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

Spoke 3 - Astrophysics and Cosmos Observations

Spoke Status

U. Becciani - P. Lubrano

Third Spoke 3 Technical Meeting
Perugia
26-29 Maggio 2025

Spoke 3 - Meetings, Schools, Conferences.... (highlights...)

- General Meetings
  - Spoke3 Meeting, May 2024, https://indico.ict.inaf.it/event/2752
  - Kick Off Cascade Funding, Sep 2024 https://indico.ict.inaf.it/event/3002
  - 2nd Technical Workshop, Dec 2024, https://indico.ict.inaf.it/event/3001
  - 2nd Kick Off Cascade Funding, Jan 2025, https://indico.ict.inaf.it/event/3110
  - Weekly and bi-weekly WP regular meetings, https://indico.ict.inaf.it/category/205
- Computing Schools
  - Course in Computing and HPC for Astronomy and Astrophysics, Jun 2024, https://indico.ict.inaf.it/event/2785
- Conferences
  - International Conference on Machine Learning for Astrophysics, Jul 2024, https://indico.ict.inaf.it/event/2690
  - Participation to Supercomputing 2024, Nov 2024, https://sc24.supercomputing.org
  - A special thematic session (ACOHPC) proposed by Spoke3 has been accepted by the Parallel Distributed Computing IAC (PDP2025), Mar 2025. 14 presentations for PhD students and young post-docs. http://pdp2025.org
- Events and Press
  - Event: Osservatorio ICSC (Catania) 24 October 2024: https://osservatorio.supercomputing-icsc.it/eventi/evento-dellosservatorio-icsc-24-ottobre-2024-catania/
  - Press release: https://www.wired.it/article/mappa-rischi-terremoti-alluvioni-italia/?utm_source=pocket-newtab-it-it
  - Press release: https://www.ansa.it/canale_scienza/notizie/frontiere/2024/04/17/lo-studio-dei-segreti-delluniverso-passa-dai-supercomputer-_53ee3694-c31f-44df-aa08-2604c14b0102.html

Recent and future events

Supercomputing 2024 (SC24- Atlanta 17-22 Nov.)

https://sc24.supercomputing.org/

Scientific Program: The Program is designed to share best practices in all areas of high performance computing. Become a contributor and share your excellence. **Presenting at SC is extremely rewarding.**

Exhibits: It is an opportunity to meet and interact with industry, research organizations, universities, and startups joining us from across the globe.

INAF Booth space INAF reserach on HPC and Centro Nazionale HPC, Big Data e Quantum Computing

SC25 ST. LOUIS NOV 16–21

PDP 2025

33° Euromicro International Conference on Parallel, Distributed and Network-based processing

Special Session 13 March 2025 Spoke3

ACOHPC (Astrophysics and Cosmos Observations: HPC and Big Data Management)

Submitted 23 articles and accepted 14 that has been presented at the session

Future events

- Basic and Intermediate FPGA programming school, Jun 2025 (65 students enrolled)
- Advanced FPGA programming techniques school, Nov 2025
- HPC School (Catania), 22-26 Sept. 2025
- Containerization Course (Bologna) Nov. 2025

XC Final PNRR Meeting: Elba 5- 8 October 2025 (TBV)

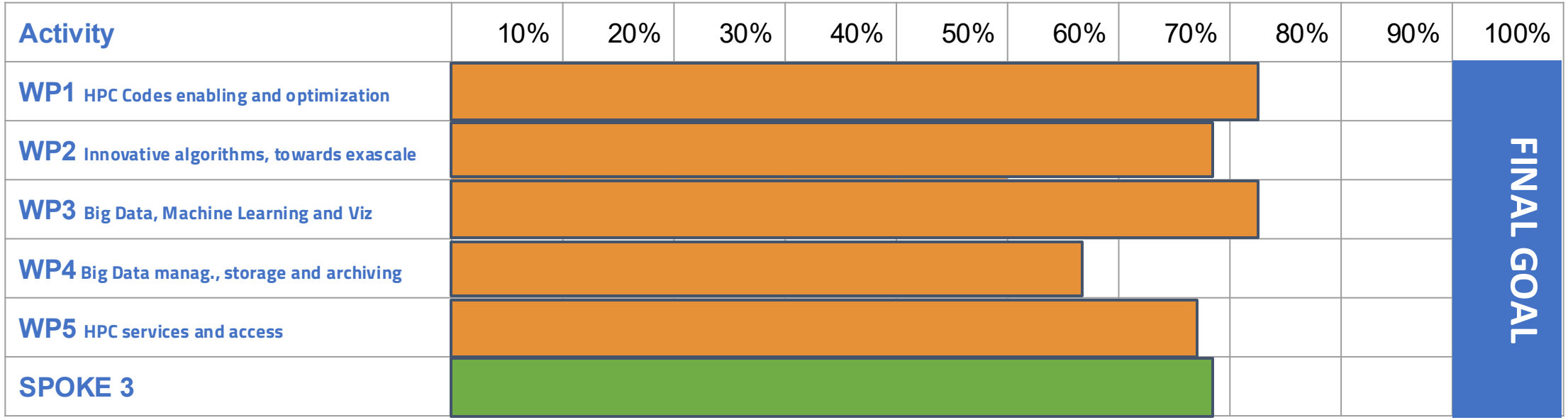
Spoke 3 Final Meeting in preparation: Sexten 14-19 December 2025

