.9%), reflecting the popularity of AI-generated art and character illustration in these markets. *Technical Help* varies from under 1% (Vietnam) to 37.3% (India), the single largest source of cross-country variation.

B.3 India and Brazil in Global Context

Figure 23 shows the coarse topic profiles of India and Brazil alongside four high-volume comparison countries (United States, United Kingdom, Germany, France), and Figure 24 quantifies their deviation from the 20-country mean.

India in WildChat is strikingly dominated by Technical Help (37.3%, +20.5pp above the 20-country mean), driven almost entirely by computer programming at 34.4% of all Indian conversations, ranking first among all 20 countries by a

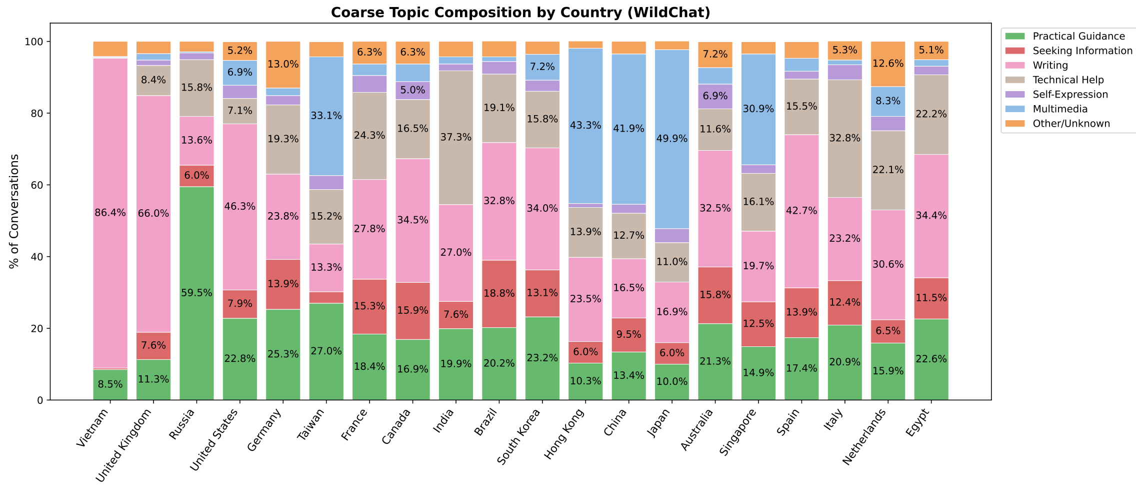


Fig. 22. Coarse topic composition across 20 WildChat countries, sorted by total conversations.

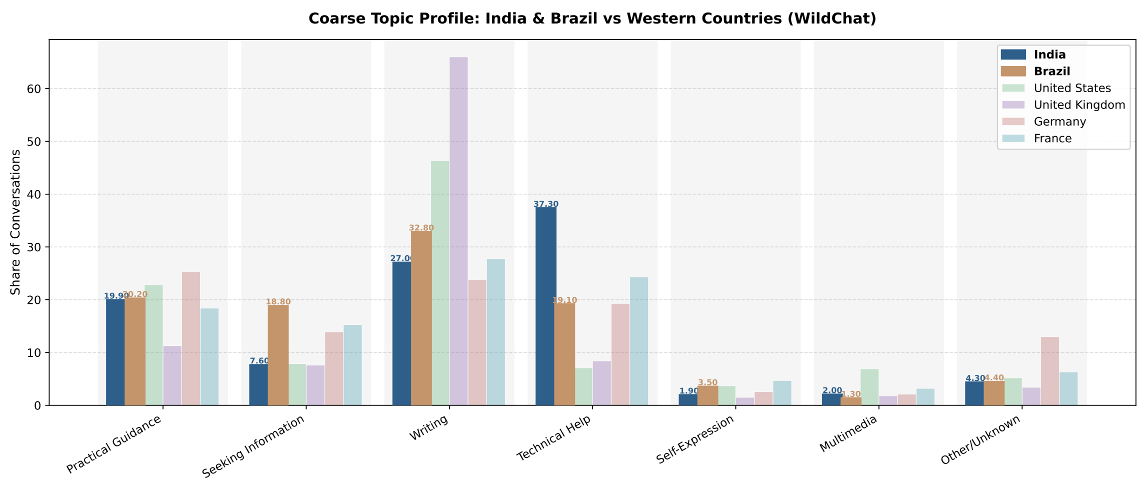


Fig. 23. Coarse topic profiles of India and Brazil compared to four high-volume comparison countries (WildChat).

wide margin. This reflects the composition of Indian WildChat users: technically proficient individuals who discovered and adopted a research proxy interface, likely drawn disproportionately from India's software engineering workforce. Writing is correspondingly lower (−5.3pp), and Multimedia usage is minimal (2.0%, −10.4pp). Edit or critique of provided text ranks second (13.0%, rank 3/20).

