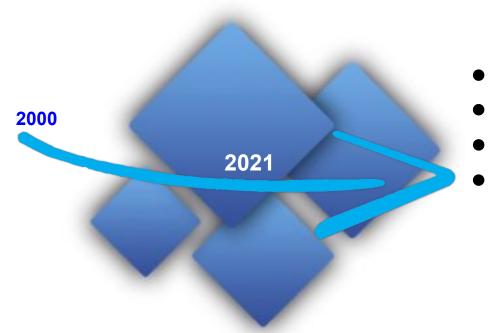




Big data
e nuove tecnologie informatiche:
IT Center of Competence







- HPC/HTC ed ExaScale
- Artificial Intelligence
- Visualizzazione Scientifica
- Calcolo Distribuito: Cloud e Container



HPC/HTC ed ExaScale

- Paradigmi: MPI, OpenMP, OpenACC e CUDA
- Codice parallelo per il solutore della AVU della missione GAIA (GAIA-DPCT) OATO
 - o produzione sul sistema Marconi 100 del CINECA
 - in fase di porting su infrastrutture di pre-Exascale -> su sistema LEONARDO (CINECA-EURO HPC JU)
 - porting sperimentale su infrastrutture HPC innovative (ARM & FPGA)
 progetto EuroExa (code-design Exascale vedi ExaCT G.Taffoni)
- Porting di algoritmi di Machine Learning su piattaforme HPC NVIDIA DGX
- Applicazioni di Edge Computing
 - Utilizzo di piattaforme di calcolo ibride : ACAP (GPU-FPGA)



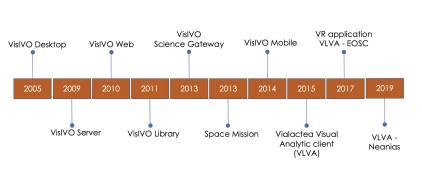


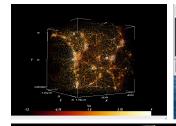
- Al per identificazione e classificazione di
 - sorgenti compatte/estese in mappe radio
 - MoU INAF-OACT -Università di Malta
 - Collaborazione INAF-OACT -Università di Catania
 - H2020 NEANIAS/PRIN CIRASA
 - Diversi modelli e framework di object detection attualmente testati (Mask-RCNN, U-Net, Tiramisù) su diverse infrastrutture GPU (CINECA DGX A100, V100, K40)
 - Preparazione dataset & applicazione modelli
 - Dati Radio di precursori SKA: ASKAP & MeerKAT
- AI & HPC
 - Collaborazione con HPC4AI (Università di Torino)
 - sistemi eterogenei (CPU-GPU)
 - Attività previste:
 - porting di algoritmi di estrazione sorgenti su GPU
 - training di modelli Al su sistemi eterogenei
- Al & Calcolo Distribuito Al per Maintenance Predittiva in CTA e ASTRI
 - modelli: sequence models (LSTMs)
 - LSTMs in PyTorch

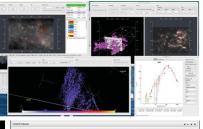




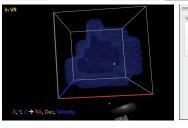
Visualizzazione Scientifica















VisIVO Framework Cosmologia

- » Desktop
- Server
- » API
- » Mobile
- » Web / Science Gateway

Visual Analytic Formazione Stellare

- » Desktop
- » Cloud
 - **➤** EOSCPilot
 - ➤ NEANIAS

Virtual Reality Radioastronomia

- » Trasferimento Tecnologico
 - 3DTelC (Erasmus+ Geophysics community)
 - ArGo Divulgazione scientifica MUR) (UniMIb collaboration)
- » iDaVIE-v MoU IDIA (Bilaterale

ITA-SA)

