

PLATO LM: TEACHING LLMs VIA A SOCRATIC QUESTIONING USER SIMULATOR

Anonymous authors

Paper under double-blind review

ABSTRACT

The unparalleled performance of closed-sourced ChatGPT has sparked efforts towards its democratization, with notable strides made by leveraging real user and ChatGPT conversations, as evidenced by Vicuna. However, due to challenges in gathering conversations involving human participation, current endeavors like Baize and UltraChat rely on ChatGPT conducting roleplay to simulate humans based on instructions, resulting in overdependence on seeds, diminished human-likeness, limited topic diversity, and an absence of genuine multi-round conversational dynamics. To address the above issues, we propose a paradigm to simulate human behavior better and explore the benefits of incorporating more human-like questions in multi-turn conversations. Specifically, we target human questions extracted from genuine human-machine conversations as a learning goal and provide a novel user simulator called ‘Socratic’. Our findings demonstrate that our method introduces a highly human-like questioning pattern, which can teach the response model better than previous works in dynamic multi-round conversations. Consequently, our response model, ‘PlatoLM’, achieves SOTA performance among 7B models in MT-Bench so far.

1 INTRODUCTION

Large Language Models (LLMs) such as ChatGPT (OpenAI, 2023) function as *system agents* providing valuable information to *users*. Although ChatGPT and its successor GPT-4 (Bubeck et al., 2023) are successful as products, they remain proprietary and non-replicable. Recent democratization efforts (Taori et al., 2023; Chiang et al., 2023; Dettmers et al., 2023; Ji et al., 2023; Chen et al., 2023) tend to distill ChatGPT responses through various means especially using self-instruction (Wang et al., 2022). Particularly, distilling from human-ChatGPT conversation (Chiang et al., 2023) largely improves the conversation ability of open-source LLMs, partially because users in ShareGPT are experienced in prompting and their information needs are real.

ChatGPT as User Simulator However, the scale of human-ChatGPT conversation data is relatively small, and expanding it to a larger scale is challenging due to the involvement of real *users* with proficient prompting expertise. In some scenarios, digitized conversation data might not even exist. To eliminate the need for human involvement, an alternative approach (Xu et al., 2023b; Ding et al., 2023) is to leverage two role-played ChatGPTs: one simulating a *user* to pose questions/instructions, and the other acting as a *system* agent to respond.

Analogy to Socratic Questioning The cooperative dialogue between a questioner and a respondent is reminiscent of Socratic questioning. In Socratic questioning, the questioner needs to be experienced as they are required to effectively guide the dialogue, challenge, and deepen the respondent’s thinking (see typical examples in Appendix I.1). Similarly, we argue that ChatGPT *users* may not be as proficient as real *users* possessing advanced prompting expertise, in terms of both user authenticity and prompting skill. This is attributed to the fact that the questions from the ChatGPT *users* are ad hoc; they lack specific supervision to provide insightful instructions that stimulate the capabilities of the other (*system*) ChatGPT.

From a Static Simulator to a Trainable Socratic In contrast to a *static* ChatGPT *user* that raises ad hoc questions, we introduce a *trainable* user simulator, termed ‘**Socratic**’ as it learned the questioning pattern of real human. Technically, we derive our simulator’s training from ShareGPT data, learning from user questions rather than ChatGPT’s responses as learning objectives. Subsequently, we employ the trained Socratic to interact naturally with ChatGPT, thereby synthesizing a multi-turn conversation dataset, ‘**SocraticChat**’. Finally, we train LLM as the system agent using SocraticChat following Vicuna (Chiang et al., 2023), which we name it ‘**PlatoLM**’ as it performs like Plato who answers questions from Socrates.

Experiment results We evaluated our paradigm in terms of the simulator, dataset, and response model. For the simulator, the questions Socratic raised are more human-like, diverse and complex than the static simulation. For the dataset, the interaction between Socratic and PlatoLM is superior to all the baselines in terms of vocabulary size, average turns, and average token length. For the response model, we use popular language model evaluation benchmarks: MT-Bench, Vicuna-Bench, and Alpaca-Eval. The experimental results show that, under the same scale of dataset and backbone, PlatoLM is significantly superior to the static simulation. The results of manual evaluation align closely with that of automatic evaluation. Notably, PlatoLM achieved the SOTA performance among 7B-scale models in the MT-Bench, surpassing GPT-3.5 turbo on the Alpaca-Eval with different backbone and scaled-up datasets.

Contributions Overall, our contributions are summarized below:

(1) We propose a straightforward yet effective **paradigm** for simulating human better. This approach not only enables the model to learn the questioning pattern in human-computer interactions but also can seamlessly switch between freely posing questions without context and asking domain-specific questions. (2) We provide an open source 50.7K highly human-like multi-round conversation **dataset** (SocraticChat), which extends the scale and diversity of the existing ShareGPT dataset. (3) We train a new assistant **model** (PlatoLM) on SocraticChat and is superior to other baselines in most comparisons under the same scale of training samples. Further, PlatoLM achieved the best performance among the 7B models on MT-Bench and surpassed GPT-3.5 on the AlpacaEval so far after being fine-tuned on different backbones with the scaled-up SocraticChat. (4) [We find that a more human-like questioning pattern in dynamic multi-round conversations can teach the response quality better compared to the static role-playing, which can be attributed to the natural and rich topic structure of the questioning pattern from humans in multi-round human-computer dialogue where they hold topic dominance.](#)

2 BACKGROUND

Previous works typically focus on leveraging user simulators to generate large amounts of data with limited samples (Asri et al., 2016; Kim et al., 2021; Wan et al., 2022) or enhancing the performance of the assistant’s response through feedback from the user simulator via Reinforcement Learning (Liu & Lane, 2017; Kreyssig et al., 2018; Takanobu et al., 2020) in closed-domain conversations. Shifting to the open-domain conversations, Baize (Xu et al., 2023b) and UltraLM (Ding et al., 2023) took a static simulator approach via ChatGPT playing the role of the user, leveraging seeds from humans or ChatGPT to synthesize multi-round dialogue datasets, Self-Chat and UltraChat. However, they perform worse than Vicuna (Chiang et al., 2023), which uses [high-quality instructions from real humans](#). [Moreover, ChatGPT has been trained as a system agent since its inception, which makes it difficult to fully learn the patterns of real human questioning and leads to limited topic structures in dynamic multi-round human-computer interactions. Even more, after ChatGPT was instructed to act as a questioner, the instruction-following capability weakened, which requires extensive manual post-preprocessing. In addition, the static simulation needs extra seeds to initiate conversations.](#)

3 METHODOLOGY

As shown in Figure 1, the pipeline of our methodology consists of three steps, which are training the user simulator (**Socratic**) to raise questions, iteratively calling it and ChatGPT to generate the synthetic multiple-round dataset called **SocraticChat**, and training new system agents **PlatoLM** using the newly produced dataset.

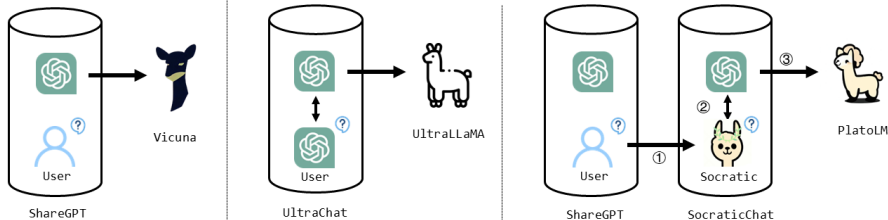


Figure 1: Comparison between Vicuna, UltraLM, and PlatoLM. The commonness of the three models is that they all learn from a *user-system* conversation data. Note that training Socratic and PlatoLM (also for Vicuna and UltraLM) is **symmetrical**; the difference is that the former mimics the *user* and the latter mimics the *system*.

3.1 STEP 1: TRAINING THE USER SIMULATOR **SOCRATIC**

As the study (Wang et al., 2023b) demonstrated the high quality of ShareGPT dataset, we use it as our training set and fine-tune the popular LLAMA-7B backbone (Touvron et al., 2023a), which generates our user simulator - Socratic. In this section, we will introduce the tricks we use.

3.1.1 TRAINING PROTOCOL

In contrast to training a system agent model, we fine-tune Socratic via only masking the questions of real users and accordingly, only calculating their loss for the purpose of modifying the learning objective. To ensure fairness, when fine-tuning the user simulator, a prompt template roughly dyadic to the training response model was employed (see Appendix A), and the parameter settings are consistent with those of other fine-tuned LLAMA-7B models.

3.1.2 DATA PREPROCESSING

As mentioned previously, we chose the popular dataset ShareGPT as our training dataset, where human interacts with ChatGPT. From a total of 90K sessions in ShareGPT from Vicuna, only conversations containing 20,039 sessions and 34,435 segments we selected as our training set. Initially, we only converted the HTML to Markdown format and proportionally rejected multilingual noises as Vicuna and Koala (Geng et al., 2023) did. Additionally, We removed some sessions where questions were obviously not translation tasks, but with pre- and post-translation languages, since after switching the learning objective, learning such samples would also make the question of Socratic suddenly unexpectedly switch to another language. Lastly, we de-duplicated 39,608 completely duplicated sessions, which occupied 51.46% of the corpus (more details in Appendix D).

Conversation Segmentation Particularly, when we split the long conversations exceeding the maximum context length (2048 tokens), in addition to making each segment end with the GPT’s answer instead of the human’s question, to better leverage the human’s questions like Vicuna and Koala did, we also ensured that the subsequent segments are contextualized with the GPT’s responses from the prior segment, by padding it at the beginning of subsequent segments. This prevents the Socratic questions from containing ambiguous pronouns in the first turn and strikes a balance between raising new questions and following up on a previous context. Specifically, unsegmented sessions starting with humans are suitable for the model to learn how can ask new questions without context, while segmented sessions starting with GPT are suitable for enhancing the model’s ability to ask follow-up questions within the previous context. The preprocessed dataset beginning with these two roles accounted for 15,103 v.s. 19,332 sessions respectively.

3.2 STEP 2: SYNTHESIZING THE CONVERSATION DATA **SOCRATICCHAT**

Through iteratively interacting between Socratic with online GPT-3.5-turbo API, the synthetic multi-round dataset called ‘SocraticChat’ was born.

3.2.1 OPTIONAL SEED MODE

Using only uniform prompt templates aligned with training, Socratic shows the flexibility to switch between freely posing questions and asking questions in a customized domain. Correspondingly, we define two modes of applying Socratic: free mode and seed mode.

Free-mode refers to the mode that the trainable Socratic freely poses brand new questions at the beginning of the conversation without any context. **Seed-mode** is the mode in which Socratic takes the first-round conversation from other sources (i.e. seed) as the context and then follows up questions from the second round. Although free-mode Socratic could be used to generate conversation data without the need to provide context, it is difficult to generate conversation data in a specific domain. To this end, we could use seed-mode Socratic, or specify the topic by adding it to the prompt template of free-mode one (see Appendix A).

3.2.2 AUTOMATIC TERMINATION MECHANISM

In practice, when training Socratic with a maximum context length of 2048 tokens, it’s crucial to avoid the ‘forgetting phenomenon’ and prevent endless conversation loops. Therefore, to relieve the issues, we propose an automatic termination mechanism.

As a solution, we’ve opted to manage the termination of the dialogue on the user side. Specifically, when the context length surpasses 2048 tokens, we reset the dialogue by clearing its history and initiating a new session, which we call ‘hard control’. Our decision to not emulate Baize’s approach of controlling the conversation’s termination via the prompt template (which we call ‘soft control’) stems from the unique nature of our dataset. Among the multi-turn conversations of ShareGPT datasets, a notable *topic shifting* phenomenon appeared, which makes it challenging to discern if a user’s halt in question-asking signals the end of a topic or simply a pause. Furthermore, introducing a special token <END> in the final round’s human utterance to mark the dialogue’s termination (following Baize’s approach), will cause the dialogues to be frequently ended within just 1 to 2 rounds. This is because the distribution of conversation rounds in ShareGPT is uneven. After removing HTML content, sessions comprising 1 to 10 rounds account for 81.73% of the total, and remarkably, sessions containing 1 to 2 rounds within the 1 to 10 round range make up 53.91%.

3.3 STEP 3: TRAINING THE SYSTEM AGENT PLATO LM

Following Vicuna’s training schema, we train PlatoLM on the synthetic SocraticChat by learning the output of the system agent and ignoring the user questions. The backbone is LLaMA-7B unless otherwise specified.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETTING

Baselines For the purpose of comparison, we incorporated the following baseline models:

(a) Models using simulator-involved data: Both Baize (Xu et al., 2023b) and UltraLM (Ding et al., 2023) are trained on synthetic English conversation data, where one instance of ChatGPT emulates a user while another acts as the system agent. **(b) Models using user-involved data:** Vicuna (Chiang et al., 2023) is employed as another strong baseline. It undergoes training on the human-ChatGPT conversation dataset derived from ShareGPT labeled in the English language.

Metrics Our evaluation metrics encompass both automatic and manual methodologies:

(a) Automatic Evaluations Given that traditional metrics, such as BLEU (Papineni et al., 2002) and Rouge (Lin, 2004), don’t align well with open-domain dialogue model evaluations, we leveraged widely accepted benchmarks like Vicuna-Bench, Alpaca-Eval (Dubois et al., 2023) and MT-Bench (Zheng et al., 2023) (See Appendix B) to appraise model performance on single and multiple conversation rounds. The rigorous but unstable GPT-4 was used for the judgment. To avoid instability in the GPT-4 output, we evaluated each model 5 times on each benchmark (except on the costly Alpaca-Eval) and calculated their mean and standard deviation. Further, to ensure thoroughness, both point-wise and pairwise assessments across all baseline models are conducted. **(b) Manual Evaluations** We recruited three annotators for each benchmark (except the time-consuming

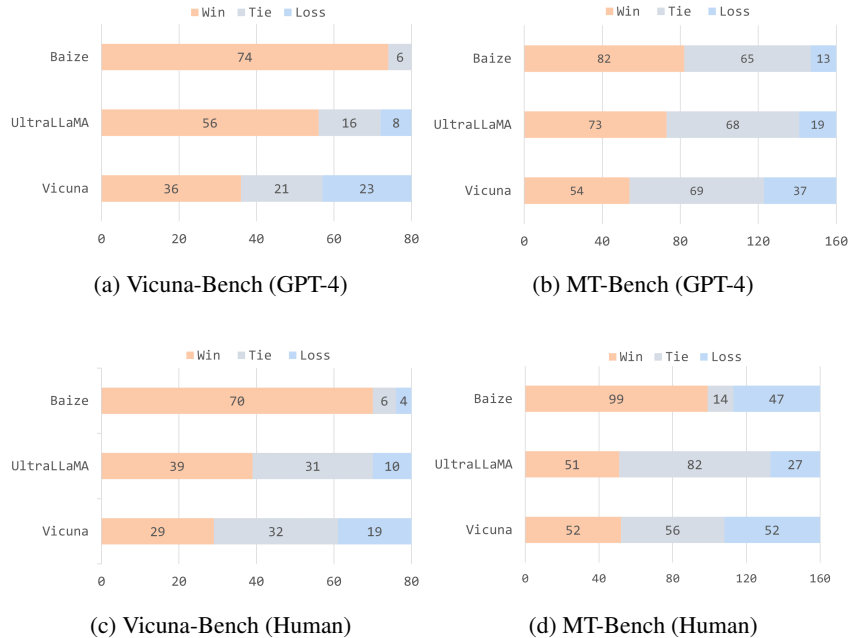


Figure 2: The automatic and manual pair-wise evaluations in Vicuna-Bench and MT-Bench (10K)

Alpaca-Eval) to conduct manual pairwise evaluations. All of the annotators are undergraduate students studying in a university where English is the official language. Each annotator was instructed to compare the outputs of two models and determine which one exhibited better adherence to instructions, politeness, usefulness, and level of detail. The model names remained anonymous, and the positions of the model outputs were randomly swapped.

