FRONTIER TOPICS MINING METHOD VIA AI-AGENT

Bin Ge, Chunhui He, Chong Zhang, Jibing Wu

Laboratory for Big Data and Decision, National University of Defense Technology, Changsha 410073, China

ABSTRACT

Identifying high-quality frontier topics from massive scientific research data to assist researchers in accurately conducting scientific research is of paramount importance. Traditional analysis methods face bottlenecks such as limited cross-domain adaptability, high resource consumption, and low efficiency. To address these challenges, this study proposes an AI-agent-based frontier topic mining method. An innovative generative-verification dual-agents (D-Agents) architecture is constructed. Specifically, prompt engineering is employed to develop a generative agent (G-Agent), which leverages the semantic understanding capabilities of large-scale pre-trained language models to automatically generate candidate frontier topics. Subsequently, a verification agent (V-Agent) is introduced to establish a multi-dimensional evaluation system, which automatically verifies candidate topics from dimensions including academic novelty, topic accuracy, and completeness of frontier topics. The effectiveness of the proposed method is validated through three manually labeled datasets in computer vision (CV), natural language processing (NLP), and machine learning (ML). Experimental results demonstrate that the D-Agents framework can simultaneously perform frontier topic mining tasks across multiple domains. On the three labeled datasets (CV-DataSet, NLP-DataSet, and ML-DataSet), the D-Agents achieve a precision exceeding 74% while maintaining a recall over 85%.Compared with the traditional bibliometric method, this method significantly improves the precision and recall of frontier topic mining in recommendation system, and the performance reaches 86%. The D-Agents framework effectively mitigates the hallucination issue of G-Agent through its automatic generation and self-verification mechanism, thereby substantially enhancing the efficiency of frontier topic mining.

KEYWORDS

LLMs, Frontier Topics, Prompt Engineering, D-Agents, G-Agent, V-Agent, RAG

1. INTRODUCTION

The exponential proliferation of scientific and technological information resources urgently requires the development of efficient methodologies to accurately extract and identify frontier topics from vast scholarly datasets. Systematically elucidating the developmental trajectory of disciplines, discerning pivotal research topics, and projecting future trends through comprehensive analysis of these resources constitutes a critical academic task. This analytical framework provides crucial theoretical guidance and practical references for optimizing research prioritization, and enhancing innovation efficiency in scientific exploration [1]. A frontier topic typically refers to the most innovative and frontier research focus within a specific discipline or field. In contrast, research trends are more concerned with describing the general patterns and temporal shifts in research activities across a particular domain. The task of frontier topic mining essentially involves conducting an in-depth analysis of the dynamic evolution of the human knowledge system using

David C. Wyld et al. (Eds): CSITEC, NLPI, CMCA, NECOM, DMML, SP, ADCO – 2026
pp. 45-64, 2026. CS & IT - CSCP 2026 DOI: 10.5121/csit.2026.160805

advanced technologies. This process not only provides robust support for scientific research decision-making but also plays a pivotal role in expanding human cognitive boundaries and accelerating exploration in uncharted areas. Figure 1 illustrates the general workflow of traditional frontier topic mining methods.

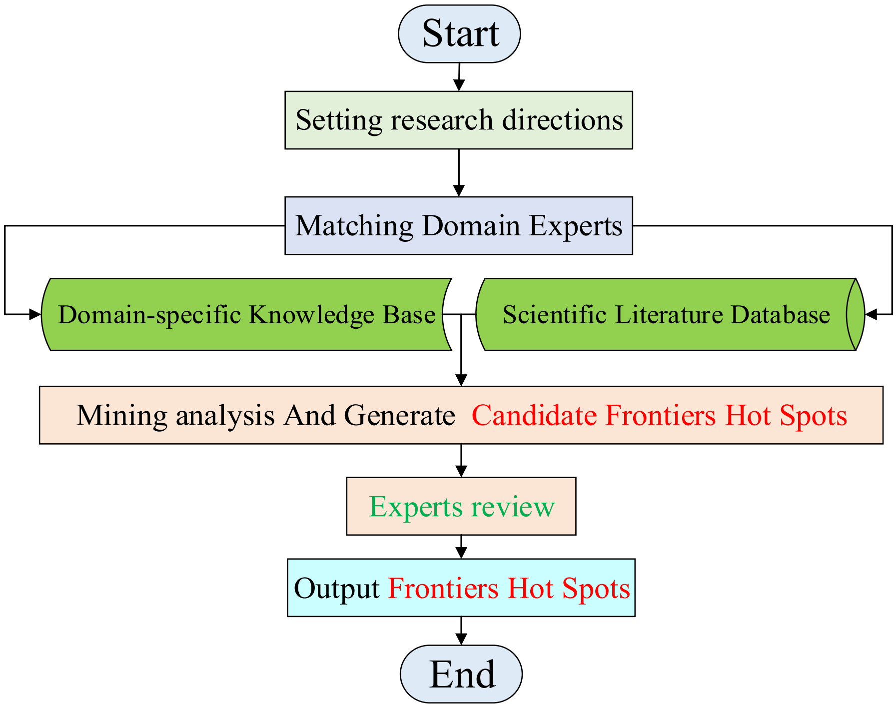

Figure 1. The general workflow of traditional frontier topics analysis and mining method.

It can be seen from Figure. 1 that the traditional frontier topic mining methods can generally only carry out frontier topic mining analysis for a certain discipline direction. Firstly, the corresponding domain experts are matched according to the set subject direction, and then the domain knowledge base and scientific literature database are collected. Then, different mining methods are combined to generate candidate frontier topics, and finally, the generated candidate frontier topics are manually reviewed by domain experts, and then the final frontier topics are output. Traditional methods are roughly divided into two categories because of different mining methods adopted. The first category is to combine bibliometric features, including word frequency analysis [2], literature citation [3], word co-occurrence analysis [4] and other statistical techniques to realize the mining of frontier topics. The second broad category is to identify frontier topics using topic models such as KPLDA [5] and BertTopic [6] in text mining. It is worth noting that although the frontier topics mined by traditional methods are highly readable and authoritative. However, this kind of method needs the support of domain experts, and relies on a large number of static scientific literature databases and high-quality domain knowledge bases. Therefore, traditional methods generally have the following shortcomings. ① Usually, it can only support a single discipline category, and cannot be compatible with frontier topic mining tasks of multiple disciplines at the same time. ② It takes a lot of manpower resources to collect and analyze literature data to achieve the goal of frontier topic mining, which is a typical low-efficiency and long-period method. ③ It is difficult to capture implicit semantic relevance information, which will affect its ability to deeply mine frontier topics. ④ Insufficient modeling of time series dynamics of unstructured data will lead to the problem of lag in frontier topic identification. ⑤ Because the frontier topics of emerging disciplines are usually sudden and uncertain, it is difficult for traditional methods to capture the early frontier topics of emerging disciplines in time.

