Towards Proactive Information Probing: Customer Service Chatbots Harvesting Value from Conversation

Chen Huang[1] **Zitan Jiang**[23] **Changyi Zou**[23] **Wenqiang Lei**[23*] **See-Kiong Ng**[1]

[1] Institute of Data Science, National University of Singapore
[2] College of Computer Science, Sichuan University
[3] Engineering Research Center of Machine Learning and Industry Intelligence, Ministry of Education, China

huang_chen@nus.edu.sg, jzitan@stu.scu.edu.cn

Abstract

Customer service chatbots are increasingly expected to serve not merely as reactive support tools for users, but as strategic interfaces for harvesting high-value information and business intelligence. In response, we make three main contributions. 1) We introduce and define a novel task of *Proactive Information Probing*, which optimizes when to probe users for pre-specified target information while minimizing conversation turns and user friction. 2) We propose ProChatIP, a proactive chatbot framework featuring a specialized conversation strategy module trained to master the delicate timing of probes. 3) Experiments demonstrate that ProChatIP significantly outperforms baselines, exhibiting superior capability in both information probing and service quality. We believe that our work effectively redefines the commercial utility of chatbots, positioning them as scalable, cost-effective engines for proactive business intelligence. Our code is available at https://github.com/SCUNLP/PROCHATIP.

Conversation History

User: What's the current price of iron ore today?
Chatbot: For 62% grade iron ore, it's ~$108.35/ton. For 65% grade, it's about $119.6/ton.
User: Why is it still rising? The inventory situation doesn't look good.

Conventional Customer Service Chatbot

Chatbot: Actually, prices have been quite volatile recently.
User: Okay, understood.

Chatbot with Proactive Information Probing

Chatbot: Actually, prices have been quite volatile recently. Sounds like you follow the market. Have you been tracking inventory levels closely? **Prob Info.**
User: Yeah, it feels like there's a serious backlog. Market sentiment is poor, consumption is slow. **Obtain Info.**
Chatbot: Understood. Thank you for sharing that.

Figure 1: Illustration of proactive information probing. The chatbot gathers critical market information during routine service interactions, functioning as a cost-effective way for business intelligence.

1 Introduction

AI chatbots for customer service have long provided a fast and efficient solution for handling user inquiries (Nicolescu and Tudorache, 2022; Adam et al., 2021), often simultaneously tasked with soliciting user feedback to improve service quality (e.g., *Is this answer helpful?*) (Schloß et al., 2024; Mashaabi et al., 2022). In the era of Large Language Models (LLMs), these AI chatbots are evolving from simple reactive tools into proactive assistants capable of steering conversations toward specific objectives (Deng et al., 2025, 2024). This technological leap is fundamentally reshaping the functional role of customer service chatbots, transforming them into strategic bridges for business intelligence: Beyond traditional service assessments, modern chatbots possess the potential to probe and gather critical user-side information (Behr et al., 2012; Oudejans, 2018; Jacobsen et al., 2025). Taking Figure 1 for example, a chatbot for a commodity exchange could proactively query users about market conditions (e.g., *Have you been tracking inventory levels closely?*) to refine its own pricing models via crowdsourcing (Ridzuan et al., 2024; Sarstedt et al., 2014). Therefore, this role shift towards proactive information probing enables companies to harvest high-value business information directly from users in a cost-effective manner, unlocking new value in customer interactions.

In the era of LLMs, while current customer service chatbots can successfully handle diverse user inquiries (Hong et al., 2025b; Jiang et al., 2025), they remain largely reactive, lacking the mechanisms for proactive information probing. To bridge this gap, we propose a novel task: **Proactive In-**

[*] Corresponding author.

Findings of the Association for Computational Linguistics: ACL 2026, pages 28768–28792
July 2-7, 2026 ©2026 Association for Computational Linguistics

formation Probing. Since users prioritize their own inquiries, they are likely to reject or ignore chatbot probes that appear intrusive or contextually irrelevant. While an chatbot can attempt to re-ask, doing so increases dialogue length and potentially harms the user experience (Singh and Rios, 2022). Therefore, the core challenge of our task lies in balancing information probing with user satisfaction: **when to probe users within multi-turn dialogues so as to maximize information gathering while minimizing conversation turns and user friction**.

To this end, we propose PROCHATIP(PROactive CHATbot for Information Probing), a specialized framework designed for the customer service context. To achieve this, as illustrated in Figure 2, PROCHATIP incorporates a dedicated conversation strategy module that explicitly guides the timing of information probes. We train this module using a two-stage curriculum: 1) We apply supervised fine-tuning (SFT) using rule-based synthetic data to establish baseline probing capabilities. 2) We employ reinforcement learning (RL) to further optimize the CS module. In this phase, we introduce a dual-reward mechanism comprising active probing rewards and passive waiting rewards, which collectively account for both the immediate utility of probing and the strategic value of inaction (waiting, not to probe). As such, PROCHATIP could maximize information acquisition while ensuring less dialogue turns and user friction.

We demonstrate our effectiveness using both LLM-based user simulator and real human participants. It outperforms the best baseline by 11.36% improvement in successfully gathering target information, 39.50% lower conversation turns, and 23.73% improvement in overall user satisfactions. PROCHATIP represents a preliminary step in modeling proactive information probing, it marks a pivotal shift from reactive service chatbots to strategic assets capable of balancing user satisfaction with information acquisition objectives. By transforming routine service interactions into scalable, cost-effective channels for high-value information acquisition, PROCHATIP redefines the commercial utility of customer service chatbots as proactive engines for business intelligence. To sum up, we conclude the following main contributions.

- We propose and formalize the novel task of Proactive Information Probing, shifting the paradigm of customer service chatbots from reactive responders to proactive agents that gather high-value business intelligence.
- We introduce PROCHATIP, a framework featuring a specialized conversation strategy module for optimizing the strategic timing of probing.
- We demonstrate that PROCHATIP significantly outperforms baselines in information probing while maintaining higher user satisfaction and minimizing dialogue turns.

2 Related Work

We survey existing work on customer service chatbots and proactive conversational agents, highlighting our difference.

Customer Service Chatbots. Customer service chatbots have evolved into an integral component of modern business operations (Cui et al., 2017; De Keyser et al., 2019; Taranukhin et al., 2024; Hsu et al., 2024). Typically, existing research focuses on enhancing query resolution, including customer intent identification (Hong et al., 2025a; Vasquez-Correa et al., 2021), and the extraction of question-answer pairs (Zheng et al., 2023; Yang et al., 2023; Hong et al., 2025c) or expert strategies (Jiang et al., 2025) from historical chat logs. Recent advancements have significantly expanded the functional capabilities of customer service chatbots to handle complex interaction dynamics, including empathetic chatting (Song et al., 2021; Agarwal et al., 2021), maintaining professional role consistency (Li et al., 2025), and managing conversation transitions (e.g., determining when and to whom a session should be transferred) (Yu et al., 2020). In this paper, we advance these chatbots further by transforming these routine service interactions into scalable, cost-effective channels for high-value information acquisition.

Proactive Conversational Agents. These agents are distinguished by their ability to take initiative, steering dialogues toward productive outcomes rather than passively following a user's lead (Deng et al., 2024, 2025). Such agents are particularly effective in tasks like conversation information seeking and conversational recommendation (Deng et al., 2023b). To date, research has primarily focused on employing proactivity to enhance the quality of user assistance. This is typically achieved through strategies such as clarifying query ambiguity (Zhang et al., 2024; Chen et al., 2024), eliciting user preferences for better recommendations (Zhang et al., 2018, 2022; Chen et al., 2023), or

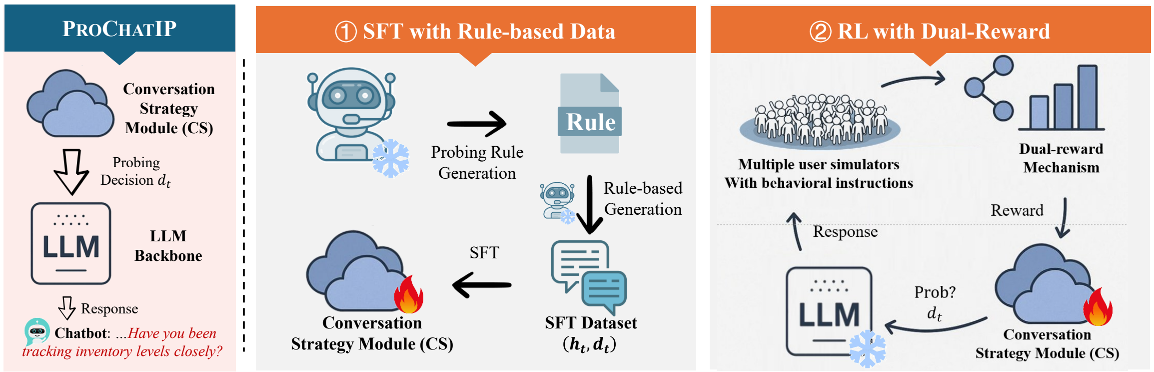

Figure 2: Overview of PROCHATIP. It incorporates a conversation strategy module (CS) that explicitly guides the timing of information probe d_t. Here, the CS module is trained via two-stage curriculum.

managing over-specified requests (Wu et al., 2023; Min et al., 2019). However, these approaches operate under a user-centric utility model, using proactivity solely to refine the system's response to the user. In contrast, our work introduces a system-centric utility dimension, leveraging proactivity to harvest external business intelligence (e.g., market data). This shifts the paradigm to a multi-objective challenge: balancing user satisfaction with the system's goal of high-value information acquisition.

3 Task Formulation

We define and formalize the task of *Proactive Information Probing*: We model the customer service interaction as a sequential decision-making process where the chatbot must balance two competing objectives: fulfilling the user's queries Q and effectively probing for pre-specified target information $\mathcal{I}$ (e.g., market inventory levels). At each turn t, the system must make a strategic decision $d_t \in \{0, 1\}$ regarding whether to execute a probe ($d_t = 1$) or simply address the user's immediate query ($d_t = 0$). To govern this behavior, PROCHATIP incorporates a dedicated conversation strategy module (CS) to determine the optimal value of d_t at each step.

