Contribution ID: **23** Type: **Paper presentation**

Towards the implementation of FAIR principles on an earthquake analysis platform

Friday 5 November 2021 15:00 (20 minutes)

FAIR principles of scientific data represent a relatively new direction in research that is best described as an initiative to encourage and help researches in making their data more available and easier to reproduce and reuse. Most importantly, the principles emphasize machine actionability, meaning the data requires strict guidelines so that automatic systems can read, retrieve and use the data with minimal human intervention. FAIR stands for Findable, Accessible, Interoperable, and Reusable, defined as below:

- Findable : data must be easy to find for humans and machines. Databases must be described with an abundance of metadata and clear, explicit identifiers so that they are machine readable. The storage system must be capable of discovering the data by intuitive query.
- Accessible : after finding it, the users must know how to access it, possibly through authentication and authorization, using a standardized communications protocol. This protocol must be open, free and universally implemented.
- Interoperable : metadata and databases need to respect the same format, so that they can be integrated with each other and that you can use the same analysis, storage or processing workflows, interchangeably.
- Reusable : data must be well-described and commented so that they can pe reproduced or combined. They will be described by the same relevant attributes and will be released with a clear and accessible data usage license. Finally, the details of their collection must meet the standards of the scientific community which they are meant for.

In our work on earthquake analysis the first step was that of collecting, parsing, curating and storing the seismic databases available online. Because of the small differences in the structure of the publicly available earthquake databases, our job was to select only the relevant information present in all databases. To this end we developed a parsing code which downloaded the databases, retrieved the data, and created new databases on a new format that allowed total interoperability between them.

The next step was to develop the analysis tools. Here, through the theory of complex networks, we elaborated codes which call the databases and use the information to create a seismic network by splitting the seismic region into small 3D cubes, and placing each earthquake in their respective cube, based on its epicenter geographical position. Building the network chronologically, each subsequent earthquake represents a link in the network, whereas the cubes represent the nodes. For these seismic networks we developed a series of codes which analyse the degree of connectivity, the structure of the motifs, and allow for automatic visualizations using open source third party software.

Summing up, we report a new analysis framework for earthquakes in which data gathered from public sources is structured in a database format that allows automatic processing through a series of codes developed in house. This process represents a "FAIRification"of the available seismic data, as the databases that we created preserve only a fraction of the initial information, but allow Interoperability and Reusability of these databases, in addition to the Findability and Accessibility that was already provided for by the original seismic catalogues. This opens new areas of research associated to FAIR and Open Science, the research on the IT and Computing infrastructure supporting these principle, research on designing faster and reliable networks for supporting access and the design and implementation of new cybersecurity standards and controls that could support the users.

Authors: PANA, Gabriel (Faculty of Physics, University of Bucharest); IVANOAICA, Teodor (Extreme Light Infrastructure European Research Infrastructure Consortium); RAPORTARU, Carina (The Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI)); NICOLIN, Alexandru (Faculty of Physics, University of Bucharest); BĂRAN, Virgil (Faculty of Physics, University of Bucharest)

Presenter: PANA, Gabriel (Faculty of Physics, University of Bucharest)

Session Classification: Network Management && Open Source and GNU in Education and Research

Track Classification: Open Source and GNU in Education and Research