Big Data, organisation and analysis

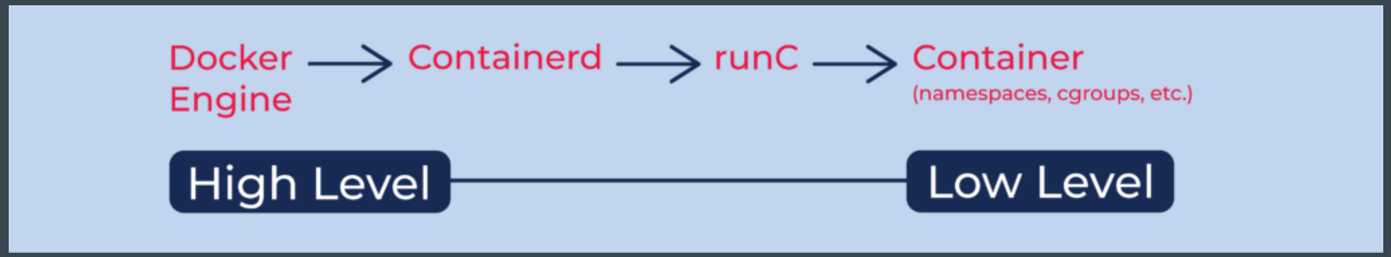
Containers

Steffen M. Noe, Spring 2023

Containerising applications

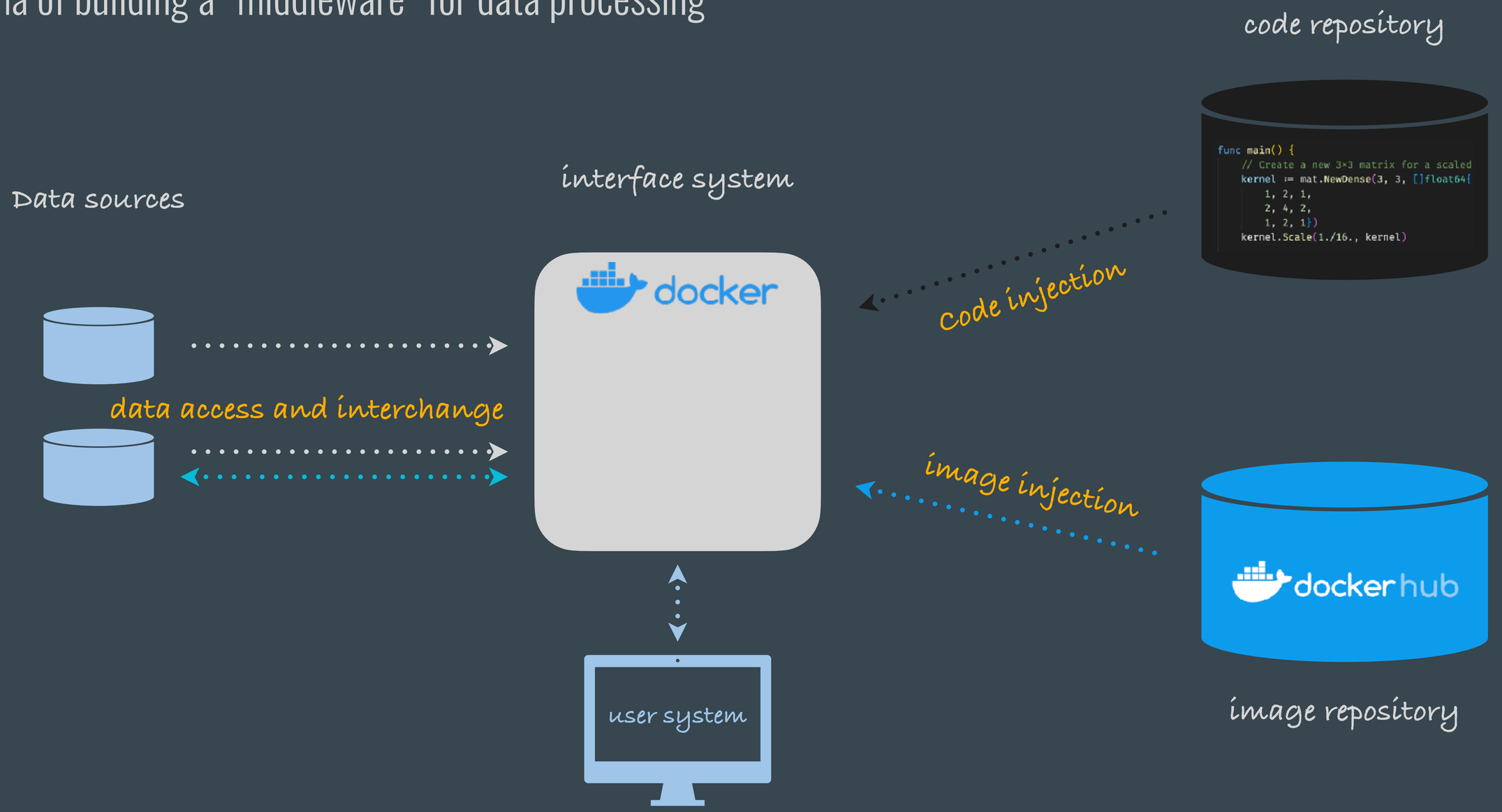
Container solutions

- Docker - available for Linux, MacOS, Windows (container engine, high level system)
- podman - available for Linux, MacOS, Windows (container engine, high level system)
- containerd - only Linux (system daemon, high level system)
- runC - command line interface to start containers (low level system)