The deep integration of artificial intelligence into scientific research has created new opportunities in frontier topic mining. The rapid advancement of large language models and AI agents enables dynamic and in-depth knowledge discovery. After pre-training on vast general-domain datasets and distilling massive amounts of data, these models can effectively overcome the limitations of traditional approaches in frontier topic detection by incorporating feedback-based reinforcement learning, incremental learning, and dynamic data updating strategies. However, it is worth noting that this type of method also has an obvious shortcoming. Often due to the common "hallucination" problem of large language models, it is difficult to verify the authenticity of the excavated frontier topics. Considering that the large language models or AI agents need to provide multi-dimensional evidence chain support in frontier topic identification, existing methods do not have an automatic verification mechanism. Aiming at the above problems, inspired by references [7-8], Joint prompt engineering [9] proposed a frontier topics mining method via AI-agent (D-Agents). This method combines G-Agent and V-Agent at the same time to realize the automatic generation and verification of frontier topics. The G-Agent mainly responsible for mining potential frontier topics, while V-Agent is responsible for automatically verifying the mined frontier topics to ensure the accuracy and integrity of frontier topics. The core contributions of this paper are as follows:

- Generation-verification dual-agent architecture is innovatively designed to be compatible with frontier topic mining tasks across multiple domains.
- A hallucination suppression mechanism is explored, wherein an agent-based automatic verification process is introduced to effectively mitigate the issue of "hallucination" in the automatic generation of frontier topics, thereby enhancing the reliability of the identified topics.
- D-Agents method exhibits the capability to identify early emerging frontier topics, significantly improving the efficiency of frontier topic mining by leveraging large language models and integrating RAG technology.

2. RELATED WORKS

A frontier topic refers to the latest knowledge system or research direction with development potential within a specific discipline or domain, typically involving high-tech or interdisciplinary integration. These topics usually exhibit three key characteristics: modernity, interdisciplinarity, and practical applicability. Analyzing and mining frontier topics is of great significance for grasping the development trend of technology and the evolution of discipline context [10]. From the perspective of discipline division, it should be included in the research category of science. Due to differences in technical paths, the frontier topic analysis and mining methods are different. Bibliometrics and citation analysis, co-occurrence network and keyword analysis, interdisciplinary analysis, dynamic tracking, visualization and AI-driven analysis, and multi-source data collaborative verification belong to commonly methods.

Bibliometrics and citation analysis methods mainly explore frontier topics by combining literature features and citation network analysis. Gao Yunfeng [11] used bibliometric method combined with VOSviewer and CiteSpace visual analysis software to quantitatively analyze the literature related to mining ecological environment restoration published in CNKI from 2000 to 2018, and successfully excavated the top 5 frontier topics in this field as "Land reclamation", "Ecological compensation and deposit", "Phytoremediation of heavy metal pollution", "Geological disaster management", and "Green mine and mine park construction". Zhang Yali [12] conducted data mining and visualization of 3830 documents related to agricultural multispectral research published between 2002 and 2021 through citation analysis tools, and successfully got frontier topics and research institutions with the most published papers in this field as well as the most influential journals. Han Qi [13] visualized and analyzed 3018 anticancer network pharmacology research literatures published between January 2008 and May 2023, and successfully excavated frontier topics such as "Tumor microenvironment", "Anticancer drugs", "Traditional Chinese medicine", "Calycosin", "Molecular mechanism", "Molecular docking" and "Anticancer agents", which have been widely recognized by the industry.

Co-occurrence network and keyword analysis methods combined with keyword co-occurrence map and high-frequency word dynamic monitoring technology are widely used in frontier topics mining tasks. In order to comprehensively understand the research status in the field of big data, Zhang Niuniu [14] adopted technologies such as co-word analysis and social network analysis combined with SPSS and Gephi software to analyze and mine related articles published in Web of Science (WOS) databases between 2017 and 2019. The top 8 frontier topics in the field of big data have been successfully excavated as “Text mining”, “Data fusion”, “Distributed computing”, “Industry 4.0”, “Data privacy”, “Artificial intelligence”, “Energy big data”, “Bioinformatics”. Zhang B [15] took the literature about science and technology industry collected in CNKI database as the data source, and analyzed the research topics and development directions by using co-word analysis and social network analysis. Finally, the relevant frontier directions are summarized into four aspects: university science and technology industry, continuous innovation of science and technology industry, policy-oriented science and technology industry and evaluation system of science and technology industry.

Due to their inherent advantages in multi-disciplinary integration, interdisciplinary analysis methods have the potential to identify emerging frontier topics within disciplines by integrating multi-term combination retrieval with cross-domain data fusion techniques. Nie Jia [16] focuses on altitude sickness, aiming at knowledge discovery, and jointly applies bibliometrics, knowledge graphs, data mining, network pharmacology and other methods to construct a multi-dimensional literature research model for altitude sickness. A comprehensive analysis was carried out on the research literature on altitude sickness published between 2006 and 2016. From the 31 research directions involved in altitude sickness, the top 6 research frontier topics in the world are "Altitude polycythemia", "Acute altitude sickness", "Chronic altitude sickness", "Apoptosis", "Altitude hypoxia" and "Pulmonary hypertension". Mi Lingyun [17] used CiteSpace as a bibliometric analysis tool, combined with cross-domain data fusion technology, systematically evaluated the development status of pro-environmental behavior (PEB) from macro, meso and micro levels, and summarized the frontier topic, trends and frontiers based on 4032 related articles in the WOS database, which provided effective theoretical guidance for effectively promoting the sustainable development of society.

Dynamic tracking methods primarily leverage various academic literature publishing platforms to construct dynamic tracking models based on time series, enabling the automatic prediction of emerging frontier topics in the future. Li Bo [18] collected the literature data of more than 10 core journals including Chinese Management Science, Management World and Journal of Management Science by selecting important journals with high authority in the field of management science. Based on bibliometrics, this paper comprehensively analyze the research features and dynamic development trend of management science, and establishes a dynamic tracking model combined with time series analysis, which reveals the frontier topics and dynamic evolution process of this field.

Visualization and AI-driven analysis methods are mainly based on knowledge graphs (KG) and semantic mining technology to automatically identify emerging concepts and visually display frontier topics and evolution paths. Xu Jia [19] collected 2882 literatures related to physiology teaching reform, and used bibliometrics and graph visualization methods to excavate the top 5 frontier topics in the field of physiology teaching reform as "Rehabilitation medicine", "Neuroscience", "Infectious diseases", "Pathogenesis of nervous system diseases" and "Treatment strategies of nervous system diseases", and revealed the close relationship between these frontier topics and brain research programs. Shao Bilin [20] took 5391 documents related to the topic of recommendation system from 2009 to 2018 and included in the WOS core database as the research object, and analyzed them by combining bibliometric methods and semantic mining technologies. With the help of VOSviewer software, the knowledge graph is constructed, and valuable knowledge is intuitively discovered through visual approaches. Finally, according to the keyword concurrency graph, it is summarized that the top 7 frontier topics in the field of recommendation system are “Collaborative filtering and matrix factorization”, “Information technology and recommendation system”, “Recommendation algorithm and performance evaluation”, “User feature representation technology”, “Cold start and data sparsity”, “Personalized recommendation” and “privacy protection”. This discovery points out the scientific research direction for the development of recommendation system field.

Multi-source data collaborative verification methods can effectively verify the authenticity of frontier topics and reduce the bias of single data source through cross-comparison and topic correlation analysis. Cheng Jie [21] performed a visualization and cross-alignment analysis based on 1051 articles related to oyster reef ecosystems from 1981 to 2022. The top 3 frontier topics in this field between 2014 and 2022 were excavated as "Habitat protection and restoration", "Ecosystem services" and "Climate change", providing effective technical support for scholars and regulators concerned about oyster reef protection. Zhang Peng [22] screened 1305 articles on SA neuroimaging from January 1998 to December 2023 in Web of Science and Scopus, and successfully excavated the top 5 frontier topics in this field using multi-source data collaborative verification and topic correlation analysis technology are "Cognitive behavioral therapy", "Machine learning", "Transcranial direct current stimulation", "Depression", and "Brain imaging technology". The above results provide effective direction guidance for the research of SA brain mechanism.

