

## Data Retention Policy at TSL, FSU College of Medicine (Jan 05, 2026)

### Purpose of the Data Retention Policy

As a core research service facility, the Translational Science Laboratory (TSL) at the FSU College of Medicine generates and delivers high-throughput data—particularly genomic sequence data and proteomics data—on behalf of investigators. While primary ownership and long-term stewardship of the data rest with the users, TSL retains a copy of the data for a defined period following release for the following purposes:

1. Facilitate User Access and Redelivery

To provide a buffer period during which users can re-download, verify, or retrieve lost data without regenerating it. This is especially valuable during downstream analysis, publication preparation, or data transfer failures.

2. Support Quality Control and Technical Assistance

To enable internal review, troubleshooting, or reanalysis in case users report issues related to sequencing quality, sample tracking, or data format inconsistencies. Retaining a verified copy allows TSL to provide high-quality, responsive technical support.

3. Ensure Temporary Backup for Critical Research Data

To act as a short-term safeguard against unintentional loss of user data during the handoff period. This temporary custodianship ensures that researchers have time to confirm successful receipt and local storage of their data.

4. Optimize Storage and Infrastructure Usage

To manage finite storage capacity by retaining service-generated data only for a defined and justifiable period. This policy allows TSL to maintain performance, reduce unnecessary storage costs, and continue serving the research community efficiently.

### Policy Statement

TSL retains service-generated data temporarily after releasing to users. TSL does not serve as a long-term data repository. Users are responsible for downloading and archiving their data. After the specified retention period, TSL will permanently delete the data.

### Data Retention Schedule

Data Type	Examples	Retention Period	Purpose
Sequencing BCL Data	Raw BCL dataset directly from the sequencer	3 months after data releasing	Allows time for re-demultiplexing, in case of modifications in sample sheets
Raw Genomic Data	FASTQ files (e.g., WGS, RNA-Seq)	12 months after data (<5TB in size) releasing 3 months after data (>5TB in size) releasing	Allows time for users to verify and download data, and for TSL to respond to quality concerns or delivery issues.
Processed Genomic Data	BAM/CRAM, VCF, count matrices	12 months after data (<5TB in size) releasing 3 months after data (>5TB in size) releasing	Supports downstream troubleshooting, publication needs, or re-analysis requests.
Raw Proteomics Data	Mass spectrometer vendor files	12 months after data releasing	Accommodates data transfer delays and enables basic troubleshooting or re-export if needed.

Processed Proteomics Data	Identified peptides/proteins, quantification tables	12 months after data releasing	Enables users to access results again for interpretation, reprocessing, or publication support.
QC Reports and Service Metadata	FastQC, alignment logs, acquisition metadata, run summaries	12 months after data releasing	Useful for TSL's internal QA tracking and responding to user inquiries.

### **User Responsibilities**

1. Download and archive data promptly upon releasing
  2. Confirm data integrity after receipt
  3. Store data in secure, institutionally compliant locations
  4. Request extended retention before standard period expires, if needed
- Users are strongly advised to download and archive all service-generated data promptly upon release notification.

### **Data Access During Retention**

Users may request data release or support within the defined retention period by contacting TSL. TSL will make reasonable efforts to restore data within this window. After these retention periods, data will be permanently deleted from TSL systems. Notices may be sent to the users before the data purging.

### **Data Disposal**

At the end of the retention period, data will be permanently deleted from TSL-managed storage systems using secure, irreversible methods. Deleted data cannot be recovered.

### **Exceptions and Extensions**

Extensions may be granted for:

1. Multi-institutional or grant-funded collaborations
2. Active publication or peer review processes
3. Legal holds or IRB-directed retention requirements

Requests must be submitted in writing prior to the expiration of the standard retention period.