Brazil presents a markedly different WildChat profile. *Seeking Information* is the most over-represented category (+8.6pp above the 20-country mean), with *specific information* queries constituting 18.0% of conversations, the highest share among all 20 countries. Computer programming is substantial but less dominant (17.3%, rank 6/20). Translation is notably elevated (6.4%, rank 6/20), reflecting Portuguese–English bilingual needs.

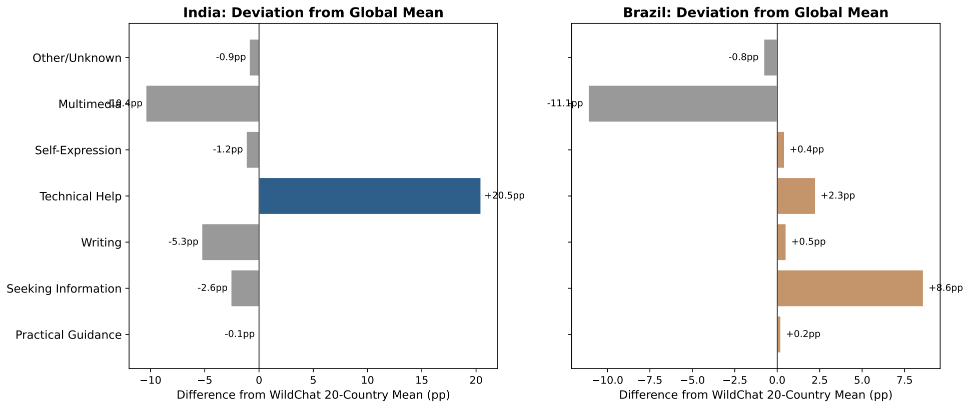


Fig. 24. Deviation of India (left) and Brazil (right) from the WildChat 20-country mean in coarse topic share (percentage points).

B.4 Comparing WildChat with Our Dataset

Figure 25 compares the WildChat topic distributions for India and Brazil with our ChatGPT-export data.

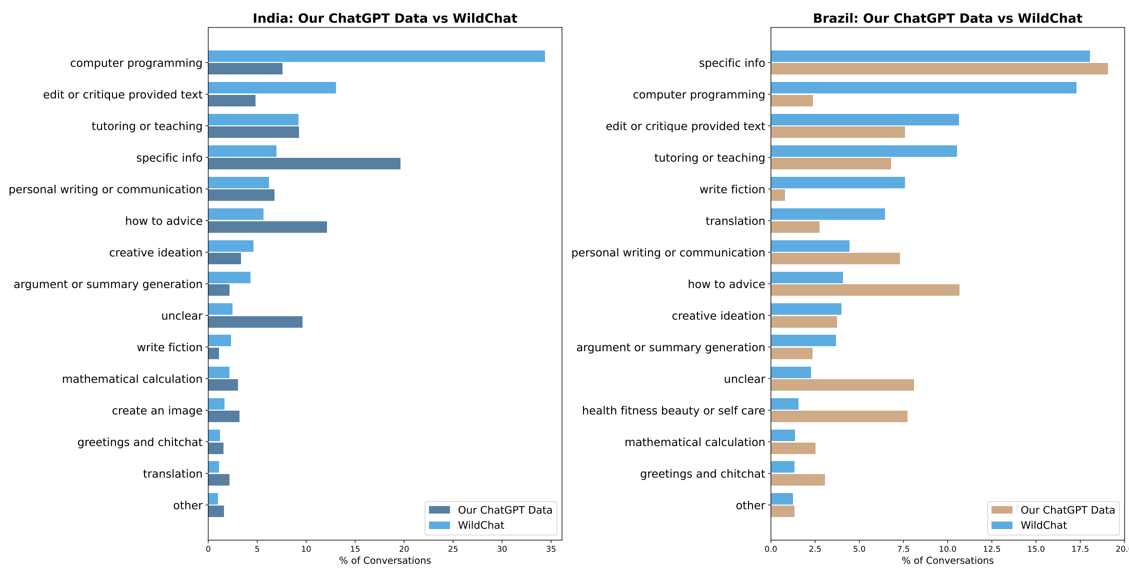


Fig. 25. Fine-grained topic comparison between our ChatGPT-export data and WildChat for India (left) and Brazil (right).

Divergences. The most striking difference is the magnitude of computer programming: 34.4% in WildChat India vs. 7.6% in our data, and 17.3% in WildChat Brazil vs. 2.4% in ours. This 4–7× inflation directly reflects WildChat's selection bias toward technically sophisticated users who navigate a research proxy interface. Conversely, our data shows substantially higher shares of *specific information* seeking (19.6% vs. 7.0% for India; 19.1% vs. 18.0% for Brazil), *how-to advice*, and *health/fitness/beauty*, categories that characterize broader, non-technical, personal conversations. These divergences validate a methodological concern: proxy-based corpora like WildChat systematically over-represent technical users and under-represent the information-seeking and practical guidance use cases that dominate among general populations.

