National Clinical Cohort Collaborative

N3C Getting Restarted Guide

A reorientation to the reopened N3C

https://unite.nih.gov

Version 1.1, April 2026

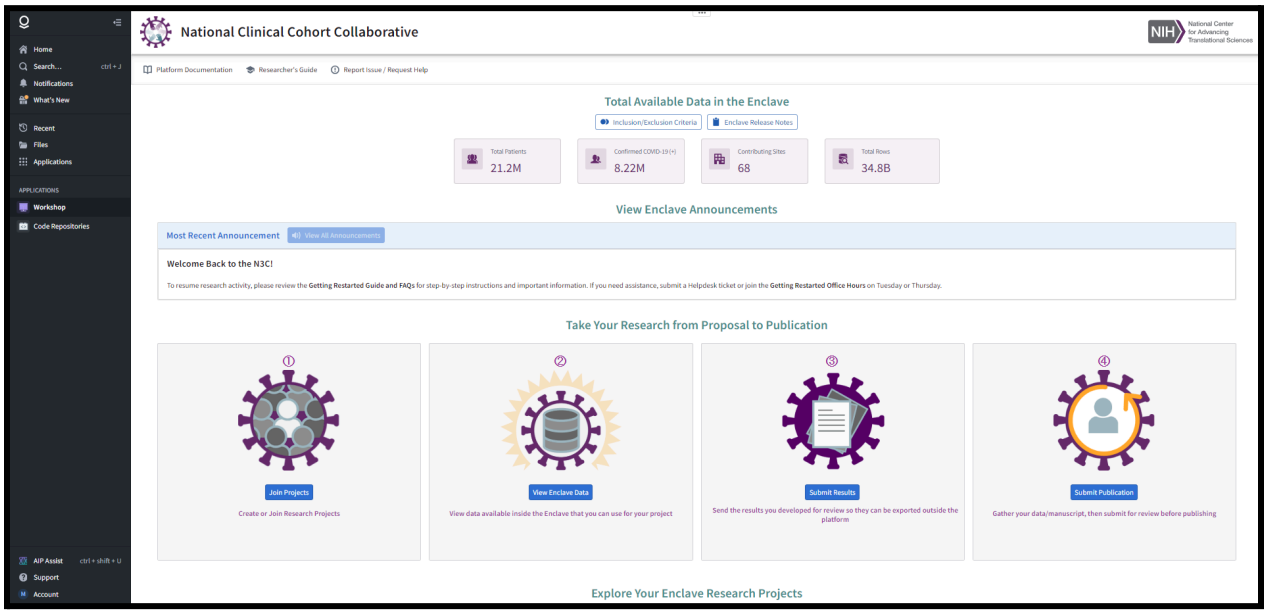


Table of Contents

Welcome back, N3C Community!

We are pleased to share that we are reopening the N3C Data Enclave, including an enhanced infrastructure and the data and core tools in Palantir Foundry's trusted research environment. You can access the N3C at https://unite.nih.gov.

A phased implementation. Over 2026, the N3C will continue to transform the platform to provide the Dynamic Workspace model to support researchers, additional governance options for institutions, federal Controlled-Access Data Repository (CADR) standards, and expansion from COVID to the full spectrum of clinical contexts and data modalities. As these updates and enhancements are implemented, the full set of N3C tools, features, and workflows will return in phases as components are tested and validated.

Priority access for the N3C Community. To help you return to your research as quickly as possible, we are providing priority access to all current N3C members while this transformation and expansion is underway. Once the updates are fully implemented, N3C will open again to new users with a new set of agreements (Data Transfer and Data Use) that will need to be executed..

Unchanged security and access protocols. As always, N3C data and researcher workspaces remain secured and accessible only to authorized project personnel. All data access, analytic workflows, and exports continue to require compliance with established N3C privacy, security, and governance protocols, including formal N3C Data Access Committee (DAC) approval, role-based permissions, audit monitoring, and public posting of approved project rationales. These protections remain unchanged and fully enforced.

The following N3C Getting Restarted Guide will help you:

- Quickly resubmit your DUR, rebuild your N3C workspace, and return to your research;
- Understand which tools and features are available now and which are coming soon;
- Quickly submit publications for fast-track N3C Publication Committee review; and
- Find help, support, and more information.

Getting Restarted Support:

- N3C Support Desk
- Office Hours - Tuesday, March 3 | 1:00–2:00 PM ET | Teams link
- Office Hours - Thursday, March 5 | 1:00–2:00 PM ET | Teams link

We recognize that this transition has been challenging and has required patience and flexibility from the N3C Community. Your commitment to collaborative, high-impact science remains central to this effort, and we look forward to moving forward together to continue to transform real-world data insights into meaningful health and health care solutions.

Onward, together! - Your N3C Team

2025 Pause & 2026 Expansion

The National Clinical Cohort Collaborative (N3C) represents one of the largest and most impactful real-world data resources assembled to advance clinical and translational research, and to improve health in the United States. Through the sustained collective efforts of clinicians, scientists, leaders, communities, academic institutions, health systems, and data partners across the country, N3C has enabled discoveries that would not have been possible within any single institution. Five years later, the N3C continues to transform real-world clinical data into actionable insights that inform biomedical research, public health strategy, and patient care.