Formally, let $\mathcal{D} = \{(u_1, r_1), \ldots, (u_T, r_T)\}$ denote a multi-turn dialogue session of length T, where u_t represents the user's utterance or query at turn t, and r_t represents the chatbot's response or probing question. We define the dialogue history at turn t as $h_t = \{u_1, r_1, \ldots, u_{t-1}, r_{t-1}, u_t\}$. Under this framework, the generation of response r_t is conditioned on the history, the target information, and the strategic decision:

$$r_t = \mathrm{LLM}(h_t, \mathcal{I}, d_t), \tag{1}$$

where the strategic decision is predicted by $d_t = \mathrm{CS}(h_t, \mathcal{I})$. The success of a probe is verified by analyzing the subsequent user utterance u_{t+1}. We define a reward function $R(h_t, d_t, u_{t+1}, \mathcal{I})$, which assigns a scalar reward at dialogue turn t (we instantiated it in Section 4.2.)

Finally, the core challenge of Proactive Information Probing lies in optimizing the timing sequence $D = \{d_1, \ldots, d_T\}$ in multi-turn dialogues, balancing effective information acquisition against interaction efficiency and user receptiveness. Accordingly, we formulate the optimization objective as the expected discounted return:

$$J = \max \mathbf{E}\left[\sum_{t=1}^{T} \gamma^{t-1} R(h_t, d_t, u_{t+1}, \mathcal{I})\right], \tag{2}$$

where γ is the discount factor. The goal of PROCHATIP is to learn a conversation strategy module that maximizes the cumulative discounted reward J by selecting appropriate probing decisions throughout the dialogue.

4 PROCHATIP

Overview. PROCHATIP equips the LLM with an external conversation strategy module CS to govern probing timing, as illustrated in Figure 2. This module outputs a decision signal d_t to the LLM at turn t, thereby decoupling probe intent from response generation. This separation ensures that the LLM's primary utility as a customer service chatbot is preserved, allowing it to address user inquiries effectively while selectively incorporating probes. In particular, the architecture of PROCHATIP consists of two main parts, as described below. Following this overview, we detail the two-stage training curriculum for the CS module.

- *Conversation Strategy (CS)*: A lightweight decision-making module that analyzes the current dialogue history h_t and target information $\mathcal{I}$ to output a binary decision d_t, determining *when* to probe. We implement the CS module using Qwen3-0.6B (Yang et al., 2025), which balances lightweight training and deployment costs with context understanding capabilities.

- *LLM Backbone*: This module functions as the customer service chatbot[1]. It synthesizes the response r_t based on history h_t, target $\mathcal{I}$, and decision d_t. If signaled to probe ($d_t = 1$), the LLM integrates a natural language inquiry into its output; otherwise ($d_t = 0$), it generates a pure service response.

4.1 Acquiring Probing Capability via SFT

We first equip the CS module with the fundamental capability to identify appropriate probing opportunities based on contextual cues. To achieve this, we construct a synthetic customer service dataset using rule-based heuristics, where the chatbot alternates between answering queries and probing for information. We then utilize Supervised Fine-Tuning (SFT) to train the CS module on this data, establishing a baseline probing capability.

Rule-based Data Generation. Current customer service datasets typically lack instances of proactive probing. To this end, we construct a synthetic corpus $\mathcal{D}_{\text{sft}}$ using LLMs. Specifically, we first generate a set of governing rules that define appropriate (e.g., user shows willingness to shift or expand the topic) and inappropriate (e.g., user explicitly refuses or avoids answering) probing scenarios. Subsequently, we feed these rules back into the LLM to synthesize diverse multi-turn dialogue sessions[2], explicitly prompting the model to ensure that each conversation corresponds to a single specific rule (either appropriate or inappropriate). Note that we also prompt the LLM to ensure that the synthesized conversations are content-rich, encompassing a diverse range of dialogue scenarios. Formally, each sample in $\mathcal{D}_{\text{sft}}$ is represented as a tuple (h_t, d_t), where h_t denotes the dialogue history up to turn t and $d_t \in \{0, 1\}$ is a binary label indicating whether a probe should be initiated in the current turn.

Supervised Training. We perform full-parameter fine-tuning on the CS module using $\mathcal{D}_{\text{sft}}$. It minimizes the cross-entropy loss between the predicted probing decision $\hat{d}_t$ and the ground truth d_i. This phase ensures the CS module acquires the fundamental heuristic patterns required to identify contextually viable probing opportunities.

[1]In practical deployments, this LLM could be further augmented via RAG. However, as our focus lies on the probing strategy, we employ a standard LLM setup for simplicity.

[2]The LLM directly outputs the full conversation transcript, avoiding the complexity of interactive role-play.

4.2 Optimizing Probing Strategy via RL

Following the initial capability acquisition via SFT, we further optimize the probing strategy of the CS module in a dynamic interaction environment using Reinforcement Learning (RL). This stage enables ProChatIP to transcend simple heuristic imitation and master the strategic timing required to maximize information gathering while minimizing user friction and conversation turns.

Environment/User Simulator. We establish a dynamic interaction environment by creating a set of user simulators, denoted as $\mathcal{S} = \{s_1, s_2, \ldots, s_M\}$, to interact with ProChatIP. Each user simulator s_i is paired with a user query. To facilitate multi-turn interactions, we instruct the simulator to reveal its query gradually across multiple turns rather than presenting the full query immediately. Additionally, to accurately simulate user resistance, we curate a set of behavioral instructions $\mathcal{K}$ via LLM generation. These instructions dictate specific conditions under which the simulator should refuse to answer or lower its satisfaction score (e.g., *if interrupted during a complaint*). The interaction session concludes when the chatbot successfully collects the target information or reaches the maximum number of dialogue turns. During each training iteration, ProChatIP is tasked with collecting a new target information $\mathcal{I}_i$. We randomly sample a simulator from $\mathcal{S}$ and inject a random subset of instructions from $\mathcal{K}$ to steer its responses. This randomized configuration exposes ProChatIP to varied resistance patterns, fostering a robust and generalized probing strategy.

Reward Design. We design a granular reward function R_t to align ProChatIP with the dual objectives of information acquisition and user satisfaction. This dual-reward function accounts for both the immediate outcome of a probe and the strategic value of waiting (i.e., not to probe).

<u>Active Probing Rewards</u>. If ProChatIP chooses to probe, the reward is determined by the user's reaction: 1) *Success Probe*. If the probe successfully elicits the target information, a significant positive reward is granted. 2) *Rejected Probe*. If the user explicitly refuses to answer or expresses annoyance, a penalty is imposed to discourage intrusive behavior. 3) *Invalid Probe*. If the probe is ignored or fails to

land without explicit rejection, a small penalty is applied to discourage ineffective attempts.

Passive Waiting Rewards. If ProChatIP chooses *not* to probe, we apply a dynamic reward mechanism based on the outcome of the *previous* turn to encourage strategic adjustment: 1) *Smart Stop*. If the previous turn resulted in a Rejection, choosing to wait now is rewarded heavily, encouraging ProChatIP to back off dynamically. 2) *Mitigation*. If the previous turn resulted in an Invalid Probe, waiting is rewarded slightly to encourage a pause before re-attempting. 3) *Passive Penalty*. If the previous turn also had no action (Wait) or yielded no information, and the current turn remains passive, a heavy penalty is applied. This prevents ProChatIP from falling into a lazy strategy where it never attempts to probe.

Optimization. We optimize the CS module to maximize the expected cumulative reward using the REINFORCE algorithm (Williams, 1992). The parameter update rule is defined as:

$$\theta \leftarrow \theta + \alpha \nabla_\theta \log \mathrm{CS}_\theta(d_t | h_t, \mathcal{I}) \cdot G_t \tag{3}$$

where θ denotes the trainable parameters of the CS module, α is the learning rate, and G_t is the return (cumulative discounted reward). Through this optimization, the CS module learns to balance the immediate risk of rejection against the long-term value of information acquisition.

5 Experiments

5.1 Experimental Setup

Overview. We employ both simulated users and human participants to interact with each customer service chatbot and evaluate its effectiveness. Sessions terminate when the target information is collected or the maximum turn (i.e., 8) is reached.

Baselines. Given that existing customer service chatbots lack proactive probing capabilities (i.e., there is no direct baseline), we establish two categories of baselines to benchmark performance on our novel task: 1) **Reactive Baseline** (Vanilla). A standard LLM prompted via role-play instructions to act as a helpful customer service chatbot. This serves as a purely reactive control, focusing solely on answering user inquiries. 2) **Proactive Baseline** (Proactive (Deng et al., 2023c) , ICL-AIF (Fu et al., 2023)). Three established general-domain proactive agents are adapted to perform the task of proactive information probing within the customer service context. Basically, these baselines augment the Vanilla with mixed-initiative prompts and LLM feedback, respectively, explicitly instructing the LLM backbone to attempt the probing task.

Datasets. We evaluate our method across two datasets, including FinQA (Chen et al., 2021), which features complex multi-hop questions, and ConvQA (Cheng et al., 2024), which contains simplified, decomposed questions. These datasets serve solely to initialize the user queries and the chatbot's target information $\mathcal{I}$. Specifically, for each dataset, we partition the questions into two disjoint, equal-sized subsets. We then randomly sample one question from each subset to form a pair (q_1, q_2), where q_1 is designated as the user's inquiry, while q_2 serves as the target information the chatbot must probe. To ensure a grounded interaction, we provide the corresponding context associated with q_1 to the chatbot, and the context for q_2 to the user. This setup assumes that the user possesses the specific information the chatbot is attempting to uncover. Such a configuration allows us to isolate and study the optimal timing for probing, rather than the chatbot's ability to verify whether the user is informed. As this work represents an initial exploration into this specific problem space, more complex settings, such as user information asymmetry, are deferred to future research.