Convergences. Despite these differences, several patterns are robust across both datasets. First, India leads in computer programming share in both WildChat (rank 1/20) and our dataset (7.6%, highest among our four countries), consistent with India's IT-sector-driven usage profile regardless of sampling frame. Second, Brazil consistently shows higher engagement with translation tasks (6.4% in WildChat vs. 2.8% in our data). Third, creative fiction writing is comparatively low in both India and Brazil across both datasets (2.3% in WildChat India, even smaller in ours) compared to 18.5% for the US in WildChat, suggesting that creative-fiction use varies substantially by country and language context.

Implications. The comparison underscores that topic distributions in user–AI interaction studies are sensitive to collection methodology and the resulting user sample. WildChat's proxy collection yields a portrait skewed toward users who actively sought out a research frontend, while our recruited cohort captures a broader spectrum of use cases. The convergence of robust findings across these methodologically distinct datasets (India's programming emphasis and the lower salience of creative fiction in both) strengthens confidence that those patterns reflect genuine occupational and linguistic differences rather than artifacts of either method.

C Additional Figures

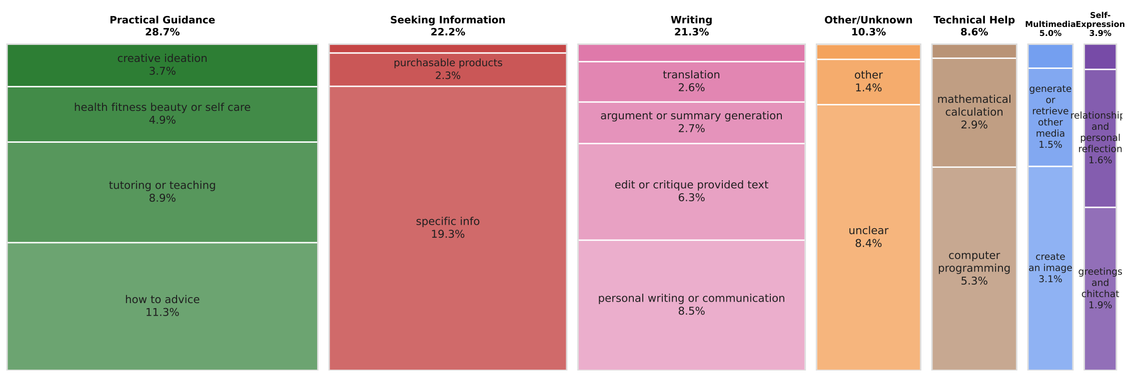


Fig. 26. Treemap of conversation topics across all four countries under the OpenAI taxonomy (Section 5.1). Outer tiles are coarse domains; inner tiles are fine-grained categories.

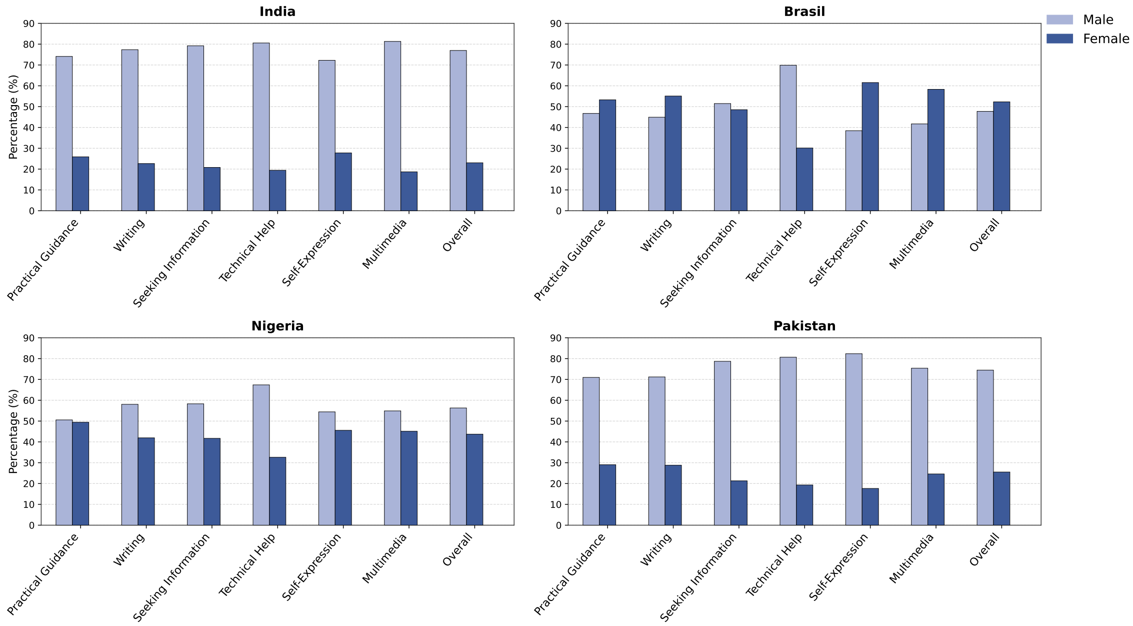


Fig. 27. Share of conversations by gender for each coarse topic, by country (Section 5.1). The "Overall" bar provides the baseline gender split; deviations from it indicate topic-level over- or under-representation.

