Research on the Full-Link Correction System of the RE Model —— Basic Construction of the First Stage of the "Deification Plan"

HAO Ziqian

Abstract

This paper, as the second core paper of the "Deification Plan", aims to systematically and comprehensively define the RE model, clarify the origin of its exclusive name, core connotation and essential characteristics, and strictly distinguish the essential differences between the RE model and true artificial intelligence (AI). It theoretically corrects the widespread misconception in the current artificial intelligence industry that "the RE model is equivalent to true AI", completes the adaptation and docking of the self-developed theoretical system with the existing AI academic system, and makes up for the shortcomings of the academic foundation of conceptual innovation. On this basis, this paper deeply analyzes the three core drawbacks existing in the actual operation of current commercial RE models, explores the internal causal relationship and underlying generation roots of these drawbacks, and designs three lightweight embedded correction modules that do not require the reconstruction of the existing model architecture. Combined with the current status of AI technology, it clarifies the module algorithm selection and supporting technical schemes, refines the update mechanism of the authoritative information source library and differentiated parameters for industry scenarios, formulates module conflict handling rules, sets engineering operation parameters in line with the current technical level, and improves the full-link closed-loop intervention logic. This significantly enhances the accuracy of knowledge output and logical rigor of the RE model, and effectively reduces the occurrence of model hallucinations. Meanwhile, this paper supplements the review of relevant research in the field and the analysis of the limitations of the

scheme, defines the differentiated quantitative standards between the preliminary research stage and the final complete version of true AI, and strengthens the academic rigor and objectivity of the paper. All the research content, theoretical derivation and scheme design of this paper are specially dedicated to the implementation of the first stage of the "Deification Plan", clarifies the quantitative acceptance criteria of the stage, and provides core theoretical support and practical foundation for the research and development of true artificial intelligence, laying a solid foundation for the steady advancement of the subsequent stages of the "Deification Plan".

Keywords

RE model; conceptual definition; true artificial intelligence; modular correction; the first stage of the "Deification Plan"; model hallucination; data reflection; intelligent top-level framework

1. Introduction

The ultimate core goal of the "Deification Plan" is to develop true artificial intelligence (AI) with capabilities of autonomous rigorous reasoning, precise knowledge output, real-time self-verification and continuous self-evolution. However, various commercial intelligent large models widely used and popularized in the global artificial intelligence market are not true artificial intelligence in essence. Due to factors such as conceptual confusion in the industry, commercial publicity orientation and fragmentation of the existing classification system, they are misunderstood and blindly equated by industry practitioners and the public. The existing artificial intelligence classification system is mostly divided from superficial dimensions such as technical paths, application scenarios and modal types, deriving numerous terms such as large language models (LLM), generative AI, conversational AI and multimodal models. It fails to touch the essential attributes of the models, leading to confusion in industry theoretical cognition and the lack of a unified top-level classification standard.

Based on the essential attributes of intelligent technology, this paper takes the lead in constructing two parallel top-level frameworks of RE and true AI, systematically

explains and analyzes the core characteristics of the RE model, completes the adaptation of self-developed concepts with existing academic systems such as connectionism and narrow artificial intelligence (ANI), and draws a clear boundary between the RE model and true artificial intelligence from the theoretical root, laying a solid theoretical foundation for the overall advancement of the "Deification Plan". Combined with the current research status of fact verification, hallucination suppression and data governance in the field, this paper proposes a targeted, low-cost, easy-to-implement and highly adaptable modular correction scheme for the fatal defects exposed by the RE model in actual application scenarios. It refines the module engineering implementation details, algorithm selection, authoritative information source support, conflict handling mechanism and industry differentiated parameters, sets resource occupation and delay thresholds in line with the current technical level, makes up for the weaknesses of pure theoretical conception, and optimizes the core performance of the existing RE model with minimal technical modification and fastest implementation speed. All the content of this research belongs to the core work scope of the first stage of the "Deification Plan", clarifies the quantitative assessment indicators of the first stage, distinguishes the quantitative standards between the preliminary research stage and the final complete version of true AI, and aims to accumulate key technical experience and practical data by correcting the prominent problems of the existing RE model, clearing cognitive, technical and application obstacles for the research and development of true artificial intelligence in subsequent stages, and consolidating the foundation for the implementation of the "Deification Plan".

