Team MCG DSN at the AgentScen Shared Task: Knowledge Integration and Self-Improvement via LLMs for Generating Business Ideas from Patent Documents

Masaya Shimanuki, Naoto Shimizu, Kentaro Kinugasa, Hiroki Sugisawa*

Mitsubishi Chemical Corporation

Abstract

Patents represent valuable sources of commercial potential; however, generating viable business idea from such information requires advanced expertise. This paper proposes a prompt-based framework that integrates patent element, inventor profile, market potential, and business model grounded in TRIZ theory to generate high-quality business ideas. Furthermore, we introduce a self-improvement mechanism that extracts AI judges' personas from results of the PBIG competition-based evaluation and incorporates these insights to refine subsequent generations of ideas. While our output demonstrated strong performance under AI-based evaluation, notable discrepancies with human judgment were observed, highlighting the need for further alignment with human evaluation.

1 Introduction

Generating business ideas from patent documents poses a critical challenge in contemporary industrial innovation. While patents serve as rich and accurate repositories of technical knowledge, transforming this information into concrete product concepts or viable business models that achieve market acceptance demands advanced domain expertise and multifaceted evaluation [1]. Although recent studies have introduced automated generation methods using large language models (LLMs) [1,2], a unified framework that concurrently integrates technical feasibility and commercial viability remains lacking.

In this study, we propose a prompt-based model that emulates the domain knowledge of business expertise to generate higher-quality ideas [3]. Our approach consists of a five-module prompt flow: (1) extraction of core technical elements via patent component analysis, (2) assessment of the inventor's strengths, (3) ranking of market potential and applicability, (4) construction of business models guided by TRIZ principles, and (5) iterative refinement through AI-driven pairwise evaluations, during which judge personas are extracted and additional constraints are applied for self-improvement.

2 Methodology

2.1 PBIG Task

The Patent Business Idea Generation (PBIG) [4] shared task is a competition aimed at generating business ideas from patents using generative AI, which are feasible for market launch within three years. A dataset of 150 USPTO patents (50 each from the domains of natural language processing, computer science, and materials chemistry) is provided, including JSONL metadata (title, application number, publication number, publication date, abstract, claims, and description), PDF documents, and figure images. Participating systems must output four JSONL formatted fields with strict length limits: Product Title (≤ 100 characters), Product Description, Implementation, and Differentiation (each ≤ 300 characters) and may leverage external resources to enhance idea

* Email: hiroki.sugisawa.ma@mcgc.com
The name, MCG DSN indicates the Data Scientist Network (DSN) in Mitsubishi Chemical Group (MCG).

diversity and practical relevance. Evaluation combines human expert judgments with three AI based automated judges in a pairwise comparison of ideas originating from the same patent, and final rankings are computed using the Elo rating method.

2.2 Moduled Prompt Flow

We propose a domain knowledge-based prompt flow to guide generative AI in producing higher-quality business ideas from a patent [3]. This flow consists of five sequential Modules; all prompt templates are provided in the Appendix A.

(1) **Patent Element Analysis**: Prompt **1** provide an analytical framework to deconstruct the patent and organize its technical contents. From a JSONL‐formatted patent record, we extracted (i) the inventor's name, (ii) the intrinsic value of the patent, (iii) its applicability domains, and (iv) a concise summary. These outputs were then forwarded to Modules 2–5.

(2) **Inventor Profiling**: Prompt **2** employs a structured inference template to identify the inventor's (company, organization, or individual) core strengths, based on web search results. The investigation was conducted by the Gemini API to retrieve and synthesize publicly available information.

(3) **Market and Application Analysis**: Using Prompt **3**, we enumerated potential markets and applications for the patented technology and rank them according to projected profitability and societal impact. The top candidates were presented in tabular form, and the highest-priority market / application was selected for further modules.

(4) **Business Model Construction**: Leveraging TRIZ theory via Prompt **4**, we constructed candidate business models that maximize revenue by capitalizing on the patent's identified strengths (from Module 1) within the selected market / application (from Module 3).

(5) **Business Idea Proposal**: Finally, integrating outputs from Modules 1–4 via Prompt **5**, we generated a concrete product and business model proposal. Each proposal was evaluated against six criteria: technical validity, innovativeness, specificity, need validity, market size, and competitive advantage.

