**National Institute of Standards and Technology (nist.gov): NIST: National Institute of Standards and Technology Data Management Plan**

**Summary of activities for data generation**

summary of activities that generate data

> *Guidance*:
>
> Data sharing in this policy refers to final research data. These data are the recorded factual materials commonly accepted in the scientific community as necessary to document and support research findings. This policy applies to new data collection as well as to data obtained through transforming or linking existing datasets. For most studies, an electronic file will constitute the final research data. This dataset will include both raw data and derived variables, which will be fully described in accompanying documentation section.
>
> Consider these questions:
>
> - How will you capture or create the data?
> - Are you pulling from existing data sources?

**Data types and classification**

a summary of the data types generated by the identified activities. Data should be categorized, at a minimum, according to the data categories presented in the NIST Data Taxonomy and Actions/Consequences for Data Categories, provided in Appendix A of this Order, as applicable.

> *Guidance*:
>
> Describe how your data will be "documented." Think about what information is needed for the data to be read and interpreted in the future. What would someone else need to be able to use these files? The documentation should include a summary of the purpose of the data collection, methodology and procedures used to collect the data, timing of the data collection, as well as details of the data codes, definition of variables, variable field locations, and frequencies. The data documentation should be a comprehensive and stand alone document that includes all the information necessary to replicate the analysis performed by the original research team.
>
> Consider these questions:
>
> - What data will you collect or create in the research?
> - What data types will you be creating or capturing and what data will be shared?
> - What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?
> - What contextual details (metadata) are needed to make the data you capture or collect meaningful?
> - How will you create or capture these details?

**Preservation**

a plan for storage and maintenance of the data generated by the identified activities, in both the short-term and long-term (if relevant). Data should be preserved, at a minimum, according to the preservation consequence levels defined in the NIST Data Taxonomy and Actions/Consequences for Data Categories, provided in Appendix A of this Order, as applicable, and in accordance with applicable records retention requirements.

> *Guidance*:
>
> - Is a short-term storage plan provided?
> - Is a plan provided for transition from data acquisition to short-term storage, to long-term storage and preservation of the data (if relevant)?
> - Is the short-term plan consistent with NIST P 5700.00 and O 5701.00 Managing Public Access to Results of Federally Funded Research?
> - Is the long-term plan consistent NIST P 5700.00 and O 5701.00 Managing Public Access to Results of Federally Funded Research?

[NIST Policy 5700.00](#)
[NIST Order 5701.00](#)

## Review, Discoverability, and Access

a plan describing whether and how the data generated by the identified activities will be reviewed and made available to the public and how the metadata describing it will be entered into the NIST Enterprise Data Inventory (EDI). The plan should describe any known access restrictions for the data and/or metadata, if appropriate. Data should be made discoverable, at a minimum, according to the discoverability consequence levels defined in the [NIST Data Taxonomy and Actions/Consequences for Data Categories, provided in Appendix A](#), as applicable.

*Guidance*:

Describe whether and how the data generated by the identified activities will be reviewed and made available to the public. The plan should describe any known access restrictions for the data and/or metadata, if appropriate.  Note that NIST does not currently require deposit in any specific repositories. Data is expected to be made publicly available unless such a release is not permitted (e.g., data contains personally identifiable information (PII) or business identifiable information (BII)).  If data will not be made publicly available, justification must be provided.

- How will data quality be assured?
- Will the data be made publicly available? (If not, why not?)
- Will the data be published? Where?
- If it is not published, how will it be made discoverable and accessible to the public?
- Who will review the data to ensure its quality before it is released to the public?
- Will the data be provided in a non-proprietary, machine-readable, and machine-actionable format?
- Will supporting documents be provided (e.g., a data dictionary) to allow for re-use of the data?