

From Competition to Collaboration: Designing Sustainable Mechanisms Between LLMs and Online Forums

Extended Abstract

Niv Fono
Technion—Israel Institute of
Technology
Haifa, Israel
niv.fono@campus.technion.ac.il

Yftah Ziser
University of Groningen, NVIDIA
Research
Groningen, Netherlands
y.ziser@rug.nl

Omer Ben-Porat
Technion—Israel Institute of
Technology
Haifa, Israel
omerbp@technion.ac.il

ABSTRACT

While Generative AI (GenAI) systems draw users away from (Q&A) forums, they also depend on the very data those forums produce to improve their performance. Addressing this paradox, we propose a sequential interaction framework in which a GenAI system proposes questions to a forum that can publish some of them. Our framework captures several intricacies of such a collaboration, including non-monetary exchanges, asymmetric information, and incentive misalignment. We instantiate the framework via large-scale simulations on Stack Exchange data and open-source LLMs. We demonstrate the incentive misalignment empirically, yet show that players can achieve about half of the estimated full-information benchmark. Our results highlight the potential for sustainable collaboration.¹

KEYWORDS

Generative AI; Strategic Behavior; Collaboration

ACM Reference Format:

Niv Fono, Yftah Ziser, and Omer Ben-Porat. 2026. From Competition to Collaboration: Designing Sustainable Mechanisms Between LLMs and Online Forums: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. <https://doi.org/10.65109/STKG8016>

1 INTRODUCTION

Online Q&A forums are facing a participation decline as users increasingly rely on generative AI for immediate answers. For instance, Stack Overflow saw a sharp drop in posts following ChatGPT’s release [5]. This trend is not only a community-level challenge. It is also a systemic risk for GenAI itself, since Q&A forums supply high-quality human data that is widely used for training, evaluation, and benchmarking of large language models [2, 4, 9, 12]. A sustained decline may therefore weaken the knowledge sources that GenAI depends on, creating an adverse feedback loop.

Most proposed responses frame the relationship between GenAI firms and forums as adversarial, focusing on data restrictions or

compensation [1, 6]. In contrast, we study a collaboration mechanism that does not rely on monetary transfers and treats both sides as strategic agents with distinct utilities. Concretely, we consider a non-monetary exchange of goods: a GenAI provider routes questions it fails to resolve to a forum, and the forum selectively publishes those questions, enabling human experts to answer them. This can generate valuable supervision signals for GenAI while restoring engagement and traffic for the forum. A key constraint is that such collaboration must respect confidentiality: GenAI providers may not want to reveal their full set of uncertain queries, and forums may not want to reveal their curation policy.

To address this challenge, we model the collaboration as a repeated two-stage interaction. In each round, the GenAI provider proposes a limited set of candidate questions for publication, and the forum applies a selection rule to decide which questions to publish. Our work makes several key contributions toward understanding and modeling the interaction between generative AI systems and Q&A forums:

- **Realistic collaboration design.** We survey the key dimensions that shape GenAI-forum collaboration and identify three design principles: (i) no monetary transfers, (ii) incentive misalignment, and (iii) asymmetric information.
- **Game-theoretic framework.** We develop a game-theoretic framework that models GenAI providers and Q&A forums as strategic parties, each optimizing distinct yet interconnected utilities.
- **Data-driven game simulation.** We empirically instantiate the proposed framework through large-scale simulations of a realistic GenAI-Forum game, using real data from multiple Stack Exchange communities and open-source LLMs. Our analyses reveal a systematic misalignment between GenAI’s improvement potential (perplexity) and forum engagement value (view counts), supporting a strategic-interaction view (rather than a single-objective optimization). Our findings demonstrate that adaptive, game-theoretic interaction can produce substantial mutual gains even under asymmetric information.

2 PROBLEM FORMULATION

We consider two players: a GenAI provider (Player G) and a forum (Player F). The interaction proceeds for T rounds. Each round t has two stages. In Stage I, Player G receives a candidate pool Q_t and submits a subset $A_t \subseteq Q_t$ with $|A_t| \leq M$ (where $M \ll |Q_t|$). In Stage II, Player F observes A_t and applies a selection rule \mathcal{R} to produce the published set $S_t = \mathcal{R}(A_t)$, subject to a capacity constraint $|S_t| \leq K$.

¹The full paper is available online [7].



This work is licensed under a Creative Commons Attribution International 4.0 License.

