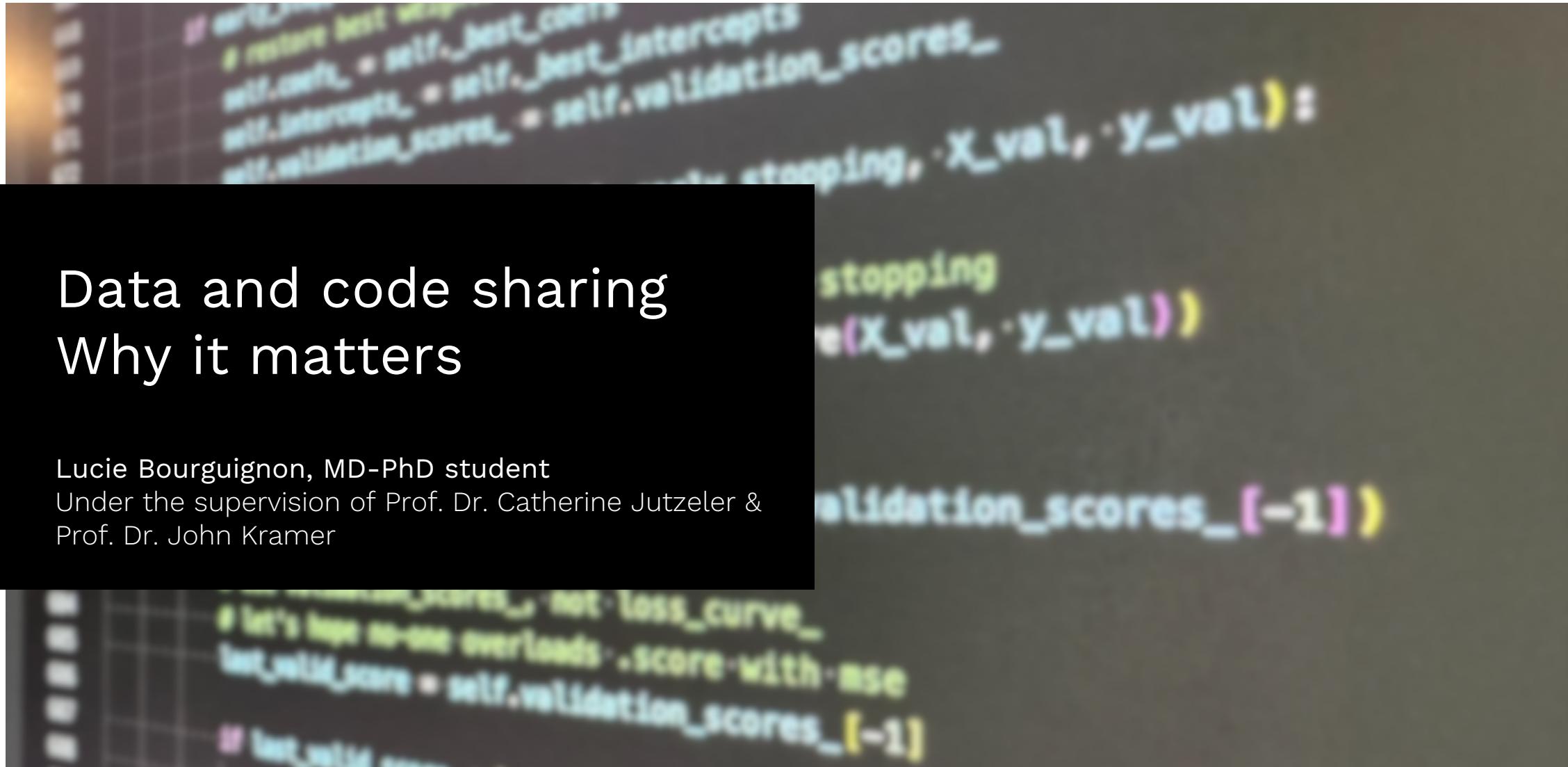


Data and code sharing Why it matters

Lucie Bourguignon, MD-PhD student

Under the supervision of Prof. Dr. Catherine Jutzeler &
Prof. Dr. John Kramer



```
# return best score, best intercepts
    self.best_score = self.best_intercepts = None
    self.intercept_ = self.best_intercept_
    self.validation_scores_ = self.validation_scores[-1]
    self.validation_score_ = self.validation_scores_[-1]

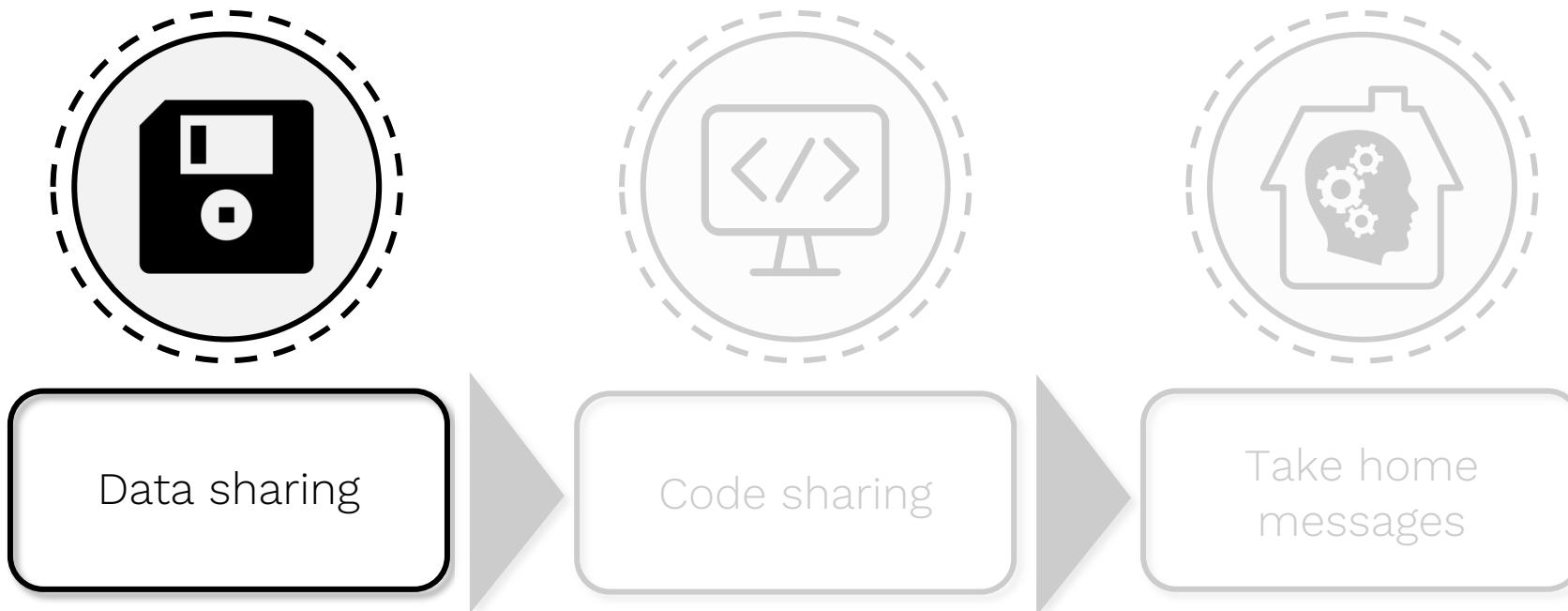
    # stopping
    while stopping(x_val, y_val)):

validation_scores_[-1])
```

```
# let's hope we can overrule .score with mse
best_val_score = self.validation_scores[-1]
```

```
# best_val score
```

Outline



Data sharing

Introduction



Process of making the same data resources available to multiple applications, users, or organizations, while maintaining data fidelity across all entities consuming the data. [[source](#)]

1

Have you ever used open source data?

2

Have you ever shared your data with someone?