Joint meeting with SPACE (Scalable Parallel and distributed Astrophysical Code for Exascale)

The **proceedings** of this meeting **will be published in a special issue of the journal Astronomy&Computing**

Sesto Pusteria (Bz) https://www.sexten-cfa.eu/event/icsc-spoke3-fm

WPs Progress Summary (February 2025)



Scientific Publications (Feb 2025)

More than **150 papers** have been published on peer-reviewed Journals within the R&D activities of Spoke 3

The list (updated at the beginning of December 2024) is available at:
https://drive.google.com/drive/folders/1s5cFVAH12R0Xx75yvyAy97Iou7SAtvHG

Radio U-Net: a convolutional neural network to detect diffuse r sources in galaxy clusters and beyond

C. Stuardi, C. Gheller, F. Vazza and A. Botteon

ABSTRACT

The forthcoming generation of radio telescope arrays promises significant advancements in sc identification and characterization of many new faint and diffuse radio sources. Conventiona are anticipated to be insufficient to exploit the capabilities of new radio surveys. Radio interl present a challenge for image segmentation tasks due to noise, artifacts, and embedded challenges, we introduce Radio U-Net, a fully convolutional neural network based on the designed to detect faint and extended sources in radio surveys, such as radio haloes, relic U-Net was trained on synthetic radio observations built upon cosmological simulations an clusters, where the detection of cluster diffuse radio sources relied on customized data Low-Frequency Array Two metre Sky Survey (LoTSS) data. The 83 per cent of clusters ex accurately identified, and the segmentation successfully recovered the morphology of the In a test sample comprising 246 galaxy clusters, we achieved a 73 per cent accuracy rate with and without diffuse radio emission. Our results establish the applicability of Radio U sets, probing its efficiency on cutting-edge high-performance computing systems. This app optimizing the exploitation of forthcoming large radio surveys for scientific exploration.

Key words: techniques: image processing – galaxies: clusters: intracluster medium – softw

The Gaia AVU–GSR solver: CPU+GPU parallel solu for linear systems solving and covariances calculation t Exascale systems

Valentina Cesare, Ugo Becciani, Alberto Vecchiato, Mario Gilberto Lattanzi, Aldinucci, and Beatrice Bucciarelli

ABSTRACT

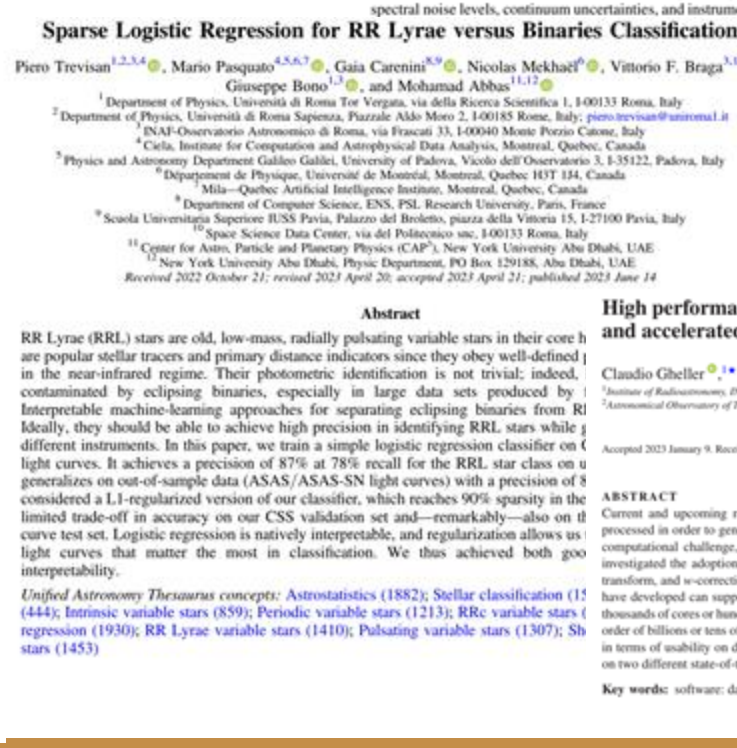
21CMEMU: an emulator of 21CMFAST summary observables

