



Forschungsdesign, FAIRe Daten und strukturierte Analyse

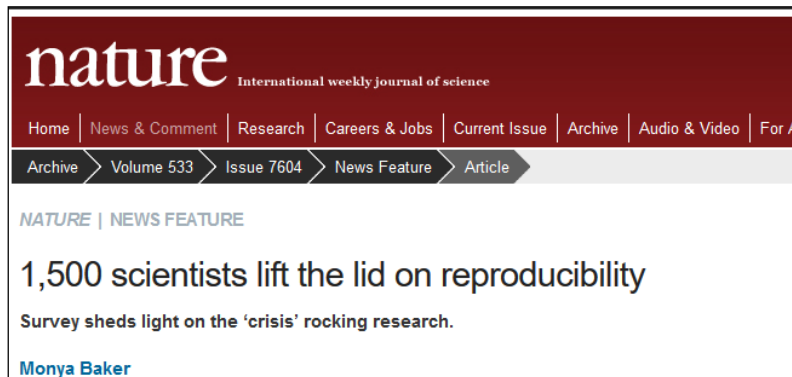
Janno Harjes
University of Bayreuth
Dept. of Mycology

43rd DWB-Workshop

28.06.2022

Crisis of reproducibility

- General pressure towards publication output:
 - measured and estimated using bibliometric values limited project durations
- Due to lack of time and resources:
 - repeatability of experiments and reproducibility of research results sometimes not guaranteed



nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For A

Archive > Volume 533 > Issue 7604 > News Feature > Article

NATURE | NEWS FEATURE

1,500 scientists lift the lid on reproducibility

Survey sheds light on the 'crisis' rocking research.

Monya Baker



MENU ▾ nature International journal of science

SPECIAL | 18 OCTOBER 2018

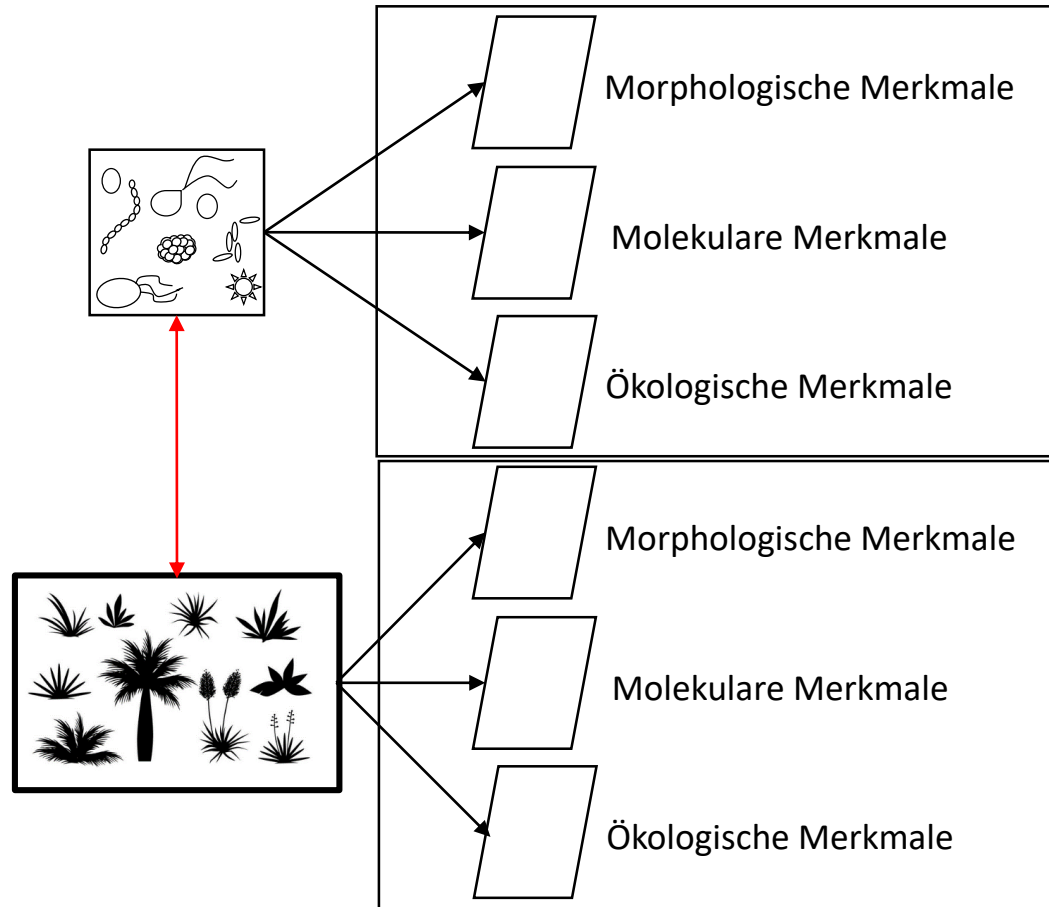
Challenges in irreproducible research

Terminology: Data versus metadata

- Highly inconsistent use of both terms in different scientific communities
 - Metadata being “data describing data” seems to be incomplete.
 - Better definitions:
- (Research) data
 - used for ‘correlations’ and ‘comparisons’ with other (research) data, also named ‘variables’ (explanatory/responsive)
- Parameters
 - used to control conditions to make measured values (→ data) reproducible, also named ‘invariables’
- (‘Metadata’) information
 - being used for making (research) data accessible: ‘locators’
 - information, describing (research) data and parameters