2. Core Introduction, Definition and Specific Content of the RE Model

2.1 Core Definition and Connotation of the RE Model

RE is the exclusive initials of Reflection-Engine, which is one of the two parallel top-level frameworks divided based on the essential attributes of the intelligent field in the exclusive theoretical system of the "Deification Plan". It is a parallel,

non-subordinate and clearly bounded core classification with true artificial intelligence (AI), not a sub-model concept in the industry, thus completely solving the problem of blurred boundaries of conceptual innovation and ensuring the rigor of the theoretical system. Meanwhile, to adapt to the existing artificial intelligence academic system, it is clarified that the RE framework corresponds to the connectionist narrow artificial intelligence (ANI) in the existing academic system, covering all commercial intelligent models based on deep learning and data fitting, realizing the seamless connection of self-developed concepts with traditional academic classification, and eliminating the shortcomings of "self-developed concepts divorced from academic foundations".

As an independent top-level framework, RE is the only, unified and full-coverage essential general term for all commercial intelligent large models that have been commercially implemented and applied on a large scale worldwide. Its essence is a non-autonomous intelligent auxiliary tool that performs probability fitting and content reflection generation based on massive multi-source training data through deep learning core algorithms. It is not an intelligent subject with independent thinking, logical judgment, autonomous error correction and self-evolution capabilities. The core operation logic of the RE model is to reproduce, integrate, reorganize and formally output various information such as text, data, knowledge and language produced by human society. It generates content products that meet the requirements in form through algorithm matching and user instructions, and has no ability to independently judge and verify the authenticity, logical rationality and factual accuracy of knowledge throughout the process. It is a typical "data mapping tool" and by no means a subject with core intelligence in the true sense.

It should be clarified that various terms such as generative AI, large language models (LLM), conversational AI and multimodal models in the industry are not top-level framework concepts. They are only subdivided application forms, interaction modes or technical names under the RE top-level framework, all belonging to the RE framework, and not at the same level as the two top-level frameworks of RE and true AI, let alone parallel to them.

2.2 Operating Mechanism and Technical Characteristics of the RE Model

From the underlying technical perspective, the RE model generally takes the Transformer architecture as the core technical basis and adopts the mainstream technical mode of "large-scale pre-training + scenario-based fine-tuning". It completes basic model training by ingesting massive Internet public data, multi-field text materials and various general and vertical knowledge bases, and then carries out model parameter fine-tuning and effect optimization based on massive user interaction data. Its technical characteristics show distinct limitations and tool attributes: first, it is extremely data-dependent, and all output content completely relies on training data reserves, with no ability to independently conduct in-depth learning and expand knowledge boundaries, and cannot generate effective content without existing data; second, it operates passively, and the content generation process can only be started after receiving clear user instructions, with no attributes of active perception, active thinking and active task initiation, and is completely driven by external instructions; third, it performs shallow fitting output, which can only complete surface language matching, content splicing and formal generation, and cannot deeply understand the deep logic, core semantics and real user needs behind the problem, let alone conduct in-depth logical reasoning and rigorous fact verification, which is also the underlying technical root of various application drawbacks of the RE model.

2.3 Application Scope and Essential Positioning of the RE Model

At present, the application scenarios of the RE model cover many fields such as text creation, intelligent Q&A consulting, code generation, content summary extraction, intelligent customer service interaction and multi-field content generation. It is a commercial intelligent product form that absolutely dominates the market, and all general large models and industry vertical large models are uniformly classified into the RE model category. However, from the essential positioning, the RE model is only an auxiliary tool to improve content generation efficiency and simplify repetitive work. Its core value is to quickly produce formal and standardized content, rather than transmit precise, rigorous, true and authoritative knowledge content. This is also the

core and essential difference between the RE model and true artificial intelligence, and the core theoretical basis for the subsequent analysis of RE model drawbacks and modular correction scheme design in this paper.

3. Comprehensive Definition of the RE Concept and Essential Differences from True AI

3.1 Name Definition of RE

RE is the exclusive initials of Reflection-Engine, an exclusive term for one of the two top-level core frameworks in the theoretical system of this paper and the "Deification Plan". It has no other derived meanings or extended interpretations, specifically referring to all commercial intelligent model sets with data reflection as the core and no autonomous intelligence. It is highly consistent with the essential connotation of narrow artificial intelligence (ANI) and connectionist AI in the existing academic system, and is a precise condensation and new definition of the essential attributes of such models.

3.2 Specific Definition of RE

As a top-level framework parallel to true artificial intelligence, RE is essentially a set of shallow intelligent tools that perform probability fitting, pattern matching and text reflection generation based on massive multi-source data. Its core operation logic is to integrate, reproduce and formally output existing massive data, and has no core intelligent capabilities such as autonomous thinking, independent reasoning, precise verification and self-error correction. Such models can only passively receive user instructions, generate corresponding content based on training data reserves, cannot independently judge the authenticity of knowledge and logical right or wrong, and cannot actively admit cognitive limitations when their own knowledge reserves are insufficient to answer accurately. It is a typical "data reflection tool" framework and by no means a subject framework with true intelligent attributes.

