

## CAPITOLATO TECNICO DI MASSIMA

**INDAGINE ESPLORATIVA DI MERCATO VOLTA A RACCOGLIERE PREVENTIVI INFORMALI FINALIZZATI ALLA RACCOLTA DI DATI A LIVELLO DI IMPRESA E PER L'ANALISI CONTROFATTUALE E DELLE RETI NELL'AMBITO DEL PIANO NAZIONALE RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 «ISTRUZIONE E RICERCA» – COMPONENTE 2 «DALLA RICERCA ALL'IMPRESA» – LINEA DI INVESTIMENTO 3.1 «FONDO PER LA REALIZZAZIONE DI UN SISTEMA INTEGRATO DI INFRASTRUTTURE DI RICERCA E INNOVAZIONE» – PROGETTO «FOSSR», CUP B83C22003950001**

### 1. Introduzione

L'oggetto del servizio riguarda le prestazioni che l'appaltatore selezionato dovrà svolgere nei work-package 3 e 5 (WP3-WP5) nell'ambito delle attività di progetto di seguito elencate:

1. Raccolta di dati a livello di impresa
2. Testing e raccolta dati
3. Analisi delle reti tra imprese e tra imprese e università
4. Analisi controfattuale

Work Package 3 - Task 1 (WP 3.1)

Attività 1: Analisi Preliminare per Raccolta Dati

L'appaltatore collaborerà con il CNR-IRCrES per comprendere le esigenze di raccolta dati, garantendo: Raccolta automatizzata e non di dati aziendali italiani, con particolare attenzione alle informazioni geografiche, settoriali e contabili sulle start-up che hanno ricevuto investimenti di venture capital negli ultimi 15 anni. Sviluppo di algoritmi di web scraping personalizzati per estrarre dati in modo efficiente e accurato dalle fonti identificate, ottimizzando i modelli per una raccolta dati di alta qualità.

Attività 2: Verifica e Ottimizzazione degli Algoritmi

Prima della raccolta dati effettiva, l'appaltatore verificherà la corretta implementazione dei modelli attraverso test interni, inclusi stress test e robustness testing. Basandosi sui risultati, saranno apportate ottimizzazioni per migliorare l'efficienza, ridurre il tempo di esecuzione e garantire una maggiore affidabilità nella raccolta dati.

Work Package 5 - Task 2 (WP 5.2)

Analisi delle Reti tra Imprese e Università

Nel contesto del WP 5.2, l'appaltatore svilupperà modelli di social network analysis multi-livello adatti alle caratteristiche specifiche dei dati dell'infrastruttura di ricerca. Ciò comporta: Progettazione e applicazione di modelli per affrontare dati relazionali complessi e implementazione di algoritmi avanzati per l'analisi delle reti. Applicazione concreta dei modelli per ottenere informazioni dettagliate sulle interazioni tra le entità all'interno dell'infrastruttura.

Work Package 5 - Task 4 (WP 5.4)

## Analisi Controfattuale

L'appaltatore, in collaborazione con la struttura appaltante, identificherà iniziative di policy specifiche oggetto dell'analisi controfattuale. Questo includerà:

Utilizzo dei dati dell'infrastruttura per misurare l'impatto sulle performance d'impresa.

Sviluppo di modelli controfattuali che simulano scenari alternativi senza l'implementazione dell'iniziativa di policy.

Valutazione accurata dei risultati, inclusa la verifica della robustezza delle stime controfattuali e la considerazione di potenziali fonti di bias o confondimento nei dati.

### **Consegna dei Risultati**

Il lavoro sarà consegnato al CNR-IRCrES, che supervisionerà l'intero processo garantendo l'aderenza alle specifiche richieste e la massima qualità dei dati raccolti.

Tempi di consegna entro il 1° dicembre 2024.

Segue breve presentazione del progetto FOSSR.