Daniela Breitman, Andrei Mesinger, Steven G. Murray, David Prelogović, Yuxiang Qin and Roberto Trotta

ABSTRACT

Recent years have witnessed rapid progress in observations of the epoch of reionization (EoR). These have enabled high-dimensional inference of galaxy and intergalactic medium (IGM) properties during the first billion years of our Universe. However, even using efficient, seminumerical simulations, traditional inference approaches that compute 3D lightcones on-the-fly can take 10^5 core hours. Here we present 21CMEMU: an emulator of several summary observables from the popular 21CMFAST simulation code. 21CMEMU takes as input nine parameters characterizing EoR galaxies, and outputs the following summary statistics: (i) the IGM mean neutral fraction; (ii) the 21-cm power spectrum; (iii) the mean 21-cm spin temperature; (iv) the sky-averaged (global) 21-cm signal; (vi) the ultraviolet (UV) luminosity functions (LFs); and (vii) the Thomson scattering optical depth to the cosmic microwave background (CMB). All observables are predicted with sub- per cent median accuracy, with a reduction of the computational cost by a factor of over 10^6. After validating inference results, we showcase a few applications, including: (i) quantifying the relative constraining power of different observational data sets; (ii) seeing how recent claims of a late EoR impact previous inferences; and (iii) forecasting upcoming constraints from the sixth observing season of the *Hydrogen Epoch of Reionization Array* (HERA) telescope. 21CMEMU is publicly available, and is included as an alternative simulator in the public 21CMMC sampler.

Key words: cosmology: theory – dark ages, reionization, first stars – methods: statistical – methods: data analysis.



Sparse Logistic Regression for RR Lyrae versus Binaries Classification

Piero Trevisan, Mario Pasquato, Gaia Carenini, Nicolas Mekhaël, Vittorio F. Braga, Giuseppe Bono, and Mohamad Abbas

Abstract

RR Lyrae (RRL) stars are old, low-mass, radially pulsating variable stars in their core h are popular stellar tracers and primary distance indicators since they obey well-defined in the near-infrared regime. Their photometric identification is not trivial; indeed, contaminated by eclipsing binaries, especially in large data sets produced by Interpretable machine-learning approaches for separating eclipsing binaries from RI Ideally, they should be able to achieve high precision in identifying RRL stars while different instruments. In this paper, we train a simple logistic regression classifier on light curves. It achieves a precision of 87% at 78% recall for the RRL star class on generalizes on out-of-sample data (ASAS/ASAS-SN light curves) with a precision of considered a L1-regularized version of our classifier, which reaches 90% sparsity in the limited trade-off in accuracy on our CSS validation set and—remarkably—also on th curve test set. Logistic regression is natively interpretable, and regularization allows us light curves that matter the most in classification. We thus achieved both goo interpretability.

Unified Astronomy Thesaurus concepts: Astrostatistics (1882); Stellar classification (15 (444); Intrinsic variable stars (859); Periodic variable stars (1213); RRc variable stars (regression (1930); RR Lyrae variable stars (1410); Pulsating variable stars (1307); Sh stars (1453)

Parameter estimation from the Lyα forest in the Fourier space using an information-maximizing neural network

Soumak Maitra, Stefano Cristiani, Matteo Viel, Roberto Trotta, and Guido Cupani

ABSTRACT

Aims. Our aim is to present a robust parameter estimation with simulated Lyα forest spectra from Sherwood-Relics simulations suite by using an information-maximizing neural network (IMNN) to extract maximal information from Lyα 1D-transmitted flux in the Fourier space.

Methods. We performed 1D estimations using IMNN for intergalactic medium (IGM) thermal parameters T_0 and γ at $z = 2$–4, and cosmological parameters σ_8 and n_s at $z = 3$–4. We compared our results with estimates from the power spectrum using the posterior distribution from a Markov chain Monte Carlo (MCMC). We then checked the robustness of IMNN estimates against deviation in spectral noise levels, continuum uncertainties, and instrumental smoothing effects. Using mock Lyα forest sightlines from the publicly

High performance *w*-stacking for imaging radio astronomy data: a parallel and accelerated solution

Claudio Gheller, Giuliano Taffoni and David Goz

ABSTRACT

Current and upcoming radio-interferometers are expected to produce volumes of data of increasing size that need to be processed in order to generate the corresponding sky brightness distributions through imaging. This represents an outstanding computational challenge, especially when large fields of view and/or high-resolution observations are processed. We have investigated the adoption of modern high performance computing systems specifically addressing the gridding, fast Fourier transform, and *w*-correction of imaging, combining parallel and accelerated solutions. We have demonstrated that the code we have developed can support data set and images of any size compatible with the available hardware, efficiently scaling up to thousands of cores or hundreds of graphic processing units, keeping the time to solution <1 h even when images of the size of the order of billions or tens of billions of pixels are generated. In addition, portability has been targeted as a primary objective, both in terms of usability on different computing platforms and in terms of performance. The presented results have been obtained on two different state-of-the-art high performance computing architectures.