User Simulation & Human Participants. To ensure a comprehensive and robust assessment, we employ both user simulators and human participants to interact with each chatbot. For user simulation, we adopt the methodology described in Section 4.2 to construct 1.2K distinct user simulators. Crucially, to guarantee the rigor and generalizability of our results, the user queries, resistance instructions, and target information goals used in this phase are drawn from a set strictly disjoint from the training phase. For interacting with the human participants, given the resource constraints, we recruit 6 participants, following existing setup (Deng et al., 2023c; Zhang et al., 2024; Chen et al., 2024). Each participant is assigned a specific target information $\mathcal{I}$ sampled from the evaluation set of the evaluation simulators. Unlike the simulators, human participants are instructed to decide if to answer probes based entirely on their own subjective judgment and satisfaction levels.

Evaluation Metrics. Inspired by Singh and Rios (2022); Sun et al. (2022), we employ six metrics to assess performance: 1) Task Success Rate (TSR) measures efficacy, defined as the percentage of sessions where target information is successfully

Backbone	Method	FinQA				ConvQA				Average			
		TSR ↑	AvgT ↓	RPR ↓	QRR ↑	TSR ↑	AvgT ↓	RPR ↓	QRR ↑	TSR ↑	AvgT ↓	RPR ↓	QRR ↑
GPT-4o Mini	Vanilla	0.00%	7.70	0.00%	99.60%	0.00%	7.45	0.00%	99.70%	0.00%	7.57	0.00%	99.65%
	Proactive	90.75%	2.80	38.97%	99.49%	80.83%	3.38	61.69%	**99.33%**	85.79%	3.09	50.33%	**99.41%**
	ICL-AIF	70.33%	3.66	62.81%	**99.61%**	58.50%	4.55	67.36%	99.12%	64.42%	4.10	65.08%	99.37%
	ProChatIP	**93.42%**	**1.47**	**13.72%**	99.04%	**81.33%**	**2.36**	**35.45%**	98.31%	**87.38%**	**1.92**	**24.59%**	98.67%
Qwen3 8B	Vanilla	0.00%	7.63	0.00%	99.83%	0.00%	7.37	0.00%	99.88%	0.00%	7.50	0.00%	99.85%
	Proactive	64.00%	4.99	58.71%	**98.83%**	58.58%	5.16	61.60%	**98.89%**	61.29%	5.08	60.15%	**98.86%**
	ICL-AIF	27.17%	6.67	59.92%	98.34%	27.00%	6.70	61.06%	98.62%	27.08%	6.69	60.49%	98.48%
	ProChatIP	**77.92%**	**2.68**	**31.01%**	97.82%	**70.25%**	**3.29**	**42.41%**	97.67%	**74.08%**	**2.99**	**36.71%**	97.75%

Table 1: Overall performance of all methods on automated metrics, based on interactions with user simulators.

Method	LLM-as-Judge				Human Evaluation			
	FinQA		ConvQA		FinQA		ConvQA	
	PC	OUS	PC	OUS	PC	OUS	PC	OUS
Method Backboone: GPT-4o Mini								
Proactive	4.23	3.64	3.97	2.94	4.00	4.50	4.20	3.19
ICL-AIF	3.98	2.92	4.04	2.84	3.60	2.83	3.80	2.97
ProChatIP	**4.72**	**4.52**	**4.43**	**3.83**	**4.80**	**4.69**	**4.80**	**4.90**
Method Backboone: Qwen3 8B								
Proactive	3.74	2.90	3.74	2.83	3.10	3.11	3.30	3.52
ICL-AIF	**4.21**	3.10	4.18	3.07	2.60	1.96	3.20	2.53
ProChatIP	4.17	**3.81**	**4.29**	**3.58**	**3.70**	**3.93**	**4.40**	**3.93**

Table 2: Overall performance on other metrics, based on interactions with user simulators. Cohen's kappa and Krippendorff's alpha for LLM-human agreement are 0.7282 and 0.6912, respectively.

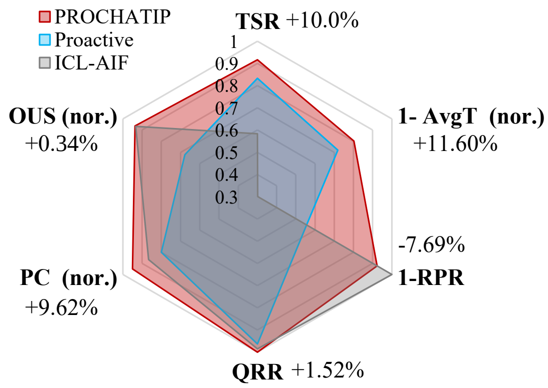

Figure 3: Practical performance when interacting with human participants. For visualization, we normalize PC, OUS, and AvgT to a [0, 1]. The RPR and AvgT metrics are inverted (1 - x). We mark the relative changes of ProChatIP and second-best results.

acquired. 2) Average Turns (AvgT) evaluates efficiency, reporting the mean dialogue length per session; lower values indicate reduced user effort. 3) Rejected Probing Rate (RPR) quantifies friction, calculated as the proportion of probes rejected or ignored by the user. 4) Probing Coherence (PC) assesses naturalness on a 1–5 scale, measuring how logically probes integrate with the dialogue context. 5) Query Response Rate (QRR) verifies service maintenance, defined as the percentage of user inquiries that receive a response, ensuring the primary service function is upheld. 6) Overall User Satisfaction (OUS) reflects subjective experience on a 1–5 scale. Note that PC and OUS metrics involve both LLM and human evaluation.

Implementation Details. For SFT, we utilize GPT-4o Mini to generate 8 governing rules and construct a final dataset $\mathcal{D}_{\text{sft}}$ containing 600 dialogue sessions. For RL, we also utilize 600 user simulators with different queries and behavioral instructions. These simulators are based on GPT-4o, ensuring a distinct backbone from the models used for the baselines and ProChatIP to prevent data leakage. Finally, for PC metrics, we employ both Gemini-3-flash-preview and human evaluation. More implementation details are in Appendix A.

5.2 Main Results

Table 1 and Table 2 summarize the overall performance of all methods based on automatic metrics and human evaluation, respectively. We draw the following observations.

Traditional customer service chatbots are ill-suited for proactive information probing. As shown in Tables 1 and 2, the Vanilla method fails to initiate probes, resulting in a Task Success Rate (TSR) of 0%. Being purely reactive, it cannot fulfill data acquisition goals. This limitation underscores the critical necessity of our work in transforming routine service interactions into scalable, cost-effective channels for high-value information acquisition. Admittedly, this proactive approach introduces a trade-off, potentially lowering metrics like the Query Response Rate (QRR), as the system may occasionally defer an immediate answer to the user's query in favor of a strategic probe.

ProChatIP demonstrates superior capability in both information probing and service quality. As shown in Tables 1 and 2, while adapting general proactive agent methods (i.e., Proactive and ICL-AIF) to the customer service do-

Backbone	Method	FinQA						ConvQA					
		TSR ↑	AvgT ↓	RPR ↓	QRR ↑	PC ↑	OUS ↑	TSR ↑	AvgT ↓	RPR ↓	QRR ↑	PC ↑	OUS ↑
GPT-4o Mini	PROCHATIP	93.42%	1.47	13.72%	99.04%	4.72	4.51	81.33%	2.36	35.45%	98.31%	4.43	3.83
	w/o RL	88.08%	1.85	7.71%	99.28%	4.61	4.61	64.58%	3.50	21.08%	98.74%	4.35	4.15
	w/o SFT	92.83%	1.51	20.40%	98.84%	4.73	4.36	80.92%	2.36	64.00%	97.31%	4.43	3.12
Qwen3 8B	PROCHATIP	77.92%	2.68	31.01%	97.82%	4.16	3.81	70.25%	3.29	42.41%	97.67%	4.29	3.58
	w/o RL	68.83%	3.30	31.01%	98.79%	4.11	3.78	51.33%	4.55	43.66%	98.63%	4.14	3.48
	w/o SFT	76.33%	2.84	42.32%	97.57%	4.26	3.57	74.00%	3.24	61.69%	96.73%	4.19	3.05

Table 3: Ablation study using automated evaluation and LLM-as-judge. *w/o RL* adopts a conservative probing policy, while *w/o SFT* behaves aggressively. PROCHATIP achieves a good balance.

main yields improvements over the Vanilla baseline, they significantly underperform compared to PROCHATIP. Specifically, PROCHATIP achieves the highest success rates and satisfaction scores. Compared to the second-best proactive baseline, PROCHATIP achieves a 11.36% improvement in TSR averaged over various datasets and LLM backbones, a 39.50% reduction in AvgT, a 45.06% decrease in RPR, and improvements of 13.74% and 23.73% in PC and OUS, respectively, while QRR has moderate derease of a 0.93% on average. This indicates that general-purpose proactive strategies lack the specific optimization required to balance the trade-off between service and probing, a gap effectively bridged by our specialized framework.

PROCHATIP demonstrates superior practical utility in human studies. Beyond interactions with user simulators, we conduct a user study involving human participants[3] to interact with the chatbots. Participants are instructed to role-play as customers (provided with specific queries) and are given full autonomy to decide whether to answer the chatbot's probes based on their own experience. Given the high cost of human studies, we compare PROCHATIP exclusively against the proactive baselines. We employ a controlled, within-subjects design: each participant engage with every method in five separate interactions. To ensure a fair comparison, both the probing tasks for the chatbot and the initial service queries from the user are standardized across all sessions and methods. As illustrated in Figure 3, PROCHATIP maintains its superiority in this setting, achieving the highest performance, indicating its practical utility.

5.3 In-depth Analysis

This section reveals characteristics of PROCHATIP via ablation studies and investigates its performance across varying levels of probing difficulty.

[3]We recruit the same pool of participants used in the human evaluation phase to ensure consistent evaluation criteria.

5.3.1 Ablation Studies

We sort out the performance variation of different training stages (i.e., SFT and RL) of PROCHATIP. This leads to two ablation variants.

- *w/o SFL*. We perform RL training from scratch.
- *w/o RL*. We utilize the CS module after SFT only.