2025 Contract Term Negotiations

N3C went offline on September 27, 2025 as part of a Health and Human Services (HHS)-wide effort to consolidate pricing for platform technology and infrastructure. In alignment with this directive, the N3C Data Enclave paused while contract payment terms were negotiated. This pause did not reflect any concerns about N3C itself, in fact, during the subsequent government shutdown, special exceptions were approved to ensure N3C data safeguarding was able to continue. During this time no data was moved outside of N3C and no data was used or accessed in a way that does not align with N3C governance. Protections of N3C data were and remain NCATS' top priority and have never been compromised.

The pause also allowed for N3C to accelerate the timetable for infrastructure updates needed for the transformation and expansion of N3C to cover the clinical continuum and for the integration of new data modalities. The next Phase of N3C will advance several strategic objectives:

Resilient Architecture
A modular structure to reduce single points of failure and strengthen operational continuity

Infrastructure Flexibility
Core components designed to improve portability and reduce long-term infrastructure risk

Alignment with the CADR/DWS Framework
Implementation of the federal Controlled-Access Data Repository and Data Workspace framework to operationalize FISMA and modern cybersecurity requirements

---

CRITICAL ASSURANCE: Throughout the pause and this transition to the next phase of N3C, NCATS has and will continue to maintain exclusive stewardship of the data. Privacy, security, and institutional trust continue to be non-negotiable safeguards. NCATS and N3C leadership are working closely with HHS and NIH to expedite the reopening and return of feature parity and establish a sustainable, stable path forward.

2026 Expansion

The next phase of N3C's development focuses on continuing the planned growth and expansion of N3C.

Across the Clinical Continuum
Expanding the N3C beyond its origins to support research across all therapeutic areas.

Dynamic Workspaces
Creating secure, researcher defined phenotype-based workspaces tailored to approved research projects and roles, supporting evolving scientific needs across clinical contexts.

Increased Community Engagement
Establishment of a Strategic Partners Committee to engage members, partners, and patients, in providing structured and transparent community input.

As we reopen the environment, N3C will continue to serve as a durable national platform for collaborative, data-driven discovery.

Please note that the N3C and NCATS websites will be updated soon to reflect the reopened platform, access to legacy features, links to tutorials, and new guidelines. Current approval processes for federal web page content updates have delayed content updates slightly and therefore we ask for your patience in dealing with mismatches between this Guide and what is on our website. This Guide should be seen as the latest and most up-to-date information if there is a discrepancy between it and the website.

What This Means for Returning Users

As you return to the N3C, you will notice both continuity and change. Core principles such as NCATS data stewardship, enclave-based access, and DAC-governed data use remain unchanged. However, the platform has transitioned to a more resilient, modular architecture with important changes to tools and workflows.

To help our Community rapidly resume N3C research, the team has focused on ensuring core infrastructure, data, tools, and key features are available on day 1. We will introduce additional tools and capabilities in the coming months.

Available Now

READY TO USE

- Login for Current Users
- Personal & Project Workspaces
- Largest and Most Recent COVID Data Releases (192 and 197) - in both De-identified and Limited Data Set (LDS) formats
- Synthetic & Public Datasets
- OMOP Data Model Structure & Ontology
- Code Workspaces (JupyterLab, VS Code, RStudio IDE) - replaces Code Workbooks
- Code Repositories: Available with PySpark and Spark SQL support
- Data Download Requests
- Publication Review Requests
- Palantir Foundry research environment
- N3C Office Hours & Support Desk

Coming Soon

IN DEVELOPMENT

- Registration for New Users in the CADR/DWS Model
- Training Portal
- New Institutional Data Transfer Agreements across Clinical Contexts
- 3 Governance Options
- Data Releases Across Clinical Contexts
- Dynamic Workspaces for Researchers
- Data Access Management Environment (DAME)
- SetArchitect
- PhenoEx
- Wayfinder
- Strategic Partners Committee
- Community Forum
- Domain Teams
- Refreshed N3C and NCATS websites with updated resources

Materials that Need to be Regenerated in the Reopened Environment

Access to environment elements in N3C has been changed to enhance sustainability for this next Phase of N3C. This means that sometimes you will need to recreate or redeploy elements that are not currently accessible. The following list provides details about each of them.

Research Project DURs
Investigators will need to submit their DURs for reapproval (expedited review) and rebuild their projects using their code and concept sets.

Intermediate or Cached Datasets
If you created derived or intermediate datasets in your project, you will need to regenerate them in the reopened platform. Retention at scale of these artifacts was not feasible for the new phase of N3C.

Materials that Are not Available in the Reopened Environment

Certain environment elements were not able to be made accessible in the reopened N3C.

Historical Download and Publication Records
Download Request and Publication Intent records from late 2024 onward are not available.

Historical Data Releases
All data have remained untouched, secure, and have not been moved outside of N3C during the offline period or since. The last data release prior to Data Transfer Agreement expiration [192] and the most recent data release [197] are available and accessible. However, other data releases (both LDS and de-identified) have been archived and user access removed to make N3C sustainable for this next Phase. These datasets are retained to preserve scientific integrity of research.

COVID-19 Data Submissions
N3C is no longer accepting new COVID-19 data submissions. COVID-19 data contributions are frozen as of October 2025 and preserved in archive form to support research continuity. Again, all data have remained untouched, secure, and have not been moved outside of N3C during the offline period or since.

Note: Archived COVID-19 datasets remain accessible to approved researchers to support completion of ongoing analyses and manuscripts.