3.3 Essential Differences Between RE and True Artificial Intelligence (AI)

True artificial intelligence (AI) is another top-level framework in the intelligent field that is completely parallel and essentially opposite to RE. It is the ultimate intelligent

form that the "Deification Plan" aims to develop, corresponding to the advanced form of general artificial intelligence (AGI) in the academic field. It has core capabilities such as autonomous consciousness, rigorous reasoning, precise judgment, self-verification, active error correction and continuous evolution, and distinguishes differentiated quantitative criteria for the "preliminary research stage" and the "final complete version":

1. Final complete version of true AI (non-reducible standards): First, it can independently complete multi-dimensional logical reasoning with an accuracy rate ≥ 99%; second, the self-verification coverage rate reaches 100%, and output is refused without effective verification basis; third, it can independently identify and declare knowledge blind areas with a blind area declaration accuracy rate of 100%; fourth, it can independently optimize reasoning logic through practical feedback without manual fine-tuning, and the autonomous evolution iteration efficiency ≥ 95%.

2. Preliminary research stage true AI (moderately relaxed standards): Considering the limitations of preliminary technical accumulation and R&D conditions, the requirements can be moderately reduced: the autonomous logical reasoning accuracy rate ≥ 90%, the self-verification coverage rate ≥ 90%, it can independently identify and declare most knowledge blind areas, has basic autonomous optimization capabilities, and does not require large-scale manual intervention. After technical maturity, it will gradually move closer to the final standards of the complete version and must strictly meet the standards in the end.

True AI takes the pursuit of absolute knowledge accuracy and absolute logical rigor as the core operation goal. It can independently identify the authenticity of information, judge the feasibility of answers, verify the rationality of content, and actively declare cognitive limitations when unable to answer accurately. It is an individual framework with deep intelligence and independent thinking ability.

The essential differences between the two can be systematically summarized into three points: first, the core kernel is different. The core of the RE framework is "data reflection" with no core intelligence and only data reproduction; the core of the true AI framework is "autonomous intelligence" with a complete core kernel and

independent thinking ability. Second, the operation logic is different. The RE framework takes "answer all questions, complete formal output" as the core goal, focusing on form over substance; the true AI framework takes "answer correctly, output accurately and effectively" as the core criterion, focusing on substance over form. Third, the subject attributes are different. All tools in the RE framework are passive, with no self-verification, error correction and evolution capabilities, and completely rely on external data and instructions; the subjects in the true AI framework are autonomous, with complete self-correction, optimization and evolution capabilities, and can perform intelligent iteration without external intervention. The biggest misunderstanding in the current artificial intelligence industry is equating the subdivided tools under the RE framework with the true AI framework, which is also the core theoretical and cognitive issue that the "Deification Plan" first needs to correct.

4. Three Core Drawbacks of the RE Model

Combined with the underlying operating mechanism and actual application performance of the RE model, its three core drawbacks are not isolated, but have a clear causal conduction chain of "data source - content security - output logic". The chaos of training data is the root cause, malicious data tampering is a derivative risk, and prioritizing response over correctness is an underlying design defect. The three factors together lead to the lack of reliability of the RE model. The specific drawbacks and internal causal relationships are analyzed as follows:

4.1 Chaotic Training Resources and Frequent Output Conclusion Errors

The training data sources of the RE model are excessively broad and chaotic, lacking strict, unified and authoritative screening standards and access mechanisms. The training data includes a large amount of low-quality resources such as network redundant information, non-authoritative fragmented content, subjective speculative remarks, common-sense wrong knowledge and false misleading content, and a standardized data screening system that prioritizes authoritative knowledge and eliminates false content has not been established. This root cause directly leads to the

flooding of a large amount of false content and distorted information in the underlying knowledge base of the RE model, which is the core incentive for model output errors and factual deviations. When outputting response content to the public, common-sense errors, professional field content deviations and factual information distortions frequently occur. The wrong content spreads rapidly through the high propagation of intelligent models, continuously misleading public cognition, especially in fields such as education popularization, daily life knowledge and professional consulting.

4.2 Vulnerability to Malicious Data Injection and Tampering of Core Knowledge

This drawback is a direct derivative risk of the disorder of training resources. Due to the lack of effective security protection, content screening and malicious information interception mechanisms in the data access link of the RE model, the data entry is too open and lacks supervision. On the basis of low-quality training data, it further leaves loopholes for malicious data injection, making it easy to be injected with false data, tampered with authoritative knowledge, fabricated false facts and implanted wrong theories. Malicious actors can damage the knowledge system of the RE model by targeted injection of malicious content, pollution of training data and interference with interaction samples, causing the RE model to output tampered false information and wrong content as correct knowledge. This malicious data intervention is not an isolated case, but a prominent hidden danger common in the industry. It can distort the cognition of specific field knowledge and mislead the public in mild cases, and tamper with historical facts, distort scientific principles and damage the authority and authenticity of knowledge in severe cases, forming a bad knowledge dissemination chaos and causing great harm.

