

ET site repository

Massimiliano Razzano
(University of Pisa & INFN-Pisa)

On behalf of the WG1

Main goals

- **Data repository**
 - Have a place to store the data
 - Accessible from the site and other places
 - Doing some interactive analysis
- **Software repository**
 - Standard development tools online
 - Software versioning
 - Developers groups etc

Software repository

- **Developed on Github**
 - Git-based, (almost) easy to use
 - Interface with CI tools (e.g. CircleCI)
 - Public and (special conditions) private repositories
 - Offers great interface with other services (notebooks, overleaf, etc..)
- **Our sw repo**
 - Created an ET organization (in Github jargon, shared account for group of projects/developer)
 - Everyone in our group can get access (see later)
 - Special bonus: private repositories (under unipi license)
 - Created some test projects, and group of projects

- <https://github.com/et-sw>

The screenshot shows the GitHub interface for the repository 'et-sw'. At the top, there is a navigation bar with the GitHub logo, a search bar, and links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. The repository name 'et-sw' is displayed with a profile picture of a sunset over a telescope. Below the name, there are tabs for 'Repositories 5', 'Packages', 'People 4', 'Teams 1', 'Projects', and 'Settings'. A search bar for repositories is present, along with filters for 'Type: All' and 'Language: All', and a 'New' button. The repository list shows two items: 'et-installer' (Private) and 'etsite-wp1' (Private). The 'et-installer' repository is described as 'ET software installer' and is written in Python, using the GPL-3.0 license. The 'etsite-wp1' repository is described as 'Work Package 1'. On the right side, there are sections for 'Top languages' (showing Python) and 'People' (showing 4 contributors).

Search or jump to... Pull requests Issues Marketplace Explore

et-sw

Repositories 5 Packages People 4 Teams 1 Projects Settings

Find a repository... Type: All Language: All Customize pins New

et-installer Private
ET software installer
Python GPL-3.0 0 0 0 0 Updated on Jun 1

etsite-wp1 Private
Work Package 1
0 0 0 0 Updated on May 29

Top languages
Python

People 4 >

Invite someone

Data repository

- **Virtual machine dedicated to our analysis**
- Hosted @ Computing center of the University of Pisa computing center (request by F. Fidecaro & M. Razzano)
- 8 cores (→ upgraded to 16)
- RAM: 16 Gb → upgraded to 32
- Large disk space: 4.6 Tb
- OS: Centos7 64 bit
- Performance can be increased in need, so far so good
- Can be accessed remotely

Data Repository organization (I)

- **Data space (/etrepo)**

- Temporary data-sandbox (writable by users), used to temp store input data
- data-sites: permanent data storage (read only)
 - Automatic script to move from sandbox → data-sites
- Under sandbox and data-sites contains subdir named as sites
 - Names so far present and agreed with Jan: sosenattos, terziet
 - Work to re-organize datasets



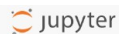
Data Repository organization (II)

- **ET software (/etrepo/et-software)**
 - Meant to host general-purpose sw and ET-specific software
- **User work directories (now under /etrepo/etuser-work/\$USER)**
 - Workspace for each user.
 - Use this space for your analysis, not your /home (here you have more space!)

Interactive notebooks

- **What about looking at the data?**
 - We have set up a JupyterLab server to allow analysis
 - Each user can access the interface
 - Can run Linux terminal Python IDE, and of course Python Notebooks
 - Using SSL encryption-based protocol

Accessible at <https://etrepo.df.unipi.it:8000>



Sign in

Username:

