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

TEcniche non Convenzionali per l'anaLisi di immagini Astrofisiche (TECLA)

UNIVERSITÀ DEGLI STUDI DI PALERMO

Salvatore Micciché, UNIPA-DiFC

salvatore.micciche@unipa.it

DIFC Dipartimento di Fisica e Chimica - Emilio Segrè

Spoke 3 Progetti Bandi a Cascata, 24/09/2024

Project Overview

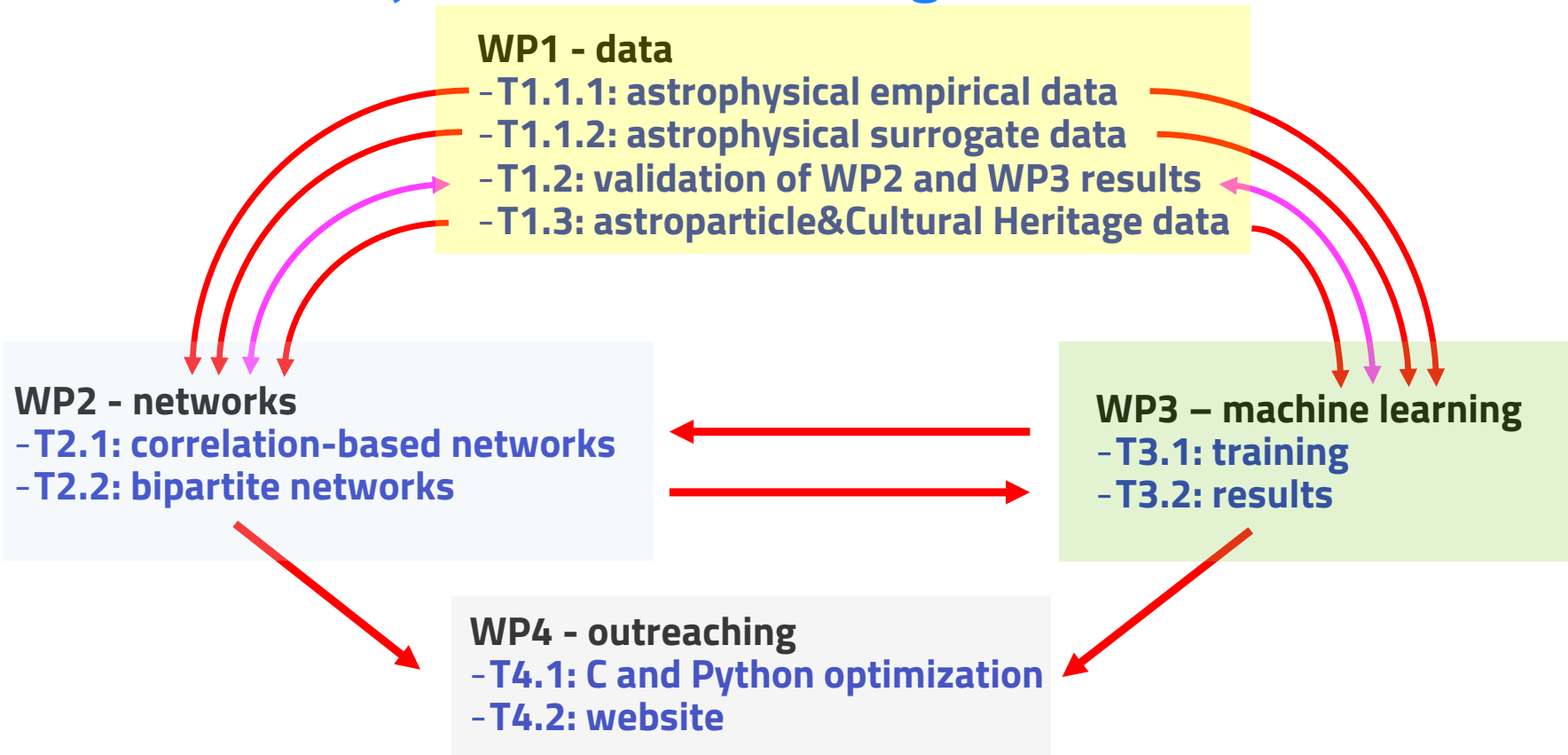
The project aims at developing innovative techniques for the detection, recognition and classification of coherent structures in astrophysical images. It is planned to study astrophysical images using statistical physics techniques, time series analysis techniques, techniques based on information theory and techniques based on Artificial Intelligence.

On the one hand, we will be interested in investigating the correlations existing between pixels present in the images of X-ray emissions of different metals. This will be done starting from the construction of a correlation-based network of pixels and the subsequent use of filtering techniques that allow us to identify the most significant correlations between pixels.

