

HSD-AI-Newsletter-2025.09.03

Exploring AI, ML, and Generative Technologies in Heliophysics

Introduction: What Is AI, ML, and Generative AI?

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-5 or Claude Opus 4.1 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.

Weekly Highlights

[Thursday, August 28 at 1pm ET](#)

Andrés Muñoz-Jaramillo (SwRI), Madhulika Guhathakurta (NASA HQ), Daniel da Silva (UMBC/671) and the Surya team

Introduction to Surya: A Foundation Model in Heliophysics

This talk introduced *Surya*, a NASA-funded foundation model built with Solar Dynamics Observatory data. Unlike traditional AI tools designed for single tasks, Surya was developed as a large, multi-purpose model that could be fine-tuned for diverse heliophysics applications—from space weather forecasting to data analysis. The presenters shared preliminary results, validation strategies, and their vision for enabling the community to leverage Surya for scientific discovery.

Surya Model Paper: [Model details](#)

| Attribute | Details |
|---------------------|---|
| Name & Origin | "Surya" (Sanskrit for Sun); developed by NASA & IBM + partners |
| Architecture & Size | Spatiotemporal transformer; ~360–366M parameters |
| Training Data | ~9 years (\approx 200 TB) of SDO imagery across AIA & HMI instruments |
| Key Capabilities | Zero-shot forecasting; fine-tunable for flare prediction, segmentation, solar wind, EUV spectral modeling |
| Performance Gains | ~16% better than prior methods; up to 2-hour lead time for flares |

| Attribute | Details |
|-----------------|--|
| Outputs | High-resolution visual forecasts, digital twin simulations |
| Availability | Open source on Hugging Face , GitHub , TerraTorch + dataset (SuryaBench) |
| Broader Impacts | Empowers space-weather defense; sets a template for scientific models |

Key Terms of the Week

| Term | Definition |
|--------------------|--|
| LLM | A Large Language Model trained on massive text datasets. Designed to generate text based on the most probable response to input. |
| Generative AI | AI that creates content—text, code, images, audio, etc. |
| Temperature | Controls randomness in text generation (lower = more predictable). |
| Context Window | How much text the model can consider at once (measured in tokens). |
| Prompt Engineering | The craft of structuring inputs to get the best model output. |
| Token: | A <i>token</i> is a chunk of text—typically a word, part of a word, or even punctuation—that a language model reads and processes. |

ChatGSFC AI Models Overview

| Model | Standard Use | Example Use Cases | Best Pick For |
|---------------------|--|--|---|
| claude-4-sonnet | Strategic analysis, quantitative reasoning, planning | Long-form reports, scenario planning, policy briefs | Data-driven strategy & structured analysis |
| claude-4-thinking | Deep, multi-step reasoning and problem-solving | Research synthesis, theoretical exploration, troubleshooting | Deep analysis & complex reasoning |
| claude-3.7-sonnet | Balanced general-purpose reasoning and writing | Drafting, brainstorming, technical explanations | Everyday balanced assistant |
| claude-3.7-thinking | Detailed, nuanced analysis | Academic reviews, multi-perspective debates, problem decomposition | Careful, step-by-step reasoning |

| Model | Standard Use | Example Use Cases | Best Pick For |
|-----------------------------|---|---|--|
| claude-3.5-sonnet-v2 | Creative + precise communication, enhanced reasoning | Proposals, technical docs, polished presentations | Clear technical + creative communication |
| claude-3.5-sonnet | Reliable all-rounder for professional writing & workflows | Emails, summaries, structured workflows | Professional productivity |
| claude-3.5-haiku | Concise, elegant creative outputs | Poetry, taglines, social posts, compact summaries | Creative brevity |
| o3-mini | Fast, lightweight responses | Routine Q&A, quick clarifications, simple instructions | Speed & cost efficiency |
| o1 | General-purpose efficiency | Notes, summaries, conversational tasks | Everyday reliability |
| gpt-4o | Versatile, advanced comprehension & reasoning | Complex projects, cross-domain analysis, creative problem-solving | Default all-round powerhouse |
| gpt-4o-mini | Faster, lighter 4o | Email drafting, brainstorming, quick summaries | Fast everyday helper |
| gpt-4.1 | Cutting-edge reasoning, long context handling | Research reports, technical analysis, code reviews | Extended reasoning & long-context tasks |
| gpt-4.1-mini | Efficient with moderate complexity | Team workflows, recurring reports, mid-level content | Balanced efficiency |
| gpt-4.1-nano | Maximum speed & efficiency | Automations, embedded apps, lightweight pipelines | Embedded use & automation |