Use of containers in Big Data applications

Schema of building a “middleware” for data processing



Virtual laboratories

- In the recent years, more and more so called virtual laboratories have been established
- some examples are:

EARTH CONSOLE

https://earthconsole.eu/virtual-labs/

esa European Space Agency

https://eo4society.esa.int/virtual-labs/

ICOS Integrated Carbon Observation System

https://www.icos-cp.eu/data-services/tools/jupyter-notebook

EUMETSAT

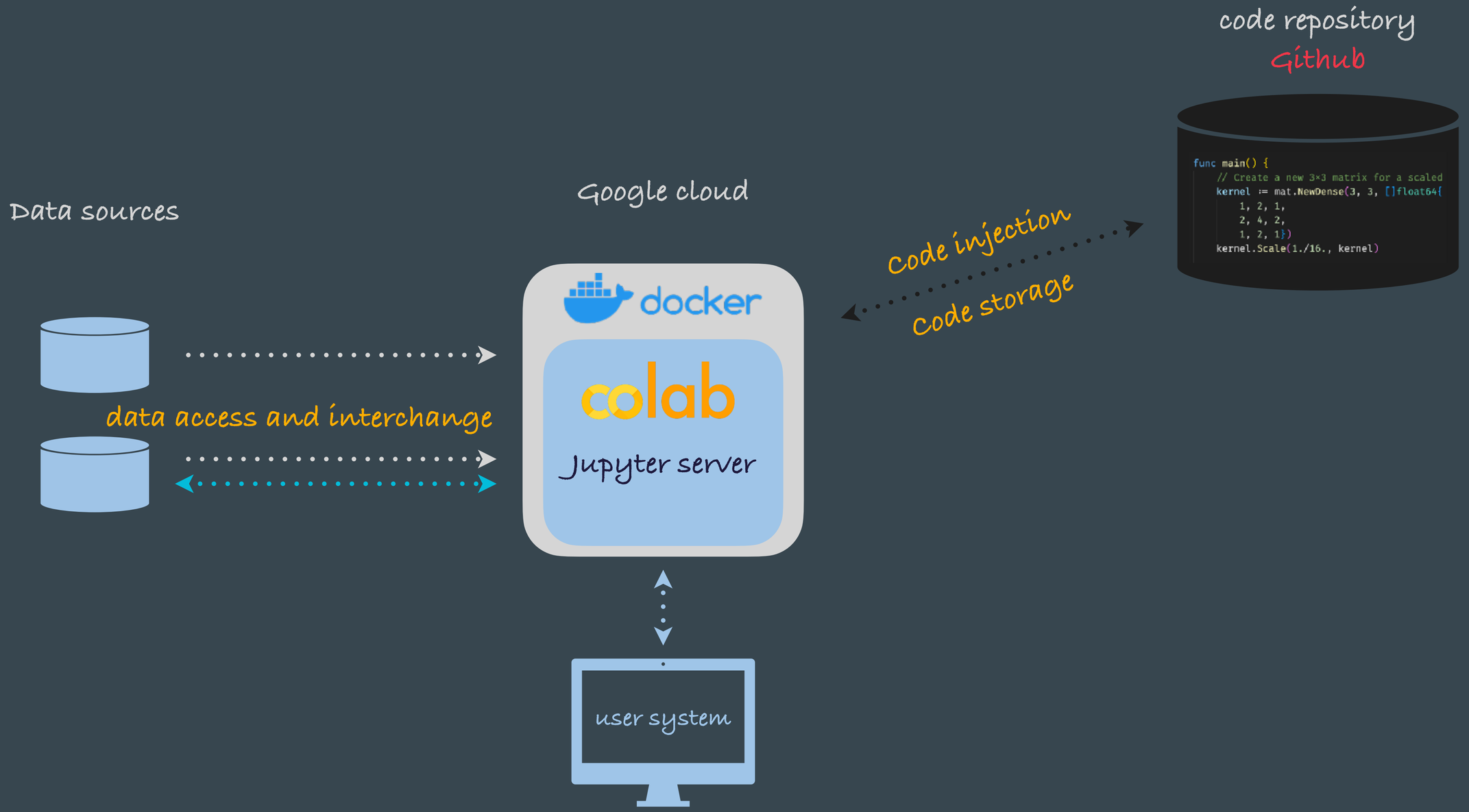
https://www.eumetsat.int/virtual-laboratory

Out of the box solutions

Google Colab as example

- Google offers with Colab a service that can be used to test some things.
- It is free but with limited resources
- Paid versions available
- Cloud based
- Jupyter notebook based
- Python as language

Using Colab as a middleware solution



Example

- See the Jupyter notebook in Github:
  - BDOA/Notebooks/example_middleware.ipynb