



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani

PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing



Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

WP5 @ Spoke2: Status report

Elvira Rossi (Università Federico II di Napoli & Daniele Spiga (INFN – sez. Perugia

"ICSC and Spoke2 - Where Are We Now?" - Physics Dept and INFN, Catania - 11 Dec 2024

Agenda of the Talk

- WP5 quick overview
- Activity, achievements and issues
 - during the past 12 months
- Few Thoughts on the past two years and Future Priorities Some thoughts after 2 years and next priorities

WP5 Objectives and high level view

Proposes:

- 5.1 Support of the adaptation of existing applications on the data-lake distributed infrastructure, and via innovative computational models
- 5.2 Competence center for the design, implementation and test of computing models

Objectives:

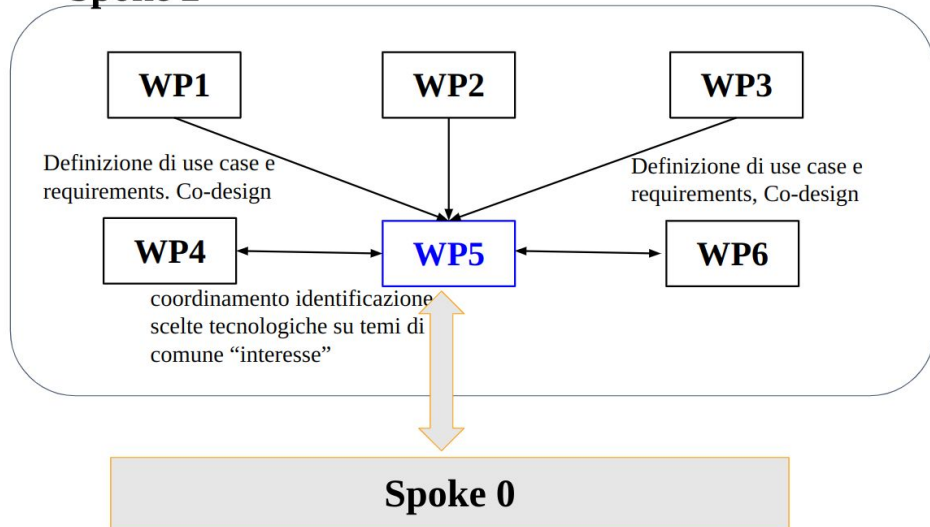
- ✓ 05.1: document and report best practices for integrations with the CN datalake
- ✓ 05.2 prepare tools to ease integration with the CN infrastructure
- ✓ 05.3 Offer support for transitioning the computing models
- ✓ 05.4 organize training opportunities open to external users

WP5		
	Nome	Mail
National coordinators	Elvira Rossi Daniele Spiga	elvira.rossi@unina.it daniele.spiga@pg.infn.it
Institute coordinators		
INFN	Daniele Spiga	daniele.spiga@pg.infn.it
UNIMIB	Mattia Bruno	Mattia.Bruno@mib.infn.it
UNINA	Elvira Rossi	elvira.rossi@unina.it
UNITS	Andrea Bressan	andrea.bressan@ts.infn.it
UNIBO	Alessandra Fanfani	Alessandra.fanfani2@unibo.it
UNIFE	Luca Tomassetti	luca.tomassetti@unife.it

WP5: indico pages & email list and meeting

Spoke 2: <https://agenda.infn.it/category/1774/>

Spoke 2



Limiting the number of meetings mainly in favor of the **more operative phase** (but also because of a limited active participation)

