



SCHOOL OF DATA ANALYSIS

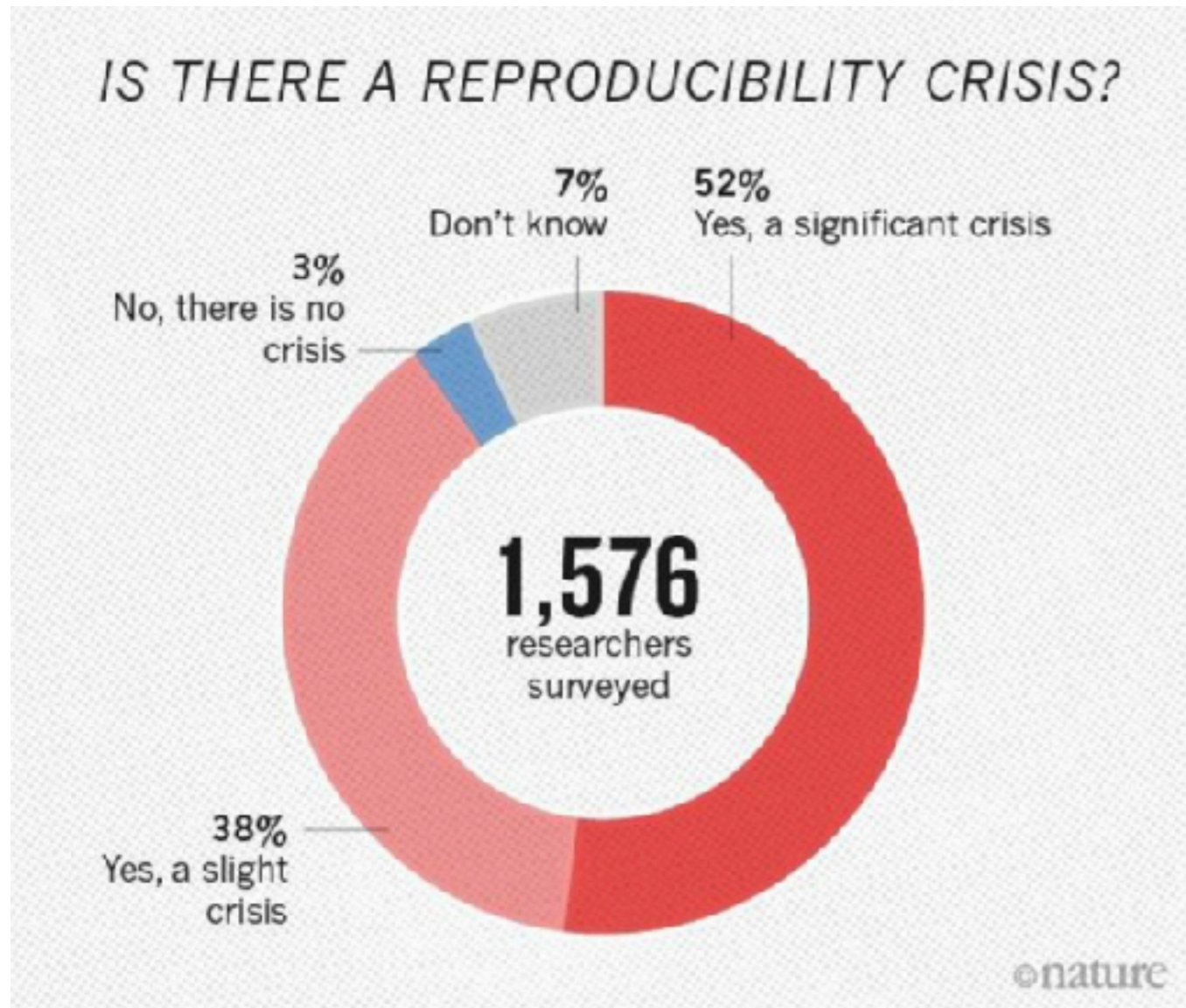


# Everware - lowering reproducibility barriers

Andrey Ustyuzhanin

Yandex School of Data Analysis and  
Higher School of Economics, Moscow, Russia

# Scientific Reproducibility



Nature: 1,500 scientists lift the lid on reproducibility by Monya Baker

# Challenge-driven education

Imagine Cup, <http://imaginecup.com/>

Hackathons, e.g., <http://webfest.web.cern.ch/>

Open data days, <http://opendataday.org/>

Guide to Challenge Driven Education, <https://www.kth.se/social/group/guide-to-challenge-d/>

