☀️ HSD-AI-Newsletter-2025.09.17

Exploring AI, ML, and Generative Technologies in Heliophysics

A bi-weekly brief on AI, ML, and LLMs for the Heliophysics Science Division

---

📌 Introduction

Artificial Intelligence (AI) is the broad field of building smart machines. Machine Learning (ML) is a subset focused on training models to learn patterns from data. Large Language Models (LLMs) like GPT-4o or Claude Opus are a recent type of generative AI, capable of producing human-like text, code, and analysis. How do these technologies intersect with heliophysics research? Let's explore.

📅 Upcoming Events

🧠 NASA AI Center of Excellence – Monthly Webinar

📅 **Wednesday, October 8, 2025**
🕚 **11:00 AM – 12:00 PM ET**
📍 **Microsoft Teams**
🔁 *Occurs the second Wednesday of every month*

👉 [**Join the Meeting**](#)
📞 Dial-in: `+1 256-715-9946,,850427523#` (US – Huntsville)
🔢 Phone Conference ID: `850 427 523#`
🆔 Meeting ID: `217 293 044 622`
🔑 Passcode: `F4bi3tK9`

🎓 NVIDIA GTC Conference

🗒️ October 27–29, 2025 | 📍 Washington, DC
One of the most important AI events of the year.

- **Training Day:** October 27 (paid courses)
- **Main Conference:** October 28–29
- **Free registration** for those with a `.gov` email
- **NCTS#:** `52889-26`
  🔗 [Register here](#)

📚 Key Terms of the Week

Term	Definition
Fine-tuning	The process of retraining a pre-trained model on a specific dataset to adapt it for a targeted task.
Retrieval-Augmented Generation (RAG)	Combines a language model with external documents to improve accuracy and factual grounding.
Hallucination	When an AI model generates false or misleading information that sounds plausible.
Chain-of-Thought Prompting	A technique where the user asks the model to reason through steps before giving an answer.
Embeddings	Numeric vector representations of words, phrases, or documents that capture semantic meaning.
MoE (Mixture of Experts)	A model architecture that activates only parts ("experts") of a large model per input, saving compute.
Multi-modal AI	AI systems that can process more than one type of data at a time (e.g., text + images or video).
Zero-shot / Few-shot	Refers to how well a model can complete a task with no or very few examples.
Instruction-tuning	Training models to follow natural language instructions more accurately and safely.
Inference Time	The time it takes a model to generate a response after receiving a prompt.

💬 What are tokens?

Tokens determine how much you can input/output at once. For example:

- `"NASA"` → 1 token
- `"Heliophysics2025"` → likely 2–3 tokens

- A sentence like `"Solar flares are energetic."` → ~5–6 tokens

Tokens in Language Models

Solar | flares | are | energetic.

~5-6 tokens

Token A chunk of text (e.g., a word or part of a word)

Models analyze text by breaking it into tokens

NASA | Heliophysics | 2025

🔧 How to talk to an AI: What is Prompt Engineering?

Prompts are the way that you share information and instructions with an AI. They are particularly important when interacting with an LLM or GPT.

There are many potential elements to a prompt, and most are optional. However, it is best to understand each of these elements because they can all contribute to getting a much better result.

Prompt Engineering is refining instructions to improve performance, provided by the user or the developer/provider.

Prompt design and prompt engineering can seem daunting at first. Once you familiarize yourself with the aspects of prompt engineering, however, you'll begin to see that they are similar to the way you provide instructions to humans. Your experience and intuition working with people can help get you started!

Name	Description	Examples
Task	The task, or goal, is the instruction telling the AI what you want it to do. This is usually the one non-optional item in a prompt. The clearer you are about the task, the better the expected result.	"Take this list of terms and give me every page in this document that mentions these terms." / "What are the key points about prompt engineering?" / "Take these bullet points and turn them into professional-sounding text."
Data	This is the input provided to the AI to supplement the information already contained in the model.	Data may not be required for simple questions, but can include files, attachments, or text pasted directly into the chat window.
Format	It is often useful to give instructions on the format that you would like for your output.	"A JSON file with the following headings." / "2-3 paragraphs of text." / "A diagram with the accompanying code used to create the diagram." / "A list of conclusions with citations and source references appended in parentheses after each item."
Role	This provides details to the AI on why and how you want the task to be performed. Role is one of the most common instructions provided to a GPT model.	"Your role is that of a third-grade teacher, providing an accurate but easy to understand summary." / "You are an expert programmer, and the code you write is sophisticated and elegant."
Tone	Especially for the case of text and chat output, this refers to the conversational tone, or how it "sounds" to the reader. Tone helps the user adjust the output for the audience.	"formal" or "casual" language / "friendly" or "professional"
Audience	Sometimes it is best to provide the AI with information on who will be receiving the output.	"Colleagues on a financial planning team" / "A class of third-graders" / "friends in my book club" / "Management reviewing my job application"

