



Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

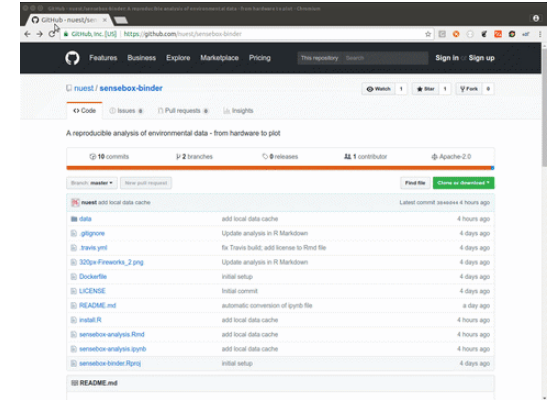
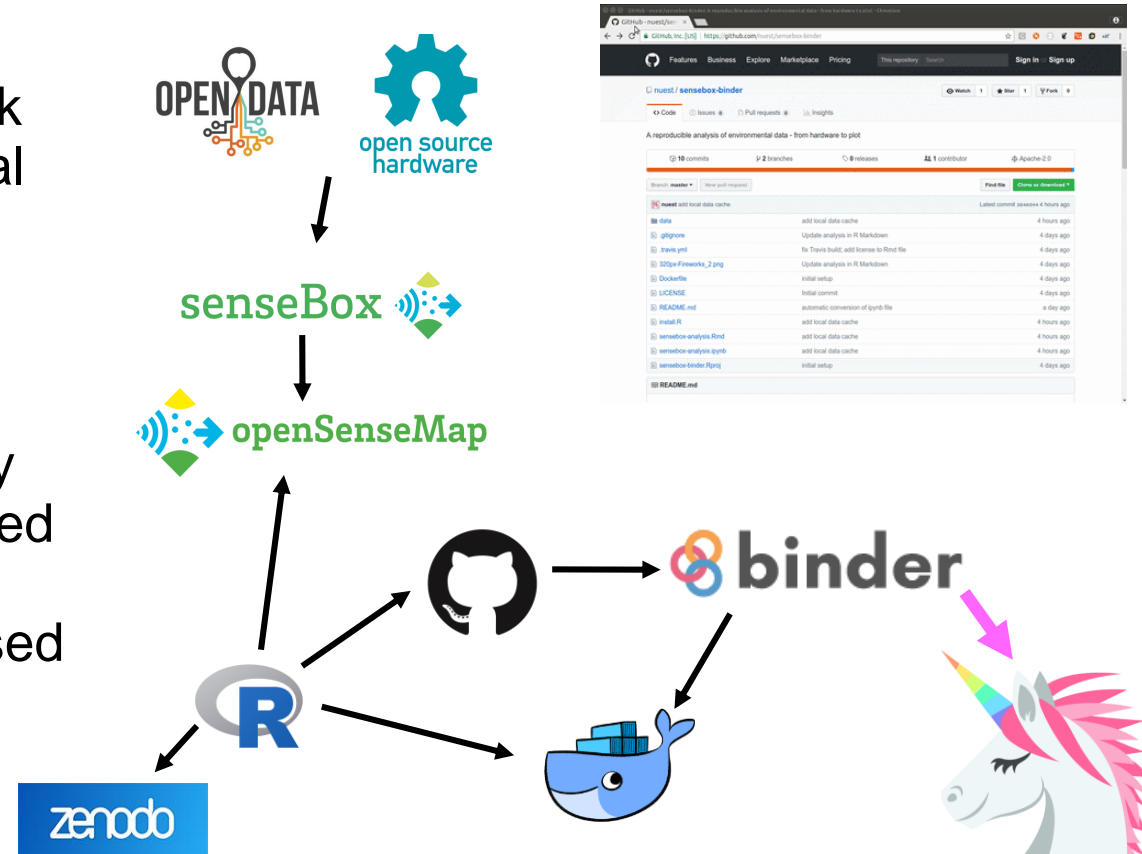
Binder

Video

Abstract + References



Documented, portable, archivable, and one-click interactive computational environments for reproducible research based on Open Source software analysing Open Data measured by Open Hardware published online as interactive website using cloud-based workflows and Linux containers. For free. Awesome.



Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek
Institute for Geoinformatics
{daniel.nuest,bartoschek}@uni-muenster.de

Open Science

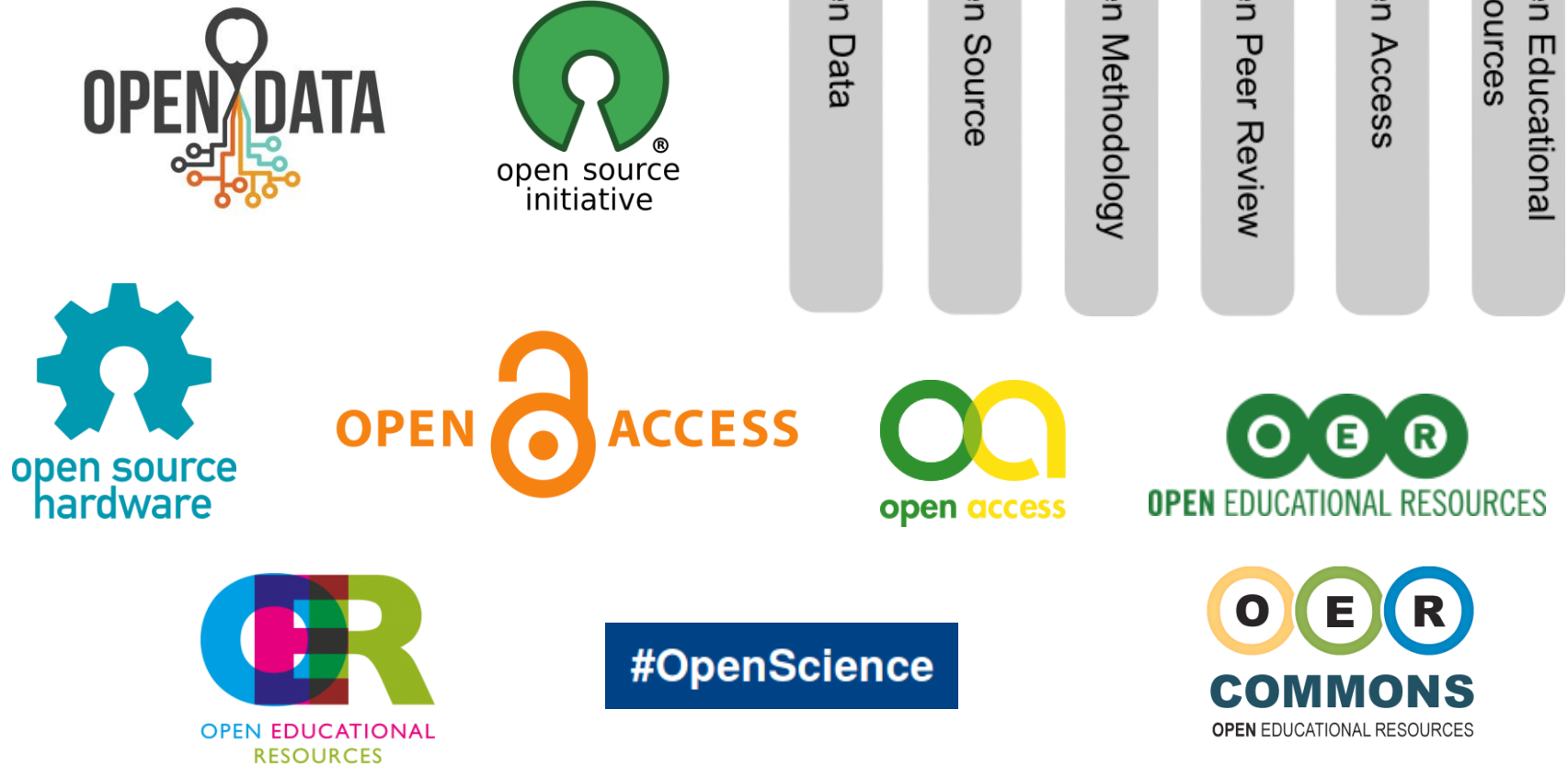
https://en.wikipedia.org/wiki/Open_science

- Overview
- Open
- senseBox
- OSeM
- R Markdown
- Jupyter
- {D,R}ocker
- Binder
- Video
- Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code Blog



All icons copyright by the respective organisations!

Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek
Institute for Geoinformatics
{daniel.nuest,bartoschek}@uni-muenster.de



<https://sensebox.de/en>

Overview

Open

senseBox

OSem

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI: [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code



Blog



senseBox:home



Citizens with senseBox:home can use the technology for their own local research, to collect data by themselves or to contribute to the openSenseMap sensor network.

senseBox:edu



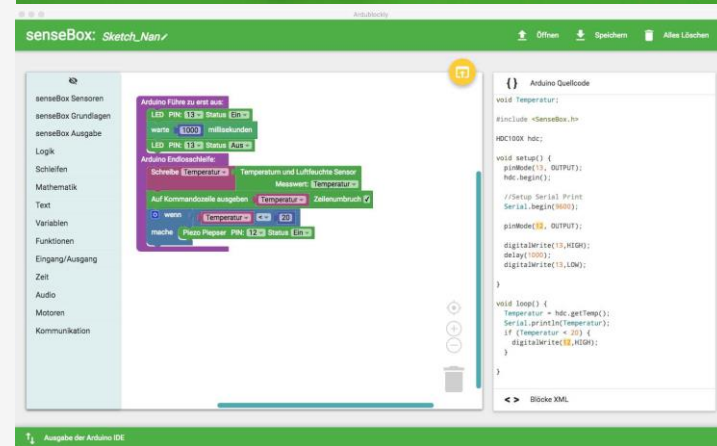
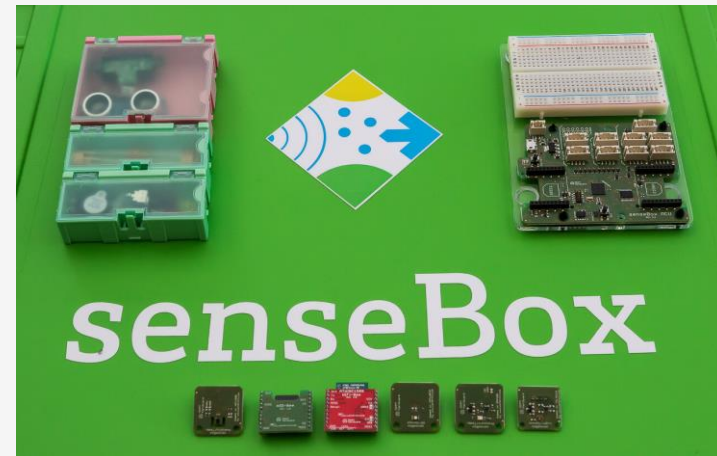
Additionally, we created senseBox:edu for schools and junior scientists, which is an experimental box with didactical concepts, instructions and project ideas included.



Twitter

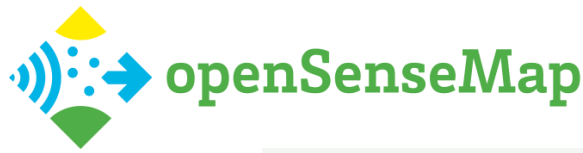
Facebook

Github



Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek
Institute for Geoinformatics
{daniel.nuest,bartoschek}@uni-muenster.de



Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI 10.5281/zenodo.1135139



Code



Blog



Browse Analyse Download API

Interpolation

Here you can visualize senseBox measurement data for areas. The Inverse Distance Weighting algorithm is used.

Phenomenon: Luftdruck

Begin: 01/03/2018 14:40 End: 29/03/2018 14:45

Time steps: 1 Exposure: outdoor

Cell width in km: 1 Power: 3

Create Interpolation Remove Interpolation

Legend: 3.49 - 22.79, 22.79 - 42.09, 42.09 - 61.39, 61.39 - 80.69, 80.69 - 99.99, > 99.99

1561 senseBoxes
748029324 measurements

Filter

Select criteria by which you want to filter stations. When you change the time frame and phenomenon in combination with the time frame, a new set of senseBoxes will be fetched from the server. This might take some time. Searching for senseBoxes by name, group, exposure and phenomenon without changing the map.

Measuring time frame: 03/02/2018 → 03/29/2018

Time range: Last 24 hours Last week Last month

The box search filters the results in the map.