Data sharing

Sharing locally

1

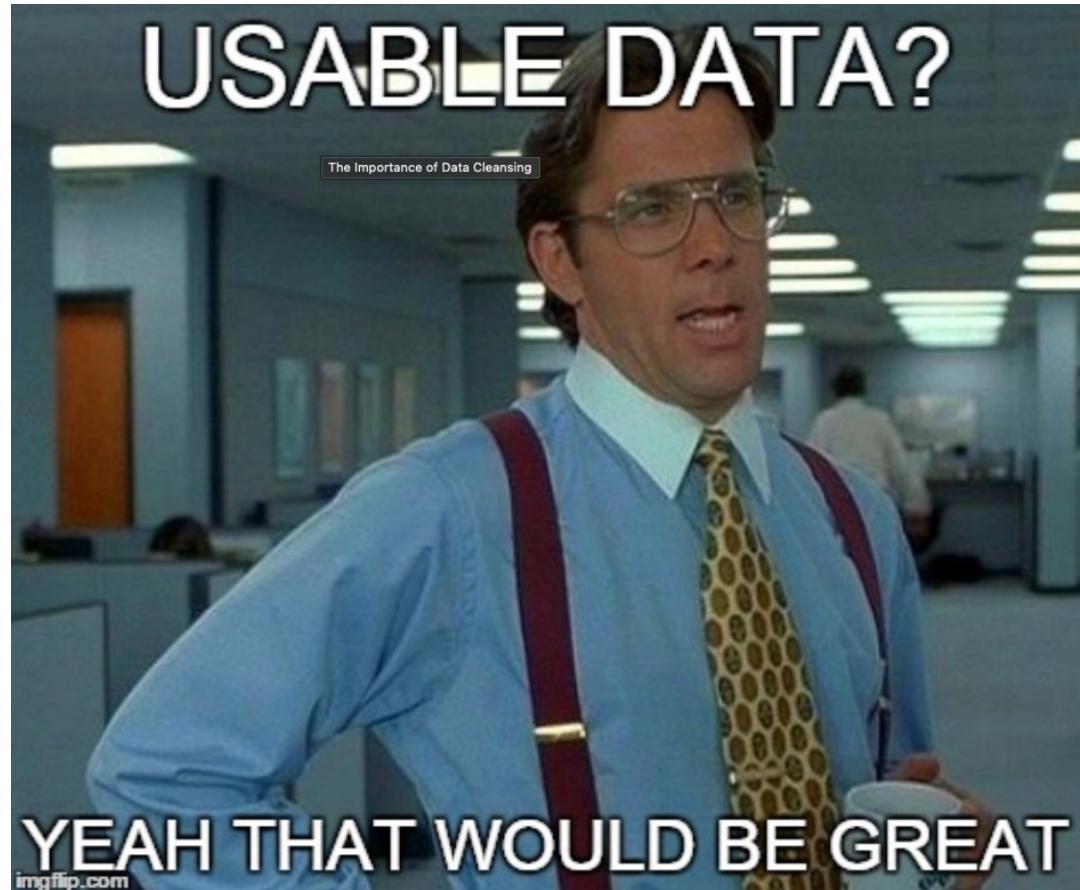
The data you collect is a property of the lab.

2

Common formatting is important (cf Jessie's lab meeting presentation).

3

Document your data.



Data sharing

Sharing broadly – FAIR principles

Findability

Accessibility

Interoperability

Reuse of digital assets



Guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. [[source](#)]

[FAIR principles](#) and [original paper introducing FAIR Data Principles](#)

Data sharing

Sharing broadly – Data sharing agreement



Agreement that sets out the purpose of the data sharing, cover what happens to the data at each stage, set standards and help all the parties involved in sharing to be clear about their roles and responsibilities. [[source](#)]

The further your collaborator,
the harder it gets

Data sharing

Important considerations

1

Do not share data on GitHub.