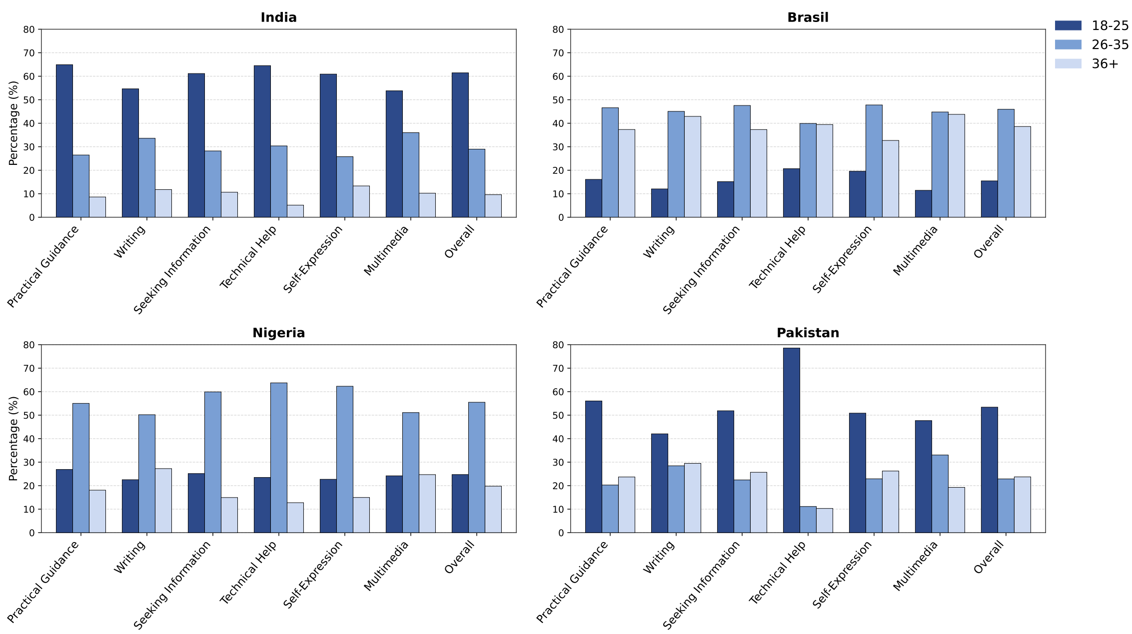


Fig. 28. Share of conversations by age group for each coarse topic, by country (Section 5.1). The "Overall" bar provides the baseline age distribution.