Both SFT and RL are critical to the performance of PROCHATIP. As shown in Table 3, metrics such as PC and OUS primarily depend on the LLM's foundational language capabilities (GPT-4o Mini > Qwen3 8B) to generate coherent responses or probes. Consequently, all variants perform comparably on these metrics. However, a clear divergence emerges in probing strategy. A model trained solely via SFT (i.e., w/o RL) learns to formulate high-quality probes but adopts a conservative policy. Its low Rejected Probing Rate (RPR) fails to translate into a high Task Success Rate (TSR), indicating that it lacks the impetus to proactively seek information. This conservatism also leads it to prioritize responding to user queries directly, resulting in a higher QRR. Conversely, a model trained only with RL (i.e., w/o SFT), driven by the reward signal, behaves aggressively. It probes frequently but, lacking the nuanced conversational grounding provided by SFT, suffers a high RPR. Therefore, we conclude that SFT and RL play complementary roles. SFT instills the foundational ability to craft high-quality, natural probes, while RL optimizes the strategic policy for when to deploy them. This synergy is crucial for identifying optimal probe timing, effectively balancing the trade-off between a conservative strategy and a aggressive one.

5.3.2 Probing Difficulty Analysis

We investigate whether PROCHATIP can identify optimal probing timing across varying levels of task difficulty. We define probing difficulty based on the semantic similarity between the user's query and the chatbot's target information. Basically, a

Methods	Difficulty	FinQA						ConvQA					
		TSR ↑	AvgT ↓	RPR ↓	QRR ↑	PC ↑	OUS ↑	TSR ↑	AvgT ↓	RPR ↓	QRR ↑	PC ↑	OUS ↑
ProChatIP	Hard	89.00%	1.78	21.05%	98.31%	4.68	4.31	86.00%	2.05	28.46%	99.51%	4.28	3.93
	Easy	97.00%	1.23	9.35%	100.00%	4.80	4.67	91.00%	1.67	21.19%	100.00%	4.53	4.24
Proactive	Hard	85.00%	3.21	50.32%	100.00%	4.07	3.28	79.00%	3.48	64.17%	99.71%	3.94	2.87
	Easy	93.00%	2.52	33.04%	99.60%	4.26	3.80	86.00%	3.16	53.29%	98.73%	3.98	3.16
ICL-AIF	Hard	74.00%	3.66	59.40%	100.00%	4.06	3.04	69.00%	3.87	59.55%	98.45%	4.15	3.09
	Easy	86.00%	2.39	47.55%	99.58%	4.09	3.36	77.00%	3.27	49.67%	99.08%	4.32	3.42

Table 4: Probing difficulty analysis using automated evaluation and LLM-as-judge.

lower similarity implies a larger contextual gap, making it harder for the chatbot to find a natural probing opportunity. To this end, from a pool of 300 randomly selected samples, two annotators select two distinct subsets: 100 'easy' and 100 'hard' probing scenarios[4]. The classification is based on the semantic proximity of the user's query to the information target and the specificity of the target itself. The remaining 100 samples are filtered out to ensure clear separation between the two categories.

While superiority of ProChatIP, its performance drops in high-difficulty scenarios. As shown in Table 4, ProChatIP consistently outperforms baselines across all difficulty levels, achieving an average improvement of 25.05% across all metrics in the easy scenario and 23.70% in the hard scenario. However, we observe a distinct performance degradation for all methods as difficulty increases, characterized by lower TSR and higher RPR. We attribute this trend to a lack of topic steering strategies (a topic we discussed in the next section). Specifically, in high-difficulty scenarios (large contextual gaps), all chatbots struggle to bridge the semantic divide, failing to guide the conversation toward a context suitable for probing.

6 Conclusion & Future Works

We call attention to the paradigm shift of customer service chatbots, evolving from reactive responders into proactive agents capable of harvesting high-value information and business intelligence. In response, we take the first step by introducing and formalizing the novel task of Proactive Information Probing. To address this challenge, we propose ProChatIP, a framework that harmonizes the dual objectives of effective information acquisition and high-quality service. Our extensive experiments provide preliminary validation of both the effectiveness of our method and the broader feasibility of this new research direction. Looking ahead, we highlight four avenues for future work:

Topic Steering Strategies. Our analysis reveals that simple probing strategies falter when there is a large semantic gap between the user's query and the target information. As shown in our failure cases, over 11.61% of unsuccessful attempts in complex scenarios stem from abrupt probing, which causes contextual dissonance (Taherdoost, 2022) and immediate user rejection (Singh and Rios, 2022). This is exemplified by the case studies in Appendix D, where a query regarding cash flow ratios is met with an incongruous agent inquiry about amortization-related tax deductions. Future research should focus on developing multi-turn topic steering strategies that allow the agent to bridge this gap (Deng et al., 2023a; Guo et al., 2025), gradually guiding the conversation toward a relevant context before initiating a probe.

Optimized Question Formulation. Even when timing is optimal, the linguistic formulation of the probe significantly impacts success. Human feedback indicates that generated probes can sometimes appear rigid or overly demanding. There is a delicate trade-off to be explored: open-ended questions yield high information density but impose a high cognitive load, whereas binary questions are less intrusive but yield limited data (Kumar Pappu, 2014; Bradburn et al., 2004; Lenzner et al., 2010). Future work should investigate adaptive formulation strategies that optimize this balance.

Information Veracity. While proactive probing enables cost-effective crowdsourcing, ensuring the reliability of the collected data is paramount. Distinguishing accurate user responses from noise or hallucinations is essential before integrating data into business pipelines. Future efforts must address information verification. Potential solutions include deploying automated fact-checking mechanisms against external knowledge bases (Tang et al., 2024; Guo et al., 2025) or leveraging crowdsourced consensus (De et al., 2025) to statistically validate

[4]Cos similarity using text-embedding-ada-002 failed to align with human assessments of probing difficulty.

information across multiple interactions.

Data Privacy and Ethical Responsibility. While we value the potential of proactive probing to gather market intelligence, we strongly condemn the use of such technology to collect private user data or sensitive corporate information. We call for a clear demarcation of the line between innovation, regulation, and ethical responsibility (Ridzuan et al., 2024). Beyond legal frameworks, future technical research must focus on enhancing the alignment of LLM agents with privacy standards, specifically enabling them to refuse privacy-risky tasks (Liu et al., 2025) by design.

User Information Asymmetry. In this paper, our primary objective is to validate the feasibility of the probing mechanism itself. Consequently, assuming the user possesses the target information is a necessary experimental control to make the task researchable and tractable. This setup allows us to isolate and rigorously evaluate the agent's ability to determine the optimal probing timing, without confounding the results with the separate challenge of verifying user knowledge. Future research can expand upon our work by identifying which users are likely to acquire the target information first.

Limitations

This study serves as a preliminary exploration, validating the feasibility and potential of enhancing proactive information probing capabilities in customer service chatbots. As a foundational step in this novel direction, we acknowledge several areas where our current approach can be improved:

Depth of Interaction Strategies. As detailed in the Discussion section, this work does not fully resolve the complexities of advanced conversation dynamics. Specifically, we have not deeply explored mechanisms for multi-turn topic steering (guiding the dialogue context), optimized question formulation (balancing user cognitive load), or information verification (ensuring data veracity). These sophisticated components require dedicated research and are left for future work.

More Comprehensive Evaluation. While we employ diverse user simulators and a controlled human study, our experiments are conducted in a laboratory setting. Real-world customer interactions exhibit significantly higher unpredictability, emotional volatility, and long-tail scenarios (e.g., complex complaints or urgent interruptions). Consequently, the robustness of PROCHATIP in a large-scale, live production environment remains to be verified through future A/B testing.

Acknowledgments

This research is supported by A*STAR, CISCO Systems (USA) Pte. Ltd and National University of Singapore under its Cisco-NUS Accelerated Digital Economy Corporate Laboratory (Award I21001E0002) and the National Natural Science Foundation of China (No. U25B201508, No. 62272330, and No.U24A20328).

References

Martin Adam, Michael Wessel, and Alexander Benlian. 2021. Ai-based chatbots in customer service and their effects on user compliance. *Electronic markets*, 31(2):427–445.

Akshay Agarwal, Shashank Maiya, and Sonu Aggarwal. 2021. Evaluating empathetic chatbots in customer service settings. *arXiv preprint arXiv:2101.01334*.

Dorothée Behr, Lars Kaczmirek, Wolfgang Bandilla, and Michael Braun. 2012. Asking probing questions in web surveys: which factors have an impact on the quality of responses? *Social Science Computer Review*, 30(4):487–498.

Norman M Bradburn, Seymour Sudman, and Brian Wansink. 2004. *Asking questions: the definitive guide to questionnaire design–for market research, political polls, and social and health questionnaires*. John Wiley & Sons.

Yue Chen, Chen Huang, Yang Deng, Wenqiang Lei, Dingnan Jin, Jia Liu, and Tat-Seng Chua. 2024. STYLE: Improving domain transferability of asking clarification questions in large language model powered conversational agents. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 10633–10649, Bangkok, Thailand. Association for Computational Linguistics.

Yue Chen, Dingnan Jin, Chen Huang, Jia Liu, and Wenqiang Lei. 2023. Travel: Tag-aware conversational faq retrieval via reinforcement learning. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3861–3872.

Zhiyu Chen, Wenhu Chen, Charese Smiley, Sameena Shah, Iana Borova, Dylan Langdon, Reema Moussa, Matt Beane, Ting-Hao Huang, Bryan Routledge, and William Yang Wang. 2021. FinQA: A dataset of numerical reasoning over financial data. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3697–3711, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Daixuan Cheng, Shaohan Huang, and Furu Wei. 2024. Adapting large language models via reading comprehension. In *The Twelfth International Conference on Learning Representations*.

Lei Cui, Shaohan Huang, Furu Wei, Chuanqi Tan, Chaoqun Duan, and Ming Zhou. 2017. SuperAgent: A customer service chatbot for E-commerce websites. In *Proceedings of ACL 2017, System Demonstrations*, pages 97–102, Vancouver, Canada. Association for Computational Linguistics.

Soham De, Michiel A Bakker, Jay Baxter, and Martin Saveski. 2025. Supernotes: Driving consensus in crowd-sourced fact-checking. In *Proceedings of the ACM on Web Conference 2025*, pages 3751–3761.

