

A Business Idea Generation Framework Based on Creative Multi-Agent Discussions

Mizuki HOSHINO¹, Fuminori NAGASAWA¹, Shun SHIRAMATSU¹

¹Nagoya Institute of Technology

Abstract

Recent advances in large language model (LLM) have enabled various applications in idea generation. However, generating business ideas from patent information remains under-explored. We participated in the PBIG 2025 shared task, which required generating business ideas from patents in three fields.

In this study, we propose a multi-agent framework in which five types of agents cooperate in stages to support the generation of business ideas that include diverse perspectives. Each phase involves one or more agents with different roles. This framework begins with a discussion between agents who are given personas, leading to the generation, selection by ranking, and refinement of ideas. Compared to conventional method, we show that the proposed method can promote the creation of more diverse and in-depth ideas.

In comparison with the output of other teams by the organizers, our system performed well in terms of specificity and Innovativeness. Also compared with the baseline, idea refinement phase is effective to improve quality of idea. However, generating ideas solely with a single agent may restrict the diversity of idea.

1 Introduction

In recent years, the accuracy of the large language model (LLM) has improved dramatically, and research on applying LLM to idea generation has been actively conducted. In general, it has been shown that the quality of idea generation for a theme can be improved by acquiring and utilizing domain information through RAG (Retrieval-Augmented Generation). However, there has not yet been sufficient research on a framework for using patent information to think of business ideas.

In this paper, we report the results of participating in the Shared Task: Product business idea generation from patents (PBIG) and working on the

task of generating business ideas from patent information. Participants will be given 150 USPTO patents extracted from three fields: NLP, Computer_Science, and Material_Chemistry, and will develop a system that outputs JSON for each patent with four items (Product Title, Product Description, Implementation, and Differentiation) that make up a "product business idea that can be realized within three years." We participated in the Shared Task as Team Shiramatsulab.

We applied a multi-agent system as an approach to this shared task. In a multi-agent system, multiple agents interact with each other to solve problems. Multi-agent system is used as a group discussion to generate business ideas from patent information. We developed the system based on the hypothesis that better ideas can be generated by multiple agents who express their opinions from their own perspectives and broaden the range of ideas.

2 Related Works

2.1 LLM-based Multiagent System for Ideation

Su et al. (2025) introduce *Virtual Scientists (VirSci)*, an LLM-driven multi-agent framework that forms a virtual research team to *generate, evaluate, and iteratively refine* scientific ideas. Their five-stage workflow—collaborator selection, topic discussion, idea generation, novelty assessment, and abstract drafting—outperforms single-agent baselines in both novelty and impact metrics.

Nomura et al. (2024) implement a brainstorming support system in which multiple LLM-based agents each assume an ISSUE-IDEA-PROS-CONS (IBIS) role. By mimicking human group dynamics while a single user interacts with the agents, the system boosts the quantity and diversity of ideas without the production-blocking effects often seen in conventional group brainstorming.

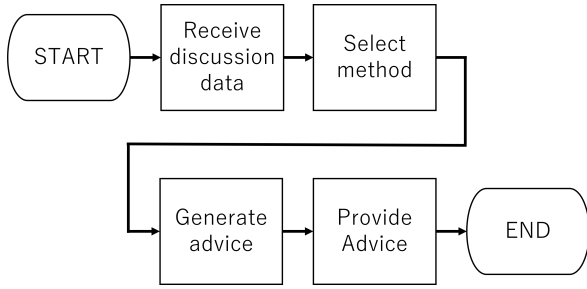


Figure 1: The process flow of ICS Agent

While the above systems each rely on a single ideation paradigm, our research integrates multiple, well-established idea-generation technique within a unified multiagent discussion environment. By orchestrating agents that embody these complementary heuristics and mediating their dialogue, we aim to provide broader creative coverage and finer-grained support for idea-centric debates.

2.2 Idea Creation Support Agent: ICS Agent

We developed an Idea Creation Support Agent (ICS Agent) that participates in group discussions to generate ideas, with the goal of enabling participants to consider ideas from a various perspectives (Hoshino et al., 2025). ICS Agent provides advice based on the state of the discussion. The advice is generated using an idea creation support method. One of three idea generation support methods (synectics method, search lighting method, and checklist method) is selected based on the state of the discussion. We included this idea generation support agent in a multi-agent discussion to enable discussions from more diverse perspectives and knowledge. The process flow of ICS Agent is shown in Figure 1.