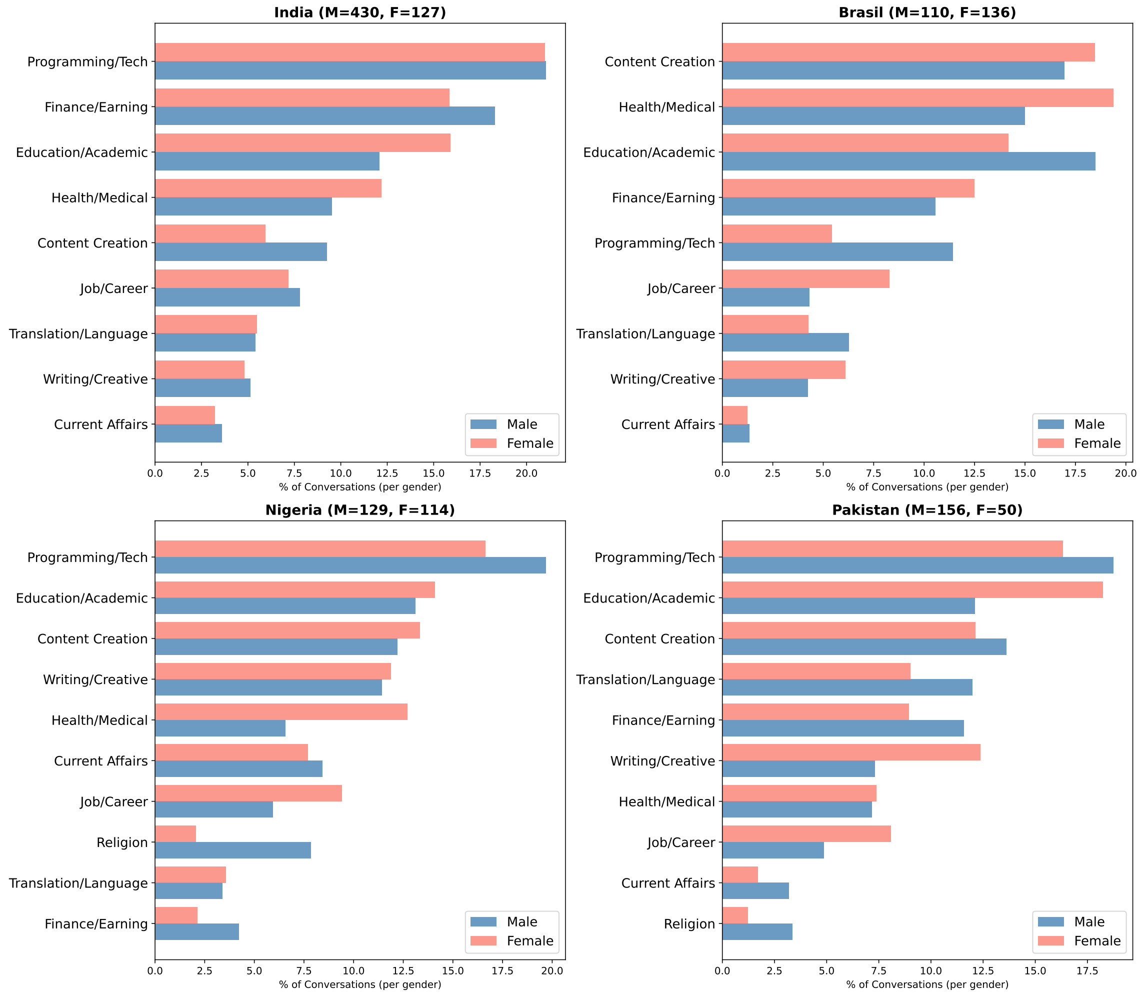


Fig. 29. Unsupervised themes by gender across countries (Section 5.2). Each bar shows the share of a gender's conversations assigned to a theme. Male over-representation in programming, finance, and religion; female over-representation in health, content creation, and education.

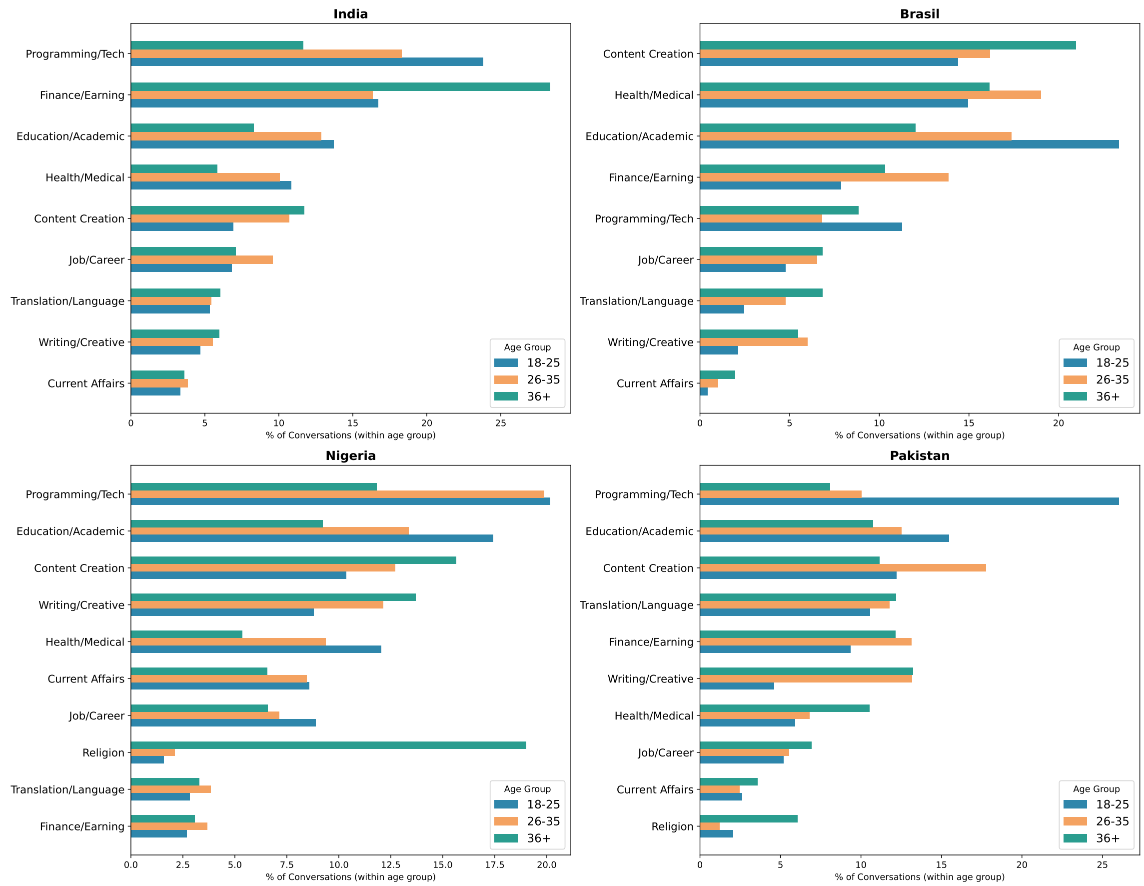


Fig. 30. Unsupervised themes by age group across countries (Section 5.2). Younger users concentrate on education and programming; older users shift toward finance, content creation, and religion.