The existing literature has proposed some frontier topic mining methods in specific fields. However, it should be noted that related methods basically rely on domain experts to retrieve and analyze a large amount of literature in order to complete a small number of cutting-edge topic mining tasks in specific fields. Existing methods generally have the problems of long cycle, low efficiency, and poor domain migration. In order to improve the above shortcomings, an efficient automatic mining framework for frontier topics is designed by combining the most advanced large language model and prompt engineering technology and retrieval-augmented generation (RAG) [23-24] technology. Considering that the phenomenon of generating "hallucinations" is common in single agents. In order to effectively deal with the above problems, by combining dual-agent architecture, this paper proposes a frontier topics mining method via AI-agent. Our method can realize automatic generation and verification of frontier topics in multiple fields based on online search and local knowledge base, combining generative agents and verification agents at the same time, and can greatly improve the mining efficiency and generation quality of frontier topics.

3. MATERIALS & METHODS

3.1. Frontier Topic Mining Method via Generative Agent (G-Agent)

AI agents are intelligent systems that can autonomously perceive the environment, perform specific actions, and make decisions to achieve preset goals. Its essence is to combine artificial intelligence technology (such as large language models, reinforcement learning) and tool sets to form a set of programs that can run independently. An independent AI agent usually includes a brain, a memory module, planning module, an environment perception module, an execution module and a tool calling module. The above modules need to be deeply customized and designed and tuned in combination with prompt engineering and different business models in order to generate a complete AI agent. The trained AI agent generated have a clear goal, a set of executable action instructions and some optional tool sets. Figure 2 shows the architecture of G-Agent.

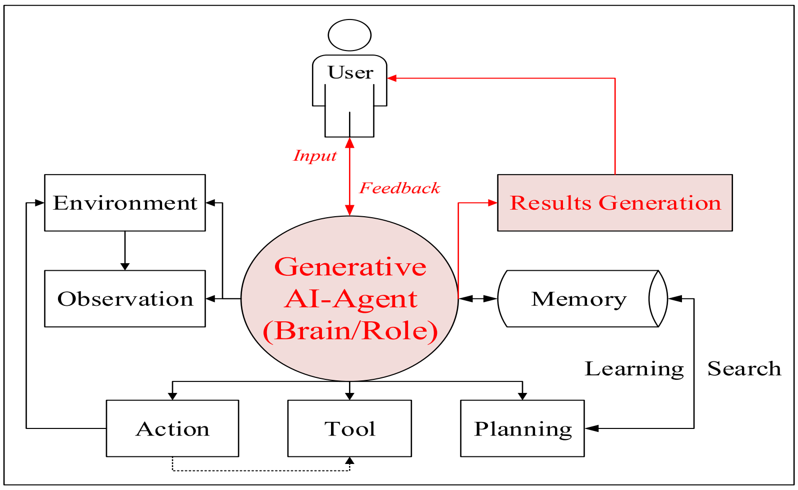

Figure 2. Architecture design of G-Agent.

According to Figure.2, the G-Agent in addition to brain, memory module, planning module, environment perception module, execution module and tool calling module. It also includes a user input and feedback module and a result generation module. Learning and search instructions are added between the memory module and the planning module for optimizing performance. After in-depth analysis of the mining requirements of frontier topics, a high-quality G-Agent was deeply customized. Its goals, roles, brains, prompts and optional tool set are set up in Table 1.

Table 1. Module name and instruction settings of G-Agent.

Agent module name	**Module specific functions or corresponding instructions**
User module	Guide users to enter frontier topic mining prompt instructions or give feedback instructions in specified fields.
Brain of the G-Agent	ERNIE 4.0
Role of the G-Agent	Senior experts in the field of intelligence.
Perception of the G-Agent	Automatically capture domain frontier topic mining goals and requirements from user-input prompt instructions.
Memory of the G-Agent	Long-term memory
Action of the G-Agent	① Default question ② Automatic questioning ③ Tool call ④ Feedback learning ⑤ Clear historical conversation
Planning of the G-Agent	① Automatically complete the search and learning of knowledge according to memory and user feedback. ② Collect and interpret the corresponding authoritative literature according to the captured frontier topics mining goals and requirements. ③ Generate no more than K candidate frontier topics after in-depth interpretation of all data in chronological order. ④ The literature collected comes from published academic papers and patent documents, research reports, and conference call for papers.

	⑤ Ensure that frontier topics are highly relevant and accurate to the field input by the user. ⑥ Ensure that the number of words in a single frontier topic is strictly prohibited from exceeding 10 words, and it is not repeated. ⑦ Only use the list form to return topics one by one, without any explanatory text. ⑧ The output format refers to the style of the knowledge base, and the output results are automatically aligned with the frontier topics in the knowledge base as much as possible.
Optional tool set for G-Agent	① Online search ② Micro single domain knowledge base
Result generation of the G-Agent	Candidate frontier topics are returned to users one by one in strict accordance with the list format.

According to the modules and instructions set in Table 1, the agent build platform [26] provided by Baidu is used to create a semi-supervised basic frontier topics automatically generated AI agent called G-Agent. The online platform supports efficient agent creation, editing, analysis, tuning, publishing, deleting, and online experience functions. After the agent is constructed and published based on the above platform, the user can automatically generate the corresponding candidate frontier topics list by constructing the mining instruction in the field according to the preset questioning mode and input it into the agent, and the structure of the list is $\mathrm{FHS} = [\mathrm{hs}_1, \mathrm{hs}_2, \ldots, \mathrm{hs}_n](1 \leq \mathrm{n} \leq \mathrm{K}, \mathrm{n} \in \mathrm{N})$, so as to obtain the candidate frontier topics in a specific field. K is an adjustable integer parameter that can be set by the user. The recommended value range is between $[3,30]$, and the default value is $\mathrm{K} = 20$. The generative agent brain in this article uses the completely free Baidu ERNIE 4.0 [27] large language model by default. Generative agent can achieve RAG goals through online search tools and local knowledge bases, and ensure the generation quality of frontier topics. The G-Agent contains retriever, generator, decoder, topic-generator four components. All core modules are calculated as follows:

$$p_\eta(z|x) \propto \exp\left(d(z)^\intercal q(x)\right), \quad d(z) = \mathrm{ERNIE}_d(z), \quad q(x) = \mathrm{ERNIE}_q(x) \tag{1}$$

$$p_\theta(y_i|x, z, kb, y_{1:i-1}) \tag{2}$$

$$p'_\theta(y_i|x, kb, y_{1:i-1}) = \sum_{z \in top-k(p(\cdot|x))} p_\eta(z_i|x) p_\theta(y_i|x, z_i, kb, y_{1:i-1}) \tag{3}$$

$$p_{topic-generator}(y|x) \approx \prod_i^M \sum_{z \in top-k(p(\cdot|x))} p_\eta(z|x) p_\theta(y_i|x, z, kb, y_{1:i-1}) \tag{4}$$

where, in the retrieval component $p_\eta(z|x)$, $d(z)$ is a dense representation of a document produced by a ERNIE document encoder, and $q(x)$ a query representation produced by a query encoder, also based on ERNIE, Calculating $top - k(p(\cdot\,|x))$, the list of k documents z with highest prior probability $p_\eta(z|x)$, in the generator component $p_\theta(y_i|x, z, \mathrm{kb}, y_{1:i-1})$, we use ERNIE pre-trained model and combination a local knowledge base kb, in the decoder component $p'_\theta(y_i|x, \mathrm{kb}, y_{1:i-1})$, we plug it as a standard decoder via combination a local knowledge base kb, in the topic-generator component $p_{topic-generator}(y|x)$, the top M documents are retrieved using the retriever, and then the generator produces a distribution via a local knowledge base $kb$for the next output topic for each document, before marginalizing, and repeating the process until output K-th topic.