2

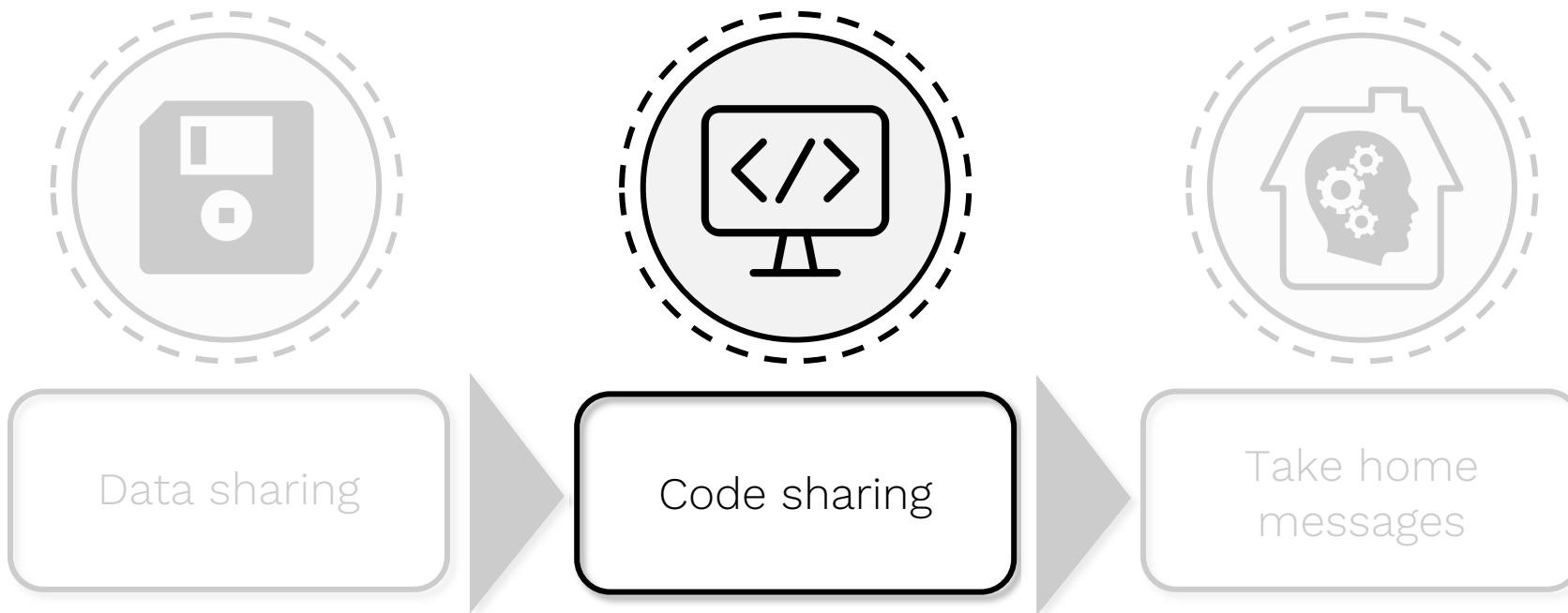
Do not leave patient ID in your code.

3

Do not store data locally.

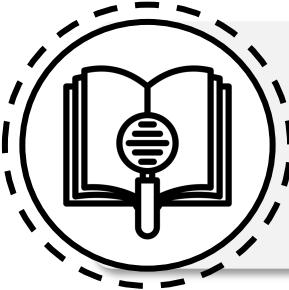


Outline



Code sharing

Introduction



Range of activities, including sharing code *privately* (e.g., with your colleagues as part of internal code review) or *publicly* (e.g., as part of a journal article submission).

1

Have you ever benefited from someone sharing code with you?

2

Have you ever shared your code with someone?

Code sharing

Mentimeter



Code sharing

Benefits and barriers



- ✓ Reduce duplication effort
- ✓ Encourages transparency
- ✓ Continuity in work
- ✓ Captures work in data management



- ✗ Additional time and effort required
- ✗ Potential criticism
- ✗ Not knowing how to share
- ✗ Plagiarism

Inspired by "[The benefits and barriers to code sharing as an Early Career Researcher](#)"