3 System Overview

The overview of our system is shown in Figure 2. This system is divided into five phases: a persona generation phase, a multi-agent discussion phase, an idea generation phase, an idea evaluation phase, and idea refinement phase. The system uses gpt-4.1-mini.

3.1 Persona Generation Phase

In this phase, patent information is read and customer personas for discussion are created. Personas are created using LLM, and the following elements are determined (Table 1). In our system, three personas are generated and they join discussion as

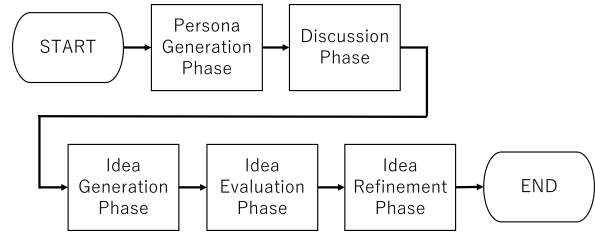


Figure 2: The process flow of our system

Table 1: Persona’s attributes

Element	description
id	Identification number
name	name in discussion
age	Persona’s age
gender	Persona’s gender
occupation	Persona’s occupation
lifestyle	Persona’s lifestyle
value	Important points
needs	Persona’s needs
pain_points	Issues persona thinks
purchasing_behavior	Buying patterns

Discussion Participating Agents (DP Agents). The Personas are used only DP Agents.

3.2 Discussion Phase

In this phase, a discussion is held using ICS Agent and DP Agents that are provided personas determined in the Persona Generation Phase. DP Agents think of business ideas from patent information and express their idea. The discussion proceed for a total of five turns, with each turn defined by all agents express idea. Between each turn, ICS Agent provides advice to DP Agents to encourage ideas from the discussion history. In turns 2–5, each Discussion Participating Agents thinks their ideas considering the discussion history and advice of the ICS Agent.

3.3 Idea Generation Phase

In this phase, Generation Agent generates multiple business ideas based on the discussion history. In our system, three ideas are generated from discussion history.

3.4 Idea Evaluation Phase

In this phase, Evaluation Agent evaluate the multiple ideas generated in Idea Generation Phase and select best idea. The evaluation was performed using LLM-as-a-Judge on six criteria, the same as the evaluation criteria for the shared task (technical

validity, innovativeness, specificity, need validity, market size, competitive advantage). Each idea is compared pairwise against the others and the final ranking is determined using Elo-based scoring.

3.5 Idea Refinement Phase

In this phase, DP Agent and Refinement Agent improves the idea selected in the Idea Evaluation Phase to make them more viable. DP Agents give their opinions on the selected idea from the persona’s perspective. DP Agents review idea from two perspectives: technical and business issues. Refinement Agent then uses these opinions to further update the idea. At the end of this phase, the final idea is output.

4 Evaluation / Results

We evaluated our system output and compared it with the baseline method provided in the Shared Task.

First, we compare the final output of our system against two baselines: (1) the output of baseline method provided by the organizers, and (2) the output before the Idea Refinement Phase.

Second, the final output submitted by our system were evaluated by the task organizers via pairwise comparisons against outputs of other participating teams, using Elo scores derived from LLM-as-a-judge.

4.1 Comparison of System Output with Baseline

To assess the effectiveness of our system, we evaluated the ideas generated by the following three methods: (1) Baseline: the baseline method uses baseline prompt provided by the task organizers, (2) Intermediate: our system’s intermediate output generated before Idea Refinement Phase, and (3) Final: our system’s final output submitted to the shared task.

We selected randomly 10 patents selected from each of the three fields and generated ideas for each patents. And compared with ideas generated by same patents each other. Ideas were evaluated using pairwise evaluation by LLM. To prevent bias in the LLM output, we conduct the evaluation of the ideas twice, with the order of the ideas swapped. In other words, the total number of comparisons is 20. The LLM model used was gpt-4.1-mini.

Tables 2 - 4 show the number of ideas that were judged to be superior when the ideas generated by each method were evaluated in pairs.

Table 2: Comparison result Baseline with Final

	Baseline	Final
cs	1	19
mc	12	8
nlp	2	18
total	15	45

Table 3: Comparison result Intermediate with Final

	Intermediate	Final
cs	15	5
mc	15	5
nlp	14	6
total	44	16

As shown in Table 2 and Table 4, Final output of our system may be more effective than Baseline method and Intermediate output. However, limited to the material_chemistry field, the output of the baseline method output is superior than Final output. Also, Table 4 shows that Intermediate output of our system was not more effective than Baseline method output. Therefore, it was suggested that the Idea Refinement Phase may contribute to improving the quality of ideas.