3.2. Frontier Topics Mining Method via Dual Agents (D-Agents)

The frontier topic mining method of D-Agents belongs to a semi-supervised serial method with strict logical sequence. The entire method covers two different components: a generative AI agent (G-Agent) and a verification AI agent (V-Agent). The candidate frontier topic list $[\mathrm{hs}_1, \mathrm{hs}_2, \ldots, \mathrm{hs}_n](1 \leq \mathrm{n} \leq \mathrm{K}, \mathrm{n} \in \mathrm{N})$ output by the G-Agent is to be used as the input of

V-Agent. The module name and instruction settings of the G-Agent are shown in Table 1. The core function of the V-Agent is to verify the novel, complete, accurate and authoritative of all elements in the input candidate frontier topic list. Then, all the frontier topics that pass the verification are returned to the user according to the specified format, so as to complete the verification of the whole frontier topics. The function of automatic verification is similar to the expert review step in traditional frontier topic mining method. It is worth noting that although the basic architecture of each module of the V-Agent is consistent with G-Agent. However, there are essential differences in their functions and goals, so there are great differences between the instructions and brain of each module in the V-Agent and the G-Agent. Figure 3 shows the overall architecture of D-Agents.

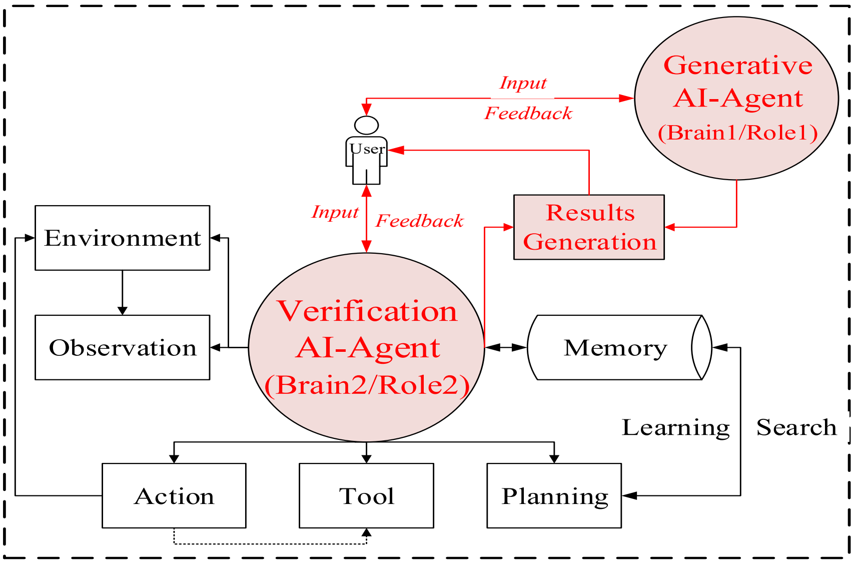

Figure 3. Overall architecture of D-Agents.

According to Figure 3, the D-Agents method is a semi-supervised serial method with strict logical sequence. It can be seen that the human-agent collaboration framework is designed based on the idea of "human-in-the-Loop" (HITL). First, the user needs to call G-Agent to output a list of candidate frontier topics, and then input this list into V-Agent to complete automatic verification and return it to the user. After review, the feedback is obtained and frontier topics are output. To achieve the above functions and goals, it is necessary to separately set corresponding instructions and brain for each module of the V-Agent. The detailed settings of the goal, role, brain, prompt instruction and optional tool set of the V-Agent are shown in Table 2.

Table 2. Module names and instruction settings of V-Agent.

Agent module name	**Module specific functions or corresponding instructions**
User module	Guide the user to input verification prompt instructions or give feedback instructions according to the candidate frontier topic list.
Brain of the V-Agent	ERNIE 3.5
Role of the V-Agent	Senior experts in the field of intelligence.
Perception of the V-Agent	Automatically capture a list of candidate frontier topics and domains to which frontier topics belong from an input prompt instruction.
Memory of the V-Agent	Long-term memory
Action of the V-Agent	① Default question ② Automatic questioning ③ Tool call ④ Feedback learning ⑤ Clear historical conversation
Planning of the V-Agent	① Automatically complete the search and learning of knowledge

	according to memory and user feedback. ② Call knowledge to complete reasoning and verification according to the capture list, domain mining goals and requirements. ③ Strictly verify each frontier topics to make it novel, complete, accurate and authoritative. ④ Ensure that each verified frontier topic is independent and complete, with concise word count and strictly abiding by academic norms. ⑤ Ensure that the frontier topics that have passed the verification cannot be repeated, and automatically filter the frontier topics that have not passed the verification. ⑥ Allows to expand and supplement undiscovered but closely related and authentic frontier topics in the above list. ⑦ The output format strictly refers to the style of the local knowledge base, and the frontier topics covered by the knowledge base should be strictly consistent with it. ⑧ Ensure that the verified frontier topics are directly output one by one without explanatory text.
Optional tool set for V-Agent	① Online search ② Large cross-domain knowledge base
Result generation of the V-Agent	The verified frontier topics are uniformly returned to the user in strict accordance with the specified format.

According to the modules and instructions set in Table 2, the agent construction platform [26] provided by Baidu is used to create the V-Agent. The user inputs the candidate frontier topic list FHS according to the preset questioning mode and constructs frontier topic verification instructions and requirements and inputs them into the agent to automatically verify the novel, complete, accurate and authoritative of the candidate frontier topic and output the verified frontier topics according to the specified format, thereby completing the whole frontier topics mining task. Considering that there are differences in different large language model training corpus and underlying frameworks. In order to support the local knowledge base, ERNIE 3.5 [28] large language model was selected as the brain to build V-Agent. This design can effectively alleviate the "hallucination" problem of generating a single large language model. The V-Agent can achieve the goal of RAG through online search tools and local knowledge base, and ensure the generation quality of frontier topics. The V-Agent contains checker and aligner two core components. The formulas are calculated as follows:

$$Candidate_{Topic}(y) = NCAA(y) \wedge L(y) \wedge I(y) \wedge R(y),\ \forall y \in p_{topic-generator}(y|x) \quad (5)$$

$$Topic(y_j) = \min_{K} \sum_{i=1}^{K} (1 - d(x_i, y_j)),\ \forall y_j \in Candidate_{Topic}(y) \quad (6)$$

$$d(x_i, y_j) = 1 - \frac{x_i \cdot y_j}{||x_i|| \times ||y_j||} = 1 - \frac{\sum_{i=1}^{n} x_i y_j}{\sqrt{\sum_{i=1}^{n} x_i^2} \times \sqrt{\sum_{j=1}^{n} y_j^2}} \quad (7)$$

where, in checker component $\text{Candidate}_{Topic}(y)$, we use $NCAA(y), L(y), I(y), R(y)$ to check the novel & complete & accurate & authoritative ($NCAA$), length, independence and repeatability of each topic. In aligner component $Topic(y_j)$, at first, x_i and y_j are transformed into vectors of the same length using pre-trained model API, and then use cosine similarity to calculate the cosine distance between the current candidate topic y_j and the topics x_i in the local knowledge baseone by one, thus completing the automatic alignment of each topic. When the minimum cosine distance $d(x_i, y_j)$ is less than the threshold $\alpha, (\alpha = 0.2)$, the current candidate topic is replaced with the corresponding topic in the knowledge base, otherwise, the current candidate topic is unchanged.