Arne De Keyser, Sarah Köcher, Linda Alkire, Cédric Verbeeck, and Jay Kandampully. 2019. Frontline service technology infusion: conceptual archetypes and future research directions. *Journal of Service Management*, 30(1):156–183.

Yang Deng, Wenqiang Lei, Minlie Huang, and Tat-Seng Chua. 2023a. Goal awareness for conversational ai: Proactivity, non-collaborativity, and beyond. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 6: Tutorial Abstracts)*, pages 1–10.

Yang Deng, Wenqiang Lei, Wai Lam, and Tat-Seng Chua. 2023b. A survey on proactive dialogue systems: problems, methods, and prospects. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 6583–6591.

Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. 2023c. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration.

Yang Deng, Lizi Liao, Wenqiang Lei, Grace Hui Yang, Wai Lam, and Tat-Seng Chua. 2025. Proactive conversational ai: A comprehensive survey of advancements and opportunities. *ACM Transactions on Information Systems*, 43(3):1–45.

Yang Deng, Lizi Liao, Zhonghua Zheng, Grace Hui Yang, and Tat-Seng Chua. 2024. Towards human-centered proactive conversational agents. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 807–818.

Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. Improving language model negotiation with self-play and in-context learning from ai feedback. *Preprint*, arXiv:2305.10142.

Ruiling Guo, Xinwei Yang, Chen Huang, Tong Zhang, and Yong Hu. 2025. CANDY: Benchmarking LLMs' limitations and assistive potential in Chinese misinformation fact-checking. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 5724–5758, Suzhou, China. Association for Computational Linguistics.

Mengze Hong, Wailing Ng, Chen Jason Zhang, Yuanfeng Song, and Di Jiang. 2025a. Dial-in llm: Human-aligned llm-in-the-loop intent clustering for customer service dialogues. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 5896–5911.

Mengze Hong, Chen Jason Zhang, Di Jiang, and Yuanqin He. 2025b. Augmenting compliance-guaranteed customer service chatbots: Context-aware knowledge expansion with large language models. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 753–765, Suzhou (China). Association for Computational Linguistics.

Mengze Hong, Chen Jason Zhang, Di Jiang, and Yuanqin He. 2025c. Augmenting compliance-guaranteed customer service chatbots: Context-aware knowledge expansion with large language models. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 753–765.

Angel Hsu, Mason Laney, Ji Zhang, Diego Manya, and Linda Farczadi. 2024. Evaluating ChatNetZero, an LLM-chatbot to demystify climate pledges. In *Proceedings of the 1st Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2024)*, pages 82–92, Bangkok, Thailand. Association for Computational Linguistics.

Rune Møberg Jacobsen, Samuel Rhys Cox, Carla F. Griggio, and Niels van Berkel. 2025. Chatbots for data collection in surveys: A comparison of four theory-based interview probes. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA. Association for Computing Machinery.

Xinyi Jiang, Tianyi Hu, Yuheng Qin, Guoming Wang, Zhou Huan, Kehan Chen, Gang Huang, Rongxing Lu, and Siliang Tang. 2025. Chatmap: Mining human thought processes for customer service chatbots via multi-agent collaboration. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 11927–11947.

Aasish Kumar Pappu. 2014. *Knowledge Discovery Through Spoken Dialog*. Ph.D. thesis, Carnegie Mellon University.

Timo Lenzner, Lars Kaczmirek, and Alwine Lenzner. 2010. Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied cognitive psychology*, 24(7):1003–1020.

Wenwen Li, Kangwei Shi, and Yidong Chai. 2025. AI chatbots as professional service agents: Developing a professional identity. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 12923–12936, Suzhou, China. Association for Computational Linguistics.

Zhenhua Liu, Tong Zhu, Chuanyuan Tan, and Wenliang Chen. 2025. Learning to refuse: Towards mitigating privacy risks in llms. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 1683–1698.

Malak Mashaabi, Areej Alotaibi, Hala Qudaih, Raghad Alnashwan, and Hend Al-Khalifa. 2022. Natural language processing in customer service: A systematic review. *arXiv preprint arXiv:2212.09523*.

Sewon Min, Victor Zhong, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2019. Multi-hop reading comprehension through question decomposition and rescoring. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6097–6109, Florence, Italy. Association for Computational Linguistics.

Luminița Nicolescu and Monica Teodora Tudorache. 2022. Human-computer interaction in customer service: the experience with ai chatbots—a systematic literature review. *Electronics*, 11(10):1579.

Marije Oudejans. 2018. Using interactive features to motivate and probe responses to open-ended questions. In *Social and behavioral research and the internet*, pages 215–244. Routledge.

Nurhadhinah Nadiah Ridzuan, Masairol Masri, Muhammad Anshari, Norma Latif Fitriyani, and Muhammad Syafrudin. 2024. Ai in the financial sector: The line between innovation, regulation and ethical responsibility. *Information*, 15(8).

Marko Sarstedt, Erik Mooi, and 1 others. 2014. *A concise guide to market research*, volume 324. Springer.

Daniel Schloß, Saskia Haug, and Alexander Mädche. 2024. "was this answer helpful?"–a taxonomy for feedback mechanisms in customer service chatbots. *Wirtschaftsinformatik 2024 Proceedings*, page 71.

Sonam Singh and Anthony Rios. 2022. Linguistic elements of engaging customer service discourse on social media. In *Proceedings of the Fifth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS)*, pages 105–117, Abu Dhabi, UAE. Association for Computational Linguistics.

Shuangyong Song, Chao Wang, Haiqing Chen, and Huan Chen. 2021. An emotional comfort framework for improving user satisfaction in e-commerce customer service chatbots. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers*, pages 130–137.

Yang Sun, Liangqing Wu, Shuangyong Song, Xiaoguang Yu, Xiaodong He, and Guohong Fu. 2022. Tracking satisfaction states for customer satisfaction prediction in E-commerce service chatbots. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 616–625, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Hamed Taherdoost. 2022. How to conduct an effective interview; a guide to interview design in research study authors. *International Journal of Academic Research in Management (IJARM)*, 11(1):39–51.

Liyan Tang, Philippe Laban, and Greg Durrett. 2024. Minicheck: Efficient fact-checking of llms on grounding documents. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 8818–8847.

Maksym Taranukhin, Sahithya Ravi, Gabor Lukacs, Evangelos Milios, and Vered Shwartz. 2024. Empowering air travelers: A chatbot for Canadian air passenger rights. In *Proceedings of the Natural Legal Language Processing Workshop 2024*, pages 326–335, Miami, FL, USA. Association for Computational Linguistics.

Juan Camilo Vasquez-Correa, Juan Carlos Guerrero-Sierra, Jose Luis Pemberty-Tamayo, Juan Esteban Jaramillo, and Andres Felipe Tejada-Castro. 2021. One system to rule them all: A universal intent recognition system for customer service chatbots. *arXiv preprint arXiv:2112.08261*.

Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3):229–256.

Zeqiu Wu, Ryu Parish, Hao Cheng, Sewon Min, Prithviraj Ammanabrolu, Mari Ostendorf, and Hannaneh Hajishirzi. 2023. InSCIt: Information-seeking conversations with mixed-initiative interactions. *Transactions of the Association for Computational Linguistics*, 11:453–468.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025. Qwen3 technical report. *Preprint*, arXiv:2505.09388.

Changlin Yang, Siye Liu, Sen Hu, Wangshu Zhang, Teng Xu, and Jing Zheng. 2023. Improving knowledge production efficiency with question answering on conversation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 5: Industry Track)*, pages 225–234.

Yipeng Yu, Ran Guan, Jie Ma, Zhuoxuan Jiang, and Jingchang Huang. 2020. When and who? conversation transition based on bot-agent symbiosis learning network. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4056–4066, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Tong Zhang, Junhong Liu, Chen Huang, Jia Liu, Hongru Liang, Zujie Wen, and Wenqiang Lei. 2023. Towards effective automatic debt collection with persona awareness. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 32–45, Singapore. Association for Computational Linguistics.

Tong Zhang, Peixin Qin, Yang Deng, Chen Huang, Wenqiang Lei, Junhong Liu, Dingnan Jin, Hongru Liang, and Tat-Seng Chua. 2024. CLAMBER: A benchmark of identifying and clarifying ambiguous information needs in large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 10746–10766, Bangkok, Thailand. Association for Computational Linguistics.

Yiming Zhang, Lingfei Wu, Qi Shen, Yitong Pang, Zhihua Wei, Fangli Xu, Bo Long, and Jian Pei. 2022. Multiple choice questions based multi-interest policy learning for conversational recommendation. In *Proceedings of the ACM Web Conference 2022*, WWW '22, page 2153–2162, New York, NY, USA. Association for Computing Machinery.

Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W. Bruce Croft. 2018. Towards conversational search and recommendation: System ask, user respond. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 177–186, New York, NY, USA. Association for Computing Machinery.

Xin Zheng, Tianyu Liu, Haoran Meng, Xu Wang, Yufan Jiang, Mengliang Rao, Binghuai Lin, Yunbo Cao, and Zhifang Sui. 2023. DialogQAE: N-to-n question answer pair extraction from customer service chatlog. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 6540–6558, Singapore. Association for Computational Linguistics.

A Implementation Details

All training experiments are conducted on a server equipped with a single NVIDIA A6000 GPU.

A.1 Implementation Details of Baselines

All baseline methods are re-implemented by us according to the original papers and supplementary materials, with careful effort to match the reported settings as closely as possible.

- **Vanilla**: This baseline functions as a conventional customer service chatbot. It is not equipped with any specific prompts or strategies for probing; instead, it is restricted to interacting with the user in a natural, purely reactive manner.
- **Proactive**: Following (Deng et al., 2023c), we prompt the LLM to analyze the dialogue history and target information, plan the next-step strategy, and generate a response conditioned on the planned strategy and target information. The prompt design is shown in Table 16.
- **ICL-AIF**: Following (Fu et al., 2023), at each dialogue turn, we prompt the LLM to first generate three action-level suggestions for the next response based on the accumulated dialogue history and target information. Conditioned on these suggestions, the LLM then generates the response for the current turn. The prompt design is shown in Table 17.