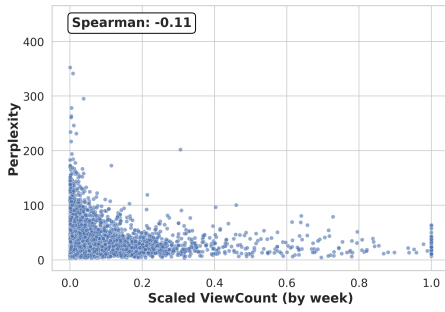


Figure 1: Perplexity vs Views for LLaMA 3.1 8B-Instruct and the Ubuntu community data.

Each published question q yields utilities $U_G(q)$ and $U_F(q)$ for Players G and F , respectively, which are privately known. We extend utilities to be additive over sets, so that the cumulative utilities are $U_G(\mathbf{A}, \mathcal{R}) = \sum_{t \in [T]} U_G(S_t)$ for $\mathbf{A} = (A_1, \dots, A_T)$, and similarly $U_F(\mathbf{A}, \mathcal{R}) = \sum_{t \in [T]} U_F(S_t)$ for Player F .

2.1 Full-information and Estimated Utility Recovery Rate

To contextualize achievable performance, we define an idealized full-information cooperative benchmark. In round t , if both utilities and the full pool Q_t were known, a natural cooperative objective is the Nash product [15], i.e., picking $S_t^* \in \arg \max_{S \subseteq Q_t, |S|=K} U_G(S) \cdot U_F(S)$. We prove that finding S_t^* is NP-hard, and thus resort to heuristic solutions. We estimate the optimal collaborative solution by $\tilde{U}_G = \max_{h \in \mathcal{H}} U_G(S_h)$ for Player G and similarly $\tilde{U}_F = \max_{h \in \mathcal{H}} U_F(S_h)$ for Player F , where \mathcal{H} is a set of efficient heuristics. Consequently, we define as our benchmark the *Estimated Utility Recovery Rates (EURR)*

$$EURR_G = \frac{U_G(\mathbf{A}, \mathcal{R})}{\tilde{U}_G}, \quad EURR_F = \frac{U_F(\mathbf{A}, \mathcal{R})}{\tilde{U}_F}. \quad (1)$$

3 EXPERIMENTAL SETUP

We instantiate the framework using data from five distinct Stack Exchange communities and open-source LLMs. We use 396,408 questions from five Stack Exchange communities, merged into a single time-aligned corpus. The interaction is simulated over weekly rounds from July 23, 2024 to July 23, 2025. For each round t , we construct the candidate pool Q_t as the set of questions posted during that week. The same questions in Q_t also provide the raw signals from which we compute both players’ utility proxies (defined next).

We proxy Player F ’s utility by (normalized) view count. Player F has a capacity constraint: at most $K = 50$ questions can be published per round, i.e., $|S_t| \leq K$. For Player G , we evaluate three open-source LLMs: **Pythia 6.9B** [3], **LLaMA 3.1 8B** [14], and **LLaMA 3.1 8B-Instruct** [11]. To the best of our knowledge, all evaluated models were released before the start of our data window (July 23, 2024) and thus could not have been trained on the Stack Exchange questions we use, reducing data contamination risks. We proxy GenAI utility by *perplexity* computed over the first 64 tokens (title + question content). Higher perplexity indicates higher model uncertainty [10, 17, 18] and correlates with prompt failure [8], OOD detection [16],

Table 1: Performance under asymmetric information.

LLM	Strategy	$EURR_F$	$EURR_G$
Pythia 6.9B	G-Greedy	0.55	0.505
Pythia 6.9B	G-Utility	0.664	0.521
Pythia 6.9B	Random	0.489	0.123
LLaMA 3.1 8B	G-Greedy	0.4	0.377
LLaMA 3.1 8B	G-Utility	0.555	0.463
LLaMA 3.1 8B	Random	0.45	0.113
LLaMA 3.1 8B-Instruct	G-Greedy	0.475	0.416
LLaMA 3.1 8B-Instruct	G-Utility	0.6	0.5
LLaMA 3.1 8B-Instruct	Random	0.482	0.115

and informative data selection [13].² We assume $M = 100$, i.e., Player G can propose at most $|A_t| = M$ questions per round.