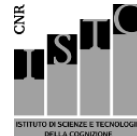


Consiglio Nazionale delle Ricerche



# FOSSR

**Fostering Open Science in Social Science Research**  
Innovative tools and services to investigate economic and societal change



# Aim of the project

## Creation of an *Italian Open Science Cloud for the Social Sciences*

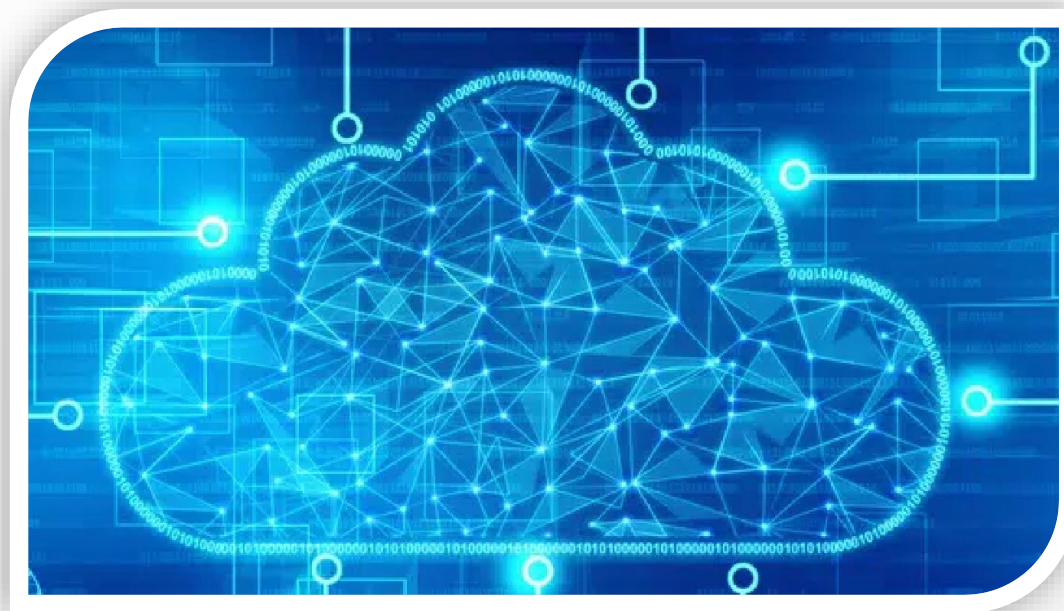
which shall provide **innovative tools and services** to investigate issues related to the **economic and societal change of contemporary societies**

### Demography and society

ageing, wealth distribution, inequalities, education, migration, etc.

### Social behaviors

Models for social simulation



### Structure of economy

innovative firms and fast-growing firms, innovation processes and outcome, new modes of knowledge production

### Public policies

Design, implementation and assessment of e.g., R&I policies, health policies, human resources policies

These are **key topics** addressed by the National Recovery and Resilience Plan (PNRR).

Network of Research Infrastructures with open resources of high-quality FAIR data, information, services, indicators, are key resources to improve the evidence-based decision-making capabilities in different policy domains.

# Context

The proposal adopts the common theme of the **development of Open Science in the Italian context** with the goal of creating a framework of tools and services for the social science scholar community, involving all the **RIs in social sciences coordinated by CNR**, namely CESSDA, SHARE and RISIS.



CESSDA ERIC Consortium of European Social Science Data Archives European Research Infrastructure Consortium, of which Italy joined in July 2021, which provides the scientific community with facilities, tools, datasets, and certified services to conduct research activities of excellence in the social sciences domain.



SHARE ERIC Survey of Health, Aging and Retirement in Europe European Research Infrastructure Consortium, an interdisciplinary and longitudinal survey on the economic, social, health and well-being conditions of the 50+ population in twenty-seven European countries (plus Israel).



RISIS Research Infrastructure for Research and Innovation (R&I) Policy Studies, which provides data and services to support the development of a new generation of analyzes, and indicators for the study of science, technology and innovation processes on the base of three main perspective: actors involved to understand the role they play, topics addressed to understand the directionality of the R&I efforts, and geography of science and innovation.

