

Plan Overview

A Data Management Plan created using DMPTool

DMP ID: <https://doi.org/10.48321/D18340>

Title: DMSP for "Common Rubric development methodology."

Creator: James Lipuma - **ORCID:** [0000-0002-9778-3843](https://orcid.org/0000-0002-9778-3843)

Affiliation: New Jersey Institute of Technology

Principal Investigator: James Lipuma

Contributor: Jeremy Reich, Cristo Leon, John Wolf, Maurie Cohen, David Crowell

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 21-588

Grant: <https://new.nsf.gov/funding/opportunities/ehr-core-research-ecrcore>

Template: Digital Curation Centre

Project abstract:

This document analysis rubric project aims to identify common areas in rubrics used for oral communication in the Department of Humanities and Social Sciences (HSS) courses and to compare these to nationally recognized best practices that informed the development of the New Jersey Institute of Technology (NJIT) General Education Requirement (GER) outcomes indicator. In addition, the project will gather and collate departmental resources and exemplars and hold a discussion group to analyze the collected materials and document analysis results.

This project aims to create outcome criteria to guide future outcomes assessments for the NJIT GER Programs, a crucial step toward attaining Middle States accreditation. This project is significant for

NJIT's ongoing accreditation work and will be helpful for other higher education institutions.

Analyzing the oral communication protocol is essential to the NJIT GER Program. HSS professors have volunteered to test out some of the elements of the system, providing valuable feedback in focus groups. The expert knowledge of the faculty is an essential data source, and this project aims to generate an understanding of processes and procedures, along with working artifacts and related resources. This will give everyone clear milestones, starting points, and guidance toward effective best practices for GER or any program assessment.

The results of this project will be shared with the CSLA Committee, OIE, and other faculty senate committees working on GER and Middle States accreditation. In addition, the resulting reports and aggregated data will be published to assist others at NJIT and in any educational group to evaluate and develop rubrics. Overall, this project is an essential step in NJIT's ongoing efforts to improve its educational programs and achieve accreditation.

Start date: 05-01-2023

End date: 04-30-2031

Last modified: 06-27-2023

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

DMSP for "Common Rubric development methodology."

Data Collection

What data will you collect or create?

The data collected for this project include oral communication rubrics used in Humanities and Social Sciences courses at the New Jersey Institute of Technology (NJIT). The Rubrics will be collected in informal and formal formats for various types of oral communication, such as class discussions and video recordings. The volume of data collected will depend on the number of rubrics the HSS faculty provides.

The rubrics will be collected through email or uploaded to a Google Drive provided by the coordinator of Assessment and Evaluation for HSS, Dr. Jim Lipuma. The Rubrics will be analyzed and compared to nationally recognized best practices that informed the development of the NJIT General Education Requirement (GER) outcomes indicator. The methodology developed to identify and write outcomes and the supporting materials will also be documented, ensuring anonymity by removing identifiers.

We have decided to collect data electronically for our project using Microsoft Word. The rubrics used in the project will be stored in Google Drive, allowing interested stakeholders to access and share the data in the long term. Additionally, we can reuse any existing data, such as rubrics used in prior assessments and other data collected by the Office of Institutional Effectiveness (OIE) for NJIT's ongoing accreditation work.

An MP4 video recording and (.TXT) metatext file will be created during the annual focus group meeting. The raw data of the video will be kept confidential because it can not be anonymized. The text transcript will anonymize by removing identifiers and uploaded to Google Drive.

The project aims to create a Google survey to collect the rubrics after the first academic year cycle.

How will the data be collected or created?

The project will create a small amount of data in documents (MS WORD). In addition, focus group WebEx videos as MP4s will also be made, and (.TXT) transcripts will be generated. Eventually, a Google form will be produced to gather rubrics and accompanying contextual items like use case and origin. That instrument and the subsequent spreadsheets will be data generated by the project.

Folder and file structure: To ensure efficient and effective data management, a hierarchical structure containing the following will be used

- The main folder for the project
- 01 Codebook
- 02 Data collected (e.g., rubrics, video recordings, text transcripts, survey responses)

- o 02.1 Rubrics

- § Sub-folders for each academic year cycle: Fall and Spring

- o 02.2 Focus Group: Video Recordings and Text Transcripts

- o 02.3 Survey Results

- § Sub-folders for any additional contextual information (e.g., use case, origin)

- Naming convention: Files named with clear and descriptive titles and labeled with metadata to indicate important information such as the: Last two digits for the year_Academic Semester_ Instructor code_Contextual information_Item name if needed, and or versioning.

Versioning: Version control is an essential aspect of data management to keep track of changes made to files over time.

- One approach is to adopt a naming convention that includes version numbers or dates in the file name.

- Another method is to use version control tracking through a data collection instrument that will enable the comparison of different versions of files.

Quality assurance processes: To ensure the accuracy and validity of the data, quality assurance processes may include checking for completeness and accuracy of the data during collection, performing data cleaning and validation, and conducting quality checks before analysis.

Guidelines for the quality check will be:

- Accessibility
- Correctness
- Validity check
- Correction notification
- Segregated

Documentation and Metadata

What documentation and metadata will accompany the data?