Name: [input]
Model: [input]
Exposure: [input]
Phenomenon: [input]

Results (875) [button]

Boxes - Get one senseBox

<https://api.opensensemap.org/boxes/{senseBoxId?format=:format}>

Field	Type	Description
format	string	The format the sensor data is returned in. If 'geojson', a GeoJSON Point Feature is returned. Default value: json Allowed values: json, geojson
senseBoxId	String	The ID of the senseBox you are referring to.

Example data on success:

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [
      10.833333333333334,
      49.58333333333333
    ]
  },
  "properties": {
    "timestamp": "2018-06-02T11:22:51.817Z",
    "type": "Point",
    "name": "00111111111111111111",
    "measurements": [
      {
        "timestamp": "2018-06-02T11:22:51.817Z",
        "value": 101
      }
    ]
  }
}
```

1526 senseBoxes 733788676 measurements

Search for boxes and places

Geologengasse Feinstaub Temp Luftdruck 1030 Wien

Group: waste outdoor Refreshing in 25 sec

Some information about this senseBox is shown here. This senseBox owner has not yet provided additional information.

PM10: 2.53 µg/m³ a minute ago

PM2.5: 2.30 µg/m³ a minute ago

02/28/2018 → 03/29/2018

Last 24 hours Last week Last month

550.0
500.0
450.0
400.0
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0

3/18/2018 12:00 AM

For details move the mouse over the datapoints

Temperatur: 14.17 °C a minute ago

02/28/2018 → 03/29/2018

Last 24 hours Last week Last month

22.0
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
-2.0
-4.0
-6.0

3/28/2018 12:00 AM

For details move the mouse over the datapoints

rel. Luftfeuchte: 40.89% a minute ago

Luftdruck: 9868.48 Pa a minute ago

<https://opensensemap.org>

Open Environmental Data Analysis



Daniel Nüst and Thomas Bartoschek
Institute for Geoinformatics
{daniel.nuest,bartoschek}@uni-muenster.de

Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code



Blog



Rendered by Travis CI & published on GitHub Pages:
<https://nuest.github.io/sensebox-binder/sensebox-analysis.html>

```
### Load data on senseBoxes

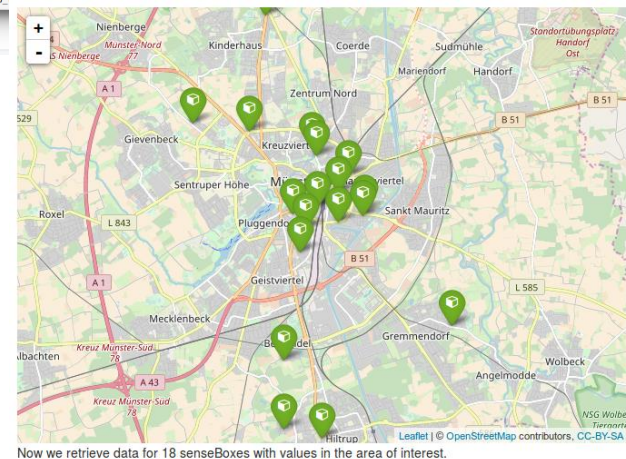
```{r load_box_data}
library("opensensmapr")
online <- TRUE # access online API or use local data backup?
analysis_date <- lubridate::as_datetime("2018-01-01 00:00:00")

if (online) {
 # retrieve data from openSenseMap API
 all_boxes <- osem_boxes()
 pm25_boxes <- osem_boxes(
 exposure = 'outdoor',
 date = analysis_date, # ±4 hours
 phenomenon = 'PM2.5'
)
} else {
 # update local data
 all_json <- toJSON(all_boxes, digits = NA, pretty = TRUE)
 write(all_json, file = here("data/all_boxes.json"))
 pm25_json <- toJSON(pm25_boxes, digits = NA, pretty = TRUE)
 write(pm25_json, file = here("data/pm25_boxes.json"))
}

load data from file and fix column types
all_boxes_file <- fromJSON(here("data/all_boxes.json"))

library("leaflet")
leaflet() %>%
 addTiles() %>%
 addAwesomeMarkers(data = ms_boxes,
 popup = ~paste0("Name: ", name, "
Id: ",
 "<a href='https://opensensemap.org/explore/', X_id, " ",
 "target='_blank'>, X_id, ""),
 label = ~name,
 icon = sense_icon)
```

The screenshot shows the RStudio interface on the left and a browser window on the right. The RStudio window displays the R source code for loading data and creating a map. The browser window shows the rendered HTML output, which includes a plot titled "Particulates measurements (PM2.5) on New Year 2017/2018" and a map of Münster, Germany, with 18 green location markers indicating the senseBox locations. The plot shows a significant spike in PM2.5 concentration around New Year's Eve.