# Objectives and ambition of FOSSR

FOSSR has the general aim of promoting a **widespread knowledge and awareness of the data and methodologies employed in empirical social science**, by providing:

- (i) systematic and organized knowledge about available *social science data resources in Italian data archives*, especially the CESSDA Archive, already object of the grand infrastructural proposal;
- (ii) resources supporting methodological advancement as to data collection and data analysis, especially important for RISIS to understand the design, the implementation, and the outcome of *research and innovation policies*, which can improve the robustness of empirical evidence produced for policy makers and to deal with new research questions;
- (iii) tools and services to make publicly available advanced *probability panels for longitudinal analyses to support important survey* such as SHARE, complementing them with a network of online laboratories.

The integration of this pool of resources shall contribute to the realization of **open science for scholars in social sciences**, going with scientific training for the production and analysis of social science based on FAIR empirical data.

One further key objective is setting a **new generation of young researchers in social sciences**, by hiring 36 researchers and technologists with fixed time contract, which will become highly skilled human resources in data science and data management in social science research, and by funding 20 PhD positions to training early career researchers.

# FOSSR added value

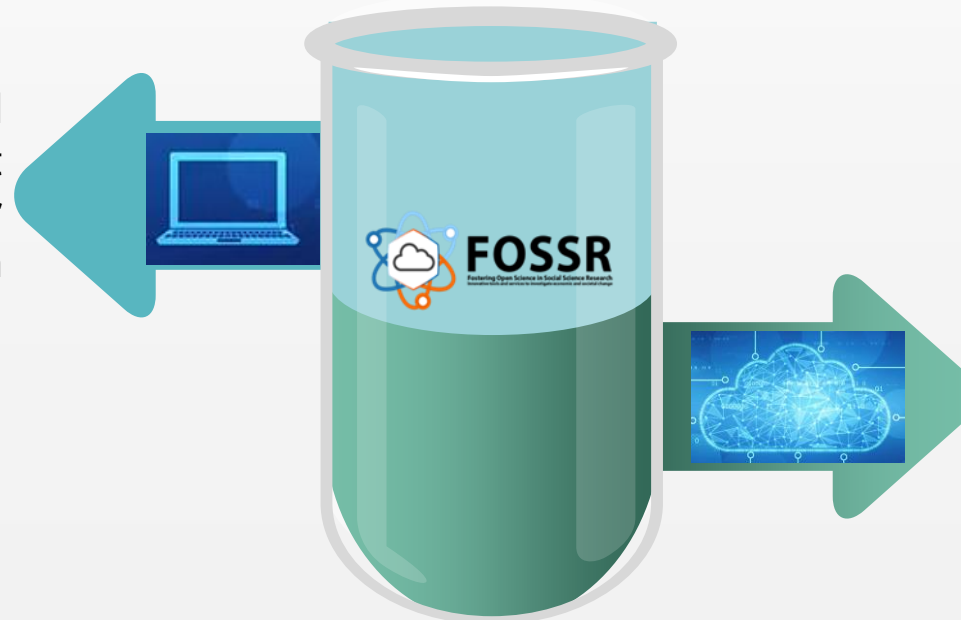
The importance of **creating linkages and synergies between RIs** dealing with social sciences is not new.

The most relevant experiences of creating integration between different European RIs can be related to the initiatives:

- Social Sciences & Humanities Open Cloud (SSHOC) Project;
- Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS);
- European Open Science Cloud Initiative (EOSC).