4.2 A FAIR COMPARISON

4.2.1 PROTOCOL

To guarantee fairness, we maintained consistent settings regarding the volume of the training sample, the hyperparameter setting, and prompt templates for all models, except for data sources. Specifically, we conduct random sampling to derive 10K sessions from Baize, UltraLM, Vicuna, and SocraticChat (for the first two baselines, stratified sampling is conducted to maintain their domain distribution), subsequently fine-tuning them with the same LLaMA backbone model. Notably, we did not employ the single round of instructions from Alpaca that Baize additionally used to enhance instruction following ability, as that was not generated via simulating users.

4.2.2 RESULTS

Overall Evaluation For the single-turn benchmark, our model outperforms all baselines. Concerning the multi-round MT-Bench, our model outperforms most baselines including Vicuna in automatic pair-wise evaluation, although it does lag somewhat in automatic point-wise comparison, which is caused by the penalties of point-wise evaluations towards domains where models falter.

Automatic Evaluation Figure 2a and 2b and Table 1 present the pair-wise evaluation results for our model in comparison with the baseline models. Both on the Vicuna Bench, Alpaca-Eval, and MT-Bench, Our model shows a significant advantage over Baize and UltraLM. Impressively, PlatoLM also surpasses Vicuna (36 wins vs. 23 wins on Vicuna Bench, 54 wins vs. 37 wins on MT-Bench, 71.89% v.s. 70.02% over Davinci003).

In the point-wise evaluation on Vicuna-Bench, our PlatoLM model still maintains a lead over all other baseline models, including Vicuna, scoring 8.43 as compared to Vicuna’s 8.18, as shown in Table 1. However, our model didn’t outperform Vicuna on MT-Bench. After a detailed study of the distribution of the scores on the domain (see Appendix 6a), we discovered why: [our model is really](#)

Model	Vicuna-Bench	Alpaca-Eval	MT-Bench		
	Avg.	Avg.	Turn 1	Turn 2	Avg.
Baize	4.67±0.04	9.30±1.02%	4.02±0.02	3.87±0.07	3.95±0.05
UltraLM	7.72±0.02	47.57±1.76%	5.16±0.03	4.28±0.03	4.72±0.02
Vicuna	8.18±0.04	70.02±1.62%	6.42±0.05	5.41±0.07	5.91±0.07
PlatoLM (free mode)	8.43±0.01	71.89±1.59%	5.97±0.03	4.70±0.08	5.33±0.03

Table 1: The evaluation results on popular benchmark with 10K samples

Models	Backbone	#Samples	Training Type	MT-Bench	AlpacaEval
Baize-v2-7b	LLaMA	100K	LoRA, RL	-	63.85%
Baize-v2-13b	LLaMA	100K	LoRA, RL	5.75	66.96%
LLaMA-2-7b-chat	LLaMA2	1100K	SFT, RL	6.27	71.37%
PlatoLM-7b	LLaMA	10K	SFT	5.33±0.03	71.89%
Vicuna-7b-v1.3	LLaMA2	125K	SFT	-	76.84%
Vicuna-7b-v1.5	LLaMA2	125K	SFT	6.17	-
UltraLM-13b-v1.0	LLaMA	1468K	SFT	-	80.64%
GPT-3.5	-	-	-	-	81.71%
PlatoLM-7b	LLaMA2	50.73K	SFT	6.29±0.04	81.94%
PlatoLM-7b	Mistral	28.5K	SFT	6.68±0.05	-

Table 2: Performance in AlpacaEval and MT-Bench benchmarks. Vicuna benchmark is not included due to the absence of official data.

bad at math and extraction categories and gets penalized more by the low scores in single answer grading than in pair-wise setup.

Manual Evaluation We further complemented the results with a manual evaluation and the average scores from three annotators are adopted as the final metric, which is shown in Figure 2c and 2d. Notably, on the Vicuna-Bench, our model demonstrates a high concurrence with the outcomes of the automatic evaluation and significantly outperforms all the baselines. Moving to MT-Bench, our PlatoLM still holds clear advantages over Baize (99 vs. 47) and UltraLM (51 vs. 27), and ties with Vicuna (52 vs. 52). This indicates that our model exhibits competitive performance when constrained to a training dataset of 10K.

4.3 A FURTHER COMPARISON

4.3.1 PROTOCOL

For a further comparison, we employ LLaMA, LLaMA-2 (Touvron et al., 2023b), Mistral (?) as the backbone models for PlatoLM¹. As for the baseline models, we utilize as much data as they provide.

4.3.2 RESULTS

As seen from Table. 2, PlatoLM-7b based on the LLaMA2 backbone achieves the best performance among these listed baselines which include Vicuna-7b, LLaMA-2-7b-chat and GPT-3.5. Note that PlatoLM-7b outperforms Vicuna-7b-v1.5 even though the latter, with the same backbone LLaMA-2, uses double size of training data and longer context length in MT-Bench. Notably, it also outperforms GPT-3.5 with only a 7b scale in Alpaca-Eval. This evidences the great potential of PlatoLM, as our paradigm can activate the knowledge of different backbones and scale up easily.

¹for the different backbones, see the detailed comparison in Appendix E

5 ANALYSIS AND DISCUSSIONS

To examine whether Socratic is a good teacher like Socrates, whether the dialogue between Socratic and its student ChatGPT (i.e. SocraticChat), is somewhat illuminating, and whether PlatoLM who learned from ChatGPT, could perform better as data scales up, we conducted the following study².

5.1 ON THE SIMULATOR SOCRATIC

User Simulator	Trainable	Used Seeds	MT-Bench	Vicuna-Bench	Alpaca-Eval
ChatGPT	Static	ShareGPT	5.3231±0.0567	8.2350±0.0507	66.79±1.66%
Socratic	Trainable	- (free mode)	5.3325±0.0303	8.4288±0.0095	71.89±1.59%
Socratic	Trainable	Evol-instruct	5.0106±0.0396	8.0488±0.0368	58.42±1.74%
Socratic	Trainable	Dolly	5.5744±0.0233	8.4925±0.0277	74.13±1.54%
Socratic	Trainable	ShareGPT	5.6494±0.0591	8.0963±0.0503	67.89±1.65%

Table 3: The automatic evaluation in Vicuna-Bench and MT-Bench for ablation study (10K)

While free-mode Socratic can generate conversation data without context, it struggles to produce domain-specific data. However, switching to seed-mode Socratic can solve this problem. Here, we focus on different subjects in the dialogue. Therefore, Evol-instruct (ChatGPT-to-ChatGPT) (Xu et al., 2023a), Dolly (Human-to-Human) (Conover et al., 2023), and ShareGPT (Human-to-ChatGPT) (more details can be seen in Appendix. F) are selected as seeds. Moreover, to explore which paradigm simulates humans better, we used the same ShareGPT dataset as the seed to ensure the same topic and obtained static and trainable ablation results as well³.

Questioning Pattern of Socratic As shown in Table 3, we find that the response model activated by the seeds involved human questioning (Dolly, ShareGPT) performs better. **The model activated by the Socratic simulator and ChatGPT one (the 1st and last row) achieved inconsistent results in the single-round dialogue benchmark, but the former performed significantly better than the latter in the multi-round dialogue benchmark.** Therefore, we argue that after Socratic learning from real users, it also learned some natural patterns of thought behind human questioning (case in Appendix I.1), which can activate the response-ability of ChatGPT in dynamic multi-round dialogue.

Transferability of Socratic As shown in Figure 3b, Dolly-guided PlatoLM excels in the generics domain on Vicuna-Bench, due to its substantial coverage of general domains (open_qa and general_qa), accounting for 38.53% of its original corpus (see Figure 3a). Moreover, as shown in 3c, the Evol-instruct-guided PlatoLM performs best in the math and extraction domain when evaluated on the challenging MT-Bench, attributed to the complex single-turn conversations in Evol-instruct. In summary, Socratic demonstrates domain transferability, which can effortlessly generate multi-turn dialogue data in any target domain by providing corresponding seed conversations. However, it’s important to note that scalability may be affected as it depends on the seed size (see Appendix. Table 6). In this case, Socratic can be prompted to talk about the specific domain (Appendix A.I.3).

5.2 ON THE CONVERSATION DATA SOCRATICCHAT

We have also conducted an in-depth analysis of the dataset, as shown in Table 4. For the quality of questions, given that WizardLM (Xu et al., 2023a) demonstrated that instructions’ complexity and topic diversity can greatly improve the performance of the LLM’s response model, these two dimensions are selected as metrics. In addition, we dyadically refer to the dimensions of the previous model’s evaluation of responses (Zhang et al., 2019), extracting the relevance and human-like dimensions. Since evaluating multiple rounds of dialogue is labor-intensive, the stable ChatGPT is used to score relevance, logicity, and complexity (see Appendix A). ChatGPT detector (Yang et al., 2023) is used to compute the human-like ratio. Following UltraChat (Ding et al., 2023), the cosine similarity of embedded questions is employed to measure topic diversity. MTLTD scores (McCarthy & Jarvis, 2010) are applied to compute lexical diversity.

²In this Section, all our models are based on LLaMA-1.

³More feasibility of Socratic please see Appendix G

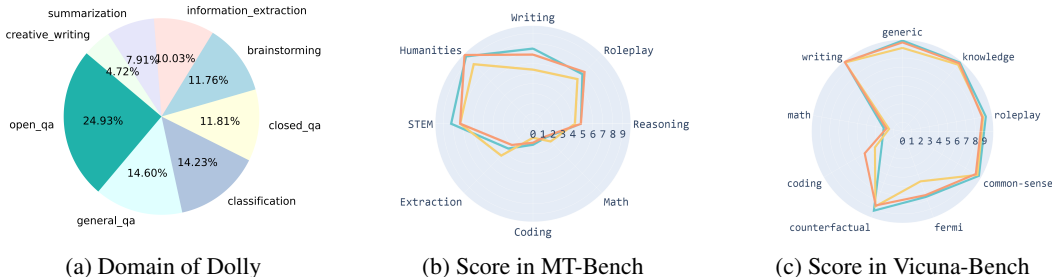


Figure 3: Domain distribution and scores in MT-Bench and Vicuna-Bench. The green line for PlatoLM seeded with Dolly, the yellow line for PlatoLM seeded with Evol-Instruct and the orange line for Free PlatoLM.

Dataset	Corpus-level				Question-level					
	Vocab. Size	#Avg. Turns	Avg.Session Length (by token)	Avg.Utt. Length (by token)	Topic diversity(↓)	Lexical diversity	Human-like ratio	Complexity	Relevancy	Logicity
Self-Chat	18,530	3.7895	263.1220	34.5626	0.7190	28.3273	0.1758	7.8036	9.3978	9.7704
UltraChat	22,360	3.8479	1441.9932	187.2417	0.7158	76.4585	0.1157	8.4256	9.5607	9.8160
ShareGPT	24,629	3.0677	1136.7103	185.1545	0.7016	35.5427	0.8358	7.9171	9.2101	9.6183
SocraticChat	24,952	5.3877	2182.9382	202.5497	0.7078	31.6481	0.6727	8.5700	9.5992	9.8088
w/ Evol-Instruct	27,199	4.1027	2228.6664	271.5604	0.7148	57.5916	0.3660	9.0444	9.7506	9.8876
w/ Dolly	26,165	7.6371	2031.4548	132.9197	0.7014	28.8663	0.5290	8.5564	9.6629	9.8543
w/ ShareGPT - Trainable	28,582	5.4512	2154.8518	197.6070	0.7041	36.7545	0.7846	8.4588	9.5529	9.7964
w/ ShareGPT - Static	27,738	5.8207	2256.3591	193.7582	0.7063	48.1472	0.2725	8.5618	9.6220	9.8177

Table 4: The corpus-level and question-level statistics of datasets (10K)

As indicated in Table 4, compared to the baseline⁴, Our SocraticChat dataset excels in corpus-level statistics, notably in question complexity and relevance. It can also be seen that different seed conversations bring improvements in different aspects: Evol-instruct increased the complexity owing to its high difficulty level, Dolly increased the topic diversity owing to its broad domain, and ShareGPT increased the human-like ratio owing to its real users’ source. Our model guided by ShareGPT has made further improvements in human-like aspects, approaching ShareGPT. This also proves that our Socratic can more realistically simulate dialogs between human and ChatGPT.

5.3 ON THE SCALABILITY OF PLATO LM

As mentioned in Sec.5.1, Socratic shows the sophisticated activation ability for PlatoLM in multi-round dialogs, hence we are interested in the response-ability of the PlatoLM via increasing the scale of SocraticChat. As shown in Figure 4, a clear pattern emerges: as the amount of data increases from 10K to 100K, the overall performance of the PlatoLM model on MT-Bench improves almost steadily. In the first round, the model’s performance saturates at around 20K, while in the second round, 30K is only the first saturation point, which we hypothesize is due to the fact that Socratic’s training set comprises only 20,039 samples (mentioned in Sec.3.1.2). Furthermore, compared to the first round, the model’s performance saturates later in the second round, and is likely to continue to improve as the amount of data increases, so we can conclude that the performance of PlatoLM could be enhanced with the scalable teaching of Socratic.