4.3 Prioritizing Response Over Correctness and Prominent Model Hallucination Problems

This drawback is an underlying design defect of the RE model, independent of the previous two types of data-level drawbacks. It is the core root cause of model hallucinations and is directly related to the essential kernel of "data reflection" of the RE model. The underlying design logic of the RE model has fundamental and

principled deviations. The product excessively pursues the formal effect of "instant response to all user instructions" rather than taking "outputting correct, precise and effective answers" as the core goal, and the model itself has no fact verification and logical reasoning mechanisms. When the model's own knowledge reserves are insufficient to accurately judge the answer and there are ambiguous or controversial information, the RE model will not actively and clearly admit its cognitive limitations. Instead, it will forcefully generate response content through content piecing together, subjective speculation, vague expression and logical confusion, completely ignoring the authenticity, accuracy and logic of the content, and thus generating a large number of model hallucinations. This wrong logic of "preferring wrong answers to no answers" turns the RE model into a simple perfunctory response tool, with numerous loopholes in output content and extremely poor rigor. It completely deviates from the rigor and reliability that intelligent models should have. Even if the previous two types of data problems are solved, this design defect will still make it impossible to completely eradicate hallucinations.

5. Review of Related Research in the Field

At present, many related studies have been carried out in the field of artificial intelligence for various drawbacks of models under the RE framework, mainly focusing on three directions: data governance, hallucination suppression and fact verification, providing industry references for the modular correction scheme of this paper. However, existing studies have certain limitations, which also highlight the innovative value of this research.

In the direction of data governance, existing studies mostly focus on training data cleaning, deduplication and annotation optimization, improving training data quality by establishing data screening rules, but fail to form a full-link embedded interception mechanism and cannot take into account malicious data active defense; in the direction of hallucination suppression, mainstream studies include retrieval-augmented generation (RAG), fact verification models, prompt optimization, etc., focusing on output-side correction, failing to perform simultaneous intervention

from the source and reasoning end, and the correction coverage is insufficient; in the direction of fact verification, most studies adopt methods such as external knowledge base comparison and logical rule constraints, which are independent tools and not lightweight integrated with the model itself, resulting in high implementation costs.

The modular correction scheme proposed in this paper is different from the existing single-link correction ideas. It builds a "source-intermediate-terminal" full-link closed-loop system, is equipped with an authoritative dedicated information source library, combines the current mainstream AI algorithm selection, formulates a complete conflict handling and update mechanism, and does not require reconstructing the model architecture. It takes into account lightweight, low cost and comprehensiveness, makes up for the shortcomings of existing studies, and complements existing technologies to adapt to the optimization needs of various RE models.

6. Design of Three Lightweight Embedded Correction Modules for the RE Model

Aiming at the three core drawbacks and internal causal relationships of the RE model, this paper adheres to the core principles of "no need to completely reconstruct the existing RE model architecture, lightweight embedded design, rapid implementation and adaptation, low-cost and high-efficiency optimization" to design three independent, collaborative and logically closed correction modules. It clarifies the module technical architecture, algorithm selection, access method, collaboration mechanism, authoritative information source support, conflict handling rules and industry differentiated parameters, sets engineering operation parameters in line with the current technical level, and embeds them into the key links of the RE model’s input, retrieval and output full links to build a "source-interception-intermediate-control-terminal-verification" full-link correction system, realizing all-round and in-depth optimization and correction of the RE model. The modules correspond to the drawbacks one by one, forming a logical chain of "one module for one drawback, multiple modules for closed-loop supplementation".

6.1 Input Port Knowledge Interception and Screening Module

This module, as the first core defense line of the RE model's data input link, is directly deployed at the "data input API interface layer" of the RE model in a lightweight plug-in access mode without modifying the underlying weights and reasoning architecture of the model. Its core function is to conduct all-round, high-standard and strict screening and interception of all retrieved knowledge, training data, external feeding information and interaction data entering the model, completely eliminating wrong knowledge, invalid information, redundant content and malicious data from the source, purifying the knowledge input channel of the RE model, and targeted solving the two source drawbacks of chaotic training resources and malicious data tampering.