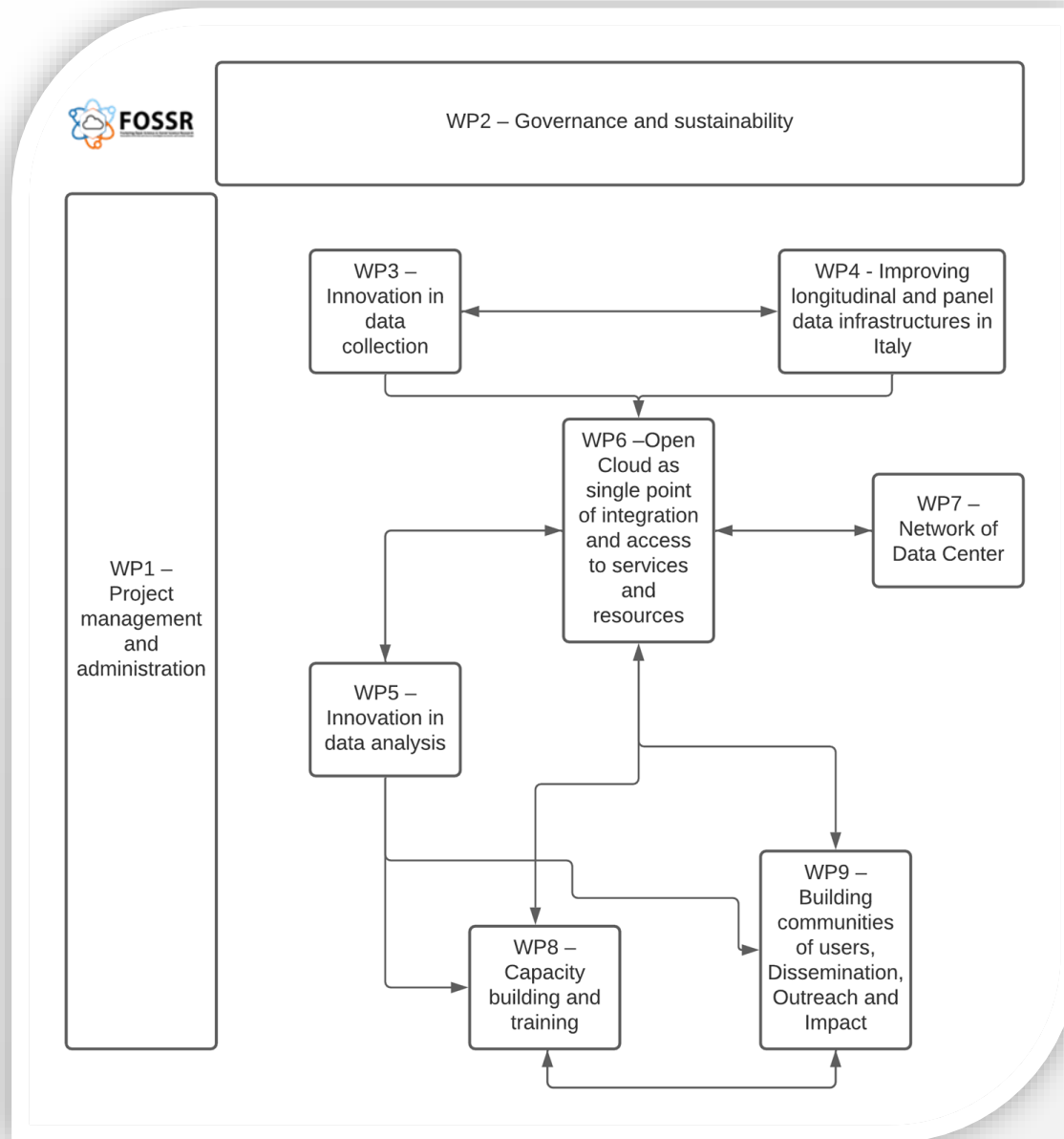
The **additional value of FOSSR** mainly rests in the creation of a network of the RIs in social sciences that

develops new open **services and resources**, highly innovative not yet existing in Italy and **essential for robust analysis** and research investigations in social sciences



builds an **Open Cloud** linked to a **network of data centers** mainly located in the South of Italy, which are supposed to improve the computing facilities existing at local level

# Structure of the project



**Scientific coordinator:**  
Emanuela Reale  
(CNR-IRCrES)

WP	WP leader name	Rank	Organization
WP1	Daniela De Gregorio	Technologist	CNR-ISMED
WP2	Massimiliano Saccone	Senior technologist	CNR-DSU
WP3	Antonio Zinilli	Researcher	CNR-IRCRES
WP4	Mario Paolucci	Director	CNR-IRPPS
WP5	Andrea Giovanni Nuzzolese	Researcher	CNR-ISTC
WP6	Mario Ciampi	Senior technologist	CNR-ICAR
WP7	Mario Sicuranza	Technologist	CNR-ICAR
WP8	Andrea Orazio Spinello	Researcher	CNR-IRCRES
WP9	Emanuela Reale	Director	CNR-IRCRES