4. DATA SETS AND EVALUATION METRICS

4.1. Dataset

For the performance evaluation of semi-supervised methods, it is common practice to employ high-quality, small-scale standard datasets. Given the current lack of standardized, high-quality evaluation datasets in the field of frontier topic mining, an objective assessment of the comprehensive performance of the proposed method is conducted. This evaluation is based on three core domains—computer vision, natural language processing, and machine learning—by collecting the most recent call-for-papers directions from leading international conferences in these fields. All data sets used in the experiment are constructed in the data annotation stage by using the iterative correction implementation scheme of "annotation → review → correction". Finally, three high-quality small-scale frontier topics evaluation data sets are labeled after manual statistics and processing. An overview of the relevant evaluation datasets will show in Table 3, Table 4 and Table 5.

Table 3. Top 20 Frontier topics labeled data sets in the field of computer vision (CV).

Dataset name	Data Source	Top 20 frontier topics
CV-DataSet	CVPR ICCV ECCV	Image 3D, Autonomous driving, Adversarial attack and defense, Biometrics and Biometric Vision, Posture recognition, Computational imaging, Datasets and evaluation, Deep learning architecture innovation, Explainable computer vision, Image and video synthesis, Low level vision, Representation learning, Scene analysis and understanding, Optimization methods, Vision applications and systems, Vision + other modalities, Transfer learning, Vision ethics, Embodied vision, Segmentation grouping and shape analysis

Table 4. Top 20 Frontier topics labeled datasets in the field of natural language processing (NLP).

Dataset name	Data Source	Top 20 frontier topics
NLP-DataSet	ACL EMNLP NAACL CoNLL	Computational social sciences and sociolinguistics, Dialogue and interaction systems, Efficient/low resource methods for NLP, NLP model interpretability and analysis, Multilingualism and linguistic diversity, Question answering systems, Information retrieval and text mining, Machine translation, Machine learning for NLP, Multimodality and linguistic foundations, Ethical bias and fairness, Information extraction, Pre-trained Models, Resources and evaluation, Semantics and syntax, Sentiment analysis, Text summary, Language understanding and generation, Linguistic theory, Speech processing

Table 5. Top 20 Frontier topics labeled data sets in the field of machine learning (ML).

Dataset name	Data source	Top 20 frontier topics
ML-DataSet	ICML NeurIPS COLT	General machine learning, Game theory and statistical learning theory, Deep learning, Neuroscience and cognitive science, Large language models, Reinforcement learning, Online learning, Time series analysis, Optimization methods, Probabilistic methods, Algorithm design and analysis, Active and interactive learning, Data geometry and topology learning, Kernel methods, Complex data learning, Trusted machine learning, Machine learning applications, High dimensional & nonparametric statistics, Bayesian methods, Supervised learning

Table 3, Table 4 and Table 5 list the top 20 frontier topics at the current stage, which have high authority in the three fields of computer vision, natural language processing and machine learning. In the experimental will use these three data sets to complete the performance evaluation of the proposed method.

4.2. Evaluation Metrics

To objectively evaluate the comprehensive performance of our method. Precision and Recall are selected as performance evaluation criteria. The specific calculation formulas are as follows:

$$Precision = \frac{Len(A \cap B)}{Len\ (B)} \times 100 \quad (8)$$

$$Recall = \frac{Len(A \cap B)}{Len\ (A)} \times 100 \quad (9)$$

among them, A is the frontier topics set contained in the labeled data set, B is the frontier topics set mining by the proposed method, $Len(A \cap B)$ represents the number of common elements between the two sets (frontier topics implement semantic alignment), $Len(B)$ represents the total number of frontier topics set elements mining by the proposed method, and $Len(A)$ represents the total number of frontier topics set elements contained in the labeled data set.

5. EXPERIMENTAL AND RESULTS

5.1. Baseline Methods

To fully and objectively demonstrate the comprehensive performance of the proposed method, two groups of comparative experiments are conducted using different datasets to evaluate its effectiveness. In order to compare the differences between the G-Agent and the D-Agents, the first group of experiments carried out comparative experiments on three datasets: CV-DataSet, NLP-DataSet and ML-DataSet, and used Precision and Recall to evaluate the performance of the methods. In order to evaluate the performance difference between the frontier topics mining method via D-Agents and the traditional method, the second group of experiments is compared with the visualization and AI-driven analysis method proposed in reference [20], and the performance of related methods is also evaluated by Precision and Recall. It should be noted that the set of labeled frontier topics in the second group of experiments is directly extracted from the experimental results of the comparative reference, and is not labeled manually.Furthermore, in the third group of experiments is comparedwith the KPLDA [5] model and BERTopic [25] model.

5.2. Results and Analysis

The experimental results of the comparison between the G-Agent and D-Agents carried out on CV-DataSet, NLP-DataSet and ML-DataSet data sets are shown in Table 6. Relevant indicators are calculated in combination with the top 20 frontier topics marked in Tables 3, Table 4 and Table 5. The complete frontier topics generated by the G-Agent and D-Agents are shown in Figure A1, Figure A2 and Figure A3 in Appendix A.

Table 6. Comparative experimental results of our method on three labeled datasets

.

Dataset name	Method	Precision	Recall
CV-DataSet	G-Agent	$\frac{17}{20} \times 100\% = 85\%$	$\frac{17}{20} \times 100\% = 85\%$
	D-Agents	$\frac{19}{21} \times 100\% = \mathbf{90}\%$	$\frac{19}{20} \times 100\% = \mathbf{95}\%$
NLP-DataSet	G-Agent	$\frac{13}{20} \times 100\% = 65\%$	$\frac{13}{20} \times 100\% = 65\%$
	D-Agents	$\frac{17}{23} \times 100\% = \mathbf{74}\%$	$\frac{17}{20} \times 100\% = \mathbf{85}\%$
ML-DataSet	G-Agent	$\frac{14}{20} \times 100\% = 70\%$	$\frac{14}{20} \times 100\% = 70\%$
	D-Agents	$\frac{18}{22} \times 100\% = \mathbf{82}\%$	$\frac{18}{20} \times 100\% = \mathbf{90}\%$

It is worth noting that, in order to make the experimental results more directly compared, we extracted a small number of manually labeled frontier topics in the agent to construct a small-scale local knowledge base. The above knowledge base is mainly used to complete the generation style guidance and alignment guidance of frontier topics. According to the experimental results shown in Table 6 and Figure. A1, Figure. A2 and Figure. A3 in the Appendix A, it can be seen that the comprehensive performance of the frontier topic mining method via D-Agents on three data sets is better than G-Agent. The Precision of the D-Agents on the three datasets of CV-DataSet, NLP-DataSet and ML-DataSet reach 90%, 74% and 82%, respectively, which are 5%, 9% and 12% higher than the G-Agent, respectively. The Recall of the D-Agents on the three datasets reaches 95%, 85% and 90% respectively, which is 5%, 20% and 20% higher than G-Agent, respectively. The reason for the above results may be related to some high-quality planning strategies set in V-Agent. It can further verify the accuracy, completeness, novelty and authority of the results by combining different brain on the basis of G-Agent. Thus, filtering out some frontier topics that cannot pass verification. In addition, it can automatically expand some high-quality additional frontier topics that have not been discovered by G-Agent but are closely related to the field. The above strategies can effectively improve the comprehensive performance of D-Agents. From three different datasets, the D-Agents has the best comprehensive performance on the CV-DataSet dataset. The performance on the ML-DataSet dataset is second. The overall performance is the weakest on the NLP-DataSet dataset. According to the above results, it can be seen that there are certain fluctuations in the comprehensive performance of the proposed method in different fields, and the core reason may be related to the underlying brain performance. Although in the construction of agents, the Baidu ERNIE large language model in the general field is selected as the agent brain. However, there are certain differences in the training strategies and training corpus of large language models corresponding to different brains. This will directly lead to some deviations in the domain adaptability of the method. In order to further reveal the performance difference between the D-Agents and the traditional method. In order to further reveal the performance difference between the D-Agents method, the traditional expert mining method and the open source DeepSeek-R1-671B large language model [29], in addition to carrying out comparative experiments on the above three labeled datasets, the traditional mining methods proposed in references [20] has carried out in-depth comparative experiments for three specific fields. The relevant experimental results are shown in Table 7.