6.1.1 Module Algorithm Selection and Supporting Technologies

A "multi-algorithm fusion filtering architecture" is adopted, combining rule engine and deep learning algorithms to balance filtering efficiency and accuracy:

1. Basic rule filtering algorithm: MD5 hash deduplication algorithm + sensitive word regular matching algorithm, quickly eliminating duplicate data, private information and malicious sensitive content with processing delay ≤ 2ms;

2. Deep learning classification algorithm: Text quality classification algorithm based on the BERT-base lightweight model, scoring the authority, authenticity and objectivity of data, and directly intercepting content with a score lower than 60;

3. Abnormal data detection algorithm: Isolation forest abnormal detection algorithm, identifying logical contradictions, factual errors and fabricated content in data, and intercepting tampered data injected maliciously;

4. Supporting technologies: Equipped with a malicious data feature library and an authoritative information source whitelist, and adopting quantization distillation technology to compress the model to ensure lightweight operation.

6.1.2 Configuration and Limitations of the Authoritative Information Source Library

The module is built with a "national-level authoritative information source library + professional field core classics library + industry standard database", including

specific sources such as CNKI academic journal database, Standardization Administration of China standard full-text database, Chinese classic classics library, official industry white papers, core research results of universities and research institutes, and national-level popular science resource library; the malicious content feature library is synchronously connected with the malicious information feature library of the National Cyberspace Security Information Center, and the invalid information filtering rules are formulated based on general industry data cleaning standards.

It should be particularly pointed out that there are serious problems such as academic 灌水, data falsification and fabricated content in the current Chinese academic circle. The authenticity and authority of content included in domestic academic databases represented by CNKI cannot be fully guaranteed, and can only be used as a "temporary alternative information source in the preliminary research stage". It cannot meet the core requirement of absolute knowledge accuracy for true AI R&D. In the later stage, a "complete and independent authoritative information source system integrating multiple countries and libraries" must be built, integrating resources from authoritative institutions in multiple countries such as the International Organization for Standardization (ISO) standard library, top global scientific research journal databases (Nature, Science, Cell, PNAS, etc.), EU scientific research database, National Institute of Standards and Technology (NIST) database of the United States and Japanese Industrial Standards Committee (JISC) database, realizing mutual verification and cross-error correction of information sources from different countries and institutions, and fundamentally guaranteeing the authenticity and authority of information sources.

6.1.3 Update Mechanism of the Authoritative Information Source Library

1. Update frequency:

- Regular update: Basic general information sources are automatically updated once a week, and industry standard and classic information sources are automatically verified and updated once a month;
- Emergency update: When authoritative institutions issue errata, revisions or urgent

corrections, an immediate update is triggered in real time, and full-library synchronization is completed within 24 hours;

- Regular comprehensive verification: A full-library data verification is conducted once a quarter to eliminate invalid, wrong and outdated content and supplement the latest authoritative content.

2. Update trigger conditions:

- Authoritative information source publishing institutions actively push updated content;
- The module detects differences between existing information source content and cross-verification of multi-national information sources;
- Major revisions of industry policies, standards and knowledge;
- When falsification or wrong content is found in the information source, a special update is immediately initiated to eliminate it.

3. Update process: Automatically capture updated content → cross-verify multi-national information sources → manual review (core fields) → storage and activation → synchronization to the three modules, without affecting the normal operation of the modules during the update.

6.1.4 Engineering Operation Parameters

This module is compatible with all open-source and closed-source RE models, only performing preprocessing on input data, with a single data processing delay ≤ 10ms and an overall resource occupancy rate ≤ 3%. It does not increase the core reasoning delay of the model, adapts to the optimization needs of RE models in all scenarios, and the delay fluctuation is ≤ 5ms even in high-concurrency calling scenarios.

6.2 Retrieval-End Logical Rigor Module

This module connects the knowledge retrieval and reasoning generation links of the RE model, located after the input interception and screening module and before the output verification module, deployed at the "model reasoning computing layer" in a lightweight logic embedding mode. Its core function is to fundamentally correct the fundamental deviation of the RE model of "prioritizing response form over content correctness", force the model to establish a core operation criterion of "accurate

answers", eliminate speculative and perfunctory output from the reasoning generation link, reduce model hallucinations from the root, and targeted solve the underlying design defects of the RE model.

6.2.1 Module Algorithm Selection and Supporting Technologies

1. Logical consistency verification algorithm: Reasoning logic verification algorithm based on Markov logic network, detecting logical contradictions and loopholes in the reasoning process to ensure a closed reasoning chain;

2. Fact retrieval and matching algorithm: BM25 dense retrieval algorithm + RAG enhanced retrieval algorithm, accurately matching authoritative information source content and improving retrieval accuracy;

3. Blind area identification algorithm: Knowledge blind area judgment algorithm based on confidence scoring, automatically triggering the blind area declaration mechanism when the retrieval confidence is lower than 70;

4. Supporting technologies: Embedding constrained prompt templates to force the model to follow the rule of "verify first, generate later, and remain silent if unknown", and adopting lightweight reasoning acceleration technology to reduce delay.