## Platforms

Kaggle, <https://www.kaggle.com/>

Codalab, <https://competitions.codalab.org/>

**Complication and boost factors are similar to research reproducibility.**

# Reproducibility boost factors

## Practical

better mentoring/supervision

more within-lab validation

simplified external-lab validation

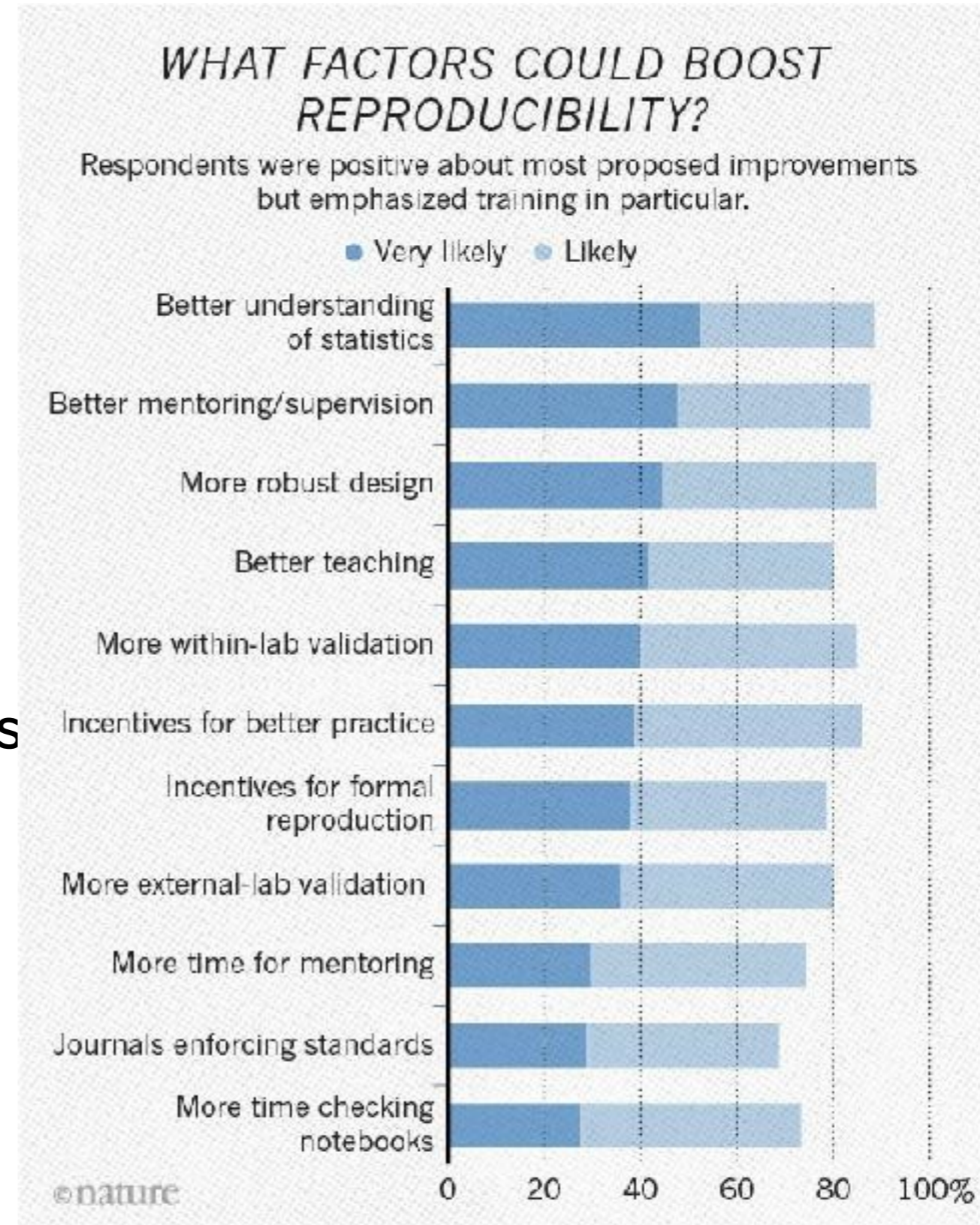
incentive for better practice

robust design

## Educational

- wider access to the best practices

- better teaching





# Everware

- › ... about re-usable science, it allows people to jump right into your research code. Lets you launch Jupyter notebooks from a git repository with a click of a button.



