R01 - RNS - SOP

## Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Editor** | **Date** | **Comments** |
| 0.1 | Brian P. | 03/14/23 | Initial Draft |
|  |  |  |  |
|  |  |  |  |

## Purpose

Collaborate with remote sites to create a central data repository.

## Definitions

* Gold Path – UCSF, NYU Collaborators
* Silver Path – Collaborators who need processing aid

## Teams

### Administrative Team

Team responsible for administrative oversight of the project.

#### Members

* Michelle Dunn
* Carolyn Wilkinson

### Data Science Team

Team responsible for data management, data transformations, and data enrichment.

#### Members

* Brian Prager
* Joshua Asuncion

### Pennsive Team

Team responsible for portions of the workflow that interact with the Pennsieve data platform.

#### Members

* Darrell De Freitas

## Sizing Scale

These sizing scales are meant to provide an upper-limit on the time it may take to complete.

* **Small:** 1 to 3 Business Days (Sub-work week)
* **Medium:** 5 to 10 Business Days(1 to 2 work weeks)
* **Large:** 1 month+
* **X-large:** Ongoing

When possible, we also include whether a project includes discovery. Discovery based tasks tend to be larger, as investigation of methods may reveal significant hurdles to overcome.

## Workflow

1. Data collection at remote site
   1. **Pipeline Item:** 1.b.M
   2. **Host or Client Responsibility:** Client
   3. **Sizing:** Small
   4. **Software or Documentation Requirements:** Redcap survey for Natus details
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. For Gold-path collaborators, we will provide the SoP and rely on their researchers to conform to the requirements based on their own workflow.
      2. For Silver-path collaborators, we may need to provide more support to ensure data is recorded correctly.
2. Clinician annotates EEG
   1. **Pipeline Item:** N/A
   2. **Host or Client Responsibility:** Client
   3. **Sizing:** Small
   4. **Software or Documentation Requirements:** N/A
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. Site is responsible for any annotations it deems necessary.
3. Annotation data capture
   1. **Pipeline Item:** 2.c.J
   2. **Host or Client Responsibility:** Client
   3. **Sizing:** Medium (Discovery)
   4. **Software or Documentation Requirements:** Channel Mapping
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. We need to decide how to store seizure information from clinical notes. A few options are as follows:
         1. Store Annotation Layer + No additional metadata:
            1. Pros: We retain the clinician notes.
            2. Cons: PHI concerns.
         2. Remove Annotation Layer + No additional metadata:
            1. Pros: Easy to implement the pipeline. No PHI concerns.
            2. Cons: Clinician required to reidentify seizures.
         3. Remove Annotation Layer + Additional metadata:
            1. Pros: No PHI.
            2. Cons: Research coordinator needed to enter metadata about seizures into Redcap or a bids compliant file.
4. EDF Conversion
   1. **Pipeline Item:** 3.b.J
   2. **Host or Client Responsibility:** Host \*(See Notes)
   3. **Sizing:** Medium to Large
   4. **Software or Documentation Requirements:** Natus, Persystm Bash Scripts, PMACS Pipeline, Redcap documentation about raw data format
   5. **Data Transfer:** 
      1. *Destination:* PMACS Server (if Silverpath)
      2. *PHI Present:* Yes
   6. **Notes:**
      1. Gold-path sites will be doing this on their own. Silver-path may or may not be able to do this, so we need to ensure we can process their data.
      2. We have existing scripts that can scrub PHI from the header after conversion to EDF.
      3. We need to create a PMACs pipeline that is automatically kicked off when data arrives to do this step and subsequent steps. This will require a fair bit of work with the PMACs team and may increase development time.
      4. We will want to dockerize this process to ensure long-term stability behind the PMACs firewall.
5. Prebids Formatting
   1. **Pipeline Item:** 3.a.M
   2. **Host or Client Responsibility:** Client
   3. **Sizing:** Small
   4. **Software or Documentation Requirements:** SOP Documentation
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. We need to provide documentation for Silver-path sites to put data into a format we can convert to BIDS
6. Deep Storage
   1. **Pipeline Item:** 2.b.M
   2. **Host or Client Responsibility:** Client
   3. **Sizing:** Small
   4. **Software or Documentation Requirements:** N/A
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. Need to provide estimate for retention policy. This could be based on pipeline timing, or on a flat time period provided by stakeholders.
      2. Should confirm with sites about their storage capabilities.
7. PHI Check
   1. **Pipeline Item:** 2.a.B
   2. **Host or Client Responsibility:** Host
   3. **Sizing:** Extra Large
   4. **Software or Documentation Requirements:** Script to flag if any potential PHI is present.
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. This is a step to mitigate risk. It should be done regardless of path.
      2. Sites should still aim to remove PHI, this is just to be a final check.
      3. This can be an ongoing feature improvement, as we create or implement better code.
8. Data Quality Check
   1. **Pipeline Item:** 3.a.B, 3.b.B, 3.c.B
   2. **Host or Client Responsibility:** Host
   3. **Sizing:** Large/Extra Large
   4. **Software or Documentation Requirements:** TBD
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. Clinician/Researcher input needed to provide the best QA options.
9. User Alert (Pass/Fail)
   1. **Pipeline Item:** 4.a.J
   2. **Host or Client Responsibility:** Host
   3. **Sizing:** Medium
   4. **Software or Documentation Requirements:** PMACs Pipeline, Clickup Trigger
   5. **Data Transfer:** N/A
      1. *Destination:*
      2. *PHI Present:*
   6. **Notes:**
      1. We need to inform the user if their data can safely be uploaded to Pennsieve, or whether we require additional input/a new upload from the client.
10. BIDS Conversion
    1. **Pipeline Item:** 4.b.J
    2. **Host or Client Responsibility:** Host
    3. **Sizing:** Small
    4. **Software or Documentation Requirements:** BIDS conversion script
    5. **Data Transfer:** N/A
       1. *Destination:*
       2. *PHI Present:*
    6. **Notes:**
       1. We already have a bids converter, but it needs to be operationalized.
11. Upload to Pennsieve
    1. **Pipeline Item:** 4.a.D
    2. **Host or Client Responsibility:** Host
    3. **Sizing:** Medium to Large (Discovery)
    4. **Software or Documentation Requirements:** PMACS Pipeline, Pennseive Install, Firewall setup
    5. **Data Transfer:** 
       1. *Destination:* Pennsieve
       2. *PHI Present:* No
    6. **Notes:**
       1. We need to ensure that the PMACs instance (a computer or VM) can install and run Pennsieve.
       2. We need to automate an upload in the PMACs pipeline
       3. We need to ensure that Pennsieve settings (guest status, tags, duplications) are all correctly configured.