



BACKGROUND

Reproducibility of studies is essential for validation of scientific results. Studies involve raw data collected from instruments and software applications used to analyse data.

Authenticity of the study results depends on the **integrity of data** and **trustworthiness of software installations** used for data analysis.

Long-term reproducibility of studies requires long-term management of data and long-term maintenance of software installations.



Data Integrity

Data Integrity is the extent to which data are complete, consistent and accurate throughout the data lifecycle.

- **Complete** – All the required data needs to be collected and stored
- **Secure** – Data must not be destroyed
- **Unaltered** – Data must not be changed
- **Confidential** – Data must not be disclosed to unauthorised individuals
- **Usable** – Study metadata needs to provide context for experts reproducing study results.

Software Integrity

Software Integrity is the extent to which a software installation is functional, reliable, and usable throughout its lifecycle from operational use (i.e., applied in new studies) to study reproducibility (i.e., applied to archived data).

- **Functional** – Software installation must stay operational
- **Unaltered** – Software must not be changed
- **Secure** – Software installation must not present risks from cyberattacks and confidentiality breaches
- **Accessible** – Software installation needs to be easily accessed for repeated use.

OECD
Organisation for Economic Co-operation and Development

No. 17 Application of GLP Principles to Computerised Systems long term readability of raw data (section 3.2 point 75)

- ❑ The **maintenance of the raw data** associated with a specific study and the specimens generated from that study are **the only means that can be used to reconstruct the study.**
- ❑ **Software should be retained** in the archive if necessary to read or reconstruct data.

PROBLEM

System decommissioning process removes the instruments from the active use in the operations of an organization.

Software associated with the instrument is also decommissioned and removed from active use. Software is, however, required for data readability.

Currently, there is no systematic and principled approach to maintaining decommissioned software for study reproducibility.

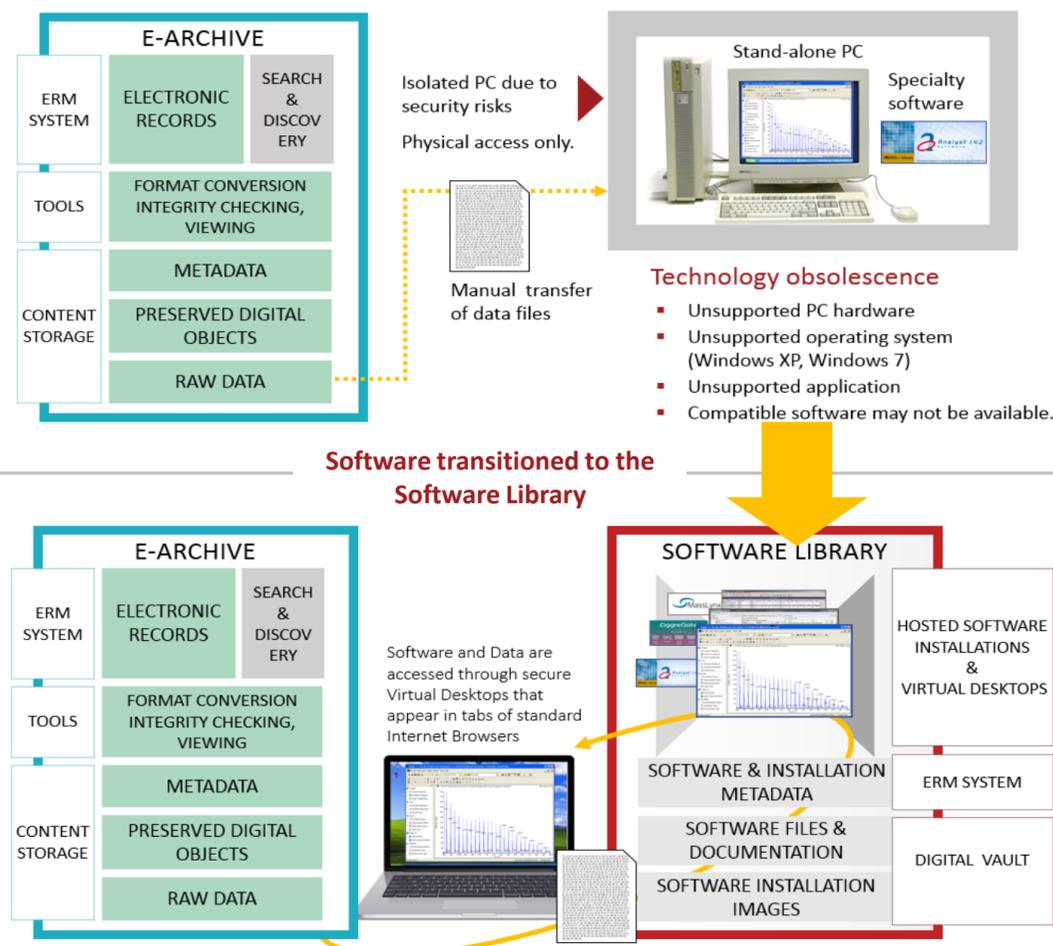
METHOD

Software transitioning is a process of creating and validating software installations for long-term use in virtualized computing environments.