Table 7. Comparative experimental results of frontier topic mining with different methods in three specific fields (red indicates inconsistency with expert mining results).

Domain Frontier Topic Mining	**Expert Mining Results**	**DeepSeek-R1-671B Mining Results**	**D-Agents Mining Results**	**D-Agents Precision**	**D-Agents Recall**
Mining Top7 frontier topics in the field of "Recommendation System" from 2009 to 2018[20]	Collaborative filtering and matrix factorization, Information technology and recommendation system, Recommendation algorithm and performance evaluation, User feature representation technology, Cold start and data sparsity, Personalized recommendation, privacy protection	Explainable recommendation, Fairness & diversity,Reinforcement learning for recommendation system,Multi-modal recommendation, Cold start & data sparsity, Knowledge graph for recommendation system, Privacy & federated learning	Collaborative filtering and matrix factorization, Information technology and recommendation system, User feature representation technology, Cold start and data sparsity, Personalized recommendation, privacy protection,AI recommendation system	$\frac{6}{7} \times 100\% = 86\%$	$\frac{6}{7} \times 100\% = 86\%$

According to the experimental results in Table 7, the comprehensive performance of D-Agents is excellent compared with the traditional baseline methods. It is worth noting that in order to better verify the effect of our method, a small number of frontier topics mined by the traditional methods are also constructed into a local knowledge base. The above knowledge base is mainly used to complete the alignment guidance and the generation style guidance of frontier topics. In the second group comparative experiment of Top 7 frontier topics mining in the field of recommendation systems from 2009 to 2018, except for the "recommendation algorithm and performance evaluation" topic that were not successfully identified, the rest of the topics were accurately generated, and an early potential topic of "AI recommendation system" was additionally generated, with both Precision and Recall were 86%, while DeepSeek only mined two correct frontier topics: cold start and data sparsity, privacy protection and federated learning. To further validate the performance differences between the D-Agents method and the topic model, we compared the mining results with the KPLDA [5] model and BERTopic [25] model. The comparative results of frontier topics mining are presented in Table 8.

Table 8. Comparative experimental results of frontier topic mining with KPLDA and BERTopic (red indicates inconsistency with expert mining results).

Domain Frontier Topic Mining	KPLDA [5] Mining Results	BERTopic [25] Mining Results
Mining Top7 frontier topics in the field of "Recommendation System" from 2009 to 2018 [20]	Explainable recommendation, Reinforcement learning for recommendation system,Collaborative filtering and matrix factorization, Fairness & diversity,Knowledge graph for recommendation system, Cold start & data sparsity, AI recommendation system	Collaborative filtering, Knowledge graph for recommendation system, Personalized recommendation, Fairness & diversity, Factorization Machines with libFM, Privacy protection, AI recommendation system

According to the experimental results in Table 7 and Table 8, We can observe that in the same task, the KPLDA and BERTopic model exhibits certain performance gaps compared to the D-Agents method. Specifically, in the extraction of the Top 7 frontier topics, KPLDA only correctly identified two (Collaborative filtering and matrix factorization, Cold start & data sparsity) frontier topics, and BERTopic just correctly identified three (Collaborative filtering, Personalized recommendation, Privacy protection) frontier topics, while the remaining ones were inaccurately recognized. It should be pointed out that this study mainly focuses on the new paradigm of end-to-end mining based on AI agents, so a representative static topic model is selected in the baseline comparison. Dynamic topic model can capture the evolution of research trends and is an important technical path in frontier topic mining. Incorporating it into a wider comparison is a direction to improve the evaluation system in the future.From the perspective of comprehensive performance, efficiency and domain adaptability, the experimental results can fully show that the D-Agents has the ability to mining frontier topics in general field. Even in some fields, the comprehensive performance is close to the level of human experts. In terms of efficiency, because it is a semi-supervised method, it does not need to rely on domain experts to label data. Directly via agent, it can quickly complete a frontier topic mining task in a specified domain in a short period of time, which can greatly improve mining efficiency. In terms of domain adaptability, it uses large language models in general domains as the brain of the agent, which can simultaneously adapt to the mining requirements of frontier topics in multiple domains. The above results can fully illustrate that the DeepSeek model lacks high-quality prompt information and self-verification mechanism, as well as the style guidance and alignment guidance of the local knowledge base. Its comprehensive performance in frontier topic mining task has some gaps with the D-Agents method.

6. DISCUSSION

The D-Agents method has demonstrated superior comprehensive performance and high mining efficiency in frontier topic mining experiments. However, during the experimental process, we identified several aspects that warrant further optimization. It can be summarized as follows: ① Before creating the agent, in-depth analysis according to the task requirements. Then, combined with the analysis results, setting clear task objectives and high-quality planning schemes for the agent, equipping some optional tool sets and selecting appropriate memory strategies will have a great impact on the overall performance of the agent. ② The overall performance of D-Agents is better than G-Agent. Because the D-Agents have prompt instructions to enhance execution functions. When the G-Agent performs poorly in executing a certain instruction, the generation quality of the model will be significantly improved after the secondary prompt of the D-Agents. In the experiment, it is found that this phenomenon will exist in the generation style and alignment instruction of frontier topics, which is called indirect reinforcement feedback learning, which is an emerging learning mechanism automatically realized by model without manual intervention. ③ The experiment shows that high-quality local knowledge base has an important influence on frontier topics generation. Because combining the local knowledge base can effectively guide the agent to generate frontier topics according to the user-specified style and complete automatic alignment. It is suggested that in the future task, according to the actual requirement of users, building a high-quality local knowledge base and integrating it into the agent can effectively improve the generation quality of frontier topics. ④ Setting the maximum K value of candidate frontier topics returned by G-Agent will also have a certain impact on the overall effect. For different fields, during multiple rounds of experiments, it was found that when the K value is set too small, important frontier topics may be missed. When the K value is set too large, some noisy topics may be generated, which will affect the generation quality. In the CV domain, the generation results of frontier topics corresponding to K = 10 and K = 30 are shown in Figure. A4 in the Appendix A. Therefore, it is recommended to set the initialized K value to an integer between [3, 30]. In the experiment, the initialized K value is set to 20 by default. It is recommended that users conduct in-depth analysis and multiple rounds of feedback fine-tuning according to the selected domain before determining the optimal K value, which will help to improve the quality of frontier topic generation. To sum up, D-Agents are fully competent for frontier topic mining tasks after multiple rounds of iterative optimization.⑤We follow the human-in-the-loop (HITL) approach when designing the architecture, fully considering the uniqueness and security of the task. Considering that the D-Agents method introduces serial generation and verification steps, compared with a single G-Agent, the computational overhead and time cost of the whole mining process will increase. With the experimental setup and task size of this study, we observe that this increase in overhead is controllable, and its main cost is reflected in the additional computation of the HITL validation step. However, the improvement of generation quality brought by this design (such as higher accuracy and recall rate) is considered to be a higher priority goal. Future work will conduct a more accurate quantitative assessment of time overhead and performance gain at different task scales by designing control experiments. ⑥It is worth noting that regarding the universality of the method, although this study mainly builds agents (G-Agent and V-Agent) based on Baidu ERNIE series of large language models and verifies their effectiveness, the proposed D-Agents architecture itself is model-independent. The core of this architecture lies in the "generate-verify" collaborative paradigm, task planning based on prompt engineering, and knowledge utilization mechanism combined with RAG.

