Open Research Award

Open Data in Open Software by *Giulio Genova & Mattia Rossi*

Dear Jury of Eurac's Open Research Award 2021,

We are Giulio Genova and Mattia Rossi, early career scientists at Eurac Research, and we would like to participate to the Open Research Award sharing our "Open Data in Open Software" story.

Within the last few years more and more environmental data has been shared from public institutions and private companies creating something that could be called a "data Tsunami". Such amount of information promoted data-driven decision making and problem solving which are now the backbone of several industries worldwide. Open data, together with Open-Source Software, are the driving innovation and contribute to technological and, possibly, social progress. An increasing amount of data, however, requires smart tools to distribute them in an efficient way and to ease the analysis conducted by users with different backgrounds. These users now require tools that help to work efficiently with different data sources by simultaneously lowering the complexity in accessing and importing them to their analysis software of choice. In other words, a better user experience is needed and here is where our work comes into play. As researchers in the environmental sector, we often use and need meteorological, hydrological or other environmental data for our studies. Therefore, we decided to develop tools, written in the open-source language R, that help researchers to quickly access analysis ready data from diverse data sources.

One important open data source for research is the Open Data Portal (http://daten.buergernetz.bz.it/) South Tyrol provides several meteorological and hydrological variables coming from 120 real-time monitoring stations. The interface they provide (CKAN based RESTFul API) is easily accessible for computer scientist or software professionals with a higher programming skill, but it is quite cumbersome to use by without prior training in these IT disciplines. As researchers, we need a quick and easy way to understand a database and its entries as well as to preview parameters and locations that may fit in a scientific study. As a next step, we are also interested in automating our processes to feed and test our environmental models seamlessly and recurrently.

Back in 2017, starting with his PhD thesis, Mattia started to work with the Monalisa database (http://monalisasos.eurac.edu/sos/) and tried to create an automated software solution to directly import the huge amount of diverse environmental variables in the programming language R. With a first R-package prototype showing very promising results the access has been expanded to the Open Data catalogue of the Autonomous Provice South Tyrol. This work resulted in the openly accessible MonalisR package (https://github.com/mattia6690/MonalisR) with the capabilities to access both databases with the same software and added by a detailed description and tutorials to get started with the developed functionalities (https://mattia6690.github.io/MonalisR/). Nowadays the package is still used by researchers and has continued to be curated within Eurac's DPS4ESLAB project (https://noi.bz.it/de/dps4eslab) alongside the newly created Environmental Data Platform (https://edp-portal.eurac.edu).

The robust approach and ease of use of MonalisR drove Giulio to use its functionalities as a backbone and build a web application on top of it, allowing people with no programming experience to visualize and

download the data. The web application was named Meteobrowser (Figure 1) and published as an R-package (https://github.com/GiulioGenova/MeteoBrowser), as well as a peer-reviewed publication (https://doi.org/10.3897/rio.5.e35894) .
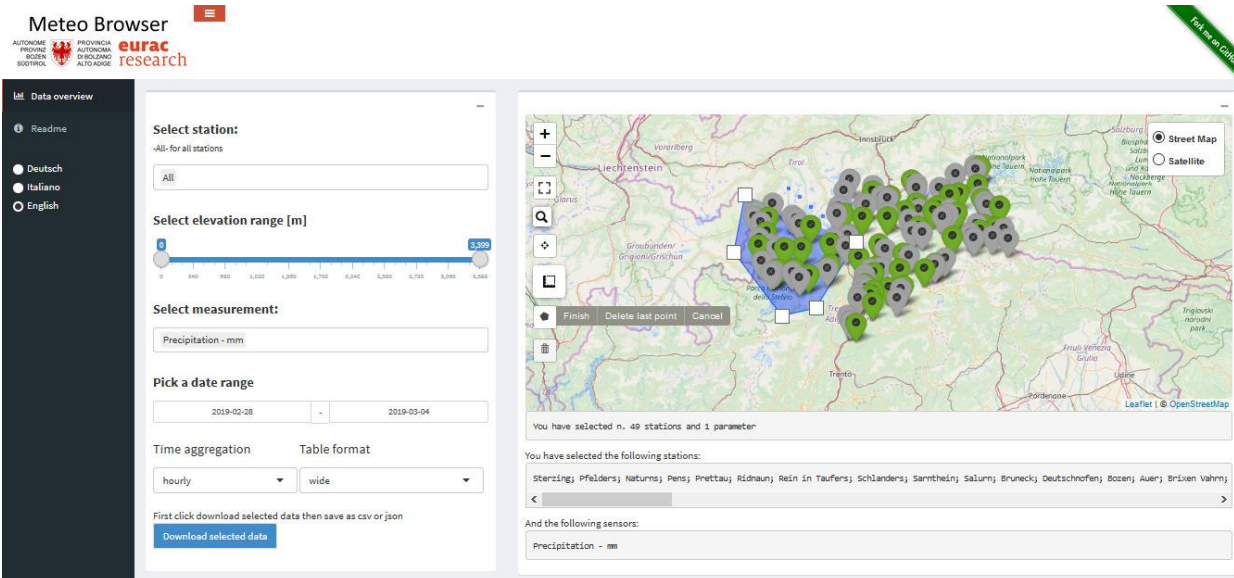

Figure 1 Main page of the MeteoBrowser with an example of spatial selection of the monitoring stations