Quick Restart Guide:

The following steps will help you quickly rebuild your workspace and research:

- **Step 1: Log in to the reopened N3C**
  - Navigate to https://unite.nih.gov
  - Use your existing institutional credentials to log in
- **Step 2: Request Your Exported Code**
  - Submit a Code Export Request if you didn't export your code during the December reopening
  - Open a ticket: Link
- **Step 3: Resubmit Your Project DUR for Expedited Review**
  - PIs: Submit a new DUR using "Create a New Project"
  - Team Members: Request to join your PI's approved DUR
  - Returning users qualify for rapid review (3-4 business days)
- **Step 4: Familiarize Yourself with Code Workspaces and Code Repositories**
  - Code Repositories: For large-scale data transformations (distributed Spark)
  - Code Workspaces: For interactive analysis (JupyterLab, RStudio, VS Code)
  - See Appendix A for detailed documentation
- **Step 5: Adapt Your Code for the reopened N3C Platform**
  - Import datasets and create tables using Fusion sheets or Pipeline Builder
  - Organize code into folders (Initial Builds → Domain Tables → Final Tables)
  - Build datasets in dependency order
- **Step 6: Publication Intent and Data Download Requests**
  - Complete your analysis in Code Workspaces
  - Submit Data Download Requests as needed
  - Submit Publication Intent forms for manuscripts

Where to Get Help?

Getting Started Office Hours: Tuesdays & Thursdays, 1-2pm ET | Teams Link

Platform Support - Submit platform support tickets through the Foundry support tickets

General Questions & Support Needs - Submit a ticket to the N3C Support Desk

Step-by-Step Guide

Step 1: Log in to the reopened N3C

1. Navigate to the N3C portal URL (unite.nih.gov)
2. Use your existing institutional credentials to log in
3. Complete any additional authentication steps, if prompted
4. Verify your profile information is up-to-date and correct

If you encounter issues logging in:

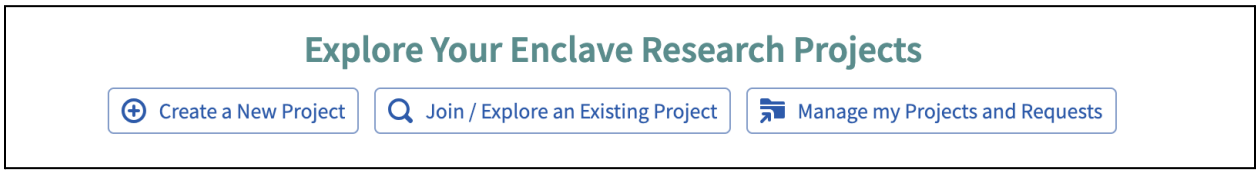
1. Verify your institutional email is still active
2. Check that your institution's Data Use Agreement (DUA) is executed
3. Contact the N3C Support Desk

Step 2. Resubmit Your Project DUR for Expedited Review

To uphold our N3C policies and security procedures, all Principal Investigators (PIs) will need to resubmit their DURs and, once approved, re-add their approved team members.

Use the "Create a New Project" button to resubmit your DUR

1. On the N3C Home Page
2. Click on "Create a New Project" and request a new DUR



3. Wait for DUR approval from the DAC. Approval may take 3-4 business days after your submission. If you have not heard back, feel free to open a ticket.
4. Add team members and create your analytic environment

TIP: You can review your previous projects's titles, IDs, and abstracts by navigating to 'Manage my Projects and Requests' and clicking on the 'Renew' tab. These are placeholders to refresh your memory—you won't be able to rejoin these projects directly. Everyone will need to submit a new DUR form, using the same project information, and it will be rapidly reviewed for reapproval and reactivation.

Note: the DURasks for project rationale. If you do not have a copy of your project rationale, write "continuation of previously approved project" and include your prior project ID. The review committee can reference your previous submission.

Join a DUR (for Team Members)

Wait for your PI to complete DUR approval, then request to join.

1. Confirm your PI has received DUR approval
2. Obtain the approved DUR name from your PI
3. Request to join at: N3C Explore Projects.
4. Wait for PI and DAC approval of your join request.
5. Approval may take 3-4 business days. You can open a ticket if you haven't heard back.

Request Your Code

- If you did not have a chance to export your code yourself during N3C's brief re-opening in December 2025, the N3C team can help you retrieve it. Contact the N3C team to request your exported code:
- Open a ticket: Link
- Subject Line: N3C Code Request - [Your Name]
- Include: Your project name/ID and specify which code assets you need

Step 3. Familiarize Yourself with Code Workspaces and Code Repository

The reopened N3C platform uses Palantir Foundry's **Code Workspaces** and **Code Repositories** instead of Code Workbooks. These tools provide a professional development environment with your preferred IDE (Integrated Development Environment), Git version control, and production-ready code capabilities.

See Appendix A for Palantir Foundry's detailed documentation on Code Workspaces and Code Repositories.

TIP: If you originally worked in Code Repositories (not Code Workbooks), your code may run with fewer modifications. Code Repositories are fully supported in the reopened environment, and you may not need to follow all of Steps 5 and 6.

