Introduction

Introduction. NNIPHQ would like to provide more opportunities for partners to learn from one another and guest experts about data management, specifically data intake, storage, processing, and cloud usage. To help our planning, we would like to gather information about your current practices and challenges. Please have someone familiar with your data management practices complete the survey. The full set of questions are posted here [add link to activity on website].

We expect this survey to take about 10-15 minutes. Thank you for your time!

Note: If you are a unit or center within a larger institution, we would like you to answer the questions for your unit or center, not your institution.

Background

Q1. Name

Q2. Title

Q3. Organization

[   ]

Q4. Email

[   ]

Data Intake

Q5. What methods does your organization use to take in data? Check all that apply.

Manual data downloads

Scheduled data delivery from third parties

Downloading from a Database (Postgres, SQL Server, etc.)

Downloading from an API (i.e. Census APIs)

[   ] Other (please specify)

Q6. What computer software do you use for data intake (how your organization collects and downloads data from external sources)? Check all that apply.

R

SAS

STATA

Excel

ArcGIS

Tableau

SQL

Python

Shell (i.e. Bash scripts)

[   ] Other (please specify)

Q7. Please share any additional details related to data intake methods or software that you would want NNIP to understand.

```
┌─────────────────────────────────────────────────────────┐
│                                                         │
│                                                         │
│                                                         │
│                                                         │
│                                                         │
│                                                         │
│                                                         │
└─────────────────────────────────────────────────────────┘
```

Q8. Do you follow any metadata standards?

No

Yes, Federal Geographic Data Committee Standard

Yes, Data Documentation Initiative (DDI)

Yes, In-house standard

Yes, Dublin Core

```
┌──────────────────────────┐
│                          │ Yes, other (please specify)
└──────────────────────────┘
```

Q9. For which of the following files types do you have a standardized folder/directory structure? Check all that apply.

Raw data

Processed data

Metadata

Related documents (data use agreements, documentation from source)

No standard folder/directory structure

Data Storage

Q10. What file formats are your processed and raw data typically stored in? Check all that apply.

CSVs/text files

Excel files

SPSS/Stata/SAS files

Shapefiles/GeoJSONS/other geographic filetypes

In relational databases (i.e. Postgres, MySQL, SQL Server)

Source format

[   ] Other

Q11. Where are your organization's data stored? Check all that apply.

In-house servers/shared drives

Local computer drives

Google Drive/Box/Dropbox

Other Cloud services (i.e. Amazon S3, Azure Blob, Google Storage)

[   ] Other

Q12. Please share any additional details related to where your organization's data are stored that you would want NNIP to understand.

[   ]

Q13. Do you maintain an overall data inventory or catalog?

Yes

No

Q14. What procedures do you use to safeguard confidential data? Check all that apply.

Encryption at rest (i.e. drive encryption, encrypted cloud storage)

Encryption in transit (i.e. secured SFTP transfers, encrypted cloud transfer)

Password protection

Multi-factor authentication

Signed data confidentiality forms

MOU clauses specifying minimum cell size or other standards for release

Storage in locked cabinets

No formal data protections

We don't store any confidential data

☐ Other

Q15. What is your organization's procedure for backing up your data? Check all that apply.

Scheduled backups to a local drive

Scheduled backup to the cloud

We use cloud storage which have automated backups

We don't have specific backup procedures

☐ Other

Q16. Please share any additional details related to data storage that you would want NNIP to understand.

Data Processing

Q17. What tools do you use for data processing? Check all that apply.

R

SAS

Stata

ArcGIS

Tableau

SQL

Python

Shell (i.e. Bash)

Other

> [   ]

Q18. Does your organization face any of the following challenges when performing data processing? Check all that apply.

Limited computation power

Memory limits (i.e. datasets are too large to fit in memory)

Long runtimes for scripts

Licensing costs (i.e. costs of SAS or ArcGIS licenses)

[   ] Other

Q19. How often do you store code or documentation for cleaning/processing data?

Always

Frequently

Sometimes

Rarely

Never

Q20. Please share any additional details related to data processing that you would want NNIP to understand.

[   ]

Cloud Usage

Q21. How would you rate your organization's familiarity with cloud storage and computing on Amazon Web Services (AWS), Azure, or Google Cloud Platform (GCP)?

Extremely familiar

Very familiar

Somewhat familiar

Not so familiar

Not familiar at all

Q22. Which of these cloud providers does your organization use? Check all that apply.

AWS

GCP

Azure

[   ] Other

We don't use cloud providers

Q23. Are there any barriers that limit your familiarity with or use of cloud providers? Please describe.

[   ]

Q24. For which of the following procedures does your organization currently use cloud storage and computing services? Check all that apply.

Data intake

Data storage

Data processing

[   ] Other

We don't use cloud services for any of the above

Q25. What specific cloud services do you use (i.e. EC2 instances for virtual computers, Azure Blob for storage, etc.)?

[   ]

Q26. Please share any additional details related to cloud storage that you would want NNIP to understand.

Organization

Q27. If your unit/center is part of a larger institution (such as a university), what role does the larger institution play in your data intake, storage, and processing? Please briefly describe any constraints you face or benefits you receive from your larger institution.

Q28. Would you be willing to share your procedures and/or standards for data intake, storage, and processing?

Yes, with NNIP partners upon request

Yes, with NNIP partners through the NNIP website

Yes, with the public through the NNIP website

No

Q29. How has COVID-19 and remote work changed your data intake, storage, and processing procedures, if at all?

Q30. What would you like to learn more about or have trainings from NNIP regarding data intake, storage, processing, or cloud usage?

Q31. Do you contract out services for data management (intake, storage, processing, maintenance, cloud usage, etc.)?

Yes, please describe

No

Q32. Is there anything else we should understand about your organization to understand your data management practices and context?

Powered by Qualtrics