



SCHOOL OF DATA ANALYSIS

Everware - effortless research booster

Igor Babushkin¹, Tim Head², Andrey Ustyuzhanin^{3,4}

¹University of Manchester, ²EPFL, ³YSDA, ⁴NRC "Kurchatov Institute"

Intro

› 2013 - rootnotes.py

<https://gist.github.com/mazurov/6194738>

› 2014 - reproducible experiment platform

<https://github.com/yandex/rep>

› 2015 - everware

<https://github.com/everware>

<http://everware.xyz>

Everware

- › simplify sharing of your ideas & code
 - › jupyterhub + docker (swarm)
- › re-think collaborative patterns
 - › reproduce
 - › learn
 - › reuse

everware

Sharing is caring

URL of git repository

Launch

Paste the link to the git repository you want to try out. If you need some inspiration try one of the following repositories:

- <https://github.com/everware/everware-dimuon-example>
- <https://github.com/betatim/everware-demo>

Read the documentation to [learn how to make your repositories work with everware.](#)

Powered by [Jupyter](#) and [Docker](#)

Logged in as [anaderi](#) (Logout)

`https://everware.rep.school.yandex.net`

Use-case

- › develop your analysis in a comfortable environment, e.g.
 - › notebook + plane
 - › ipad/smartphone
 - › server-based
- › sync with github + Dockerfile
- › impress / present slides
- › share single link with your peers to have it running &
 - › behave in expected way
 - › get same result
 - › ready for reuse

Contexts

- › education
- › hackathon
- › research

Behind the scenes

- › Docker (+swarm)

- › docker swarm spawner based on dockerspawner
<https://github.com/everware/dockerspawner>
- › containers <https://github.com/everware/container-tools>

- › JupyterHub

- › GitHub

- › tools & docs to simplify docker image creation & playing with analysis locally

Ideas for the future, roadmap

- › attach custom CPU & Disk storage, e.g.
 - › e.g. DigitalOcean or Amazon AWS
 - › university clusters
 - › Rackspace Carina
- › transparent parallel execution
 - › ipyparallel, CERN grid, HPC(?)
- › integrate with disk storage
 - › CERN grid
 - › dropbox/Google drive/CERN EOS/CERN box
- › integrate with knowledge exchange platforms
 - › openml.org
 - › peerj.com
 - › open-data CERN portal
 - › docker-compose support

Birds of a feather

- › <https://cloud.sagemath.com/>, <http://wakari.io>,
<https://juliabox.org/>
 - › limited support for containerization & customization of environment
- › <http://mybinder.org>
 - › no support for parallel execution
 - › some Docker limitations
 - › lack of execution on own hardware
- › <http://pachyderm.io>
 - › lacks github integration
- › <https://www.dominodatalab.com>
 - › private, closed platform

Wish-list for Jupyter team

- › notebook diff (nbdiff?)
- › transparent integration with git (merging?) simplify sharing & showing presentation
 - › PDF generation
 - › sockets.io (supported by reveal.js)
- › support for docker swarm for ipyparallel
- › documentation
 - › remove outdated documentation
 - › overall guide on Jupyter family projects for new practitioners (jupyter, kernels, ipyparallel, jupyterhub, tmpnb, nbconvert, Hydrogen, ...)
- › simultaneous notebook editing
 - › keep notebook state to server
- › Make JupyterHub a bit more extensible
 - › add attributes to UserObject without modifying code of JupyterHub (by inheritance?)

Thank you!

Contacts

Andrey Ustyuzhanin

andrey.ustyuzhanin@cern.ch



anaderiRu@twitter

Tim Head

thead@cern.ch



betatim@twitter