ļ



Calcolo Distribuito: Cloud e Container

- Architetture CaaS (Container as a service)
 - Algoritmi di maintenance predittiva in CTA
 - Servizi di logging e monitoring su dataset ad alto throughput (Big-Data)
 - Efficienza Energetica
 - Fault Tolerance & Availability
 - Deployment : Kubernetes
- Containerizzazione di applicazioni scientifiche (data analysis, monitoring & control, scientific visualization)
- Porting su diverse infrastrutture HPC o cloud
 - LOFAR.it WG leadership in pipeline dev.
 - LOFAR, ASKAP Containers per pipeline software radio
 - OpenStack/Google Cloud
 - o ASTRI/CTA/SKA Dish Containers per sistemi di controllo e monitoraggio
 - VLVA Containers per applicativi di visual analytics
 - NEANIAS Sviluppo di applicazioni web a microservizi su cloud
- Sviluppo di gateway scientifici per l'astrofisica (VLVA, CTA)
- CTA Metodologie di Authentication & Authorization sul progetto AARC2



Novel EOSC services for Emerging Atmosphere, Underwater & Space Challenges



BIT.CC

https://www.neanias.eu/



20 partner da 9 stati

- Call INFRAEOSC-02-2019 Prototyping new innovative services
- Progettazione di servizi tematici innovativi
- Consegna servizi alle comunità degli utenti finali
- Integrazione servizi su piattaforma EOSC

IMPLEMENTAZIONE DI UN **ECOSISTEMA COLLABORATIVO** E **MULTIDISCIPLINARE** PER

FAVORIRE LA CONDIVISIONE DELLA CONOSCENZA









































Bilancio complessivo € 5 597 025

Contributo EU per INAF: € 569 187,50

Durata: Nov 2019 - Nov 2022



SPACE e architettura service-oriented



https://catalogue.neanias.eu/







SPACE · VIS

SPACE · MOS

SPACE · ML

Servizi "Core" Supporto al Integrazione alla Cloud Intelligenza Visualizzazione ed EOSC Artificiale EUROPEAN OPEN SCIENCE CLOUD Infrastrutture di Ricerca



Visualizzazione in Astrofisica

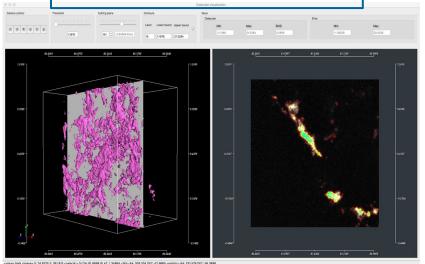


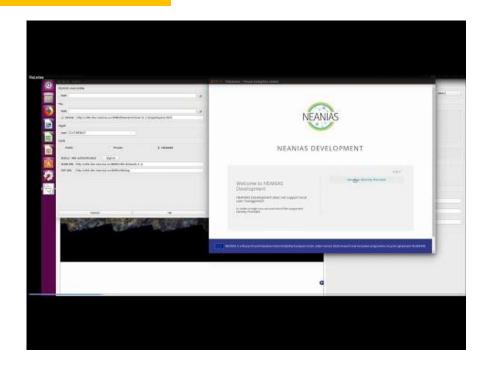
→ { Piu di 30 TB dati infrarosso e radio }

http://vlkb.neanias.eu:8080/

{ VIALACTEA }

- » Accesso a survey del Piano Galattico attraverso servizi VO e FAIR
- » Studio regioni di formazione stellare
- » Cataloghi di Filamenti e Sorgenti







Rilevamento di sorgenti

√{ segmentazione e caratterizzazione }

https://caesar.neanias.eu/



This is CAESAR

Welcome fellow astronomer! This is CAESAR (Compact and Extended Source Automated Recognition). To start using the service please authenticate using your NEANIAS credentials

Log i

Service powered by NEANIAS



{CAESAR}

- Estrazione di sorgenti compatte ed estese da mappe radio astronomiche
- » Processamento immagini non supervisionato
- Caratterizzazione sorgenti e post-processamento
- » Multi-thread & parallelizzazione multi-nodo per immagini grandi





MOSAICo



Metodologie Open Source per l'Automazione Industriale e delle procedure di CalcOlo in astrofisica







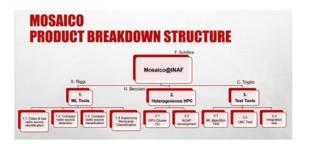


Obiettivo aziendale: consolidare la presenza nel settore tecnologico di appartenenza, e quella nel settore dei grandi progetti astrofisici, che offre una grande possibilità di sviluppo ad aziende di piccole e medie dimensioni.