Step 4. Adapt Your Code for the reopened N3C Platform

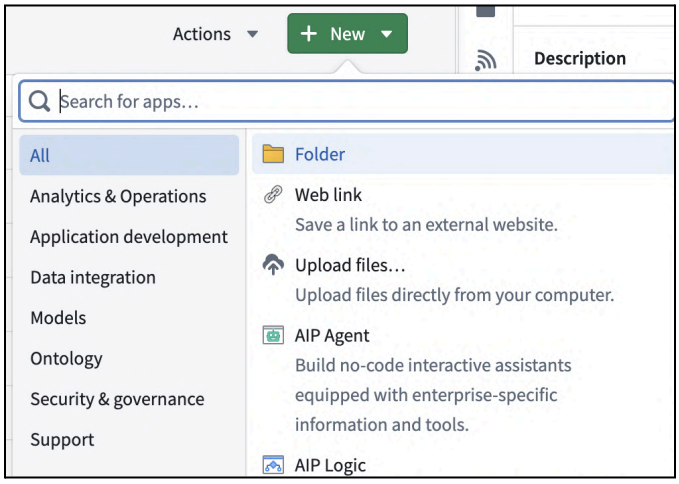
Your exported code from Code Workbooks will need slight modifications to work in the reopened environment. Planning and organization before copying files into Code Repositories will help ensure this transition is as easy as possible. Failure to properly inventory datasets, and sequence dependencies can result in build failures.

We recommend running larger distributed transforms in Code Repositories and then reading the cleaned and filtered tables to Code Workspaces for further analysis. The Code Workspace environment is best used for statistical analysis, modeling, and data visualization.

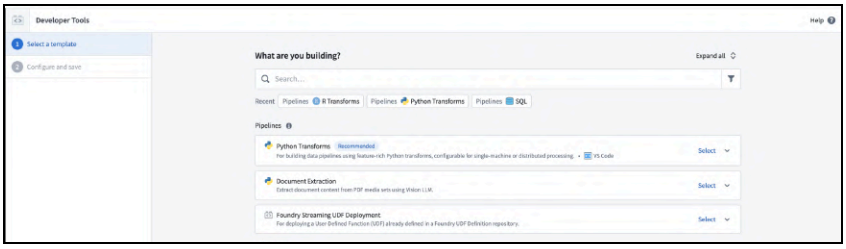
Creating a Code Repository

To begin, open your DUR and create a new Code Repository:

1. Select the "New" button in the top right corner
2. A dropdown will appear - type Code Repositories and click on the icon



3. Select Python Transforms Recommended (image below) if you are working in Python. If your analysis will be in R or SQL, select the appropriate option. Examples in the following tutorial will use Python Transforms.

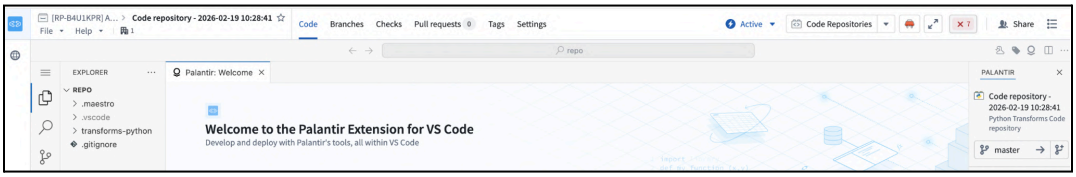


4. Name the Repository and ensure it is located in the correct DUR.

5. Select "Open in VS Code."

6. Once you have VS Code open, navigate to the top right corner of the webpage.

7. Select Code Repositories.



8. This will activate a new workspace where you will begin inputting your code.

NOTE: Code Repositories use distributed Spark environments to transform large datasets using SparkSQL, or PySpark. Your code from Code Workbooks will work as expected in this new environment. The only change is that you will need to retype the inputs and outputs so that the datasets reference the correct location of your data.

Before You Begin:

Prior to setting up your repository, ensure:

- All datasets are identified
- Supporting assets are recreated (Selected concept sets)
- File structures are organized by function and type
- Execution order is clearly defined

A disciplined, front-loaded approach significantly reduces rework and ensures a stable, reproducible build.

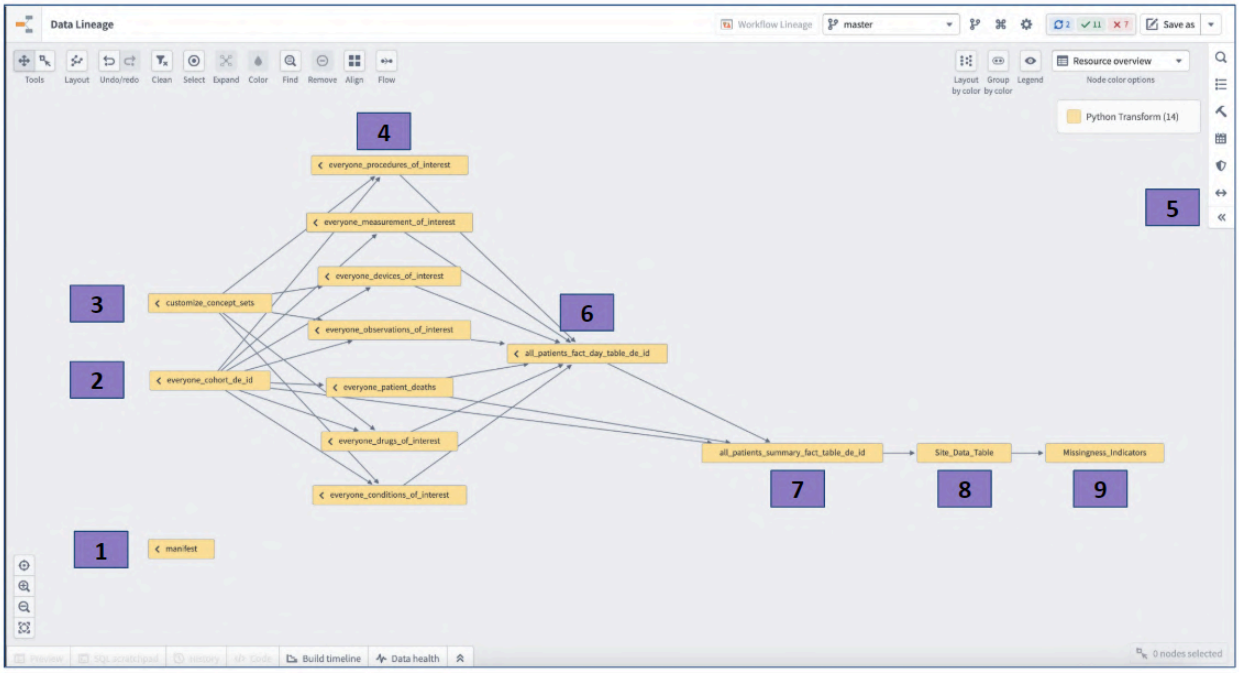
Recreating Previous Work in Code workbook to Code Repository: Example Using Logic Liaison Template