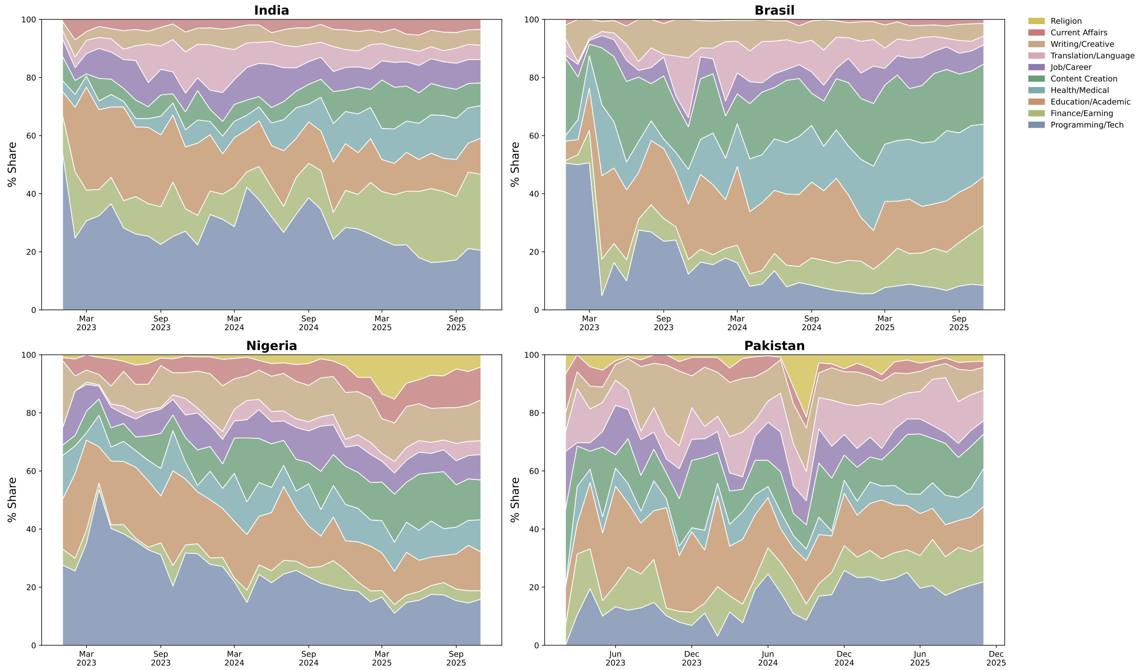


Fig. 31. Temporal shift in unsupervised theme share (stacked area) by country (Section 5.2).

