Event series: Open Science and Covid-19. Working together to fight the pandemic

Organized in collaboration by the Italian nodes of [OpenAIRE](), [ELIXIR](), [RDA]() and [EOSC Pillar]().

Scope

The series of events is intended to offer the participants a deep understanding of the motivations behind the sharing of data and other research results (publications, software, etc), and how they can perform it in a simple way.

The series includes six online training events consisting of three webinars followed by Q&A sessions and three practical tutorials. The first event will be an introductory webinar aimed at raising awareness about the crucial importance of sharing COVID-19 data rapidly and systematically; the second and third webinars will have a similar format and goal as the first one but will focus on epidemiological and clinical data, respectively. Two out of the three tutorials will provide learners with knowledge and skills to effectively manage and share different types of omics data. Last tutorial will be about best practices on how to share research software. These tutorials will help researchers, clinicians and technicians to produce results that can be immediately replicated, evaluated and reused.

The series consists of a set of webinars and tutorials.

The **webinars** aim at raising awareness in the scientific community of the importance of data sharing to monitor the evolution of the Covid-19 pandemic and to find an effective cure. The tools and good practices of Open Science and Open Access will be showcased. Besides some of the inconsistencies in the collection of molecular, epidemiological and clinical data will be highlighted.
Tutorials will be training events showing specific solutions to share different types of homogeneous data and other research products, such as analysis software and protocols.

Format and delivery of sessions 1-3: these sessions will last up to two hours and will consist of a webinar, followed by a Q&A and discussion session. Delivery will be made highly interactive by means of tools such as Mentimeter [www.mentimeter.com], which will be used to ensure frequent formative feedback, and shared notes taking. Participants' questions and comments will be collected during the webinar in the shared document and addressed during the Q&A
session.

Format and delivery of tutorials 4-6: The tutorials will be delivered remotely and will be very interactive and practical. Participants will be encouraged to bring a sample of their data and will be accompanied through the pre-processing phase (data annotation, formatting, etc) and the deposition phase. Example data and datasets to work with will be provided anyway. Very short bits of theory will be followed by the presentation of practical examples and

activities carried out by learners with the supervision/support of the trainers.

Aims:

Participants will learn about:
- The efficacy of the Open Science model and its cooperation method in the context of the ongoing pandemic.
- Specific aspects of open science and access to research data on COVID-19 and SARS-CoV-2.
- The main European initiatives for the sharing of research data and results within the scientific community.

Learning outcomes:

At the end of instruction, learners will be able to:
- Tell why and how the open data approach will increase the visibility of their work and the chance for collaborations, and enhance the reliability of their research results
- Explain why sharing their data on COVID-19 is of outstanding importance to ward-off current and future outbreaks
- Identify available resources (databases, repositories, and tools) for their specific type of data
- Pre-process their data prior deposition (data annotation, formatting, anonymization. etc.)
- Describe where and how to deposit their data in public repositories
- Explain how to protect their data
- Explain one or more ways of evaluating compliance with FAIR principles
- Apply best practices for software sharing to their software

Introductory Webinar

Prerequisites:

The webinar aims at giving an overview of the motivations behind the need to share research results (data, publications, software, etc..) related to COVID-19 disease and SARS-CoV-2 virus, given the current global outbreak. No specific knowledge or prerequisites are needed, except the ones defined by the target audience.

Target audience:

The introductory webinar is dedicated to:

- Researchers, in particular those involved in the collection, production and analysis of COVID-19 SARS-CoV-2 research results (data, publications, software, etc…).
- Clinicians and doctors involved in the current outbreak management and assessment.
- Technicians and other medical personnel involved in the collection and management of COVID-19 and SARS-CoV-2 related data and information.

- RDA in Italy node members

Teaching goals:

The introductory webinar aims at highlighting how the traditional way of doing science is changing given the current COVID-19 outbreak, in particular concerning scientific communication and sharing of results in the international context and among the different disciplines.

The webinar will inform participants about:

- the efficacy of Open Science model and its cooperation method in the context of the ongoing pandemic;.
- specific aspects of open science and access to research data on COVID-19 and SARS-CoV-2;.
- the main European initiatives for the sharing of research data and results within the scientific community..

Epidemiological and Clinical data

The first part will be dedicated to highlighting the non-homogeneity of the current data, showing some of the biggest inconsistencies in the current management of COVID-19 epidemiological and clinical data. The second part will illustrate guidelines for data collection and management proposed for example by WHO (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/surveillance-and-case-definitions) or suggestions from COVID-19 data portal (https://www.covid19dataportal.org/submit-data) and RDA recommendations (https://www.rd-alliance.org/groups/rda-covid19-epidemiology).

These two sessions are aimed at raising awareness on:

- the fact that epidemiological and clinical data needs to be collected on a strong statistical basis
- the fact that a country needs a common strategy to collect epidemiological and clinical data in order to successfully merge data coming from different regions
- the importance on the two above aspects that reflects on decision making protocols
- the fact that a country needs a strategy on clinical data digitalisation, storage and sharing
- how this strategy can be implemented thanks to the standards realised by ELIXIR and platforms developed in other countries.

Target Audience (session on epidemiological data)

- Policy makers and health sector administrators, such as those involved in the collection of epidemiological data and decision making related to the current outbreak.

This webinar will be dedicated to COVID-19 epidemiological data. No specific knowledge or prerequisites are needed, except the ones defined by the target audience.

Teaching goals (epidemiological data):

The aim of the event is to inform and motivate the audience to adopt standard and shared guidelines for the collection of COVID-19 epidemiological data.

Participants will develop skills on:

- Knowledge of the main inconsistencies of COVID-19 epidemiological data collection.
- Awareness of the extreme importance for consultation and decision making of reliable and comparable data.
- knowledge of the principles of WHO protocols: ie. the use of a case-based reporting form (and a related data dictionary), and the aggregated weekly reporting form.

Tutorials on how to share omics data

Two tutorials will focus on different types of omics data (Part I : Virus/Host genomics, transcriptomics, GWAS. Part II: Virus/Host proteomics and structural data, metabolomics, interactomics)
For each specific type of data, the tutorial will describe:
- Advantages of data sharing
- Challenges of data sharing and how to handle them
- How to make data findable, accessible, interoperable, and reproducible
- Tools and resources for COVID-19 data collection and management
- Procedures to manage data prior deposition
    o How to protect data
    o Data ownership
    o Data format(s)
    o Dara annotation
- Procedures to make data available in public resources
- Tools, best practices, examples based on the Research Data Alliance (RDA) COVID-19 recommendations and guidelines that are under development by the RDA-COVID-19 Working Group.

Tutorial on best practices for sharing software/workflow for the analysis of data related to COVID-19

This tutorial will teach a selection of materials relevant for COVID-19 and extracted from the lesson on four simple recommendations to encourage best practices in open research software (4OSS-lesson) and RDA recommendations for software sharing. Materials from the

4OSS-lesson will be adapted for remote delivery.