Model Decision Matrix by Task Type

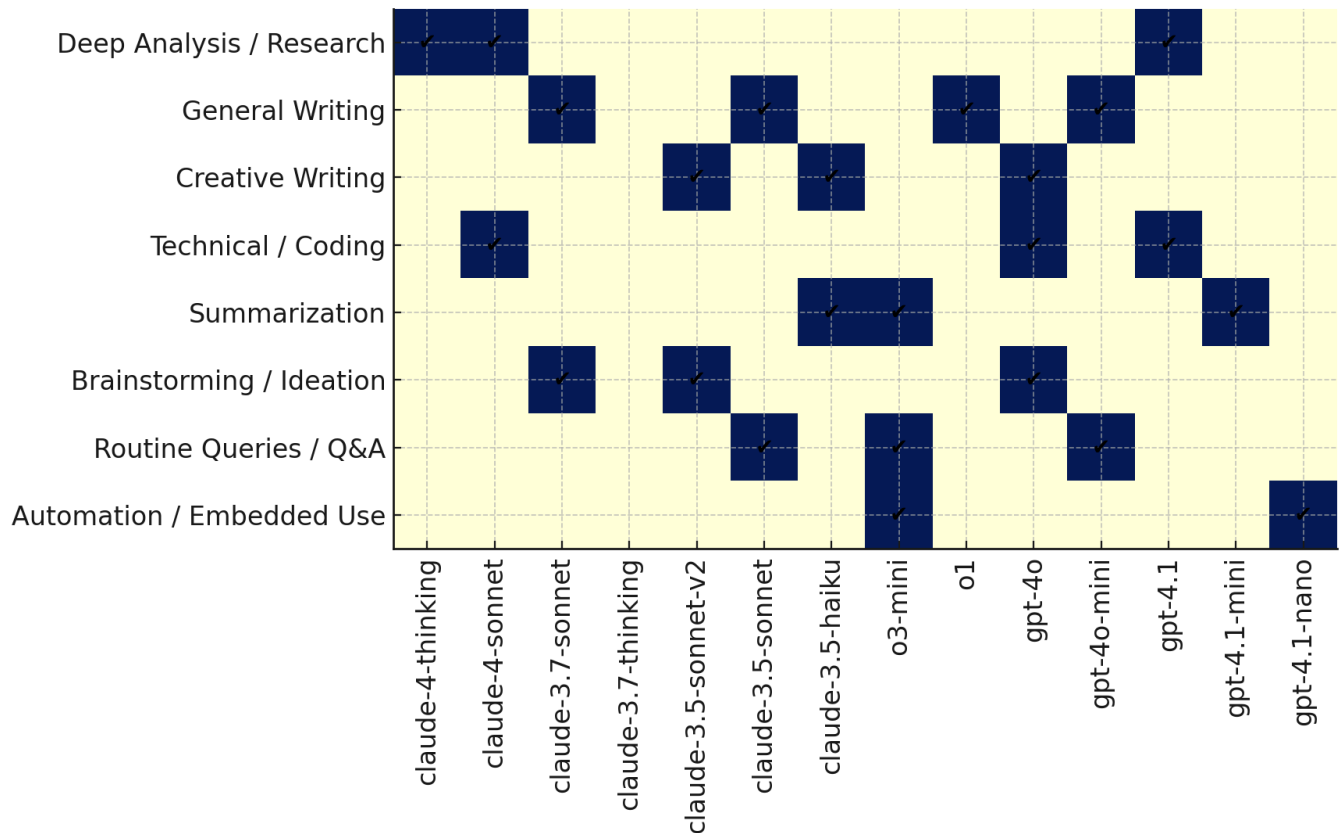
| Task Type | Best Models | Why |
|--|--|---|
| Deep Analysis / Research | claude-4-thinking, claude-4-sonnet, gpt-4.1 | Handle multi-step reasoning, long contexts, and nuanced problem-solving |
| General Writing (emails, notes, docs) | claude-3.5-sonnet, claude-3.7-sonnet, gpt-4o-mini, o1 | Reliable balance of clarity, speed, and cost efficiency |