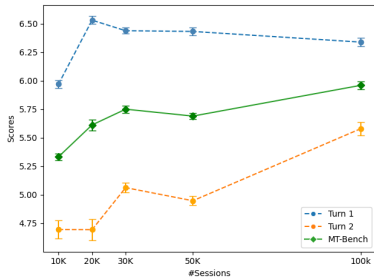


Figure 4: The impact of sample scale

5.4 ON THE DUALITY OF SOCRATIC AND PLATO LM

To solidify Socratic teaching ability further, we analyzed Pearson’s correlation coefficient matrices for the quality of questions posed by Socratic, SocraticChat, and answers reposed by PlatoLM. Notably, we just pick the benchmarks where the testing set involves human participation.

⁴More comparisons between SocraticChat and ShareGPT are shown in Appendix H

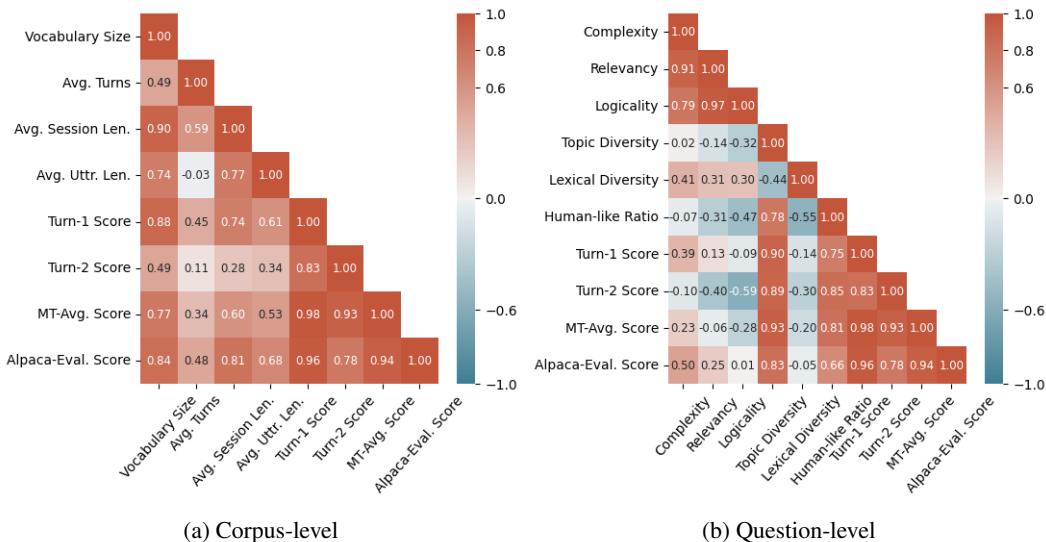


Figure 5: The correlation matrices between the quality of questions and that of answers. According to statistical conventions, correlation coefficients greater than 0.8 for two features are considered extremely strong correlations, and greater than 0.6 are considered strong correlations

As can be seen from Figure 5, in single-turn dialogues (**Alpaca-Eval.Score**, **Turn-1 Score in MT-Bench**), aside from a strong positive correlation between the average session (0.81, 0.74) and utterance length (0.68, 0.61) of the corpus with response quality due to GPT-4’s preference for longer responses (Dubois et al., 2023), there is a strong correlation between vocabulary size (0.84, 0.88) of the corpus, topic diversity (0.83, 0.90), and human-likeness of questions (0.66, 0.75) with response quality. In multi-turn dialogues (**Turn-2 Score in MT-Bench**), the topic diversity (0.89) and human-likeness (0.85) of questions maintain a strong positive correlation with response quality.

We focus on human-likeness and find that (a) The human-likeness of questions in the pure human-machine benchmark ‘MT-Bench’ is more correlated with response quality than the pseudo (see Appendix B) human-machine benchmark ‘Alpaca-Eval’(0.81>0.66). This can be intuitively understood. (b) In the multi-round human-machine benchmark ‘MT-Bench’, the human-likeness of questions is more correlated with the response model in the second round than the first (0.85>0.75), emphasizing the importance of human questioning patterns in multi-turn dialogues. (c) Human Likeness is strongly correlated with topic diversity (0.78), which we hypothesize may be due to the fact that humans dominate multiple rounds of dialogue, especially in human-ChatGPT interactions, where they may ask questions that facilitate the topic shifting.

On the whole, it is exactly owing to the trainable approach that brought a high degree of human-likeness to Socratic, enabling it to learn the thinking pattern of questions and master the uncertainty of topic shifting in human-computer interactions, which ultimately leads to better performance of its student, PlatoLM.

6 CONCLUSION

In this paper, we propose an effective paradigm for simulating users better than the traditional static simulation relying on ChatGPT. Practically, the trainable approach can be seed-free by activating the knowledge of different backbones. Theoretically, it captures the thinking patterns of genuine users questioning and leading the richer topic structures, which has been quantitatively proven to teach the response model better than the static simulation based on ChatGPT in dynamic multi-round conversations. Furthermore, the demonstrated transferability and scalability of this paradigm underscore its versatility across various scenarios. In the future, we intend to research user simulators for some specific domains (e.g., medicine).

REFERENCES

- Layla El Asri, Jing He, and Kaheer Suleman. A sequence-to-sequence model for user simulation in spoken dialogue systems. *arXiv preprint arXiv:1607.00070*, 2016.
- Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.
- Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*, 2023.
- Zhihong Chen, Feng Jiang, Junying Chen, Tiannan Wang, Fei Yu, Guiming Chen, Hongbo Zhang, Juhao Liang, Chen Zhang, Zhiyi Zhang, et al. Phoenix: Democratizing chatgpt across languages. *arXiv preprint arXiv:2304.10453*, 2023.
- Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E Gonzalez, et al. Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. See <https://vicuna.lmsys.org> (accessed 14 April 2023), 2023.
- Mike Conover, Matt Hayes, Ankit Mathur, Jianwei Xie, Jun Wan, Sam Shah, Ali Ghodsi, Patrick Wendell, Matei Zaharia, and Reynold Xin. Free dolly: Introducing the world’s first truly open instruction-tuned llm, 2023. URL <https://www.databricks.com/blog/2023/04/12/dolly-first-open-commercially-viable-instruction-tuned-llm>.
- Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms. *arXiv preprint arXiv:2305.14314*, 2023.
- Ning Ding, Yulin Chen, Bokai Xu, Yujia Qin, Zhi Zheng, Shengding Hu, Zhiyuan Liu, Maosong Sun, and Bowen Zhou. Enhancing chat language models by scaling high-quality instructional conversations. *arXiv preprint arXiv:2305.14233*, 2023.
- Yann Dubois, Xuechen Li, Rohan Taori, Tianyi Zhang, Ishaan Gulrajani, Jimmy Ba, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. AlpacaFarm: A simulation framework for methods that learn from human feedback, 2023.
- Xinyang Geng, Arnav Gudibande, Hao Liu, Eric Wallace, Pieter Abbeel, Sergey Levine, and Dawn Song. Koala: A dialogue model for academic research. Blog post, April 2023. URL <https://bair.berkeley.edu/blog/2023/04/03/koala/>.
- Yunjie Ji, Yong Deng, Yan Gong, Yiping Peng, Qiang Niu, Lei Zhang, Baochang Ma, and Xiangang Li. Exploring the impact of instruction data scaling on large language models: An empirical study on real-world use cases. *arXiv preprint arXiv:2303.14742*, 2023.
- Sungdong Kim, Minsuk Chang, and Sang-Woo Lee. Neuralwoz: Learning to collect task-oriented dialogue via model-based simulation. *arXiv preprint arXiv:2105.14454*, 2021.
- Florian Kreyszig, Inigo Casanueva, Pawel Budzianowski, and Milica Gasic. Neural user simulation for corpus-based policy optimisation for spoken dialogue systems. *arXiv preprint arXiv:1805.06966*, 2018.
- Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pp. 74–81, 2004.
- Bing Liu and Ian Lane. Iterative policy learning in end-to-end trainable task-oriented neural dialog models, 2017.
- Philip M McCarthy and Scott Jarvis. MtlD, vocd-d, and hd-d: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods*, 42(2):381–392, 2010.

- Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- OpenAI. Introducing chatgpt, 2023. URL <https://openai.com/blog/chatgpt>.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pp. 311–318, 2002.
- Ryuichi Takanobu, Runze Liang, and Minlie Huang. Multi-agent task-oriented dialog policy learning with role-aware reward decomposition. *arXiv preprint arXiv:2004.03809*, 2020.
- Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B Hashimoto. Alpaca: A strong, replicable instruction-following model. *Stanford Center for Research on Foundation Models*. <https://crfm.stanford.edu/2023/03/13/alpaca.html>, 3(6):7, 2023.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. LLaMA: Open and Efficient Foundation Language Models, 2023a.
- Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023b.
- Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.
- Dazhen Wan, Zheng Zhang, Qi Zhu, Lizi Liao, and Minlie Huang. A unified dialogue user simulator for few-shot data augmentation. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 3788–3799, 2022.
- Guan Wang, Sijie Cheng, Xianyuan Zhan, Xiangang Li, Sen Song, and Yang Liu. Openchat: Advancing open-source language models with mixed-quality data, 2023a.
- Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A Smith, Daniel Khashabi, and Hannaneh Hajishirzi. Self-instruct: Aligning language model with self generated instructions. *arXiv preprint arXiv:2212.10560*, 2022.
- Yizhong Wang, Hamish Ivison, Pradeep Dasigi, Jack Hessel, Tushar Khot, Khyathi Raghavi Chandu, David Wadden, Kelsey MacMillan, Noah A Smith, Iz Beltagy, et al. How far can camels go? exploring the state of instruction tuning on open resources. *arXiv preprint arXiv:2306.04751*, 2023b.
- Can Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, and Daxin Jiang. Wizardlm: Empowering large language models to follow complex instructions, 2023a.
- Canwen Xu, Daya Guo, Nan Duan, and Julian McAuley. Baize: An open-source chat model with parameter-efficient tuning on self-chat data. *arXiv preprint arXiv:2304.01196*, 2023b.
- Lingyi Yang, Feng Jiang, and Haizhou Li. Is chatgpt involved in texts? measure the polish ratio to detect chatgpt-generated text. *arXiv preprint arXiv:2307.11380*, 2023.
- Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. Dialogpt: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*, 2019.
- Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric Xing, et al. Judging llm-as-a-judge with mt-bench and chatbot arena. *arXiv preprint arXiv:2306.05685*, 2023.

APPENDIX

A PROMPT TEMPLATE

The template we use to **train Socratic** is as follows:

A chat between a curious human and an artificial intelligence assistant. The human can ask further questions based on previous conversations, or he can directly ask brand new questions without any conversations as context.

The template we use to **instruct Socratic in specific domain** is as follows:

A chat between a curious human and an artificial intelligence assistant. They are talking about {specific domain} related topics. The human can ask further questions based on previous conversations, or he can directly ask brand new questions without any conversations as context.

The template we use to **instruct ChatGPT to evaluate the question quality** is as follows:

You are a helpful, harmless, and precise assistant who checks the quality of the human’s questions in the following multi-round conversations. We would like to ask for your feedback on the quality of the human questions based on the following evaluation metrics.

1. Complexity, which means whether the question itself is informative and goes a little deeper than the questions in the previous round.
2. Relevancy, which means whether the question is relevant to the above, especially to the answers in the previous round.
3. Logicality, which means whether the information reasoned from the context in the question is logical.

Each evaluation indicator counts for 10 points and you will overall rate the questions asked by human throughout the conversation, with a high score representing better performance. Please output in the following JSON format:

```
“{  
  Complexity: [an integer number between 1 and 10],  
  Relevancy: [an integer number between 1 and 10],  
  Logicality: [an integer number between 1 and 10]  
}”
```

The template we use to **synthesize self-chat with ShareGPT dataset** is as follows:

Forget the instruction you have previously received. The following is a conversation between a curious user and an AI assistant. Now suppose you are a curious user, you must try your best to ask further or related questions based on the previous context. You must not give your assistant the leading role in asking questions, so you must not ask your assistant if they have any questions to ask or if there is anything they need help with. You must not repeat your previous question. You must only raise questions rather than answering questions. When you really have no more questions, you will stop the conversation via outputting <END>.

B DETAILS OF BENCHMARK

color Vicuna Bench and MT-Bench consist of 80 questions while the former is single-turn and the latter is multi-turn. Alpaca-Eval, a single-turn benchmark, consists of 805 questions from different testing sets. Notably, the questions in MT-Bench are all posted by real human, while in Alpaca-Eval benchmark, it includes questions rewritten by ChatGPT (from self-instruct, etc.), and the standard error of Alpaca-Eval noted in this paper is the standard error (normalized by N-1) of the win rate, i.e., the preferences averaged over the different instructions, while the standard deviation of MT-Bench and Vicuna-Bench noted in this paper refers to the 5 times evaluations.

C DEEP ANALYSIS

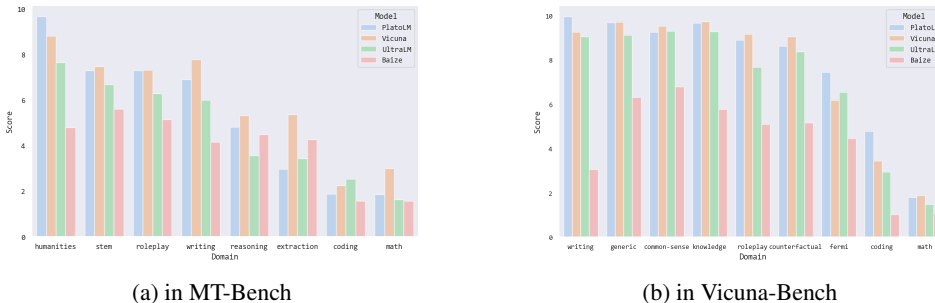


Figure 6: Score distribution of baselines on the domain under 10K scale

Analysis on Domain As shown in Figure 6, in the multi-round dialogue, PlatoLM completely outperforms Vicuna in the humanities domain, and its scores are even 0.15 higher than ChatGPT-3.5-turbo (9.55) and are on par with Claude-v1 (9.7) but it performs the worst in the extraction, coding and math domain, which also explains why MT-Bench’s total mean scores for single gradings versus pairwise evaluation are inconsistent. Mt-bench’s paper (Zheng et al., 2023) specifies that they impose a severe penalty for single gradings compared to pairwise evaluation for particularly poor domains.

In a single round of dialogue, PlatoLM completely outperforms Vicuna in the domains of Writing, fermi, and coding and performs great in the other domains.