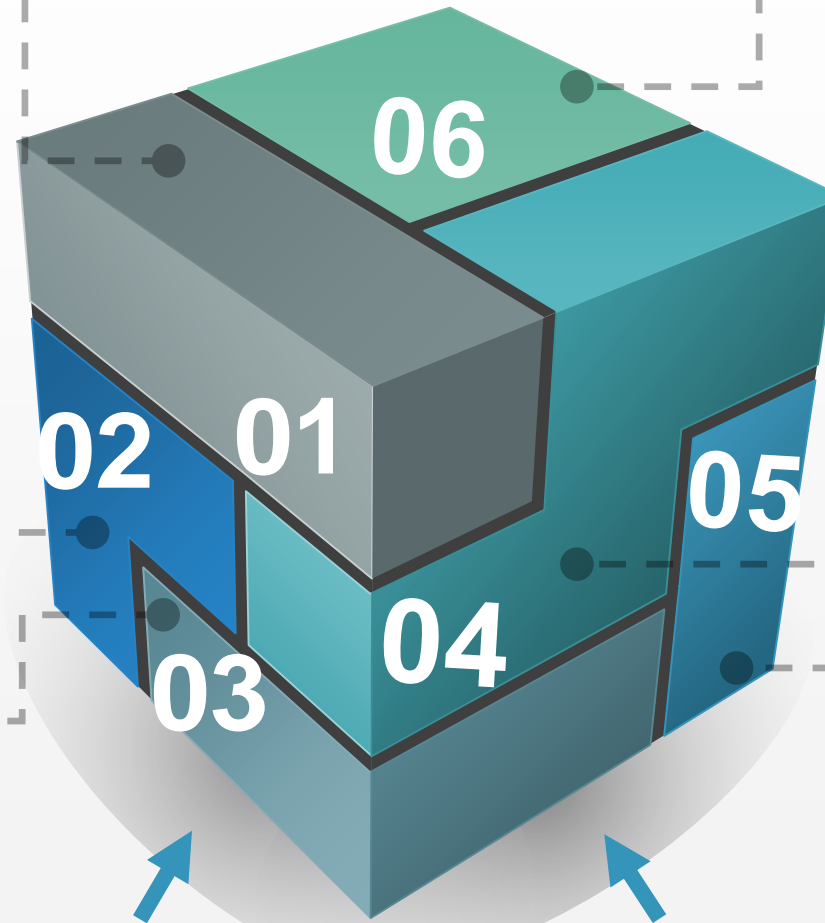
# Expected outcomes

Facilitating and promoting the **remote access to high quality data for the social sciences** by using the Italian Open Science Cloud

**Promoting the use of data** for the social sciences among scholars and non-academic users and assist data users to exploit data and resources

Providing **innovative tools** to collect, explore, analyze, and harmonize data

**Developing a governance model** that can support the long-term activity of the thematic network beyond the end of the project



Training **new generations of researchers/ data users** and disseminate the results deriving from projects developed by using the resources provided

Building a **distributed set of data centers that help to storage and distribute data** and resources derived from the combined activities of the thematic network

Strengthening the **ability of decision makers to interpret data** and use them for policy design (ex-ante evaluation of the different options) and for ex post evaluation of policies and their impact