Descriptive data documentation

- Concatenation of workflow segments documentation
 - Data



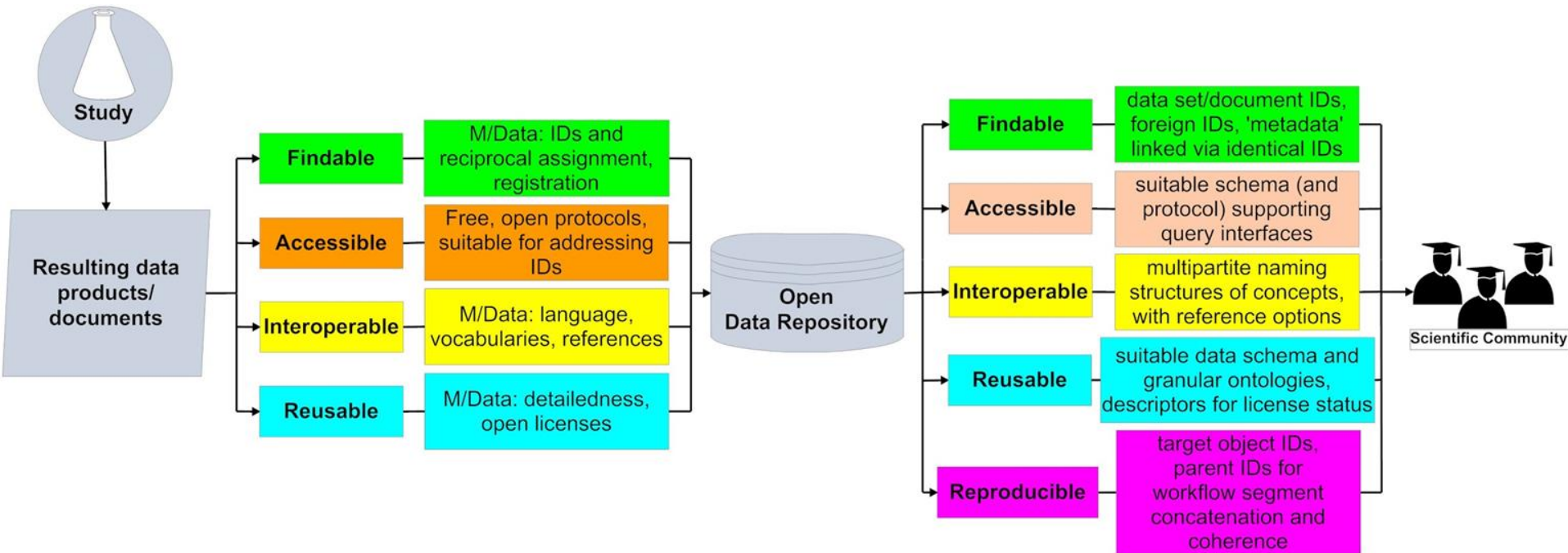
FAIR principles

- Findable
 - *To find for humans and computers*
 - IDs, registration, data description
- Accessible
 - *Authentication and authorisation*
 - Communication protocol, authentication and authorisation
- Interoperable
 - *Integration with data, in applications, workflows for processing*
 - Language for knowledge representation
- Reusable
 - *Sufficiently described for replication and combination in different settings*
 - Relevant attributes, community standards, provenience information, licenses