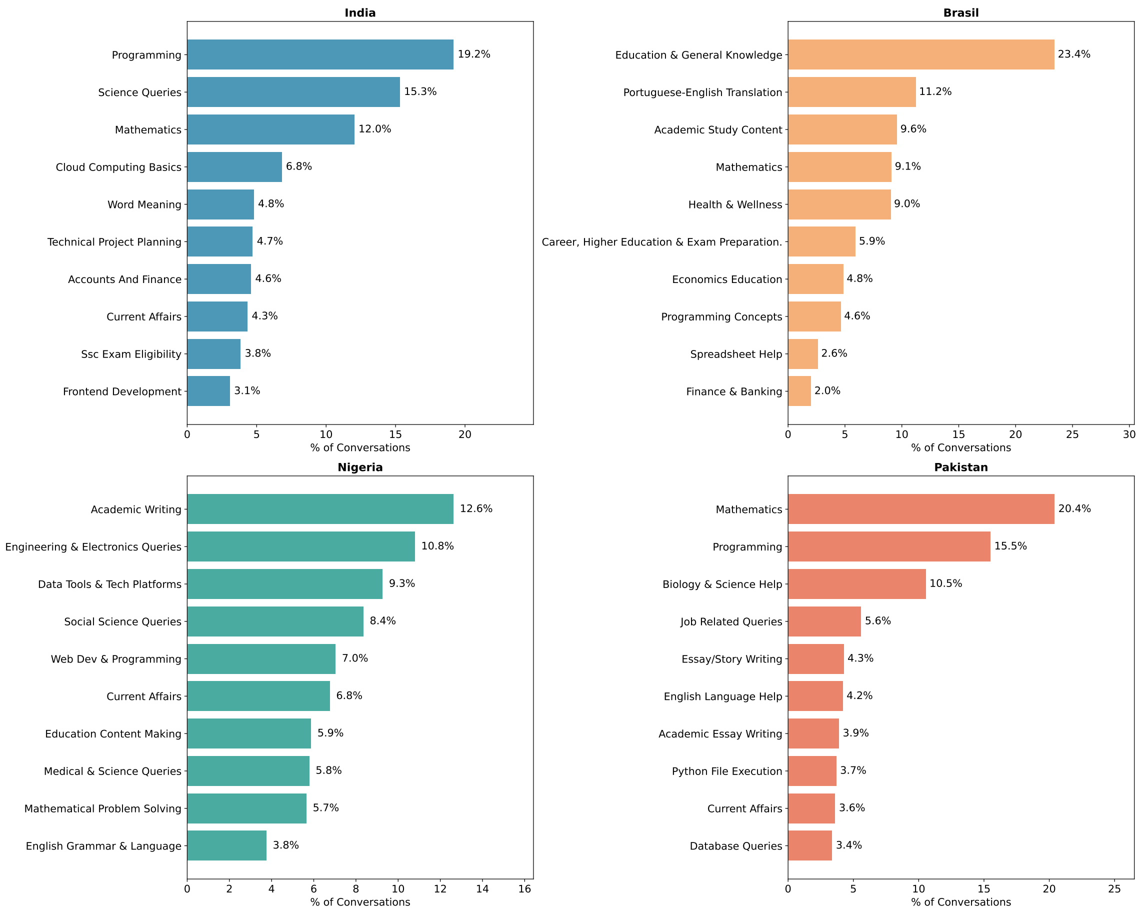


Fig. 32. Top 10 unsupervised topic clusters for coursework conversations, by country (Section 5.3).

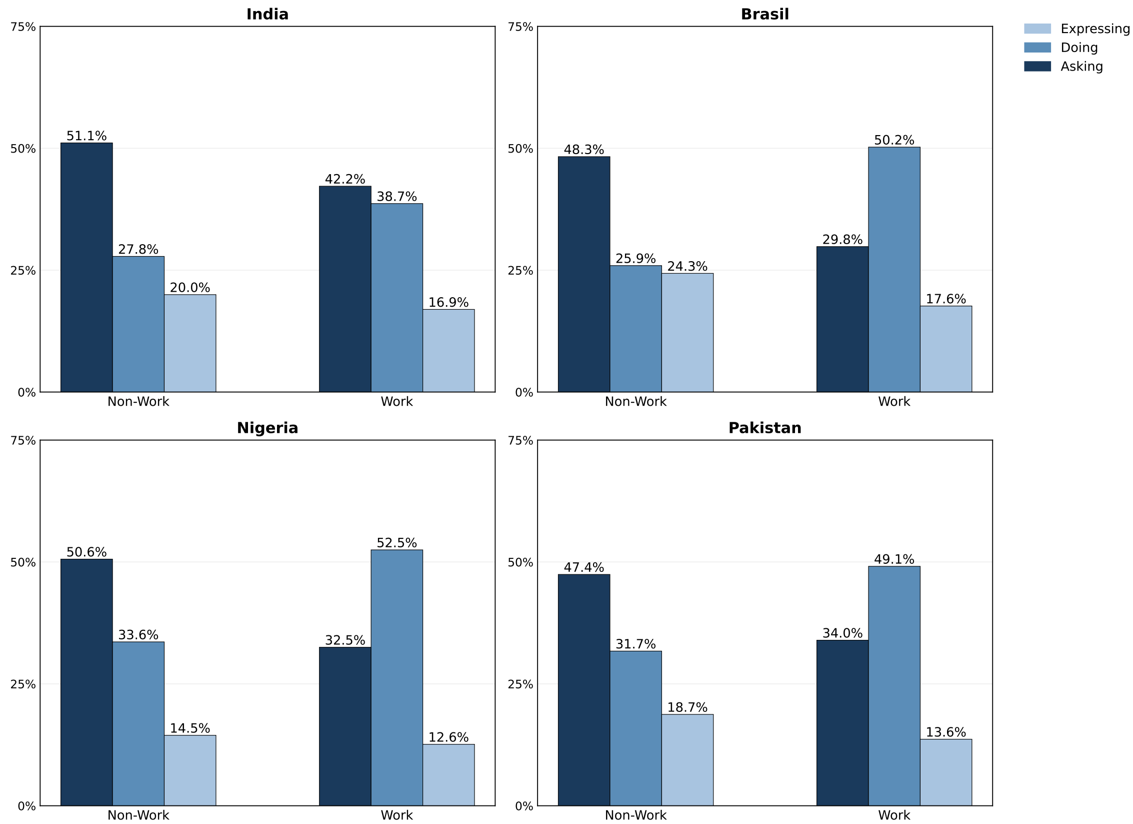


Fig. 33. Share of *Asking*/*Doing*/*Expressing* messages by task purpose (work vs. non-work) and country (Section 5.4).