Code sharing

Addressing barriers – Time & effort

1

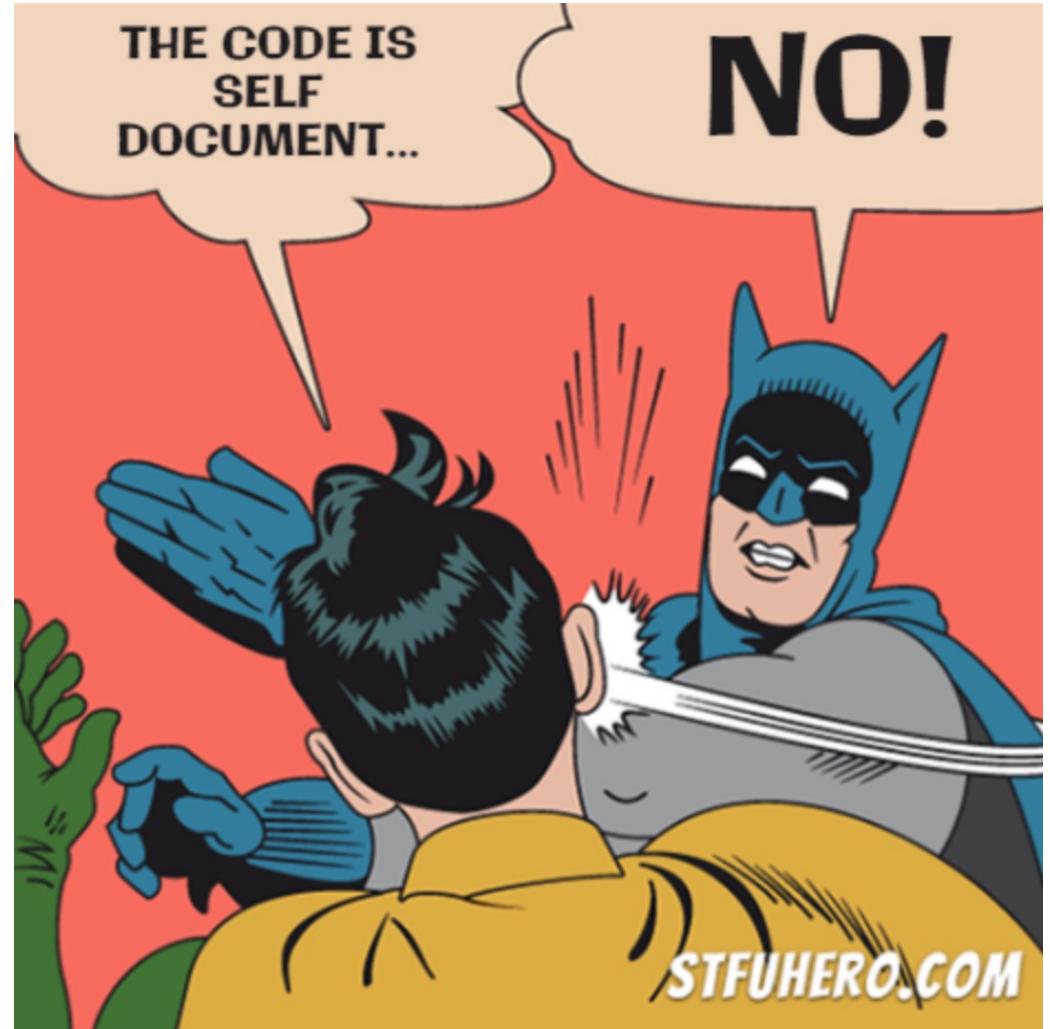
Cleaning and documenting your code is part of the job.

2

Do it for others but also for your future self.

3

Sharing messy code is better than not sharing.



Code sharing

Addressing barriers – Potential criticism



Only very few people are coding experts.

1

Start with good practices as early as possible.

see Additional resources

2

If there is a bug, it is better if someone finds it.

3

Code sharing

Addressing barriers – How to share

1

Use GitHub and GitHub desktop.

2

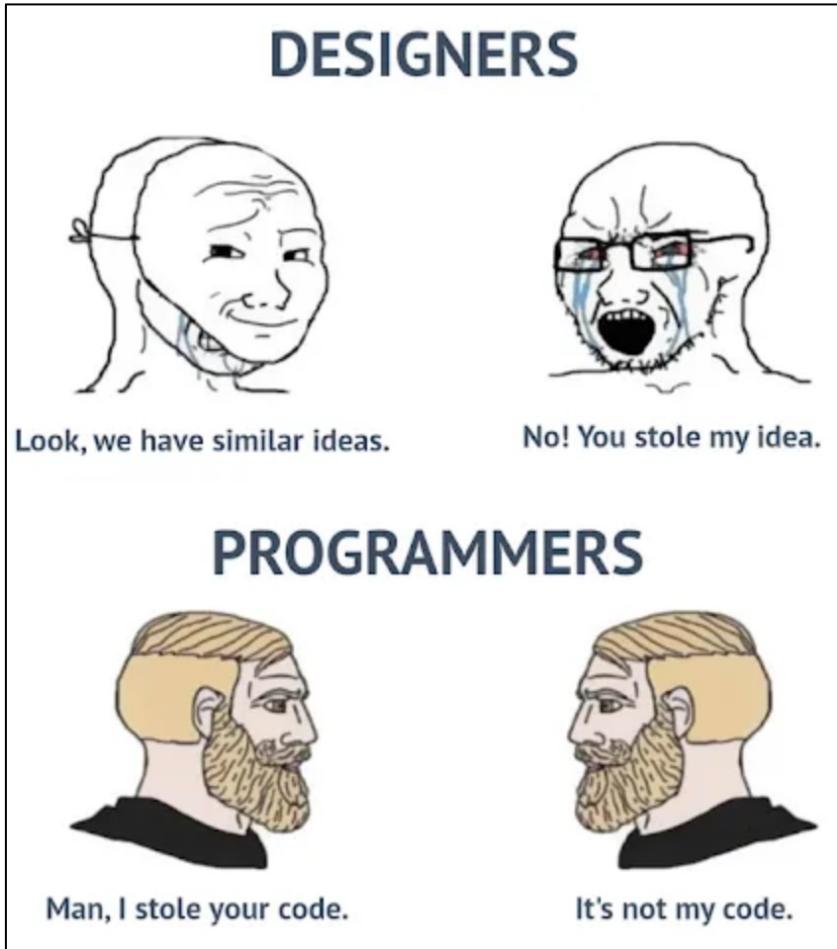
Get your code reviewed
the same way you would
get it for your manuscript.

