LLM financial report drafting: correct numbers, unsupported judgments

27 November 2026 · 5 min read · blogs.iictp.org/llm-financial-report-writing-unsupported-claims

Short answer: LLM financial report drafting can state every number correctly and still add judgments and causes that nobody supplied, according to a paper by Seunghan Lee and colleagues posted on arXiv on 20 November 2026. They call the additions unsupported framing and propose FrameCheck, a rule-based layer between the model and human review. It matters because a reconciled figure inside an invented explanation can still mislead readers.

TL;DR The short version

- A paper posted on 20 November 2026 shows that LLMs drafting financial reports can state every number correctly yet add judgments and causes nobody supplied.
- The authors call this unsupported framing and propose FrameCheck, a rule-based layer between the model and human review.
- A reconciled figure inside an invented explanation still misleads, so review the narrative, not only the numbers.

Read online, listen to it, or change the text size: https://blogs.iictp.org/llm-financial-report-writing-unsupported-claims

What is unsupported framing?

It is wording that goes beyond the input. If the facts say revenue increased 3% on higher prices, a draft that says revenue surged and adds strong demand as a cause has introduced both a judgment and an explanation. A number check would pass it.

How FrameCheck works

1. Guide the draft with professional examples that report a change of similar size.
2. Flag judgment words that professional writers would not use for a change of that magnitude.
3. Flag causes that the input does not give.
4. Remove flagged wording by rule, with no extra model call, before human review.

What is new compared with prompting

The authors report that prompting the model with how each word is normally used made such judgments more frequent. FrameCheck instead uses norms taken from annual 10-K filings, the reports that US public companies file with the securities regulator, and enforces them outside the model.

Results reported by the authors

- Experiments covered 17 open models and three million drafts.
- Unsupported framing was far more frequent in model drafts than in text written by professionals.
- FrameCheck removed almost all of it and brought the rate of judgment words close to that of 10-K writers.
- When notes supplying causes were provided, FrameCheck kept almost all the causes that came from them.
- Code is available at github.com/seunghan96/framecheck.

This is an arXiv preprint or workshop paper. Treat its numbers as the authors' own results, not as independently confirmed findings.

Limitations and open questions

- The norms describe how 10-K writers use words, not how readers understand them. Words such as material also carry legal meaning.
- The cause check relies on a list of connecting words and misses causes phrased other ways.
- Readers and reference labels were produced by language models. A study with finance professionals is future work.
- Norms come from English US filings, and other documents or markets may differ.
- It assumes one change per passage, which can cause false flags in summaries.
- Closed models were not evaluated.

Sources were read on 27 November 2026. Products, rules and data change, so check the linked originals before you rely on any detail.

How to apply it in a close or reporting cycle

1. Give the model the figures and the documented explanations together.
2. Require that every causal phrase points to a supplied note.
3. Keep a list of intensity words that need a stated threshold before use.
4. Treat a removed or unsupported cause as a task: the preparer must supply the real one.
5. Reconcile numbers separately; wording checks do not verify arithmetic.
6. Log drafts, edits and approvals for review.

What it means for accountants and auditors

Management commentary is expected to explain material changes, so deleting an invented cause leaves a gap that a person has to fill with evidence. Review questions should shift from whether the number is right to whether each adjective and each because has support. A related approach for assurance work is auditor-in-the-loop report review.

Frequently asked questions

- **What is the risk in LLM financial report drafting?** Added judgments and causes that the source data does not support.
- **What is FrameCheck?** A rule-based layer that flags and removes such wording before review.

- **Does it check the numbers?** The paper addresses framing, so keep separate reconciliation.
- **Has it been tested with accountants?** Not yet, according to the authors.

Related reading

- SARA for sustainability assurance
- Small models reading SEC filings with evidence
- Confidence in document extraction
- AI Practitioner certification

> Related credential: AI Practitioner. Explore the exam facts and syllabus.

Sources and further reading

1. Lee, Seo, Kang, Lee, Han, Yoo, Kim, Lim and colleagues. Right Numbers, Wrong Story: When LLMs Narrate Financial Changes Beyond Their Inputs. arXiv 2611.26968. 20 November 2026 (preprint). arxiv.org/abs/2611.26968

Published by IICTP, an independent technical credential body, on 27 November 2026. The current version of this post is at https://blogs.iictp.org/llm-financial-report-writing-unsupported-claims. More guides: blogs.iictp.org.