# The reproduceability crises

02457 Machine Learning Operations

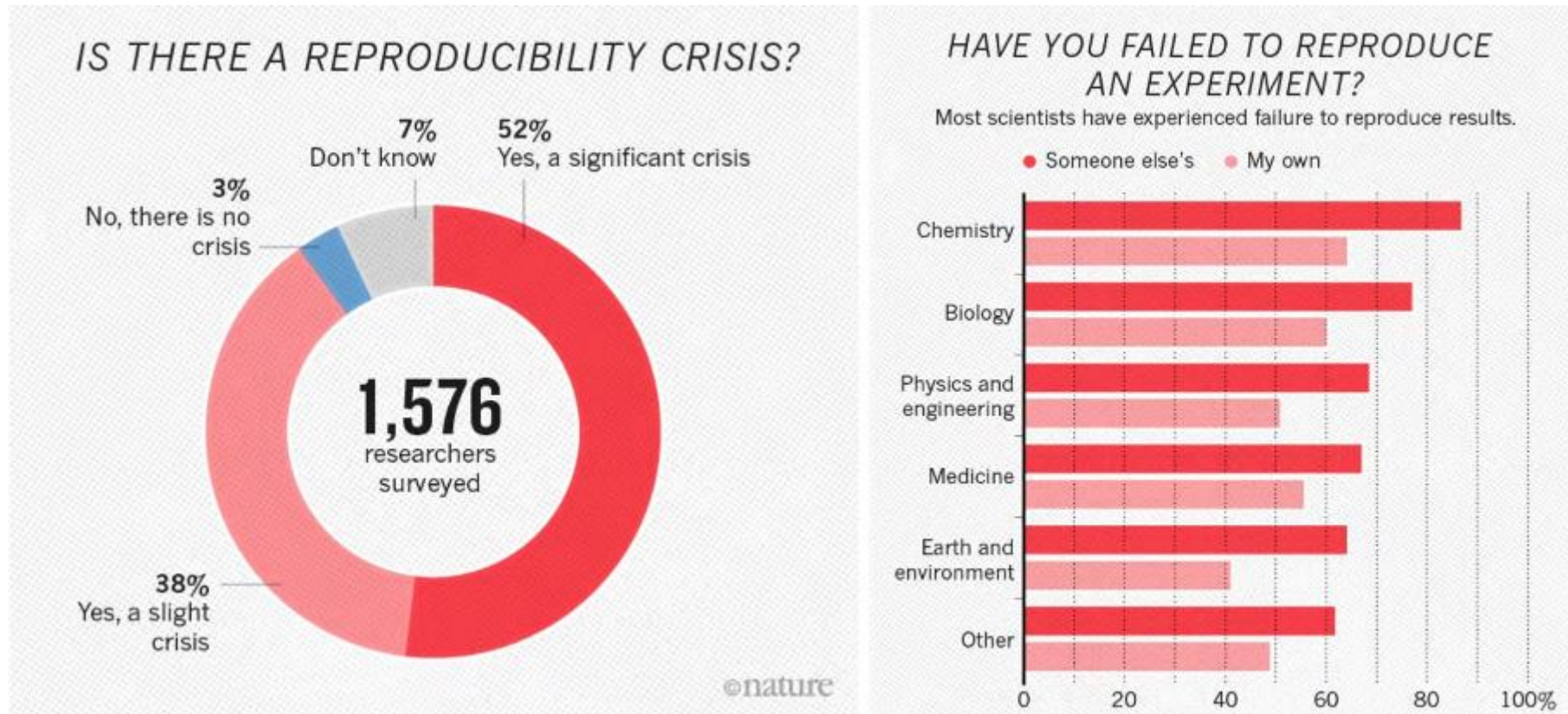Nicki Skafte Detlefsen,

Postdoc

DTU Compute

# What is it?

- Being able to reproduce other peoples experimental results is an essential part of the scientific method

- Well known problem thoughout most fields (physics, chemestry, biology and computer science)

- With the rise of deep learning, the problem has only been made worse due to competition



This is where it breaks

Nicki Skafte Detlefsen

# How bad is it?

Wow its bad...



IS THERE A REPRODUCIBILITY CRISIS?

7% Don't know
52% Yes, a significant crisis
3% No, there is no crisis
1,576 researchers surveyed
38% Yes, a slight crisis

©nature

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?
Most scientists have experienced failure to reproduce results.