Key words: software: data analysis – software: development – techniques: image processing – techniques: interferometric.

Spoke 3 and IG Projects

Spoke 3 on the **first call** received **an allocation of approximately 1MEuros for funding**.
Spoke 3 is **involved in five main projects**, acting as the proposing Spoke for some of them.

Spoke 3 allocated funds: **approximately €850,000** were **assigned to public institutions** that participated (RomaTOV and INAF)

The second call closed in December 2023 . Two additional projects were proposed:

- Proposing companies: IFAB (HMMA) and ISP (ATS with INFN and INAF)

The third call closed in March 2024. Two more (support) projects were proposed by Spoke 3:

- Proposing company: Leonardo (IGUC), and Leithà/UnipolSai (HAEQ with INAF)

In calls 2 and 3, Spoke 3 was allocated an additional €300,000 to public institutions.

Cascade Funding

N.	Thematich Area and Spoke WP correlations	Approved Euros
1	Scientific Visualization with Artificial Intelligence Support. **WP 3 – 10 Projects Collaborative Activity for all of these !!**	**1.585.800**
2	Accelerated Alghorithms **WP 2– 2 Projects**	**332.642**
3	Parallelization and acceleration of scientific codes **WP 2 – 4 Projects 2 Collaborative Activities**	**517.400**
4	Intelligenza Artificiale applicata alla features detection e classificazione **WP 1 and 3 - 3 Projects 2 Collaborative Activities**	**448.371**
5	Intelligenza Artificiale applicata a missioni spaziali **WP 1 3 and 4 - 7 Projects 3 Collaborative Activities**	**910.800**
6	Metodologie avanzate di analisi dati **WP1 2 and 3 - 10 Projects 2 Collaborative Activities**	**1.471.500**

- **Initial funds assignment 3,200,000 Euros**
- **Additional funds on October 2024. 2,066,513 Euros**

First KO Meeting 24 Sept. 2024

- **22 Projects started on Sept.**

Second KO Meeting 9 Jan. 2025

- **13 Project Started on Dec.**

The projects are synergic with the research activities of the WPs and contribute to their outcomes.

22 Projects are submitted from private PMI

An internal Spoke 3 Reviewer or collaborator is supervising the activity.

Cascade Funding

Sotto-Tematica	Ente	Progetto	Agevolazione richiesta
6A	Università degli Studi di Napoli Parthenope	AI-Legs	€ 150,000.00
1C ed 1E	Intellisync S.r.l.	Astroview	€ 229,905.49
1A	Università degli Studi di Perugia	AI-VISIO-LAB	€ 198,513.00
3B	Università degli Studi G. D'Annunzio Chieti - Pescara	PaGUSci	€ 191,000.00
3D	Università degli Studi di Camerino	OASIS	€ 99,479.00
5D	Università degli Studi di Milano	MLS	€ 129,580.00
1D	Koexai S.r.l.	DeepCosmonet	€ 170,000.00
4A	Università degli Studi di Palermo	Tecla	€ 149,884.00
6E	Università degli Studi di Palermo	CatMad	€ 172,357.50
6C	Koexai S.r.l.	ASTRAI	€ 90,000.00
1B	Auticon S.r.l.	Astrovisio	€ 179,866.02
3A	Università Cà Foscari Venezia	SDEGNO	€ 80,093.75
2A	E4 Computer Engeneering S.p.A.	RISC-V4Astro	€ 173,635.00
2B	Nurjana Technologies S.r.l.	Nebula	€ 159,006.66
6B	Beta80 S.p.A.	Dashiboard	€ 80,000.00
3C	ExactLab S.r.l.	BOG	€ 146,832.00
5B	Università degli Studi Guglielmo Marconi	SPECTRA	€ 95,436.00
6D	Università degli Studi di Milano	GRAVITY	€ 166,240.00
5C	Itahca S.r.l.	CANDELA	€ 97,000.00
4B	Fuko S.r.l.	AstroClass	€ 148,500.00
5A	Nuclear Instruments S.r.l.	SPARTAN	€ 150,000.00
	AGEVOLAZIONE TOTALE		€ 3,057,328.42
1A	Auticon S.r.l.	AstroTool	€ 198,847.37
	AGEVOLAZIONE TOTALE Incluso AstroTool		**€ 3,256,175**

