Fermipylot

An AI Assistant for Automated Fermi-LAT Data Analysis

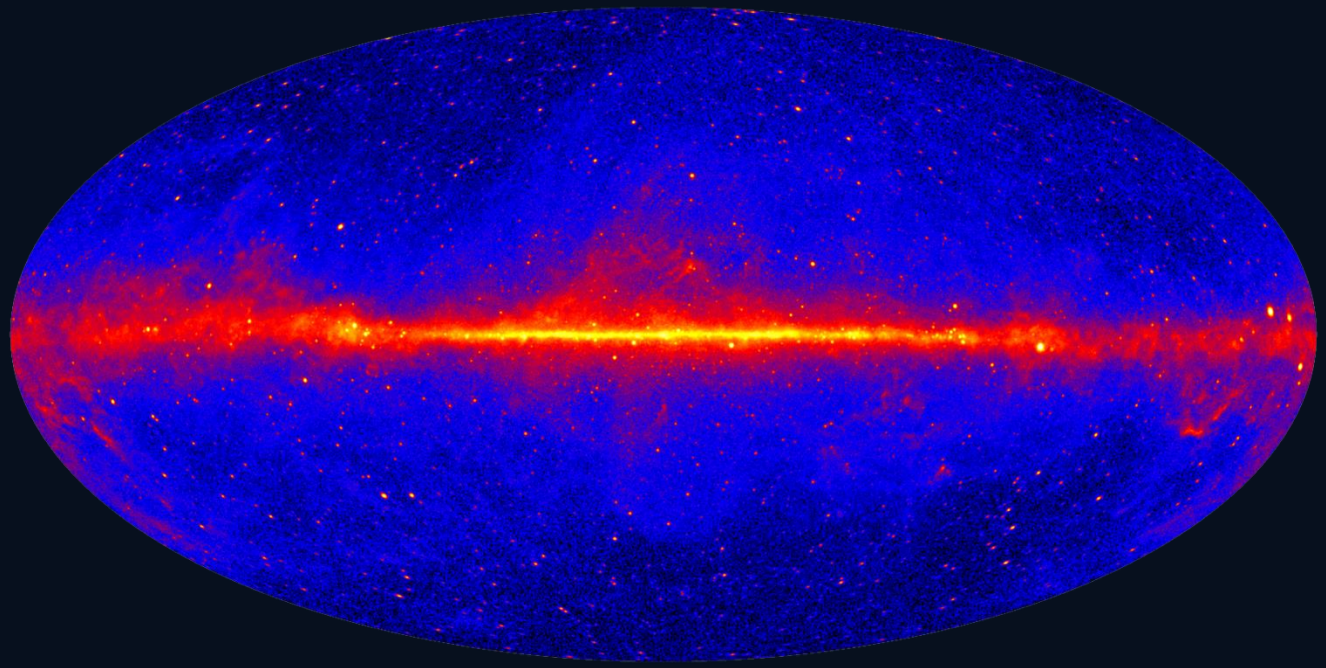
Ermes Aviano
University & INFN Trieste; Trieste, Italy

Saptashwa Bhattacharyya, Dmitry Malyshev,
Giacomo Principe, Gabrijela Zaharijas



The gamma-ray sky

Gamma rays are light with much higher energy than visible light.



15 years of LAT data above 1 GeV. >9.2 Million gamma rays. Bright band: our Milky Way.

Credit: NASA/DOE/Fermi LAT Collaboration

Allow the study of extreme processes, including jets near black holes and stellar explosions.

Fermi-LAT: how gamma rays become data

The Large Area Telescope (LAT) is one of two instruments onboard the Fermi satellite launched in 2008 and still operating.

Tracker

Tracks particles created by a gamma ray to reconstruct its incoming direction.

Calorimeter

Measures the energy deposited by those particles.

Anticoincidence shield

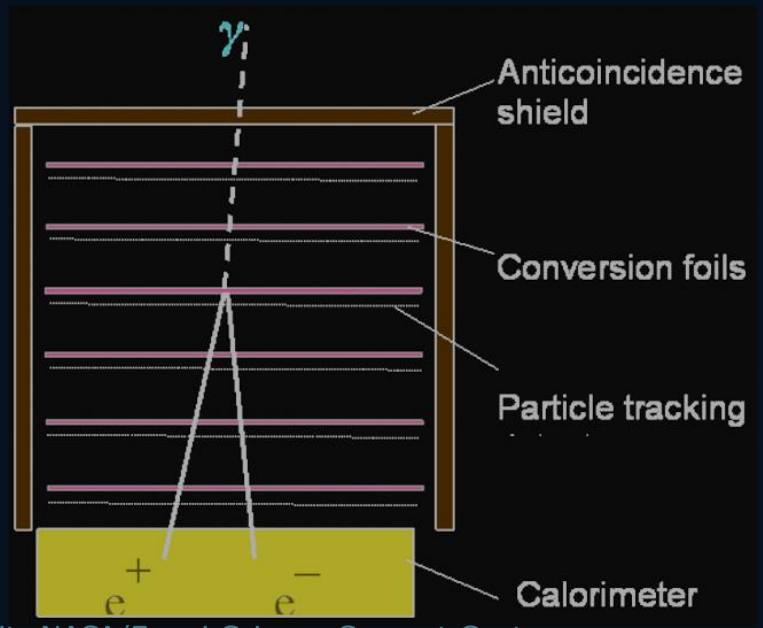
Rejects incoming charged-particle backgrounds.