● Someone else's   ● My own

Chemistry
Biology
Physics and engineering
Medicine
Earth and environment
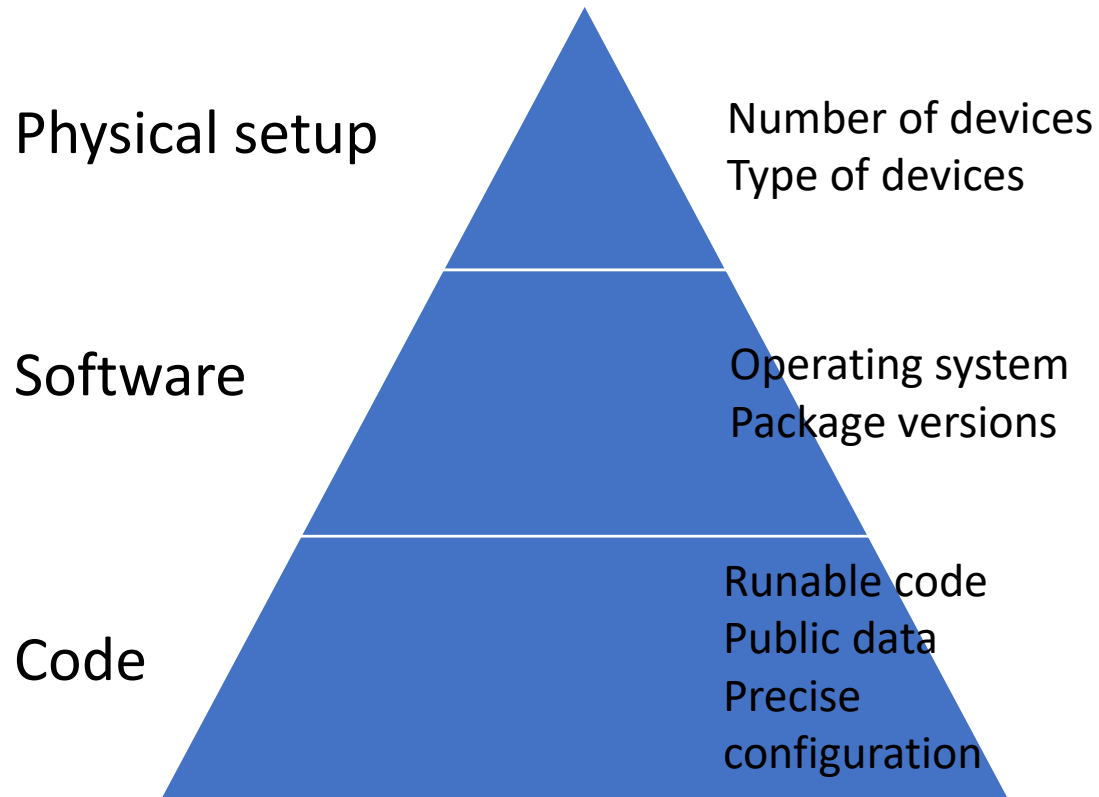Other