4.2 System Leader Board Results

The leaderboard results for the Share Task is shown in Table 5. These score were evaluated by the organizer. Table 5 displays the highest scores, our team’s scores, and the average scores for only five teams that submitted ideas in three fields. Each team’s evaluation score is calculated by averaging the scores in the three criteria for each patent field. In this comparison, we did not include teams that submitted ideas in only one or two fields for calculation. These results are based on evaluations by only LLM-as-a-Judge.

As shown in Table 5, generated ideas of our system is superior in Specificity and Innovativeness. On the other hand, evaluation scores in Market size and Technical validity are lower evaluation than

Table 4: Comparison result Baseline with Intermediate

	Baseline	Intermediate
cs	8	12
mc	13	7
nlp	7	13
total	28	32

Table 5: Evaluation score using LLM for each criteria

	Top	Our	Average
tech_valid	1110.7	993.3	1012.3
spec	1189.7	1035.3	1008.0
need_valid	1084.7	1011.3	1008.3
market_size	1060.7	985.7	1014.2
innov	1086.7	1027.7	1000.6
comp_adv	1140.0	1004.7	1001.7

the average.

5 Discussion

5.1 Idea generation ability against baseline

As shown in Table 2 - Table 4, while Idea Refinement Phase is effective to make quality of idea higher, intermediate idea is as same quality as baseline idea. This result likely reflects the design of Idea Generation Phase. In this phase, the ideas are generated by agent who is not participating in the discussion. As mentioned in Section 3, Idea Generation Agent does not have a specific persona. In addition, ideas are generated by a single agent. Because the Idea Generation Agent is not provided a persona, three similar ideas are generated, which may have prevented the idea from reflecting discussion as collective knowledge. Therefore, it is considered to have reduced diversity of the generated ideas. Similarly, a single agent that no persona has is used in the Idea Refine Phase. However, Refinement Agent refines idea by considering feedback from the DP Agents. It might be that by gaining insights into idea challenges and improvement suggestions from the diverse perspectives of DP Agents, Refinement Agent was able to improve quality of idea.

5.2 Evaluation in Shared Task Leaderboard

As mentioned in Section 4.2, ideas generated by our system is high scores in Specificity and Innovativeness. One contributing factor for Specificity may be DP Agents. Each DP Agents express their thoughts on business idea and discuss them. Through this discussion, idea details may become more specific.

In contrast, the scores in Market size and Technical validity are low. It may be considered that Market size evaluation is related to Specificity. As the discussions become more specific, the target market may become more limited and the market size may become smaller. Also regarding Technical validity, because our system did not conduct

technical analysis, it is possible that ideas difficult to realize within three years remained.

6 Conclusion

We proposed a multi-agent framework using LLM for generating business ideas from patents as part of the PBIG: Shared Task. Our system outputs were high score in Specificity and Innovativeness. In addition, compared with baseline method, it may have been suggested that our system output is superior than baseline. We found that views from agents with different personas are effective for quality of idea. However, it may be that generating idea by a single agent caused loss of diversity of idea. For future work, we plan to : (1) verify the effects of generating idea by each multi agents to its quality. (2) investigate the effects of the discussion structure and dialogue format between agents on the quality and diversity of ideas generated.

Acknowledgments

This study was supported by JST CREST (JP-MJCR20D1).

References

- Mizuki Hoshino, Sora Matsumoto, and Shun Shiramatsu. 2025. Development and evaluation of llm-based ideation support agent for idea generation discussions. In *Proceedings of the 39th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI 2025)*, Osaka, Japan. Japanese Society for Artificial Intelligence. Paper ID: 3J4-GS-5-02 (in Japanese).
- Moeka Nomura, Takayuki Ito, and Shiyao Ding. 2024. Towards collaborative brainstorming among humans and AI agents: An implementation of the IBIS-based brainstorming support system with multiple AI agents. In *Proceedings of the ACM Collective Intelligence Conference (CI 2024)*, Boston, MA, USA.
- Haoyang Su, Renqi Chen, Shixiang Tang, Zhenfei Yin, Xinzhe Zheng, Jinzhe Li, Biqing Qi, Qi Wu, Hui Li, Wanli Ouyang, Philip Torr, Bowen Zhou, and Nanqing Dong. 2025. Many heads are better than one: Improved scientific idea generation by a LLM-based multi-agent system. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)*.