

Leveraging a Compliant Data Repository to Significantly Improve Biostatistics Workflow



Certara Integral is a scalable cloud-based software as a service (SaaS), validated data and analysis repository with an intuitive user interface for collecting, managing, and storing multiple types of data sets for analysis, sharing, and reporting of all types of clinical data. This regulatory compliant repository significantly improves the biostatistics workflow, providing analyses that are traceable, reproducible, and secure.

The Certara Integral data and analysis repository is:

- Scalable
- Cloud-based
- Software as a service (SaaS)
- Validated

Certara Integral is designed to be a 21 CFR Part 11 compliance bolt-on solution to an existing statistical computing environment within biopharmaceutical organizations. It is compatible with both cloud-based and on premise computing environments, with all modern operating systems, and is agnostic to the analytical software being used (SAS, R, Python, etc.). Integral is built on Amazon Web Services (AWS), uses OKTA for authentication, and has an Oracle database.

SECURITY

Roles and Permission groups govern access in Integral and are assigned by Customer Administrators. Permission Groups are configurable and are designed to facilitate easily assigning permissions broadly across studies or projects based on a user's function. By attaching simple to build Permission Rules and associating them with these Permission Groups, an Administrator could instantly assign its users en-masse to hundreds of projects whilst tightly restricting content according to business rationale.

In the same way that access can be granted quickly, a user's access can likewise be promptly and completely removed. Additionally, project level and folder based permission assignment or removal is also available to enable granular access control. Integral Security Group permissions include the ability to Blind and Unblind repository content.

STORAGE ENVIRONMENT/FILE SYSTEM

The folder structure hierarchy and naming conventions in Integral are completely configurable according to internal programming best practices or standard operating procedures. Generation of complete folder structure hierarchy for projects can be automated to enable standardization across users.

Integral provides configurable metadata that can be associated with folder types so that complex folder hierarchies are not required to organize content. These metadata can be used to filter the content of the database to facilitate searching. Integral also leverages elastic search to enable keyword searching of all readable file content within the repository.

Integral enables the viewing of the most common file types directly through the application interface. It additionally allows for a side-by-side comparison of script files to highlight changes between versions.

ARCHIVING

Integral is a GxP compliant archive. Archiving is accomplished in the same environment in which active studies are worked on, simply by locking a project, folder, or file. Once content is locked, it cannot be changed in any way yet still remains accessible for viewing or copying. Only a user with the Archivist Role has the ability to unlock content in Integral. Locking and unlocking of content becomes a permanent part of the audit trail, indicating the user, the date and time, as well as the reason.

SYSTEM VALIDATION

Integral is validated by Certara prior to deployment of each release. Integral validation is governed by Certara's SOP driven software development lifecycle. The Integral validation summary report is published to the Integral help website.

PROGRAMMING LIFE CYCLE/VERSION CONTROL

A statistical analysis includes input datasets from various sources, mature reusable code, bespoke programming, intermediate results, and final output. It is the collection of these various assets that represent an analysis that is attributable, traceable, and reproducible. For this reason, Integral is designed to version entire statistical analyses and not just versions of files. Analyses are intentionally committed by the user to the repository as Save Points at a point in time of the user's choosing. Each Save Point can be viewed independently, detailing all files associated with the analysis and all dependencies. It is possible to roll-back to a previous version of a Save Point to make it the most recent version again.

QC/REVIEW PROCESS

For companies that utilize parallel programming as part of their quality system, Integral's access control can be leverage to ensure independence. Code review can be accomplished directly in the interface via the viewing panes, including the file comparison view pictured above.

Integral can be used as a global code repository to help manage mature reusable code. Users can be sure they are using the most recent version of a program, can be automatically notified if the source program has changed since they started using it, and can easily see the provenance of the code (who updated it, when it was updated, what was updated, and why it was updated).

AUDIT TRAIL/TRACEABILITY

Audit Trails within Integral are created on-demand for files saved within the repository. Audit trails can be generated for a single file or at the folder level for the creation of all files within a folder structure. Audit trails show the who, when, why, and what for content history. Deltas are performed for all text file types (including .csv, .xpt, etc.) using various mechanisms to highlight differences, with

custom selection of versions to compare. All previous versions of files are retained within the database and are retrievable. The same technology (OKTA) used for authentication is used to apply electronic signatures to create, modify, and delete actions performed by the user.

DEPENDENCIES

A critical component of regulated drug development is ensuring reported analyses are based on accurate study observations. Data are frequently updated, which can result in statistical analyses becoming out of date. Integral uses a database to track relationships between input datasets and statistical analyses to ensure analyses are current. Dependencies are easily set at the time analyses are committed to the repository. When the dataset upon which an analysis depends is updated the user will receive a notification and the analysis will become visibly out of date (by color code).

Statisticians will often need to use materials outside of the project being worked on in order to complete an analysis, including datasets, programs, etc. Making copies of files is a common way to introduce untraceable errors into an analysis. Integral allows the user to make “linked copies” of files that can be auto updated when the original file is changed.