- Focused on flagships
- Co-hosted meetings

WP5: <https://agenda.infn.it/category/1781/>

Month	Date	Event
October 2024	23 Oct	Biweekly Meeting Spoke2 - WP2.5
October 2024	02 Oct	Biweekly Meeting Spoke2 - WP2.5
July 2024	05 Jul	Mini-Workshop on Data Management
June 2024	07 Jun	Biweekly Meeting Spoke2 - WP2.5
March 2024	27 Mar	Biweekly Meeting Spoke2 - WP2.5
February 2024	02 Feb	Biweekly Meeting Spoke2 - WP2.5
January 2024	17 Jan	Biweekly Meeting Spoke2 - WP2.5
December 2023	13 Dec	Biweekly Meeting Spoke2 - WP2.5
November 2023	22 Nov	Biweekly Meeting Spoke2 - WP2.5
November 2023	10 Nov	Biweekly Meeting Spoke2 - WP2.5
November 2023	08 Nov	WPS - Analysis Facility
November 2023	03 Nov	Biweekly Meeting Spoke2 - WP2.5
October 2023	13 Oct	Biweekly Meeting Spoke2 - WP2.5
July 2023	07 Jul	Biweekly Meeting Spoke2 - WP2.5
June 2023	09 Jun	Spoke 2 - WP5 & WP2 - ATLAS
May 2023	24 May	Biweekly Meeting Spoke2 - WP2.5
April 2023	28 Apr	Biweekly Meeting Spoke2 - WP2.5
April 2023	12 Apr	Biweekly Meeting Spoke2 - WP2.5
March 2023	22 Mar	Biweekly Meeting Spoke2 - WP2.5
March 2023	09 Mar	Meeting Spoke2 - WP2.5

25 events in 2022-24: general WP5, operational and work meetings with WP5 contacts and WP5 enthusiast
Many more informal discussions among WP5 and other WPs peoples

Monthly/Biweekly/Weekly meetings as needed

[Mini-Workshop on Data Management in July 2024](#)

July 2024

05 Jul Mini-Workshop on Data Management

Email list:

cn1-spoke2-wp5-all@lists.infn.it

Link to subscribe to the email list:

<https://lists.infn.it/sympa/info/cn1-spoke2-wp5-all>

Please contact us if you are interested in contribute to WP5

- **Training activities and documentation**
 - Schools, Heterogeneous resources exploitations, Data management
- **Technical support to Spoke2 Flagships**
 - WP5 is not owning them. It provides support to scientific WP flagships
- **WP5 and Innovation grants**
 - Where we exploit our technological assets
- **Contribute to the overall CN infrastructure**
 - Aka to Spoke-0 infrastructure

WP5 main assets

Or..

Where we spent our effort in the past 12 months

WP5 exploitation activities include participation in various National and International conferences (CHEP, ACAT, Communities Workshops, WLCG...)

Training in 2024 and plans for the upcoming events

SOSC (School on Open Science Cloud) a only in person events held in Bologna the past week!

- 28 students, not only Spoke2, not only ICSC!

A dedicated training on Data Access and Data Sharing (Data Management in the Data Lake)

- A fully online event

SOSC 2024 Sixth International School on Open Science Cloud

2-6 Dec 2024
University of Bologna, Department of Physics and Astronomy
Europe/Rome timezone

- Overview
- Timetable
- Contribution List
- My Conference
- My Contributions
- Registration
- Travel & Accommodation
- Venue

The theme of the sixth International School on Open Science Cloud is "Computing Models for Scientific Experiments"

The SOSC24 is organized by



The 6th edition of the international from 02 to 06 December 2024, T "Augusto Righi" of the University of Perugia and the ICSC Foundat

[Details here](#)

Mini-Workshop on Data Management

5 July 2024
Europe/Rome timezone

- Overview
- Timetable
- Contribution List
- My Conference
- My Contributions
- Registration
- Participant List
- Scientific Committee



[Details here](#)

Workshop on "Quasi-Interactive Analysis of Big Data with High Throughput"

8-10 Jan 2025
Bologna
Europe/Rome timezone

- Overview
- Timetable
- Contribution List
- Registration
- Participant List
- Visit to the technopole

Contact

- ✉ tommaso.dotelevi@bo.it
- ✉ francesco.giuseppe.grev
- ✉ lenzop@fn.infn.it
- ✉ daniele.spiga@pg.infn.it
- ✉ elly@na.infn.it



[Details here](#)

Upcoming Events

In our pipeline:

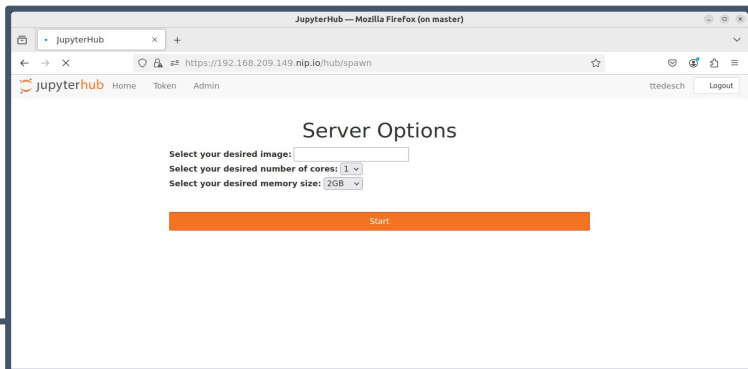
- Co-organizing the high rate workshop 8-10/01/2025: **PLEASE REGISTER ASAP if interested**
- Hight Technical meeting on commonalities between WP5 co-developed platforms **TBD**
- Co-organizing GPU and FPGA events → WP4/Inter Spoke