Limitations: The datasets used for validation were limited in size, this may affect the generalizability of our findings, particularly for rare or domain-specific phenomena. Future work could explore larger-scale benchmarks or synthetic data augmentation to mitigate this issue. We recommend replicating experiments with multiple runs in future studies to assess result consistency, though we note that our reported trends align with established literature in similar settings. The thresholds for decision-making were chosen based on preliminary analysis and domain expertise, but their sensitivity to minor data shifts remains unquantified. To improve robustness, we plan to incorporate sensitivity analysis or adaptive thresholding methods. In the future, we will release a D-Agent online tool with interpretable features (e.g., including deep thinking and reasoning process).

7. CONCLUSIONS

A frontier topic mining method based on an AI-agent framework is proposed by integrating a general large language model with prompt engineering techniques. The study explores an automatic cross-validation mechanism for knowledge sharing among multiple agents and confirms the effectiveness of employing AI-agents to autonomously generate candidate frontier topics. Experimental results demonstrate that the D-Agents method is capable of effectively identifying frontier topics across multiple domains, achieving strong comprehensive performance on benchmark datasets. This approach overcomes the limitation of traditional frontier topic mining methods, which heavily rely on domain experts to collect and review extensive literature. It establishes a novel paradigm—using a G-Agent to automatically generate candidate frontier topics, followed by a V-Agent to validate them—offering an efficient and intelligent solution for frontier topic discovery in diverse academic fields. The experiments were repeated on three small-scale independently labeled datasets, and the results showed a consistent trend, thus indirectly supporting the effectiveness of the method.

It is worth noting that although the D-Agents method shows strong comprehensive performance in frontier topic mining, there is still room for further optimization. However, the proposed method has only been experimentally verified on agents with Baidu ERNIE large language model as the brain, focus on verifying the universality and performance of this method in other mainstream large language models (such as GPT series, LLaMA series, etc.) as the brain of the agent in the future, so as to further confirm its model-independent design advantages in the future. In addition, due to the diversity of the underlying large model, the agent generation results may cause slight changes in the generated frontier topics. Future work can focus on exploring the universality of the method and proposing high-quality knowledge base auxiliary agents in general domains to complete frontier topics alignment and generation style guidance, thereby further improving the mining quality of frontier topics. In addition, considering that the citation network contains the inheritance and development of academic research, it can mine novel research directions and evidence that may be missed based on text retrieval, and help judge the accuracy of topic generation from the perspective of academic development. In the future, it is also an important research direction to integrating other dynamic topic models for subtask validation and further optimize the proposed method in combination with the above technologies.

Acknowledgments

This research was supported by the research project on teaching reform in general colleges and universities of Hunan Province (HNJG-2022-0429). Thanks to the agent online management platform provided by Baidu Company for helping to build relevant agents and complete experimental.

CONFLICTS OF INTEREST

All authors declare that there are no financial or personal relationships with other people or organizations that could inappropriately influence (bias) their work. No competing interests exist in the conduct of this research or the preparation of this manuscript.

APPENDIX A

Figure A1. Frontier topics generation results of G-Agent and D-Agents on CV-DataSet.

Figure A2. Frontier topics generation results of G-Agent and D-Agents on NLP-DataSet

.

Figure A3. Frontier topics generation results of G-Agent and D-Agents on ML-DataSet.

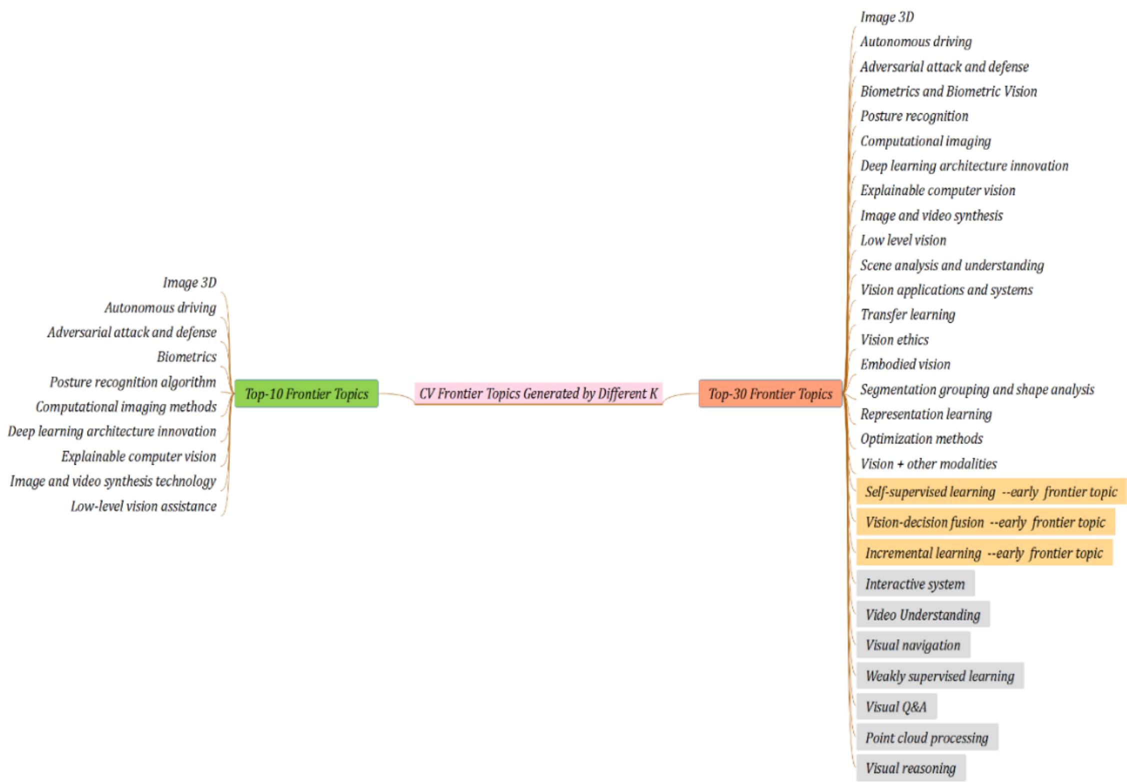

Figure A4. Frontier topics generated by D-Agents in the CV domain according to different K.

REFERENCES

[1] Zhang H, Yang X, Zhao X, et al. Subject Frontiers Hot Spots Mining Based on Social Network Attention[J]. Journal of Zhengzhou University (Natural Science Edition), 2018, 50(03):46-52. https://link.oversea.cnki.net/doi/10.13705/j.issn.1671-6841.2017201.

