

From HTC to LLMs

Julia Gray
julia.ann.gray@gmail.com

Original slides by:
Jaehoon Yu <jaehoonyu1@gmail.com>

The Pattern You Already Know

Your Grid Analysis

Data

Multiple WLCG nodes

Work

Z→mumu analysis runs on each node

Combine

Histograms merged → Z peak emerges

Scheduler

HTCondor

same
pattern



LLM Training

Data

Billions of text samples

Work

Model forward/backward pass on each GPU

Combine

Gradients aggregated → weights updated

Scheduler

Kubernetes / SLURM

Today's Exercise

1 Run the script

```
python3 asp_assistant.py
```

2 Describe your plot

Plain English — no code needed yet

3 Get PyROOT code

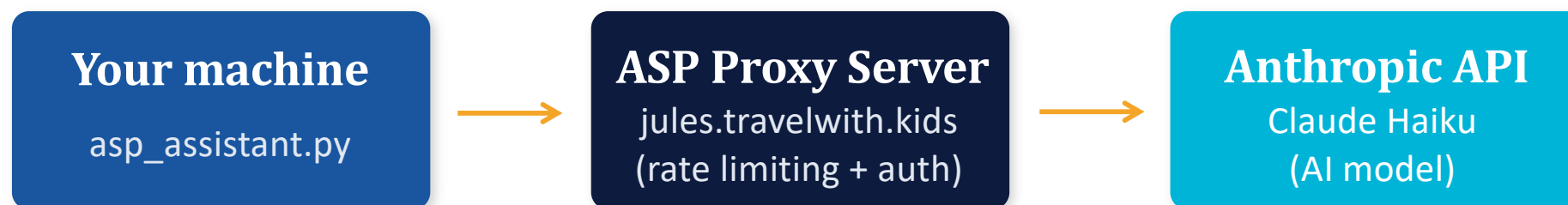
Copy it into your analysis script and run it

4 Iterate

Ask for improvements, observe what works and what doesn't

The script is available directly on the class page: `asp_assistant.py`

What's Happening Under the Hood



```
# What gets sent (JSON)
{ "prompt": "Plot zMass with a Gaussian fit",
  "token": "your_token" }
```

Same pattern as your HTCondor job submission — a structured request sent to a remote system that does the heavy work

After Today: Your Own Free Access

RECOMMENDED

Google Gemini

- Go to aistudio.google.com
- Sign in with any Google account
- Click Get API key → Create API key
- Free: 15 req/min, 1M tokens/day
- No credit card needed

NO GOOGLE NEEDED

OpenRouter

- Go to openrouter.ai
- Sign up with any email address
- Go to Keys → Create Key
- Use free models: Llama 3.1, Mistral
- No credit card needed

Change 2 lines in `asp_assistant.py` and you're using your own key forever — see the exercise documentation for the exact code

AI Tools for Physicists



AI in HEP today

Graph Neural Networks for jet tagging, ML triggers at LHC, anomaly detection in detector data



AI coding assistants

GitHub Copilot, Cursor, Claude — write analysis code faster, debug ROOT macros, generate boilerplate



AI for data analysis

Natural language queries over datasets, automated plot interpretation, systematic uncertainty estimation



African physics computing

The same grid infrastructure you used today supports climate modelling, genomics, and AI training across the continent