We consider the following strategies for Player G : **G-Greedy** (top perplexity), **G-Utility Maximization** (maximize $U_G(q) \cdot \hat{\pi}(q)$, where $\hat{\pi}$ estimates $\Pr(q \in S_t \mid q \in A_t)$ from observed accept/reject feedback), and **Random**. For Player F , we construct \mathcal{R} using a classifier learned offline C and threshold θ , i.e., $S_t = \mathcal{R}(A_t) = \text{Top}_K \{q \in A_t : C(q) \geq \theta\}$. Finally, we compute the EURRs (Equation (1)) by estimating \tilde{U}_F and \tilde{U}_G via an idealized cooperative experiment in which players jointly select S_t .

4 RESULTS

Misalignment evidence. Across domains and models, perplexity (GenAI utility proxy) exhibits a weak negative association with normalized views (forum utility proxy). The trend is exemplified in Figure 1, where each point corresponds to a single question (with its perplexity on the vertical axis and normalized view count on the horizontal axis), and the reported Spearman coefficient is consistently negative across domains (-0.11 in this example). It indicates that questions that are most uncertain for the LLM are not typically the ones that maximize forum engagement, motivating the need for strategic interaction rather than naive maximization of a single objective.

Interaction outcomes. Table 1 reports utilities and EURR values for the asymmetric-information interaction under different Player G strategies and multiple LLMs. Across all evaluated LLMs, **G-Utility** yields the strongest performance for both players relative to **G-Greedy** and **Random**. For the forum, **G-Utility** increases normalized views substantially, reaching $EURR_F$ values between 0.555 and 0.664. For the GenAI side, **G-Utility** also improves cumulative perplexity and achieves $EURR_G$ between 0.463 and 0.521. In contrast, the random baseline yields very low GenAI utility recovery ($EURR_G \approx 0.11$ – 0.12) and does not reliably improve forum engagement. Overall, these results indicate that even under asymmetric information and incentive misalignment, a simple adaptive policy recovers 46% – 66% of the estimated benchmark.

ACKNOWLEDGMENTS

This research was supported by the Israel Science Foundation (ISF; Grant No. 3079/24). We thank the anonymous reviewers for their valuable feedback.

²While high perplexity might reflect low-quality or noisy inputs, we assume such cases are relatively rare in our setting, given the platform’s active moderation

