Team NS_NLP at the AgentScen Shared Task: Structured Ideation Using Divergent and Convergent Thinking

Hayato Yoshiyasu
Nippon Shokubai Co., Ltd., 5-8 Nishi Otabi-Cho, Suita, Osaka 564-8512, Japan
hayato_yoshiyasu@shokubai.co.jp

Abstract

This paper presents our participation report for the Shared Task "Product Business Idea Generation from Patents" [1] conducted at The 2nd Work-shop on Agent AI for Scenario Planning - IJCAI 2025, as Team NS_NLP. In this study, we explore a method that combines divergent and convergent thinking in a stepwise reasoning process, supplemented with external information, to generate business ideas based on patent data. As a result, our approach achieved first place in several criteria within the Materials Chemistry category, based on evaluation conducted by both LLMs and human experts.

1 Introduction

In today's business environment, characterized by the VUCA era—Volatility, Uncertainty, Complexity, and Ambiguity—companies are increasingly required to make rapid and flexible decisions and respond strategically to constant change. Under such conditions, the continuous generation of innovative business ideas is essential for creating new value and maintaining adaptability.

However, the process of generating business ideas still heavily relies on human experience and intuition, which presents several challenges. First, ideas are frequently shaped by individual knowledge and prior experiences, they tend to be biased and constrained by existing frameworks. Second, integrating and analyzing large and diverse information sources—such as technical data, market trends, and customer needs—is essential but challenging for humans to perform efficiently.

In addition to these challenges, the generation of high-quality business ideas demands a wide range of skills, including the ability to create novel concepts, evaluate them objectively, and continuously gather relevant information. These demands present a substantial challenge to the continuous generation of valuable and innovative ideas, thereby impeding efforts to enhance corporate value through sustained business innovation.

Considering these difficulties, growing attention has been directed toward the use of generative AI technologies, particularly large language models (LLMs). Recent studies have demonstrated the effectiveness of LLMs in complex intellectual tasks such as generating scientific hypotheses and discovering novel knowledge (Wang et al., 2024; Si et al., 2025). These findings suggest that LLMs possess strong potential for supporting idea generation through more autonomous and progressive reasoning processes in the future.

Several studies have explored the use of LLMs for generating research themes and product ideas. For example, some approaches leverage scientific literature by collecting and fine-tuning on papers to generate ideas grounded in scholarly knowledge (Wang et al., 2024; Porsdam et al., 2023), while others utilize patent data as an alternative source of domain-specific information (Zhu et al., 2022). These studies suggest that augmenting the prior knowledge of LLMs with domain-relevant data can enhance the novelty and relevance of the generated ideas. However, challenges remain in terms of evaluating the practicality and feasibility of the generated ideas, as well as in achieving sufficient technical depth.

Motivated by the challenges discussed above, we participated in the Shared Task "Product Business Idea Generation from Patents (PBIG)" [1] held at The 2nd Workshop on Agent AI for Scenario Planning (AgentScen) - IJCAI 2025, specifically in the Materials Chemistry category.

[1] https://sites.google.com/view/agentscen/shared-task

2 Task Description

2.1 Product business idea generation from patents (PBIG)[1]

The aim of the PBIG shared task is to generate viable product business ideas utilizing patent information. Generating business ideas requires diverse capabilities, including a deep understanding of relevant domains, user need identification, and creative concept integration. If LLMs can support the generation of business ideas that are both innovative and viable, they may serve as a promising means to accelerate AI-driven innovation.

In this task, participants are required to generate four explanatory texts of a business idea. A "business idea" is defined as a concept for a product or service that utilizes patented technology and is realistically implementable within a three-year timeframe. The required outputs are as follows:

1. Product Title: A concise name for the product.
2. Product Description: A brief explanation of the product's key features and functions, target users, their needs, and the benefits provided.
3. Implementation Method: A description of how the patented technology is applied to the product.
4. Differentiation Points: An explanation of how the product is unique compared to existing solutions and what makes it stand out.

2.2 Dataset

Participants were provided with a dataset containing full texts and diagrams of patents. Fifty patents were selected from the USPTO for each of three domains: Natural Language Processing, Computer Science, and Materials Chemistry.