Now we retrieve data for 18 senseBoxes with values in the area of interest.

# Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek  
Institute for Geoinformatics  
{daniel.nuest,bartoschek}@uni-muenster.de



Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code



Blog



## The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

<https://jupyter.org>

Open Standards for Interactive Computing

The Jupyter Notebook is based on a set of open standards for interactive computing. Think HTML and CSS for interactive computing on the web. These open standards can be leveraged by third party developers to build customized applications with embedded interactive computing.

- Notebook Document Format
- Interactive Computing Protocol
- The Kernel

**Synchronise** R Markdown and Jupyter Notebook with **ipyrm**d on every Rmd save:

`ipyrm --to ipynb --from Rmd -y -o sensebox-analysis.ipynb sensebox-analysis.Rmd`

<https://github.com/chronitis/ipyrm>

```
knitr::kable(data.frame(nrow(all_boxes), nrow(pm25_boxes)),
 col.names = c(
 "# senseBoxes",
 "paste1('# senseBoxes with PM2.5 measurements around', format(analysis_date, '%Y-%m-%d %T %Z'))"))

senseBoxes | # senseBoxes with PM2.5 measurements around 2018-01-01 00:00:00 UTC
-----|-----
1184 | 35
```

Exploring openSenseMap

The openSenseMap currently provides access to `r:nrow(all_boxes)` senseBoxes of which `r:nrow(pm25_boxes)` provide measurements of PM2.5 around `r:format(analysis_date, "%Y-%m-%d %T %Z")`.

The following map shows the PM2.5 sensor locations.

```
In [4]: plot(pm25_boxes)
```

# Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek  
Institute for Geoinformatics  
{daniel.nuest,bartoschek}@uni-muenster.de

Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code



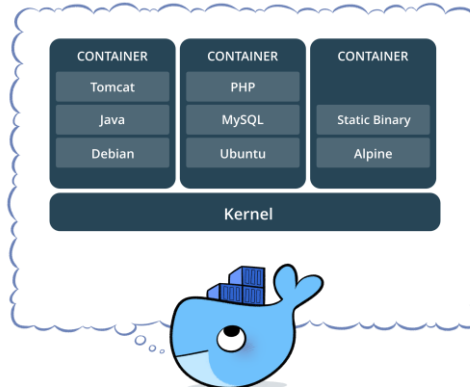
Blog



<https://github.com/rocker-org/binder/blob/master/Dockerfile>

```
1 FROM rocker/geospatial:3.4.2
16 ENV HOME /home/${NB_USER}
17 WORKDIR ${HOME}
18
19 RUN apt-get update && \
20 apt-get -y install python3-venv python3-dev && \
45 CMD jupyter notebook --ip 0.0.0.0

1 FROM rocker/binder:3.4.2
2
3 # Copy repo into ${HOME}, make user own $HOME
4 USER root
5 COPY . ${HOME}
6 RUN chown -R ${NB_USER} ${HOME}
7 USER ${NB_USER}
8
9 ## run any install.R script we find
10 RUN if [-f install.R]; then R --quiet -f install.R; fi
```



<https://docker.com>

# Docker & Rocker



Documented,  
portable,  
archivable, and  
one-click  
interactive  
**computational  
environments**  
for reproducibility  
based on  
Dockerfiles

The versioned stack

image	description
r-ver	Specify R version in docker tag. Builds on <code>debian:stable</code>
rstudio	Adds rstudio
tidyverse	Adds tidyverse & devtools
verse	Adds tex & publishing-related packages
geospatial	Adds geospatial libraries

```
daniel@gin-nuest:~$ docker pull rocker/r-ver:3.1.0
3.1.0: Pulling from rocker/r-ver
4176fe04cefe: Already exists
50c17ea00d6a: Pull complete
Digest: sha256:ca237b5b5f23cbb7465ff60777acce0de8d3293079541ff36814e5c7e730f03c
Status: Downloaded newer image for rocker/r-ver:3.1.0
```

```
daniel@gin-nuest:~$ docker run -it rocker/r-ver:3.1.0
```

```
R version 3.1.0 (2014-04-10) -- "Spring Dance"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-unknown-linux-gnu (64-bit)
```

```
> getOption("repos")
CRAN
"https://mran.microsoft.com/snapshot/2014-09-17"
> 1 + 1
[1] 2
>
```

