Image credits: S T on Unsplash; visual design: Rita Giuffredi.

Exploring research infrastructure data through advanced statistical applications

September 26-27, 2024
online training course

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

Agenda

Set. 26

10:00 — **Module 1.** Research Infrastructures and Open Science

Andrea Orazio Spinello, Emanuela Varinetti (CNR-IRCrES)

The module will outline the contours of the new paradigm for scientific knowledge production, emphasizing openness, the value of collaboration, and data availability. It will explore the features of data-intensive science, which underpin the creation and strengthening of Research Infrastructures (RI). Thse characteristics of these sociotechnical platforms will be discussed by analyzing both the Italian and European contexts, with reference to FOSSR's efforts in developing the Italian Open Cloud for Social Sciences.

12:30 — *lunch break*

14:00 — **Module 2.** Accessing and querying interoperable RI data

Lucio Morettini, Andrea Orazio Spinello, Emanuela Varinetti (CNR-IRCrES)

The module will present examples of accessing and querying data from research infrastructures, highlighting the importance of data interoperability. Specific cases will illustrate the use of the RISIS infrastructure for science, technology and innovation studies. We will present cases in which research questions are addressed differently depending on the nature of the data, analysing the ways of managing datasets and their enrichment with data external to the infrastructure. The content presented will highlight the value of shared database access.

16:00 — *end of day 1*

Set. 27

10:00 — **Module 3.** Network models applied to RI data

Antonio Zinilli (CNR-IRCrES)

The module aims to illustrate the basic concepts and statistical measures of network science and provide an overview of the main statistical network models. The module will conclude with two applications where networks are analysed using data from research infrastructures. The two applications that will be covered in this module are:
- **Application 1**: Complex networks and academic project funding
- **Application 2**: Research collaborations and research productivity

12:30 — *lunch break*

14:00 — **Module 4.** Causal Bayesian networks and applications to RI data

Lorenzo Giammei (CNR-IRCrES)

The module aims to illustrate the basic concepts and statistical measures of network science and provide an overview of the main statistical network models. The module will conclude with two applications where networks are analysed using data from research infrastructures. The two applications that will be covered in this module are:
- **Application 1**: Complex networks and academic project funding
- **Application 2**: Research collaborations and research productivity

16:30 — *end of the training*

Course description

The concept of '**open science**', which encompasses knowledge sharing within scientific communities and the interaction between science and society, plays a crucial role in contemporary research, particularly in the field of Social Sciences and Humanities (SSH). Open sharing of knowledge, including data and services for research, enables the exploration of new research questions, drives interdisciplinary collaboration, and fosters the development of innovative analytical tools. From a societal perspective, **it is essential to ensure that research findings are accessible not only to the scientific community but also to social and political stakeholders, thereby amplifying the impact of scientific work on decisions that affect citizens**. A significant boost to the development of 'open science' comes from the creation and enhancement of **Research Infrastructures** (RIs).

RIs facilitate 'open science' by **enabling researchers and other stakeholders to access high-quality data, tools, and services**. In Italy, the SSH community is still in the early stages of transitioning toward a knowledge production model based on widely shared research infrastructures. **The FOSSR project aims to sustain and encourage this process by enhancing the Italian nodes of three existing European RIs and implementing new services, tools, and resources to create a conducive environment for the development of open science in SSH in Italy**. Effective dissemination regarding the opportunities linked to RIs is essential to ensure that individuals are aware of available data and resources. Additionally, users must have the skills and expertise to use infrastructure data appropriately.

The course will guide participants through the process of **engaging with a RI and making reasoned use of its resources for research purposes**. Furthermore, it will present the application of advanced statistical techniques on infrastructure data, particularly **Network Models and Bayesian Modeling**.

Training objectives

The main aim of this methodological course is to provide an overview of the opportunities for exploiting shared data within RIs.

Specifically, the course aims to raise awareness of these opportunities and **encourage the implementation of data-driven approaches through the application of advanced statistical techniques**. The focus is primarily on the **RISIS European Infrastructure, on topics related to Science, Technology, and Innovation (STI)**, one of the three infrastructures that initiated the FOSSR project, and on advanced statistical techniques that FOSSR is treating in its research work packages. Drawing inspiration from the data available within RISIS, the course will explore aspects of **data access, interoperability and processing in depth**. From data exploration to data processing, the process will be substantiated by the exemplary application of advanced statistical techniques, particularly Network Science and Bayesian Modelling, that will be treated in theory and practise.

After completing this course, learners will be aware of the opportunities linked to RIs in SSH, particularly the RISIS infrastructure, and will gain in-depth knowledge of techniques developed within the FOSSR research Work Packages.

Audience

The course is designed for individuals who are or wish to be involved in **creating, capturing, analysing, or generally managing research data within the social science disciplines**. The target audience includes, but is not limited to, early-career researchers, researchers aspiring to advance their careers, technicians, data stewards, and data managers.

How to apply

Applications should be submitted **within September 19th, 2024** through the form available at: ***https://l.cnr.it/fossr-training-set24***

More information: ***www.fossr.eu/eventi/fossr-training-sept24/***
Contact person: **andrea.spinello@ircres.cnr.it**

The online training is open to everyone, with no specific requirements on previous knowledge or competences. However, due to the format of the training course, **the number of participants is limited to a maximum of 35**. A selection will be made on the basis of the closeness of the scientific field covered by the submitted curriculum vitae to the subject matter of the course. Notification of acceptance: September 20th.

Organisational details

The course will be structured in four on-line training modules distributed in two days, each one based on frontal lessons and several interactions. The course will be held online; the link will be provided to participants in due time before the beginning of the course.
The course is conducted in Italian, while the materials will be published in English.

The training is organised by CNR-IRCrES in the frame of the FOSSR project.
Local scientific/organising committee: Andrea Orazio Spinello (CNR-IRCrES), Serena Fabrizio (CNR-IRCrES), Alessia Fava (CNR-IRCrES), Rita Giuffredi (CNR-IRCrES), Alessandra Maria Stilo (CNR-IRCrES).