Many N3C researchers used Logic Liaison templates as the foundation for their analyses. The steps below use a Logic Liaison template pipeline as an example to demonstrate how to reformat code for Code Repositories. These same steps apply to any code migration from Code Workbooks to Code Repositories, regardless of whether you used Logic Liaison templates.This particular example uses Python code, however, the steps for using Code Repositories will remain the same whether you are working in R, Python, or SQL. The only difference will be the environment type that you select when opening the application.

This tutorial follows an example of a legacy Logic Liaison template to create the following output tables:

- all_patients_fact_day_table_de_id
- all_patients_summary_fact_table_de_id
- Site_Data_Table
- Missingness_Indicators

To view this sample repository in Foundry, click on the link.



Data lineage helps guide you through the process visually. Note: In your Code Repository, a data lineage will only be visible after you build each dataset.

Step 4A: Import Datasets

When you submit your DUR, LDS (Limited Data Set) and De-identified datasets version 197 (the last release of COVID data) is automatically made available in your DUR. If you need to work with another version, we are able to make v192 available. V192 is the pre-DTA extension data which was released in May 2025.

This example uses the following De-identified tables:

- condition_occurrence
- drug_exposure
- device_exposure
- drug_era
- measurement
- observation
- observation_period
- person
- procedure_occurrence
- visit_occurrence
- death

Confirm you can access these tables in your active DUR.

Step 4B: Create Fusion Sheets or use Pipeline Builder (if needed)

Fusion sheets allow you to define concept sets and other configuration data that your code will reference. For this example, we will create three Fusion sheets:

- manifest_fusion
- LL_concept_sets_fusion_everyone
- LL_DO_NOT_DELETE_REQUIRED_concept_sets_all

To create Fusion sheets for concept sets:

1. Open a new Fusion sheet in your workspace
2. Input your concept sets and column headers
3. Convert the Fusion sheet into a table
4. Verify the table name matches the inputs in your Python repository code

Once your Fusion sheets are created and converted to tables, they will be available as input datasets for your Code Repository transformations. Pipeline builder could also be used in this step.

NOTE: **Pipeline Builder** provides much of the same power of spark-backed scalable data transformation in a low-code interface. Researchers can still accomplish complex pipelines at a high scale of data with this tool. **Fusion Spreadsheets** could be swapped for manually entering data in Pipeline builder which can then similarly be instantiated as a dataset.

Step 4C: Organize Your Code Repository

Copy your code into repositories. Builds must be executed in order; therefore, we recommend organizing them into directories based on each stage of execution. In this example, the initial three datasets are built from input datasets and fusion sheets. Subsequent intermediate datasets are dependent on these. It is recommended to run a Preview on each step of your code to check transform logic. Once the preview accurately populates data, click Build. If a build fails but the preview works, this points to an environment issue rather than a coding issue.

Folder: Initial Builds (Numbers correspond to image above)

1. manifest
2. customize_concept_sets
3. everyone_cohort_de_id

Next, add your code for each of the domain tables below. We recommend putting them all in a folder to keep intermediate datasets from the same step together.

Folder: Domain Tables (Number 4 in image above)

- everyone_conditions_of_interest
- everyone_devices_of_interest
- everyone_drugs_of_interest
- everyone_measurements_of_interest
- everyone_observations_of_interest
- everyone_procedures_of_interest
- everyone_patient_deaths

Next, create a folder to input code for the final four tables and click Build. Check that all tables show up as expected in the data lineage.

Folder: Final Tables (Numbers correspond to image above)

6. All_patients_fact_day_table_de_id
7. all_patients_summary_fact_table_de_id
8. Site_Data_Table

9. Missingness_Indicators

Step 4D: Moving from Code Repositories to Code Workspaces

Once your major transforms are complete in Code Repositories, you may want to switch to Code Workspaces to continue your analysis.

NOTE: Large Data Best Practices - given the scale of some N3C tables we recommend using Polars with the lazy load option in Code workspaces. This will minimize the memory pressure on the containers and often speed up exploratory code. Researchers may configure the compute resources available in Settings -> Compute Resources.

Steps to transition:

- Open Code Workspaces
- Select JupyterLab, RStudio, or VS Code
- Go to the Data tab and click Add > Read data
- Select the final datasets you created in Code Repositories
- Define an alias (a short reference name for your code)
- Choose your read strategy (pandas DataFrame, Polars LazyFrame, raw file, etc.)
- Code Workspaces generates a code snippet ready to use

A sample template with step by step instructions for moving data from Code Repositories into Jupyter Notebook, for further analysis in R or Python can be found here. Similar steps can be used to transition to working in VS Code or RStudio. A secondary example was created with synthetic Synpuf Data. This repository and example using a logic liaison template can be found here.

