THE NEW LAB ASSISTANT: UTILIZING MULTI-AGENT AI ARCHITECTURES IN MEDICAL RESEARCH

Alim Aliyev
Simon Kuznets Kharkiv National University of Economics
Scientific Advisers – Tetiana Lysenkova; Olena Aliyeva, Associate Professor, ZSMPU
Language Adviser – Tetiana Lysenkova, Senior Lecturer

In modern scientific research, the primary bottleneck is no longer a lack of data, but an overwhelming abundance of it. Navigating complex literature in specialized fields requires AI to evolve from a passive search tool into an agentic research partner [1]. The aim of this study was to investigate the use of Multi-Agent Systems (MAS) to optimize literature review, data extraction, and statistical validation in medical research, as well as to support *in silico* approaches.

Current methodologies demonstrate that relying on a "one-size-fits-all" AI model is insufficient for rigorous scientific inquiry. Instead, researchers employ a "triangulation" approach, orchestrating specific models for their unique architectural strengths. Broad synthesizers such as ChatGPT are effective for rapid data mapping and identifying cross-disciplinary trends, while deep readers such as Claude are well suited for critical audits of multiple full-text papers simultaneously, identifying methodological nuances that metadata searches often miss. Evidence-oriented tools such as Perplexity and NotebookLM strengthen this process by grounding responses in retrieved or uploaded sources, helping filter out clinically irrelevant data or animal-only trials. Such triangulation improves both the speed and the reliability of literature-based medical research.

An advanced form of AI-assisted research involves orchestrating specialized models into a unified Multi-Agent System that functions as a virtual laboratory team [2]. In our practical testing, several AI tools were given the same task: to identify relevant literature on a medical problem - specifically, the search for pharmacological agents for neuroprotection in cerebral ischemia/hypoxia - to prepare a brief review and provide an analytical synthesis of the findings. The comparison included ChatGPT, Claude, Perplexity, and NotebookLM. The results showed that ChatGPT and Claude were stronger in synthesis and structured interpretation, while Perplexity and NotebookLM were especially useful for source-based retrieval and transparent evidence tracking. A typical MAS workflow begins with a “Scout” agent scanning databases such as PubMed for recent peer-reviewed studies. This data is then evaluated by a “Critic” agent, which analyzes methodologies and flags studies with low statistical power. Concurrently, a “Bio-Statistician” agent reviews effect sizes and confidence intervals, marking weak findings as low-confidence evidence [3]. Finally, an “Orchestrator” model synthesizes these outputs into a cohesive evidence map, highlighting the compounds with the most robust support.

AI also expands the role of *in silico* research in medicine. By simulating mechanisms, screening candidate compounds, and narrowing the range of plausible hypotheses before wet-lab testing, AI helps reduce the volume of unnecessary laboratory experiments and makes the research process more efficient.

Conclusions.

The integration of MAS transforms the modern researcher from a traditional data gatherer into a “Strategic Auditor.” By delegating knowledge extraction - the what and the how much - to an AI team, the scientist can focus on interpretation - the why and the what next. Ultimately, the researcher’s new primary objective is to verify the logic behind AI-synthesized conclusions before advancing to physical validation or practical application [4].

REFERENCES

1. Wu, S., et al. (2023). «AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation Framework.» arXiv preprint arXiv:2308.08155. URL: https://arxiv.org/abs/2308.08155
2. Boiko, D. A., MacKnight, R., & Gomes, G. (2023). «Emergent autonomous scientific research capabilities of large language models.» Nature, 624(7991), 127-134. URL: https://www.nature.com/articles/s41586-023-06792-0

3. Viceconti, M., et al. (2021). «In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products.» Methods, 185, 120-127. URL: https://doi.org/10.1016/j.ymeth.2020.01.011

Ekins, S., et al. (2019). «Exploiting machine learning for end-to-end drug discovery and development.» Nature Materials, 18(5), 435-441. URL: https://doi.org/10.1038/s41563-019-0338-z