KU LEUVEN

Reproducibility Package:

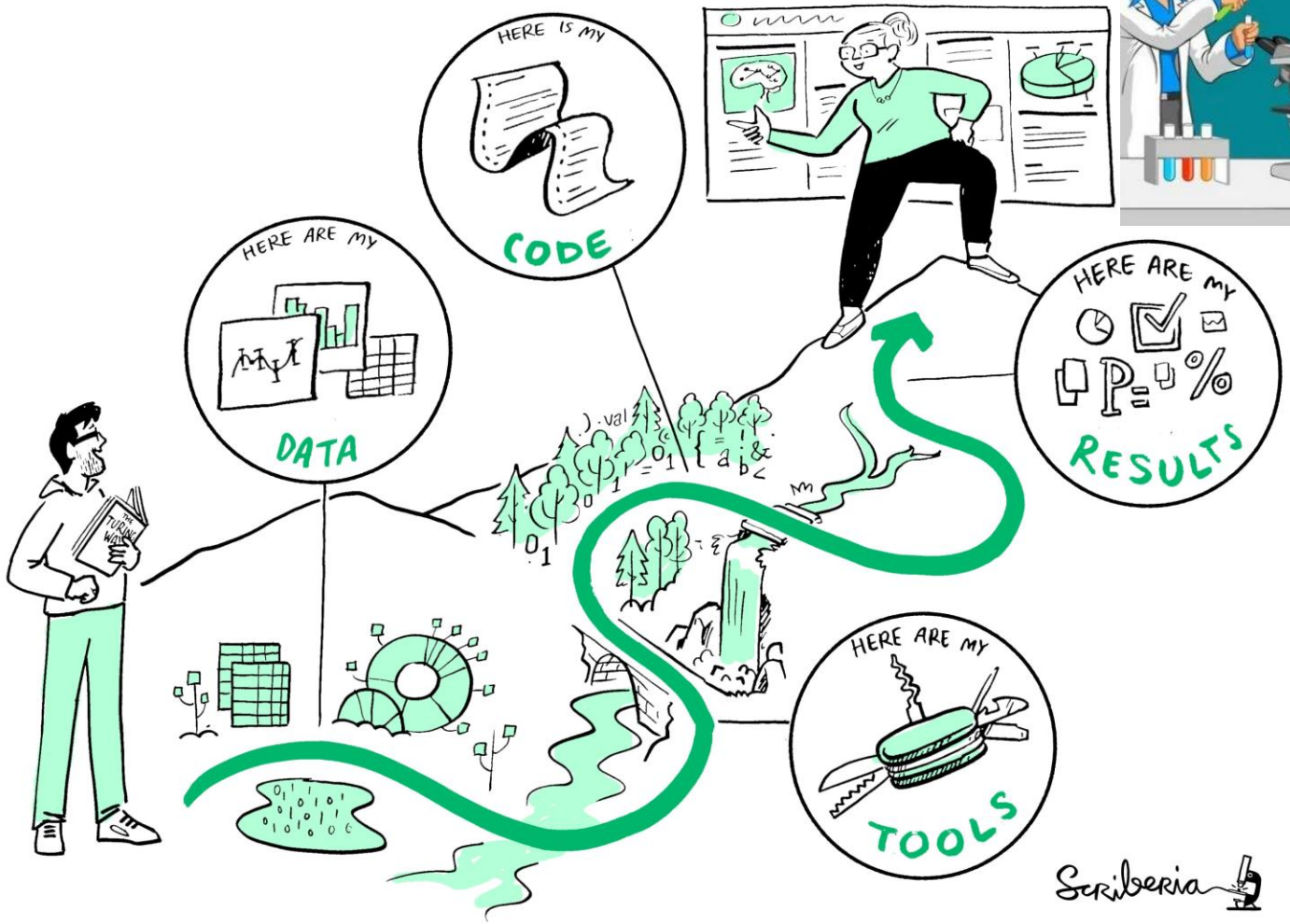
A practical approach to FAIRness in research data

