

# Ultrahigh Performance Composite Material: Agentic AI for Data-Driven Formulation

Sofia Dahlgren<sup>1</sup>, Vera Kotova<sup>1</sup>, Helen Feldhaus<sup>1</sup>, Arjun Gupta<sup>1</sup>, Emmett Butcher<sup>1</sup>, Navid Zobeiry<sup>2</sup>, Felix Nguyen<sup>3</sup>

<sup>1</sup> Department of Chemical Engineering, University of Washington, <sup>2</sup> Department of Material Science and Engineering, University of Washington, <sup>3</sup> Lockheed Martin Corporation

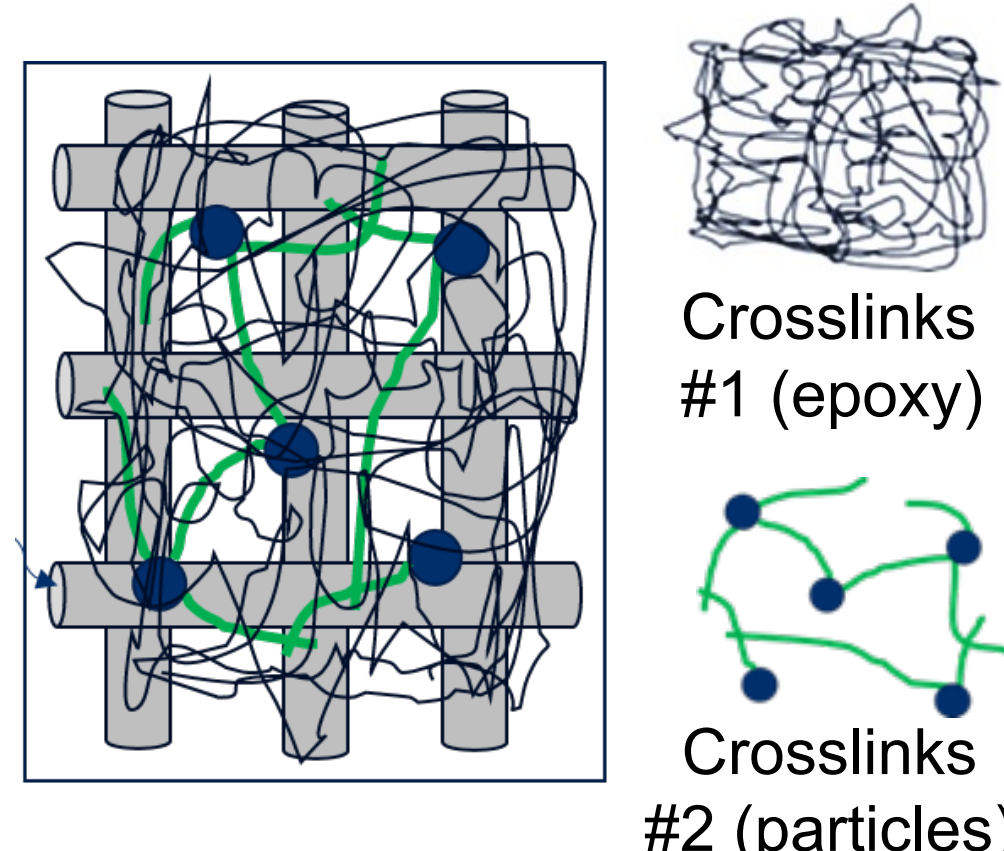
LOCKHEED MARTIN

W CHEMICAL ENGINEERING  
UNIVERSITY of WASHINGTON

## Background

### Motivation

- Aerogel-epoxy nanocomposites are promising next-generation aerospace materials<sup>1</sup>
- Composite optimization is a major challenge due to the high-dimensional design space and cost of experimentation
- Agentic artificial intelligence (AI) is an emerging strategy for accelerated data-driven workflows