Sotto-tematica	Ente - Acronimo Progetto	Agevolazione
1E	Tinvention - DISCORSIVO	€ 180,000.00
6E	UniCamerino - Grab-IT	€ 174,917.00
6D	UniInsubria - FaB-HPCc	€ 168,150.00
5D	Datrix - PADDLE	€ 101,902.13
6A	Nuclear Instrument - LEGIMAC	€ 150,000.00
1C	HOB - PRESOL	€ 49,996.00
5B	Koexai - Grais	€ 100,000.00
5C	Alten - NSM	€ 146,500.00
4A	Koexai - Espai	€ 150,000.00
6A	UniMarconi - GRAIL	€ 149,780.00
1E	Auticon - Astrodata	€ 178,703.51
6D	Koexai - FastParam	€ 170,000.00
1A	Consorzio SCIRE - VIRTUOSO	€ 200,000.00
5C	UniUrbino - AIDA	€ 90,389.00
	Total	**€ 2,010,337**

36 NEW projects for Spoke 3

➔ **Big Opportunity**

Some projects on the same topic have agreed to collaborate on common activities.

Cascade Funding: Overview and progress status

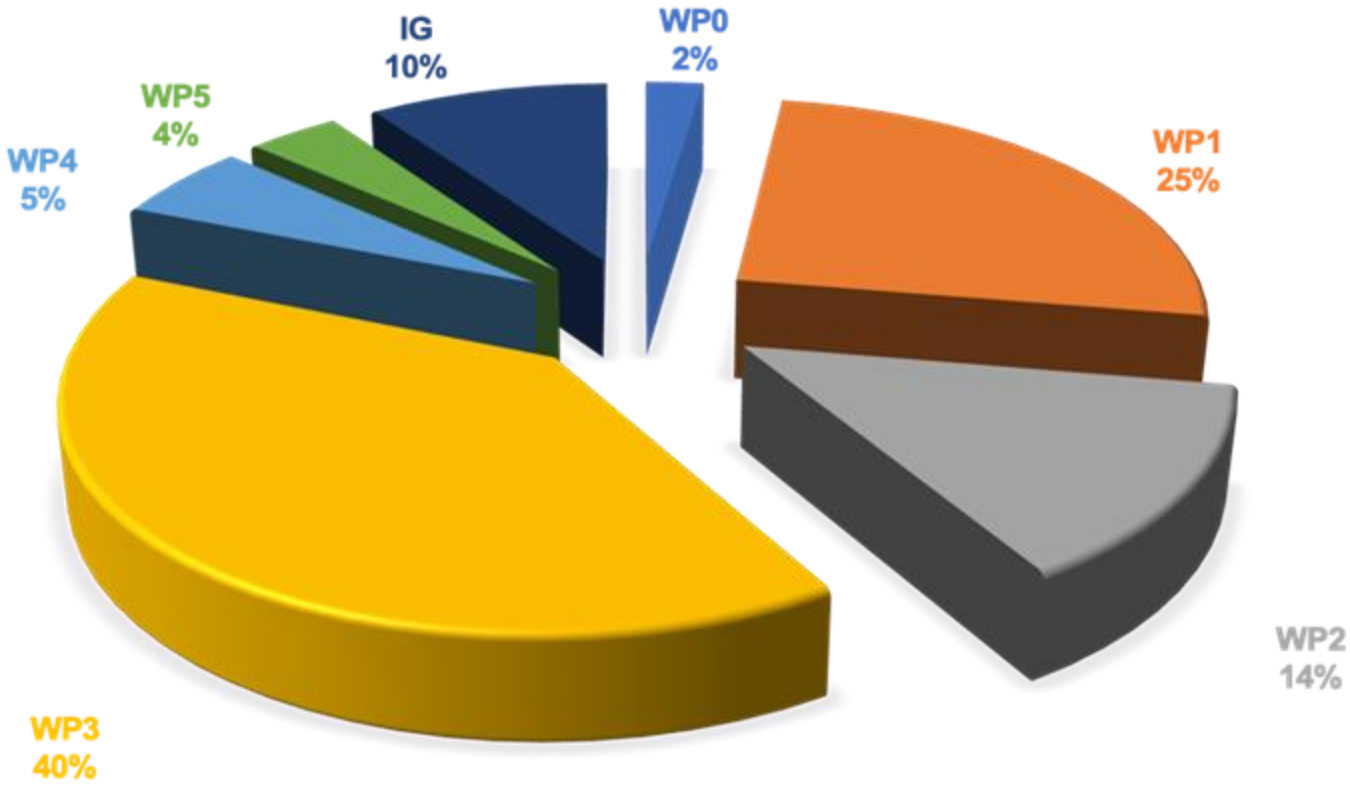
- **35 on 36 projects are currently active**. *No critical issues* have been reported except PADDLE project has been withdrawn . Generally, an internal project milestone is completed every three months, with the corresponding SAL payment.
- Some companies and universities, following our request, **are working on a convergent approach and results mainly for visualization ad ML techniques.** All activities **complement the work carried out in the WPs**.
- **All projects will be completed by November 2025**