The videos will be created with standard metadata tags. The documents submitted by faculty will be an uncontrolled sample. After data reduction and aggregation, the documents will be labeled and tagged with

the study protocol information, date time, and NJIT standard data.

- Description of the data: A detailed description, including what it represents, how it was collected, and any relevant contextual information.
- Data format: Information on the file format, including the software used to create and view the data.
- File naming convention: A consistent naming convention will help users identify and locate specific files.
- Data structure: Information on the data structure, including any data dictionaries or codebooks used.
- Variable descriptions: Detailed descriptions of the variables included in the data.
- Methodology: Information on the methods used to collect and analyze the data.
- Quality assurance processes: Information on the quality assurance processes used to ensure the accuracy and validity of the data.
- Ethics and privacy: Information on any ethical or privacy considerations related to the data.

To capture this documentation and metadata, a variety of methods will be used, including:

- Creating a Data Management Plan (DMP) that outlines the data collection, management, and sharing processes.
- Developing a Codebook that describes the variables and their values in the data.
- Creating a README file that provides a brief description of the contents of each file.
- Including metadata tags in the videos and documents to capture relevant information.

The metadata standards will depend on the data collection type and the relevant metadata standards. Various standards, such as the Dublin Core, ISO15836-1:2017, Data Documentation Initiative (DDI), and Metadata Encoding and Transmission Standard (METS), will be used for appropriate research data. The chosen standards will ensure that the metadata is interoperable and easily discoverable by other researchers. We will use Zotero software to assist us in collecting, organizing, annotating, citing, and sharing our research.

Ethics and Legal Compliance

How will you manage any ethical issues?

To ensure that our study adheres to the ethical guidelines of the New Jersey Institute of Technology, the College of Science and Liberal Arts, and the Department of Humanities and Social Sciences, we will take the following steps:

- **Informed consent:** Before collecting any data, we will obtain informed consent from participants to ensure that they have a comprehensive understanding of the study's nature and purpose and their rights concerning their participation. Participants will be informed of the potential risks and benefits of the analysis, the data collection and storage methods, and any possible data sharing. They will also be informed of their right to withdraw from the study at any time.
- **Protecting participant identity:** If necessary, we will safeguard participant identity by anonymizing their data. Any identifiable information, such as names, email addresses, or student identification numbers, will be removed from the data before analysis. In cases where participants cannot be fully anonymized, such as in video recordings, we will obtain their consent for using and storing the recordings.
- **Secure data storage and transfer:** We will ensure secure data storage and transfer to prevent unauthorized access or disclosure. We will store the data using Google Drive, an encrypted platform with appropriate access controls. Access to the data will be restricted to authorized personnel, and we will use passwords and/or encryption to protect the data during transfer.
- **Compliance with relevant regulations:** Our project will comply with all applicable laws and regulations, such as the General Data Protection Regulation (GDPR) or the Health Insurance Portability and Accountability Act (HIPAA), depending on the data collection type. We will ensure that we obtain any necessary approvals, such as Institutional Review Board (IRB) approval, before commencing data collection.

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

The participants will hold the IP and Copyright on their rubrics in this study. No other IP or copyrighted material will be generated within the study.

- NJIT will own the data generated in this study and will be licensed for reuse under the Creative Commons (CC) license. This license will allow interested stakeholders to access and use the data generated in the study for non-commercial purposes.
- If third-party data is used in this study, I will ensure that any restrictions on the reuse of the data are respected. I will also seek permission from the data owners before using it in the study.
- In this study, we will not delay or limit data sharing. Nonetheless, any publication or patent-seeking activities will only occur after the data has been thoroughly analyzed and interpreted. This

approach will guarantee that the conclusions drawn from the data are precise and dependable before they are disseminated to the public.

Storage and Backup

How will the data be stored and backed up during the research?

NJIT's digital services will handle all aspects of data management. Through Google Education, we have infinite storage that can remain indefinitely. Minimal data is generated, and duplicates will be stored offline on an external hard drive to ensure safety.

Responsibility for backup and recovery will be shared among the research team. Each team member will ensure their data is backed up regularly and stored securely. We will use automated and manual backup processes to ensure our data is secure and accessible during an incident.

In the event of an incident, we will follow our established data recovery procedures. We will first identify the cause of the incident and then restore the data from the most recent backup. If necessary, we will engage with IT professionals to ensure our data is recovered successfully.

How will you manage access and security?

As part of our data management plan, we will take several measures to ensure access and security.

Initially, we will analyze and evaluate the potential threats that could affect data security. Subsequently, we will establish appropriate measures to manage these risks. These strategies could involve implementing strong passwords and encryption protocols, restricting access to authorized personnel only, and performing routine surveillance to detect unauthorized access.

Access to the data will be controlled through appropriate permissions and access controls. Only authorized personnel, the project team, has approved will be granted access to the data. Collaborators who require access will be provided with secure login credentials and must sign a data-sharing agreement outlining their data management and security responsibilities.