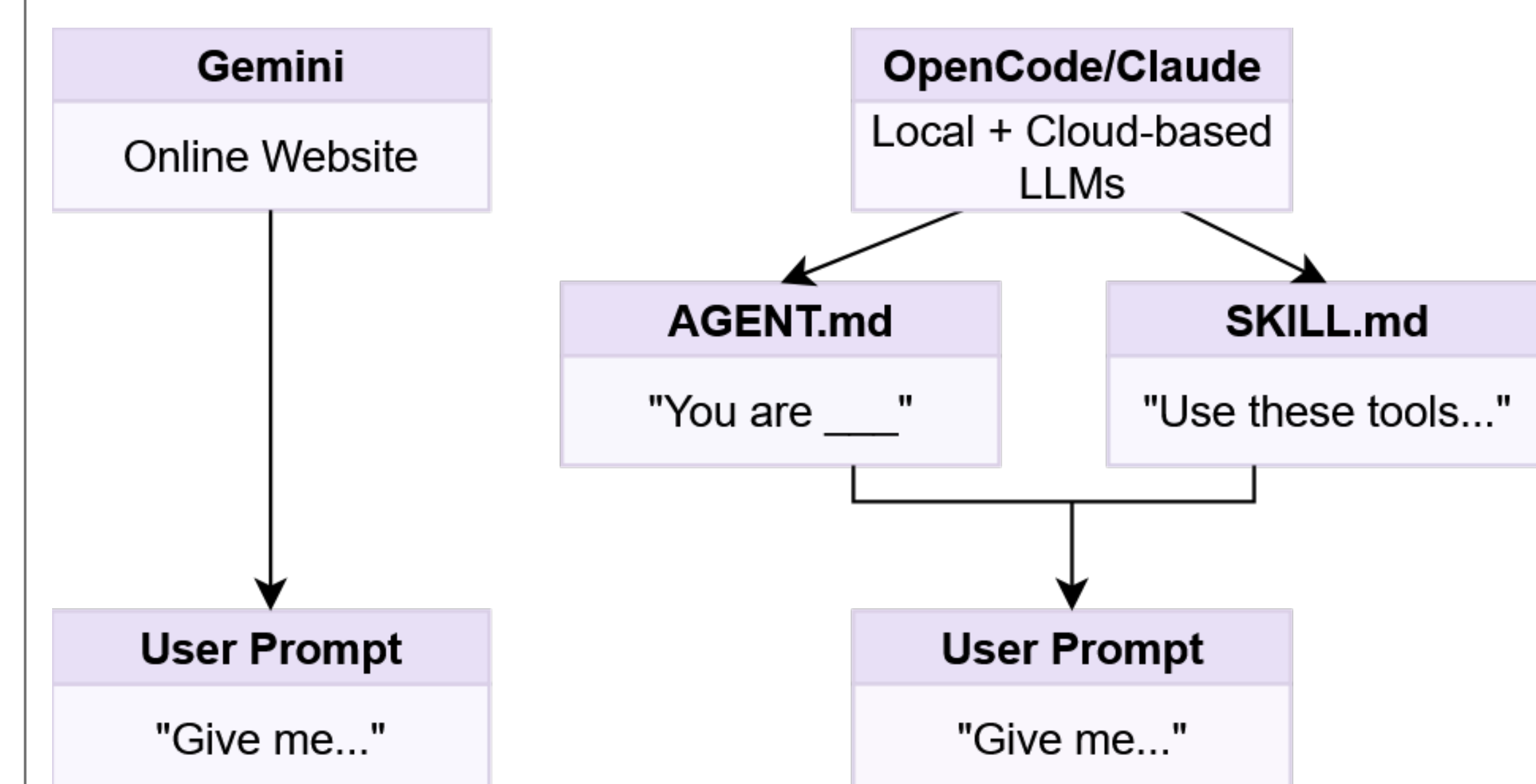
### Objectives

- Automate source discovery from a research literature fingerprint.
- Perform high-throughput data extraction from source PDFs.
- Propagate results to a single data table where materials are ranked according to established Ashby's indices.