Fund distributions

⇒ **SME**: **2.85 MEuro** (**57%** of the global fund)

⇒ **South Investements**: **2.01 MEuro** (**40%** of the global fund - ***50%*** first trance)

Spoke 3 Recruitment: person/years/WP



TOTALE		
WP	Posizioni	Annualita
WP0	1	2.5
WP1	12.5	27.5
WP2	6.5	15
WP3	20.5	43
WP4	3	6
WP5	2.5	4
IG	7	10.5
Totale	53	108.5

	Assunti	Donne	SUD
INAF	23	9	11
INFN	7	2	2
Uni TOV	6	2	0
Uni TS	2	2	0
Uni CT	3	1	3
Uni TO	3	1	0
SISSA	3	2	0
SNS	6	1	0
	53	20	16
		38%	30%

CN-2…il Centro Nazionale … che verrà!!
Post 2025

➔ **VISION**

Superamento degli Spoke ma possibili ancora verticalizzazioni su aree tematiche
Aree tecnologiche trasversali (AI, Big Data, Cloud ecc)

➔ **FINANZIAMENTO**

Diversificazione dei Canali: Pubblico MUR, Privato, Europeo
AI Factory
Infratsruttura ➔ canali diretti MUR ed EURO-HPC

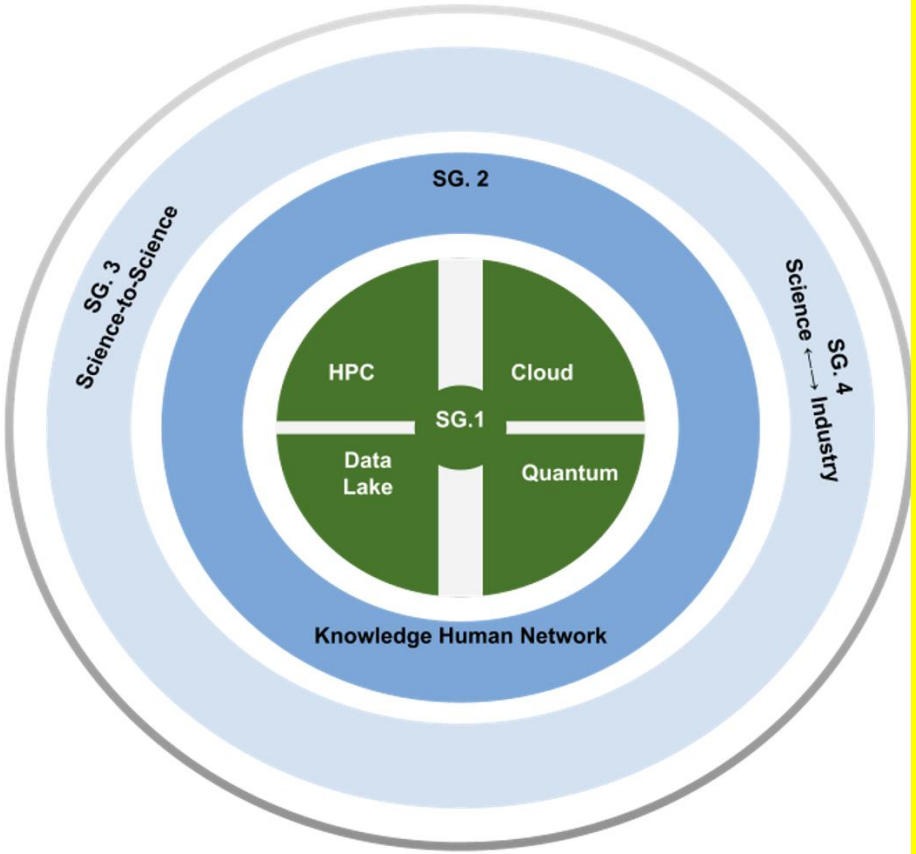
➔ **SUPPORTO** alla Ricerca di base e alla ricerca delle diverse aree degli spoke

➔ **COLLABORAZIONE** tra enti di **Ricerca e Impresa**

➔ **FORMAZIONE**

CN-2 VISION Spoke Board

Scientific Goals



SG.1 - Creazione e mantenimento dell'Infrastruttura Nazionale Italiana Permanente per il Calcolo, i Dati e il Quantum Computing
Sviluppare, implementare, mantenere ed espandere l'infrastruttura CN-2 per supportare un'ampia gamma di applicazioni scientifiche e industriali.