NOTE: By default, the alias will match the dataset name. Foundry stores the alias mapping in a hidden file under: /home/user/repo/.foundry

---

PySpark Setup (if needed): Install pyspark and openjdk (versions 8.x, 11.x, or 17.x only) from the Libraries tab. Unsupported OpenJDK versions cause errors.

Collaboration: Technical Considerations

Each user has an isolated environment. Share code via **Sync Changes** and work on independent branches when multiple users collaborate.

Package Restrictions: Python packages folium and pandasgui are blocked for security reasons.

Learn More: Review Palantir's documentation for data connectivity and integration details.

Common Migration Issues: Code running in Code Workbook but not in Code Repositories

Sometimes, moving code from a code workbook to a code repository will not work without modifying the code to run in a code repository.

To troubleshoot, perform the following steps:

- Have you verified your transform decorator is correct? That is, @transform vs. @transform_df?
- Have you verified your inputs are all declared and passed as inputs to your compute function?
- Are the names of your dataframes the same as your workbook? Is your code repository actually returning the dataframe?
- Check that the libraries being used in the code workbook code are also available in the code repository.
- Verify the inputs and branches to the code workbook cell are in fact the same datasets used in the code repository.

Review Palantir's FAQ on builds and checks errors for more detail.

Step 5: Publication Intent and Data Download Requests

The publication intent and data download request processes remain unchanged. The same workflows, timelines, and requirements apply.

EXPEDITED PUBLICATION REVIEWS
Have urgent manuscript deadlines? Use the Google Form for expedited review during the platform transition. The Committee reviews against the former DUR list, meaning no DUR reapproval needed first. Review takes 1–2 weeks.

For publication inquiries, contact: n3c.pubs@gmail.com

What You Need to Know:

- Submit publication intent before any external dissemination.
- All data exports require privacy review and approval.
- Results must comply with N3C disclosure requirements (e.g. no cell sizes <20, etc.).
- Review processes typically take 3–7 business days.

Publication Intent Process:

- Register manuscripts, abstracts, or presentations using N3C data.
- Include the working title, abstract, author list, and target venue.
- Ensure proper attribution and coordinate research across N3C.
- Verify compliance with N3C dissemination policies.

Download Request Process:

- Submit requests to export aggregate results, tables, or figures.
- The privacy review verifies compliance with disclosure requirements (no small cells, no PHI).
- Approved files available through secure download portal.
- Cannot export row-level patient data or identifiable information.

Important Reminders:

- Submit early to allow time for review before deadlines.
- Double-check all cell sizes before requesting downloads.
- Maintain clear documentation of what you're exporting and why.
- Follow up if you don't receive a response within the expected timelines.

Help & Support

N3C provides multiple channels for assistance as you navigate the reopened N3C platform and resume your research.

Contact N3C:

- **Email:** ncats_n3c@mail.nih.gov for all questions and support needs

- **Office Hours:** To support you during the reopening, we are offering special Getting Restarted Office Hours on Tuesday, March 3 and Thursday, March 5 at 1-2pm EST. Join using the link, here!
- **Platform Support:** Submit platform support tickets through the Foundry support tickets, here!

When Requesting Help:

Include the following information to have your request be addressed as quickly as possible:

- Your project name and DUR number
- Detailed description of the issue, question, and error log
- Screenshots if applicable
- Steps to reproduce the problem (for technical issues)

Stay Informed:

- Monitor N3C community forums for announcements and updates
- Office hour schedules will be sent via email
- Participate in training sessions and webinars
- Review platform documentation and FAQs

Frequently Asked Questions (FAQs)

Access & Registration

Q: Do I need to create a new account?

A: No, your existing N3C credentials should work. If you have issues, contact N3C Support Desk: here.

Q: Do I need to resubmit my DUR?

A: Yes, you need to resubmit your DUR for the reopened N3C platform. However, returning users with previously approved projects qualify for rapid review. This is a required step to help ensure the next Phase of N3C is sustainable; it is not a re-review of your science.

Q: Can I submit my DUR before my institution's DTA is finalized?

A: Yes! Returning users can submit new DURs and access pre-DTA expiration data releases while their institution completes the DTA process. You don't have to wait—submit your DUR and proceed in parallel.

Q: Is my institution's DTA still valid?

A: No, all institutions will need to execute new DTAs and DUAs which address the new CADR (Controlled Access Data Repository) requirements. The new DTAs are in the process of being released and if your organization is jumping to get back involved with N3C please let us know at ncats_n3c@mail.nih.gov. We are excited to engage with any Institution who wants to contribute data again. You do not need to terminate previous DTAs, but new ones must be executed before data can be sent again.

Q: Can government agencies access N3C data through alternative pathways?

A: No. All users, including other government agencies, HHS, other HHS organizations, and academic researchers, must go through the same DUR submission and Data Access Committee (DAC) review process. There are no back doors or special access pathways. All approved DURs will have their abstracts posted publicly for transparency.

Q: Can I swap out the PI on an existing DUR when I resubmit?

A: It depends on a case-by-case basis. Due to changes in which institutions signed DTA extensions, data access permissions may have changed. Contact ncats_n3c@mail.nih.gov with your specific situation, and the team will work through your scenario.

Q: If my old projects are no longer relevant, do I need to close them out?

A: No. Old projects will automatically expire. You don't need to take any action to close them out.