Photon data = list of photons (direction, energy, arrival time)



PHOTON ENERGIES

~20 MeV to >300 GeV



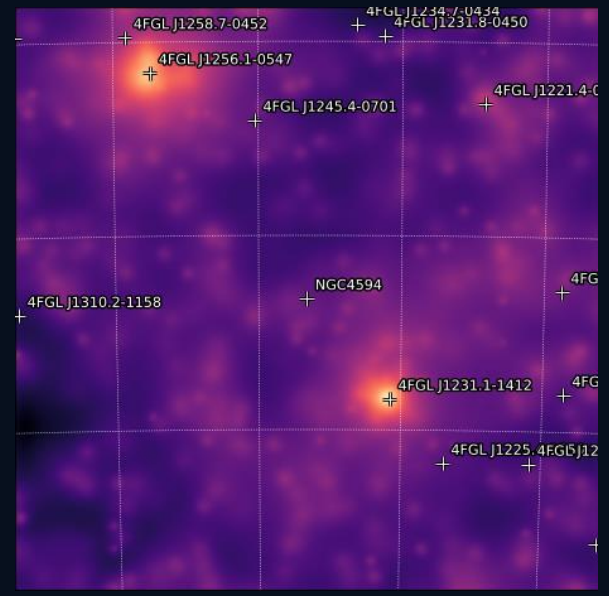
Credit: NASA/Fermi Science Support Center

From photons to a source measurement

Standard binned likelihood analysis with the Fermi tools

Analysis step	Tools and inputs
Select photons and good observing times	gtselect, gtmktime
Bin counts and compute exposure	gtbin, gtltcube, gtexpcube2
Model the target, nearby sources and diffuse emission	Catalogue + diffuse models gtsrcmaps
Fit source properties and inspect residuals	gtlike / pyLikelihood, gtmodel

A Region Of Interest (ROI)



15°×15° around NGC4594

The Fermi-LAT project provides the analysis resources through NASA FSSC:
data, software, catalogues, diffuse models and extensive documentation / tutorials.

Fermipy: one configuration, the same tools

A Python wrapper that organises the Fermi analysis workflow.

Configuration file – illustrative example

```
data:
  evfile: events.txt
  scfile: spacecraft.fits
selection:
  target: NGC4594
  emin: 100
  emax: 300000
```

One file records the data, target, time and energy cuts, binning and sky-model choices.

A few Python calls run the workflow

```
gta = GTAnalysis("config.yaml")
gta.setup()
# choose parameters to fit
gta.fit()
```

Fermipy runs the Fermi tools and provides fits, spectra, light curves and diagnostic maps.

The researcher chooses the settings and checks the results.

Extensive documentation: configuration reference, tutorials, notebooks and API.

The research goal

Can an AI assistant turn a scientific request into a reproducible analysis that a researcher can inspect?

THE CURRENT EFFORT

Researchers translate a question into specialised settings, code and a sequence of tasks.

THE PROPOSED SUPPORT

Fermipylot uses technical sources to prepare the workflow and record its choices.

Fermipylot architecture

Natural-language request



Interpret the goal and coordinate tools with LangGraph

Retrieve documentation, code and notebooks

Typed inputs that compile into a Fermipy configuration

Run analysis, check fit status and save outputs

Analysis Specification: target, time, energy ranges and requested products.

Results validation: run status and fit checks → implemented.
Scientific validation against expert analyses → next step.

METHOD

Models, RAG and bounded actions

Local LLM: Qwen2.5-72B-Instruct-GPTQ-Int4 (vLLM)

RETRIEVAL AND FEEDBACK

Hybrid retrieval: FAISS + BM25
Qwen3-Embedding-8B

↓

Reciprocal Rank Fusion + reranking
Qwen3-Reranker-8B

↓

Context and response critique
Up to three attempts per stage

Sources: Fermipy docs, code and notebooks.
Refine the search when needed.

MODE-DEPENDENT PERMISSIONS

MODE	ALLOWED EFFECT
Explain / plan	Answer or specification
Prepare	Job files only
Acquire data	Download data
Run analysis	Execute the fit

Ambiguous requests default to answering without execution.

PROTOTYPE EXAMPLE

OQ 334: a concrete analysis request

«I want to analyze OQ 334 in June 2026.»