SG.2 - Creazione e mantenimento di una rete di conoscenze di specialisti
Costruire una rete di esperti per promuovere il progresso nei settori dell'HPC, dei Big Data e del Quantum Computing attraverso collaborazioni solide e programmi educativi efficaci.

SG.3.1 - Facilitare la collaborazione tra diverse comunità scientifiche
Affrontare le sfide comuni nei settori dell'HPC, dei Big Data e del Quantum Computing creando una piattaforma per esperti e scienziati di settore, favorendo lo scambio di idee, la condivisione di risorse e la ricerca collaborativa. Questo ambiente faciliterà le collaborazioni interdisciplinari, essenziali per innovazioni significative.

SG.3.2 - Rafforzare la collaborazione tra settori scientifici e industriali
Promuovere la cooperazione tra ricerca e industria su sfide comuni relative alla gestione e all'elaborazione di HPC e Big Data. Ciò contribuirà a colmare il divario tra ricerca e mercato attraverso l'implementazione di progetti a medio-lungo termine con rilevanti implicazioni industriali. Le iniziative strategiche sono fondamentali per allineare l'Italia alle nazioni leader nella ricerca applicata.

SG.3.3 - Favorire la collaborazione tra i settori industriali
Stimolare la cooperazione tra le industrie su problematiche condivise relative alla gestione e all'elaborazione dei Big Data e dell'HPC.

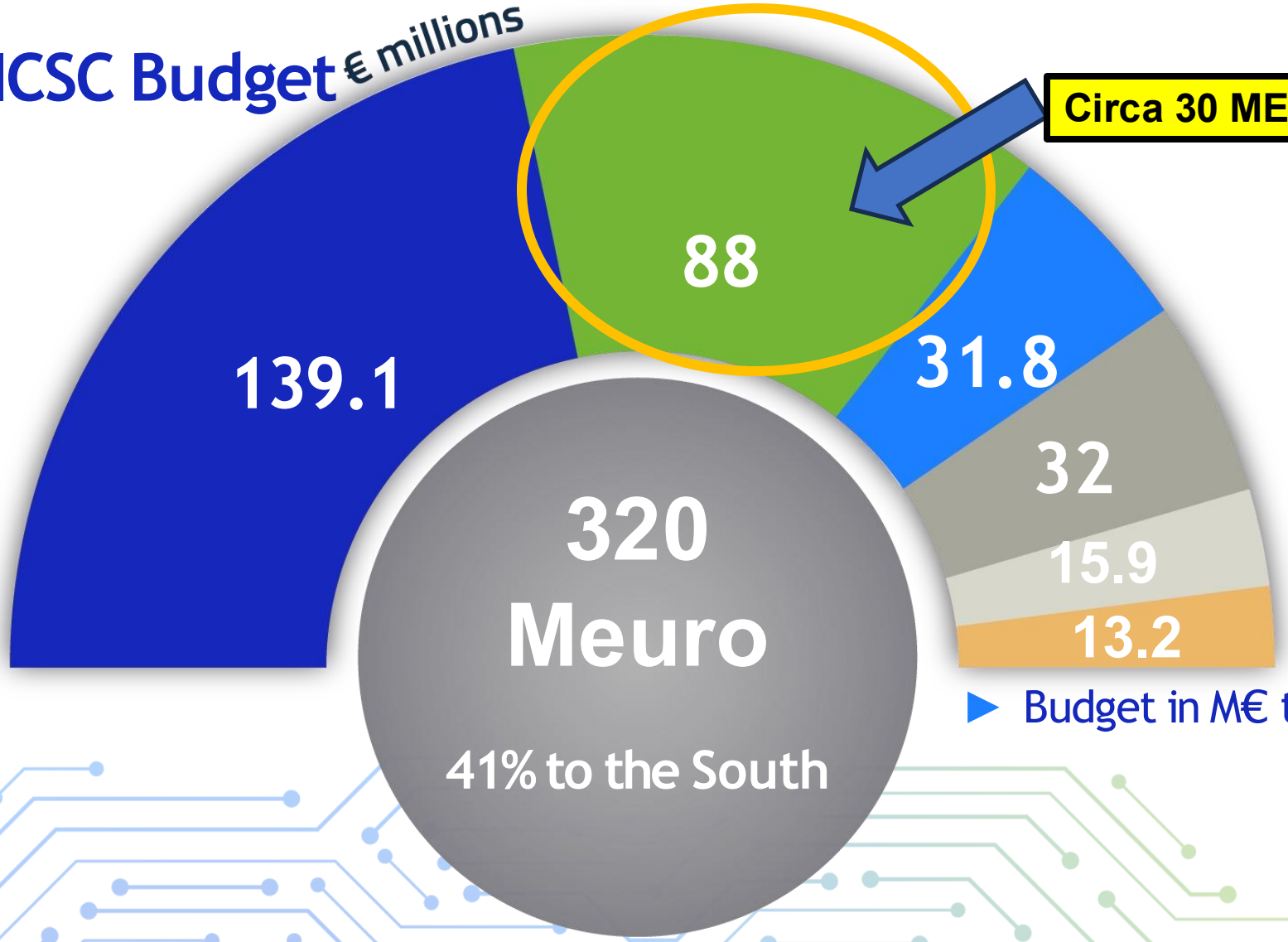
SG.4 - Impatto sulle politiche
Il CN-2 sarà l'ecosistema di riferimento a livello governativo nella definizione di strategie di medio-lungo termine relative al calcolo su larga scala.