REFERENCES

- [1] Ines Akaichi and Sabrina Kirrane. 2025. A comprehensive review of usage control frameworks. *Computer Science Review* 56 (2025), 100698. <https://doi.org/10.1016/j.cosrev.2024.100698>
- [2] Ebtesam Almazrouei, Hamza Alobeidli, Abdulaziz Alshamsi, Alessandro Cappelli, Ruxandra Cojocaru, M erouane Debbah,  tienne Goffinet, Daniel Hesslow, Julien Launay, Quentin Malartic, Daniele Mazzotta, Badreddine Noune, Baptiste Pannier, and Guilherme Penedo. 2023. The Falcon Series of Open Language Models. arXiv:2311.16867 [cs.CL] <https://arxiv.org/abs/2311.16867>
- [3] Stella Biderman, Hailey Schoelkopf, Quentin Anthony, Herbie Bradley, Kyle O'Brien, Eric Hallahan, Mohammad Aflah Khan, Shivanshu Purohit, USVSN Sai Prashanth, Edward Raff, Aviya Skowron, Lintang Sutawika, and Oskar van der Wal. 2023. Pythia: A Suite for Analyzing Large Language Models Across Training and Scaling. arXiv:2304.01373 [cs.CL] <https://arxiv.org/abs/2304.01373>
- [4] Damai Dai, Chengqi Deng, Chenggang Zhao, R. X. Xu, Huazuo Gao, Deli Chen, Jiashi Li, Wangding Zeng, Xingkai Yu, Y. Wu, Zhenda Xie, Y. K. Li, Panpan Huang, Fuli Luo, Chong Ruan, Zhifang Sui, and Wenfeng Liang. 2024. DeepSeekMoE: Towards Ultimate Expert Specialization in Mixture-of-Experts Language Models. arXiv:2401.06066 [cs.CL] <https://arxiv.org/abs/2401.06066>
- [5] R Maria del Rio-Chanona, Nadzeya Laurentsyeva, and Johannes Wachs. 2024. Large language models reduce public knowledge sharing on online Q&A platforms. *PNAS Nexus* 3, 9 (09 2024), pgae400. <https://doi.org/10.1093/pnasnexus/pgae400> arXiv:<https://academic.oup.com/pnasnexus/article-pdf/3/9/pgae400/59316621/pgae400.pdf>
- [6] Mike Felgenhauer. 2024. Competition for publication-based rewards. *Economics Letters* 244 (2024), 112017. <https://doi.org/10.1016/j.econlet.2024.112017>
- [7] Niv Fono, Yftah Ziser, and Omer Ben-Porat. 2026. From Competition to Collaboration: Designing Sustainable Mechanisms Between LLMs and Online Forums. arXiv:2602.04572 [cs.AI] <https://arxiv.org/abs/2602.04572>
- [8] Hila Gonen, Srini Iyer, Terra Blevins, Noah A. Smith, and Luke Zettlemoyer. 2024. Demystifying Prompts in Language Models via Perplexity Estimation. arXiv:2212.04037 [cs.CL] <https://arxiv.org/abs/2212.04037>
- [9] Xinyi Hou, Yanjie Zhao, Yue Liu, Zhou Yang, Kailong Wang, Li Li, Xiapu Luo, David Lo, John Grundy, and Haoyu Wang. 2024. Large Language Models for Software Engineering: A Systematic Literature Review. arXiv:2308.10620 [cs.SE] <https://arxiv.org/abs/2308.10620>
- [10] Hsiu-Yuan Huang, Yutong Yang, Zhaoxi Zhang, Sanwoo Lee, and Yunfang Wu. 2024. A Survey of Uncertainty Estimation in LLMs: Theory Meets Practice. arXiv:2410.15326 [cs.CL] <https://arxiv.org/abs/2410.15326>
- [11] Hugging Face. 2024. *Meta-Llama-3-8B-Instruct*. <https://huggingface.co/meta-llama/Meta-Llama-3-8B-Instruct> Accessed: 2025-10-08.
- [12] Haoyu Lu, Wen Liu, Bo Zhang, Bingxuan Wang, Kai Dong, Bo Liu, Jingxiang Sun, Tongzheng Ren, Zhuoshu Li, Hao Yang, Yaofeng Sun, Chengqi Deng, Hanwei Xu, Zhenda Xie, and Chong Ruan. 2024. DeepSeek-VL: Towards Real-World Vision-Language Understanding. arXiv:2403.05525 [cs.AI] <https://arxiv.org/abs/2403.05525>
- [13] Katerina Margatina, Timo Schick, Nikolaos Aletras, and Jane Dwivedi-Yu. 2023. Active Learning Principles for In-Context Learning with Large Language Models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, Houada Bouamor, Juan Pino, and Kalika Bali (Eds.). Association for Computational Linguistics, Singapore, 5011–5034. <https://doi.org/10.18653/v1/2023.findings-emnlp.334>
- [14] Meta AI. 2024. *Introducing Meta Llama 3.1: Our most capable models to date*. <https://ai.meta.com/blog/meta-llama-3-1/> Accessed: 2025-10-08.
- [15] John Nash. 1953. Two-Person Cooperative Games. *Econometrica* 21, 1 (1953), 128–140. <http://www.jstor.org/stable/1906951>
- [16] Jie Ren, Jiaming Luo, Yao Zhao, Kundan Krishna, Mohammad Saleh, Balaji Lakshminarayanan, and Peter J. Liu. 2023. Out-of-Distribution Detection and Selective Generation for Conditional Language Models. arXiv:2209.15558 [cs.CL] <https://arxiv.org/abs/2209.15558>
- [17] Pablo Tur n and Montse Cuadros. 2026. Perplexity, Uncertainty, and the Limits of Active Learning. In *Hybrid Artificial Intelligent Systems*, Emilio Corchado, H ctor Quinti n, Alicia Troncoso Lora, Hilde P rez Garc a, Esteban Jove P rez, Jos  Luis Calvo Rolle, Francisco Javier Mart nez de Pis n, Pablo Garc a Bringas, Francisco Mart nez  lvarez,  lvaro Herrero, Paolo Fosci, and Ramos S rgio Filipe (Eds.). Springer Nature Switzerland, Cham, 315–327.
- [18] Fanghua Ye, Mingming Yang, Jianhui Pang, Longyue Wang, Derek F. Wong, Emine Yilmaz, Shuming Shi, and Zhaopeng Tu. 2024. Benchmarking LLMs via Uncertainty Quantification. In *Advances in Neural Information Processing Systems*, A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (Eds.), Vol. 37. Curran Associates, Inc., 15356–15385. <https://doi.org/10.52202/079017-0491>