Exercise break



Code sharing

Addressing barriers – Plagiarism



Copying code from others
is normal.

1

You can put a license on
your code to protect it.

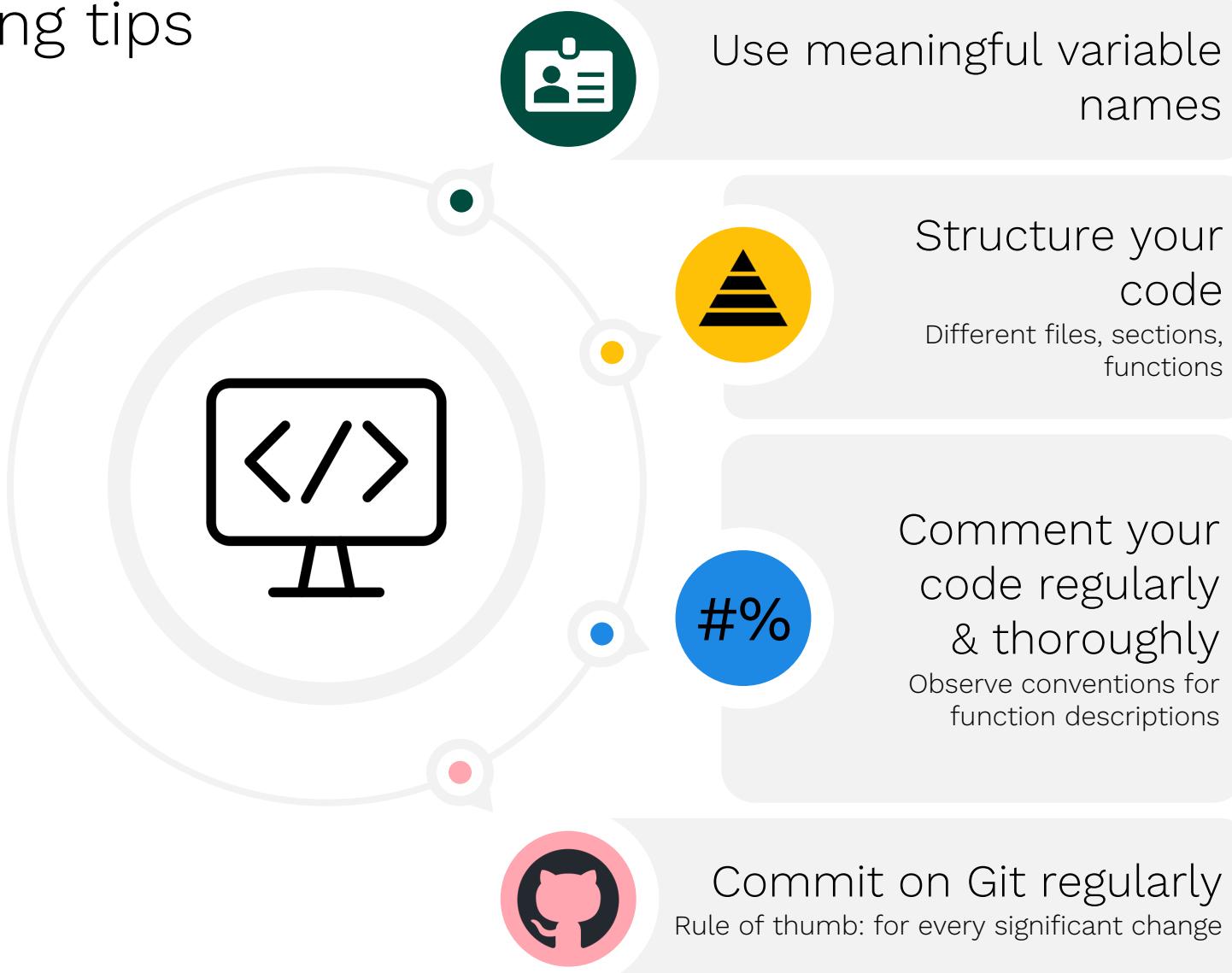
2

Code sharing can also be
private.