OQ 334 is a catalogued gamma-ray source.

The user supplies the source and month. Other choices remain explicit and reviewable.

INPUT	RECORDED VALUE
Source	4FGL J1422.5+3223 (OQ 334)
Period	1 June to 1 July 2026
Energy	100 MeV to 300 GeV
Query radius	15 degrees
Product	Likelihood fit

Saved in AnalysisSpec and configuration files

What the prototype has demonstrated

STAGE	REPORTED OUTCOME
Data access	Data downloaded
Job checks	Validation passed
Fermipy execution	Run completed
Fit report	fit_success = true

OQ 334, JUNE 2026

One likelihood-fit job completed through the full technical workflow.

Scientific reliability and usefulness still require comparison with expert analyses.

This example reports a fit. It does not establish a general success rate.

The next step is systematic evaluation

QUESTION	PROPOSED EVIDENCE
Are the choices appropriate?	Compare settings and outputs with expert analyses
Does grounding help?	Assess source support, retrieval and configuration errors
Does the assistance help users?	Measure completion, corrections, time and compute cost

Current scope CLI prototype, cluster-specific setup, limited case coverage.

CONCLUSIONS

Fermipylot as a case study for scientific AI

Fermipylot aims to prepare configurations, run analyses and support results validation from natural-language requests.

Configurations, outputs and assumptions remain inspectable.

WHAT WE HAVE

A working prototype tested on a concrete analysis example (e.g. OQ 334).

WHAT WE WILL EVALUATE

Agreement with expert analyses, manual corrections and task time.

Broader goal: preserve and share scientific expertise.

FUTURE PERSPECTIVE

Combined Fermi LLM tools development collaborating with Clemson University, Stanford University and Louisiana State University


Thank you for your attention!

Back up slides

Prototype deployment and data access

REPORTED MODEL SETUP

Language model

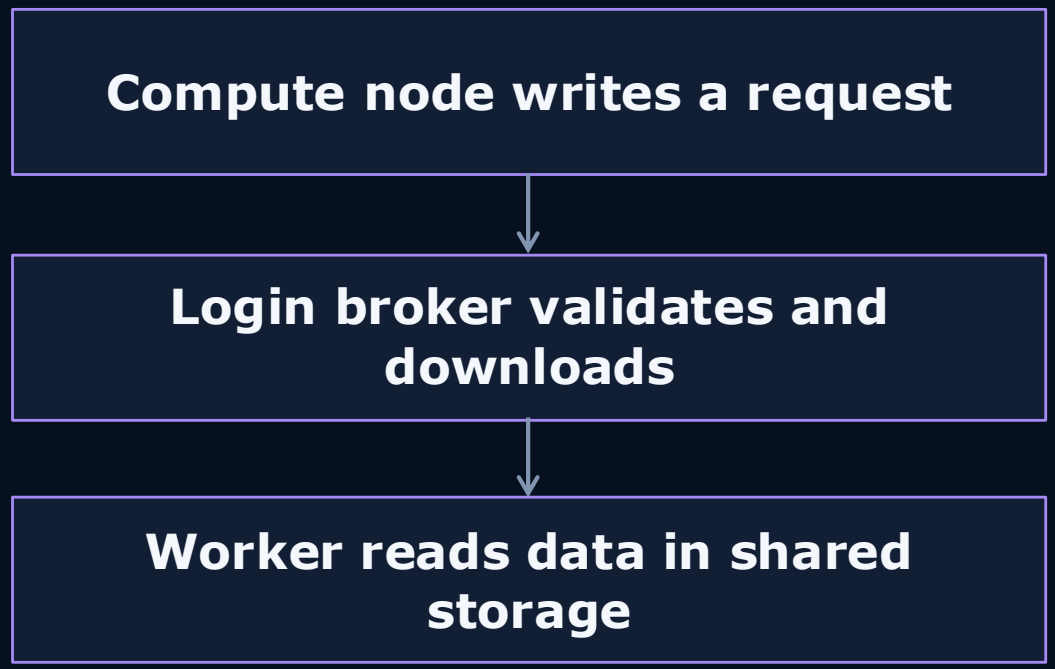
Qwen2.5-72B-Instruct-GPTQ-Int4

Embedding and reranking

Qwen3-Embedding-8B
Qwen3-Reranker-8B

One node: 32 CPUs, 4 GPUs, 256 GB
Separate agent and worker environments

CLUSTER DATA PATH



Downloads: photon events + spacecraft history

BACKUP

OQ 334 run record

REPORTED FIELD	VALUE
status	completed
fit_success	true
fit_quality	3
fit_status	0
logL	−33910.8636

SAVED ANALYSIS RECORD

Analysis Specification file
Configuration file
Manifest, scripts and logs
Worker report and fit outputs

Scope of this example

Likelihood fit only. The reported status and likelihood value do not establish scientific validity.