WP5 and Spoke2 Flagships

Key technologies integrated so far:

- Identity federation (IAM, JWT)
- Orchestration (K8s)
- DAG/Workflow management (DASK)
- Interfaces/access (GUI, API)
- Data Management (experiment RUCIO)

“Extended Computer Vision at High Rate” (WP6)



High Rate Analysis From (WP2)

Warning
Work in progress documentation! In case of comments, questions or issues, please contact [Francesco G. Gravili](#) or refer to the GitHub repository linked on top of the page.

Acknowledgements

The content of this guide reflects the efforts of many people: many thanks to everyone who contributed to it!

Example on INFN analysis facility with systematic variations

Here we are going to perform of the VBS with a lepton and a tau as final state reconstruction on an OpenData sample of W+Jets events.

Basic imports and definitions of number of parallel tasks and the Dask scheduler port

```
(1) import ROOT
import os
from dask.distributed import Client

[ ] #fill this cell with "+" button of Dask labextension

Declare custom functions defined in utils.h
```

```
(2) text_file = open("utils.h", "r")
data = text_file.read()

def my_initialization_function():
    ROOT.gInterpreter.Declare("{}".format(data))
ROOT_ROOT.Experimental.Distributed.initialize(my_initialization_function)

Define chain of rootfiles to analyze
```

```
(3) chain = [
    "#root://eospublic.cern.ch/eos/root-oss/benchmark/CMSpenDataM1ggs7au/au/W1Jets10Mu.root",
    "#root://eospublic.cern.ch/eos/root-oss/benchmark/CMSpenDataM1ggs7au/au/W2Jets10Mu.root",
    "#root://eospublic.cern.ch/eos/root-oss/benchmark/CMSpenDataM1ggs7au/au/W3Jets10Mu.root",
]
```

Do analysis

More on dedicated talks

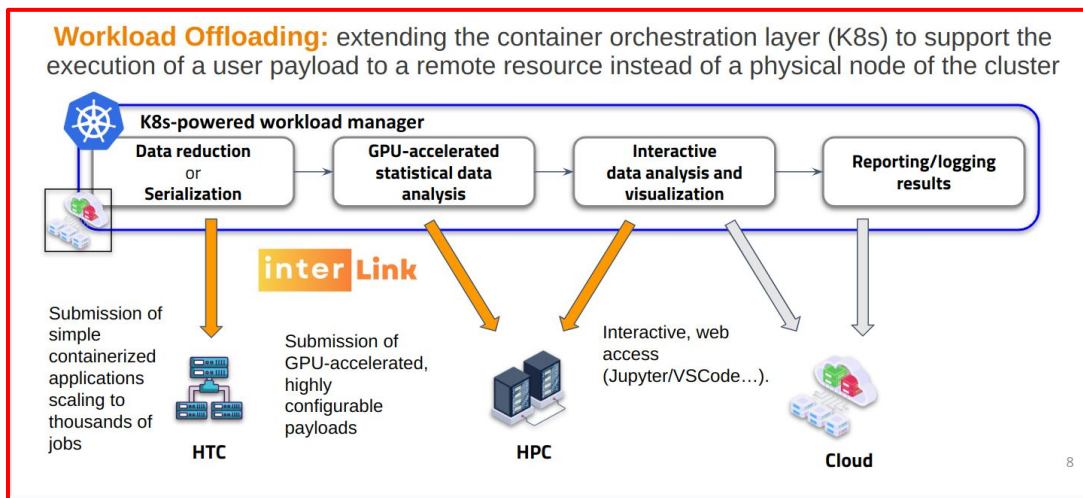
- [Francesco G.](#)
- [Tommaso. T](#)
- [Alessandro B.](#)

Advanced Machine Learning. Flash Simulation and bleeding edge applications (WP2)