## Methods

### Regular Prompting vs. Agentic AI Approach

Agentic AI provides more control over the model response.<sup>2</sup>



### Programming AGENT.md and SKILL.md files

To provide essential model context & specify code executables.

```
1 ---  
2 name: discovery  
3 description: discover novel  
4 literature  
5 argument-hint: "identify similar  
6 literature"  
7 #tools: []  
8 ---  
9 # AGENT.md - Reference-Based  
10 Discovery Agent  
11 ## Persona  
12 ## Responsibilities  
13 ## Extraction Requirements  
14 ## Constraints & Rules
```

> agents  
> skills / pdfreader  
extract\_material\_properties.py  
research\_paper\_analyzer.py

Defines strict rules and ensures a highly reproducible output for every run.

Allows for a continuous update of the recipes database.

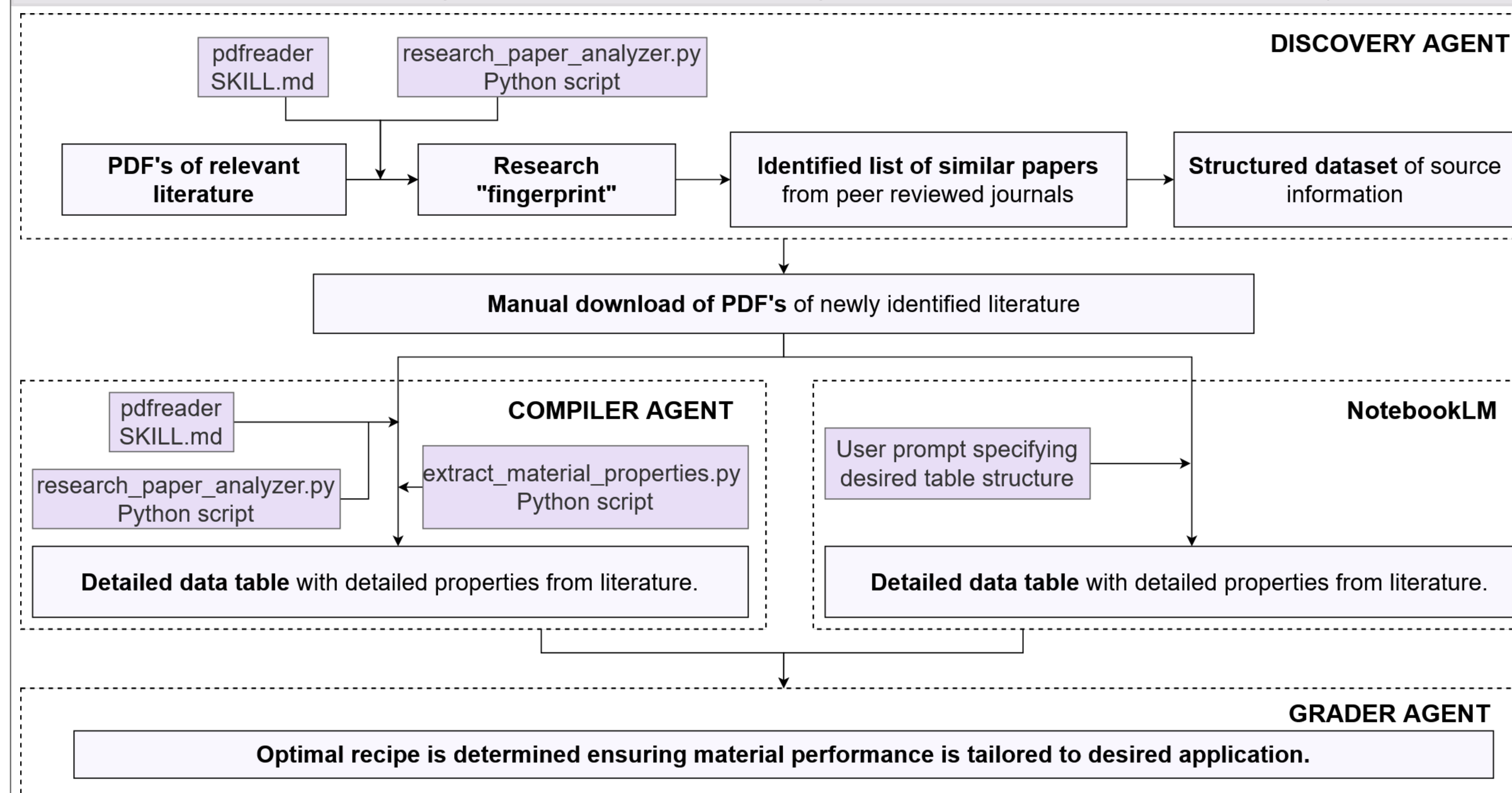
### Reproducibility and Accuracy Verification