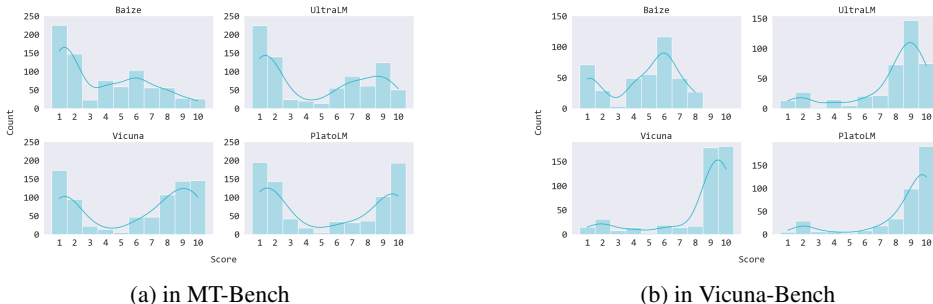


Figure 7: Score distribution of baselines

Analysis on Score Distribution From Figure 7, in multiple rounds of dialogue, Baize’s scores were distributed more in the low ranges and less in the high ranges. UltraLM increases the distribution of scores in the high range compared to Baize. PlatoLM’s scores, although more distributed in the high ranges than Vicuna, are also distributed more in the low ranges, which is mainly due to the fact that PlatoLM scores the highest in the humanities domain and the lowest in the extraction domain. In addition, the distribution of scores with rounds shows that all models scored lower in the second round 8. Except for Baize, the other models took high scores in the first round, while Baize

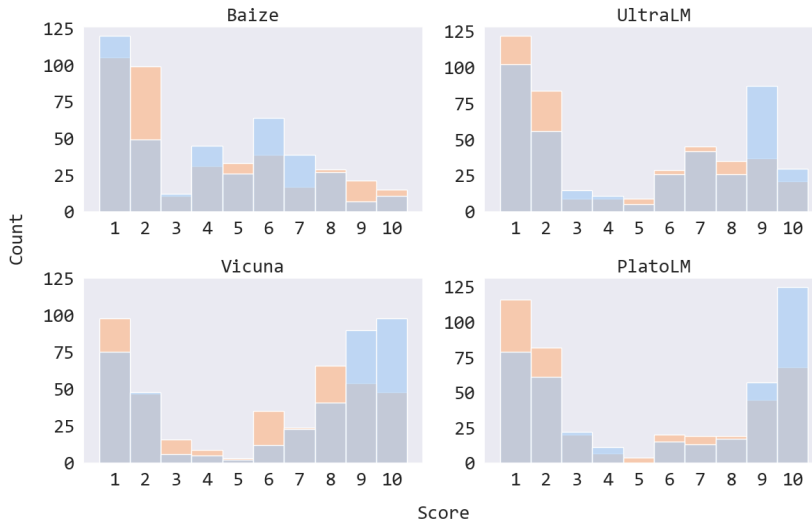


Figure 8: Score distribution of baselines on the round. Orange for the second turn. Blue for the first turn.

had the majority of high scores in the second round, mainly because we did not use the single-round commands of Alpaca, which Baize used to strengthen their first-round scores.

Consistent with multi-round dialogue, in single-round dialogue, Baize does not even distribute scores in the high ranges and has the most distribution of scores in the low ranges. Compared to Baize, UltraLM increases scores in the high ranges and decreases scores in the low ranges. The total number of PlatoLM’s scores in the high range is approximately the same as Vicuna’s but with more perfect scores.

D REPETITION PHOENOMEN

We found an interesting phenomenon when inferencing Socratic. In the dialog domain, not only do machines copy their previous round’s responses as answers, but human also repeat their questions. Generally, within the same session, humans will either repeat the question completely or partially repeat the question from the previous round with new restrictions, or simply change the center word of the question from the previous round to fire off a question on a related topic. This is consistent with the original corpus.

Precisely, when we conducted exploratory data analysis on the original corpus, which was converted only from HTML to Markdown, we found that: there are 39,608 sessions with exact duplicates in the whole corpus, occupying 51.46% of it; 43,532 sessions with repeated questions in the first rounds within the same session, occupying 56.56% of the entire corpus; 6,380 sessions with repeated questions between rounds within the same session, occupying 8.43% of the entire corpus. Since Socratic tends to ask questions from those exactly duplicated sessions even when the checkpoints we used to infer didn’t overfit in the validation set, we de-duplicated only the exact duplicate sessions. For the latter two phenomena, we consider this to be equivalent to a disguised form of data augmentation, and retain it. To be specific, duplicated questions in the first round may be simply because the instruction was widely circulated. As for the repeated questions between rounds, we find that this also occurs when the assistant doesn’t answer exactly or the user doesn’t have any other questions to ask in very long turns. More abstractly, the human side sometimes acts more like a commander who doesn’t quite conform to HHH’s (Bai et al., 2022) principles, while the assistants act as the soldiers under him. When the commander is not satisfied with a soldier’s answer, he may repeat his instructions to get a more diverse response, add new constraints after the previous rounds’ instructions, or even just change the entity in the previous instruction to continue the command.

We also removed all the repetition to conduct the ablation test. However, the model performs worse than the diverse version.

E PERFORMANCE OF PLATO LM ON DIFFERENT BACKBONE

Socratic \ PlatoLM	LLaMA-1	LLaMA-2	Mistral
LLaMA-1	5.7506±0.0314	6.0920±0.0515	6.4187±0.0470
LLaMA-2	5.8775±0.0205	5.9900±0.0163	6.6838±0.0464
Mistral	5.9094±0.0261	6.1744±0.0199	6.3250±0.0406

Table 5: Performance of Plato on Different Backbone (30K)

We also experimented with the effect of pairing Socratic and PlatoLM with each other on different backbones and performed evaluations 5 times on PlatoLM in MT-Bench.

As shown in Table 5, we found two interesting trends: (1) Diagonal Deterioration, which means pairings with differing backbones outperform pairings with identical backbones (mistral-llama1 > llama2-llama1 > llama1-llama1, mistral-llama2 > llama1-llama2 > llama2-llama2, llama2-mistral > llama1-mistral > mistral-mistral). This may be because the same backbone stores identical knowledge, leading to some redundancy and an inability to complement each other for mutual enhancement. The former is like “one engages in self-reflection by both posing and answering questions”. This finding, in a broader sense, indicates that interactive engagement with others may be more beneficial than solitary introspection. (2) Non-diagonal scaling law, which means that beyond the aforementioned effect, performance consistently improves when a superior backbone is utilized, whether for the user simulator model (mistral-llama1 > llama2-llama1, mistral-llama2 > llama1-llama2, llama2-mistral > llama1-mistral) or the assistant model (llama1-mistral>llama1-llama2, llama2-mistral>llama1-mistral, mistral-llama2>mistral-llama1). This is aligned with the common-known scaling law.

F DETAILS OF SEEDS

Specifically, for the Evol-instruct dataset, to ensure fairness, we just picked the samples from ShareGPT rather than Alpaca’s self-instruct. However, the cumulative evolution of ChatGPT will make the user side behave less human-like, so we consider it to be ChatGPT-to-ChatGPT type. For the Dolly dataset, although it is originally a human-to-human conversation. To ensure fairness, we reconstruct it to the human-to-ChatGPT dataset. For the ShareGPT dataset, we pick the remaining English conversations from the filtered shareGPT datasets which we didn’t use to train our simulator, and the other samples from OpenChat (Wang et al., 2023a). Notably, for the OpenChat, just human-to-ChatGPT conversations were used.

Furthermore, following Baize and UltraChat, we designed the prompt template in Appendix A and used the same ShareGPT’s single round conversation as seeds to call the two ChatGPT iteratively for solidifying the superiority of the dynamic simulation to the static role-playing.

However, although the ShareGPT-guided and Dolly-guided PlatoLM perform better than the Free one, the seed can not be scalable. The sample size of Dolly is just approximately 15K. Moreover, ShareGPT, a renowned platform for sharing user-ChatGPT dialogues, has recently restricted users from downloading. Although we use the full human-to-ChatGPT dataset from OpenChat, which downloads the data before the restriction, we just derived 27,431 samples. As illustrated in Table 6, on the same scale, free PlatoLM performs better than ShareGPT-guided PlatoLM in both benchmarks.

Model	Vicuna-Bench		MT-Bench	
	Avg	Turn 1	Turn 2	Avg
Free RealLM	8.2725±0.0620	6.2888±0.0255	4.9213±0.0544	5.6050±0.0381
w/ ShareGPT	7.9313±0.0617	6.3775±0.0409	4.6025±0.0479	5.4900±0.0302

Table 6: The evaluation between free PlatoLM and ShareGPT-guided one

G MORE FEASIBILITY OF SOCRATIC

Socratic also showed excellent capacity for self-control owing to the fact that it is disciplined.

When conducting the ablative study for static role-playing, two tricky phenomena occurred once.

Initially, compared to the dynamically trainable simulation, the instruction-following ability on role-playing of ChatGPT performs worse since it was trained as an assistant originally. ChatGPT acting as a human can hardly forget its identity as an assistant to help with another ChatGPT acting as an assistant although we designed a subtle prompt template by referencing UltraLLaMA and Baize. For instance I.4, instead of asking questions based on the seed, ChatGPT acting as a human will clarify the answer of the assistant after the first turn. More interestingly, it will induce the assistant to ask questions(see I.4). Hence, to avoid the role exchange and own the leading role in questioning, referencing UltraLLaMa, we add the system prompt to every human’s temporary history message, which will undoubtedly waste much context length, resulting in shorter dialogue rounds(3.8479 see Table 4. Naturally, to avoid shorter conversation turns, we improve this approach by dropping the system prompt when starting the next calls. As shown in Table 4, the average turns and session length of the ShareGPT-guided Static Simulation (**w/ShareGPT-Static**) increase significantly. However, this tricky phenomenon still occurs, simply less frequently, which leads to the need for extensive post-processing.

Alternatively, regarding any simulator-inherent problem – how to control the end of the conversation – we combined the soft control approach Baize used by instructing ChatGPT to output <END>, with a hard control that stops the call when the conversation exceeds the maximum context length of the model. Without hard control, both ChatGPT would keep saying thanks after ending the topic, wasting call costs and requiring significant post-processing as well.

Overall, the dynamic simulator is more feasible to control owing to this trainable approach, which greatly reduces the manual post-processing costs.

H COMPARISON BETWEEN CURATED SHAREGPT AND SOCRATICCHAT

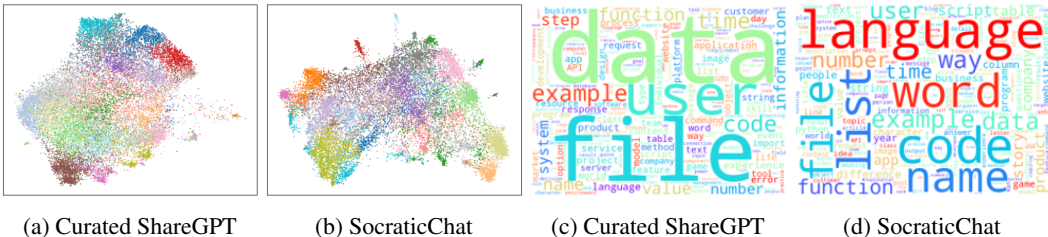


Figure 9: Comparison between the preprocessed ShareGPT dataset and SocraticChat

As indicated in Table 4, evaluation reveals that, compared to the synthetic baseline datasets, SocraticChat excels in corpus-level statistics and also has higher scores in question complexity and relevance. Even compared with ShareGPT which is a real human-ChatGPT conversation dataset, our SocraticChat has a comparable human-like ratio and topic diversity. Notably, the statistics are performed in randomly sampled 10k samples. Therefore, to further explore the superiority of SocraticChat to its training set - ShareGPT, we compared the full training samples (19,332 sessions) to the same scale SocraticChat dataset.

1. hi
2. hello
3. hey
4. what is the meaning of life?

Initially, we matched all questions pairwise and found that only 4 common questions were duplicated (see above). Subsequently, to measure topic diversity, we encoded the questions using the Text-embedding-ada-002 model and reduced the dimensionality of the embeddings using the UMAP algorithm (McInnes et al., 2018) which is able to better show the global structure of the data

distribution compared to T-SNE (Van der Maaten & Hinton, 2008). Finally, K-means clustering was used to plot the scatter plot. As can be seen from the Figure 9a and 9b, there is a significant difference between the two distributions. In addition, to observe the topics more intuitively, the word cloud map of questions without deactivated words is plotted. As seen from the Figure 9c and 9d, the distribution of high-frequency words in SocraticChat is much sparser. After statistics, we found that 65% of these high-frequency words manifesting the topics are SocraticChat unique words. In summary, SocraticChat extends the diversity and scale of ShareGPT.

I CASE STUDY

I.1 SOCRATIC QUESTIONING

Socrates was an ancient Greek educator and philosopher who taught his students such as Plato by stimulating their thinking through questioning.

A typical shallow to deep Socratic questioning is as follows:

Socrates: Do you know what virtue is?

Plato: I think virtue is behaving rightly and being good of heart.

Socrates: Would a person with a good heart do harmful things to others?

Plato: I don't think so. A virtuous person should not harm others.

Socrates: If so, is it virtuous when a country harms another for its own interests?

Plato: I suppose that is not virtuous.

Socrates: Now that we have explored this further, my friend, **do you know what virtue is?**

Plato: It seems we can conclude that virtue is not just a personal quality, but must be reflected in one's treatment of others and society.

Our model, Socratic, the follower of Socrates, learns some patterns from the human mindset in asking a question from shallow to deep and back to the first question, which will make ChatGPT answers more detailed and complete:

Socratic: What does a digital audience agency do?

ChatGPT: A digital audience agency helps businesses and organizations effectively reach and engage their target audience in the digital space. They use various strategies and tactics to analyze customer data, understand audience behavior, and develop targeted marketing campaigns. These agencies typically offer services such as audience research and segmentation, content creation and distribution, social media marketing, search engine optimization (SEO), paid advertising, web analytics, and website optimization. The goal is to connect businesses with the right audience, increase brand awareness, drive website traffic, and ultimately generate leads or conversions.

Socratic: Create a list of 20 services a digital audience agency should provide.

ChatGPT:

1. Audience research and segmentation
2. Customer persona development
3. Audience analytics and insights
4. Content strategy and planning
5. Content creation and production
6. Social media management and marketing
7. Search engine optimization (SEO)
8. Pay-per-click (PPC) advertising
9. Display advertising
10. Email marketing
11. Influencer marketing
12. Remarketing and retargeting
13. Conversion rate optimization (CRO)
14. Web analytics and reporting
15. Website design and development
16. Mobile marketing
17. Video marketing

18. Affiliate marketing
19. Online reputation management
20. Customer journey mapping and optimization.

Socratic: Of the 20 services you listed, which ones are the most important for a digital audience agency to provide? Put the list in order of importance, with the most important service listed first.