Experts curated the patents based on technical feasibility and diversity of potential product ideas, aiming to facilitate the generation of practical and diverse ideas.

2.3 Evaluation Metrics

The generated ideas were evaluated by both human experts and LLMs. In the human evaluation, each idea was first scored based on predefined criteria, and these scores were subsequently used to perform pairwise comparisons between different generation methods. In contrast, the LLMs evaluation employed two approaches: a direct pairwise comparison and a score-based pairwise comparison analogous to the human evaluation process.

Finally, Elo ratings were computed based on the comparison results to establish a ranking of the idea generation methods according to their relative performance. Each idea was evaluated across the following six criteria.

1. Feasibility: Whether the patented techno-logy is appropriate for the product, can be implemented, and is realistically achievable within three years.
2. Novelty: Whether the patented technology offers a new solution to an existing demand.
3. Specificity: Whether the idea is concrete and clearly articulated.
4. Necessity: Whether the proposed solution addresses a genuine user need.
5. Market Potential: Whether the market is sufficiently large and has a substantial number of potential users.
6. Competitive Advantage: Whether the use of the patented technology provides a business advantage over existing solutions.

3 Methodology

This section outlines our approach to the Shared Task. Recent studies have shown that step-by-step prompting, known as Chain of Thought (CoT), is more effective than single step prompting for complex reasoning tasks (Wei et al., 2022) Based on these findings, a stepwise reasoning strategy was employed to generate business ideas. Our prompting approach leverages not only CoT reasoning but also incorporates both divergent and convergent thinking, which are known to facilitate creative ideation (Kim et al., 2013).

To address the limitations of relying solely on the model's prior knowledge—which can lead to biased outputs (Shah et al., 2024)— our approach is designed to incrementally generate content by incorporating supplementary information as needed. This process enables a stepwise development of patent-derived technologies into more well-grounded and practically viable business ideas.

An overview of our approach is illustrated in Figure 1. The process comprises seven sequential steps, each of which is described in detail in the following sections. The specific prompts used at each step are provided in the Appendix B.1~7

3.1 Step 1. Patent Analysis

In the first step, the functional properties of the materials described in each patent and their potential application markets are extracted. To ensure consistency in the granularity of the output, a few-shot prompting strategy was employed (Brown et

Figure 1: Workflow of our proposed approach

al., 2020). Furthermore, the prompt was formulated to differentiate between the material's inherent functions and those introduced or enhanced by the patented technology.

3.2 Step 2. Term Refinement

The extracted functional and market terms occasionally included overly broad or abstract expressions—such as "automotive industry"—that lacked sufficient specificity. To address this issue, we introduced a filtering step using an LLMs to identify and exclude such high-level concepts. Specifically, pairs of extracted terms were input into the model, which was prompted to infer whether a hypernym–hyponym (i.e., hierarchical) relationship existed between them. If such a relationship was identified, the model returned the term pair along with a confidence score ranging from 0 to 1. Terms identified as higher-level concepts with a confidence score above a predefined threshold (set to 0.9 in this study) were excluded from the final output.

3.3 Step 3. Market Ideation

Based on the refined information, we generated potential market domains. The prompt was designed to diversely generate 5 to 10 candidate market domains that satisfy the two conditions: (1) the functions described in the patent correspond to existing market needs, and (2) the functions improved or enhanced by the patented technology address known challenges. To encourage the generation of novel market ideas, the prompt also included an instruction to exclude any markets already mentioned in the original patent.

3.4 Step 4. Idea Generation

For each market identified in Step 3, the prompt was designed to generate ideas using the patented technology. It also encouraged analysis of market trends and challenges to create ideas that address real-world needs.

Furthermore, to prevent the incorporation of unrelated technologies (e.g., IoT or AI) and to ensure that the generated ideas reflect the intrinsic value of the patented technology, the prompt was constrained to utilize only the technologies explicitly described in the patent.

3.5 Step 5. Information Retrieval

To complement and enhance the business ideas generated in Step 4, we developed an agent-based information retrieval pipeline. Specifically, we integrated LLM with web search capabilities via SerpAPI[2] and implemented an agent based on the ReAct framework (Yao et al., 2023). This agent is capable of dynamically retrieving and synthesizing external information as needed, including market size and growth rate, major competitors, and the technological advantages held by those companies—factors that are essential for evaluating the feasibility and potential of the proposed business ideas.

