

Publikationen und Forschungsdaten - Einführung

PubMan Days 2013

Michael Franke, Max Planck Digital Library

 Except where otherwise noted, this work is licensed under
<http://creativecommons.org/licenses/by/3.0/de/>



MAX PLANCK

digital library



Was sind Forschungsdaten?

Forschungsdaten sind jene Daten, auf denen wissenschaftliche Publikationen gründen.

„Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“ (Allianz AG „Forschungsdaten“)

„(...) researchers know within their projects what research data is“ (*Leibniz-Gemeinschaft*)

Verschiedene Sichten auf Forschungsdaten

Nach dem Charakter der Daten

Bilder (Scans, Fotos, Mikroskopdaten, Teleskopdaten, Satellitendaten)
Multimedia (Audiodaten, Videodaten, 3D-Daten, 4D-Daten)

} Qualitative Daten

Zahlenreihen (Umfragedaten, Experimentaldaten, Sensormessreihen, Gensequenzen, Geodaten)
Born-digitals (Simulationsdaten, Algorithmen, Websites)

} Quantitative Daten

Nach dem Format der Daten

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Messdaten (TXT, XLSX, ...)

Born-digitals (netCDF, grib, ...)

Nach der Quelle der Daten bzw. der Art der Forschung

Observational: data captured in real time that is usually unique and irreplaceable. For example, remote sensing data, survey data, field recordings, sample data.

Experimental: data captured from lab equipment that is often reproducible. For example, gene sequences, chromatograms, magnetic field data.

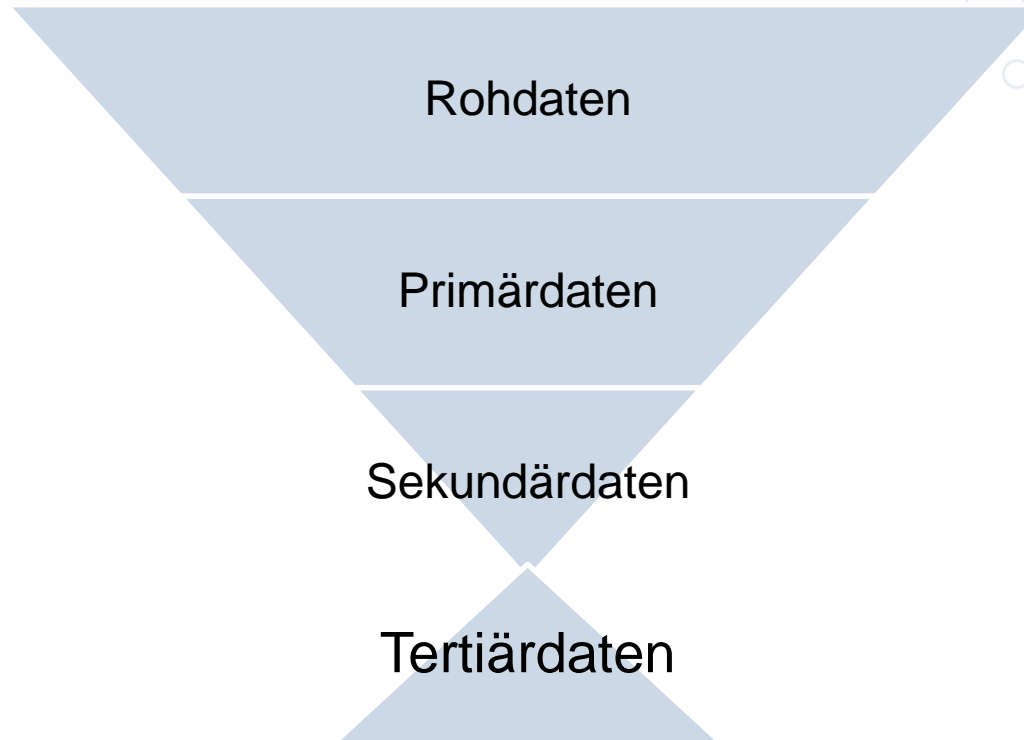
Models or simulation: data generated from test models where model and metadata may be more important than output data from the model. For example, climate models, economic models.

Derived or compiled: resulting from processing or combining 'raw' data. For example, text and data mining, compiled databases, 3D models.