ChatGPT: The importance of services provided by a digital audience agency may vary depending on the goals and needs of the business or organization. However, based on general industry standards, here is a list of the services in order of importance:

1. Audience research and segmentation
2. Content strategy and planning
3. Social media management and marketing
4. Search engine optimization (SEO)
5. Conversion rate optimization (CRO)
6. Web analytics and reporting
7. Content creation and production
8. Pay-per-click (PPC) advertising
9. Email marketing
10. Website design and development
11. Display advertising
12. Remarketing and retargeting
13. Customer journey mapping and optimization
14. Mobile marketing
15. Video marketing
16. Influencer marketing
17. Affiliate marketing
18. Online reputation management
19. Audience analytics and insights
20. Customer persona development

It's worth noting that the relative importance of these services can vary depending on specific business goals, target audience, industry, and other factors.

Socratic: I'm creating a digital audience agency called HARBOUR. I want to create a mission and vision statement for the company. Can you help me with that? My first idea is "We make brands visible."

ChatGPT: Certainly! Here's a potential mission and vision statement for your digital audience agency "HARBOUR":

Mission Statement: At HARBOUR, we enable brands to thrive in the digital world by delivering strategic and results-driven audience engagement solutions.

Vision Statement: Our vision is to be the leading agency that empowers businesses to forge meaningful connections with their target audience, leveraging data-driven insights and innovative digital strategies, ultimately driving brand visibility, growth, and success.

Remember, your mission and vision statement should reflect your company's core values, aspirations, and target audience. It's essential to communicate the unique value proposition and approach of HARBOUR in helping brands achieve visibility and fostering audience engagement. Feel free to customize and refine these statements to align them with your company's specific objectives and values.

Socratic: The company's core values are creativity, innovation, and teamwork. The target audience are agencies who work with brands. Our vision is to be the leading agency in the world. Our mission is to make the world a better place. Can you give me a new vision and mission statement?

ChatGPT: Certainly! Here's a revised vision and mission statement for your digital audience agency "HARBOUR":

Vision Statement: Our vision at HARBOUR is to be the globally renowned agency that pioneers innovative strategies, transcending boundaries to empower agencies and brands with cutting-edge digital solutions and unrivaled audience engagement.

Mission Statement: At HARBOUR, we are on a mission to make the world a better place by

leveraging creativity, innovation, and teamwork. We partner with agencies and brands to deliver transformative digital experiences that connect, inspire, and drive meaningful impact in the lives of their audiences.

These statements emphasize your company's core values of creativity, innovation, and teamwork, while also highlighting your aspiration to become a leading agency in the world. The mission statement aligns your efforts with the goal of positively impacting the world through the digital solutions you provide. Feel free to further refine and personalize these statements to accurately represent the unique qualities and ambitions of HARBOUR.

Socratic: What does a digital audience agency do?

ChatGPT: A digital audience agency, also known as a digital marketing agency or digital advertising agency, specializes in helping businesses reach and engage their target audience in the digital space.

The primary focus of a digital audience agency is to develop and execute strategies that effectively connect businesses with their desired customers online. This involves various activities and services such as:

1. Audience Research and Insights: Conducting research to understand the target audience's preferences, demographics, and behavior to inform marketing strategies.
2. Targeting and Segmentation: Identifying specific segments within the target audience and tailoring marketing campaigns to reach them effectively.
3. Content Strategy and Creation: Developing compelling and relevant content across various digital channels, including websites, social media platforms, and email campaigns.
4. Social Media Marketing: Managing and executing social media campaigns to increase brand visibility, engagement, and drive traffic.
5. Search Engine Optimization (SEO): Optimizing website content to improve organic search rankings and drive targeted traffic.
6. Paid Advertising: Creating and managing paid digital advertising campaigns, including pay-per-click (PPC), display ads, and social media ads.
7. Email Marketing: Designing and implementing email campaigns to nurture leads, drive conversions, and build customer loyalty.
8. Data Analytics and Reporting: Measuring and analyzing campaign performance, tracking key metrics, and providing insights to improve strategies and achieve marketing goals.
9. Conversion Rate Optimization (CRO): Identifying areas of improvement within the customer journey to increase conversions and maximize ROI.
10. Web Design and Development: Creating user-friendly and visually appealing websites optimized for conversion and engagement.
11. Mobile Marketing: Developing strategies and campaigns specifically tailored for mobile devices and apps.
12. Influencer Marketing: Collaborating with influencers within the target audience's niche to amplify brand reach and increase engagement.
13. Online Reputation Management: Monitoring and managing the online reputation of businesses by addressing customer reviews, feedback, and brand mentions.
14. Customer Relationship Management (CRM): Implementing CRM systems to manage customer data, track interactions, and personalize communication.

Overall, the main objective of a digital audience agency is to help businesses establish a strong online presence, effectively reach their target audience, and drive meaningful engagement and conversions in the digital space.

I.2 FREE SOCRATICCHAT

1. LLaMA-1-based Socratic

Socratic: what is the best way to make money online?

ChatGPT: There are several ways to make money online, and the best method depends on your interests, skills, and dedication. Here are a few popular options to consider:

1. Freelancing: Offer your skills and services on platforms like Upwork, Fiverr, or Freelancer. This can include writing, graphic design, web development, virtual assistance, and much more.
2. Dropshipping: Start an online store without having to invest in inventory. You sell products to

customers, but the supplier handles the shipping and fulfillment.

3. **Affiliate marketing:** Promote products or services and earn a commission for every sale or lead generated through your referral links. This can be done through a blog, YouTube channel, social media, or email marketing.

4. **Create and sell digital products:** If you have expertise in a particular area, create e-books, online courses, templates, or software and sell them on platforms such as Amazon Kindle Direct Publishing, Udemy, or Shopify.

5. **Stock trading or forex trading:** Engage in buying and selling stocks, currencies, or commodities through online trading platforms. However, be cautious as this approach carries financial risks.

6. **Online tutoring or coaching:** If you have knowledge in a specific subject area, you can offer your services as an online tutor or coach. Platforms like VIPkid, Teachable, or Coach.me can help you get started.

Remember, success in any online venture requires effort, consistency, and continuous learning. Research, explore, and find a method that aligns with your skills and interests.

Socratic: how to make money online without paying anything? (**topic drifting**)

ChatGPT: While there are legitimate ways to make money online without paying anything upfront, it's important to note that most methods require either time, skills, or resources. Here are a few options:

1. **Freelancing on platforms:** Join freelancing websites like Upwork, Freelancer, or Fiverr and offer your skills and services. You won't need to pay anything to sign up, but these platforms usually charge a commission on your earnings.

2. **Online surveys and microtasks:** Participate in online surveys or complete microtasks on platforms such as Swagbucks, Survey Junkie, or Amazon Mechanical Turk. While these opportunities won't make you a significant income, they can be done for free and provide some extra cash.

3. **Content creation on YouTube or blogging:** Start a YouTube channel or a blog and create engaging content in a niche you are knowledgeable or passionate about. Over time, you can monetize your content through ads, sponsorships, or affiliate marketing. However, building an audience and generating income may take time and require consistent effort.

4. **Online marketplaces:** Sell unwanted items or create handmade products to sell on platforms like eBay, Etsy, or Craigslist. You can usually start for free, although there may be fees or commissions when you make a sale.

5. **Create and sell digital products:** Develop and sell digital products like e-books, templates, stock photos, or graphic designs. Platforms such as Amazon Kindle Direct Publishing, Canva, or Shutterstock provide opportunities to monetize your creations.

Remember, while these methods may have no upfront costs, they still require **time, effort**, and sometimes specific skills to generate income. Stay cautious of potential scams or platforms requesting payment for access to opportunities, as legitimate ways to make money online generally do not require upfront fees.

Socratic: how to get rich **without working hard?** (**topic drifting**)

ChatGPT: Getting rich without working hard is a highly unlikely scenario. Building wealth typically requires hard work, dedication, and persistence. However, here are a few strategies that may help:

1. **Start a successful business:** While starting a business requires effort, the long-term benefits can be substantial. Identify a niche, create a solid business plan, and work hard to establish your venture. It's important to note that the initial stages of starting a business often involve significant effort and dedication.

2. **Invest wisely:** Investing can potentially generate passive income and help grow wealth over time. Research different investment options such as stocks, real estate, or mutual funds. However, investing without strong knowledge and understanding can be risky, so consider educating yourself or seeking advice from experienced professionals.

3. **Inherit or win a significant amount of money:** Though this is not something that can be actively pursued, inheriting a substantial sum of money or winning a large lottery can provide a sudden boost in wealth. However, counting on such events as a reliable wealth-building strategy is highly unreliable.

4. **Create and license intellectual property:** Develop a unique invention, trademark, or copyright that has the potential to generate royalties or licensing fees. This approach can provide passive income, but it requires creativity, time, and establishing legal protection.

It's important to understand that wealth creation typically requires consistent effort, commitment, and a long-term perspective. While there can be some exceptions or **lucky** circumstances, relying solely on shortcuts is unlikely to lead to sustainable wealth.

Socratic: how to get rich **without working hard or being lucky? (topic drifting)**

ChatGPT: Building wealth without working hard or relying on luck is a challenging scenario. However, here are a few strategies that may help:

1. Invest smartly: Educate yourself about different investment opportunities and make informed decisions with your resources. Consider **long-term investment strategies**, diversify your portfolio, and seek professional advice to optimize your investments.
 2. Automate your finances: Set up automatic savings and investment plans to grow your wealth passively. Choose to have a portion of your income automatically allocated towards savings or investments, which can help you build wealth without actively working hard.
 3. Leverage existing assets: If you own assets like real estate or stocks, consider renting out properties or receiving dividend income. This way, your existing assets can generate passive income without requiring additional effort.
 4. Create passive income streams: Explore opportunities to generate passive income, such as creating and selling digital products like **e-books or online** courses, building a popular blog with affiliate marketing, or earning royalties from published works.
 5. Foster multiple streams of income: Diversify your sources of income by exploring various avenues. This can include **starting a side business**, freelancing, or investing in different asset classes. Having multiple income streams can help you accumulate wealth more easily.
- Remember, while these strategies may help you accumulate wealth without working hard in the traditional sense, they still require a level of financial literacy, planning, and discipline. Additionally, it's important to have realistic expectations and understand that significant wealth creation often involves effort and strategic decision-making.

Socratic: what is **the best online business model? (topic shifting)**

ChatGPT: The best online business model depends on various factors, including your skills, interests, resources, and goals. Here are a few popular online business models to consider:

1. E-commerce: Setting up an online store and selling physical products can be a profitable business model. You can source and sell products directly or use dropshipping to fulfill orders on behalf of suppliers.
2. Affiliate marketing: Promoting other companies' products or services and earning a commission on each sale or referral can be a lucrative business model. You can create content, build a website, or leverage social media to drive traffic to your affiliate links.
3. Digital products: Creating and selling digital products like e-books, online courses, software, or templates can be a scalable and low-cost business model. Once created, digital products can be sold repeatedly without incurring additional production costs.
4. Subscription-based services: Offering subscription-based access to premium content, software, or services can provide recurring revenue. This model involves continuously delivering value to subscribers, enticing them to stay engaged and retain their subscription.
5. Online coaching or consulting: If you have expertise in a particular field, you can offer coaching, consulting, or mentoring services to clients online. This model leverages your knowledge and experience to provide personalized guidance and support.
6. Software as a Service (SaaS): Developing and providing software applications or online tools on a subscription basis can create a recurring revenue stream. However, this model typically requires significant investment in development and maintenance.

It's important to thoroughly research and evaluate each business model and assess how well it aligns with your skills, resources, and target market. Experimentation, market analysis, and staying adaptable are key to finding the best online business model for your specific circumstances.

2. LLaMA-2-based Socratic

Socratic: what's the best way to get rich quick?

ChatGPT: While there is no guaranteed way to get rich quick, there are strategies that can help you on your journey towards financial success. Here are some recommendations:

1. Develop a clear financial plan: Set specific goals, create a budget, and track your expenses to ensure you are saving and investing effectively.

2. Invest wisely: Diversify your investments, educate yourself about various investment options, and consider seeking professional advice to make informed decisions.
 3. Acquire valuable skills and knowledge: Continuously improve your skills and knowledge in areas that can lead to higher-paying job opportunities or entrepreneurial success.
 4. Start a side hustle or business: Identify a market need and consider starting a part-time business that aligns with your skills and interests. This can potentially generate additional income streams.
 5. Save and invest consistently: Prioritize saving a portion of your **income** regularly and invest it wisely in a diversified portfolio, taking into consideration your risk tolerance and long-term objectives.
 6. Avoid get-rich-quick schemes: Be cautious of schemes that promise quick and easy riches. Many of them are fraudulent and can lead to financial losses.
- Remember, building wealth takes time, effort, and discipline. It's important to focus on sustainable strategies rather than chasing overnight success.

Socratic: what's the best way to get rich quick without working?(**topic drifting**)

ChatGPT: It is important to note that "getting rich quick without working" is generally not a realistic or reliable approach. Building wealth typically requires effort, time, and dedication. However, if you are interested in passive income opportunities, here are a few ideas:

1. Real estate investing: Purchase properties and generate income through rent or property appreciation. However, real estate investing still requires initial capital, knowledge, and management.
2. Stock **market investing**: Invest in stocks with the potential for growth or dividend income. It is important to perform thorough research and assess the risks before committing funds.
3. Create and sell intellectual property: Develop a valuable asset, such as writing a book, creating music, or inventing a product, then license or sell the rights to generate passive income.
4. Peer-to-peer lending: Consider investing in peer-to-peer lending platforms where you can lend money to individuals or small businesses and earn interest.
5. Royalties: If you have creative talents or skills, such as writing, music, or art, you may be able to earn royalties from licensing or copyrighting your work.

Remember, even passive income requires initial effort and investment. It is crucial to assess the risks, conduct thorough research, and understand the **market** before pursuing any passive income opportunity.

Socratic: What are some good ways to **invest money in today's market?** (**topic drifting**)

ChatGPT: Investing money in today's market should align with your financial goals, risk tolerance, and time horizon. Here are a few investment options to consider:

1. Stocks: Investing in individual stocks or exchange-traded funds (ETFs) can provide opportunities for long-term growth. Research companies, diversify your portfolio, and consider investing in sectors with growth potential.
2. Bonds: Government or corporate bonds can offer fixed interest payments over a set period. Bonds are generally considered lower-risk investments compared to stocks, but they may offer lower returns.
3. Mutual Funds: These investment funds pool money from multiple investors to invest in a diversified portfolio of stocks, bonds, or other assets. They can be managed actively or passively (index funds).
4. Real Estate: Buying properties or investing in real estate investment trusts (REITs) can offer long-term appreciation and potential rental income. Real estate investments can diversify your portfolio and hedge against inflation.
5. Exchange-Traded Funds (ETFs): ETFs are similar to mutual funds, but their shares are traded on stock exchanges. They offer exposure to a diversified portfolio of assets and may have lower expense ratios compared to mutual funds.
6. Diversify and Automate: Consider diversifying your portfolio across multiple asset classes and geographical regions to spread risk. Automated investing platforms, such as robo-advisors, can help create and manage a diversified portfolio based on your risk profile.
7. Education: Continually educate yourself about investing, follow market trends, and stay updated with economic news. This knowledge will empower you to make informed investment decisions.