Tools & Code

Q: Where are my Code Workbooks?

A: Code Workbooks are no longer available. Your code was exported, and you'll need to adapt it to run in Code Repositories and/or Code Workspaces (JupyterLab, VS Code, or RStudio).

Q: How do I get my exported code?

A: Please submit an Enclave Code Export Request here. The N3C team has access to code from 1,022 of 1,034 projects (>100,000 code assets) and can provide access to your archived code.

Q: Will my old code run without changes?

A: No, all code will need modifications to run correctly in Code Workspaces. The N3C team is developing template scripts and documentation to help with this adaptation process.

Q: What about PySpark and Spark SQL?

A: Both PySpark and Spark SQL remain available and can be used from within Code Repositories. Code Workbooks are the only application that is no longer available.

Q: What happened to Logic Liaison templates?

A: Logic Liaison templates are no longer available due to the loss of Code Workbooks. N3C is developing an RWE Research Template Library as a replacement (coming soon).

Q: Will there be a cohort-building component in the future?

A: Yes! A cohort builder is planned as part of Dynamic Workspaces. The N3C team is building a series of applications that align to the Dynamic workspace model and the more sustainable contract terms, and cohort building will be one of those components.

Data

Q: Is the same data available?

A: The core N3C datasets are available in the reopened N3C platform, harmonized and cleaned. However, only two releases are available (Releases 192 and 197)—over 190 previous releases are no longer accessible.

Q: What data releases are currently available?

A: Two releases are available: Release 192 (pre-DTA expiration, for returning investigators) and Release 197 (last release). Select the release that matches your approved DUR timeframe.

Q: What datasets can I select for my research?

A: When submitting your DUR, you can select from EHR data, mortality data, claims data, and other linked datasets based on your research needs and approved data tier.

Q: Is N3C still accepting COVID data submissions?

A: No. N3C is no longer accepting new COVID data submissions from contributors. However, existing COVID data remains available for research, and you can continue manuscripts based on COVID data.

Q: Will data uploads remain annual in the future?

A: Under the new dynamic workspaces model with disease-agnostic data submissions, N3C is hoping to have more frequent data updates than annual. The specific frequency is still being determined and will be discussed as part of the new DTA agreements.

Publications

Q: How do I submit a publication for review?

A: To submit a publication for review, complete the Publication Intent Form. The N3C Publications Committee reviews all manuscripts to ensure proper N3C data usage, attribution, and compliance with data use agreements. Publication review typically takes 1-2 weeks depending on revisions requested.

Q: Do I need to submit my manuscript for publication review?

A: Yes. All manuscripts, abstracts, posters, and presentations using N3C data must be submitted to the N3C Publications Committee for review before submission to journals or conferences. This ensures proper acknowledgment and compliance with N3C policies.

Q: Can I submit my publication for review while my DUR resubmission is pending?

A: Yes, if the DUR for that project was previously approved and you have submitted it for publication! During the platform transition, the Publications Committee has established an interim review process using a Google Form. The Committee will review against your former DUR list to confirm your approved project. You do not need to complete DUR reapproval before publication review.

Q: What information do I need to include in my publication submission?

A: Your publication submission should include: manuscript title, authors list, abstract, target journal or conference, estimated submission date, and your DUR/project ID. If continuing previous work, reference your prior project ID.

Q: How long does publication review take?

A: Publication review typically takes 1-2 weeks, depending on whether revisions are requested and how quickly authors respond. For urgent deadlines, contact the Publications Committee to discuss expedited review.

Q: Can I access legacy features to finish my publication?

A: Legacy features are available for a limited time for those who submitted publications to the Publication Committee before March 2, 2026, and need to complete reviewer-requested analyses or finalize dissertation work. If you believe you qualify and have not been notified, contact ncats_n3c@nih.gov.

Dynamic Workspaces

Q: What are Dynamic Workspaces?

A: Dynamic Workspaces is N3C's new approach that provisions research environments based on your computable phenotype rather than disease-specific silos (COVID, Renal, Cancer). Your workspace will be dynamically created to match your approved patient population.

Q: When will Dynamic Workspaces be available?

A: Dynamic Workspaces are planned for a future release. Currently, you'll work in standard Code Workspaces, but this new model is coming to make research more flexible and phenotype-driven.

Q: How does this differ from the old domain-based approach?

A: Previously, workspaces were organized by disease domains which each had a rigid patient phenotype pre-selected for users of those domains (e.g. COVID-19, Renal, Cancer, etc.). Dynamic Workspaces will instead be created based on your specific research phenotype, allowing more flexibility across disease areas, better control populations, and expanded research questions.

Q: How will the Renal and Cancer pilot enclaves work in the new structure?

A: The Renal and Cancer pilots will transition to the dynamic workspaces model. Unfortunately, institutions that participated in these pilots will need to execute new DTAs under the new governance framework. However, the new model is simpler, more efficient, and addresses feedback about defining patient populations and phenotypes. The change should ultimately make it easier to contribute and leverage data for cancer and renal research.

Data Transfer Agreements (DTA)

Q: Does my institution need a new DTA?

A: Yes. All institutions will need to execute new DTAs with NCATS. The new DTAs contain core terms and conditions from original agreements but include important updates for auto-renewal (no extensions needed), site selectable governance level, and greater transparency through the DAME system.

Q: Can I access data while my institution's DTA is pending?

A: Yes, as a returning user, you can submit your DUR and access pre-DTA expiration data releases while your institution completes the DTA process. Fundamentally, Organizations contributing data (or not) is not related/linked to researcher access to data for any Institution.