Naeem Muhammad – Research Data Manager

Research Coordination Office, KU Leuven

Agenda

- Background
- Reproducibility Package
- Use cases
- Demo

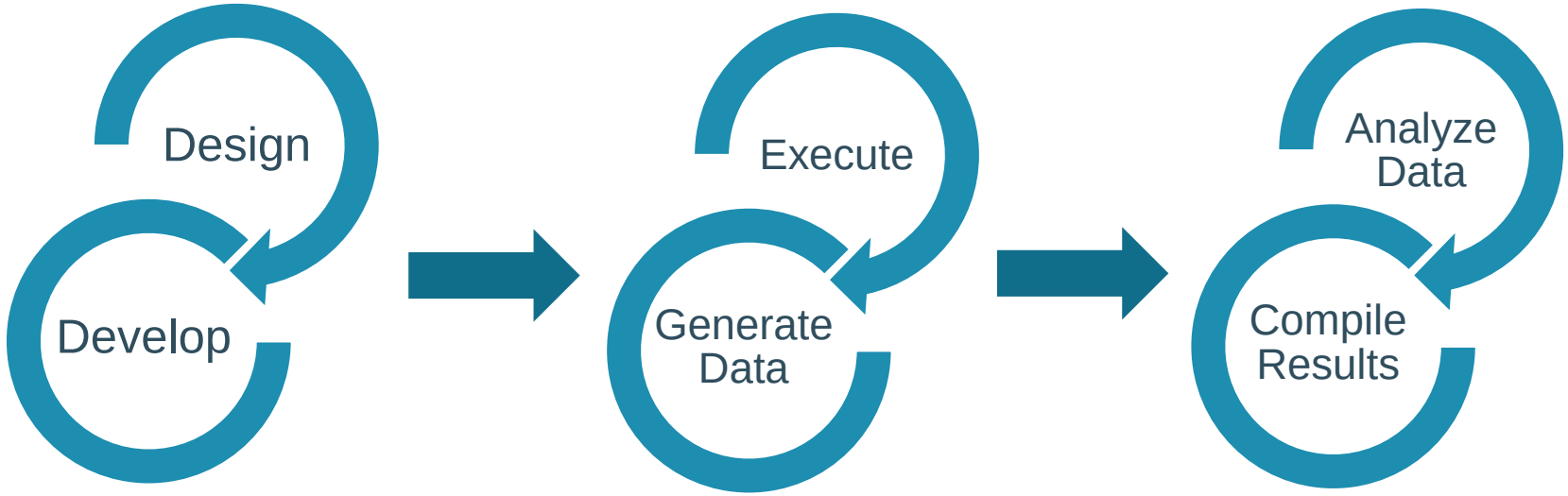


https://training.ni4os.eu/mod/url/view.php?id=774

https://www.scientific-editing.info/blog/importance-of-reproducibility-in-science/

Reproducibility in Research

- Difficult to reproduce every type of research
- Computational research a possibility
- Data analysis in research life cycle a possibility



Computational Reproducibility

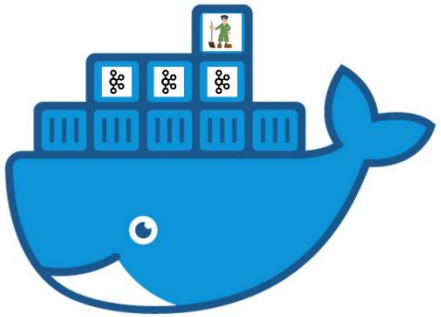
- Recreate exact runtime environment
- Exact randomness
- Fixed software/package versions
- Transportable solution
  - Small size
  - Standard technology