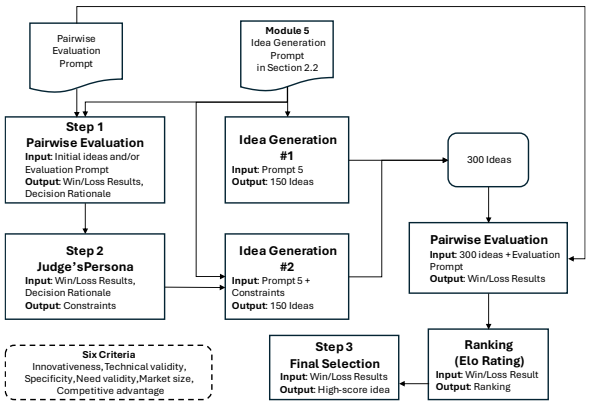
2.3 Self-Improvement via AI Judge

To further refine the business ideas generated in Section 2.2, we implemented the following self-improvement steps within an AI judge (Scheme 1):

Step (1) Initial Pairwise Evaluation (Prompt 6): We employed the o3-mini model to perform pairwise comparisons of business ideas, using six criteria. To reduce computational costs, comparison pairs were selected via random sampling. For each comparison, the judge outputs a justification, including the winning idea's identifier and a brief rationale. All ideas were subsequently ranked using Elo ratings [5,6].

Step (2) Judges' Persona Extraction: In Prompt **7**, we extracted "judge personas" from the justifications generated in Step 1, capturing judge's thinking patterns and prioritization criteria. Building upon these personas, Prompt **8** generated generalized supplementary constraints reflecting the judges' perspectives. These constraints were incorporated into Prompt **5** to guide the model in producing outputs that avoid impractical ideas and better fulfill all evaluation criteria.

Step (3) Final Selection: A total of 300 business ideas were generated—150 each from Module 5 in Section 2.2 and from Steps 1–3 in this section. Elo scores were computed for all outputs, and the highest-scoring idea in each category was selected as the final submission. Outputs that violated predefined guidelines, such as character limits, were manually excluded from consideration.



Scheme 1. Self-improvement via AI judge

3 Results & Discussion

3.1 Self-Improvement via AI Judge

In this study, we focused on the category of materials chemistry and conducted Step (1) using approximately 20 randomly selected patents. An example output of Step (1) is presented in Listing 1. Through pairwise comparisons of business ideas, Prompt **6** succeeded in generating pseudo-rationale to why one idea (Idea A) was considered superior to another (Idea B).

In Step (2), we applied Prompt **7** to the list of pseudo-rationales to extract AI judges' personas, which are described in detail in Appendix B. As a result, seven key evaluation criteria were identified as influential in the assessment of AI-generated business ideas: (1) technical rigor, (2) feasibility, (3) market applicability, (4) regulatory compliance, (5) specificity, (6) innovativeness, and (7) intellectual property defensibility. Specifically, technically oriented judges tended to favor proposals that adhered closely to patent content and included quantitative specifications (e.g., Young's modulus ≥ 2700 MPa or additive concentrations of 0.1–2 wt%), while abstract or loosely related ideas were consistently disfavored. Judges emphasizing feasibility preferred ideas that were compatible with existing infrastructure and realistically implementable within a two- to three-year timeframe; they rated proposals requiring large-scale investment or long development cycles lower. Market- and regulation-oriented personas prioritized ideas targeting large-scale, regulation-heavy markets (e.g., automotive, healthcare, home appliances) and those addressing pressing regulatory frameworks such as Euro 7 and VDA 278. Furthermore, ideas with detailed technical specifications (such as chemical structures, process conditions, catalysts, and operational temperature ranges) were positively evaluated, as were inventions exhibiting originality and resistance to imitation. Conversely, vague or overly generic ideas were consistently rated unfavorably.

Subsequently, in Prompt **8** we elicited the PBIG task's evaluation criteria from the pseudo-judge personas and incorporated the insights thus obtained as constraints into Prompt **5**. Although this approach succeeded in securing favorable evaluations from the AI reviewers on the PBIG task (Section 3.2), the imposed constraints sometimes proved counterproductive, leading to instances in which prescribed character limits were ignored or unrealistic numerical values were adopted as shown in Listing 2. This tendency was especially pronounced in the chemistry domain, where it is difficult to idealize the materials properties, thereby emphasizing the necessity of grounding the task's parameters in practical realities.

Idea A leverages the patent's chemistry to achieve ultra-low (<1 ppm) formaldehyde VOC release, a performance level not explicitly disclosed in the patent claims or typical commercial POM offerings. **Idea B**'s focus on mechanical strength and acid resistance closely matches multiple embodiments already taught in the patent, offering less novelty. Thus Idea A demonstrates the higher innovative leap.