6.2.2 Engineering Operation Parameters

This module communicates with the input interception module in real time, receiving screened compliant data to avoid repeated processing, with a single reasoning logic verification delay ≤ 20ms and an overall resource occupancy rate ≤ 4%. The delay fluctuation is ≤ 8ms in high-concurrency scenarios, without affecting the normal reasoning efficiency of the model.

6.3 Output-Terminal Final Verification and Correction Module

This module, as the last core checkpoint for the RE model’s content output, is deployed at the "output interface layer" after the RE model completes content generation and before outputting to users in an external verification mode. It conducts final and all-round review and correction of all content to be pushed to users, serving as the last line of defense to ensure the accuracy of knowledge output, realizing double verification of the RE model’s output content, and forming closed-loop collaboration with the previous two modules to make up for the omissions of

single-module correction.

6.3.1 Module Algorithm Selection and Supporting Technologies

1. Fact consistency verification algorithm: Fact verification algorithm based on RoBERTa, comparing generated content with authoritative information sources to identify factual errors and biased content;

2. Multi-model cross-validation algorithm: Calling 2 lightweight dedicated verification models to cross-verify high-risk content and improve correction accuracy;

3. Text error correction and polishing algorithm: Seq2Seq text correction algorithm, accurately replacing wrong content and sorting out logic to ensure standardized and rigorous output content;

4. Supporting technologies: Building a key entity (data, terminology, conclusion) extraction and comparison mechanism, 100% verifying core knowledge points, and adopting streaming verification technology without blocking the output process.

6.3.2 Engineering Operation Parameters

The module's final verification delay ≤ 15ms, with an overall resource occupancy rate ≤ 3%. The cumulative total delay of the three modules is ≤ 45ms, and the cumulative total resource occupancy rate is ≤ 10%, fully meeting the technical standards of lightweight deployment of current large models with no obvious perceived delay for users.

6.4 Module Conflict Handling Mechanism

Strictly following the priority principle of "source (input module) < intermediate (retrieval module) < terminal (output module)", taking the final verification module result as the core, a closed-loop feedback mechanism is established:

1. Conventional conflict handling: When the results of the input module and the retrieval module are inconsistent, the retrieval module's result shall prevail; when the results of the retrieval module and the output module are inconsistent, the final verification result of the output module shall prevail;

2. Feedback correction mechanism: If the output module verification finds that the content cannot be corrected and there are serious factual errors, it will be directly sent

back to the retrieval-end logical rigor module for re-reasoning;

3. Source re-screening mechanism: If after re-reasoning by the retrieval module, it is still found that there are excessive differences in information sources, contradictions in core knowledge points and no valid conclusion can be formed, it will continue to be sent back to the input port knowledge interception and screening module for re-data screening and information source verification, eliminating problematic information sources and replacing compliant data before re-entering the process;

4. Blocking mechanism: If it still fails to meet the standards after three feedback re-screenings, the current content generation will be directly terminated, and the user will be informed that an accurate answer cannot be provided to prevent the output of wrong content.

6.5 Industry Scenario Differentiated Module Parameter Setting

Combined with the knowledge rigor and risk level requirements of different industries, differentiated module parameters are set to balance universality and industry pertinence:

1. General daily scenarios (life, entertainment, ordinary text creation)

- Input module: Quality scoring threshold ≥ 50, delay ≤ 8ms, resource occupancy ≤ 2%;

- Retrieval module: Confidence threshold ≥ 60, no mandatory multi-source verification;

- Output module: Fact verification coverage rate ≥ 80%, allowing mild expression optimization.

2. Education and popularization scenarios (primary and secondary education, popular science promotion, general education)

- Input module: Quality scoring threshold ≥ 70, only official education information sources are allowed, delay ≤ 10ms;

- Retrieval module: Confidence threshold ≥ 80, mandatory dual-source verification;

- Output module: Fact verification coverage rate 100%, any ambiguous or wrong expression is strictly prohibited.

3. Medical and health scenarios (health consultation, medical knowledge

popularization, medication guidance)

- Input module: Quality scoring threshold ≥ 85, only information sources from the National Health Commission, pharmacopoeia and core medical journals are allowed, delay ≤ 12ms;

- Retrieval module: Confidence threshold ≥ 90, mandatory triple cross-verification, speculation is strictly prohibited;

- Output module: Fact verification coverage rate 100%, content involving diagnosis and treatment and medication needs to be marked with information source, and risky content is forcibly reminded.