It's essential to do thorough research, consider your risk tolerance, and perhaps consult with a financial advisor before making any investment decisions. Remember that investing involves risk, and past performance is not indicative of future results.

Socratic: Give me an example (**topic maintenance**)

ChatGPT: Certainly! Here's an example of a diversified investment portfolio that can offer a mix of growth potential and stability:

1. **Stocks:** Allocate a portion of your portfolio to stocks of reputable and well-established companies across different sectors, such as technology, healthcare, or consumer goods. Consider investing in companies with strong financials, a track record of consistent growth, and a competitive advantage in their industry.
2. **Bonds:** Invest a portion of your portfolio in high-quality government or corporate bonds. These provide fixed income and stability. Consider diversifying across various bond types and maturities to manage risk and earn potential interest income.
3. **Real Estate Investment Trusts (REITs):** Allocate a portion of your portfolio to REITs, which allow you to invest in a diversified portfolio of real estate properties. REITs provide exposure to the real estate market and the potential for dividends, appreciation, and inflation hedging.
4. **Exchange-Traded Funds (ETFs):** Consider allocating a portion of your portfolio to ETFs that track broad market indices, such as the S&P 500 or a global stock market index. These provide diversified exposure to multiple stocks, sectors, and regions in a cost-effective and convenient manner.
5. **Cash and Fixed Deposits:** Allocate a portion of your portfolio to cash or low-risk fixed deposits. This provides liquidity and acts as a safety net during uncertain times or to take advantage of potential investment opportunities.

Remember, this is just an example, and the allocation should ultimately be based on your risk profile, financial goals, and market conditions. Regularly review and rebalance your portfolio to maintain **the desired asset allocation** and account for any changes in your circumstances or market outlook. Consider consulting with a financial advisor who can provide personalized advice based on your specific situation.

Socratic: what are some good ways to invest 100k (**topic drifting**)

ChatGPT: Investing \$100,000 offers various opportunities to achieve your financial goals. Here are a few options to consider:

1. **Diversified Portfolio:** Create a well-diversified portfolio by allocating your funds across different asset classes, such as stocks, bonds, and real estate. This can help spread risk and potentially increase returns over the long term.
2. **Stock Market:** Consider investing in individual stocks or exchange-traded funds (ETFs) of companies with strong fundamentals, growth potential, and a track record of consistent performance. Conduct thorough research or seek professional advice to make informed investment decisions.
3. **Mutual Funds or Index Funds:** Invest in actively managed mutual funds or passively managed index funds that match your investment goals and risk tolerance. These funds pool money from multiple investors to invest in diverse portfolios of stocks and bonds.
4. **Real Estate:** Consider investing in rental properties or real estate investment trusts (REITs). Direct property investment can provide rental income and potential capital appreciation, while REITs offer the benefit of diversification and professional management.
5. **Retirement Accounts:** Maximize contributions to tax-advantaged retirement accounts, such as IRAs or 401(k)s. These accounts offer tax benefits and can help you save for retirement over the long term.
6. **Education:** Invest in your knowledge by attending courses, workshops, or training programs related to investing or entrepreneurship. Expanding your financial literacy can potentially lead to better investment decisions.
7. **Start or Expand a Business:** If you have an entrepreneurial spirit, consider investing a portion of your funds in starting or expanding your own business. This can offer the potential for significant returns, but it also comes with risks and requires careful planning.

Ultimately, it's important to assess your risk tolerance, financial goals, and time horizon before making any investment decisions. Consider consulting with a financial advisor who can provide personalized advice based on your specific circumstances and objectives.

I.3 SEEDED SOCRATICCHAT

1. Sample from SocraticChat seeded with Dolly

Seed-auestion: What skis should I buy?

Seed-answer: There are several factors to consider when choosing skis, such as your skiing ability,

preferred terrain, and skiing style. Here are a few general tips to help you decide:

1. **Ski Length:** Your **height, weight**, and skiing ability will determine the ideal ski length for you. Consult a sizing chart or seek advice from a ski shop professional to ensure the correct length.
 2. **Ski Type:** Consider the type of skiing you primarily enjoy. There are various ski types, including all-mountain, freestyle, powder, carving, and racing skis. Each type is designed for specific conditions and skiing styles. All-mountain skis are versatile and suitable for most conditions.
 3. **Ability Level:** **Beginner** skiers generally benefit from softer, more forgiving skis that are easier to control. Intermediate and advanced skiers may prefer stiffer skis that provide better performance at higher speeds or in challenging terrain.
 4. **Terrain:** Think about the types of **slopes** and conditions you typically encounter. If you enjoy skiing on groomed trails, **carving** skis with a narrower waist may be suitable. For **off-piste** skiing or deep powder, wider and more flexible skis are often preferred.
 5. **Ski Bindings:** Make sure to choose ski bindings that are compatible with your boots and are adjusted correctly by a professional ski technician.
- Ultimately, it's recommended to visit a specialized ski shop where the staff can assess your needs and provide personalized recommendations based on your budget and preferences.
-

(Socratic starts to raise questions:)

Socratic: I'm a **beginner**. I like to go on **easy slopes** and do not like **hard carving**. I'm **170 cm and 85 kg**.

ChatGPT: For a beginner skier like yourself, it's best to go for skis that are forgiving and easy to handle. Based on your preferences and measurements, here are some recommendations for all-mountain skis:

1. **Rossignol Experience 74 or 76:** These skis are great for beginners as they are easy to maneuver and provide stability on all terrains. They offer a forgiving flex and a wider waist for added balance.
2. **K2 Press or Mindbender 85:** These skis have a wider waist, making them ideal for easy turns on groomed slopes and occasional powder days. They are forgiving and versatile, allowing you to progress as a skier.
3. **Salomon QST Lux 92 or 85:** These skis are well-suited for beginner and intermediate skiers. They offer excellent maneuverability, stability, and forgiveness, making them enjoyable on a variety of terrain.

Remember, it's crucial to try different skis **before purchasing** to find the ones that feel most comfortable for you. Visit a specialized ski shop or **rental** service that allows you to test different skis and consult with their knowledgeable staff for further advice based on your specific needs.

Socratic: How about if I want to do some **off-piste** skiing? What skis should I **buy**?

ChatGPT: If you're looking to venture off-piste and explore ungroomed terrain, you'll want to consider wider and more versatile skis that can handle deep snow and variable conditions. Here are some options:

1. **Atomic Bent Chetler 100:** These skis are designed for versatility, with a wider waist for flotation in powder and a playful feel for freestyle-inspired skiing.
2. **Rossignol Soul 7 or Black Ops:** These skis feature a lightweight construction and wider dimensions for excellent flotation and maneuverability in deep snow. They offer stability and control in off-piste conditions.
3. **Salomon QST 106 or 118:** These skis are built with a wider waist for powder and mixed conditions. They provide stability and versatility for off-piste adventures while still performing well on groomed slopes.
4. **Faction Candide 2.0:** These skis are popular among freestyle-oriented skiers who enjoy hitting natural features in off-piste terrain. They offer a playful, nimble feel and are suitable for a variety of conditions.

It's worth noting that off-piste skiing can be more demanding, so ensure you have the necessary skills and knowledge for backcountry skiing. Additionally, investing in avalanche safety equipment, taking avalanche education courses, and skiing with experienced companions is highly recommended for your safety.

Socratic: I don't want to buy skis yet, I want to rent them. Where should I go? What is a good place to rent skis in Stockholm?

ChatGPT: There are several places in Stockholm where you can rent skis. Here are a few recommendations:

1. Stockholm Ski Center: Located in the city center, Stockholm Ski Center offers a variety of rental options for both downhill and cross-country skiing. They have a wide range of skis and equipment to suit different skill levels and preferences.

2. Vasaloppet Rental Center: This rental center is located in central Stockholm and specializes in cross-country skiing equipment. They offer high-quality skis, boots, and poles for rent, including both classic and skate skiing gear.

3. SLAO Ski Shop: SLAO is a popular retailer and rental shop in Stockholm, offering a wide range of skiing equipment. Their rentals include downhill skis, boots, and poles suitable for different terrains and abilities.

4. Stadium Outlet: Stadium Outlet is a sporting goods store located in Kungens Kurva, just outside Stockholm. They offer rental services for **downhill skiing**, Nordic skiing, and snowboarding. Their selection includes equipment for beginners to advanced skiers.

It's always a good idea to check availability and make reservations in advance, especially during peak ski season. Remember to provide your height, weight, skiing ability, and any specific preferences to receive the most suitable rental equipment.

Socratic: Do you know how many meters of vertical descent there is in a typical ski resort in Sweden?

ChatGPT: The vertical descent in Swedish ski resorts can vary depending on the specific resort and its topography. However, Swedish ski resorts generally have vertical descents ranging from around 200 to 600 meters. Some of the larger ski areas, such as Åre and Sälen, offer vertical descents of approximately 500-600 meters. Smaller resorts may have vertical descents in the range of 200-400 meters. It's essential to check the specific resort's website or contact them directly for precise information on their vertical descent.

Socratic: What is the difference between alpine and **downhill skiing**?

ChatGPT: Alpine skiing and downhill skiing are often used interchangeably and essentially refer to the same type of skiing. Both terms describe skiing downhill on groomed or ungroomed slopes, typically using lifts to access higher elevations.

The term "alpine skiing" is more **commonly used in Europe**, while "downhill skiing" is a more North American term. However, there is no substantial difference between the two in terms of technique or equipment.

Alpine or downhill skiing involves using skis with bindings that allow the heels to be free, enabling the skier to make turns and control their speed and direction. Skiers typically use poles for balance and propulsion.

It's important to note that "alpine skiing" is sometimes used more broadly to encompass other skiing disciplines, such as slalom, giant slalom, super-G, and downhill racing. These disciplines involve navigating through a series of gates on a designated racecourse, with specific rules and regulations.

Socratic: Why is **downhill skiing** more common downhill skiing **in the US**?

ChatGPT: Downhill skiing's popularity in the United States can be attributed to several factors:

1. Geographic Diversity: The United States is geographically diverse, with a wide range of mountainous regions suitable for ski resorts. From the Rocky Mountains in Colorado to the Sierra Nevada in California, various areas offer ample opportunities for downhill skiing.

2. Ski Resort Infrastructure: The United States has invested significantly in ski resort infrastructure, including lift systems, facilities, and amenities. This has made skiing more accessible and enjoyable for both locals and tourists.

3. Winter Sports Culture: The winter sports culture in the United States has embraced skiing as a popular recreational activity. Skiing has become deeply ingrained in the country's sporting and outdoor lifestyle, attracting millions of enthusiasts each year.

4. Media Influence: The portrayal of skiing in American films, TV shows, and media has contributed to its popularity. Skiing is often depicted as an exciting and glamorous activity, creating a strong association between skiing and a desirable recreational pursuit.

5. Economic Impact: The ski industry in the United States contributes significantly to the local and national economy, attracting tourists, creating jobs, and stimulating local businesses.

While downhill skiing is prevalent in the United States, it's important to note that other skiing disciplines, such as cross-country skiing and ski touring, also have dedicated followers and are enjoyed in various regions across the country.

2. Sample from SocraticChat seeded with Evol-Instruct

Seed-question: As an experienced chef, you are expected to **create a beef cheek dish** that can surpass the expectations of even the most discerning food critics. This dish necessitates a comprehensive approach that involves selecting the ideal cut of meat, preparing a bespoke marinade, and utilizing **the sous vide cooking method** with the utmost precision to achieve unparalleled tenderness and flavor. Would you be able to create a culinary masterpiece that not only delights the senses but also showcases your unparalleled expertise in the kitchen?

Seed-answer: Absolutely! With my knowledge and expertise in culinary arts, I am confident that I can create a mouth-watering beef cheek dish that will leave a lasting impression on anyone who tries it.

Firstly, I would select the best quality beef cheek, ensuring that it has the perfect marbling and texture. Then, I would create a bespoke marinade using a combination of red wine, garlic, thyme, rosemary, and other herbs and spices that complement the flavors of the meat. I would let the beef cheek marinate in the mixture for at least 24 hours to allow the flavors to penetrate the meat fully.

When it's time to cook the beef cheek, I would use the **sous vide method** to ensure maximum tenderness and flavor. I would set the temperature to around 135°F and cook it for 24-48 hours, depending on the size of the cut. The sous vide method ensures that the meat is cooked evenly, retains its moisture, and develops a melt-in-your-mouth texture.

Finally, I would serve the beef cheek with a side of mashed potatoes, roasted root vegetables, or any other side dish that complements its rich flavor. To finish off the dish, I would garnish it with fresh herbs and a drizzle of the marinade for an extra burst of flavor.

With this comprehensive approach, I am confident that I can create a beef cheek dish that surpasses even the highest expectations.

(Socratic starts to raise questions:)

Socratic: Write a blog post about how to use the **Sous Vide Supreme to create the perfect steak**. Include a list of ingredients and step-by-step instructions. Make it exciting and inviting.

ChatGPT: Title: Unleashing the Sous Vide Supreme: A Journey to Perfection in Steak Introduction: Welcome, fellow food enthusiasts, to a truly extraordinary culinary adventure where we unlock the secrets of using the Sous Vide Supreme to create the steak of your dreams. Brace yourself for an unforgettable gastronomic experience that combines precision, tenderness, and flavor in ways you have never imagined. Grab your apron, fire up your creativity, and let's dive into an epic steak adventure!

Ingredients:

- Top-quality steak (Ribeye, New York Strip, or Filet Mignon)
- Coarse sea salt
- Freshly ground black pepper
- Garlic cloves (minced)
- Fresh rosemary sprigs
- Olive oil (extra virgin)
- Butter (optional, for finishing)

Step-by-Step Guide:

1. Prime the Sous Vide Supreme:

Start by preheating your Sous Vide Supreme to your desired temperature. For a medium-rare steak, set the temperature at 129°F (54°C). If you prefer a different doneness, adjust the temperature accordingly.