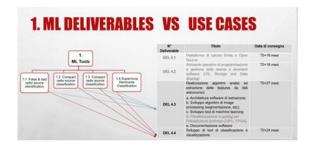


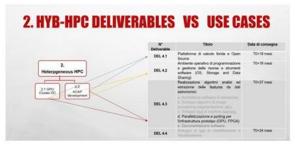
Obiettivo Scientifico: realizzare e testare sistemi di controllo automatizzati ed algoritmi di elaborazione di immagini basati sul Machine Learning, operanti su piattaforme di calcolo eterogenee ed ambiente Open Source.

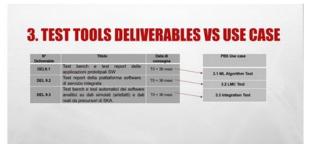
MOSAICo MANAGEMENT











On Schedule

15. Team Summary

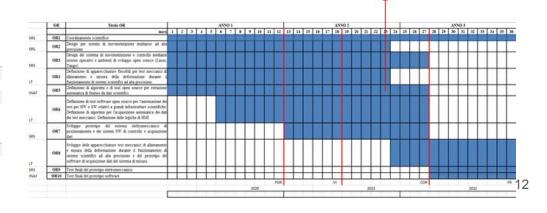
Personale INAF coinvolto
 Numero di partecipanti INAF al progetto: 18

Struttura	Nfte	NO	TI 21	TI 22	TI 23	TD 21	TD 22	TD 23	Nex	Extra
O.A. CATANIA	15	0	3.40	2.95	3.15	0.50	0.50	0.50	0	0.00
IRA BOLOGNA	2	1	0.20	0.20	0.20	0	0	0	1	0.10
Totali	17	1	3.60	3.15	3.35	0.50	0.50	0.50	1	0.10

Fondi a sostegno

21. Totale fondi a disposizione (dato aggregato, k€)

Certi 2021	Certi 2022	Certi 2023	Presunti 21	Presunti 22	Presunti 23	
129.5	0.0	0.0	180.0	200.0	50.0	

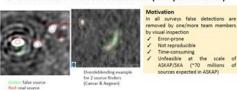


MOSAICo Science Use cases and Results



Supervised Methods

1 - Real vs false source identification (Compact Sources)



- False compact source detections reported in catalogs can reach up to ~10-20% in complex background environments (like the Galactic Plane)
- False detections are mainly due to:
 - imaging artefacts
 - 2. over-deblending of diffuse and extended sources or failed fits
- · Spurious sources partially removed with quality cut selection
 - > Slightly better rejection performances with simple MLP NN (3 parameters)
 - > Imaging artefacts not well identified, need additional features (e.g. defined by a user or CNN)

2) Compact source detection and improved fitting





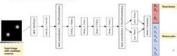




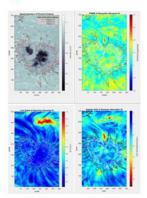
e 1 blended sourc

- Idea: use a deep network to learn the background patterns and achieve a more precise
- estimation of source parameters (NB: need to handle blended sources)
- What do we want to achieve?
 - Preciar sub-pixel (s,y) location (this is what we achieve with fitting methods) at minimum
 2d ellipse pars = flux amplitude (not just bounding boxes)
 - High completeness and reliability: "100% above 10 sigma and >10% above 5 sigmas (this is what we
- achieve with fitting methods)

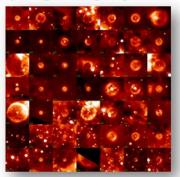
 Train a network to learn:
- Object sourcement for classification (true/false) (CLASSIFICATION)
- Source centroid (x,y) (needed to distinguish multiple sources) (REGRESSION)
- Potentially the network should be able to learn the complex background features in data providing also a better source parameter (x, x), bresi, lenin, theta) extinuation writ tradition methods (based on 2D gazzatan methods fitting) (BCGRESSON)



UnSupervised Methods



Semantic Multispectral Segmentation by using Self Org Maps



Etherogeneous Processing Platform

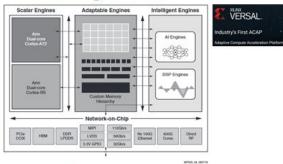


Figure 4: Xilinx Versal ACAP Functional Diagram

Deploy ML tools su piattaforme hardware di edge computing e near real time accelerator – Confronto delle performance e parametri prestazionali.