# Strengthening skills and stakeholders' interactions

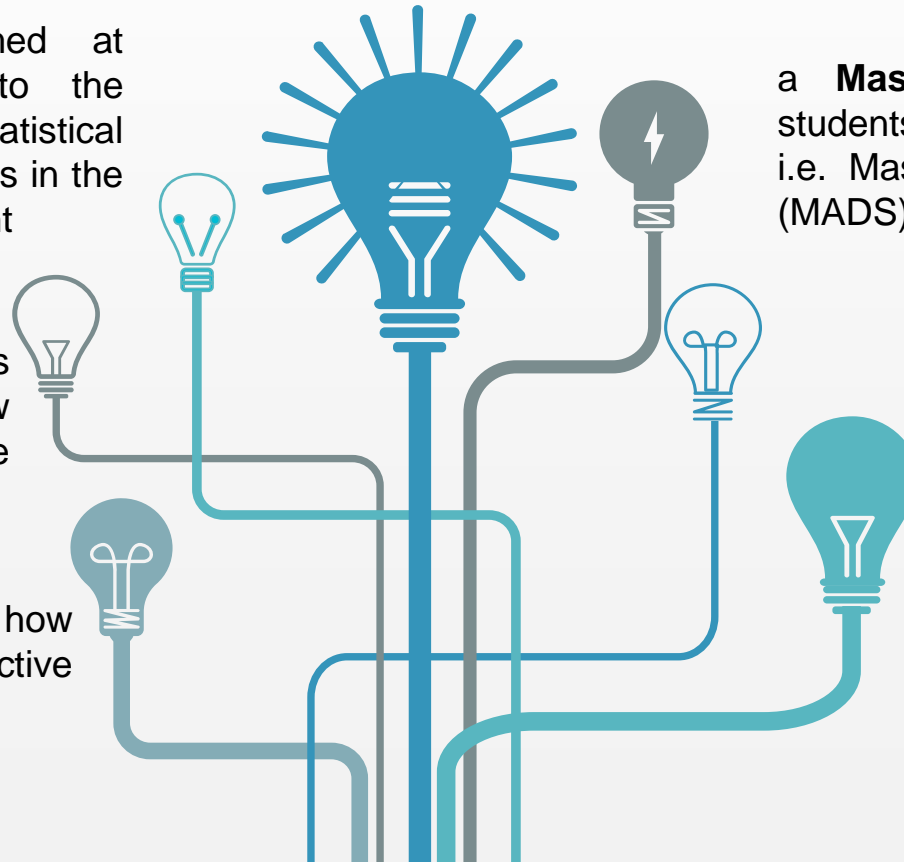
Within the WP8, FOSSR promotes activities aimed at

- (i) developing or strengthening **individual skills**, using advanced tools and techniques to address innovative questions;
- (ii) providing **expertise** in managing the research data lifecycle;
- (iii) fostering and empowering **productive interactions** between producers of knowledge and stakeholders;
- (iv) transferring **high specialistic knowledge for data science** through higher education courses preparing careers.

**Online training courses** aimed at improving knowledge related to the application of advanced statistical techniques, but also improving skills in the field of Research Data Management

**Learning sessions** to share experiences between scholars and stakeholders on how knowledge generated by data can be translated to produce social impact

**Action research sessions** to understand how skills can be exploited to develop productive interactions;



a **Master level course** to train students on data analysis techniques, i.e. Master of Applied Data Science (MADS);

**Creation of 20 PhD positions in universities** (50% in the South) to train a highly skilled new generation of data analysts

**Open 38 new positions of fix-term researchers and technologists** to increase competences on tools and methods for working with RIs in social sciences

# Long-term sustainability

We expect to have at the end of the project **new research capabilities to investigate topics of social science research** thank to the availability of innovative open services, data collections and computational resources, which shall also attract researchers from abroad, and arouse the interest of local authorities, associations, and other local actors.

In terms of sustainability, we foresee the creation of a **Joint Research Units (JRU)** between

- CNR
- The universities that are in the Italian nodes of the RIs (Milan Bicocca, Padua, Politecnico of Milan)
- The universities where the PhD positions will be opened because of existing activities on FOSSR actions – Milan Statale, Bologna, Cattolica, Sapienza, Napoli Federico II, Napoli Parthenope, Bari Aldo Moro, Catania, Sassari
- ISTAT
- INAPP
- other research organizations and universities that are likely to provide important contribution to the maintenance and improvement of the resources developed.