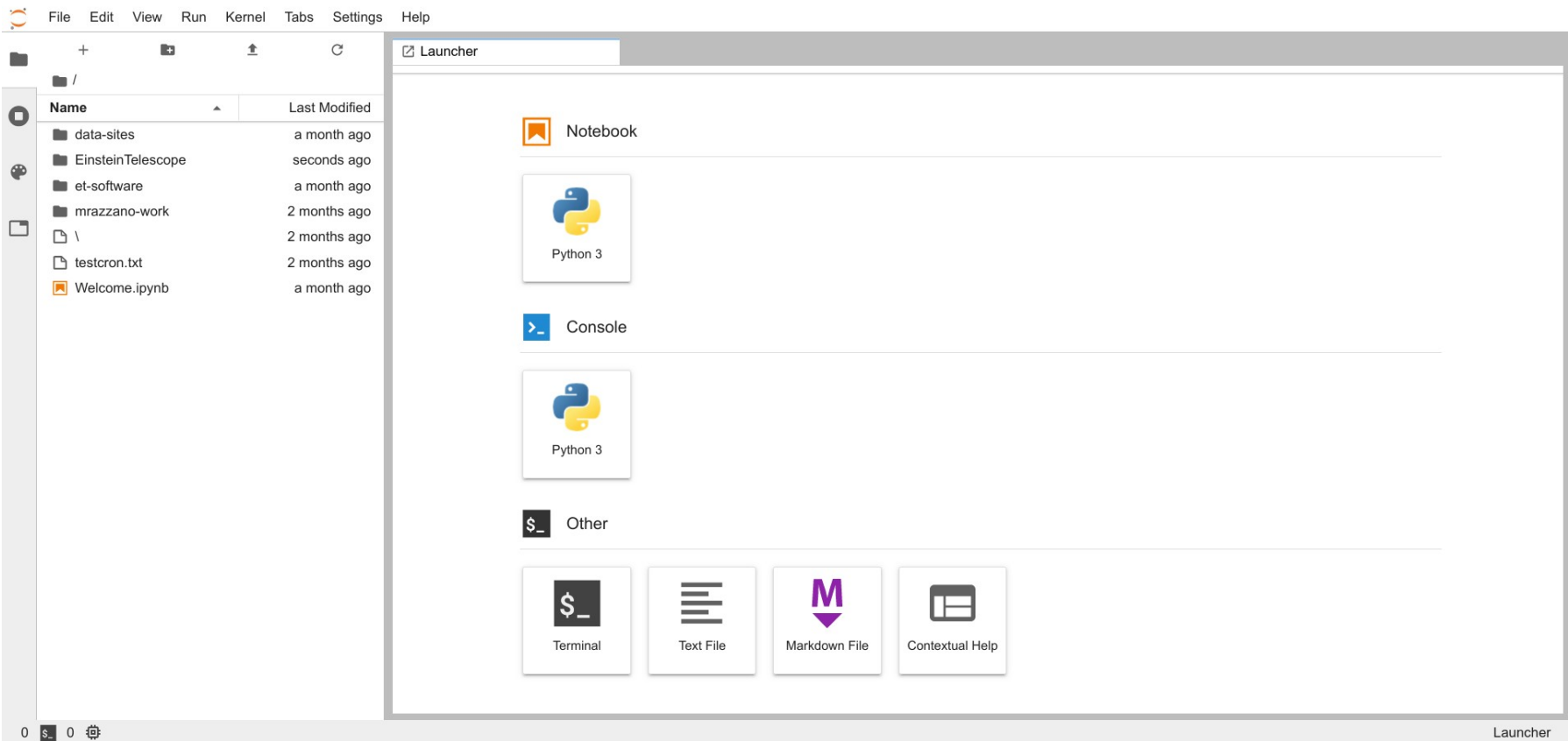
Password:

Sign In

Interactive notebooks

File Manager

Workspace



Interactive notebooks

Welcome (sample) notebook

```
[1]: #Test imports

import os

import numpy as np
import scipy as sp
import pandas as pd
import matplotlib.pyplot as plt
import obspy

print("All imports are successful!")

All imports are successful!

[2]: #define some directories
home_dir = os.environ["HOME"]
data_dir = os.path.join(home_dir, "data-sites")

[3]: import glob
glob.glob(os.path.join(data_dir, "**"))

[3]: ['/home/mrazzano/data-sites/terziet', '/home/mrazzano/data-sites/sosenattos']

[ ]:
```

Mode: Command Ln 1, Col 1 Welcome.ipynb

Howto: get access

- **Automatic accounting system**
- Fill the form at <https://forms.gle/n2MpK1cg2Mxfdz1o8> (sent around by email, will send again and put in a more convenient place)
 - Scripts will take your requests, make an account for you, set up directories and send an email to you with username and temp pwd
 - Usernames as name+surname → nsurname
 - Latency half hour, can be longer in some cases
 - Will also add github account to et-sw (I am doing it, so it will take some extra time)
 - In case of problems, let me know!

Documentation at <https://tinyurl.com/y4ukh98d>

Howto: access and upload data

- **Access**
 - Now via ssh and scp
 - Implement (if useful) web-based
- **Upload data**
 - Scp to the data-sandbox/sitename
 - The system will copy the files to the final storage space (data-sites)
- **Use JupyterLab interface**
 - Go to <https://etrepo.df.unipi.it>
 - Login with your username and pwd
 - Write code and enjoy

- **Repositories**
 - Software repository ready and working in Github
 - Data repository upgraded and running
 - Automatic accounting system
- **Status**
 - Repositories are up and running
 - Data already copied (Luca, Jan, Carlo)
- **Future perspectives**
 - Update the organization of material
 - Upgrade the packages on the Jupyter nb (e.g. obspy)