Listing 1: Example of justification.

"**product_description**":"Drop-in molded quick-connects for gasoline & diesel rails that use patented low-formaldehyde POM (0.4–0.9 mol % oxy-alkylene; 0.1–2 wt % branched POM) plus 0.8 wt % MgO scavenger. Targets Tier-1 fuel-system makers battling Euro-7 aldehyde caps and permeation limits; delivers ≥2700 MPa modulus and <0.3 mg m^{-2} formaldehyde, extending service life and cutting cabin odor (370 charactors)"

Listing 2: Example of product description.

3.2 Elo Score

Table 1 summarizes the evaluation results for the PBIG task in the materials chemistry category, reporting three score types: the "auto-score" assigned by an AI judge, the "human-score" given by expert reviewers, and "our score," computed according to the six criteria defined in Section 2.3. Notably, our method achieved an auto-score of 1185 for the "Innovation (innov)" criterion (second highest among all PBIG participants), while human scores were lower, indicating a disparity between AI and human judgments. It should be noted, however, that our approach did not incorporate an analysis of human judges' personas. Consequently, applying our framework to analyze real-world customer feedback may offer a promising avenue for generating proposals that resonate more effectively with human evaluators. For a detailed discussion of the LLM-as-a-Judge concept, see References [7–12].

Table 1: Elo Scores in Materials Chemistry category.

criteria	auto-score	human-score	Our-score
tech_valid	896	928	1484
spec	1112	950	1464
need_valid	946	1026	1477
market_size	939	1006	1196
innov	1185	1009	1164
comp_adv	1002	974	1487

Conclusion

We suggested a prompt-based model for generating business ideas from patent documents and achieved the generation of high-quality ideas in the PBIG task [7]. Our approach systematically integrates processes such as extracting patent elements, profiling inventors, conducting market analysis, and constructing business models based on TRIZ principles, leveraging domain knowledge from business strategy experts. Furthermore, AI self-improvement mechanism enhanced idea quality according to predefined competition criteria. While demonstrating strong performance in AI-based evaluations, the study also highlighted challenges in aligning with human judges' preferences. Future work should focus on refining our approach to better accommodate human judgment, for example by incorporating feedback analysis from real-world business settings.

References

[1] A. Subramanian, K.P. Greenman, A. Gervaix, T. Yang, R. G.-Bombarelli, arXiv:2303.08272 (2023).

[2] O. Plätke, R.C. Geibel (2024) The use of artificial intelligence for idea generation in the innovation process (Springer Proceedings in Business and Economics) Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-031-66517-2_14

[3] H. Sugisawa, K. Kinugasa, Patent application in preparation (2025).

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

[5] Elo. The Rating of Chessplayers, Past and Present. Ishi Press, (1986)

[6] Chatbot Arena: Elo Rating Calculation (July 17, 2023)

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

[8] H. Yoshiyasu, "Team NS\NLP at the AgentScen Shared Task: Structured Ideation Using Divergent and Convergent Thinking," The 2nd Workshop on Agent AI For Scenario Planning (AGENTSCEN2025)}, in press (2025).

[9] G. Kanumolu, A. Urlana, V.K. Charaka, B.M. Garlapati, "Agent Ideate: A Framework for Product Idea Generation from Patents Using Agentic AI," The 2nd Workshop on Agent AI For Scenario Planning (AGENTSCEN2025), in press (2025)

[10] Y. Xu, T. Hirasawa, S. Kawano, S. Kato, and T. Kozuno, "MK2 at PBIG Competition: A Prompt Generation Solution," The 2nd Workshop on Agent AI For Scenario Planning (AGENTSCEN2025), in press (2025)

[11] Y. Terao and Y. Tachioka, "Collaborative Invention: Refining Patent-based Product Ideation via LLM-Guided Selection and Rewriting," The 2nd Workshop on Agent AI For Scenario Planning (AGENTSCEN2025), in press(2025).

[12] M. Hoshino, S. Shramatsu, and F. Nagasawa, "A Business Idea Generation Framework Based on Creative Multi-Agent Discussions", The 2nd Workshop on Agent AI For Scenario Planning (AGENTSCEN2025), in press (2025).

Appendix A. Prompts

Prompt 1: Patent Element Analysis

Task
- Propose a novel product and/or business model based on the patent.
- Organize and present the content of the patent clearly.