Q: When will CTSAs and institutions receive the new DTAs?

A: NIH Office of General Counsel (OGC) review is complete. The new DTAs have been shared with Renal and Cancer pilot institutions. Distribution to all other institutions will follow. You do not need to terminate previous DTAs, but new ones must be executed.

Q: What is DAME?

A: DAME (Data Access Management Environment) is N3C's new system for managing Data Access Requests and institutional agreements. It provides improved workflow transparency and eliminates the need for DTA extensions through automatic renewals.

Coming Soon

The following features are planned for future releases:

- **DAME (Data Access Management Environment)** - Streamlined DUR management with transparency
- **Dynamic Workspaces** - Domain agnostic phenotype-driven research environments
- **Cohort Builder** - Tool for building research cohorts
- **RWE Research Template Library** - Replacement for Logic Liaison templates
- **More frequent data updates** - Moving beyond annual data submissions
- **Linkage to CMS claims data** - OMOPified and harmonized CMS data for use by researchers
- **Linkage to diverse geospatial and environmental datasets** - A variety of geospatial and environmental datasets will be linked at the zip/hex level for researchers to seamlessly and easily leverage in their studies.

Appendix A

Code Workspaces

Code Workspaces brings the JupyterLab®, RStudio® Workbench, and VS Code third-party IDEs into Palantir Foundry, enabling users to boost productivity and accelerate their data science and statistics workflows by using their preferred tools on the high-quality data available in N3C while leveraging the Foundry Ontology. Code Workspaces containers are natively integrated with the rest of the Foundry ecosystem, combining familiar IDEs with the all benefits of the Foundry platform, including: data security, branching, build scheduling, and resource management.

Code Workspaces gives platform administrators an easily-deployed, fully-managed, secure, and production-ready way to provide JupyterLab®, RStudio® Workbench, and VS Code to users, with Foundry's data governance and compliance with FedRAMP, GxP, and other standards built in. With Code Workspaces, users can securely connect to existing internal systems and build analyses, transforms, models, applications, or entire workflows on data with Foundry's access controls and data permissioning.

Key features

Key features of Code Workspaces include:

- Security: Code Workspaces is built on Foundry's core security components, which underpin the platform as a whole, including robust permissions and granular access controls. This provides Foundry's security model to the third-party IDEs available in Code Workspaces. For example, restricting access to a dataset in Foundry will also restrict it in Code Workspaces IDEs, ensuring consistent permissions across tools.
- Customizable environments: Code Workspaces allows users to define custom environment profiles and adjust the compute resources of their workspace as desired.
- Git workflow support: Code Workspaces are backed by the Code Repositories infrastructure, which provides industry-standard version control features such as branching, merging, and commit history. These features enable multiple users to operate in the same workspace more easily and safely.
- Applications: Code Workspaces currently supports Dash ↗ and Streamlit ↗ for Python applications and Shiny® ↗ for R applications. Users can create application workflows directly in Code Workspaces, with Foundry's built-in version control, branching, and data governance features.
- Model integration: Users can create model assets within a Code Workspace and track them against modeling objectives. You can create multiple models from the same workspace.

- Transforms/build integration: Code Workspaces serves as a development environment for transforms. Logic written in Code Workspaces can be published as data transformation pipelines and seamlessly integrates with Foundry's data integration toolkit, including builds, schedules, data lineage, and health checks. Code Workspaces supports both R transforms and Python/Jupyter® transforms.

When to use Code Workspaces

Foundry has a variety of applications you can use for analytical or coding purposes. For example, if you are an analyst, you may be best served by Contour, Foundry's point-and-click low-code interface for dataset analysis.

If you need to write large-scale data pipelines, set up data connections, or work with streaming data, other Foundry tools have more functionality than Code Workspaces; for these use cases, we recommend using Pipeline Builder, Data Connection, and Foundry Streaming, respectively.

Specifically, Code Workspaces run on a single node, while other Foundry applications leverage Spark infrastructure. Thus, we recommend that users performing large-scale data transformations use Pipeline Builder or Code Repositories rather than Code Workspaces.

Code Workspaces is geared for building machine learning models or those familiar with working in JupyterLab® or RStudio® Workbench.

For further information, please review Code Workspaces documentation: Link

Code Repositories

Code Repositories provides a web-based integrated development environment (IDE) for writing and collaborating on production-ready code in Foundry. The application provides a user-friendly way to interact with the underlying Git repository, and provides a range of additional features:

- All common Git version control tasks, including branching, committing, and tagging releases can be performed through the web UI
- Repositories have integrated support for code review and collaboration through *pull requests*, including support for highly configurable permissions to ensure the codebase is high-quality.
- Every repository type includes integrated features to aid with the code authoring experience, including IntelliSense, code linting and error checking, and rich help dialogs.

Repository types

Code Repositories support creating repositories of many types. The most common repository types are described below.

- **Transforms** repositories support authoring data transformation logic and include features to enable previewing and debugging transforms. Supported languages include Python, Java, and SQL.
- **Functions** repositories enable writing business logic that can be executed with low latency in an operational context, and include native support for accessing data from the Foundry Ontology. The Code Repositories environment supports autocomplete based on Ontology data types, and enables code authors to preview Functions while authoring them. Functions can be written in TypeScript or Python.
- **Model development** is supported in Code Repositories. Learn more about how to develop models in Code Repositories.

For further information, please review Palantir's Code Repositories documentation: Link