

# Pipeline Setup for Data Import

- 1. Document the Data Import Pipeline:**

Outline the steps required to set up the data import pipeline.
- 2. Identify Missing Workers:**

Compile a list of workers that are not included in the current implementation.
- 3. Determine Appropriate Module:**

Identify the module where each worker should be located.

  - If the necessary module does not exist, create a new one.
- 4. Define State File Format:**

Specify the state file format that each worker should operate with.
- 5. Evaluate Existing Options:**

Determine if the current options are sufficient to complete the scenario.

  - If not, additional development of the `generate-state` function is required.
- 6. Update Vault (if applicable):**
  - If Vault is used, add the necessary data to Vault.
- 7. Create Input Files for the Pipeline:**

Generate input files for the pipeline, considering the missing workers.
- 8. Run State Generator for the First Input File:**

Execute the state generator on the first input file.
- 9. Reach Missing Worker via Workflow Manager:**
  - Proceed to the missing worker by running the workflow manager.
- 10. Create Worker Stub for Payload Display:**
  - Develop a stub for the worker to display the payload.
- 11. Write Missing Worker Script:**

Develop the script for the missing worker.
- 12. Run Workflow Manager Post-Debugging:**
  - After debugging, restart the workflow manager.
- 13. Create Production Input Files:**
  - Create input files for production (daily and historical data).
- 14. Set Up Cron Job for Daily Tasks:**
  - Schedule the daily tasks using a cron job (pending update to version 1.9.0).

---

Revision #2

Created 2024-08-30 14:56:06 UTC by Anton Petrushkaneki

Updated 2024-08-30 14:56:56 UTC by Anton Petrushkaneki