RESEARCH DATA DEFINITIONS

- Material generated or collected during the course of conducting research.  
- Factual records used as primary sources for scientific research, commonly accepted in the scientific community as necessary to validate research findings.  
- Information collected, observed, or created, for purposes of analysis to produce original research results.  
- Any information in binary digital form derived from the research process.

RESEARCH DATA LIFECYCLE

1. Creating / Re-using: planning data collection, locating existing data sources; producing, collecting or documenting data.
2. Processing / Analyzing: validating, cleaning, transforming data; creating metadata; using, creating analysis tools; interpreting the data.
3. Preserving / Publishing: reviewing the data; getting data into a format suitable for preservation; depositing data and metadata in archive / repository; promoting data re-use.

RESEARCH DATA TYPES

- **Observational Data**: data captured in-situ, can’t be recaptured, recreated or replaced. Examples: Sensor readings, sensory (human) observations, survey results, interview notes, transcripts.
- **Experimental Data**: data collected under controlled conditions, in situ or laboratory-based, should be reproducible, but can be expensive. Examples: gene sequences, chromatograms, spectroscopy, microscopy.
- **Simulation Data**: result from using a model to study the behaviour and performance of an actual or theoretical system, models and metadata, where the input can be more important than output data. Examples: climate models, economic models, biogeochemical models.
- **Derived/Compiled Data**: reproducible, but can be very expensive. Examples: derived variables, compiled database, 3D models.
- **Reference or canonical Data**: static or organic collection [peer-reviewed] datasets, most probably published and/or curated. Examples: gene sequence databanks, chemical structures, census data, spatial data portals.

**Raw Data**
Raw data refer to data that have not been changed since acquisition, eg. a real-time GPS-encoded navigation file, and the initial time-series file of temperature values from a heat probe.

**Processed Data/Active Data**
Editing, cleaning or modifying the raw data results in processed data, eg. raw multibeam data files can be processed to remove outliers and to correct sound velocity errors.

Credits and sources

[1] https://www.ed.ac.uk/information-services/research-support/research-data-service  

Contact and info  
researchdata.epfl.ch  
researchdata@epfl.ch