[2] Prapobratanakul, Chariya. Frequency Analysis, Distribution, and Coverage of Academic Words in Materials Science Research Articles: A Corpus-Based Study [J]. LEARN Journal: Language Education and Acquisition Research Network 17.2, 2024: 793-813. https://doi.org/10.70730/sjnm9318.

[3] Kyebambe, M, et al. Forecasting emerging technologies: A supervised learning approach through patent analysis [J]. Technological Forecasting and Social Change 125, 2017: 236-244. https://doi.org/10.1016/j.techfore.2017.08.002.

[4] Klarin, Anton. How to conduct a bibliometric content analysis: Guidelines and contributions of content co-occurrence or co-word literature reviews [J]. International Journal of Consumer Studies 48.2, 2024: e13031. https://doi.org/10.1111/ijcs.13031.

[5] Ge, B, et al. Chinese news hot subtopic discovery and recommendation method based on key phrase and the LDA model [J]. DEStech Transactions on Engineering and Technology Research, ECAR, 2018. https://doi.org/10.12783/dtetr/ecar2018/26371.

[6] Lalk, Christopher, et al. Measuring alliance and symptom severity in psychotherapy transcripts using bert topic modeling [J]. Administration and Policy in Mental Health and Mental Health Services Research 51.4, 2024: 509-524. https://doi.org/10.1007/s10488-024-01356-4 PMID: 38551767.

[7] He, C, Ge B, and Zhang C. Chinese Text Open Domain Tag Generation Method via Large Language Model [C].10th International Conference on Big Data and Information Analytics. IEEE, 2024, 183-188. https://doi.org/10.1109/bigdia63733.2024.10808200.

[8] Jin, Y, et al. AGENTREVIEW: Exploring Peer Review Dynamics with LLM Agents [C]. //In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 1208–1226, Miami, Florida, USA. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.70.

[9] Giray, L. Prompt Engineering with ChatGPT: A Guide for Academic Writers [J]. Ann Biomed Eng 51, 2023, 2629–2633. https://doi.org/10.1007/s10439-023-03272-4 PMID: 37284994.

[10] Hu H, Xue W, Jiang P, et al. Bibliometric analysis for ocean renewable energy: A comprehensive review for hotspots, frontiers, and emerging trends [J]. Renewable and Sustainable Energy Reviews. 2022, 167:112739. https://doi.org/10.1016/j.rser.2022.112739.

[11] Gao Y, Xu Y, Zhu Y, et al. An analysis of the hotspot and frontier of mine eco-environment restoration doul big data visualization of VOSviewer and CiteSpace [J]. Geological Bulletin of China. 2018, 37(12):2144-53. http://dx.chinadoi.cn/10.12097/j.issn.1671-2552.2018.12.004.

[12] Zhang Y, Zhao D, Liu H, et al. Research hotspots and frontiers in agricultural multispectral technology: Bibliometrics and scientometrics analysis of the Web of Science [J]. Frontiers in Plant Science. 2022, 13:955340. https://doi.org/10.3389/fpls.2022.955340 PMID: 36035687.

[13] Han Q, Li Z, Fu Y, et al. Analyzing the research landscape: mapping frontiers and hot spots in anti-cancer research using bibliometric analysis and research network pharmacology [J]. Frontiers in Pharmacology. 2023, 14:1256188. https://doi.org/10.3389/fphar.2023.1256188 PMID: 37745055.

[14] Zhang N, Feng G. Analysis of Big Data Research Hotspots Based on Keyword Co-occurrence [C]. In IEEE 9th International Conference on Data Science in Cyberspace, 2024, 464-471. https://doi.org/10.1109/dsc63484.2024.00070.

[15] Zhang B, Zhang S.Analysis of the Sci-tech Industry Research Hot Spots Based on Keywords Co-occurrence Analysis and Social Network Analysis[J].Science and Technology Management Research, 2016(20): 93-98.http://dx.chinadoi.cn/10.3969/j.issn.1000-7695.2016.20.018.

[16] Nie J. Study on the Law and Mechanism of Tibetan Medicine's Prevention and Treatment of Plateau Disease Based on Knowledge Discovery [D]. Chengdu University of Traditional Chinese Medicine, 2017.

[17] Mi L, Zhang W, Yu H, et al. Knowledge mapping analysis of pro-environmental behaviors: research hotspots, trends and frontiers [J]. Environment, Development and Sustainability. 2024:1-35. https://doi.org/10.1007/s10668-024-05046-x.

[18] Li B, Qin Y,Xu Z. Dynamic Tracking of Research Status and Frontiers in the Field of Management Science [J]. Chinese Journal of Management Science, 2023, 31(07):276-286. https://link.oversea.cnki.net/doi/10.16381/j.cnki.issn1003-207x.2021.1830.

[19] Xu J, Sun S, Zhao Y, et al. Knowledge domain, research hotspots and frontiers in physiology teaching reforms from 2012 to 2021: A bibliometric and knowledge-map analysis [J]. Frontiers in Medicine. 2023, 10:1031713. https://doi.org/10.3389/fmed.2023.1031713 PMID: 37020677.
[20] Shao B, Li X, Bian G. A survey of research hotspots and frontier trends of recommendation systems from the perspective of knowledge graph [J]. Expert Systems with Applications. 2021, 165:113764. https://doi.org/10.1016/j.eswa.2020.113764.
[21] Cheng J, Lu D, Sun L, et al. Development Trends, Current Hotspots, and Research Frontiers of Oyster Reefs: A Bibliometric Analysis Based on CiteSpace [J]. Water. 2023, 15(20):3619. https://doi.org/10.3390/w15203619.
[22] Zhang P, Zhang J, Wang M, et al. Research hotspots and trends of neuroimaging in social anxiety: a CiteSpace bibliometric analysis based on Web of Science and Scopus database [J]. Frontiers in Behavioral Neuroscience. 2024, 18:1448412. https://doi.org/10.3389/fnbeh.2024.1448412 PMID: 39713279.
[23] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks[J]. Advances in neural information processing systems, 2020, 33: 9459-9474.
[24] Hang C, Yu P, Tan C. TrumorGPT: Graph-Based Retrieval-Augmented Large Language Model for Fact-Checking[J]. IEEE Transactions on Artificial Intelligence, 2025, 6(11), pp. 3148-3162.
[25] Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure[J]. arXiv preprint arXiv:2203.05794, 2022.
[26] Baidu. Agent Builder [EB/OL]. https://agents.baidu.com/agent, 2024-04-19/2025-02-25.
[27] Baidu. ERNIE-4.0-8k [EB/OL]. https://cloud.baidu.com/doc/WENXINWORKSHOP/s/clntwmv7t, 2025-02-20/2025-02-25.
[28] Baidu.ERNIE-3.5-128k [EB/OL]. https://cloud.baidu.com/doc/WENXINWORKSHOP/s/dlw4ptsq7, 2025-02-21/20025-02-25.
[29] Deng Z, Ma W, Han Q, et al. Exploring DeepSeek: A Survey on Advances, Applications, Challengesand Future Directions[J]. IEEE/CAA Journal of Automatica Sinica, 2025, 12(5): 872-893.

AUTHOR

Bin Ge was born in Qingdao, Shandong Province in 1979. He received the Ph.D. degree from theNational University of Defense Technology, China. He is a Researcher, and the research interests include big data analysis and information extraction. ORCID: https://orcid.org/0009-0005-5893-9082

Chunhui He was born in Yongzhou, Hunan Province in 1991, and obtained the mathematics master's degree from Xiangtan University in 2017. His main research interests are information processing and artificial intelligence application.
Corresponding author of this paper. ORCID: http://orcid.org/0000-0003-1505-1620