On the other hand, we will exploit the possibility of using machine learning techniques to enhance the power of the methodologies previously investigated.

The methodologies developed during the project will then be tested with other types of data, mainly of astroparticle nature and of Cultural Heritage.

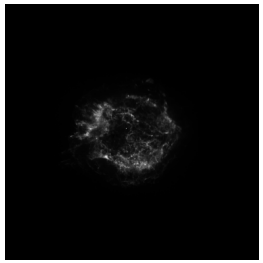
Technical Objectives, Methodologies and Solutions



Technical Objectives, Methodologies and Solutions

WP2

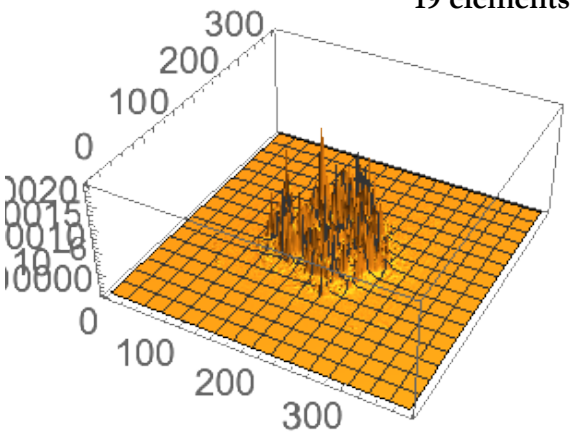
- correlations existing between pixels in X-ray emission images of different metals construction of a correlation-based network of pixels. The most scientifically relevant aspect concerns the choice of the similarity measure between the numerical matrices representing the astrophysical image. Potential candidates are the Frobenius and Kullback-Leibler distances.
- Subsequently, we will use filtering techniques that allow us to identify the most significant correlations between pixels. We will use unsupervised clustering techniques in order to identify anisotropies and typify supernova remnants.
- We will evaluate an investigation approach based on a description in terms of Bipartite Networks between elements and pixels (the problem is the threshold)



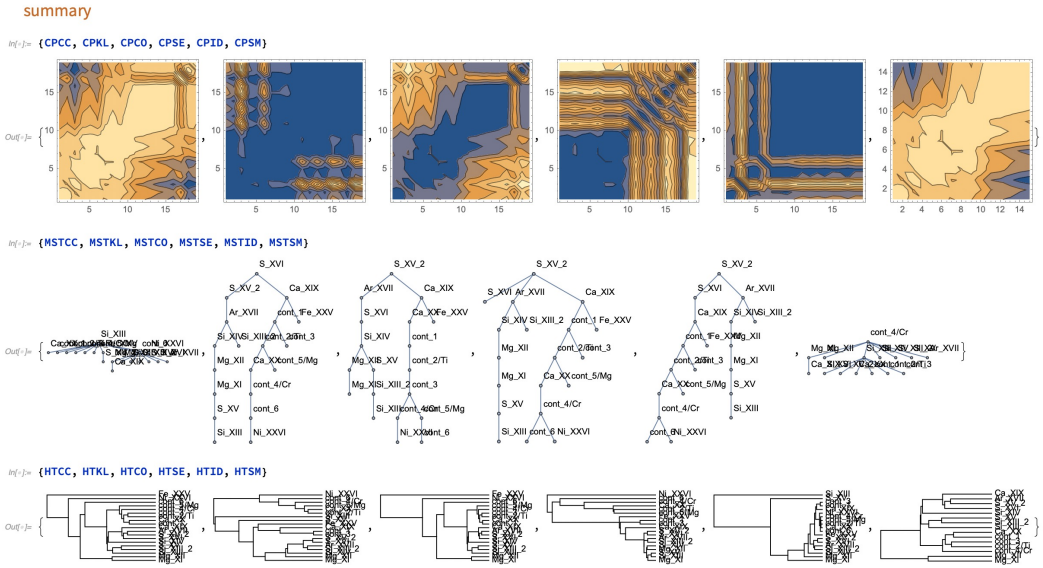
Input data

observations of the supernova remnant (SNR) Cassiopeia A (Cas A)

19 elements



{Mg_XI, Mg_XII, Si_XIII, Si_XIV, Si_XIII_2, S_XV, S_XVI, S_XV_2, Ar_XVII, Ca_XIX, Ca_X", cont_1, cont_2/Ti, cont_3, cont_4/Cr, cont_5/Mg, Fe_XXV, cont_6, Ni_XXVI}



