

Plan Overview

A Data Management Plan created using HKUL DMPTool

Title: Declining Majority of Online Adults Say the Internet Has Been Good for Society

Creator: Daijiao Zhang

Affiliation: The University of Hong Kong

Data Manager: Zhang Daijiao

Template: HKU Template

Last modified: 02-27-2026

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Declining Majority of Online Adults Say the Internet Has Been Good for Society

Data Collection

What data will you collect or create?

Data type: quantitative questionnaire survey data, sample statistics; data source: online/offline questionnaire survey of adult groups in the United States (Pew core survey)

Data size: effective sample N=1785 (Internet users), including multi-dimensional variables such as Internet use, social media behavior, social attitude, etc. The data format is structured data (such as CSV/SPSS, which fits the actual format of the data set)

How will the data be collected or created?

Collection method: standardized questionnaire survey (including online/offline channels), the questionnaire is the Topline Questionnaire officially designed by Pew to ensure data standardization;

Sorting method: de-weight the original questionnaire data, clean the abnormal values, standardize the coding (such as the unified coding of the Rickett scale for attitude problems), convert the original data into a structured format (such as SPSS/CSV), and establish a data dictionary (labeling variable meaning, coding rules, and missing value processing methods).

Documentation and Metadata

What documentation and metadata will accompany the data?

Core documentation:

- README.md: Overview, provenance, structure, versioning, and citation.
- Data dictionary: Variable definitions, coding rules, missing value handling, and scales.
- Original questionnaire PDF, Pew methodology report, and reproducible R/Python analysis scripts.

Metadata:

- Aligned with DataCite/Dublin Core standards: creator (Pew), dates (2018/01/03–10), access (CC BY-NC 4.0), formats (SPSS .sav, CSV, PDF), and DOI.
- Hosted in HKU Research Data Repository for discoverability.

Ethics and Legal Compliance

How will you manage any ethical issues?

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

Storage and Backup

How will the data be stored and backed up during the research? i. e. until stored in the final location (e.g. on your password protected laptop)?

Storage: Primary storage on my password-protected laptop.

Backup: Daily incremental backups to cloud storage.

How will you manage access and security?

Security: Because it is the social survey data of the American population, there is no sensitive personal identity information (the Pew data set has been anonymized), only the data access is open to the project research team, and the account password protection is set.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Raw dataset, cleaned/processed data, codebook, original questionnaire, and reproducible analysis scripts.

What is the long-term preservation plan for the dataset?

Preserved for ≥ 10 years via HKU Library's research data archive and Pew's permanent data repository. Converted to non-proprietary formats (CSV, PDF) for sustained accessibility; backed up per HKU's long-term storage protocols.

Data Sharing

How will you share the data?

Are any restrictions on data sharing? If yes, Why?

Responsibilities and Resources

Who will be responsible for data management?

What resources will you require to deliver your plan?