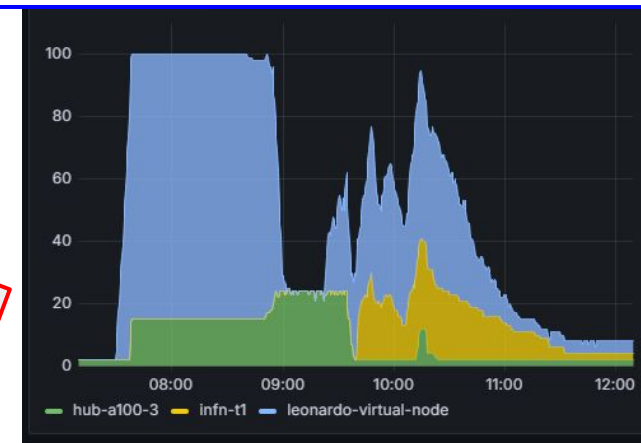
Support to Flagships (cont)

Key technology

Heterogeneous resources (k8s+interlink)



Works



On October 25th, we run a first workflow combining CPU resources from **CNAF Tier-1**, **Leonardo** and a **local node**.

Test limited to CPU-only payloads.

Offloading to Leonardo booster (with GPU payloads) coming soon.

NOTE: Selected as demonstrator for the Spoke0 CINECA-INFN Federation PoC initiative

- Adopts cloud-native platform. “Mature enough” and k8s enabled
- Requires by design access to **heterogeneous resources**
 - **Cloud and HPC** (Leonardo Booster)

More on dedicated talks

[Anderlini’s talk at WP5](#) & [Anderlini’s talk @ WP2](#)

Federated Data Management (DM)

Enable and enhance the interoperability between heterogeneous storage solutions in the ICSC distributed infrastructure

- big traditional HPC centres, Cloud, Grid/HTC

Key technology

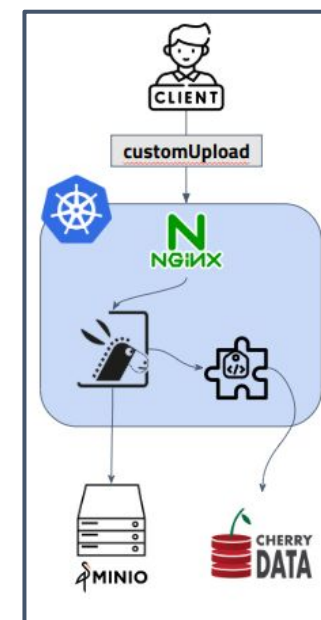


Strict collaboration with Spoke0 related activities

To promote the adoption of RUCIO based ecosystem in order to grant to the user “The ability to store and retrieve data without caring about to know where the data actually is”

Synergies with Innovation Grant

Enhancing RUCIO ecosystem.
A ongoing synergy with **interoperable Data Lake (IDL) project**



More on dedicated talks

[Luca. P](#) ; [Nicolo' M.](#); [Domingo R](#)

WP5 and the CN infrastructure (aka interactions with Spoke0)

All WP5 supported flagship are using the RAC provided resources

- INFN-Cloud @CNAF, INFN-Cloud @Napoli and HPC Leonardo @CINECA
- IG supported at INFN-Cloud@Catania
- High Level cloud-native service successfully deployed

ICSC dedicated IAM under integration

- People is registering
- Groups structure created

Activated the collaboration on several topics

- Data Management: we use, promote and enhance the Spoke0 adopted solutions
- Providing support to test and validate the ICSC PoC federation middleware

The ICSC IAM instance

The screenshot displays the ICSC IAM instance interface. The top part shows the login page with the ICSC logo and the text "Welcome to icsc". Below this, there are options to "Sign in with" (including the INFN logo), "Local credentials", and "Not a member?". A green button labeled "Apply for an account" is visible. The bottom part of the screenshot shows a "Group requests" section with a list of groups:

- icsc/admins/fastcv
- icsc/admins/spoke2analysisfacility
- icsc/priv-admins/fastcv
- icsc/users/fastcv
- icsc/users/spoke2analysisfacility

End user Documentation ([here](#))

The collage shows three overlapping screenshots of the ICSC documentation website. The top-left screenshot displays the 'Resource Access' page, which includes instructions on how to access the platform and a section for 'Account Requests'. The middle-left screenshot shows the 'Examples' page, listing various analysis types like 'CMS VBS-like analysis' and 'FCCee analysis'. The rightmost screenshot is the 'Troubleshooting and Notes' page, featuring a search bar and several warning and attention boxes regarding storage and Docker configurations.