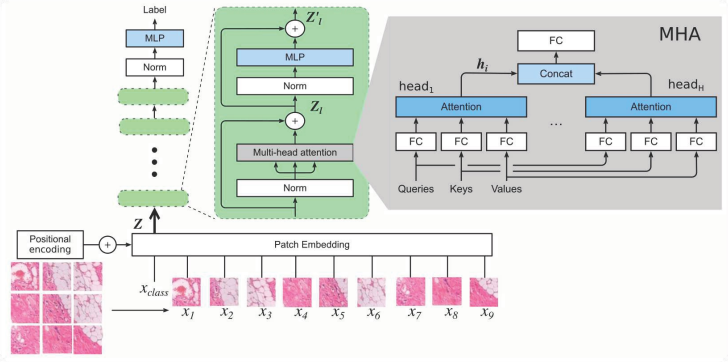
Technical Objectives, Methodologies and Solutions

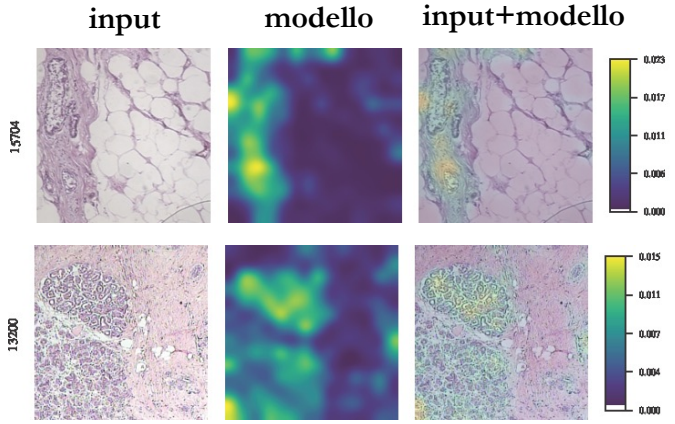
WP3

- traditional statistical analysis can be effectively combined with methods borrowed from artificial intelligence such as supervised or unsupervised Machine Learning and feature selection techniques.
- very accurate 3-D MHD simulations will be used as datasets to train AI methods in both supervised and unsupervised mode.
- Furthermore, they will also allow an effective validation of the obtained results.
- We will also explore the parameter space through feature selection techniques in order to (i) better understand the physical link between each feature and the classification and (ii) avoid or mitigate the risk of parameter ***overfitting***.

The Transformer architecture was proposed for natural language processing and is an encoder-decoder architecture characterized by the so-called attention layer. Vision Transformers (ViT) are customized versions of the original Transformer that inherit the specificities of the text transformer, namely the ability to model long-term relationships in the data: this feature could be important in the case of astronomical images where we expect long-range relationships between pixels.

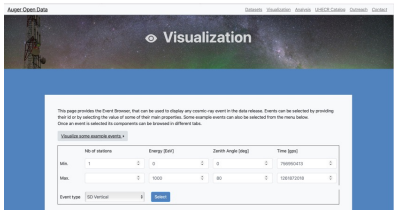
Transformers and their attention module will be used to define a model explainability mechanism that highlights which parts of the image contribute the most to pattern identification or supervised classification. Therefore, the possibility of having simulated numerical data for image generation will be crucial to augment the learning datasets, using augmentation techniques.

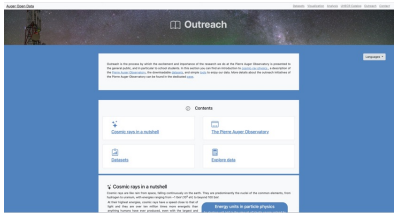


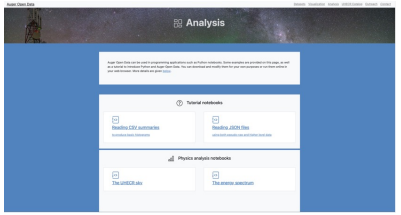


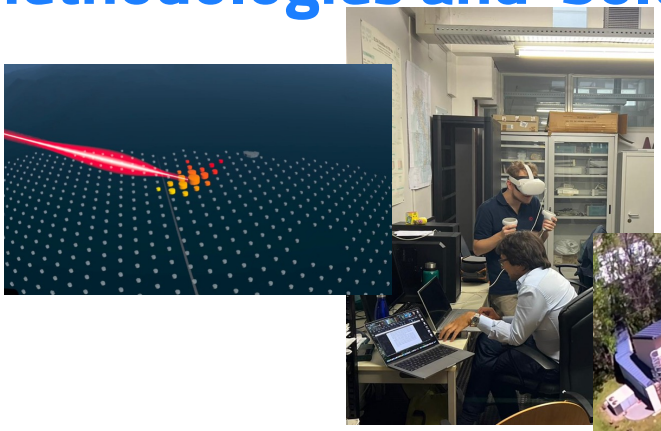
Technical Objectives, Methodologies and Solutions