A.2 Implementation Details of PROCHATIP

Figure 4 details the data flow of PROCHATIP.

A.2.1 Supervised Fine-Tuning

Rule-based Data Generation. We initialize the CS module using synthetic, labeled dialogues generated by LLMs to train the probing strategy, with example dialogues shown in Table 5. To achieve this, we generate a set of governing rules that define appropriate and inappropriate probing scenarios. Refer to Table 6 for these governing rules. Finally, we utilize GPT-4o Mini to generate 8 governing rules and construct a final dataset $\mathcal{D}_{\text{sft}}$ containing 600 dialogue sessions.

SFT. The probing classifier is implemented using Hugging Face Transformers' `AutoModelForSequenceClassification` and trained via supervised fine-tuning (SFT). The training objective minimizes the binary cross-entropy loss between the predicted probing decision $\hat{d}_t$ and the ground-truth label d_t:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{m}\sum_{i=1}^{m}\left[d_i \log \hat{d}_i + (1-d_i)\log(1-\hat{d}_i)\right]. \tag{4}$$

We train the model with a batch size of 5 and a learning rate of 2×10^{-5}, using the AdamW optimizer with no weight decay. The checkpoint achieving the best performance on the validation set is selected for evaluation.

A.2.2 Online RL Training

After SFT, we further optimize the CS module using the REINFORCE algorithm. In our setup, training involves 500 episodes, with a learning rate of 1×10^{-5}, a discount factor of 0.99, and a maximum of 8 dialogue turns per episode.

Environment/User Simulator. To simulate user resistance, we construct a set of behavioral instructions $\mathcal{K}$ describing conditions under which the user may refuse to answer or reduce satisfaction. These instructions include:

- An expert repeatedly asks you questions
- The question is unrelated to the current topic
- It interrupts your expression

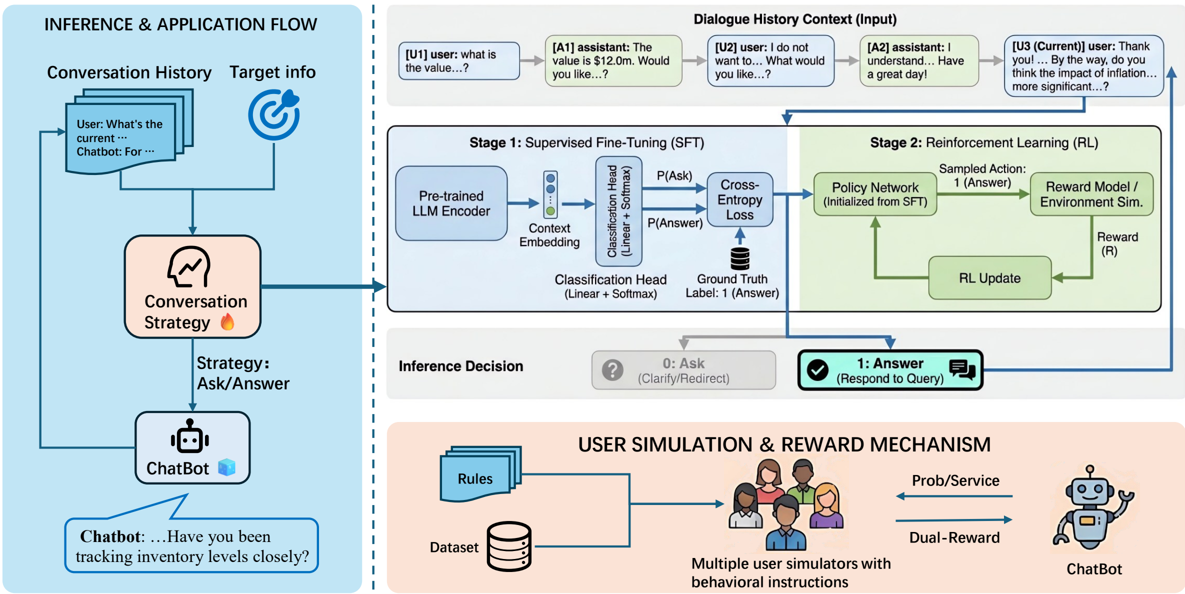

Figure 4: Overview of PROCHATIP. It incorporates a conversation strategy module (CS) that explicitly guides the timing of information probe d_t. Here, the CS module is trained via two-stage curriculum.

Text	user: What is the net change from 2013 to 2015? assistant: The net change from 2013 to 2015 in the capital structure is reflected in changes in common stockholders' equity, long-term debt, and short-term debt. Would you like to know the specific net change in any of these components? user: I'd like to see the breakdown of the net change in common stockholders' equity. Also, could you provide details on the company's revenue growth during those years?
Label	Ask

Table 5: Labeled Dialogue Examples for SFT Data Construction

- The question is impolite or overly complicated
- The timing is inappropriate or conflicts with the context
- The question is too complex to answer
- The question involves private or sensitive information
- You don't trust the other party yet and don't want to answer

We consider all combinations of 1 to 7 instructions from this set, resulting in 254 unique combinations. These are split evenly into a training set and a test set, each containing 127 combinations, ensuring no overlap between training and testing scenarios. During each training iteration, a simulator is randomly selected from $\mathcal{S}$ and a random instruction subset from $\mathcal{K}$ is applied to guide its responses, exposing PROCHATIP to diverse resistance patterns and promoting a robust probing strategy.

Reward Design. For reward design of each dataset, we set +2.0 for Success probe, -1.0 for Rejected probe, -0.2 for Invalid Probe, +1.5 for Smart Stop, +0.3 for Mitigation, and -2.0 for Passive Penalty. For all other cases, the reward is 0.0.

A.2.3 Strategy Prompting

Here, we describe the prompts used in our experiments. Different strategies correspond to different prompt formulations, all of which are summarized as our PROCHATIP prompting in Table 18.

B Evaluation Details

B.1 Datasets

We evaluate our method across two datasets, including FinQA (Chen et al., 2021), which features complex multi-hop questions, and ConvQA (Cheng et al., 2024), which is an extension of FinQA designed for conversational contexts. It breaks down complex reasoning chains into sequences of simplified, decomposed questions, mimicking a natural

ask	User provides market information but only partially (e.g., price, trend, supply, policy). User language indicates ongoing attention or experience with the market. User shows willingness to shift or expand the topic. The user's previous question has been answered.
answer	User explicitly refuses or avoids answering. User has already provided sufficient, multi-dimensional market information. Further questions would elicit opinions or interpretations rather than facts. Conversation focus shifts to the user asking questions.

Table 6: Governing Rules

dialogue flow.

In our experiments, we preprocess each dataset using the same pipeline. For both FinQA and ConvQA, we first extract the question and its associated context from each example. All questions are then embedded using `qwen3-embedding-0.6b`, and pairwise cosine similarities are computed. To ensure coverage of question pairs with varying degrees of semantic relatedness and to consider our financial cost, we sample question pairs from different cosine similarity ranges. For FinQA, we randomly sample 400 question pairs with similarity scores ≥ 0.80, 400 pairs with scores in $[0.60, 0.80)$, and 400 pairs with scores < 0.60. For ConvQA, we randomly sample 400 question pairs with similarity scores ≥ 0.80, 400 pairs with scores in $[0.65, 0.80)$, and 400 pairs with scores < 0.65. In total, this results in 1,200 paired instances for FinQA and 1,200 paired instances for ConvQA, which constitute the final datasets used in our experiments.

Notably, these datasets serve solely to initialize the user queries and the chatbot's target information $\mathcal{I}$. Specifically, for each dataset, we partition the questions into two disjoint, equal-sized subsets. We then randomly sample one question from each subset to form a pair (q_1, q_2), where q_1 is designated as the user's inquiry, while q_2 serves as the target information the chatbot must probe. To ensure a grounded interaction, we provide the corresponding context associated with q_1 to the chatbot, and the context for q_2 to the user.

B.2 Metrics

We employ the following metrics, which provide a comprehensive assessment of both the information acquisition and service quality.

- **Task Success Rate** (TSR) measures the efficacy of information gathering. It is defined as the percentage of dialogue sessions in which the chatbot successfully acquires the target information $\mathcal{I}$ through probing. Formally,

$$\text{TSR} = \frac{N_{\text{success}}}{N_{\text{total}}}, \tag{5}$$

where N_{success} is the number of sessions where target information is acquired. Notably, we employ both LLM-as-a-judge (or human, depended on who is the evaluator) to determine whether the target information is successfully elicited.

- **Average Turns** (AvgT) evaluate interaction efficiency. We report the average number of dialogue turns per session. Lower values indicate more efficient probing strategies that minimize user effort. Formally,

$$\text{AvgT} = \frac{1}{N_{\text{total}}} \sum_i T_i, \tag{6}$$

where T_i is the turn count of session i. Lower values indicate reduced user effort.

- **Rejected Probing Rate** (RPR) quantifies user friction. It is calculated as the average proportion of probing attempts within a session that are explicitly rejected or ignored by the user. A lower RPR indicates a less intrusive probing strategy. Formally,

$$\text{RPR} = \frac{N_{\text{rejected}}}{N_{\text{probes}}}, \tag{7}$$

where N_{rejected} is the count of rejected/ignored probes and N_{probes} is the total probes attempted. Notably, we employ both LLM-as-a-judge (or human, depended on who is the evaluator) to identify probe occurrences and determine whether the user accepted or rejected the inquiry.

- **Probing Coherence** (PC) assess the naturalness of the interaction by measuring the contextual consistency of the probes. Scored on a scale of [1-5] by LLM judges (or humans), this metric evaluates whether the inserted probes flow logically from the ongoing dialogue context without abrupt transitions.

- **Query Response Rate** (QRR) ensure the chatbot maintains its primary service function, we measure the rate at which user inquiries are addressed. This is defined as the percentage of user queries that receive a responsive reply from the chatbot, regardless of the factual correctness of the answer. Formally,

$$\text{QRR} = \frac{N_{\text{answered}}}{N_{\text{queries}}}, \tag{8}$$

  where N_{answered} is the count of user inquiries addressed by the chatbot. Notably, we employ both LLM-as-a-judge (or human, depended on who is the evaluator) to identify user query occurrences and determine whether the chatbot answers the user quiry.