Name	Description	Examples
Context	Background information that can help inform the AI on how to do the task.	"We are scientists trying to explain an important results to our peers. The peers are familiar with our research, so they'll understand many of the phenomena involved, but they may need to be briefed on the key phenomena first."
Examples	When teaching an intern how to do analysis, you don't just tell them what to do and send them on their way. You usually sit down with them and show them a couple of examples first. This works well for AIs as well - examples show what a "correct" answer looks like.	"For the first entry, the title would be 'Bananas' and the number of features would be 7." / "In the first case provided, the decision would be no because it lacked a backup plan."
Constraints	It can be very useful to provide specific limits or boundaries on the task, so the AI's result can be more focused on what you require.	"Consider only cases relevant to heliophysics." / Focus only on specific items or exclude certain information
Guidelines	These instruct the model on how to think and behave. Guidelines can be essential if you want to trust the model's actions.	"If you're uncertain, state the uncertainty" / "Acknowledge when multiple valid answers exist and include them in the response" / "Prioritize safety and do not provide advice that could cause physical or mental harm" / "Only use factual sources and avoid opinion-based sources"

There are many ways that prompts can be customized for an AI, but they start to become more intuitive with practice. In general, more information is better. A simplified version of the description above could be "when in doubt, leave it in!"

🗒️ If you give the model clear context, a specific role, and a good format, you'll get *much* better results.

🧮 ChatGSFC Model Overview

Model	Best For	Use Cases
GPT-4o	Versatile, high-quality reasoning	Analysis, coding, synthesis
GPT-4o-mini	Fast & lightweight	Drafts, summaries, quick tasks
Claude 4 Sonnet	Structured thinking & strategy	Long reports, workflows, planning
Claude 4 Thinking	Deep step-by-step reasoning	Synthesis, troubleshooting, discovery
Claude 3.7 Sonnet	Reliable generalist	Brainstorming, technical writing
Claude 3.5 Haiku	Compact creative output	Taglines, poetic summaries
o3-mini	Fast, low-cost	FAQs, short queries
GPT-4.1	Long-context, technical tasks	Code reviews, modeling, full papers
GPT-4.1-nano	Automation & embedded uses	Pipelines, fast response apps

✅ Tip: Start with **Claude-4 or GPT-4.1** for deep work, **GPT-4o-mini** or **Claude 3.7 Sonnet** for daily use.

💡 Tips & Tricks (New Set)

- 🧩 **Break big tasks into smaller ones:** Instead of "write a report," ask the model to generate an outline first, then sections.
- 🎯 **Use role-based prompts:** "You are a solar physicist explaining this to an intern" leads to better domain-appropriate tone.
- ✍️ **Iterate using follow-ups:** Start with a basic response, then refine with clarifying questions or improvements.
- 📄 **Give it your writing style:** Paste a paragraph you wrote and say, "write in this style." Models can mimic tone effectively.
- 🔄 **Use the same prompt with multiple models:** Great for comparing accuracy, tone, and cost.
- 🧠 **Teach the model with examples:** Show "bad" vs. "good" responses—it will often generalize correctly.

⚠️ Points of Caution (New Set)

- 🕵️ **False citations:** Some models make up fake DOIs, authors, or sources—double-check any reference!

- 🛠️ **Overconfidence in technical outputs:** LLMs may generate syntactically correct but scientifically wrong code or math.
- 🔁 **Non-determinism:** The same prompt may yield different outputs—important for reproducibility.
- 📦 **Input/output limits:** Even large models can silently truncate long prompts or responses. Be aware of context limits.
- 📉 **Performance degradation:** Model behavior can change subtly over time (e.g., after backend updates or fine-tunes).
- 🌐 **Public models ≠ secure environments:** Don't assume OpenAI, Claude, or Gemini sessions are confidential unless noted.

🧪 arXiv AI paper selections

1. AIssistant: An Agentic Approach for Human–AI Collaborative Scientific Work on Reviews and Perspectives in Machine Learning

Authors: Sasi Kiran Gaddipati, Farhana Keya, Gollam Rabby, et al.
DOI: https://doi.org/10.48550/arXiv.2509.12282

2. SciGPT: A Large Language Model for Scientific Literature Understanding and Knowledge Discovery

Authors: Fengyu She, Nan Wang, Hongfei Wu, et al.
DOI: https://doi.org/10.48550/arXiv.2509.08032

3. A Survey of Scientific Large Language Models: From Data Foundations to Agent Frontiers

Authors: Ming Hu, Chenglong Ma, Wei Li, et al.
DOI: https://doi.org/10.48550/arXiv.2508.21148

🧰 Resources

Internal

- GSFC AI Center of Excellence
- ChatGSFC Portal
- GSFC Code Assistant

External

- [OpenAI Playground](#)
- [Anthropic Claude](#)
- [Google Gemini](#)
- [HuggingFace Spaces](#)

> *"Think like a physicist. Prototype like a hacker. Document like a scientist."*

✉ Questions or contributions?

c.alex.young@nasa.gov | barbara.j.thompson@nasa.gov | christopher.bard@nasa.gov