WP4







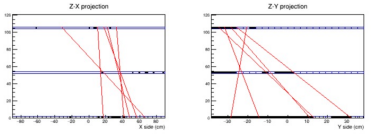


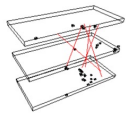
Auger experiment

Extreme Energy Events Science inside Schools

Extreme Energy Events Science inside Schools

PISA-01 DQM Event display





PISA-01
event 27063
Date 21/9/2024
Run 00004
UTC: 11:54:45

CENTRO RICERCHE ENRICO FERMI — INFN — Progetto Extreme Energy Events (EEE) La Scienza nelle Scuole

Home Ricerca EEE Monitor (DQM) Video Scuole Notizie Link utili Riservato Logo Homepage Centro Fermi

Involved Staff and new recruitments

Staff	Research Area	Role	
Salvatore Micciché	Complex Systems and Complex Networks	PO	
Rosario N. Mantegna	Statistical Physics and Complex Systems	PO	
Giosuè Lo Bosco	Machine learning and deep learning	PA	
Claudio Fazio	Physics Education	PO	
Marco Miceli	Astrophysics	PA	
Maria Luisa Saladino	Chemistry Cultural Heritage	PA	
Giovanni Marsella	Astroparticles	PO	

To be hired

Postdoc with strong IT skills

- **Python**
- **Machine learning**
- **Big data**

(ASTRO)data scientist

Consultancy

Website for

- **Public Engagement**
- **Online calculations**

Involvement in WPs
WP1
Marco Miceli
Giovanni Marsella
Maria Luisa Saladino
Postdoc
WP2
Salvatore Micciché
Rosario N Mantegna
Postdoc
WP3
Giosuè Lo Bosco
Postdoc
WP4
Claudio Fazio
Postdoc
Consultancy

Timescale, Milestones, SAL

						ore mese per mese														
docente	ruolo	costo	ORE 2024	ORE 2025	ORE TOTALE	202408	202409	202410	202411	202412	202501	202502	202503	202504	202505	202506	202507	check 2024	check 2025
micciche	PO	73	20	330	350	30	25	35	30	15	30	20	20	30	30	40	45	-115	115
mantegna	PO	73	20	180	200	20	20	25	25	20	25	30	20	15	0	0	0	-90	90
lo bosco	PA	48	100	100	200	20	30	30	15	15	10	20	10	10	10	20	10	-10	10
fazio	PO	73	20	130	150	0	15	15	0	0	10	10	0	0	30	40	30	-10	10
miceli	PA	48	60	60	120	10	25	25	10	10	0	5	5	5	5	10	10	-20	20
saladino	PA	48	85	15	100	10	30	35	5	10	0	0	0	0	0	10	0	-5	5
marsella	PO	73	85	15	100	0	45	45	0	0	0	0	0	0	0	10	0	-5	5
						90	190	210	85	70	75	85	55	60	75	130	95		
						ore		490	ore		230	ore		200	ore		300		
						5570	11745	13080	5455	4235	5225	5580	3640	4005	5100	8490	6435		
						costo		30395	costo		14915	costo		13225	costo		20025		
						costo cumulato		30395	costo cumulato		45310	costo cumulato		58535	costo cumulato		78560		0
								SAL1			SAL2			SAL3			SAL4		
										645							575		
										40085							38475		
										2024							2025		

Timescale, Milestones, SAL

Milestone	Data	descrizione	SAL (Euro)
M1	04/11/2024	**WP1:** preparation of astrophysical, astroparticle and cultural heritage data: agreements with WP2 and WP3 about the data format and their description. Start of the procedures for the recruitment of the research fellow.	30395.00 +postdoc 1/4
M2	04/02/2025	**WP1:** final delivery of astrophysical, astroparticle and cultural heritage data. **WP2:** start data analysis with correlation-based networks. **WP3:** start data analysis with AI techniques. **WP4:** start procedures for the selection of the consultancy firm.	14915.00 +postdoc 2/4
M3	04/05/2025	**WP2:** completion of correlation-based network analyses. Start of data analysis with bipartite networks. **WP3:** continuation of data analysis with AI techniques. **WP4:** start of the consulting firm activities. WP2 and WP3 requirements on the software specifications.	13225.00 +postdoc 3/4
M4	04/08/2025	**WP1:** Validation activities of WP2 and WP3 results. Validation activities on WP4 activities. **WP2:** completamento delle analisi riguardanti i network bipartiti. **WP3:** completion of analyses regarding bipartite networks. **WP4:** WEB site delivery and testing with selected schools. Delivery of software in Python and C.	20025.00 + consultancy +postdoc 4/4 +15%

TOTAL BUDGET
149844,00

Staff	78560
Postdoc	30000
Consultancy	25000
	133560
15% (S+P)	16284
TOTAL	**149844**