The intuitive way the interface is build and the openness of code and application increased interest among researchers and technicians at the province of Bolzano, and Eurac's IT department provided a server and a custom domain to host the application permanently http://meteobrowser.eurac.edu/. After reviewing the tool, the province of Bolzano listed the Meteobrowser as an official tool to download the freely available hydrological and meteorological data. Today the app is used by students, researchers, and companies working in a wide range of fields. Starting from August 2019 the app is visited by 60 users per week for about 3000 users per year (Figure 2).
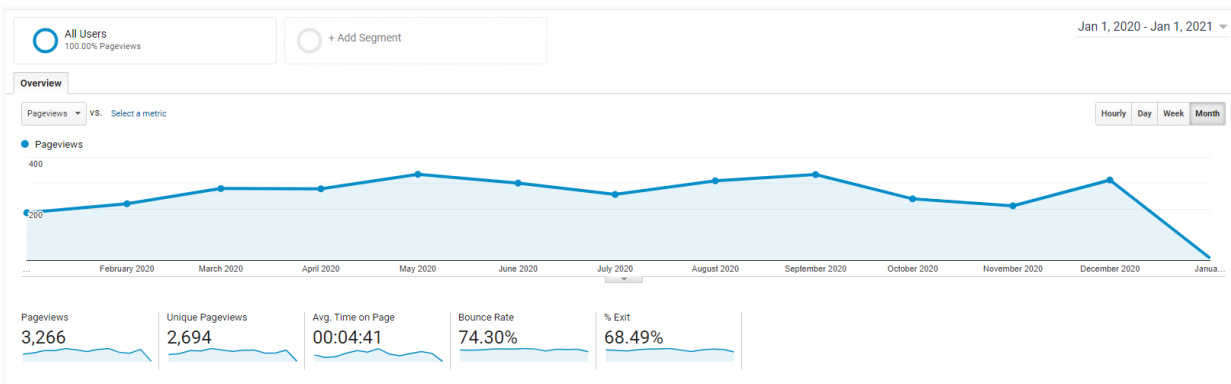

Figure 2 Average monthly users in 2020 of the MeteoBrowser

The cooperation during the development of these tools lead Mattia and Giulio to create Bolzano R (https://www.bolzanor.eu) (Figure 3), a group of R-users to facilitate the connection between local R-developers and R-users. We realized that many persons at the different research centers as well as private R-enthusiasts are present in the Autonomous Province of South Tyrol and we initiated the group as a panel to meet, foster ideas and collaborations and to keep up-to-date with the newest developments within the R world. At the same time, the website gives all the participants the possibility to create Bolgposts, disseminate the activities and list software packages and publications.
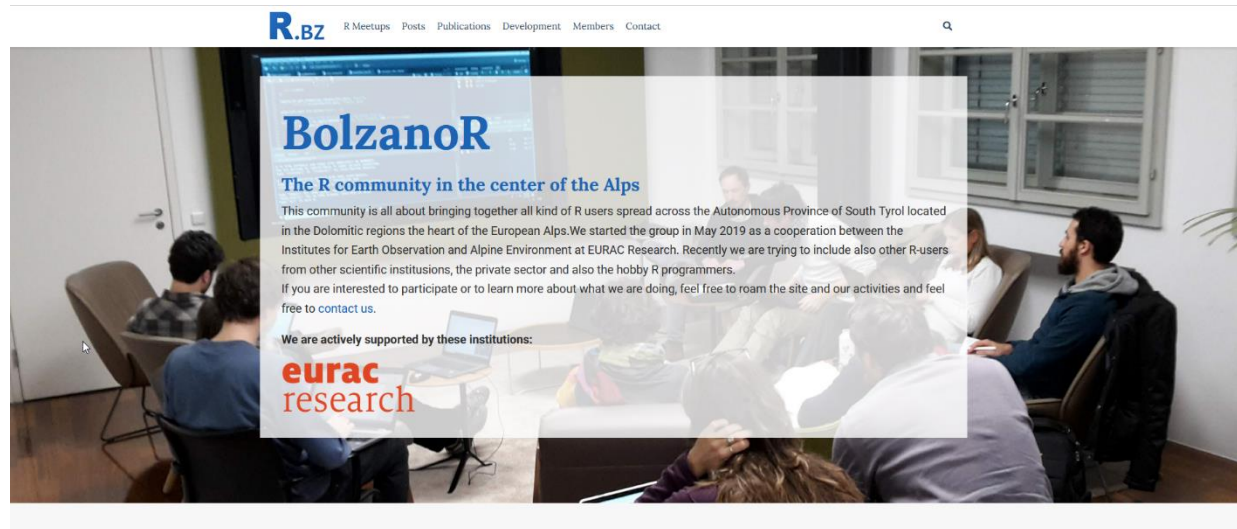

Figure 3 Homepage of BolzanoR's website

The developments done by Mattia and Giulio during the past few years contributed to enhance the capabilities to create Open Data solutions that helped the researchers with their studies while providing easy tools to access diverse and recurrently needed data streams. Moreover, the community of R-users that has developed augmented the knowledge of the R programming language and established connections between researchers. This ultimately led to interesting research ideas and cooperation between participants but also to other entities in South Tyrol such as the OpenDataHub South Tyrol or the Free University of Bolzano.

Best regards,

Giulio Genova & Mattia Rossi