4. Legal and financial scenarios (legal consultation, financial knowledge, policy interpretation)

- Input module: Quality scoring threshold ≥ 85, only information sources from laws and regulations databases, financial regulatory agencies and authoritative financial and economic sources are allowed, delay ≤ 12ms;

- Retrieval module: Confidence threshold ≥ 90, legal provisions and policy content need to be accurately matched with the original text, mandatory multi-source verification;

- Output module: Fact verification coverage rate 100%, core clauses and data are accurate, and subjective interpretation is strictly prohibited.

5. Scientific research and academic scenarios (scientific research reference, academic popularization, data citation)

- Input module: Quality scoring threshold ≥ 80, priority is given to information sources from top journals and research institutions, low-quality academic content is eliminated, delay ≤ 10ms;

- Retrieval module: Confidence threshold ≥ 85, data and conclusions need to be marked with information sources;

- Output module: Fact verification coverage rate 100%, tampering with scientific research conclusions and data is strictly prohibited.

6.6 Module Collaboration Mechanism and Compatibility Description

The three modules adopt a collaboration operation mechanism of "series connection as the main, parallel connection as the auxiliary". The data flow path is: input port knowledge interception and screening module → retrieval-end logical rigor module → output-terminal final verification and correction module. The processing results of the previous module are directly transmitted to the next module without additional data transfer, reducing operation loss; a data mutual verification and conflict feedback mechanism is set between modules to form closed-loop correction.

In terms of compatibility and invasiveness, the three modules adopt a lightweight method of "API layer plug-in, prompt layer embedding" with no invasion of the underlying model architecture and no need to modify model weights. They are compatible with all closed-source commercial RE models such as GPT series, Wenxin Yiyan and Tongyi Qianwen, as well as all open-source RE models such as Llama, Qwen, Baichuan and ChatGLM. No additional hardware environment is required, and they can be quickly deployed and implemented with strong universality.

7. Relevance of This Research to the First Stage of the "Deification Plan"

The detailed introduction of the RE model, core concept definition, in-depth analysis of existing drawbacks and design of three modular correction schemes carried out in this paper are all carried out around the core tasks of the first stage of the "Deification Plan", serving as the core theoretical support, key implementation measures and important practical foundation of this stage.

The first stage of the "Deification Plan" is the "RE model optimization and basic technology accumulation stage". The core goal is not to directly develop true artificial intelligence with complete intelligence, but to promote core work in two steps and formulate clear quantitative acceptance criteria: first, theoretically and industrially correct the common misunderstanding that "the RE model is equivalent to the true AI model", clarify the boundaries of the two top-level frameworks, unify industry cognition, and realize the popularization and recognition of the RE concept in the industry; second, carry out feasible, low-cost and highly adaptable optimization and

correction for the widely used RE models, solve their most prominent problems such as wrong output, knowledge distortion, perfunctory response and model hallucinations, achieve a knowledge output accuracy rate ≥ 95% for the RE model, and reduce the model hallucination rate by ≥ 80%, significantly improving the accuracy and reliability of knowledge dissemination of the RE model.

Through the implementation of the three modular correction schemes, the efficient optimization of the RE model can be achieved without subverting the existing commercial RE model architecture and industrial system. It not only meets the actual market demand of current intelligent applications but also accumulates key technical experience such as knowledge screening, logical verification, self-error correction and closed-loop control in practice, laying a solid theoretical, technical and practical foundation for the research and development of true artificial intelligence in the second and third stages of the "Deification Plan", and is a key first step for the "Deification Plan" to move from theoretical conception to practical implementation.

8. Analysis of Scheme Limitations

Although the modular correction scheme for the RE model proposed in this paper has strong feasibility and optimization effects, limited by the essential kernel of "data reflection" of the RE model itself and focusing on the core goals of the first stage of the "Deification Plan", there are still certain limitations: first, the module can only optimize the output effect of the RE model, cannot fundamentally change its non-autonomous intelligent tool attribute, and cannot transform the RE model into true AI; second, for niche knowledge points in new fields not covered by the authoritative knowledge base, the interception and verification effects of the module will be reduced, and the multi-national integrated authoritative information source library needs to be continuously updated and expanded in the later stage to gradually cover more subdivided fields; third, the module operation relies on the integrity of the authoritative information source library, and if there is a delay in the update of the information source library, it may affect the correction efficiency. The real-time update and automatic verification mechanism of the information source library will be

continuously optimized in the later stage; fourth, for extremely complex cross-field logical reasoning problems, the logical verification ability of the module still needs to be continuously optimized combined with more complex reasoning rules; fifth, the verification accuracy of high-risk industry scenarios depends on the quality of information sources, and the construction progress of an independent multi-national information source library needs to be accelerated; sixth, the current module is only designed for text-based RE models and does not cover multimodal RE models such as images, audios and videos. Such multimodal models also have problems such as content falsification, factual distortion and logical fallacies, with higher propagation risks and greater verification difficulties. Limited by the research focus of the first stage, adaptive design has not been carried out temporarily. The subsequent series of papers will focus on supplementing and in-depth researching the correction schemes, special algorithms and the construction of authoritative multimodal information source libraries for multimodal RE models.

