Docker for Research Software
About us

MLE @ Accelerate Science – University of Cambridge

I really like Physics...

...and machine learning...

...and race cars 🏎️
Setting the scene
Why I hate python

No clear packaging route.

No automatic environments.

So. Many. Tools.

conda

...but it’s the best we’ve got...
What do we want?

1. Reproducibility
2. Portability
3. Versioning / redundancy
4. Isolation
5. Satisfaction
Enter Docker
What is docker?
What is docker?

Dockerfile → image → container
Daily driving Docker...

...as a development environment
Docker as a development environment

It’s obviously hard and with no benefits

It’s obviously easy but there are no benefits

Docker is just hard

People don’t know
Why?

All the benefits of Docker.

Clean personal environment.

Develop a deeper understanding of Docker.
The Workflow
Examples
Course material

Running workshops can be painful

What is your favorite OS...?

...and why is it Linux?

This template enables us to easily develop consistent courses
Problems
Gotchas

Docker for desktop is not great...

...use orbstack instead.

1. Dynamic memory and compute allocation
2. Lighter and faster
Gotchas

Added complexity with GPUs...

...but not really

• Nvidia Container Toolkit
• Most cloud GPU instances come with this:

```
~$ docker run --gpus all my-project
```
Gotchas

Added complexity with remoting into cloud instances...

...but again...not really

• `ssh` into remote
• Pull image and run
• Use vscode dev containers
Gotchas

Different platforms...

...and this is actually sometimes a problem

• amd64 vs arm64
• Need to do long multi-platform builds