<https://www.rocker-project.org>

<https://www.rocker-project.org/images/>

# Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek  
Institute for Geoinformatics  
{daniel.nuest,bartoschek}@uni-muenster.de

<https://mybinder.org>



Turn a GitHub repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

Build and launch a repository

GitHub repo or URL

Git branch, tag, or commit

 Path to a notebook file (optional) File ▾ 

Copy the URL below and share your Binder with others:

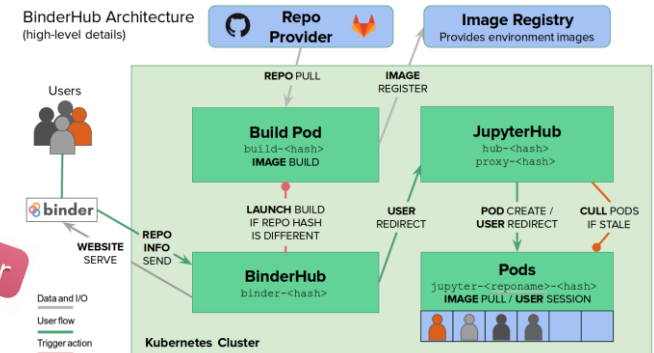
Copy the text below, then paste into your README to show a binder badge:

**Binder Team @mybinderteam · 27. März**  
A brief #humblebrag: over the last 30 days [mybinder.org](https://mybinder.org) has served >100000 sessions to ~72000 distinct users from 169 countries! 🤖

<https://twitter.com/mybinderteam/status/978603170250330112>

## How it works

- 1 Enter your repository information**  
Provide in the above form a URL or a GitHub repository that contains Jupyter notebooks, as well as a branch, tag, or commit hash. Launch will build your Binder repository. If you specify a path to a notebook file, the notebook will be opened in your browser after building.
- 2 We build a Docker image of your repository**  
Binder will search for a dependency file, such as requirements.txt or environment.yml, in the repository's root directory ([more details on more complex dependencies in documentation](#)). The dependency files will be used to build a Docker image. If an image has already been built for the given repository, it will not be rebuilt. If a new commit has been made, the image will automatically be rebuilt.
- 3 Interact with your notebooks in a live environment!**  
A JupyterHub server will host your repository's contents. We offer you a reusable link and badge to your live repository that you can easily share with others.



- Overview
- Open
- senseBox
- OSem
- R Markdown
- Jupyter
- {D,R}ocker
- Binder**
- Video
- Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



https://docs.google.com/presentation/d/1f5W4Rnezb6BRz4YXcXWYAx8t4KRfUosbCjS4Z1or7m/edit#slide=id.g25d5bc82125\_0\_53



# Open Environmental Data Analysis

Daniel Nüst and Thomas Bartoschek  
Institute for Geoinformatics  
{daniel.nuest,bartoschek}@uni-muenster.de

Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI [10.5281/zenodo.1135139](https://doi.org/10.5281/zenodo.1135139)



Code



Blog



## What happens?

Click on “Launch binder”

Open Jupyter Hub

Switch to Rstudio

Render the literate programming document

View workflow, plots, interactive map

The screenshot shows the GitHub repository page for 'nuest/sensebox-binder'. The repository description is 'A reproducible analysis of environmental data - from hardware to plot'. It has 19 commits, 2 branches, 0 releases, and 1 contributor. The license is Apache-2.0. The commit history is as follows:

Commit	Message	Time
revert add local data cache		Latest commit 4 hours ago
data	add local data cache	4 hours ago
gfignone	Update analysis in R Markdown	4 days ago
travis.yml	fix Travis build; add license to Rmd file	4 days ago
320px-Fireworks_2.png	Update analysis in R Markdown	4 days ago
Dockerfile	Initial setup	4 days ago
LICENSE	Initial commit	4 days ago
README.md	automatic conversion of ipynb file	a day ago
install.R	add local data cache	4 hours ago
sensebox-analysis.Rmd	add local data cache	4 hours ago
sensebox-analysis.ipynb	add local data cache	4 hours ago
sensebox-binder.Rproj	initial setup	4 days ago