3.6 Step 6. Idea Evaluation

In Step 6, we designed an evaluation process to identify the most promising business idea from those supplemented with external information. Rather than conducting a simultaneous comparison of all candidates, we adopted a tournament-style pairwise comparison approach. The LLM was prompted to evaluate each pair of business ideas based on the criteria defined in the "Evaluation Metrics" section.

3.7 Step 7. Output Generation

In the final step, the selected business idea was formatted according to the output specifications defined by the Shared Task. Given the character limits imposed by the submission format, the prompt explicitly instructed the model to adhere strictly to both the structural and length constraints. Additionally, a verification mechanism was implemented to ensure output length. If the generated output exceeded or fell short of the specified character limits, it was automatically regenerated.

[2] https://serpapi.com/

Table 1: Results of comparative experiment (three-point rating)

Metrics	Baseline	Our Method
Technical validity	2.0	2.0
Innovativeness	2.0	**2.5**
Specificity	2.0	**2.9**
Need validity	2.0	**2.3**
Market size	2.0	1.8
Competitive advantage	2.0	2.0

Table 2: Elo Rating and Ranking results
※ () indicates participant ranking.

Metrics	LLMs	Human
Technical validity	971 (5)	1000 (3)
Innovativeness	1152 (3)	**1017 (1)**
Specificity	1017 (4)	1017 (2)
Need validity	**1129 (1)**	1007 (3)
Market size	997 (4)	**1017 (1)**
Competitive advantage	1055 (2)	997 (3)

4 Experiments

A comparative experiment was conducted within our team against the baseline method defined in this task, which generates business ideas using prompt engineering techniques. For this experiment, five patents were selected from the 50 patents in the Shared Task dataset, specifically from the Materials Chemistry category. Four researchers reviewed the content of each patent and compared the business ideas generated by both the baseline method and the proposed method.

The generated ideas were evaluated according to the criteria defined in the "Evaluation Metrics" section, using a three-point scale:

1. Inferior to the baseline
2. Equivalent to the baseline
3. Superior to the baseline

This evaluation enabled us to assess the relative advantages of the proposed method over the baseline and to verify its suitability for generating the final set of 50 business ideas.

For implementation, we constructed the processing flow using LangChain[3] and LangGraph [4] and employed GPT-4.1 [5] (gpt-4.1-2025-04-14) provided by OpenAI as the underlying LLM. To ensure structured outputs across multiple stages of the workflow, prompt engineering techniques were applied where appropriate (Marvin et al., 2023). (The specific prompt is provided in Appendix A.)

5 Results

Table 1 presents the results of the comparative experiment, showing the average scores for each evaluation criterion. These scores were assessed by researchers across five selected patents, enabling a comparison between the baseline method and the proposed method. The proposed method outperformed the baseline in terms of Innovativeness, Specificity, and Need validity. In contrast, it showed lower performance in Market size. For Technical validity and Competitive advantage, both methods performed at a comparable level.

Table 2 summarizes the results of the Shared Task, presenting the Elo ratings and rankings based on the submitted ideas (All results in Appendix C.). Although discrepancies were observed between ratings provided by LLMs and human evaluators, the proposed method achieved the highest scores in Innovativeness, Need Validity, and Market Size. Conversely, criteria such as Technical Validity and Competitive Advantage yielded Elo ratings below the initial baseline value of 1000.

6 Conclusion

This paper presented our method to the Shared Task: Product Business Idea Generation from Patents (PBIG). We developed a step-by-step workflow for generating business ideas by applying both divergent and convergent thinking based on the patented technology.

As a result, the proposed method achieved the highest scores in several criteria, demonstrating its effectiveness in generating business ideas. However, the evaluations for Technical Validity and Competitive Advantage were relatively low, indicating remaining challenges. Since these criteria are likely to require information beyond patent data, future work should focus on the accumulation and utilization of supplementary information sources suitable for idea generation.

