



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Consiglio Nazionale
delle Ricerche

From open data to open knowledge graphs: the role of semantic technologies

February 28-29, 2024
online training



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

Agenda

Feb. 28

- 9:30 **Welcome and introduction to the course**
Claudia Pennacchiotti, Valentina Tudisca (CNR-IRPPS)
- 9:40 **Module 1. Linked Open Data**
Giorgia Lodi (CNR-ISTC)
- 12:30 *lunch break*
- 14:30 **Module 3. Querying knowledge graph with SPARQL**
Alessandro Russo (CNR-ISTC)
- 17:30 *end of day 1*

Feb. 29

- 9:30 **Module 3. Ontology design**
Andrea Giovanni Nuzzolese (CNR-ISTC)
- 12:30 *lunch break*
- 14:30 **Module 4. Knowledge extraction**
Andrea Giovanni Nuzzolese (CNR-ISTC)
Miguel Ceriani (CNR-ISTC)
- 17:30 *end of the training*



@fossrproject



fossr.eu



fossr-eu



Course description

According to the open data handbook, **open data are a tremendous resource for transparency and democratic control, participation, self-empowerment and economic growth**. Open data are one of the key tools for enabling open science, which is at the core of FOSSR and data-driven social science. However, **open data are a knowledge soup, which is inherently heterogeneous** with respect to syntax and semantics. Furthermore, open data management typically conveys **common problems associated with big data processing**. This makes **extremely complex to reuse open data and fully exploit their inestimable value in most of data-driven decision-making scenarios**. For example, how can citizens, organisations, or artificial agents make better decisions, or be engaged in collaborative and active social behaviours? Again, how can natural or artificial agents leverage open data to get only contextual data which are relevant to their specific goals?

In last decade, **linked open data (LOD)** have emerged as a **set of best practices and principles for sharing machine-readable interlinked open data on the Web**. LOD, ontologies and, more in general, **semantic technologies play a fundamental role in the implementation of FAIR principles, especially with regards to interoperability**, whose adoption is commonly overestimated or overlooked.

We will understand **how semantic technologies enable the creation of more efficient, reusable, interoperable, and accessible open datasets enhancing them to knowledge graphs**. This is utmost important in FOSSR, which has a whole work package (i.e. WP5) dedicated to innovation in data analysis based on knowledge extraction, integration, and enrichment. WP5 aims at **building the knowledge representation layer in FOSSR to fill the gap between data and knowledge, thus enabling a better interaction between machines and humans for social science research**.

Training objectives

During the course the students learn:

- (i) RDF to design linked open datasets;
- (ii) SPARQL to query the linked open data cloud;
- (iii) OWL, ontology design patterns, and Protégé to design ontologies;
- (iv) RML to extract knowledge from structured sources to knowledge graphs;
- (v) Some prompting engineering to ask large language models to perform knowledge graph construction from unstructured sources.

Audience

Students, researchers and practitioners aimed at learning how to enhance their data to knowledge graph in order to foster FAIR principles by maximizing knowledge reusability and interoperability. No particular knowledge is required as the course is intended for lay users. However, basic knowledge on open data and possibly Python programming is required.

How to apply

Applications should be submitted through the form at:

<https://l.cnr.it/fossr-oline-training-registration-form>

within **February 26, 2024**

contact person for info: andreagiovanni.nuzzolese@cnr.it



Organisational details

The course will be held online; the link will be provided to participants in due time before the beginning of the course.

Lessons will be held in Italian. Teaching materials, thematic literature and notes on the topics covered will be available to all participants in English.

The training is organised by CNR-ISTC and CNR-IRPPS in the frame of the FOSSR project.

Local scientific/organising committee: Andrea Giovanni Nuzzolese, Giorgia Lodi, Alessandro Russo, Miguel Ceriani (CNR-ISTC).