If data is collected in the field, we will ensure safe transfer into our main secured systems using encrypted

transfer protocols and secure storage devices. We will also ensure that any data collected is backed up regularly to prevent loss due to technical failures or other incidents.

Overall, we will take a comprehensive approach to data security and access control, utilizing best practices and appropriate technologies to ensure the data remains secure throughout the research process.

Selection and Preservation

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

All rubrics and codebooks are considered to be of "long-term value."

As standard practice, project personnel maintains notebooks for raw or hand-recorded data. Notebooks and included data will be supported by the specific personnel and stored within their respective office. In addition, department staff will electronically scan all notebooks annually to facilitate electronic archiving on the university server for added backup protection.

Identifying which data must be retained or destroyed for contractual, legal, or regulatory purposes is essential. This may vary depending on the specific project and its requirements. For instance, some data may need to be retained for several years to comply with regulations or contractual agreements and or all Webex recordings will be deleted 3 years after the completion of the 6-semester study. Therefore, consulting with legal and regulatory experts is essential to determine the appropriate retention period for different data types.

It is crucial to determine which other data to keep, assessing their possible uses in future research or if they hold value beyond the current project. It's also important to consider the data's worth to the project itself, as some data could be helpful in later analysis or verifying the findings. Therefore, keeping more data is suggested as long as the expenses of storing and preserving them are feasible.

Foreseeable research uses for the data may include follow-up studies, replication studies, or meta-analyses. Data may also be of value to other researchers or institutions working on related topics. Therefore, keeping track of potential research uses or interests is crucial, as this may inform decisions about managing and sharing the data.

The time that data should be retained and preserved may depend on the specific project, the type of data, and any relevant legal or regulatory requirements. However, keeping and maintaining data for as long as they have potential research value is generally recommended, as developing a plan for long-term preservation that considers the appropriate storage formats, metadata, and access controls. It may also be necessary to periodically review and update the preservation plan to ensure the data remains accessible and usable over time.

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

The long-term preservation plan for the dataset involves permanently maintaining all raw data, analyzed data, reports, lab books, and publications. These data will be held for at least three years from the end of any NSF award period or three years from the date of public release in compliance with NSF policies.

To determine where the data will be held, it is vital to consider reputable and secure data repositories or archives capable of preserving it for the required duration. The selected storage or archive should also be accessible and compatible with future technologies.

The selected data repository or archive costs should be considered and budgeted accordingly. For example, some repositories may charge fees for data deposition, curation, or long-term storage. Therefore, it is essential to carefully evaluate the costs and benefits of different options to select the most appropriate and cost-effective repository or archive.

Preparing data for sharing and preservation can be lengthy and requires careful planning and budgeting. This process typically involves cleaning and documenting the data, formatting it to conform to standard guidelines, and ensuring that all relevant metadata is included. Therefore, allocating sufficient resources and time to this task is essential to ensure that the data is properly preserved and accessible to all who need it.

Data Sharing

How will you share the data?

The project group and external educational and research community can access the data. To facilitate internal dissemination, project personnel will be encouraged to share data openly and store data needed by multiple stakeholders in appropriate folders. Shared data needs will be identified at project team meetings and through peer-to-peer interactions. Significant results will be disseminated externally

through peer-reviewed journals, national and international conferences, and depositing data onto globally accessible databases where appropriate. Regular reports to NSF or sponsors may also be required. All data, supporting information, and messages from the project funded by NSF will be made available to requesters within one month, subject to NJIT intellectual property requirements and NSF policies. A persistent identifier will be pursued to make data citation and retrieval easier. There is no mention of a separate plan for sharing the data.

Questions to consider:

How will potential users find out about your data?

Potential users will be informed about data availability through peer-reviewed publications, conference presentations, and regular reports to NSF.

With whom will you share the data, and under what conditions?

Data will be shared with external educational and research communities and project team members. Sharing will be subject to any restrictions imposed by NJIT intellectual property requirements and NSF policies.

Will you transfer data via a repository, handle requests directly or use another mechanism?

Data will be transferred via globally accessible databases, direct requests, or another suitable mechanism.

When will you make the data available?

Data will be available within one month of the request, subject to applicable restrictions.

Will you pursue getting a persistent identifier for your data?

A persistent identifier will be pursued to facilitate data citation and retrieval. (e.g., DOI, ISBN, etc.)

Do you have any plan for divulgation separate from dissemination?

A separate plan for divulgation is not mentioned in the given text

Are any restrictions on data sharing required?

Video raw data will not be shared in order to ensure confidentiality.

Responsibilities and Resources

Who will be responsible for data management?

D. James Lipuma will be responsible for implementing the DMP, management activities and responsibilities.

What resources will you require to deliver your plan?

Third-party software such as Zotero and Grammarly will be utilized in this process.

All additional needed resources and support will be provided by the existing NJIT departments and existing facilities and infrastructure.

Planned Research Outputs

Text - "Report"

A report of descriptive statistics and rubric elements

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Report	Text	2024-02-29	Open	None specified	1 MB	Creative Commons Attribution Non Commercial Share Alike 4.0 International	Dublin Core DDI (Data Documentation Initiative)	No	No