Developed sw, Artifacts, new components ([here](#))

This screenshot shows a list of GitHub repositories under the 'ICSC-Spoke2-repo' organization. The repositories listed are: 'HighRateAnalysis-WP' (Jupyter Notebook, 1 star, updated Oct 18), 'idl-parser' (Shell, 0 stars, updated Apr 8), 'wp5-custom-images' (0 stars, updated 2 hours ago), and 'idl-rucio' (Python, 0 stars, updated 26 days ago). Each repository entry includes a star icon and a 'Star' button.

Summary, Next Steps and thoughts after 2 years of Spoke2/WP5...

Most of the initial ideas have been implemented

- Different level of maturity; different level of interests from users (aka WP[1-3] and 6)

Interaction with Spoke0 and further integration of “our” services:

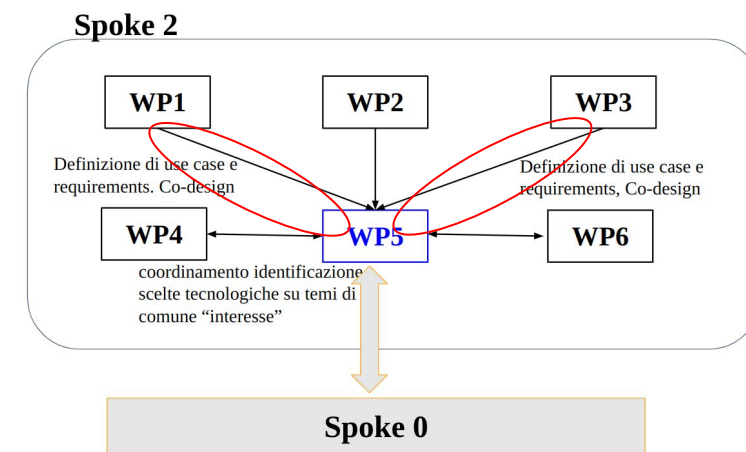
- We expect to contribute to the evolving portfolio of services. Share with other communities, grant a longer term maintenance; ease the operations

Keep co-designing solutions to enhance the “resources heterogeneity” integration

- I.e Bubble HPC; FPGA based systems

Enhance the harmonization of our technical solutions

- I.e. can we reduce the “complexity” and provide common building blocks



We didn't manage to establish all the connections we had hoped for

- Lack of communications? (i.e. DM solutions)
- Lack of effort ?
- All ok as it is now?

Backup Slides

WP5: People

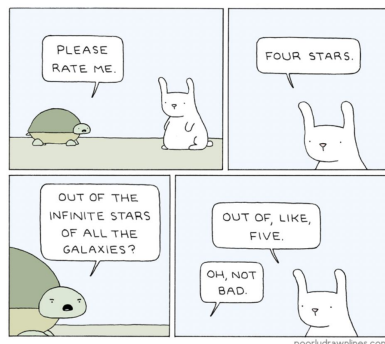
Contributors: INFN, UNIMIB, UNINA, ROMA1, UNITS, UNIBO, UNIPD, UNIFE