# Prerequisites

data: published on any website (**CERNBOX**)

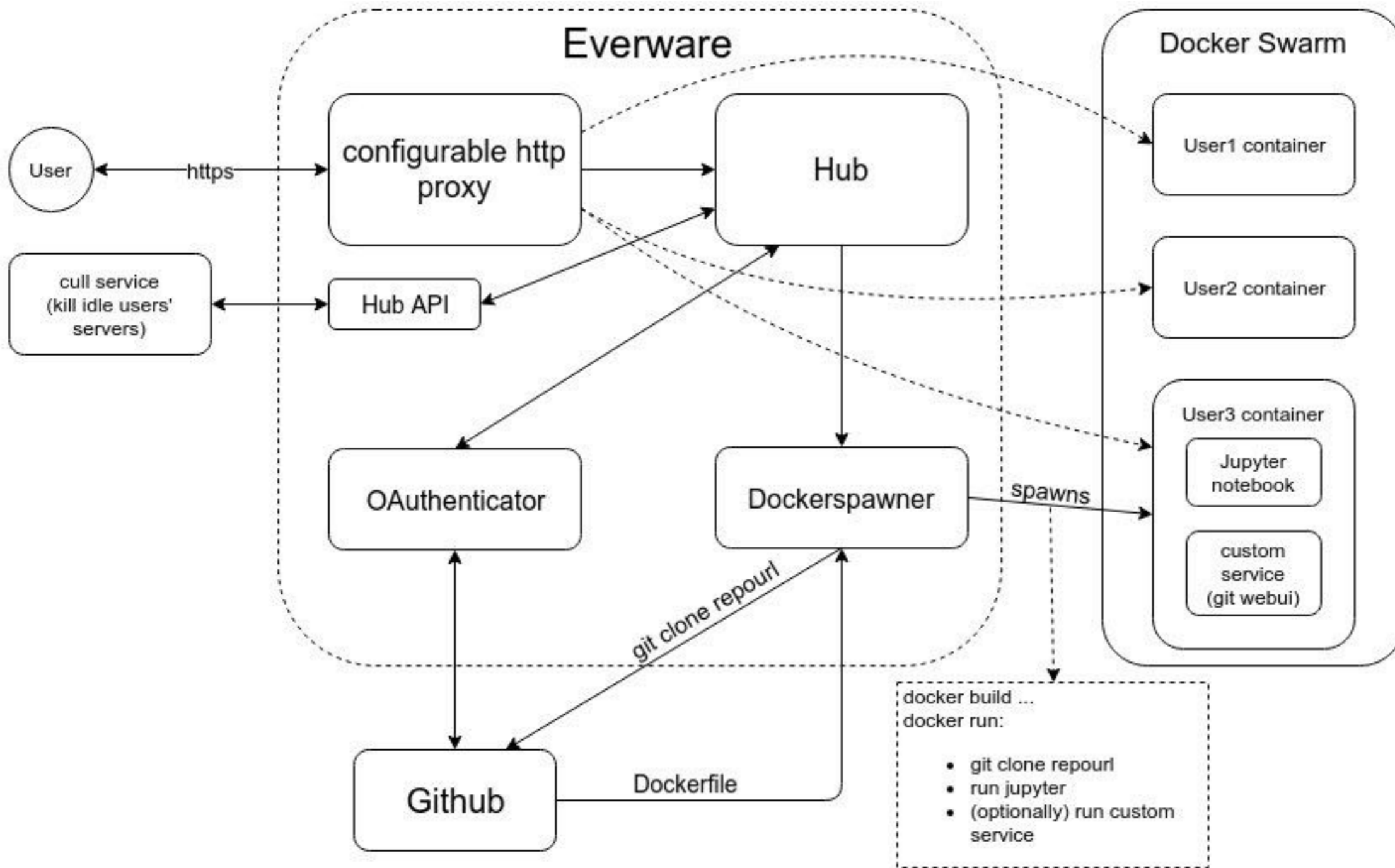
all the code: **github**

environment configuration: **github** (conda, requirements.txt)

run-time: either public or private cloud

> [everware.rep.school.yandex.net](http://everware.rep.school.yandex.net),  
[everware-01.cern.ch](http://everware-01.cern.ch)