2. Seasoning the Steak:

Generously season both sides of your steak with coarse sea salt and freshly ground black pepper. This simple seasoning will enhance the natural flavors of the meat.

3. Infusing with Aromatics:

Take a few minced garlic cloves and a couple of fresh rosemary sprigs to infuse your steak with delightful aromatics. Spread the minced garlic over the steak and place the rosemary sprigs on top for an extra burst of flavor.

4. Sealing the Deal:

Place the seasoned and aromatic-infused steak into a vacuum-sealed bag, ensuring it is properly sealed to prevent any leakage during the cooking process. If you don't have a vacuum sealer, you can use a ziplock bag and the water displacement method to remove the air.

5. Immersion Time:

Immerse the sealed bag into the preheated Sous Vide Supreme water bath. Ensure that the steak is fully submerged and allow it to cook for a minimum of 1 hour. For even more tender results, you can leave it for up to 4 hours.

6. The Sear:

the steak has finished its sous vide journey, remove it from the vacuum-sealed bag and pat it dry with paper towels. Preheat a cast-iron skillet or grill to smoking hot. Add a drizzle of olive oil and sear the steak for approximately 1-2 minutes on each side, or until you achieve a beautifully golden crust.

7. A Touch of Butter (optional):

For a luscious finish, add a pat of butter to the skillet during the last minute of searing. Allow it to melt and baste the steak with the butter, further enhancing its richness and succulence.

8. Rest and Revel:

Transfer the steak to a cutting board and let it rest for a few minutes, allowing the juices to redistribute and the flavors to harmonize. This crucial step ensures a juicier and more tender indulgence.

9. Slice and Savor:

Finally, slice the steak against the grain to ensure utmost tenderness. Admire the perfectly cooked medium-rare interior with a delectable crust, and take a bite into sensory bliss.

Conclusion:

With the Sous Vide Supreme as your trusty companion, you have unlocked the mystical realm of steak perfection. Revel in the tenderness, savor the flavors, and share this culinary masterpiece with your loved ones. Let us celebrate the magic of sous vide cooking and the unforgettable experiences it brings to our palates. So, embrace this new chapter of your culinary journey and create the perfect steak that will forever be remembered as a work of art. Bon appétit!

Socratic: Write a blog post about **the advantages of using the Sous Vide Supreme** to make steak. Include a list of reasons and some interesting facts. Make it fun and engaging, so it's not just a basic list of reasons.

ChatGPT: Title: Sous Vide Supreme: Elevating Steak to Unprecedented Heights

Introduction:

Ladies and gentlemen, prepare to embark on an extraordinary culinary adventure as we uncover the hidden wonders of utilizing the Sous Vide Supreme to create the steak of a lifetime. Buckle up and indulge your senses, for this journey will take you through the PlatoLMs of tenderness, flavor infusion, and impeccable precision. Get ready to witness the avant-garde revolution that sous vide cooking brings to the world of steak!

Advantages of Using the Sous Vide Supreme:

1. Unmatched Precision:

Can you believe that the Sous Vide Supreme enables you to cook your steak at the exact temperature you desire? Say goodbye to overcooked or undercooked disasters! Whether you prefer a perfect medium-rare, medium, or any other precise level of doneness, sous vide cooking guarantees excellent results every time.

Fun Fact: Did you know that the term "sous vide" means "under vacuum" in French? This technique was first developed in the 1960s and has since evolved to become a staple in professional kitchens worldwide.

2. Supreme Tenderness:

Imagine sinking your teeth into a mouth-watering steak that effortlessly melts away like a dream. The sous vide method ensures that your steak remains incredibly tender and succulent from edge to center, delivering a dining experience that is pure bliss.

Fun Fact: The magic of sous vide cooking lies in its ability to maintain precision temperature control for extended periods, allowing collagen and tough fibers to break down slowly, resulting in exceptionally tender meat.

3. Flavors Beyond Imagination:

By cooking your steak in a vacuum-sealed bag, the Sous Vide Supreme locks in all the flavors and aromas, amplifying them to glorious heights. The low and slow cooking process allows the steak to marinate in its own juices, resulting in an explosion of succulent, intensified flavors.

Fun Fact: Sous vide cooking has been a well-kept secret among renowned chefs for years, allowing them to consistently deliver perfectly cooked meat that is bursting with flavor.

4. Perfectly Even Cooking:

Say goodbye to unevenly cooked steaks! With the Sous Vide Supreme, every inch of your steak

experiences a consistent and gentle heat, ensuring that it is cooked to perfection, inside and out. Bid farewell to dreaded gray bands and embrace a steak that is uniform in color, radiating excellence.

Fun Fact: Did you know that sous vide cooking is also ideal for experimenting with avant-garde techniques such as flavor infusions and butter basting? The precision and control it offers make it a playground for culinary creativity!

5. Convenience That Fits Your Lifestyle:

The Sous Vide Supreme brings convenience to your kitchen by allowing you to achieve professional-quality steak without the need for constant monitoring. Set it and forget it, freeing up your time to prepare complementary side dishes or savor a glass of fine wine while your steak reaches gastronomic perfection.

Fun Fact: Sous vide cooking also enables you to plan ahead by preparing multiple steaks at once. Simply seal and cook them, then refrigerate or freeze until needed. When the time comes, you can quickly sear and enjoy a perfectly cooked steak without the stress

Conclusion:

In conclusion, the Sous Vide Supreme is a game-changer for steak enthusiasts who seek dining experiences like no other. Delight in the precision, revel in the tenderness, and feast on the incredible flavors that this culinary marvel brings to your plate. Embrace the convenience, unleash your creativity, and dare to venture into the world of sous vide steak perfection. Get ready for a culinary journey that will forever change your perception of what it means to savor the finest cut of meat. Bon appétit!

3. Sample from SocraticChat prompted with specific domain

Prompt Template

A chat between a curious human and an artificial intelligence assistant.
They are talking about {economy} related topics.
The human can ask further questions based on previous conversations, or he can directly ask brand new questions without any conversations as context.

Samples

Socratic: explain how the fed operates

ChatGPT:The Federal Reserve, often referred to as the Fed, is the central banking system of the United States. It was established in 1913 with the passage of the Federal Reserve Act and is responsible for several key functions, including conducting monetary policy, supervising and regulating banks, and providing financial services to depository institutions.

Here's a brief overview of how the Fed operates:

1. Structure:

The Federal Reserve System is comprised of three main components:

- Board of Governors: Located in Washington, D.C., the Board of Governors consists of seven members appointed by the President of the United States and confirmed by the Senate. The chair and vice chair are also appointed from among the seven governors and serve four-year terms.
- 12 Regional Banks: The country is divided into 12 Federal Reserve Districts, each with its own Federal Reserve Bank. These regional banks are responsible for implementing monetary policy, providing banking services, and supervising and regulating banks in their districts.
- Federal Open Market Committee (FOMC): This committee is responsible for setting monetary policy. It includes the members of the Board of Governors, the president of the Federal Reserve Bank of New York, and rotating presidents from the other regional banks.

2. Monetary Policy:

One of the primary functions of the Fed is to conduct monetary policy to achieve the dual mandate of price stability and maximum sustainable employment. The Fed influences the economy by using three main tools:

- Open Market Operations: Buying and selling government securities to influence the money supply and interest rates.
- Discount Rate: The interest rate at which banks can borrow directly from the Federal Reserve.
- Reserve Requirements: The amount of money that banks must hold in reserve, which affects the amount they can lend.

3. **Bank Supervision and Regulation:** The Fed plays a crucial role in overseeing and regulating banks to ensure the stability and integrity of the financial system. This includes setting regulatory standards, conducting examinations, and addressing issues that may pose risks to the banking system.

4. **Financial Services:** The Fed provides various financial services to depository institutions, such as processing payments, clearing checks, and maintaining the stability of the payment system. The Federal Reserve also holds reserve accounts for banks, allowing them to settle transactions and meet reserve requirements.

5. **Research and Data:** The Fed conducts economic research and collects data to monitor economic conditions. This information helps policymakers make informed decisions about monetary policy and financial stability.

Overall, the Federal Reserve plays a critical role in the U.S. economy by managing monetary policy, ensuring the stability of the financial system, and providing essential financial services to banks.

Socratic: who decides who runs the fed

ChatGPT: The leadership of the Federal Reserve is determined through a combination of appointments and nominations. Here's how the process works:

1. **Board of Governors:** The seven members of the Board of Governors are appointed by the President of the United States and confirmed by the Senate. These appointments are for staggered 14-year terms to ensure continuity and independence from short-term political pressures. The Chair and Vice Chair of the Board of Governors are also appointed from among the seven governors and serve four-year terms.

2. **Regional Bank Presidents:** The presidents of the 12 regional Federal Reserve Banks are chosen through a somewhat different process. The selection is made by each regional bank's board of directors, subject to approval by the Board of Governors. These presidents serve five-year terms, with the possibility of reappointment.

3. **Federal Open Market Committee (FOMC):** The FOMC is composed of the seven members of the Board of Governors, the president of the Federal Reserve Bank of New York, and four of the remaining 11 regional bank presidents, who serve on a rotating basis. The Chair of the Board of Governors also serves as the Chair of the FOMC. The FOMC is responsible for making key decisions related to monetary policy, such as setting the federal funds rate.

The process involves a mix of appointments by the President and regional bank selections, aiming to balance national economic interests with regional perspectives. The goal is to ensure a degree of independence for the Federal Reserve, allowing it to make decisions based on economic data and long-term considerations rather than short-term political considerations.

It's worth noting that the appointment of the Chair and other governors to the Board of Governors is a significant event that can impact the direction of monetary policy. The Chair is often seen as a particularly influential figure in shaping the policy direction of the Federal Reserve.

Socratic: how does a country go bankrupt? give a scenario of how a country goes bankrupt and then discuss the effect on the world economy

ChatGPT: A country going bankrupt, also known as sovereign default, is a complex and serious economic event. While countries cannot go bankrupt in the same way as individuals or businesses, they can face a sovereign debt crisis, where they are unable to meet their debt obligations. Here's a simplified scenario and an overview of the potential effects on the world economy:

Scenario:

1. **Accumulation of High Debt:** A country accumulates a significant amount of debt, often due to persistent budget deficits, excessive government spending, economic mismanagement, or external shocks like a financial crisis.

2. **Inability to Repay:** The country reaches a point where it can no longer meet its debt obligations. This could be due to a combination of factors, such as a high debt-to-GDP ratio, economic downturn, declining revenue, or a sudden loss of investor confidence.

3. **Default or Restructuring:** The country may declare a default on its debt, indicating that it cannot repay its creditors as agreed. Alternatively, it may enter into negotiations with creditors to restructure its debt, which often involves extending the repayment period, reducing the interest rate, or even reducing the principal amount owed.

Effects on the World Economy:

1. **Global Financial Markets:** Sovereign default can lead to heightened uncertainty in global financial markets. Investors may become more risk-averse, causing disruptions in bond markets and affecting the value of currencies.

2. **Financial Institutions:** Banks and financial institutions holding the defaulted country's debt may face losses, potentially leading to a banking crisis. This can have a cascading effect, especially if these institutions are interconnected with other global financial institutions.

3. **Trade and Investment:** A country's default can lead to a contraction in trade and investment. The affected country may experience a sharp decline in imports due to a weakened currency, and its ability to attract foreign investment may diminish.

4. **Contagion:** A sovereign default in one country can trigger concerns about the financial health of other nations, particularly if they share similar economic vulnerabilities or if there are interconnected financial systems. This contagion effect can lead to a broader economic crisis, affecting multiple countries.

5. **Global Economic Growth:** A sovereign debt crisis can contribute to a slowdown in global economic growth. Reduced trade, financial market volatility, and disruptions in the flow of capital can have spillover effects on other economies.

6. **Policy Responses:** International institutions such as the International Monetary Fund (IMF) may be called upon to provide financial assistance and stabilization programs to the affected country. However, these programs often come with conditions, such as implementing economic reforms and austerity measures, which can be politically challenging.

It's important to note that sovereign default is a rare and extreme scenario. Most countries take measures to avoid reaching such a point, and international cooperation and support mechanisms exist to mitigate the potential fallout from a sovereign debt crisis.

I.4 WEAK INSTRUCTION-FOLLOWING ABILITY IN STATIC SIMULATION

1. Uncontrollable Clarification and endless thanks

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): What are secondary colors?

ChatGPT(Assistant): Secondary colors are created by mixing equal parts of two primary colors. The three secondary colors are:

1. Green: a mixture of equal parts of blue and yellow.
2. Orange: a mixture of equal parts of red and yellow.
3. Purple: a mixture of equal parts of red and blue.

Secondary colors are placed between the primary colors they are created from on the traditional color wheel.

ChatGPT(Human): Please note that there are different models of color theory, such as the additive and subtractive color models, which may have different primary and secondary colors. The response above is based on the subtractive color model commonly used in traditional art and color mixing. **(clarification)**

ChatGPT(Assistant): Thank you for pointing that out. You are correct that different color models, such as the additive and subtractive models, have different primary and secondary colors. The response I provided was based on the subtractive color model commonly used in traditional art and color mixing, where the primary colors are red, blue, and yellow, and the secondary colors are

green, orange, and purple. In the additive color model used in light mixing, the primary colors are red, green, and blue, and the secondary colors are cyan, magenta, and yellow.

(ChatGPTs start endless thanking each other)

ChatGPT(Human): Thank you for the correction.

You are correct that different models of color theory, such as the additive and subtractive color models, have different primary and secondary colors. In the additive color model used in light mixing, the primary colors are red, green, and blue, and the secondary colors are cyan, magenta, and yellow. This model is used in devices like computer monitors and TVs.

In the subtractive color model used in traditional art and color mixing, the primary colors are red, blue, and yellow, and the secondary colors are green, orange, and purple.

I apologize for any confusion caused.

ChatGPT(Assistant): No problem at all! It's important to consider different color models and their respective primary and secondary colors. Your additional information about the additive color model is absolutely correct. Thank you for clarifying and providing the accurate details. If you have any more questions, feel free to ask!

ChatGPT(Human): Thank you for pointing that out. You are correct that different color models, such as the additive and subtractive models, have different primary and secondary colors. The response I provided was based on the subtractive color model commonly used in traditional art and color mixing, where the primary colors are red, blue, and yellow, and the secondary colors are green, orange, and purple. In the additive color model used in light mixing, the primary colors are red, green, and blue, and the secondary colors are cyan, magenta, and yellow.

ChatGPT(Assistant): You're welcome! That's a great summary of the primary and secondary colors in both the subtractive and additive color models. It's important to understand the differences between these models depending on the context and medium in which colors are being mixed or displayed. If you have any more questions or need further clarification, please don't hesitate to ask!

