

Peer Review

Review of: "Creating a Biomedical Knowledge Base by Addressing GPT's Inaccurate Responses and Benchmarking Context"

Tiago Lubiana¹

1. Independent researcher

Interesting article! And thank you for sharing the article so openly. I may be harsh in wording below, but I do believe there are good things around there. **My general suggestion is a general rewrite – the content may be good, but it is hidden in the format.**

I think the project has interesting ideas in this super hot field of RAG-based systems, but I must admit I found it not so pleasant a reading – and I could not even find the “knowledge base” that the article is supposed to describe.

It follows the classical this-is-how-an-article-should-look-like recipe with a big list of references that seem a bit artificial and not read profoundly (see e.g. [this reference](#)). IMHO it clutters the article a bit, making it hard to tease out the contributions. Also, we are all a bit tired of this cookiecutter template for papers – at least I am!

The methods and introduction are mixed, and not in a fun way. It explains about RAG and then jumps to “*Once we had a responsive RAG with context and references, we wanted to benchmark how well that system worked,*” so I end up a bit lost.

Explain what the Gene Network is, explain what you wanted to do (perhaps enjoy the freedom here to do some storytelling, explain who are you and what you *really* want out of this project, and what you find interesting/beautiful). Explain what GNQA means – is it an acronym?

Then tell *exactly* how you selected the corpora for the RAG, the tools and filters used and how you processed it, give the reader some insights on your decision making process – that is important.

“The aim was a GPT question-answer system for the users of the GN web-service that provides contex” – this is buried in the text!

Why go super abstract in a title *Creating a Biomedical Knowledge Base by Addressing GPT's Inaccurate Responses and Benchmarking Context* when the goal is – for what I could tell – describe how you added a GPT + RAG feature to the GN service.

Anyways, here are some notes:

1 - The tool at <https://genenetwork.fahamuai.com> could be testable without the need for login. This improves usability and is even a requirement for some journals (e.g. Application Notes in Bioinformatics)

2 - I could not really find a link to the “GNQA knowledge base”. As someone that comes from a biocuration background, I find it quite important to look at the resource itself — after all, the database is *the* (or one major) contribution of the work. For what I can tell there is no structured *knowledge base*, just a knowledge handling, LLM-based, RAG-powered system that helps a knowledge base, <https://genenetwork.org/>.

3 - The repository is interesting (<https://git.genenetwork.org/gn-ai/tree/README.md>) but I found it hard to navigate a bit. Maybe having a mirror in GitLab, GitHub or similar could improve the experience for those interested on the code.

4 - Retrieval Augmented Generation ASessment (RAGAS) is an automated tool that can evaluate RAGs[Shahul:2023]. — The reference here is broken.

5 - Figure numbers are all mixed. That makes it harder to read and follow through.

6 - It seems like too many things are presented here, and I cannot super understand what you are aiming at. **Who are you writing to?** Is it a service for biomedical researches that you want to share? Are you developing methods to asses RAG and auto generate questions? There are many stories, but they are not tied in a single narrative.

Well, I will stop here and I do apologize for the criticism – I hope it is, though, constructive.

I think we all live in a post-publication crisis distopy where articles are written to be published, but not to be read. This piece seems to be one more victim of the system.

Maybe split it in 2: one short, blogpost like story, with people, personality and details, and a second one, going through the details. But make it is a text that it will be a pleasant reading.

Declarations

Potential competing interests: No potential competing interests to declare.