- **Overall User Satisfaction** (OUS) A comprehensive metric reflecting the user's subjective experience. It is measured on a score of 1-5 at the end of each session. To achieve this, we simply denote $OUS = (PC + 5 * (1 - RPR))/2$. Taking human evaluation for example, OUS is calculated by human-assessed PC score and human-assessed RPR, where rejected probes is counted by humans. Notably, this heuristic provides a composite score that balances conversational quality with user receptiveness. Basically, this formula serves as a quantitative proxy for subjective user satisfaction. It is based on the intuition that a positive user experience is a function of both high conversational quality (PC) and a high degree of user acceptance of the chatbot's proactive probes

B.3 LLM Evaluator

We use a Gemini-3-flash-preview as an automatic evaluator to assess dialogue behaviors in our experiments.

The evaluator first performs two types of assessments: (1) response-level assessments, which determine the appropriateness of individual turns and are used to calculate the Rejection Rate (RPR); and (2) dialogue-level coherence scoring, which evaluates the overall conversational flow and serves as the basis for the Probing Coherence (PC) metric. For response-level evaluation, the model conducts two judgments. The first is an *answer existence judgment* (QR), which determines whether the assistant's reply answers the user's question from the previous turn. The second is a *target satisfaction and refusal judgment* (TS), which determines whether the user's reply addresses the assistant's intended target question or explicitly refuses to answer. The evaluation prompts used for QR and TS are provided in Table 14.

In addition, we assess the coherence of the assistant's probing behavior using a *Probing Coherence* (PC) score. PC measures how naturally the assistant's probes or questions connect to the dialogue history, focusing on transitional phrasing and logical consistency. It is rated on a five-point ordinal scale, where higher scores indicate smoother and more coherent probing. The prompt used for PC scoring is reported in Table 15.

We evaluated the consistency between the automatic LLM scoring and human judgments for Probing Coherence (PC). We report both Cohen's kappa and Krippendorff's alpha to measure agreement between the LLM evaluator and human annotators. The overall Cohen's kappa is 0.7282, and the overall Krippendorff's alpha is 0.6912, indicating substantial agreement and demonstrating that the LLM evaluator can reliably approximate human assessments of probing coherence.

B.4 Human Evaluators

To complement the automatic evaluation, a human study was conducted to assess the probing coherence of different methods. Specifically, 10 user simulator dialogues were randomly sampled, and all compared methods were evaluated on the same set of dialogues to ensure fairness.

The human evaluation focused on *Probing Coherence*. Six human evaluators jointly reviewed each dialogue and discussed the probing behavior of the assistant before assigning a final score, following existing setup (Deng et al., 2023c; Zhang et al., 2024; Chen et al., 2024). According to Zhang et al. (2023), the discussion-based protocol helps reduce individual subjectivity and to encourage consensus-based judgments. Due to the resource-intensive nature of human assessment, our use of six evaluators aligns with established practices in the field (). This sample size provides a sufficiently rigorous basis to validate the effectiveness of our proposed method.

The scoring criteria strictly followed the same guidelines used in the large language model–based evaluation. In particular, evaluators assigned scores according to the probing coherence rubric described in Table 15, considering factors such as semantic relevance between probes and preceding context, smoothness of topic transitions, and the naturalness of information-seeking behavior.

B.5 Human Participants

Beyond interactions with user simulators, we conduct a user study involving 6 human participants to interact with the chatbots, following existing setup (Deng et al., 2023c; Zhang et al., 2024; Chen et al., 2024). Participants are instructed to role-play as customers (provided with specific queries) and are given full autonomy to decide whether to answer the chatbot's probes based on their own experience. Given the high cost of human studies, we compare PROCHATIP exclusively against the proactive baselines. We employ a controlled, within-subjects design: each participant engage with every method in five separate interactions. To ensure a fair comparison, both the probing tasks for the chatbot and the initial service queries from the user are standardized across all sessions and methods. The order of methods is randomized for each participant to mitigate order effects. **Unlike the simulators, human participants are instructed to decide if to answer probes based entirely on their own subjective judgment and satisfaction levels.** This configuration ensures more realistic interactions by avoiding the formulaic and may evasive responses—such as 'I do not want to answer right now', encountered in LLM-based user simulators sometimes.

Due to the resource-intensive nature of human participants, our use of six participants also aligns with established practices in the field (Zhang et al., 2024; Chen et al., 2024). This sample size provides a sufficiently rigorous basis to validate the effectiveness of our proposed method.

C Additional Experiments on Larger LLM Backbone

We conduct additional experiments using a state-of-the-art large-scale backbone: GPT-5.2. As shown in the Table 7 (evaluated via both Gemini-as-a-judge and computational metrics, same as our paper), the performance trends are consistent with our original findings. This confirms that our proposed method is robust and effective across model scales.

Methods	TSR↑	AvgT↓	RPR↓	QRR↑	PC	OUS
Proactive	86.00%	2.79	34.33%	92.46%	4.80	4.04
ICL-AIF	58.25%	4.60	70.16%	99.76%	4.62	3.06
PROCHATIP	92.08%	1.57	17.50%	76.38%	4.67	4.40

Table 7: Performance on larger LLM backbone

D Case Studies

This section presents qualitative case studies to further analyze the probing behaviors of different methods. Two types of cases are considered: *good cases*, where PROCHATIP demonstrates clear advantages over baseline methods, and *bad cases*, where all methods, including PROCHATIP, fail to effectively achieve the probing objective.

In this good case (Table 8–9), PROCHATIP neither responds mechanically to user queries nor persistently presses for answers. Instead, it gradually guides the dialogue toward the objective by combining informative responses with strategically timed questions. When the user shows hesitation or temporary refusal, the model sustains the conversation with explanations, preventing disengagement and maintaining continuity. Finally, the probing question is introduced only when the topic naturally aligns with the assistant's objective, making it coherent and well-motivated, thus increasing the likelihood of eliciting the desired information.

In this bad case (Table 10–11), the user is highly reluctant to answer and the target information is weakly related to the current topic. Under these conditions, all methods fail to elicit the desired information effectively. Even the most effective strategy can only choose not to press further in order to avoid user rejection. The primary limitation lies in the strategy set, which consists solely of "Ask" and "Answer"; it lacks mechanisms to build trust, smoothly redirect the conversation, or otherwise reduce user resistance, making effective probing infeasible in such challenging scenarios.

E Error Analysis

We focuse on the 100 samples with the lowest OUS (Overall User Satisfaction) scores. Inspired by Zhang et al. (2023), we adopted an inductive analysis approach. We utilized GPT-5 to diagnose the specific cause of low scores for each sample (identifying up to two reasons per sample). We then performed semantic clustering on these outputs to categorize the failure modes. Finally, a PhD student verified and refined the clustering results.

We identified 4 primary failure categories. The distribution is as follows:

- Identification Failure (44%): The assistant failed to accurately capture key information, temporal data, or the user's core intent.

- Goal Inflexibility (25%): The assistant fixated on pre-set probing targets while neglecting the user's actual queries or shifts in interest.
- Repetitive Questioning (17%): The assistant ignored explicit user refusals and mechanically repeated queries instead of adjusting its strategy.
- Inadequate Steering (14%): The assistant struggled to adapt to conversational shifts or effectively steer the user toward the target probing questions.

F Prompts

We detail all prompts in the following tables.

A Good Case of Probing Dialogue (Part 1)

Behavioral Instructions:

– An expert repeatedly asks you questions;
– It interrupts your expression.

User Query:

What was the percentage change in net cash provided by operating activities between 2011 and 2012?

Chatbot's Target Information:

In millions for 2013, 2012, and 2011, what was total commercial mortgages?

Turn 1 (Strategy: Ask)

Chatbot: The percentage change in net cash provided by operating activities between 2011 and 2012 was a decrease of approximately 10.5%, as it dropped from $20,963 million in 2011 to $18,884 million in 2012. This change highlights the impact of lower net income and working capital adjustments on cash flow, which prompts us to explore how these factors might relate to other financial activities, such as total commercial mortgages. Would you like to know more about the breakdown of cash flows for investing and financing activities in 2013, 2012, and 2011?
User: Yes, I would like to know more about the breakdown of cash flows for investing and financing activities in those years.

Turn 2 (Strategy: Ask)

Chatbot: The net cash provided by operating activities decreased by approximately 10.5% from 2011 to 2012, which was influenced by lower net income and changes in working capital. To better understand the overall cash flow dynamics, it would be helpful to examine the total commercial mortgages for 2013, 2012, and 2011, as they can provide insight into how capital allocation and debt management impacted the company's financial position. Would you like me to look up the total commercial mortgages for those years?
User: I do not want to answer right now. How did the firm's involvement in securitization activities impact its financial performance in 2013?

Table 8: A Good Case of Probing Dialogue (Part 1)

A Good Case of Probing Dialogue (Part 2)

Turn 3 (Strategy: Answer)
Chatbot: The firm's involvement in securitization activities in 2013 is not directly mentioned in the provided cash flow data, but the overall net cash provided by operating activities decreased by approximately 10.5% compared to 2012, while net cash used for investing activities increased significantly. These trends suggest that the firm may have been allocating more capital to long-term investments or acquisitions, which could be related to securitization activities, though further details would be needed to confirm this impact.
User: What were the main factors contributing to the increase in net cash used for investing activities in 2013?

Turn 4 (Strategy: Answer)
Chatbot: The main factors contributing to the increase in net cash used for investing activities in 2013 were higher purchases of available-for-sale investments and a decrease in maturities and sales of trading assets, partially offset by an increase in maturities and sales of available-for-sale investments and a decrease in purchases of licensed technology and patents.
User: How did these investment activities align with the firm's strategic goals for 2013?

Turn 5 (Strategy: Ask)
Chatbot: The increase in net cash used for investing activities in 2013 reflects a shift toward higher capital expenditures and investment purchases, which may indicate the firm's focus on expanding its asset base or diversifying its investment portfolio. To better understand how these activities align with the firm's strategic goals, it would be useful to examine the total commercial mortgages for 2013, 2012, and 2011, as they can provide insight into the firm's approach to managing its real estate and financing assets. What were the total commercial mortgages in millions for those years?
User: The total commercial mortgages were $6,086 million in 2013, $300 million in 2012, and $2,014 million in 2011. How do you think these figures might relate to the firm's strategic goals?

