

Peer Review

# Review of: "Chain of Ideas: Revolutionizing Research via Novel Idea Development with LLM Agents"

Jerónimo Paiva<sup>1</sup>

1. Universidade do Porto, Portugal

*Peer Review: Chain of Ideas: Revolutionizing Research via Novel Idea Development with LLM Agents*

## Summary:

This paper introduces the *Chain-of-Ideas (CoI)* agent—a novel framework for enhancing the ideation capabilities of large language models (LLMs) by mimicking human research practices through a structured literature chain. The authors also propose *Idea Arena*, a tournament-style evaluation mechanism for assessing idea quality. While the general direction is promising and relevant, especially amid the growing integration of LLMs in scientific workflows, several critical issues in writing quality, conceptual clarity, and presentation standards need significant improvement.

## ◆ 1. Writing Quality and Clarity – Needs Substantial Improvement

Despite addressing an important topic, the paper is weakened by awkward phrasing, redundant expressions, and imprecise grammar throughout.

### *Examples:*

**Redundancy:** “Effective research ideation is a critical step for scientific research.” This is tautological. Ideation is inherently part of research; restate more concisely.

**Unidiomatic Phrasing:** “Facilitates LLMs to capture” is non-native. The correct form is “enables LLMs to understand” or “helps LLMs capture.”

**Vague Modifiers:** Terms like “trivially prompt,” “novel idea,” and “substantial potential” are vague. Specify how existing prompting is insufficient and in what way novelty is achieved.

This paper would benefit from rigorous copyediting by a native or professional academic editor. If this were submitted to a top-tier conference like NeurIPS or ACL, such a lack of polish would reflect poorly, regardless of the technical contribution.

## ◆ 2. Conceptual Framing – Strong Vision, but Execution Lags

The *CoI Agent* idea is conceptually appealing, especially the emphasis on progressive reasoning inspired by human cognition. However, the paper struggles to clearly articulate:

**What exactly distinguishes CoI from prior works like RAG or ResearchAgent beyond structural organization?**

**How chains are selected and pruned** (e.g., what makes a paper “important”? Why 5 in length and 3 in width? Are these hyperparameters?).

**What constitutes a “trend”** in the chain and how it’s inferred—is this done through LLM prompting, or pre-defined heuristics?

Much of this is only described vaguely or is deferred to appendices. A clear, self-contained overview should be included in the main text.

## ◆ 3. Abstract and Introduction – Redundant and Wordy

There is significant overlap between the abstract and the first half of the introduction. This repetition wastes space. Additionally:

The abstract includes inflated language (“revolutionizing research,” “enhancing ideation capabilities”) that is not well-supported by empirical evidence.

The introduction lacks a sharp thesis. It meanders between motivations, citations, and examples without firmly stating *how* CoI will solve the limitations of existing systems in a technically distinct way.

As a high-expectation reader, I expect the introduction to convey:

What is broken in current methods.

What your method is.

Why it is better, in one sharp paragraph.

## ◆ 4. Technical Rigor – Generally Sound, But Needs More Transparency

The paper introduces the *Idea Arena* evaluation framework, which is quite innovative. However:

There is **insufficient detail on how trends are predicted** or how “novelty” is quantified beyond embedding similarity.

The **human evaluation** is described briefly, but the methodology (e.g., inter-rater reliability, criteria enforcement) needs clarification.

The **cost efficiency claim** (“\$0.50 per idea”) is an attractive selling point, but highly sensitive to settings (LLM API version, prompt size, etc.). Details must be disclosed for reproducibility.

## ◆ 5. Language Tone – Overly Promotional

The manuscript sometimes crosses the line from confident to promotional. For example:

“Our agent is capable of generating ideas comparable to human creativity.”

This is a strong and potentially controversial claim that needs to be qualified with evidence. Use more modest, objective phrasing unless your results are overwhelmingly conclusive (which they are not).

## ◆ 6. Visuals and Structure – Mixed

Figures (e.g., Figure 2) are helpful but cluttered. The captions lack sufficient explanation. In a high-pressure review setting, unclear visuals hurt the reader’s ability to grasp your system quickly. Simplify and annotate better.

Also, your sectioning could benefit from stricter discipline. Too many subsections without clean transitions make the reading fragmented. Ensure each section answers: *what is the purpose here, and what is the take-away?*

## ✓ Strengths

Despite its shortcomings, the paper has commendable qualities:

Addresses an **important and underexplored problem**: structured ideation via LLMs.

Introduces a **practical and novel framework (CoI)** rooted in emulating human cognition.

Proposes a thoughtful evaluation method (*Idea Arena*) with human and LLM judges.

Empirical results are **competent and reasonably strong**, with good baselines.



## **Final Verdict: Borderline Accept — Major Revisions Required**

**Recommendation:** *Revise and Resubmit with Major Edits*

This paper has significant potential but falls short of the standards expected at top venues due to issues in presentation, clarity, and writing discipline. I strongly encourage the authors to:

Rewrite the abstract and introduction for conciseness and clarity.

Eliminate vague and promotional language.

Improve figure clarity and explanatory flow.

Conduct a full linguistic revision for fluency and precision.

Expand technical details and cleanly differentiate contributions from prior work.

If these revisions are made, this could become a valuable and citable contribution in the LLM research tools space.

### **Declarations**

**Potential competing interests:** No potential competing interests to declare.