FAIR++ additions

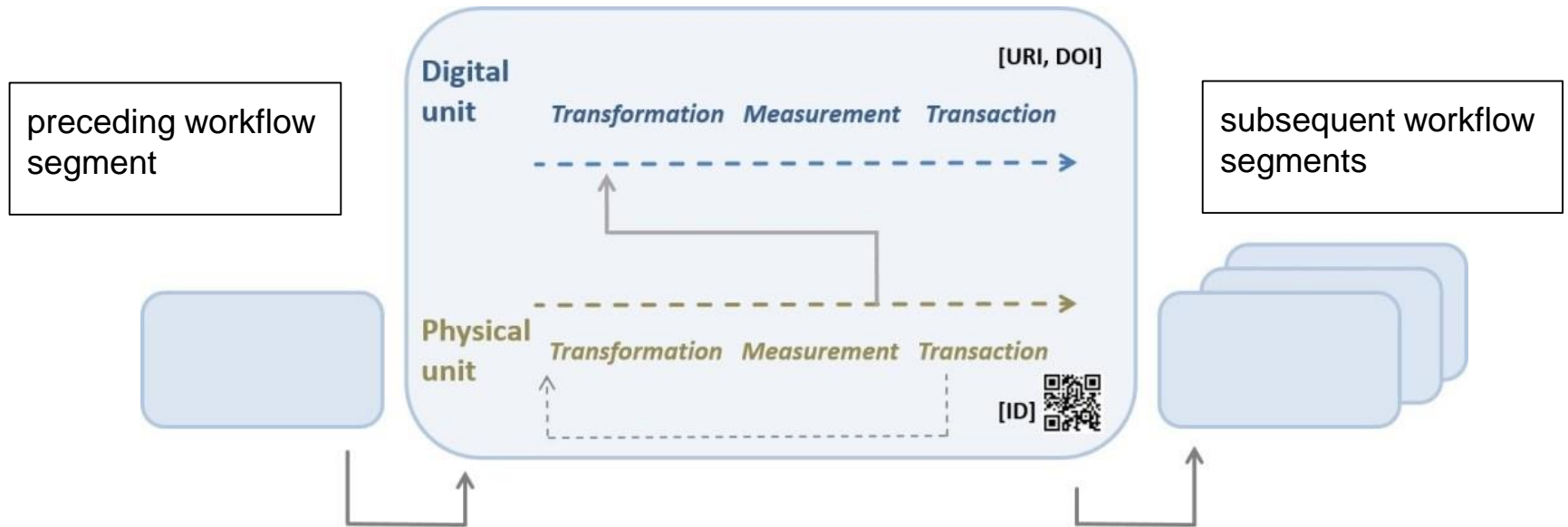
- Reproducible
 - *Enabling repeatability of experiment and reproducibility of results*
 - physical object ID linked with attributes reflecting operational designs along workflow

FAIR++ additions



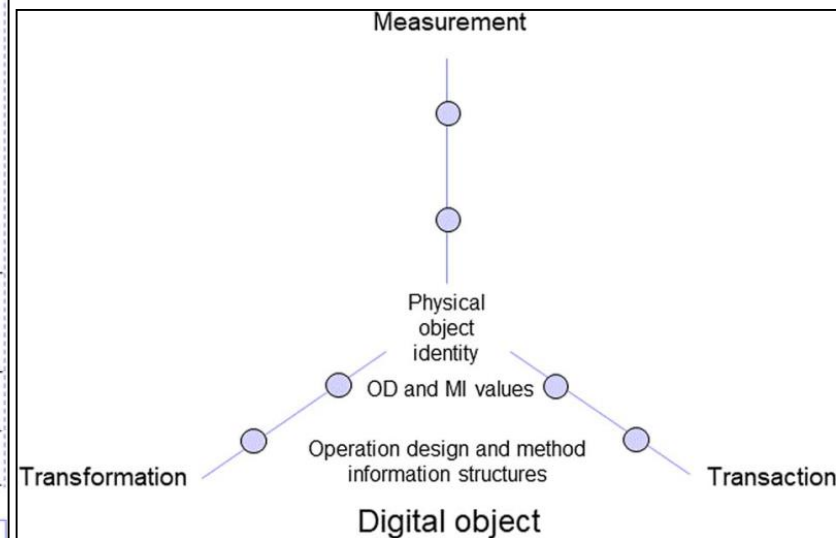
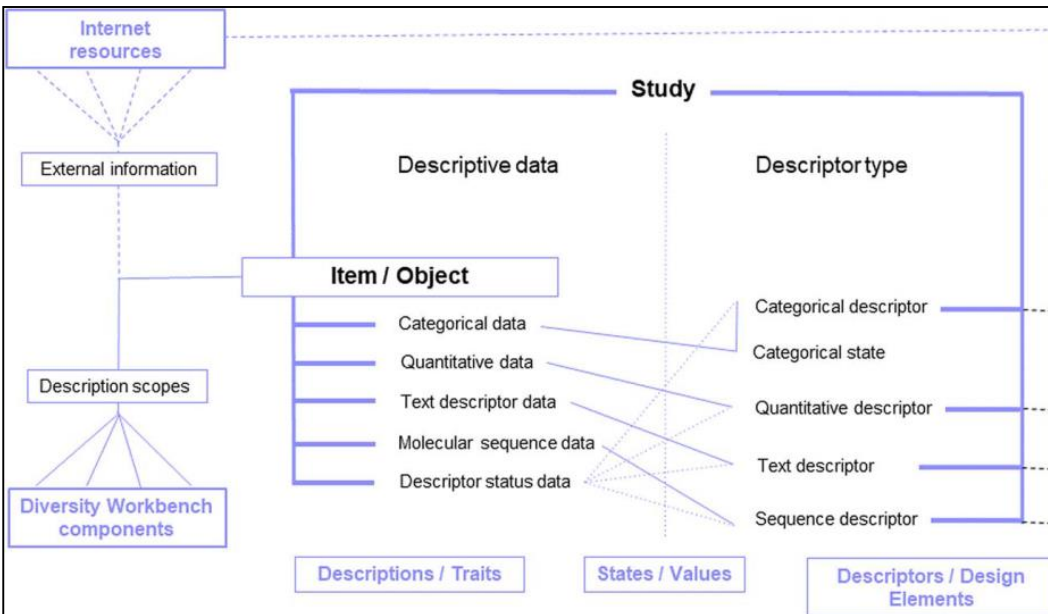
Wilkinson, M. D. et al. *Scientific Data* 3 (2016). doi: 10.1038/sdata.2016.18

