

## Exercise 6

### **Prepare and upload an open dataset to Zenodo Sandbox**

Results of analyses are generally published to advance knowledge in all areas of scientific research. But it is also useful to publish datasets for the same reasons. And to make these data accessible to the entire scientific community, it's very important to publish them in the “open” domain. When an article is published in a scientific journal, it receives a unique identifier called “DOI” for “Digital Object Identifier”. This makes it easier to identify publications and to refer to them in other publications. Following the same model, “open” datasets published in data repositories also receive a DOI.

#### **Tasks**

To complete the exercise of analyzing and publishing research results that you are carrying out this semester, we ask you here to upload the open dataset that you developed as part of Exercise 3 (Lausanne's inhabited hectares characterized by environmental data on noise, vegetation and air pollution).

You have to place your vector dataset (.gpkg or .shp) in a .zip file, then upload this file to the Zenodo data repository. We'll be using Zenodo's Sandbox test space for this exercise (<https://sandbox.zenodo.org/>). The aim is to obtain a DOI for your open dataset. When you will write your article (semester project), you'll be able to refer to your dataset and cite it as a reference. Take into account the fact that the DOI is fake when we use Zenodo Sandbox, and if you want to refer to your dataset or share it, then you have to use the direct link to the record.

1. Go to <https://sandbox.zenodo.org/>
2. Sign up to create a Zenodo account
3. You will receive an email to confirm your email address, and you will receive the documentation too
4. Choose New upload in the menu (top right of the page)
5. You can select the “EPFL” community where to submit your record
6. Get a DOI
7. Word a title for you dataset, specify a publication date
8. Add the author (you are the only author)

## Exploratory data analysis in environmental health

Dr Stéphane Joost, Dr Mayssam Nehme, Noé Fellay

---

9. Create an original text to describe your dataset
10. Choose an appropriate type of open licence (Creative Commons Attribution 4.0 International). Check the information on these types of open license [here](#)
11. Add key words,
12. Add a field (domain of research)
13. Complete the other elements required and make your dataset public

For your short report, explain the different steps necessary to publish your open dataset, insert the main information among which the descriptive text and the keywords, add appropriate screen captures if useful and at the end mention the full reference to your dataset.

For instance:

Reinhard, F., Parkan, M., Produit, T., Betschart, S., Bacchilega, B., Hauptfleisch, M. L., Meier, P., SAVMAP, C., & Joost, S. (2015). Near real-time ultrahigh-resolution imaging from unmanned aerial vehicles for sustainable land use management and biodiversity conservation in semi-arid savanna under regional and global change (SAVMAP) (2.0). Zenodo.  
<https://doi.org/10.5281/zenodo.1204408>