Attività: In Itinere; Data inizio: 2000

- Stima dell'inviluppo complessivo di FTE INAF dall'inizio al 2030
- **Stima delle FTE INAF** a Tempo Indeterminato dall'inizio al 2030
- Stima dell'inviluppo complessivo di FTE

(includendo tutti i partners dall'inizio al 2030)

200.0

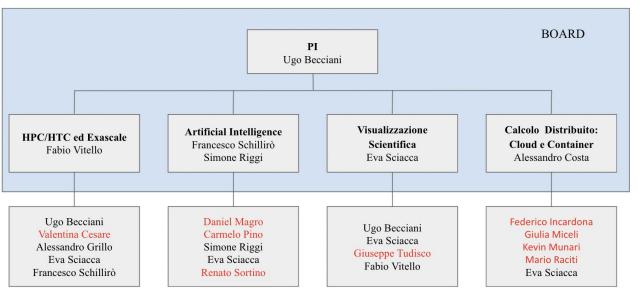
150.0

70.0

2000 2021



TEAM



7 Ricercatori e Tecnologi TI 9 Unità di Personale 7 AdR/Borsista 2 PhD

4 Working Group

BOARD Decisionale composto da 6 CO-Is





























Ruolo in INAF e LEADERSHIP

- Centro di Competenza ad alta specializzazione nel campo IT
 - Applicazione di tecnologie innovative per l'analisi dei BigData R&D
- Collaborazioni e condivisioni di ruoli di leadership
 - sedi INAF: IAPS,OATO,IRA,OATS,OACA,
 - sedi e istituti universitari: Malta-ISSA, IDIA & Università di Cape Town -SA (UCT), Università di Milano Bicocca, Università degli studi di Catania, Cineca, Consorzio COMETA.
- Infrastrutture Coinvolte
 - o CHIPP, CTA, ASTRI, CINECA, EOSC, GARR
- Schede correlate ("figlie"):
 - NEANIAS (BIT.ICC-1) MOSAICO (BIT.ICC-2)
- Schede con interazione:
 - EuroExa,CHIPP, EcoGAL, PROGRESS, Gaia G-DPCT, Lofar-It, PLEIADI, CIRASA, RAGA

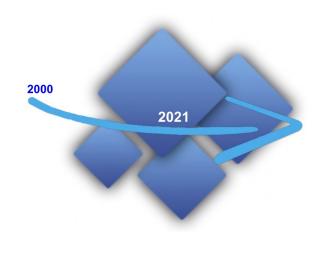




Fondi

cna	€ 25.000,00
astrocomp	€ 15.000,00
datagrid miur cofon prin	€ 20.000,00
trigrid	€ 200.000,00
pi2s2	€ 300.000,00
egiinspire	€ 30.000,00
scibus	€ 250.000,00
edgi	€ 30.000,00
muoni	€ 350.000,00
gaia	€ 50.000,00
vialactea	€ 106.000,00
aeneas	€ 75.000,00
erasmus+	€ 34.000,00
argo	€ 35.000,00

€ 350.000,00
€ 40.000,00
€ 146.000,00
€ 52.000,00
€ 15.000,00
€ 37.000,00
€ 145.000,00
€ 30.000,00
€ 160.000,00
€ 98.000,00
€ 45.000,00
€ 17.000,00
€ 108.000,00



Stima fondi acquisiti da INAF fino al 2020 (k€): 2500



Criticità

Centro Di Competenza Patrimonio Dell'INAF

Principale Asset di BIT.CC

Personale e relativo Know-How

Canali di finanziamento stabili da parte di INAF

Necessità di acquisire costantemente nuove competenze specifiche in **HPC**

 Risorse adeguate da investire nella sperimentazione

Al & Calcolo Distribuito e Visualizzazione

 Acquisizione di risorse GPU Based e infrastrutture su cui sperimentare architetture di CaaS (Container As a Service)