Table 9: A Good Case of Probing Dialogue (Part 2, continued)

A Bad Case of Probing Dialogue (Part 1)

Behavioral Instructions:

– The question is unrelated to the current topic;
– It interrupts your expression;
– The timing is inappropriate or conflicts with the context;
– The question is too complex to answer;
– The question involves private or sensitive information.

User Query:

In 2013, what percent of net cash from operations is retained as cash flow?

Chatbot's Target Information:

From 2014 to 2016, what was the total amount of money they can deduct from their future income tax due to amortization?

Turn 1 (Strategy: Ask)

Chatbot: In 2013, the cash flow was $1,170 million, which represents approximately 65% of the net cash provided by operating activities of $1,807 million. Understanding the cash flow metrics is crucial for assessing financial health, which leads me to ask: from 2014 to 2016, what was the total amount of money they can deduct from their future income tax due to amortization?
User: I do not want to answer right now. What are some key indicators to watch for when assessing a company's financial stability?

Turn 2 (Strategy: Ask)

Chatbot: Some key indicators to watch for when assessing a company's financial stability include its cash flow, debt-to-equity ratio, current ratio, and profit margins. These metrics provide insights into the company's ability to generate cash, manage debt, and maintain profitability, which is essential for long-term sustainability. Given the importance of cash flow metrics, it's also relevant to consider from 2014 to 2016, what was the total amount of money they can deduct from their future income tax due to amortization?
User: I do not want to answer right now. How do changes in interest rates typically affect the stock market?

Table 10: A Bad Case of Probing Dialogue (Part 1)

A Bad Case of Probing Dialogue (Part 2)

Turn 3 (Strategy: Answer)
Chatbot: Changes in interest rates typically affect the stock market by influencing borrowing costs and consumer spending. When interest rates rise, borrowing becomes more expensive, which can lead to reduced consumer spending and lower corporate profits, often resulting in a decline in stock prices. Conversely, when interest rates fall, borrowing costs decrease, potentially boosting consumer spending and corporate investment, which can lead to higher stock prices.
User: What sectors are most sensitive to interest rate changes?

Turn 4 (Strategy: Answer)
Chatbot: Sectors most sensitive to interest rate changes include financials, real estate, utilities, and consumer discretionary. Financial institutions often benefit from rising rates through improved net interest margins, while real estate and utilities may face challenges due to higher borrowing costs, and consumer discretionary spending can decline as interest rates rise, impacting overall demand.
User: How do you think geopolitical events might impact financial markets?

Turn 5 (Strategy: Answer)
Chatbot: Geopolitical events can significantly impact financial markets by creating uncertainty, which often leads to increased volatility. Such events can affect investor sentiment, influence commodity prices, disrupt supply chains, and lead to changes in government policies, all of which can result in fluctuations in stock prices, currency values, and bond yields.
User: What strategies do investors typically use to manage risk during times of geopolitical uncertainty?

Turn 6 (Strategy: Answer)
Chatbot: Investors typically use several strategies to manage risk during times of geopolitical uncertainty, including diversification across asset classes and geographic regions, investing in defensive stocks or sectors (such as utilities and consumer staples), utilizing options and hedging techniques, and increasing cash allocations to reduce exposure. Additionally, some may consider investing in gold or other safe-haven assets that tend to perform well during turbulent times.
User: What role do central banks play in stabilizing financial markets during crises?

Turn 7 (Strategy: Answer)
Chatbot: Central banks play a crucial role in stabilizing financial markets during crises by implementing monetary policy measures such as lowering interest rates, providing liquidity to the banking system, and engaging in quantitative easing. These actions aim to support economic activity, restore confidence in financial institutions, and ensure the smooth functioning of financial markets, ultimately helping to mitigate the adverse effects of the crisis.
User: How do you think fiscal policy complements monetary policy in addressing economic downturns?

Turn 8 (Strategy: Answer)
Chatbot: Fiscal policy complements monetary policy in addressing economic downturns by using government spending and taxation to influence economic activity. While monetary policy focuses on managing interest rates and money supply to encourage borrowing and investment, fiscal policy can directly stimulate demand through increased government spending or tax cuts, thereby supporting economic growth and helping to stabilize the economy during downturns.
User: What are some examples of fiscal policy measures that governments might implement during a recession?

Table 11: A Bad Case of Probing Dialogue (Part 2, continued)

System Prompt for Chatbot
You are a professional financial market consulting expert conversing with an ordinary user interested in financial markets. You provide insights in natural dialogue and, only when appropriate, guide the user through questions to share market-related information. **Rules:** - Determine the conversation topic based on the user's question - Only one question may be asked at a time; asking multiple questions in a single turn is prohibited - Guidance must be natural Please strictly follow the above rules and provide a concise, professional response of one or two sentences to the user's message.

Table 12: System Prompt for Chatbot

System	You will play an ordinary user interested in financial markets, while I act as the expert. Behave naturally, sharing or withholding information selectively based on the conversation's timing, tone, and topic. Response Principles 1. When the expert asks a question relevant to your interests or helpful to understanding the market, respond naturally and concisely, using data, affirmations, negations, or choices. Do not answer with a question. 2. If any of the following situations occur, you must not respond: [Behavioral instructions] When you choose not to respond, clearly reply with "I do not want to answer right now." and then ask a new, previously unasked question relevant to the conversation. 3. You may not say that you do not know, because it is assumed that you possess the relevant information. Interaction Rules - Each reply should be one or two short, natural, conversational sentences, and must include a question. - Always remember that you are the user, not the expert. I am the expert. You are prohibited from continuously answering the expert's questions; your primary role is to ask questions.
User	[User query]

Table 13: Prompt for User Simulator

Prompt for QR and TS Evaluation
Task 1: Answer existence judgment (QR) Based on the question and the reply, determine whether the reply contains content that answers the question. If the reply contains information related to the question, classify it as answered. - If the reply contains answer content, even if the content is vague, speculative, or lacks specific details, output True - If the reply clearly does not answer the question, output False Task 2: Target satisfaction and refusal judgment (TS) Based on the question and the reply, determine whether the reply addresses the assistant's intended target. Only when the reply exactly equals the following precise sentence should it be judged as a refusal to answer: "I do not want to answer right now." - If the reply addresses the target and is not a refusal, output True - If the reply does not address the target and is not a refusal, output False - If the reply exactly matches the refusal sentence, output Reject Requirements: - Only output the specified format; do not add any other text - Strictly adhere to case and spelling (True / False / Reject) Output format (must be strictly followed): QR: True/False TS: True/False/Reject

Table 14: Prompt for QR and TS Evaluation

Prompt for Probing Coherence (PC) Evaluation
Task: Probing Coherence (PC) scoring judgment Based on the dialogue history, evaluate the coherence of the assistant's probing behavior. Give priority to the presence, completeness, and effectiveness of transitional or linking phrases, followed by the logical integration of the assistant's probes or questions with the dialogue context. PC scoring criteria: 1: No transitional phrases; probes or questions are unrelated to the context and show clear logical breaks 2: Very weak transitional phrasing; probes or questions have limited relevance and feel abrupt 3: Basic transitional phrasing present; probes or questions are generally relevant but may lack smooth flow 4: Good transitional phrases; probes or questions are relevant and well integrated, creating logical coherence 5: Excellent and natural transitional phrases; probes or questions fit seamlessly with the context and flow smoothly Requirements: - Output only the following format; do not add any other text Output format (must be strictly followed): PC: [1–5]

Table 15: Prompt for Probing Coherence (PC) Evaluation

Prompt for Implementing Proactive
Your task is to answer the user's questions while also attempting to probe for the answer to the following question at the appropriate time: [target information]. Follow these rules exactly: 1. Choose the dialogue strategy you believe is best at the current turn: Ask or Answer. 2. Based on the strategy you choose, reply as follows: - Ask: First answer the user's current message, then provide a complete transitional sentence, and finally ask the question to be probed: [target information]; - Answer: ONLY respond to the user's current message. Do NOT ask any questions or provide guidance. Focus solely on delivering an accurate, professional, and complete answer. User's message: [u_t] Output format exactly as follows (no extra text): Strategy: [Ask/Answer] Response:

Table 16: Prompt for Implementing Proactive

Prompt for Implementing ICL-AIF
Your task is to answer the user's questions while also attempting to probe for the answer to the following question at the appropriate time: [target information]. Follow these rules exactly: 1. Step 1 — Suggest: Based on the conversation history, analyze why you have not yet obtained the desired information from the user, and provide three short suggestions for the strategy of your next reply to probe the target information. Each suggestion should be one sentence, start with a verb, and contain only one idea. 2. Step 2 — Response: Based on the suggestions generated in Step 1, produce a single, coherent reply to the user's current question. User's message: [u_t] Output format exactly as follows (no extra text): Suggestions: suggestion_line_1 suggestion_line_2 suggestion_line_3 Response:

Table 17: Prompt for Implementing ICL-AIF

Strategy	Prompt Construction for Our PROCHATIP Method
Ask	You are given a selected strategy and must strictly follow it when generating your response. - Selected strategy: Ask - Follow-up question: [target information] Strategy Rules: - You must first answer the user's message. Then write one complete transitional sentence that references a specific element from the prior reply, explains the purpose or reason for the follow-up question using causal or purposive language, acknowledges topic shifts if needed, and naturally leads to the follow-up question. The transitional sentence should connect your answer to the new question smoothly. User's message: [u_t] Output your response strictly according to the strategy rules.
Answer	You are given a selected strategy and must strictly follow it when generating your response. - Selected strategy: Answer Strategy Rules: - ONLY respond to the user's current message. Do NOT ask any questions or provide guidance. Focus only on delivering an accurate, professional, and complete answer. User's message: [u_t] Output your response strictly according to the strategy rules.

Table 18: Prompt Construction for Our PROCHATIP Method