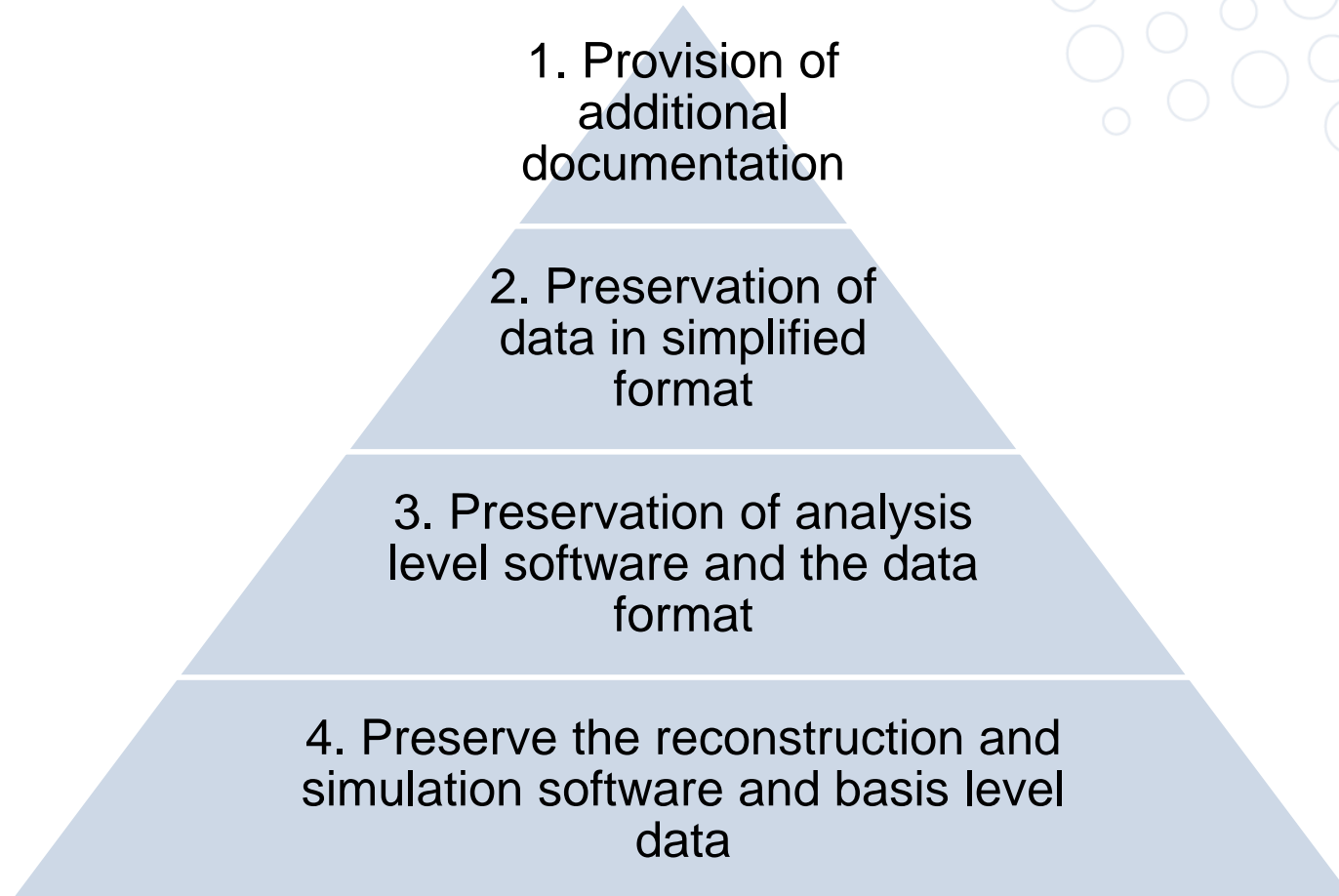
Reference or canonical: a static or organic conglomeration or collection of datasets, probably published and curated. For example, gene sequence databanks, collection of letters.

(University of Bristol)

Nach dem Grad der Aggregation/Verdichtung



DPHEP Tiers



Motivationen für Langzeitarchivierung

Gute wissenschaftliche Praxis

Motivation

- Qualitätssicherung, Reproduzierbarkeit

Interesse

- Förderer, Institutionen, Öffentlichkeit,
(Wissenschaftler)

Schwerpunkt

- Vollständigkeit der Daten und Methoden, sichere
Aufbewahrung

Nachnutzung der Daten (Data Sharing)

Motivation

- Effizienz, Erkenntnisgewinn

Interesse

- Förderer, Institutionen, Öffentlichkeit, Wissenschaftliche Community

Schwerpunkt

- Offene, auffindbare, lesbare Daten, Lizenzen, lange Zeiträume

Zitierbarkeit von Daten

Motivation

- Illustration, Nachvollziehbarkeit, Zueignung, Reputation

Interesse

- Wissenschaftler, Öffentlichkeit, Wissenschaftliche Community, Verlage

Schwerpunkt

- Offene Daten, atomare Adressierbarkeit, persistente Identifikation

Publikationen und Forschungsdaten

Forschungsdaten in wissenschaftlichen Publikationen

Als „supplementary material“

Als externe Referenz

Zur Veranschaulichung der Methoden und Ergebnisse („Enhanced Publication“)

Für die Überprüfbarkeit der Ergebnisse?

Datenpublikation

In Fachrepositorien

In peer-reviewten Data Journals

Vielen Dank für die Aufmerksamkeit!