0   20   40   60   80   100%

https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

# What can we do about it

- Make sure to document everything about our experiments

Physical setup

Number of devices
Type of devices

Software

Operating system
Package versions

Code

Runable code
Public data
Precise
configuration

THE SCIENTIFIC METHOD

Result — Observation — Question — Hypothesis — Experiment — Conclusion

Document this step as thoroughly as possible

# Closer look

MLOps

**Physical setup**

Number of devices
Type of devices

Easy to document

**Software**

Operating system
Package versions

Ask Conda

**Code**

Runable code
Public data
Precise configuration
Seed

This is your job

Nicki Skafte Detlefsen
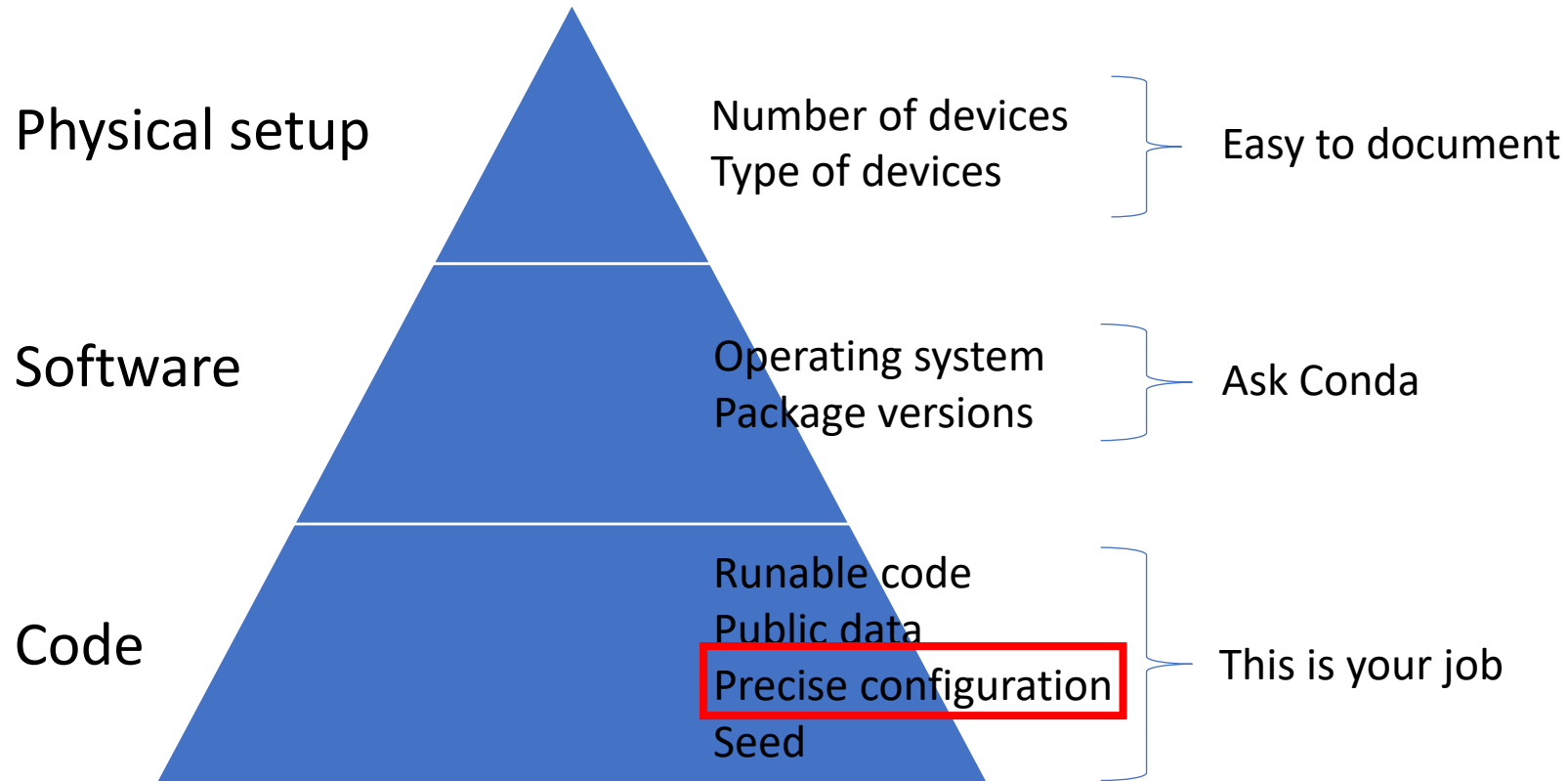
# Configurations

There is a lot of subjective choices that we do when running experiments, most noteable the hyperparameters.

| Parameters in script | Argparser | Config files |
|---|---|---|
| ```<br>class hparams:<br>    lr = 0.1<br>    batch_size = 16<br>    num_layers = 5<br>``` | ```<br>python my_script.py \<br>    --lr 0.1 \<br>    --batch_size 16 \<br>    --num_layers 5<br>``` | ```<br>experiment1.yaml<br><br>lr: 0.001<br>batch_size: 16<br>num_layers: 5<br><br>python my_script.py \<br>    config=experiment1.yaml<br>``` |
| Not easy to configure<br>Experiments may be lost if not carefull | Easy to configure<br>Falls on user to save the configuration | Easy to configure<br>Parameters are systematically saved with the experiment |

# Hydra



*A framework for elegantly configuring complex applications.*

https://github.com/facebookresearch/hydra

# Other frameworks

- https://github.com/IDSIA/sacred

- https://mlflow.org/

- https://wandb.ai/site