

# Agentic AI for Scientific Research

Olatomiwa O. Bifarin

School of Chemistry and Biochemistry  
Georgia Institute of Technology

August 15, 2024

## Abstract

**Note:** This manuscript is very much a work in progress, with currently mostly outlines. Many of the papers I will be reviewing here, I have reviewed on my AI blog <https://theepsilon.substack.com/>. This paper will allow me to congeal my readings and thoughts in one single document. The actual abstract will go here.

## 1 Introduction

- Briefly introduce the increasing role of AI in science, highlighting the potential of LLMs to accelerate discovery and innovation.
- State the purpose of the review: To provide a comprehensive overview of current research on LLMs and agentic AI in science, focusing on their capabilities, challenges, and opportunities.
- Briefly outline the three main sections of the review: AI/LLM Basics and Techniques, AI Agents for Scientific Research, and Fine-tuning LLMs for Domain Expertise Tasks.

## 2 AI/LLM Basics and Techniques

### 2.1 Foundation Models and LLMs

- Define LLMs and their significance in natural language processing.
- Explain the Transformer architecture, highlighting self-attention mechanisms and the ability to process long sequential data.
- Briefly mention the different scales of LLMs, such as GPT-3, GPT-4, PaLM, and LLaMA.

## 2.2 Pretraining and Fine-Tuning

- Explain the process of pretraining LLMs on massive text datasets to acquire general knowledge.
- Describe how fine-tuning adapts pretrained LLMs to specific scientific tasks using smaller, labeled datasets.

## 2.3 Types of Transformer Architectures

- **Encoder-Only Models (e.g., BERT):** Explain their use for property prediction, classification, and sentiment analysis in chemistry (e.g., ChemBERTa, MatSciBERT, Mol-BERT).
- **Decoder-Only Models (e.g., GPT):** Describe their capabilities for molecule generation, de novo design, and property-directed inverse design in chemistry (e.g., MolGPT, Adilov2021, cMolGPT, iupacGPT).
- **Encoder-Decoder Models (e.g., T5, BART):** Discuss their applications in synthesis prediction and retrosynthesis planning in chemistry (e.g., ReactionT5, Text+Chem T5, MolT5, T5Chem, Chemformer, Molecular Transformer).

## 2.4 Key Techniques

- **Prompt Engineering:** Explain the importance of carefully crafted prompts for guiding LLMs in scientific research. Briefly mention various prompting strategies, such as few-shot learning, Chain-of-Thought prompting, and prompt engineering for inverse design (Jablonka et al., 2024).
- **Retrieval-Augmented Generation (RAG):** Describe how RAG combines LLM generation with information retrieval to improve accuracy and ground responses in factual data (e.g., PaperQA, WikiCrow, Chemist-X).
- **Explainable AI (XAI):** Discuss techniques like SHAP and LIME for providing insights into the decision-making process of black-box models, enhancing interpretability, and building trust in AI outputs for scientific applications (e.g., XpertAI, SciAgent).

# 3 AI Agents for Scientific Research

## 3.1 What are AI Agents?

- Define AI agents as autonomous systems that can perceive, reason, decide, and act within a specific environment to achieve goals.
- Differentiate between LLM-based agents and traditional AI agents, highlighting the strengths of LLMs in natural language understanding and reasoning.

## 3.2 Key Modules of AI Agents

- Discuss the four primary modules: Perception, Memory, Reasoning, and Action. Provide examples of their implementation in scientific agents (e.g., ChemCrow’s use of perception modules for multi-modal data, PaperQA’s memory for retaining relevant information, STORM’s reasoning for generating questions, and Coscientist’s use of robotic actions).

## 3.3 Agent Architectures

- **Single-Agent Systems:** Discuss examples of LLM-based agents designed for specific tasks in chemistry, such as ChemCrow, Chemist-X, and Coscientist. Analyze their strengths and limitations, focusing on factors like task specificity, automation levels, and reliance on human feedback.
- **Multi-Agent Systems:** Introduce the concept of multi-agent systems, where multiple agents with specialized roles collaborate to solve complex tasks (e.g., ProtAgent, TAIS). Discuss different collaboration schemes, such as expert consultation, brainstorming, and roundtable discussions, highlighting their potential for interdisciplinary research.

