

Data Management Concept Note

Date: 10 May 2020

Authors: Data stewards team, datastewards@tudelft.nl, TU Delft
(see [CRediT author statement](#) for detailed contributions)

DOI: <https://doi.org/10.5281/zenodo.3819204>

Who should read this

This concept note is suitable **for any individual, group or organization** whose work involves operations with large or small volume of data, software or other types of information. It introduces clear and instructive steps to check key data management issues in every stage of data handling, such as collecting, processing, storing, analysing or disseminating.

Motivation

[TU Delft Data Stewards team](#) has been asked by different teams and organizations for expertise in providing recommendations on data management. In the spirit of open science and FAIR data, the team would like to share this essential guide for general data management. This document outlines the rationale, the main points to be considered for effective data management, and suggests sensible approaches for organizational implementations.

Rationale

Data management is the practice of collecting, storing, protecting and responsible sharing of data with others. There has been increasing demand on data management during the COVID-19 crisis, with respect to processing sensitive data, using it in an effective and efficient way, and sharing it in a multi-stakeholder environment. **Good data management is not only crucial for timely response to the COVID-19 crisis, but also a general necessary practice for responsible and trustworthy organizations and policy decisions.** Organizations who collect and process valuable and sensitive information are also accountable for responsible management of the data assets. There are multiple risks associated with handling valuable and sensitive information, for example, the data falls into the wrong hands, is misused by decision-makers, etc.

From data management perspective, all organizations need to achieve the following:

- Avoiding access to the data by people with bad intentions
- Providing reliable information to make decisions based on facts, instead of political / organizational motivations
- Being a reliable organization to compile, manage and provide the information to end-users

A solid data management strategy is thus essential not only to increase the organisational effectiveness and efficiency when it comes to making data-driven decisions, but also helps minimise the risks of data mis-handling, ownership dispute or other potential issues. These objectives could be accomplished by developing the following competencies within the organizations:

- Clear policies and procedures on data use
- Faster decision making based on well structured data
- Transparency when communicating decisions to the target group
- Being prepared for an increase in the data flow as the emergency worsens
- Being prepared for possible infrastructure failure and securing access to data

Main Points for Good Data Management

Data acquisition

- Check the type, source of the data and how to gather/collect it
 - Data types (to help define sensitivity of collected and processed data)
 - Data format (to help define the tools and infrastructure needed)
 - Data size (to help define storage and infrastructure needs)
- Check the ownership of the collected and processed data
 - Check with the data source about the ownership, attribution and use conditions (e.g. licence)
 - Check the need to make a data processing agreement with the data source on the ownership / access control
 - Are there (own) institutional policies that apply to the data?
 - Can the data be shared with other parties (if so, under which circumstances?)
- Confidentiality of the data (if applicable):
 - Register crucial information regarding data access / sharing
 - Ensure security of confidential data (personal data, commercially sensitive data, or data that would harm society when publicly disclosed)
 - Ensure compliance with General Data Protection Regulation / Algemene verordening gegevensbescherming when processing personal data
 - Ensure there are procedures in place to avoid data breaches, with the support of a privacy advisor/data protection officer

Data collection

- Establish a workflow for data collection
 - How will the data be collected?
 - Who has access to which data in short / long term?
 - What resources are needed for data analysis?
 - How will the data be exchanged / transferred among relevant stakeholders?
- Storage arrangement
 - Check available storage capacity and backup strategy

- Ensure the secure storage for confidential data
- Access control
 - Monitor the access authorization
 - Make a coordination plan for data access/exchange

Data storing / backup

- Create a clear folder structure and consistent file naming convention
- Make a backup strategy where data is stored at least two different physical locations and preferably automatically backed up
- Access control to confidential data
- Apply encryption at disk or folder level if needed
- Create a consistent and standard versioning of the data files
- Determine the minimal documentation of the data that is required to find it, understand it and use it

Data sharing

- Create proper data sharing procedures
 - Consider agreements established in the Data acquisition phase, and evaluate/assess data sharing with other parties
 - Be aware of the permission and consequence of sharing confidential data
- Copyright / Licensing
 - How should others use the data
 - Who should be attributed for creating/gathering the data

Organizational Implications

In addition to the above mentioned actions, there are also a few things to consider to make data management a standard practice in daily operations.

Organizational policy on data management

Adding data management into the organizational mandate and clarifying relevant principles help organizations to operate and grow in a professional and sustainable way. A policy for data management establishes the roles and responsibilities for data management across the multiple stakeholders across the organisation. It will ensure that procedures are integrated throughout all operations where data is involved. As a result, good data management will become a standard practice in daily workflows.

Key working principles

To ensure that data management is embedded within the organisation and the policy can be successfully implemented, it is crucial that all staff members of the organization adhere to the same key data management principles. Development of a checklist which could be easily

used and applied by all staff members would be a good starting point. Such a checklist should take into account the elements of good data management described above.

Training and awareness raising

A successful implementation of a data management policy and a strategy depends on their acceptance and adoption by all staff members. Training and awareness sessions facilitate the understanding of the different roles around data and provides a starting point for improving the data management practices.

CRedit Author Statement

Writing-original draft, Writing-review & editing	Esther Plomp, ORCID 0000-0003-3625-1357 Heather Andrews, ORCID 0000-0002-6637-2830 Marta Teperek, ORCID 0000-0001-8520-5598 Paula Martinez Lavanchy, ORCID 0000-0003-1448-0917 Santosh Ilamparuthi, ORCID 0000-0001-5101-7898 Yasemin Turkyilmaz-van der Velden, ORCID 0000-0003-2562-0452 Yan Wang, ORCID 0000-0002-6317-7546
Visualization	Santosh Ilamparuthi
Conceptualization	Kenny Meesters (Tilburg University), ORCID 0000-0002-8324-956X Yan Wang
Project lead	Yan Wang