Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

GRAIS

Gamma Ray Artificial Intelligence System

Luca Naso, Koexai S.r.l.

Spoke 3 Progetti Bandi a Cascata, 09/01/2025

GRAIS
Gamma Ray Artificial Intelligence System

Area Tematica: 5

Intelligenza Artificiale applicata a missioni spaziali

Sotto-tematica: b)

Ricerca di sorgenti astronomiche transitorie in esperimenti di raggi gamma ad alta energia utilizzando tecniche di apprendimento automatico

Project Overview: challenges and proposed solutions

Astrophysics has advanced significantly in understanding high-energy phenomena like AGNs and GRBs. Recently the interest has shifted to detecting detecting **variations on very short time scales**, particularly **orphan afterglows**—transient events that lack observable primary emissions.
Space missions such as FERMI and AGILE have provided extensive gamma-ray datasets for over 15 years, yet identifying these short transient events remains a significant challenge, with very few confirmed to date, making the search similar to finding a needle in a haystack.

The GRAIS project aims to utilize **Machine Learning** techniques to address these challenges. However, the scarcity of training datasets for short transient events poses a limitation. To overcome this, GRAIS will implement **Generative AI** to create synthetic datasets that mimic the characteristics of these elusive events. This synthetic data will be combined with **semi-supervised learning and statistical approaches** to develop an automatic detection system, enhancing the identification of transient candidates within the vast observational datasets available.

Technical Objectives and Methodologies

1. **Automatic Identification System**
   Creation of **statistical and machine learning models** for the characterization of short-duration gamma-ray transients, including orphan afterglows.
   These models will be applied to existing observational data, enabling the automatic identification of transient events. The project will explore the possibility of implementing the system for real-time usage, to be used for example with Telegrams posted by FERMI community.
   The identified sources will also be shared with the international scientific community, enhancing the catalogue of known gamma-ray transients.

2. **Synthetic Dataset**
   Development of an innovative **Generative AI framework** for creating a comprehensive dataset of short-duration gamma-ray transients, particularly focusing on **orphan afterglows**.
   This synthetic dataset will be shared with the international scientific community, facilitating further research and collaboration in the field of high-energy astrophysics.

Involved Staff and new recruitments

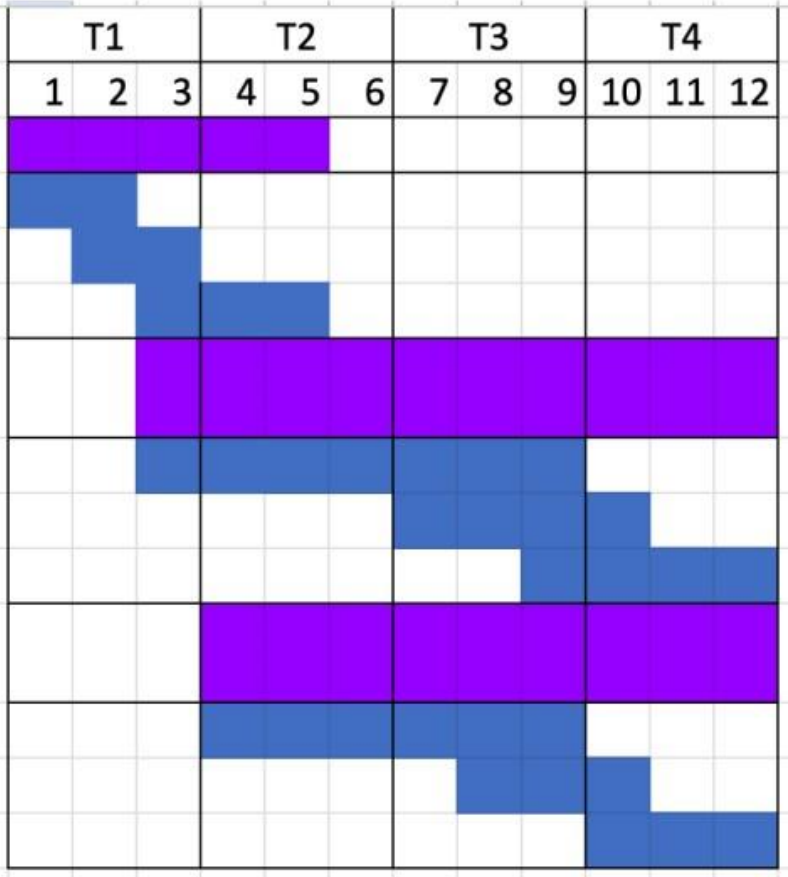
Spoke Scientific Advisors: Francesco Longo, Università degli Studi di Trieste e INFN Trieste
Sara Cutini, INFN Perugia

Koexai Team size: 1,75 FTE = 3.012 hours

1. Luca Naso: Project Lead (PhD in Astrophysics)
2. Marco Cataldo: Project Administrative Manager
3. Romina Anfuso: Project Manager
4. Vincenzo Del Zoppo: Data Scientist Team Lead
5. Stefano Calì: Senior Data Scientist (PhD in Elementary Particle Physics)
6. Fatima Modica Bittordo: Data Scientist

Gantt

OR	AR	Titolo
1		*Studio e caratterizzazione del dominio astrofisico e dei dataset disponibili*
	1.1	Studio dello stato dell'arte di sorgenti gamma transitorie e missioni spaziali
	1.2	Studio e preparazione dei dataset disponibili
	1.3	Studio e caratterizzazione delle sorgenti variabili note
2		*Studio e definizione di un framework di Intelligenza Artificiale per la generazione di dataset sintetici di transienti gamma di breve periodo*
	2.1	Studio e definizione del modello di AI generativa per la creazione del dataset sintetico
	2.2	Evoluzione e determinazione del modello di AI tramite confronto con le osservazioni
	2.3	Creazione e condivisione di un ampio dataset sintetico di eventi
3		*Studio e definizione di modelli di Statistica e Machine Learning avanzati per la caratterizzazione di transienti gamma di breve periodo*
	3.1	Caratterizzazione tramite modelli di Statistica e Machine Learning avanzati
	3.2	Ottimizzazione degli algoritmi avanzati
	3.3	Creazione e condivisione di un catalogo di sorgenti gamma transitorie candidate



Timescale, Milestones, SAL

Project start: Dec 2024

Milestone 1 – Feb 2025
Literature and dataset
- AR1.1 & AR1.2

Milestone 2 – May 2025
1st Gen AI model version
- AR1.3 & AR2.1

Milestone 3 – Aug 2025
1st characterization model
- AR2.2
- AR3.1 & AR3.2

Milestone 4 – Nov 2025
End of project
- AR2.3
- AR3.3