## 3.4 Applications in Science

- **Drug Discovery:** Discuss how AI agents are automating and accelerating drug discovery, from target identification and lead optimization to synthesis planning and clinical trial design (e.g., ChemCrow, cMolGPT).
- **Materials Science:** Explore how AI agents contribute to materials design, from candidate material generation to property prediction and synthesis planning (e.g., ChatMOF, CALMS).
- **Bioinformatics:** Illustrate the role of AI agents in bioinformatics analyses, from gene expression data processing and variant effect prediction to protein structure prediction and drug-target interaction identification (e.g., AutoBa, Geneformer).

## 3.5 Challenges and Opportunities

- **Safety and Dual Use Concerns:** Address the risks associated with autonomous scientific agents, especially concerning the potential for misuse and the creation of harmful substances. Highlight the need for safeguards, ethical guidelines, and human oversight (e.g., ChemCrow’s safety checks, TAIS’s emphasis on human validation).
- **Explainability and Trust:** Underscore the importance of transparency and interpretability for building trust in AI agents’ outputs, facilitating collaboration between scientists and AI (e.g., XpertAI’s use of NLEs, WikiCrow’s focus on verifiable citations).

- **Generalization and Adaptability:** Discuss the challenges of creating agents that can generalize to new tasks and adapt to evolving data and scientific understanding.

## 4 Fine-tuning LLMs for Domain Expertise Tasks

### 4.1 Why Fine-Tuning?

- Explain the need for domain-specific expertise in LLMs for scientific research, exceeding what can be obtained through prompting alone.
- Discuss the benefits of fine-tuning for improving accuracy, efficiency, and applicability of LLMs on specialized tasks.

### 4.2 Fine-Tuning Strategies

- **Instruction Tuning:** Describe how models are trained to follow instructions and complete specific tasks (e.g., Med-PaLM 2's use of instruction tuning on biomedical texts).
- **Reinforcement Learning with Human Feedback (RLHF):** Explain how RLHF aligns models with human preferences and values, improving their performance on tasks that require subjective judgment (e.g., scientific feedback generation, as in Liang et al., 2023).
- **Domain-Specific Pretraining:** Discuss how pretraining on specialized datasets can enhance the LLM's understanding of a specific scientific domain (e.g., ChemBERTa's pretraining on chemical data, BioMegatron's pretraining on biomedical texts).

### 4.3 Applications in Fine-Tuning

- **Optimizing Chemical Reactions:** Show how fine-tuned LLMs can be used to predict reaction yields and optimize reaction conditions (e.g., MolXPT, T5Chem).
- **Identifying Disease-Related Genes:** Illustrate how fine-tuned LLMs can analyze gene expression data to identify disease-predictive genes (e.g., Med-PaLM 2).
- **Improving Scientific Writing:** Discuss how fine-tuned LLMs can assist with scientific paper writing, from generating summaries and abstracts to providing feedback and suggesting improvements (e.g., ScholarBERT, STORM).

#### 4.4 Challenges and Future Directions

- **Data Scarcity:** Address the challenge of obtaining sufficient high-quality labeled data for fine-tuning specialized LLMs in science.
- **Overfitting and Generalization:** Discuss the risks of overfitting to specific datasets and the need for robust validation techniques to ensure generalizability of fine-tuned models.
- **Interpretability and Explainability:** Emphasize the importance of maintaining transparency and understanding the reasoning process of fine-tuned LLMs.

### 5 Conclusion

- Summarize the key findings of your review, emphasizing the transformative potential of LLMs and agentic AI for scientific discovery and innovation.
- Highlight the challenges that remain to be addressed, particularly concerning safety, explainability, and evaluation.
- Offer a perspective on the future of this rapidly evolving field, advocating for responsible development and collaboration between AI researchers and scientists to fully unlock the potential of LLMs in science.

### Acknowledgements

Acknowledgements will go here.