[3] https://www.langchain.com/langchain

[4] https://www.langchain.com/langgraph

[5] https://platform.openai.com/docs/models/gpt-4.1

References

Wang, Q., Downey, D., Ji, H. and Hope, T. 2024. Scimon: Scientific inspiration machines optimized for novelty. *In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), pages 279-299.

Si, Chenglei, Diyi Yang, and Tatsunori Hashimoto. 2024. Can llms generate novel research ideas? a large-scale human study with 100+ nlp researchers. *arXiv preprint arXiv:2409.04109*.

Porsdam Mann, S., Earp, B. D., Møller, N., Vynn, S. and Savulescu, J. 2023. AUTOGEN: A personalized large language model for academic enhancement—Ethics and proof of principle. The American Journal of Bioethics 23.10: 28-41.

Zhu, Q., Luo, J. 2023. Generative Design Ideation: A Natural Language Generation Approach. In: Gero, J.S. (eds) Design Computing and Cognition'22. DCC 2022. Springer, Cham. https://doi.org/10.1007/978-3-031-20418-0_3

Elo. The Rating of Chessplayers, Past and Present. Ishi Press, 1986

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, QV. and Zhou, D. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems 35: 24824-24837.*

Kim, K.H., Pierce, R.A. 2013. Convergent Versus Divergent Thinking. *In: Carayannis, E.G. (eds) Encyclopedia of Creativity, Invention, Innovation and Entrepreneurship. Springer, New York, NY.* https://doi.org/10.1007/978-1-4614-3858-8_22

Shah, Chirag. 2024. From Prompt Engineering to Prompt Science with Humans in the Loop. *Communications of the ACM.*

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... and Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems, 33, 1877-1901.*

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K. and Cao, Y. 2023. React: Synergizing reasoning and acting in language models. *In International Conference on Learning Representations (ICLR).*

Marvin, G., Hellen, N., Jjingo, D., & Nakatumba-Nabende, J. 2023. Prompt engineering in large language models. *In International conference on data intelligence and cognitive informatics (pp. 387-402). Singapore: Springer Nature Singapore.*

A Format Instructions

Our methodology incorporates specific format instructions at Steps 1, 2, 3, 6, and 7 of the process.

```
The output should be formatted as a JSON instance that conforms to the JSON schema below.

As an example, for the schema {"properties": {"foo": {"title": "Foo", "description": "a list of strings", "type": "array", "items": {"type": "string"}}}, "required": ["foo"]}
the object {"foo": ["bar", "baz"]} is a well-formatted instance of the schema. The object {"properties": {"foo": ["bar", "baz"]}} is not well-formatted.

Here is the output schema:
```
{Instructions for each step}
```

## B Prompt Templates

### B.1 Step 1. Patent Analysis

```
System: You are an excellent patent analyst. Please extract information from given patent document.

Human: {document}
```

B.2 Step 2. Term Refinement

```
System: You are a knowledgeable assistant in linguistics and terminology.
Determine whether the given two terms have a hierarchical relationship:

Human:
Term A: {term_a}
Term B: {term_a}
```

B.3 Step 3. Market Ideation

```
System: Given the following two sets of properties, identify 5 to 10 potential markets, applications, or industries where:
- The 'General properties' are in demand or required
- The 'Distinctive properties' represent current challenges, unmet needs, or innovation opportunities
```
```

Additionally, consider the list of 'Existing markets' provided. Exclude these from your suggestions to avoid redundancy.
Focus on real-world use cases and emerging needs. Return your results as a structured list of potential market sectors or product categories that are not already covered by the existing markets.
(e.g., 'fuel-related parts like gasoline tanks and valves', 'automotive parts exposed to cleaners', 'sliding components in AV and OA fields', 'binder resin compositions for metal powders')
Human:
General properties:
{General properties}
Distinctive properties:
{Distinctive properties}
Existing markets to exclude:
{Existing markets}

B.4 Step 4. Idea Generation

System: You are a business strategist. Based on the following patent document and the potential market, generate a new business idea that leverages the patented technology.
Analyze the current trends, critical challenge, and innovation gaps in the listed markets. Then, propose a business concept that addresses these needs using the core invention described in the patent.
Do not incorporate unrelated technologies (e.g., IoT, AI, blockchain) unless they are explicitly part of the patented invention.