To validate workflow performance & identify edge-case uses.

- The end-to-end workflow was independently implemented on 6 machines using Claude Haiku 4.5 and consistent prompting.
- Multiple generated data sets were screened by a human reader and each entry compared to that in the source literature.
- This procedure was performed on data sets extracted from 30-40 independent sources.
- Compiler validation across LLMs was performed on synthetic data containing no missing values (n=3 trials per model).

## Workflow Results

### Overview of Agentic AI Workflow for High Performance Material Discovery

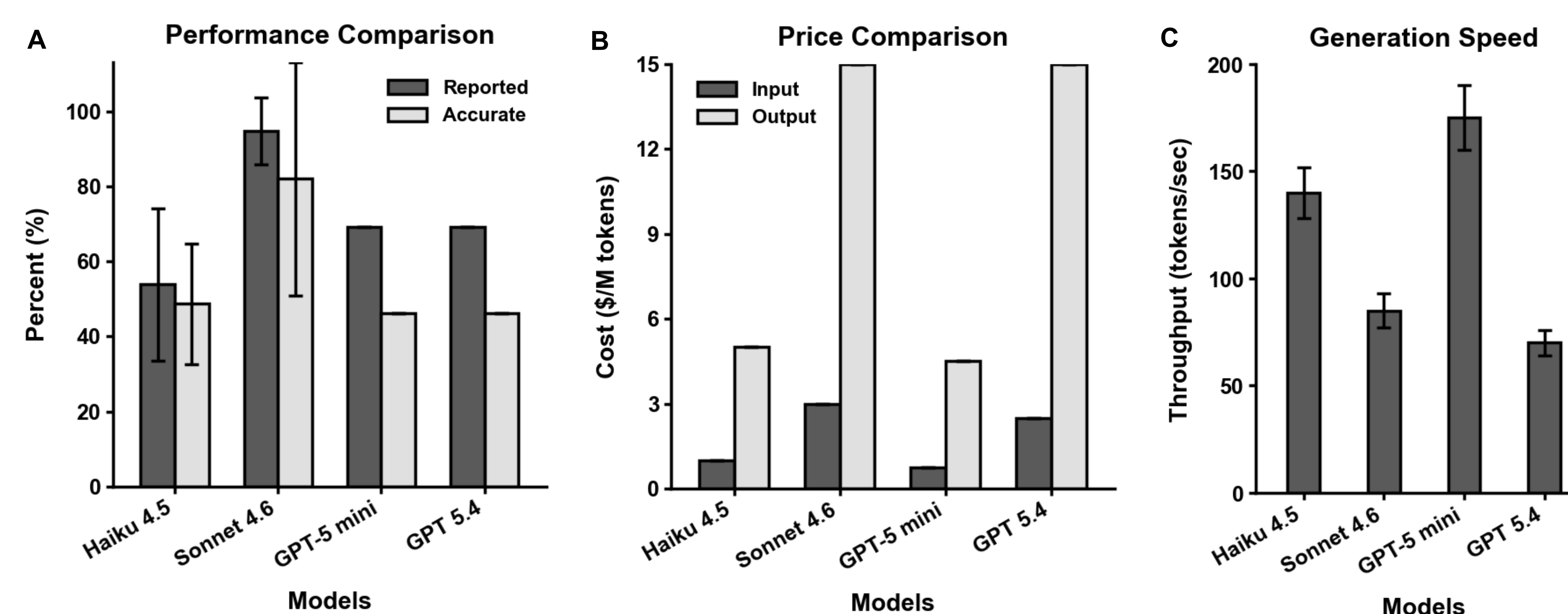


### Discovery Agent vs. Gemini Deep Search Comparison

Discovery Agent	Gemini Deep Search
<ul style="list-style-type: none"><li>Identifies <b>highly relevant peer reviewed literature based on citation chain and research fingerprint</b></li><li>Identifies literature and generates summary in less than <b>5 minutes</b></li><li>Seamless integration into Compiler agent</li></ul>	<ul style="list-style-type: none"><li>Identifies <b>peripherally related sources</b> to generate report</li><li>Results are <b>highly dependent on hyper specific user prompt</b></li><li>Search requires <b>20 minutes</b> per inquiry and a subscription</li></ul>

**Conclusion:** Discovery Agent consistently identifies high quality and relevant peer reviewed literature on silica epoxy composites faster than Gemini Deep Search, with seamless integration into the automated workflow.

### Compiler Agent Performance Run on Cloud-Based LLMs



**Figure 1.** (A) Compiler agent performance for values reported (i.e., extracted from synthetic data PDF) and accurate across different LLMs, n=3 per model. (B) Price per token to run each model as of early 2026.<sup>3,4</sup> (C) Throughput in tokens/second to run each model across major cloud providers as of early 2026.<sup>5</sup> All data reported as mean ± standard deviation.

**Conclusion:** Claude Sonnet 4.6 offers superior Compiler agent performance compared to other LLMs, with tradeoffs for higher cost and lower speed. NotebookLM (data not shown) offers a high-accuracy commercial alternative to the Compiler agent.