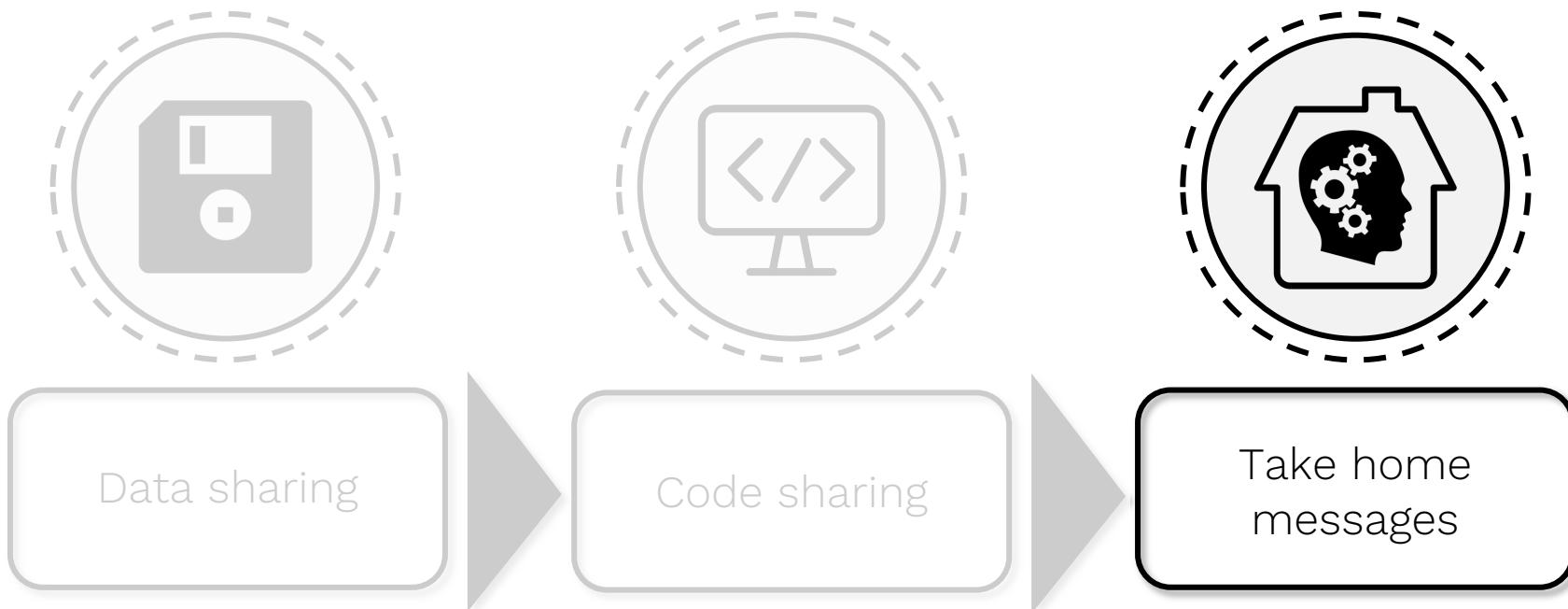
3

Code sharing

A few coding tips



Outline

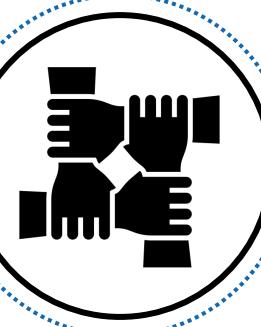


Take home messages

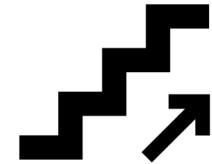
Why you should share your data & code



Contributes to making
your research more
[open](#) and [transparent](#)



Helps [you](#), your
[colleagues](#) and the
research community



Start [step by step](#)
&
set the example



Additional resources

- [FAIR principles](#)
- [Original paper introducing FAIR Data Principles](#)
- [The good research code handbook](#)
- [Learning Git branching](#)
- [GitHub Skills](#)
- [GitHub Desktop](#)