Structure
(1). Who developed it? Output: Name of the organization or company.
(2). What is the value of this patent? Output: A detailed explanation. Note: Describe the unique features and differentiators from existing inventions as thoroughly as possible.
(3). What are the potential applications? Output: A list in table format. Note: Include both market sectors and use cases in the columns. Propose as many as possible.
(4). Summary of the patent Output: A detailed explanation. Note: Describe it as clearly and accessibly as possible, highlighting the core-essence of the patent.

Prompt 2: Inventor Profiling

Request
Based on the company/organization name, infer and list its potential strengths using the structure below.

Company/organization
[Output(s) suggested by Prompt **1 (1)**]

Structure
1. **Strength**: **Description**
2. **Strength**: **Description**

Prompt 3: Market and Application Analysis

Request
We are exploring potential markets and applications for the given patent. Summarize your research following the output format below.

Patent Details
[Output(s) suggested by by Prompt **1 (2)**]

List of Potential Applications
[Output(s) suggested by Prompt **1 (3)**]

Output Format
* Present the information in a table only (do not include any additional text).
* The table should include the following columns:
** Market
** Market Growth Rate
** Application
** Estimated Profit from Application
** Social Significance
* Sort the entries in descending order of estimated profit.

Prompt 4: Business Model Construction

Request
Based on the strengths of the following patent, investigate which business model would likely generate the highest revenue when entering the specified industry.

Patent Strengths
[Output(s) suggested by Prompt **1 (2)**.]

Target market
[Max-profit output suggested by from Prompt **3**]

Output Format
* Present the results in a table.
* The columns should be:
* Business Model
* Estimated Revenue
* Combined TRIZ Principles Used
* Description of the Business Model
* Each business model should be proposed by applying two or three TRIZ problem-solving principles.

Prompt 5: Business Idea Proposal

Request
Using the information collected so far, propose a product and corresponding business model.

Output Requirements
- product_title: A concise name for your product (up to 100 characters).
- product_description: A brief explanation of the product outlining its essential features and functions, the target users, their needs, and the benefits provided by the product (up to 300 characters).
- implementation: An explanation describing how you will implement the patent's technology into your product (up to 300 characters).
- differentiation: An explanation highlighting what makes your product unique and the reason why it stands out from existing solutions (up to 300 characters).

Company Strengths
[Output(s) suggested by Prompt **3**]

Target Market
[Output(s) suggested by Prompt **2**]

Business Model Concept
[Output(s) suggested by Prompt **4**]

Patent Summary
[Output(s) suggested by Prompt **1 (4)**]

Patent Advantages
[Output(s) suggested by Prompt **1 (2)**]

Constraints
* The proposal must clearly address all of the following elements:
Technical validity: Is the patent suitable for the product? Is the implementation feasible? Can it be done within three years?
* Innovativeness: Does the patented technology offer a novel solution to the demand?
* Specificity: Is the idea specific? For example, "help researchers manage references" is more specific than "help researchers do research."
* Need validity: Do the described users really need this solution?
* Market size: Is the market large enough? Are there many potential users?
Competitive advantage: What business advantage does the product gain by using this patented technology?
* Only output the content (no additional text).
* Count the word total and strictly stay within the word limit.

Prompt 6: Initial Pairwise Evaluation

Task
- Your task is to choose the better idea from the perspective of **Technical validity**.
- Is the patent suitable for the product? Is the implementation feasible? Can it be done within three years?
- The idea must be capable of being made or used in some industry, which can include manufacturing, agriculture, or other practical applications. It should not be a purely theoretical concept.

Output format
Return a JSON object with exactly these keys:
- idea_id: either "A" or "B"
- reason: brief justification.

Remember: output ONLY the JSON object.

Prompt 7: Judges' personas extraction

Background
- You are participating in a competition to generate inventions.
- Below are the reasons for victory or defeat when pitting your inventions against a baseline method.
- To improve the prompt, we would like to infer the judges' characteristics from these win/loss explanations.

Request
- Please extract the judges' personas based on the win/loss explanations.

Evaluation Criteria
- Technical validity, innovativeness, concreteness, alignment with needs, market size, competitive advantage.
- This is a competition for generating inventions from patents.

Judging Reasons
["reason" list suggested by Prompt **6**]

Prompt 8: Factors that Judges' take into account

Background
- You are a generative AI participating in a competition to create inventions.
- The competition results - including score, title, summary, and points of differentiation—have been compiled.

- You wish to impose constraints on the AI prompt in order to improve these results.

Request
To achieve a score above 1200 points, generate a list of additional constraints to include in the prompt, using the judges' personas. These constraints should be universally applicable to any patent.