Through a systematic validation process, the installations are approved for study reproducibility and long-term data readability.

Software Library (SL) refers to a platform that hosts and provides access to the installed software through virtual desktops. Software Library complements Data Archives.

Software Integrity methods in the Software Library are aligned with Data Integrity methods in e-Archives.



CASE STUDY

Analyst 1.4.2 software by SciEX has been used by a pharmaceutical organization since 2006.

The instrument set-up and the software installation followed the best practices and produced documentation

- **Installation qualification (IQ)**
- **Operational qualification (OQ)**
- **Performance qualification (PQ)**
- **Re-qualification** after the initial IQ, OQ, PQ and in accordance with a user's Standard Operating Procedure (SOP) requirements.

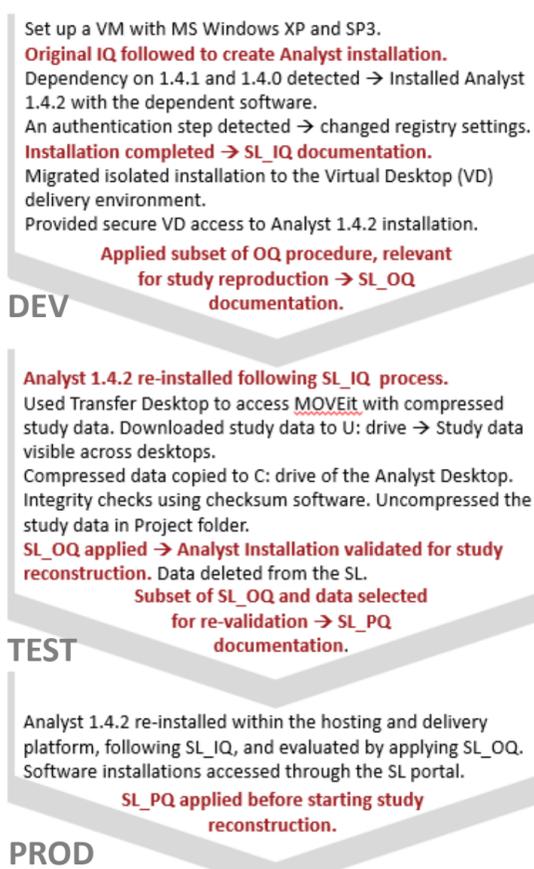
The instrument has been decommissioned and study data stored in e-Archive. The software installation on an isolated PC with Windows XP has been used to reproduce studies.

A Software Library instance was set up to transition Analyst 1.4.2 software, i.e., to create and validate software installation using IQ, OQ and PQ documentation.

The Software Library components include:

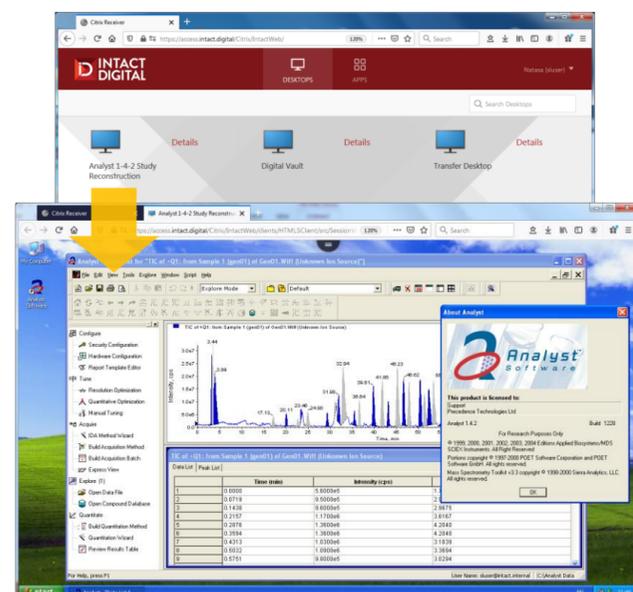
- **Digital Vault** – Stores software executable and installation files. Not accessible to end users
- **Transfer Desktop** – Enables secure transfer of data into the Software Library using MOVEit platform generally used by the pharmaceutical organization for data transfers.

Installation and validation process



Study reconstruction

Testing and PQ was performed using a copy of archived data: 1GB compressed; 9GB uncompressed. Transfer from MOVEit to SL: 1 min. From Transfer Desktop to Analyst Desktop: 2 min. Checksum: < 20 sec. Decompression of 1GB: 10 min.



Example using data from a sample included with Analyst 1.4.2 installation.