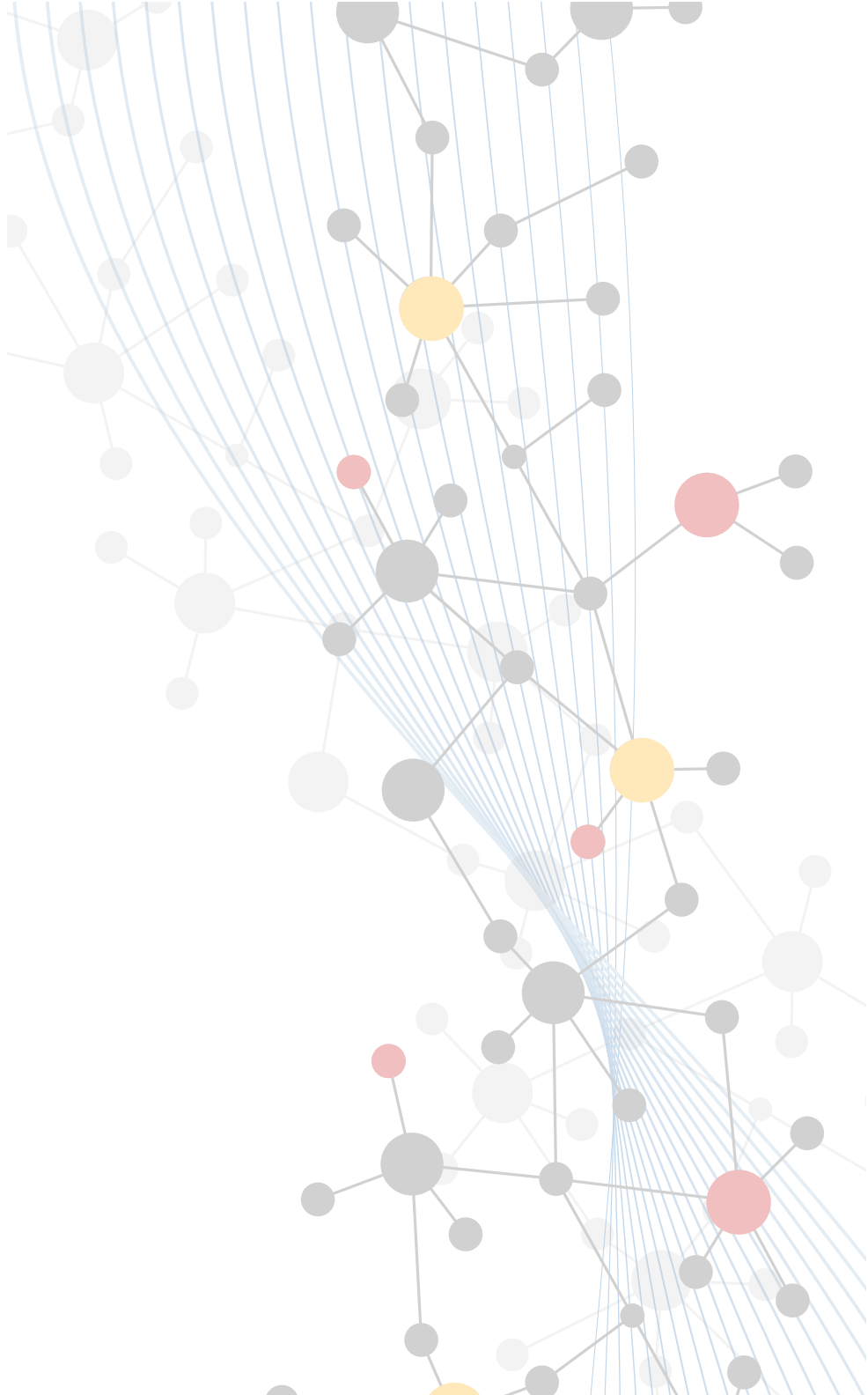
REPRODUCIBILITY

Integral ensures analysis reproducibility by saving a complete analysis as a Save Point in the repository. This methodology is analogous to the way source code control systems work in software development. An analysis inputs, programs, and outputs are linked together in a single version, irrespective of the various file versions. Save Points cannot be corrupted by updating single files within the Save Point without updating the entire Save Point. Integral makes it possible to revert to previous Save Points at specific points in time to reproduce each set of results provided that the analytical software used is still available.

IMPORTING/OUTSOURCING AND COLLABORATION

Integral enables multiple file uploads into a single folder or the upload of entire folder structures via the Integral Client Application. Integral is a web application and may be used as a secure data transfer mechanism between departments or companies, eliminating the need for sFTP sites and email data exchanges. Exchanging data directly in the repository results in a closed chain of custody and enhances traceability as well as audit readiness. Robust access controls ensure users see only relevant information for their function.

Because Integral is agnostic to operating system and analytical software and requires no VPN access, collaborating with external partners and resources is easy to set up and take down through the normal administration of the system.



About Certara

Certara is a leading provider of decision support technology and consulting services for optimizing drug development and improving health outcomes. Certara's solutions, which span the drug development and patient care lifecycle, help increase the probability of regulatory and commercial success by using the most scientifically advanced modeling and simulation technologies and regulatory strategies. Its clients include hundreds of global biopharmaceutical companies, leading academic institutions and key regulatory agencies.

For more information visit www.certara.com or email sales@certara.com.