Allegato 5 – Programma dello Spoke

The Spoke 3 "Astrophysics and Cosmos Observation" aims to

- **Objective 1**. High and Extreme performance computing, via developing software solutions for data analysis and numerical simulations able to effectively exploit HPC resources in the perspective of the Exascale era. (WP1, WP2, WP5)

- **Objective 2**. Big data processing and visualization, via adopting innovative approaches (e.g. Artificial Intelligence, inference via Bayesian statistics) for the analysis of large and complex data volumes and for their exploration (e.g. in-situ visualization), capable of efficiently exploiting HPC solutions. (WP3, WP4, WP5)

- **Objective 3**. High Performance storage, Big Data management, and archiving applying the Open Science principles and implementing them in the Big Data Archives. Experimenting and adopting novel technologies and computational approaches for fast and scalable I/O (WP4, WP5)

- **Objective 4**. Training and dissemination. Creating a community of scientists and code developers prepared to adopt and exploit innovative computational solutions. Fostering the engagement of the scientific community in Astronomy, Astrophysics and Astroparticle Physics.

The work plan of Spoke 3 is organized in 5 work packages (WP), whose goals and tasks are detailed below. An additional WP has been introduced (WP0) for the Spoke management.

- **WP0 – "Management"**
  This WP develops a sound coordination and management of the Spoke covering technical, administrative, legal and financial issues.
  Planned Tasks:
  - T0.1 – Overall coordination, governance, monitoring, communication flow and methods.
  - T0.2 – Data Management Plan
  - T0.3 – Training, dissemination plan and outreach activities
  - T0.4 – Cascade fundings, Innovation grants and key science projects

- **WP1 – "HPC Codes Enabling and Optimization"**
  This WP selects a number of codes that require intensive computational resources to face the next generation of scientific challenges and performs their redesign, reimplementation and optimisation in order to effectively exploit state-of-the-art HPC solutions.
  Planned Tasks:
  - T1.1 – Selection, Analysis and testing of codes, algorithms and programming models; development plan
  - T1.2 – Software Development, Refactoring and Optimisation according to the plan setup
  - T1.3 – [Integration, Verification and Validation

- **WP2 - "Design of innovative Algorithms, Methodologies, Codes toward Exascale and beyond"**

This WP identifies innovative algorithms and methodologies upgrading their capability to exploit, and scale on, the exascale and post exascale architectures, reintegrating the resulting improved features in codes, workflows and pipelines. The energy impact will also be specifically considered

Planned Tasks:

- T2.1 – Science cases definition, algorithms identification, parallelism level assessment and profiling
- T2.2 – Algorithms Co-design and methodologies to scale-up the capabilities of the algorithms and to find new innovative solutions
- T2.3 – Design of new architectural solutions aimed at the exploitation of post-exascale infrastructures (GPUs, FPGAs, Vector accelerators, NVM, HBM, ARM)
- T2.4 – Algorithms and methodologies integration into new big-data analysis applications

- **WP3 - "Big Data Analysis, Machine Learning and Visualization"**
  This WP develops a prototype framework of data analysis, based on Machine Learning (ML) and Visualization tools exploiting diverse computing platforms and combining them with exascale applications.. The framework will be tailored to the ACO-S community, identifying the use case best suited to tackle high-performance visualization tools and ML techniques. Furthermore, it copes with observational data coming from challenging large experiments (e.g. LOFAR, MeerKAT, SKAO data challenges, Large Simulations Datasets, CTA etc)

  Planned Tasks:
- T3.1 – Requirements from AAA community
- T3.2 – Innovative Machine Learning
- T3.2 – HPC/Cloud Visualization Services

- **WP4 - "Big Data Management, Storage and Archiving"**
  WP4 will analyze, explore, standardize and store data of different collections, characterizing them in the appropriate way in order to facilitate the respect of the FAIR principles and enable users to benefit from an innovative storage and archiving platform. A distributed archive infrastructure with hot and cold storages and proper access tools will be implemented or customized by existing ones, respecting the interoperability directives (RDA, IVOA, …). The requirements of the storage in terms of identification of data collection characteristics (data models, data dimension, accessibility methods, cold or hot storage etc…) will be analyzed, defining and implementing suitable technical solutions.

  Planned Tasks:
- T4.1 – Data management, Standardization, and Interoperability
- T4.2 – Local and Distributed Long and Short-term Storage Optimization
- T4.3 – Archive and Repository Definition and Implementation

- **WP5 - "HPC Services and Access"**
  The WP is in charge of managing, maintaining and deploying an integrated environment providing the tools for the efficient development of the work described in WPs 1-4

  Planned Tasks:
- T5.1 – Collaborative software development, management and continuous integration platform
- T5.2 – Design, implementation and validation of an interoperable service architecture
- T5.3 – ML and Visualization enabling services deployment and HPC/Cloud integration