Local coordinators

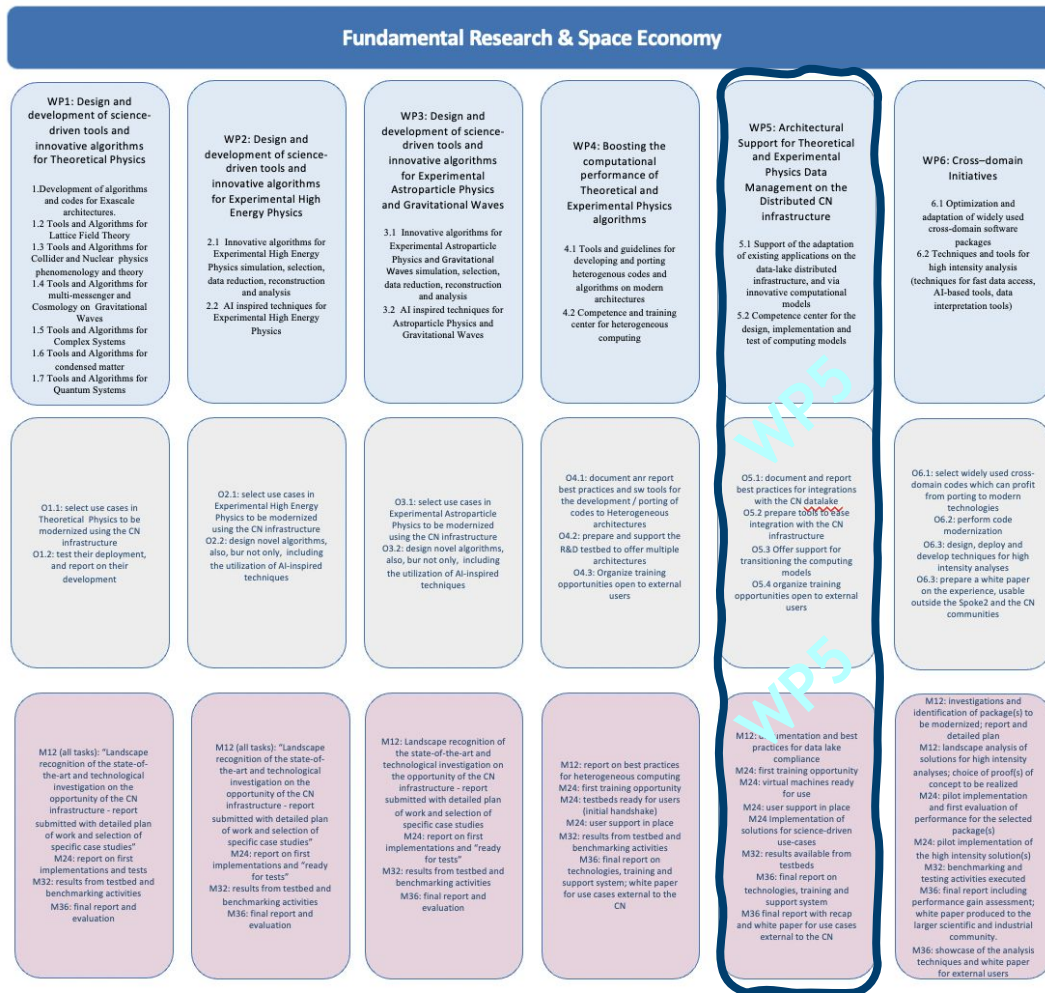
WP5 - People

WP5		
	Nome	Mail
National coordinators	Elvira Rossi Daniele Spiga	elvira.rossi@unina.it daniele.spiga@pg.infn.it
Institute coordinators		
INFN	Daniele Spiga	daniele.spiga@pg.infn.it
UNIMIB	Mattia Bruno	Mattia.Bruno@mib.infn.it
UNINA	Elvira Rossi	elvira.rossi@unina.it
UNITS	Andrea Bressan	andrea.bressan@ts.infn.it
UNIBO	Alessandra Fanfani	Alessandra.fanfani2@unibo.it
UNIFE	Luca Tomassetti	luca.tomassetti@unife.it

Adelina D'Onofrio	Federica Simone	Mattia Bruno
Alessandra Fanfani	Francesco Ciroto	Michele Pavone
Alessandro De Salvo	Francesco Gravili	Muhammad Anwar
Alessia Spolon	Francesco Noferini	Nadia Rega
Andrea Bressan	Francesco Visconti	Orso Iorio
Andrea Contu	Gianluca Sabella	Paolo Dini
Antonio Stamerra	Giovanni Della Ricca	Piergiulio Lenzi
Alessandro Bombini	Giovanni Andronico	Roberto Peron
Bernardino Spisso	Giovanni Ianniruberto	Sandra Malvezzi
Carmelo Magnafico	Giuseppe Milano	Sara Vallero
Constantinos Siettos	Guido Russo (PO)	Simone Gennai
Daniele Bonacorsi	Guido Russo (PU)	Stefano Bagnasco
Daniele Spiga	Lorenzo Rinaldi	Stefano Bagnasco
Domenico Elia	Luca Tomassetti	Tommaso Diotallevi
Eleonora Luppi	Luca Zampieri	Tommaso Tedeschi
Elvira Rossi	Lucia Silvestris	Valentina Fioretti
Fabio Garufi	Marco Landoni	



Spoke 2 - Fundamental Research & Space Economy



Spoke 2 - objectives

The Spoke 2 intends to address the needs of theoretical and experimental physics with accelerators, astroparticle physics with space- and ground-based detectors and gravitational wave investigation designing, developing and testing solutions apt to the current and next-generation experiments, and fitting the opportunities provided by the PNRR and the National Centre (CN) "Big Data, HPC and Quantum Computing".