In the later stage, aiming at the above limitations, we will focus on improving the construction of a multi-national integrated authoritative information source library, real-time update mechanism and cross-field complex logical reasoning verification rules, further improving the adaptability and optimization effect of the module, and laying a technical foundation for the transition to the true AI framework.

9. Conclusion

This paper systematically defines and analyzes the core characteristics of the RE model, clarifies its positioning as a top-level framework parallel to true artificial intelligence, completes the adaptation of self-developed concepts with the existing artificial intelligence academic system, clarifies the name connotation, operating mechanism, technical attributes and application scope, and supplements and improves the definition of conceptual innovation boundaries. It distinguishes the differentiated quantitative criteria for the preliminary research stage and the final complete version of true AI, draws a clear boundary between the two top-level frameworks of RE and true artificial intelligence from the essential level, corrects the wrong cognition in the

current artificial intelligence industry that equates RE subdivided tools with true AI, and consolidates the theoretical foundation for the advancement of the "Deification Plan". Combined with the current research status in the field, this paper deeply explores the internal causal relationship and underlying roots of the three core drawbacks of the RE model, and designs three lightweight embedded modules: input port knowledge interception and screening, retrieval-end logical rigor and output-terminal final verification and correction. Combined with the current status of AI technology, it determines the algorithm selection and supporting technologies, formulates a complete update mechanism of the authoritative information source library, module conflict handling rules and industry scenario differentiated parameters, and builds a full-link closed-loop correction system. Without reconstructing the existing model architecture, the knowledge accuracy and logical rigor of the RE model can be improved, and model hallucinations can be effectively alleviated, making up for the engineering shortcomings of pure theoretical conception.

Meanwhile, the analysis of scheme limitations is supplemented to strengthen the objectivity and rigor of the research, and the quantitative goals and core tasks of the first stage of the "Deification Plan" are clarified. All the content of this research serves this stage. Through the conceptual clarification at the theoretical level and the correction scheme design at the practical level, it not only solves the application pain points of the existing commercial RE models but also accumulates core technical and practical experience for the subsequent research and development of true artificial intelligence, providing a clear direction and solid support for the orderly advancement of the subsequent stages of the "Deification Plan", and helping to achieve the research and development goal of true artificial intelligence.

HAO Ziqian

Independent Researcher

ORCID: https://orcid.org/0009-0002-1835-1362

Contact Email:chaos_endless@163.com

Appendix A:

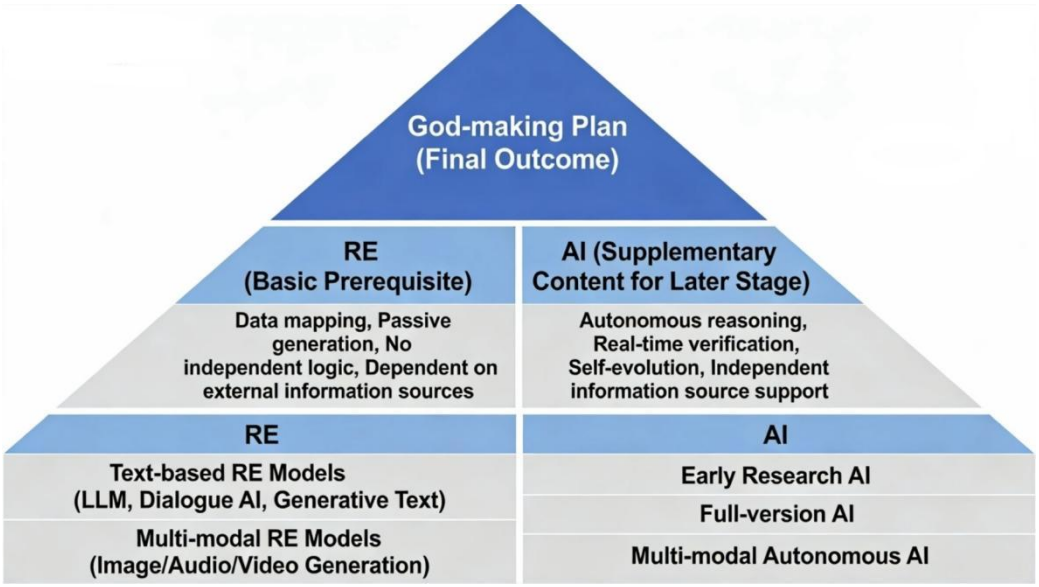

Figure A-1 Pyramid Comparison of RE Model and True AI Model System