<https://media.giphy.com/media/l49JRjO65S0WQ1Kyk/giphy.gif>

# Abstract & References

Overview

Open

senseBox

OSeM

R Markdown

Jupyter

{D,R}ocker

Binder

Video

Abstract + References

DOI 10.5281/zenodo.1135139



Code



Blog



Open Science is a major movement. It is deeply connected with other initiatives on openness, availability, and transparency: Open Access, Open Data, Reproducible Research, Free and Open Source Software (FOSS), and Open-source Hardware (OSH, [https://en.wikipedia.org/wiki/Open-source\\_hardware](https://en.wikipedia.org/wiki/Open-source_hardware)). Free to use platforms supporting Open Science become more recognised and many are FOSS themselves. They range from large, publicly-funded players such as Zenodo (<https://zenodo.org>), to community projects with first grant funding such as Binder (Pérez et al., 2017).

The senseBox community (<https://sensebox.de/en>) provides a Do-It-Yourself Toolkit for Citizen Science as Open Hardware and an Open Data API (<https://opensensemap.org>) for collecting, sharing, and visualising observations, including particulates (Pfeil et al., 2015). Both approaches were designed following usability engineering methods to enable broad use, i.e. in education or by non-professional users (Wirwahn Bartoschek, 2015).

We present an open environmental data analysis building on all these projects. A central R Markdown (<http://rmarkdown.rstudio.com/>) document provides a recipe for data retrieval (with backup in plain-text files), analysis, and visualisation with R (<https://r-project.org>).

The document was created in a containerised environment with Docker (<https://docker.com>), more specifically the Rocker (Boettiger, 2017) image (<https://hub.docker.com/r/rocker/binder/>) which provides Jupyter Notebook (<https://jupyter.org/>) and Rstudio (<http://rstudio.com/>) user interfaces.

The container's advantages are (a) portability, (b) completeness using configuration files and scripts, (c) at least limited archivability, (d) simple integration of browser-based interactive user interfaces, and (e) one click online execution with Binder. All resources are published on GitHub (<https://github.com/nuest/sensebox-binder>) and the Zenodo research data repository (Nüst, 2017).

We demonstrate an open analysis workflow for environmental data utilising numerous Open X projects. It showcases the potential of latest technology and may serve as a template towards development of best practices, but also provides insight into remaining stumbling blocks. Connecting all these parts requires some technical proficiency. In the future reproducible computational workflows should become part of researcher's education, making them aware of simple yet effective practices (e.g. text-based formats, scripted analyses). In the meantime templates and semi-automatic assistants (cf. Nüst et al., 2017) can mitigate challenges.

**Geophysical Research Abstracts**  
**Vol. 20, EGU2018-17461**

**EGU General Assembly 2018**  
Boettiger, C. and Edelbuettel, D., 2017. An Introduction to Rocker: Docker Containers for R. The R Journal, RJ-2017-065. [url](http://www.rjournal.org/)

Nüst, D., 2018. Open environmental data analysis with senseBox, openSenseMap, Jupyter Notebook, RStudio, and BinderHub (Version v2). Zenodo. doi: 10.5281/zenodo.1139929

Nüst, D., Konkol, M., Pebesma, E., Kray, C., Schutzzeichel, M., Przybytzin, H., Lorenz, J. Opening the Publication Process with Executable Research Compendia. D-Lib Magazine. 2017. doi: 10.1045/january2017-nuest

Pérez, F., Granger, B., and Ragan-Kelley, B., 2017. Binder: enabling sharing and publication of reproducible computational research. <https://figshare.com/s/e9d0ad7bdc4e405cccfa>

Pfeil, M., Bartoschek, T., and Wirwahn, J., 2015. OpenSenseMap – A Citizen Science Platform for Publishing and Exploring Sensor Data as Open Data. FOSS4G Conference Proceedings, pp. 122-139, Seoul, 2015. [url](http://www.foiss4g.org/)

Wirwahn, J. and Bartoschek, T., 2015. Usability Engineering for Successful Open Citizen Science. FOSS4G Conference Proceedings, pp. 68-78, Seoul, 2015. [url](http://www.foiss4g.org/)