Reproducibility

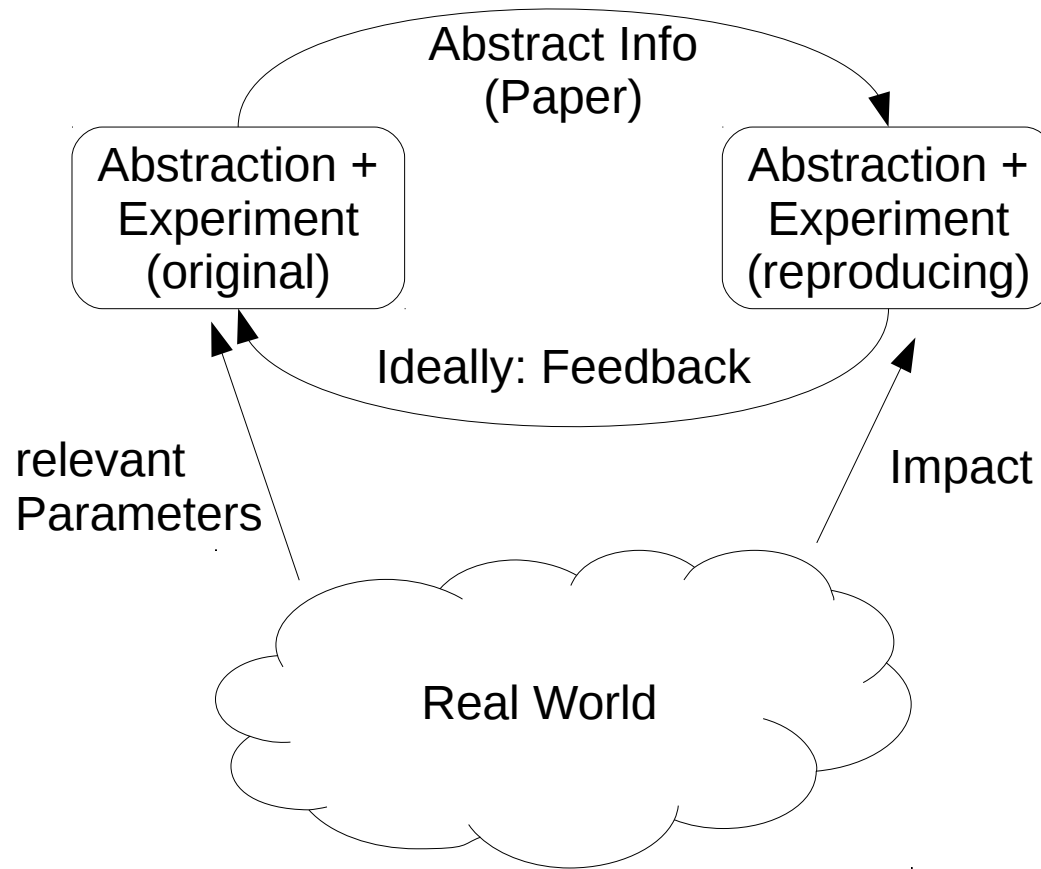Maja Curic, Christian Facchi, Stefan Neumeier, Claas Lorenz

(Some) Motivation

- operators/vendors: unexpected behaviour in real world setups → QA, Maintenance of products

- Research results (→ academic QA)

- System/Software Testing → standards, conformance, interoperability

Experimental Design

- proper (experiment) design, e.g. by well known rules, enables reproducibility

- clear definition of done: what is the degree of information provided to reproduce results?

- what non-determinism do I face?

  - User interaction

  - Physics (e.g. Air Interface)

  - Black Boxes (e.g. middleboxes on the internet)

The Experimental Cycle

Abstract Info
(Paper)

Abstraction +
Experiment
(original)

Abstraction +
Experiment
(reproducing)

Ideally: Feedback

relevant
Parameters

Impact

Real World

Some Takeaways

- not all experiments are reproducible

- Open data helps but is not sufficient → did they do their statistics correctly?

- in an ideal world reproducibility results (especially failures) are fed back to the original experimentors to improve their theory and/or experiments

- reviewers should embrace minimal standards on papers to improve result credibility and reproducibility