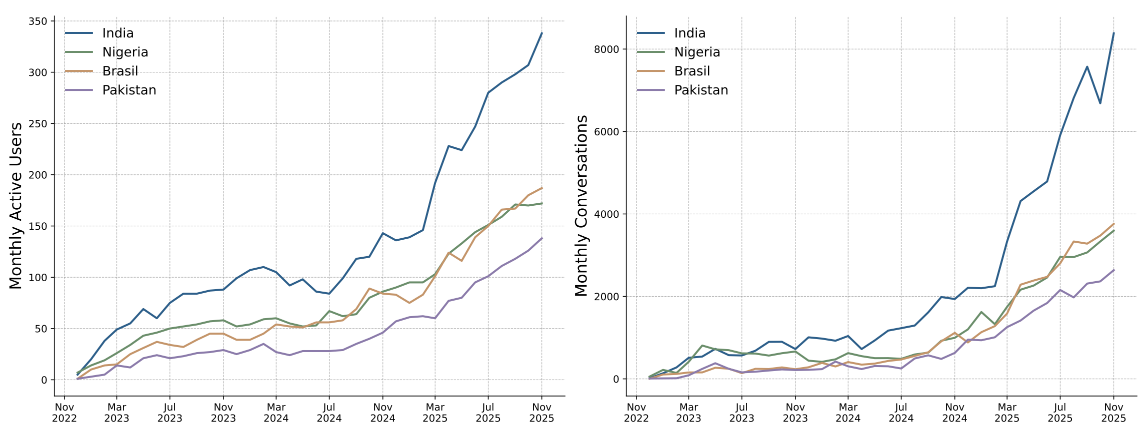


Fig. 34. Monthly active users (left) and monthly conversations (right) by country (Dec 2022–Feb 2026). See Section 3.

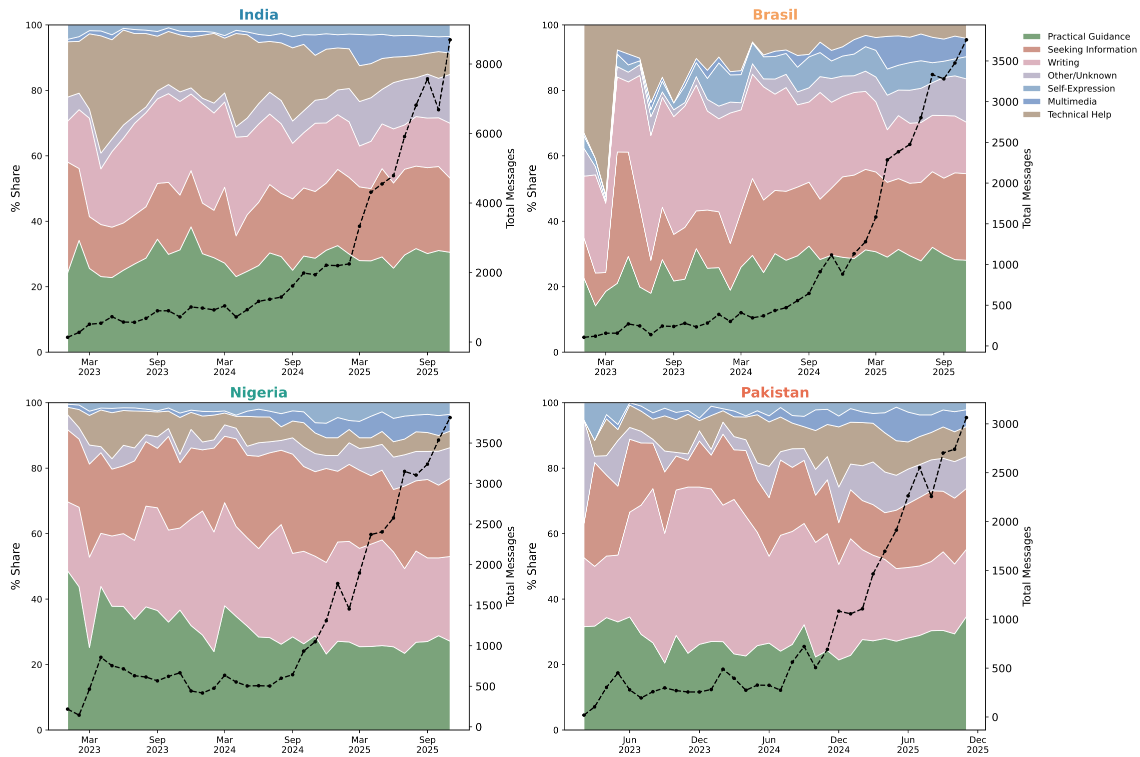


Fig. 35. Monthly coarse-topic share (stacked area, left axis) and total conversation volume (dashed line, right axis) by country, under the OpenAI taxonomy. Deferred from Section 5.1: the in-text discussion uses the line-plot version (Figure 5); this stacked-area rendering is included here for readers who prefer to see volume and share on the same chart.