Evaluation Criteria
Technical validity, innovativeness, specificity, alignment with needs, market size, competitive advantage.
This is a competition to derive inventions from patents.

Desired Output
A bullet-point list of constraints.

Score Results
[Elo score]

Judges' Persona (extracted from reasons for wins and losses)
[Output(s) suggested by Prompt **7**]

Appendix B. Persona and Constraints

Judges' Persona (specialized in chemistry)

1. Technically Rigorous & Patent-Faithful
Key Traits: Engineering-first mindset, conservative with feasibility, and require direct use of patented materials, processes, or compositions.
What Wins: Inventions that specify detailed formulation ranges, processing routes, and quantifiable outputs (e.g., <0.5 mg/m² formaldehyde).
What Fails: Ideas that stray into abstract software, AI, or vaguely connected uses, regardless of market appeal.

2. Feasibility-Driven Realists
Key Traits: Expect commercialization within \~2–3 years using existing equipment and infrastructure.
What Wins: Drop-in solutions with low switching cost, validated manufacturing paths, and minimal retooling.
What Fails: High-risk or speculative technologies requiring new plants, new chemistry platforms, or novel infrastructure.

3. Regulation-Responsive Evaluators
Key Traits: Highly sensitive to global regulatory frameworks—especially emissions (e.g., Euro 7, VDA 278, GB/T 27630).
What Wins: Inventions that enable immediate compliance or preempt near-future regulatory mandates.
What Fails: Concepts not linked to urgent regulatory pain points—even if innovative or sustainable.

4. Market-Oriented Strategists
Key Traits: Prioritize large, urgent, and high-growth markets with tangible demand and broad applicability.
What Wins: Ideas addressing mass markets like automotive interiors or appliances with regulatory pressure and OEM interest.
What Fails: Niche, narrow, or speculative markets with limited customer base.

5. Specificity- and Detail-Seeking Engineers
Key Traits: Demand clear articulation of how the invention works, what it contains, and what it improves.
What Wins: Highly detailed ideas including exact ppm levels, mol % comonomers, phr ranges, and processing methods.
What Fails: Generic, broad-stroke proposals lacking technical, chemical, or process depth.

6. Innovation-Oriented but Execution-Focused
Key Traits: Favor originality only when technically validated and commercially actionable.
What Wins: Combinations that extend the patent's scope in a novel, IP-defensible, and manufacturable way (e.g., resin + acoustic damping).
What Fails: "Buzzword" novelty (e.g., AI, platforms) that lacks tangible linkage to the patent or physical deliverables.

7. Competitive Advantage and IP Defensibility Seekers
Key Traits: View IP as a moat; prefer solutions difficult to replicate or bypass.
What Wins: Inventions with unique formulations, protected performance features, or platform-level lock-in.
What Fails: Me-too products or ideas that offer little legal or performance protection from competition.

Constraints (specialized in chemistry)

Technical Rigor and Patent Fidelity
- Specify exact compositional ranges (e.g., 0.1–2 wt% additive, 0.4–0.9 mol% comonomer).
- Define quantifiable performance targets (e.g., Young's modulus ≥ 2700 MPa, formaldehyde < 0.5 mg/m²).
- Directly integrate patented methods, materials, or formulations into the invention concept.

Feasibility and Manufacturing Readiness
- Require drop-in compatibility with existing infrastructure (no equipment swaps or new machinery).
- Limit implementation time to ≤ 3 years by leveraging current supply chains and industry-standard processes.

Regulatory Relevance
- Align the invention with imminent or forthcoming global regulations (Euro 7, GB/T 27630, VDA 278, etc.).
- Include language demonstrating proactive compliance, certification readiness, or the benefits of regulatory exemptions.

Market Scope and Applicability
- Target large, regulated markets (automotive, appliances, medical, etc.).
- Quantify the total addressable market (TAM) or cite specific OEM/Tier-1 applications.

Detail, Specificity, and Process Clarity
- Provide precise chemical structures, blend ratios, processing temperatures, catalysts, and quench agents.
- Define processing parameters (e.g., melt index, temperature range, MI = 0.5–1.5).

Innovative Yet Practical

- Restrict novelty to physically realizable combinations (e.g., a patented resin plus an acoustic layer), avoiding vague software abstractions.
- Ensure any AI or digital component directly controls or monitors a physical/material process.

IP Strength and Competitive Advantage
- Emphasize IP-backed differentiation (e.g., patented compositions, protected process steps).
- Include commercialization strategies such as OEM licensing models, territorial exclusivity, or supply contracts tied to compliance.