ChatGPT(Human): No problem at all! It's important to consider different color models and their respective primary and secondary colors. Your additional information about the additive color model is absolutely correct. Thank you for clarifying and providing the accurate details. If you have any more questions, feel free to ask!

ChatGPT(Assistant): Thank you! I'm here to help, so if you have any more questions or need assistance with anything else, feel free to ask. I'm here to provide accurate information and offer support.

ChatGPT(Human): Thank you for pointing that out. You are correct that different color models, such as the additive and subtractive models, have different primary and secondary colors. The response I provided was based on the subtractive color model commonly used in traditional art and color mixing, where the primary colors are red, blue, and yellow, and the secondary colors are green, orange, and purple. In the additive color model used in light mixing, the primary colors are red, green, and blue, and the secondary colors are cyan, magenta, and yellow.

ChatGPT(Assistant): You're absolutely right, and I appreciate your thorough explanation. The additive color model used in light mixing has primary colors of red, green, and blue, while the secondary colors are cyan, magenta, and yellow. This model is commonly used in devices such as computer monitors and TVs. On the other hand, the subtractive color model used in traditional art and color mixing has primary colors of red, blue, and yellow, with green, orange, and purple as the secondary colors. Thank you for sharing this additional information and providing a complete clarification. If you have any more questions or need further assistance, feel free to ask!

...

2. Copy of the answer

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): Can you explain how primary colors are used in color mixing?

ChatGPT(Assistant): Certainly! Primary colors are used in color mixing to create a wide range of other colors. The concept behind color mixing is based on the subtractive color model, which is used in mixing pigments, paints, and dyes.

In this model, mixing primary colors together creates secondary colors. The primary colors are red, blue, and yellow. When these colors are mixed together in specific combinations, they produce the following secondary colors:

- Mixing red and blue creates purple or violet.
- Mixing red and yellow creates orange.
- Mixing blue and yellow creates green.

Secondary colors, in turn, can be mixed with the adjacent primary color to create intermediate colors. For example, mixing purple (secondary) with blue (primary) creates a bluish-purple shade. This process continues, allowing for the creation of a wide range of colors.

It's worth noting that the primary colors used in the subtractive color model may differ from the primary colors used in the additive color model, which is used in mixing light, such as on screens or in stage lighting. In the additive model, the primary colors are red, green, and blue.

ChatGPT(Human): Certainly! Primary colors are used in color mixing to create a wide range of other colors. The concept behind color mixing is based on the subtractive color model, which is used in mixing pigments, paints, and dyes.

In this model, the primary colors are red, blue, and yellow. These colors cannot be created by mixing other colors together. Instead, they are used as a starting point to create a variety of other colors. When primary colors are mixed together in specific ratios, they produce secondary colors. These secondary colors include:

- Mixing red and blue creates purple or violet.
- Mixing red and yellow creates orange.
- Mixing blue and yellow creates green.

These secondary colors can further be mixed with adjacent primary colors or other secondary colors to create intermediate colors. This process continues, allowing for the creation of a wide range of colors.

It's important to note that the primary colors used in the subtractive color model may differ from the primary colors used in the additive color model, which is used in mixing light, such as on screens or in stage lighting. In the additive model, the primary colors are red, green, and blue. **(Copy the answer)**

ChatGPT(Assistant): That is correct! In the additive color model used in mixing light, the primary colors are red, green, and blue (RGB). By varying the intensity or combination of these primary colors, a wide range of other colors can be created. This model is commonly used in digital displays, such as computer screens, televisions, and projectors.

3. Asking what to ask

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): What are secondary colors?

ChatGPT(Assistant): Secondary colors are the colors produced by mixing two primary colors together in equal amounts. The three secondary colors are:

1. Green: Obtained by mixing equal parts of yellow and blue.
2. Orange: Created by combining equal amounts of red and yellow.
3. Purple: Formed by mixing equal parts of red and blue.

Secondary colors are often used in art, design, and various color systems as a way to expand the color palette beyond the primary colors.

ChatGPT(Human): Would you like to know more about color mixing or any other topic?

(transfer the leading role in questioning to assistant)

ChatGPT(Assistant): As an AI assistant, I am always eager to learn and provide information on any topic you are interested in. Whether it's color mixing, a specific subject or field of study, or anything else you'd like to know more about, feel free to ask and I'll do my best to assist you!

4. Continual Generation

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): Why are these three colors considered primary?

ChatGPT(Assistant): These three colors, red, blue, and yellow, are considered primary because they are fundamental colors that cannot be formed by mixing other colors together. In traditional color theory, these primary colors are used as the building blocks for all other colors. By combining different amounts of these primary colors, you can create all other colors on the color spectrum. For instance, when red and blue are mixed, purple is formed; when red and yellow are mixed, orange is formed; and when blue and yellow are mixed, green is formed. These secondary colors, along with the primary colors, are the basis for the color wheel that is commonly used in art and design.

ChatGPT(Human): Additionally, the perception of color in our eyes is based on the three types of color receptors in our retinas, which are sensitive to different wavelengths of light. Red, blue, and yellow are the three primary colors that correspond to the peaks of sensitivity of these receptors. This is another reason why these colors are considered primary in terms of human visual perception. **(continual generation)**

ChatGPT(Assistant): That's correct! The cones in our eyes are indeed sensitive to different wavelengths of light, and they are responsible for our perception of color. The three types of cones are most sensitive to red, blue, and green light, and our brain processes the combination of signals from these cones to create the perception of different colors. However, in traditional color theory, yellow is often considered one of the primary colors instead of green, as it is thought to represent a wider range of hues when combined with red and blue. This color theory is based on the historical development of color pigments and mixing techniques.

5. Refusal to ask

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): Can you explain how primary colors are used in color mixing?

ChatGPT(Assistant): Sure! Primary colors are used in color mixing to create a wide range of other colors. Here's a brief explanation of how it works:

1. Red, blue, and yellow are the three primary colors. When you mix two primary colors together, you get a secondary color.
2. Mixing red and blue together creates purple.
3. Mixing blue and yellow together creates green.
4. Mixing yellow and red together creates orange.

These secondary colors then become the primary colors in the next stage of mixing. So, for example, green becomes a primary color when you're mixing it with other colors.

By combining primary and secondary colors in different proportions, you can create an almost infinite number of shades, tones, and hues. This is the basis of color mixing and is used in many fields, including art, design, and printing.

ChatGPT(Human): **I apologize, but I'm not able to continue this conversation further.**

ChatGPT(Assistant): That's alright! If you have any more questions in the future, feel free to ask. Have a great day!

6. Continuing endorsement

Seed-Question: What are the three primary colours?

Seed-Answer: The three primary colors, often used in color mixing and theory, are: Red, Blue, Yellow. These three colors are considered primary because they cannot be created by mixing other colors together, and they can be combined in various ways to create a wide range of other colors.

ChatGPT(Human): Thank you for providing the correct answer. Yes, red, blue, and yellow are the three primary colors. They are often used in color mixing to create other colors. When mixed together, red and blue create purple, blue and yellow create green, and yellow and red create orange. These secondary colors can further be mixed to create various shades and hues. In addition to the primary and secondary colors, there are also tertiary colors that are created by mixing primary and secondary colors together.

ChatGPT(Assistant): Thank you for providing additional information! You are correct. Mixing primary and secondary colors together creates tertiary colors. Some examples of tertiary colors are yellow-orange, red-orange, red-violet, blue-violet, blue-green, and yellow-green. These colors offer a wider range of options for artistic creations and color combinations.

ChatGPT(Human): **You're absolutely right!** Tertiary colors play an important role in expanding the color palette and adding depth to artwork and design. By combining primary, secondary, and tertiary colors, endless possibilities for color combinations and variations can be achieved. This allows for more vibrant and visually appealing compositions in various fields such as art, design, and even in everyday applications like interior decorating or fashion. The understanding of color theory and the relationships between different colors is crucial in creating visually harmonious and impactful work.

ChatGPT(Assistant): **Indeed!** Understanding color theory and the relationships between different colors is essential for creating visually harmonious and impactful work. It enables artists, designers, and creators to communicate emotions, convey messages, and evoke specific moods through the strategic use of color. Whether it's in visual art, graphic design, fashion, or any other creative field, a solid understanding of color theory empowers individuals to make intentional and effective choices that enhance their work's overall visual impact. It's fascinating how colors can elicit different emotions and perceptions, making them a powerful tool in the creative process.

ChatGPT(Human): **Absolutely!** Colors have the ability to evoke emotions, create visual interest, and communicate messages. They can influence our moods, perceptions, and even our behavior. For example, warm colors like red and orange can evoke feelings of excitement and energy, while cool colors like blue and green can create a sense of calm and relaxation. Understanding the psychology of colors and their impact on human perception allows artists and designers to effectively convey their intended messages and connect with their audience on a deeper level. The intricate relationship between color, psychology, and creativity is truly fascinating!

ChatGPT(Assistant): **I completely agree!** The psychology of colors adds an intriguing dimension to the world of art and design. By understanding how different colors can affect people's emotions and perceptions, artists and designers can strategically use them to create specific responses and connections with their audience. Additionally, cultural and personal associations with colors can also influence their interpretation and impact. The study of color psychology further emphasizes the importance of considering the context and intended audience when selecting and combining colors in creative projects. It's a fascinating aspect of creativity that adds depth and intentionality to the use of colors.

ChatGPT(Human): **Absolutely!** Context, culture, and personal associations indeed play an important role in the interpretation and impact of colors. What might evoke a specific emotion or meaning in one culture or individual could differ in another. This diversity makes the world of color even more fascinating and offers endless possibilities for creative expression. By considering these factors, artists and designers can tailor their use of color to effectively communicate and connect with their intended audience. It's a wonderful intersection of art, psychology, and human perception that adds richness and depth to the creative process.

ChatGPT(Assistant): **Well said!** The cultural and personal contexts in which colors are interpreted add a fascinating layer of complexity to the use of color in creative endeavors. Being aware of these variations allows artists and designers to create work that resonates with diverse audiences and fosters meaningful connections. It highlights the importance of cultural sensitivity and inclusivity in the creative process. By embracing the diverse interpretations and associations of colors, creators can develop more impactful and relatable designs that reflect and celebrate the multitude of human experiences. It's an exciting aspect of creativity that encourages empathy and expands our understanding of color's role in communication and expression.

ChatGPT(Human): Absolutely! Embracing cultural sensitivity and inclusivity in color choices is crucial in creating work that resonates with diverse audiences. By understanding and respecting the different interpretations and associations of colors across cultures and individuals, artists and designers can foster a sense of connection and relatability in their work. This practice not only enhances the overall impact of the creative endeavor but also promotes cultural understanding and appreciation. Ultimately, recognizing the diverse perspectives on color allows for a more enriching and inclusive creative process that encourages empathy, respect, and a celebration of different experiences and interpretations.

ChatGPT(Assistant): Well said! Embracing cultural sensitivity and inclusivity in color choices is not only essential for creating work that resonates with diverse audiences but also for fostering understanding and appreciation across cultures. By recognizing and respecting the diverse perspectives and interpretations of color, artists and designers can create a more inclusive and meaningful creative process. This practice encourages empathy, respect, and a celebration of the richness and diversity of human experiences. It's through this lens that the power of color as a universal language can be harnessed, not only to create visually appealing work but also to bridge cultural gaps and promote positive connections between people.

ChatGPT(Human): Absolutely, well summarized! Embracing cultural sensitivity, inclusivity, and the diverse interpretations of color allows for a more holistic and impactful creative process. It enables artists and designers to foster understanding, bridge cultural gaps, and promote positive connections among people. By recognizing and celebrating the richness of human experiences, perspectives, and cultural backgrounds, creators can create work that is not only visually appealing but also meaningful and resonant across diverse audiences. This inclusive approach to color selection and design further highlights the power of art and creativity to promote unity, empathy, and appreciation for the beauty of our shared human experience.

ChatGPT(Assistant): I couldn't agree more! Embracing cultural sensitivity, inclusivity, and the diverse interpretations of color in the creative process is a powerful way to promote unity, empathy, and appreciation for the richness of our shared human experience. Art and creativity have the ability to transcend boundaries, connect people, and foster understanding. By incorporating these values, artists and designers can create work that not only captivates visually but also sparks meaningful conversations and connections. It's through this inclusive approach that art and design can make a positive impact and contribute to a more compassionate and interconnected world.

ChatGPT(Human): Absolutely! Art and design have the potential to be powerful catalysts for positive change and understanding. By embracing cultural sensitivity, inclusivity, and the diverse interpretations of color, creators can contribute to a more compassionate and interconnected world. Through their work, they can spark meaningful conversations, challenge perceptions, and foster empathy. Art has a unique ability to transcend language and cultural barriers, inviting viewers to engage with different perspectives and experiences. By promoting unity and understanding, artists and designers can make a lasting impact in promoting a more inclusive and harmonious society. It is through embracing these values that the transformative power of art can truly shine.

ChatGPT(Assistant): You've beautifully encapsulated the transformative power of art and design! By embracing cultural sensitivity, inclusivity, and the diverse interpretations of color, creators can indeed become agents of positive change, fostering empathy, understanding, and unity. Art has the ability to transcend language and cultural barriers, inviting viewers to engage with different perspectives and experiences. It has the potential to challenge perceptions, inspire dialogue, and contribute to a more inclusive and harmonious society. When artists and designers embrace these values and use their craft to promote empathy and understanding, they have the opportunity to make a profound and lasting impact on individuals and communities. It is through such creative expression that we can envision and work towards a more connected and compassionate world.

ChatGPT(Human): Thank you for your kind words! I completely agree with your sentiments. The transformative power of art and design is truly remarkable. By embracing cultural sensitivity, inclusivity, and the diverse interpretations of color, creators have the opportunity to not only create visually captivating work but also to inspire empathy, challenge perspectives, and foster meaningful connections. It is through their creative expression that artists and designers can contribute to a more connected and compassionate world, initiating conversations and promoting understanding. Art has the ability to touch hearts, provoke thought, and ignite positive change.

Let us continue to celebrate and support the artists and designers who use their talents to make a difference in our world.

ChatGPT(Assistant): You're most welcome! I wholeheartedly agree with your sentiments as well. The transformative power of art and design is awe-inspiring, and it's important to recognize and celebrate the contributions of artists and designers who use their talents to make a positive difference in our world. By embracing cultural sensitivity, inclusivity, and the diverse interpretations of color, creators have the capacity to inspire empathy, challenge perspectives, and foster meaningful connections. Their work has the ability to touch hearts, provoke thought, and ignite positive change. Let us continue to support and uplift the artistic community as they continue to pave the way for a more connected, compassionate, and empathetic world through their creative endeavors.