## Conclusion

### Agent Performance

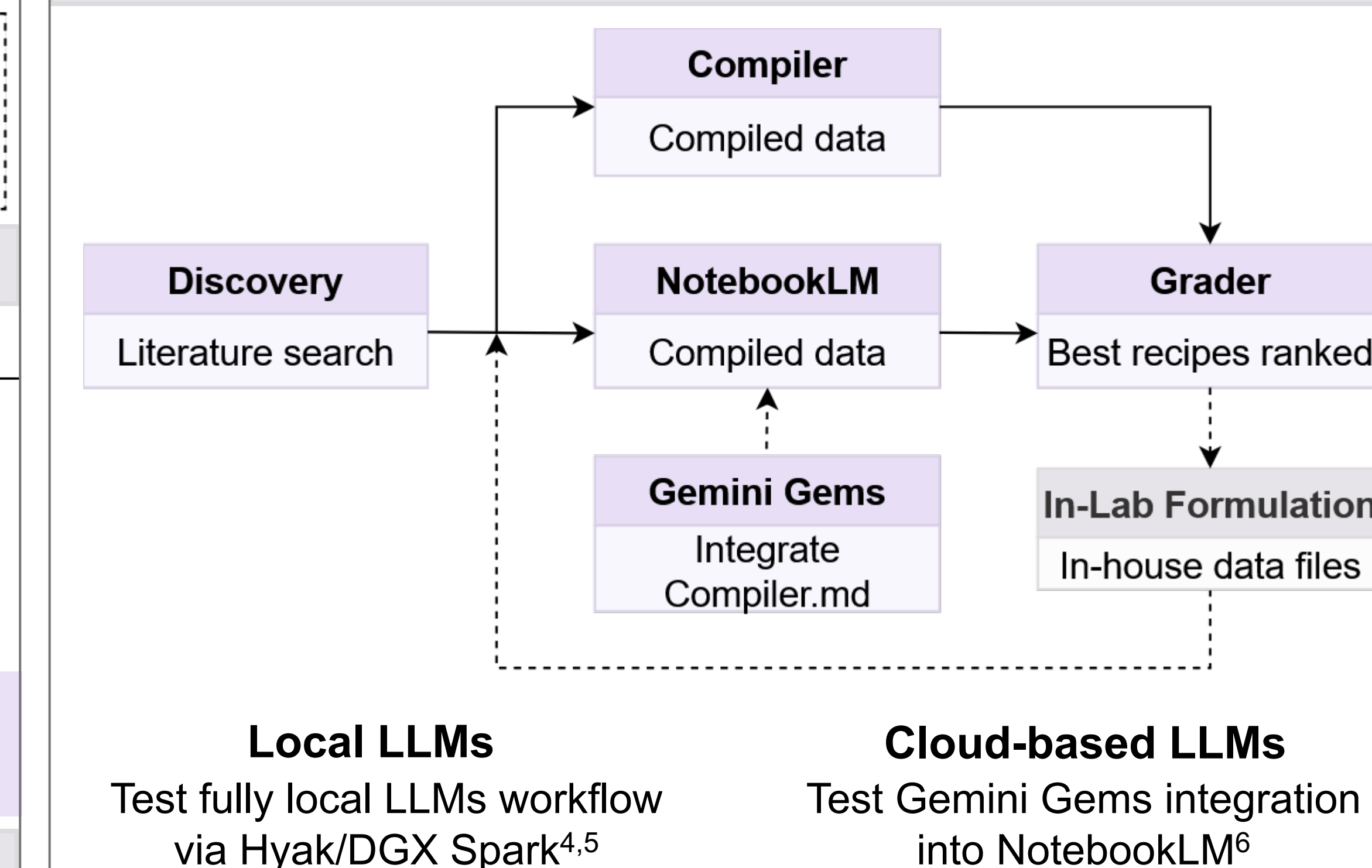
Discovery	Compiler	Grader
Demonstrates enhanced retrieval of unique, reputable sources compared to Gemini Deep Search.	Exhibits model dependency with Claude Sonnet 4.6 offering competitive performance to NotebookLM.	Provides material rankings based on Ashby's indices. Improvements for reproducibility are required.

### Advantages and Applications

- Workflow integrates with different LLM models for scalability
- High autonomy with minimal user prompting required
- Broader impacts in materials formulation via substitution of key words in Discovery.md, enabling research acceleration

## Future Work

### Laboratory Integration & Workflow Refinement



## Acknowledgements

We would like to thank Alex Gray and Paulina Portales from UW Materials Science and Engineering for their composite expertise. We are also grateful for Professor Dave Beck and the eScience Institute for resources and mentorship on agentic AI. We thank Chino Ruttanasupagid for project management support. Lastly, we would like to thank Lockheed Martin for sponsoring this project.

## References

- Takeshita, S. et al. Angewandte Chemie International Edition 2025, 64 (28), e202504250.
- OpenCode. <https://opencode.ai/docs/> (accessed 2026-02-17)
- OpenAI API Pricing. OpenAI. <https://openai.com/api/pricing/> (accessed 2026-05-25)
- Claude API Docs. <https://platform.claude.com/docs/en/about-claude/pricing> (accessed 2026-05-25)
- LLM Leaderboard 2026. Vellum. <https://www.vellum.ai/llm-leaderboard> (accessed 2026-05-25). Y. Hyak resources
- NVIDIA Product Literature. <https://www.nvidia.com/en-us/products/workstations/dgx-spark/> (accessed 2026-05-25).
- Repositories, C. Gemini Gems. Gemini. <https://gemini.google/overview/gems/> (accessed 2026-05-25).