# How everware works



# Example repositories

algorithm meta-analysis,

[https://github.com/openml/study\\_example](https://github.com/openml/study_example)

gravitational waves,

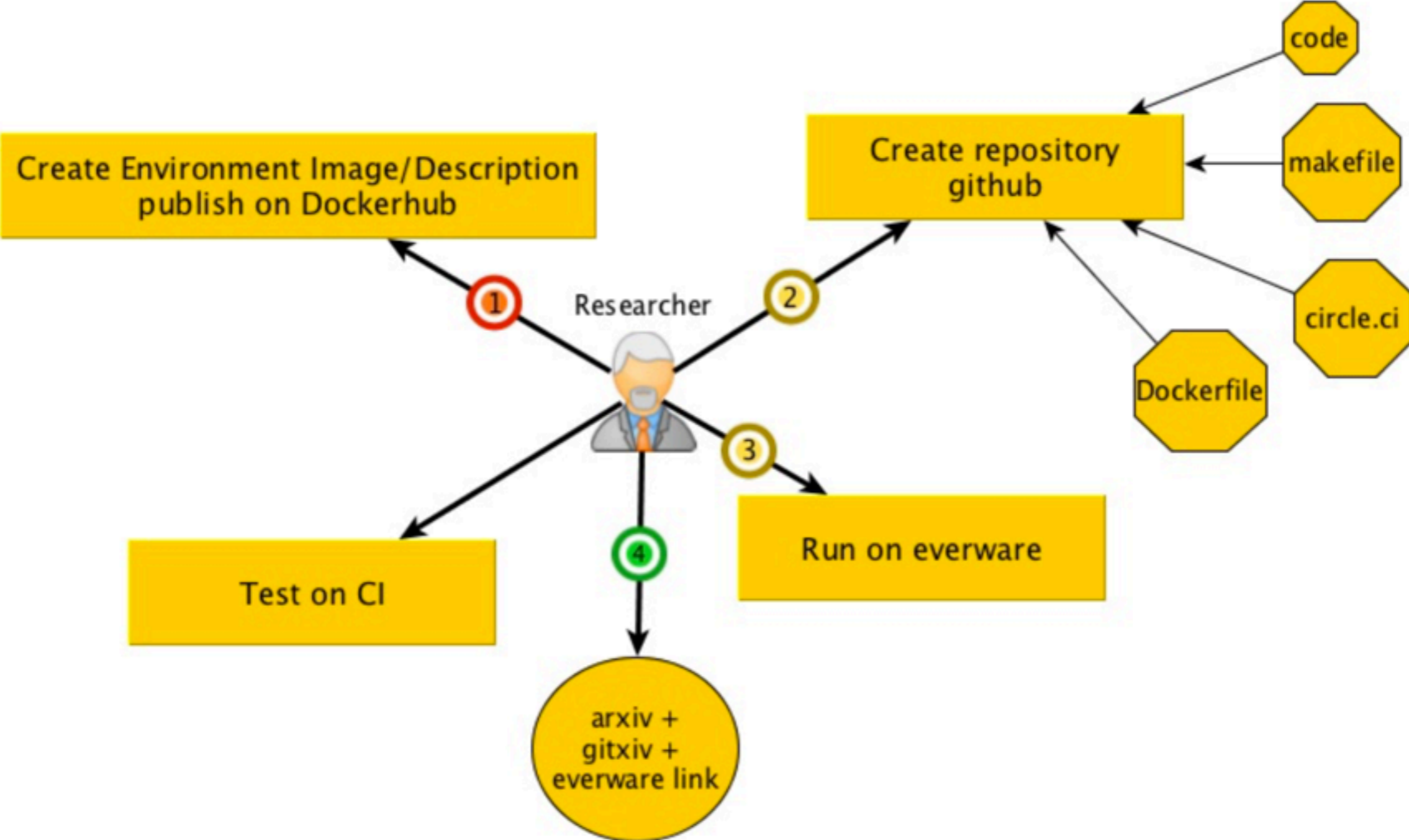
<https://github.com/anaderi/GW150914>

COMET tracking,

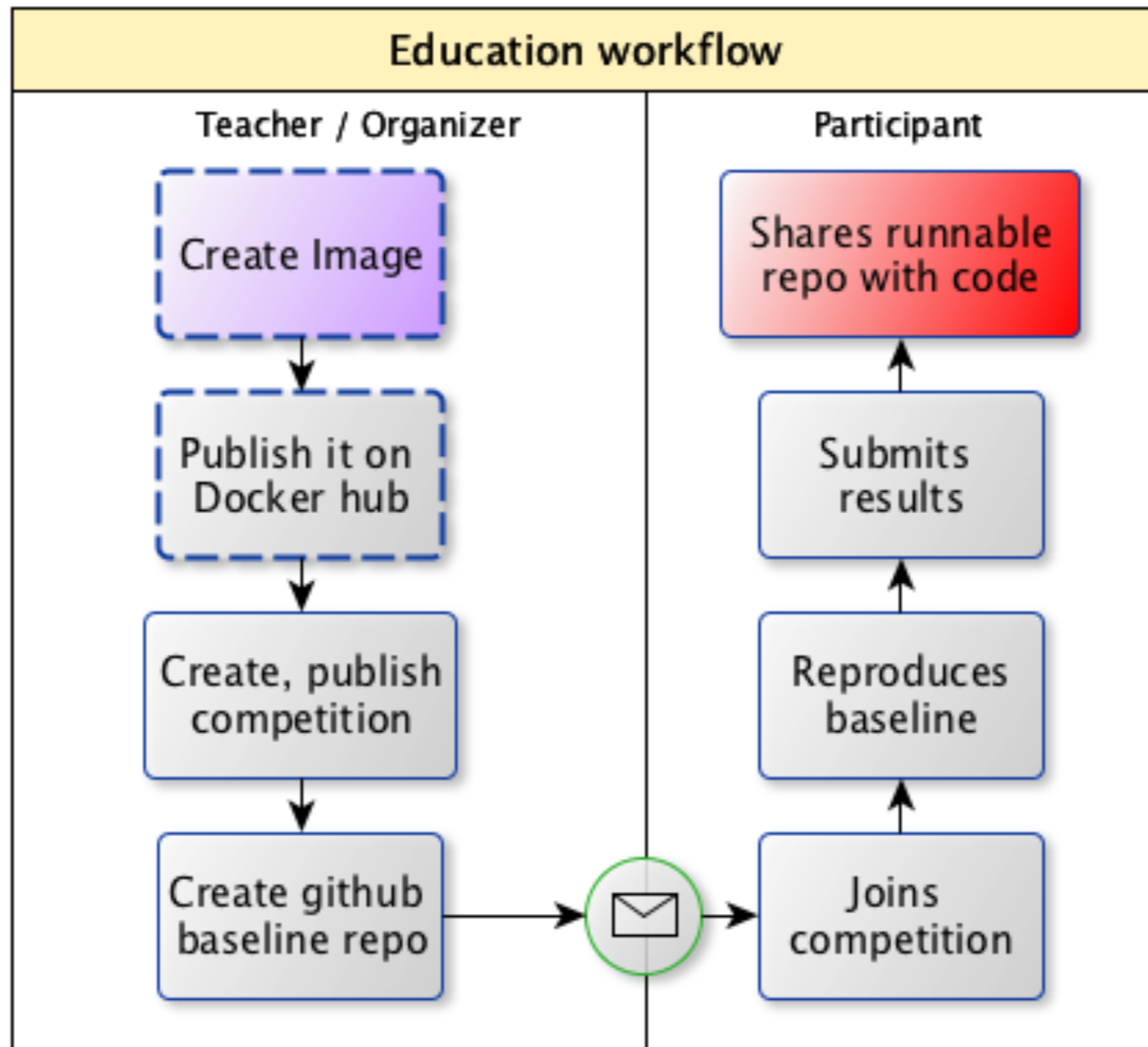
<https://github.com/yandexdataschool/comet-example-ci>



# Basic research workflow



# Education / outreach workflow



# Educational / outreach examples

Python course at YSDA 2015

Machine Learning in High Energy Physics summer school 2016

YSDA course on Machine learning at Imperial College London  
2016, 2017

Kaggle competitions 2016

Machine learning course at University of Eindhoven

LHCb open data masterclass

# Discussion / Conclusion

## Similar projects

- › SWAN, binder

## Research:

- › How to combine analysis scripts with Jupyter analysis model?

## Education:

- › LHCb masterclass?
- › other LHCb educational events?

Thanks for attention!