| Task Type | Best Models | Why |
|---|---|---|
| Creative Writing (stories, poetry, brainstorming) | claude-3.5-haiku, claude-3.5-sonnet-v2, gpt-4o | Strong creative flair with precise or imaginative language |
| Technical / Coding Tasks | gpt-4.1, gpt-4o, claude-4-sonnet | Strongest for debugging, documentation, and long-form technical reasoning |
| Summarization (compact) | claude-3.5-haiku, o3-mini, gpt-4.1-mini | Quick, concise outputs optimized for brevity |
| Brainstorming / Ideation | claude-3.7-sonnet, claude-3.5-sonnet-v2, gpt-4o | Balance of creativity and structured reasoning |
| Routine Queries / Q&A | o3-mini, gpt-4o-mini, claude-3.5-sonnet | Fast, cost-effective, straightforward answers |
| Automation / Embedded Use | gpt-4.1-nano, o3-mini | Optimized for lightweight, high-speed integration |

✅ **How to use this matrix:**

- If you need **depth & reasoning**, lean on **Claude-4** or **GPT-4.1**.
- If you need **creativity**, pick **Claude-3.5 (Haiku/Sonnet v2)** or **GPT-4o**.
- If you need **speed**, use **mini/nano/o3-mini**.
- If you need an **everyday assistant**, use **Claude-3.7 Sonnet** or **GPT-4o-mini**.

Decision Matrix: Which Model to Pick for Each Task Type



🧠 Tips & Tricks

- **Use System Prompts Wisely:** Define the assistant's tone or role at the start.
- **Chunk Long Inputs:** Break text into smaller blocks to help with long documents.
- **Try Multi-Step Prompts:** Ask for step-by-step reasoning or summaries.
- **Leverage LLMs for data cleaning and CSV parsing.**
- **Use system prompts** for setting tone/role
- **Chain multiple tasks** into one prompt
- **Use Temperature** ~0.3 for factual work; ~0.7 for creative drafting
- **Ask the model** to explain code, fix bugs, or translate formats
- **Test multiple models** side-by-side for comparison

🔧 Common Uses in Heliophysics

- 🔍 Literature summarization (e.g., from NASA ADS or arXiv)
- 🌱 Experimental pipeline planning (e.g., solar flare prediction models)
- 📄 Code explanation and refactoring for legacy Fortran/Python
- 📦 Metadata extraction from satellite or simulation datasets

-  Drafting reports, proposals, or code documentation
-

⚠ Points of Caution

- **Hallucinations:** Always verify factual outputs from LLMs.
 - **Data Leakage:** Don't share sensitive or proprietary data with public models.
 - **Model Bias:** LLMs may reflect unintended social or scientific biases.
 - **Reproducibility:** Outputs can vary across sessions—log your prompts and versions.
-

Resources

Internal GSFC/HSD Resources

- [NASA Goddard AI Center of Excellence](#)
- [ChatGSFC](#)
- [Goddard Code Assistant](#)

External Tools

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

Think like a physicist. Prototype like a hacker. Document like a scientist.

✉ Questions or tips to share?

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