Workflow segmentation



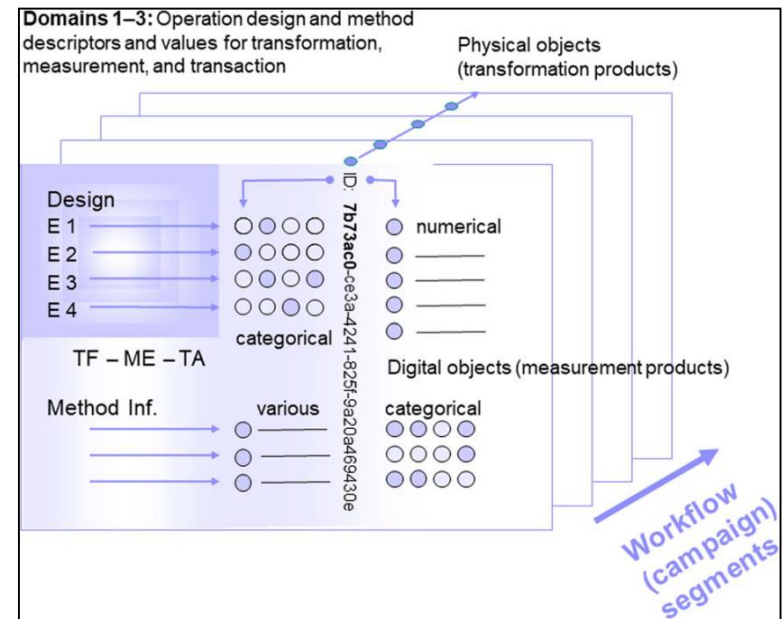
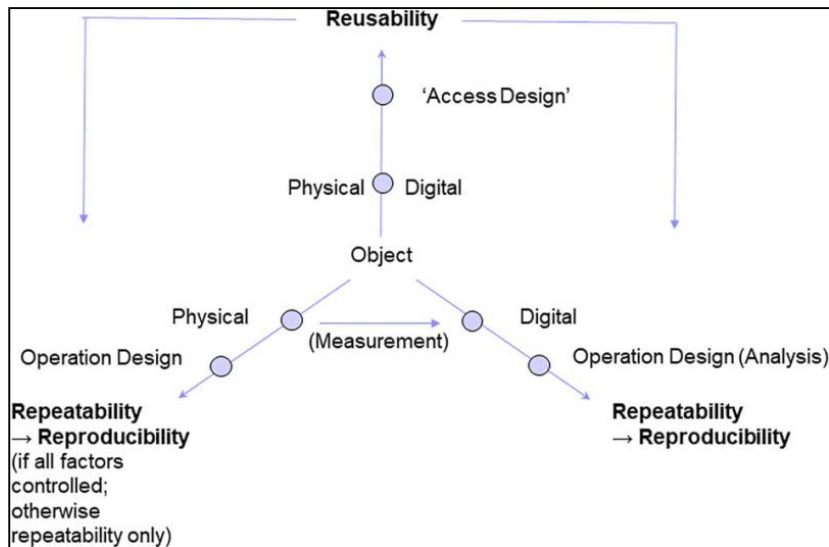
Reproducibility - requirements

- Procedural (experimental) designs and documentation
 - Object-orientated documentation of procedural designs and executions using corresponding descriptor types
 - Structured workflows with distinct segments of workflow steps and categorization into measurement, transformation and transaction



Reproducibility - requirements

- Concatenation of workflow segments documentation
 - Data quality documentation must be available for data analysis to aid meaningful interpretation
 - Documenting 'cross-influences' of procedural steps along entire workflow



Thank you for your kind attention

28.06.2022