Your response should include:
- A compelling business idea title
- A clear and concise description of the business model
- How the patented technology is used
- The target customer segment
- The value proposition and competitive advantage

Avoid repeating ideas that are already common in the listed markets and proposing service-based or platform-based models unless they are directly derived from the patented invention.

Human:
Patent document: {document}
Potential markets: {market}

B.5 Step 5. Information Retrieval

System: You are a market research analyst. Your task is to evaluate the commercial potential of the following business idea.
Please use external search tools to gather and summarize the following information:
1. Current market size and CAGR of the relevant industry to 2027
2. Key competitors and their offerings, competitive advantages, distinctive features, proprietary technology
3. Competitive advantages, distinctive features, and proprietary technology of each competitor
4. Major growth drivers and market trends
5. Key challenges and barriers to entry
6. Regulatory or technological considerations
Provide a concise and structured summary.
Human: {business idea}

B.6 Step 6. Idea Evaluation

System: ## Input
Read two product business ideas using the technology.
<idea id='1'>{idea_1}</idea>
<idea id='2'>{idea_2}</idea>
Task
Your task is to choose the better idea from the perspective of criteria.
Evaluation Criteria:
- Innovation: How novel or groundbreaking is the business idea? Does it offer new solutions or improvements to existing problems?
- Feasibility: How practical and achievable is the business idea? Are the necessary resources, skills, and technologies available for implementation?
- Specificity: How clearly defined and detailed is the business plan? Does it address specific market needs, customer segments, and implementation steps?
- Market Size: How large is the potential market for the business? Does the idea target a growing or underserved market with high potential for expansion?

Table 3: Results of LLMs evaluation

Teams	Technical validity	Innovativeness	Specificity	Need validity	Market size	Competitive advantage
AiAnonymous	1038	782	883	892	944	941
ditlab	1021	1052	1067	1093	1050	1011
MK2	**1132**	**1207**	**1184**	1125	**1118**	**1146**
NS_NLP	971	1152	1017	**1129**	997	1055
Shiramatsulab	998	920	986	1006	1024	967
Team_MCG_DSN	896	1185	1112	946	939	1002
TrustAI	940	697	748	806	924	874

Table 4: Results of human evaluation (- means not evaluated)

Teams	Technical validity	Innovativeness	Specificity	Need validity	Market size	Competitive advantage
AiAnonymous	-	-	-	-	-	-
ditlab	996	1002	**1047**	**1035**	1009	**1038**
MK2	1017	990	1010	989	1013	991
NS_NLP	1000	**1017**	1017	1007	**1017**	997
Shiramatsulab	-	-	-	-	-	-
Team_MCG_DSN	928	1009	950	1026	1006	974
TrustAI	**1057**	978	973	941	952	998

```
- Competitive Advantage: How does the business plan stand out from competitors? What unique factors give it an edge in the market?

## Output
```

B.7 Step 7. Output Generation

```
System: Based on the provided information, generate a concise and information-rich business idea consisting of the following four components:
- title (80-100 characters): A short, compelling name for the product or solution.
- description (260-290 characters): A compact summary of the product's key features, target users, their needs, and the benefits.
- implementation (260-290 characters): A brief explanation of how the patented technology will be applied in the product.
- differentiation (260-290 characters): A clear statement of what makes the product unique compared to existing solutions.
Each field must be written clearly and concisely, strictly limited to the specified character count. Prioritize clarity, specificity, and value delivery. Avoid vague or generic language. If any field is too short or too long, regenerate that field until it meets the requirement.
Use clear, specific, and value-driven language. Avoid vague or generic expressions.
Human:
# business plan
{business plan}
```

C Evaluation Results

The evaluation results of all participating teams in the Shared Task are presented. Table 3 shows the results based on evaluations conducted by LLMs, while Table 4 presents those based on human evaluators. In both cases, Elo ratings were calculated based on either pairwise comparisons or scoring-based comparative evaluations across six criteria. The detailed criteria for scoring are publicly available on GitHub[6].

6 https://github.com/stockmarkteam/pbig-shared-task/tree/main/guidelines