Computational Reproducibility

https://clipart.world/computer-clipart/cartoon-computer-clipart-transparent/

Reproducibility Package

- Docker based virtual machine
- Operating System
- Install required software, libraries etc
- Set Env Variables
- Fixed software/package versions
- Automated
- Minimum Human Interaction
- Readme file



Docker

GitLab

Reproducibility Package

Dockerfile

- Operating System
- Required Software
- Required Libraries

Research Work

- Simulation scripts
- Analysis tools/scripts
- Documentation

Automated Recreation

- Bash script
- User Manual

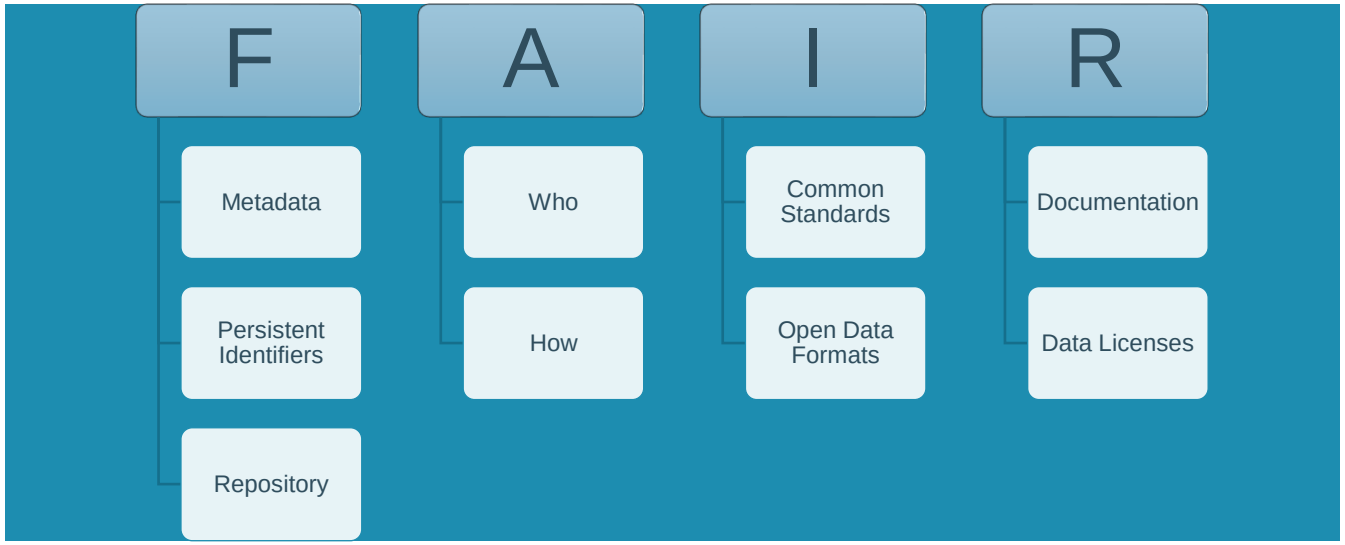
Size: ~ few kilobytes to few megabytes

Reproducibility Package

1. Execute bash script (initiated by user)
2. Create virtual machine
    - Install operating system
    - Install required software
    - Create user login
    - Set environment variables
3. Establish SSH connection
    - User lands in a directory where simulation & analysis scripts are available.
4. Execute simulation scripts
5. Execute analysis scripts
6. Output is generated

Seamless process

Reproducibility Package Advantage



Research Data FAIRness

Reproducibility Package – Sensitive Data

Data Not Available

Analysis Scripts

Data

Reproducibility Package Use Cases

- Paper reviewers
  - Manuscript
  - Executable analysis scripts

- Future reuse
  - Data archiving
    - research methodology, analysis scripts, output
  - New projects

- FAIR research
  - Data & code publishing

Reproducibility Package

- Zenodo: https://zenodo.org/record/7990608
- Gitlab: https://gitlab.com/nmrdm/reproduce/python_3_8

Questions