- **Finanziamento specifico EU ➔ AI Factory ➔ 430 MEuro**

- The European High Performance Computing Joint Undertaking …. **The IT4LIA AI Factory project**, proposed by Italy with the participation of Austria and Slovenia, was selected. The initiative is co-funded by the Italian **Ministry of Universities and Research**, the **National Cybersecurity Agency (ACN)**, the **Emilia-Romagna Region**, the **CINECA consortium**, the **National Institute for Nuclear Physics (INFN)**, the **ItaliaMeteo Agency**, the **Italian Institute of Artificial Intelligence for Industry (AI4I)**, and the **Bruno Kessler Foundation (FBK)**. The project, coordinated as the hosting entity by CINECA, also involves other national institutions, including the **ICSC Foundation – Italian Research Center on HPC, Big Data, and Quantum Computing.**

At the heart of the project is the creation of an **advanced supercomputer**, optimized for artificial intelligence, to be installed at the **Tecnopolo Manifattura in Bologna**—a European hub for supercomputing, big data, AI, and quantum computing. This infrastructure, among the most powerful in the world, will become a strategic asset for positioning Italy as a leader in AI computing in Europe. The project also includes the development of **applications and services aimed at building an ecosystem that supports businesses.**

ICSC Budget € millions



- Budget in M€ to be spent in the period **2022-25**

Finanziamento POST PNRR *Legge Finanziaria 2025 Art 1 comma 579-581*

579. **Il Ministero dell'universita' e della ricerca sostiene le attivita' dei centri nazionali e dei partenariati estesi**, nonche' le iniziative di ricerca per tecnologie e percorsi innovativi in ambito sanitario e assistenziale del Piano nazionale per gli investimenti complementari al PNRR, al fine di consentirne il consolidamento nel tempo e la sostenibilita' economico-finanziaria al termine del periodo di attuazione del PNRR. ….

580. Con decreto del Ministro dell'universita' e della ricerca, sentiti gli altri Ministri interessati, da adottare entro novanta giorni dalla data di entrata in vigore della presente legge, **sono stabiliti gli indicatori chiave di prestazione** di cui al comma 579 nonche' le modalita' per il monitoraggio del raggiungimento degli obiettivi previsti dagli indicatori stessi e la rendicontazione delle spese sostenute.

581. Per le finalita' di cui al comma 579, nello stato di previsione del Ministero dell'universita' e della ricerca e' istituito un apposito fondo con una **dotazione di 150 milioni di euro per ciascuno degli anni 2027 e 2028.**

582. Con decreto del Ministro dell'universita' e della ricerca, da adottare **entro il 30 giugno di ogni anno, sono individuati i centri nazionali e i partenariati estesi,** nonche' le iniziative di ricerca per tecnologie e percorsi innovativi in ambito sanitario e assistenziale, in possesso dei requisiti di cui al comma 579, **ammessi al riparto delle risorse di cui al comma 581.**

150 Mln/euro → circa 20-30 Meuro/anno attesi per il CN-HPC 2027 e 2028

CN – HPC ….. Post 2025.. Concludendo

COME NON DISPERDERE LE COMPETENZE FORMATE?

- **Abbiamo acquisito competenze HPC e Big Data per le sfide astrofisiche e della fisica fondamentale**
- **NECESSARIO Valorizzare l'esperienza (capitale umano) acquisita nel Centro Nazionale**

Quali Prospettive?

1. **Stabilizzazioni e nuovo piano di assunizione negli enti**

2. **Bando PNR DD 307** (18 mesi, inizio entro il 2025) **85% delle risorse destinate al SUD**
   → **Prevedibili 10-15 Mln di euro di finanziamento per l'HUB XC (solo alcuni partner sono ammissibili al finanziamento). Partner dello Spoke 3 partecipanti: INAF e INFN**

3. **Finanziaria 2027–2028**:
   → Attesi **20–30 M€/anno per CN-HPC**
   → Auspicata la **continuità di tutte le istituzioni afferenti**