



INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

GxP Data Archiving: Dynamic Records & Readability Matrix

Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-10-27 · gxp data archiving · gxp archive requirements · archiving dynamic records · metadata retention · long-term readability · electronic records archiving · alcoa+ archiving · archive migration validation · audit trail retention · 21 cfr part 11 archiving

Original article: <https://intuitionlabs.ai/articles/gxp-data-archiving-decision-matrix>



In this article

1. [Executive Summary](#)
 2. [Introduction and Background](#)
 3. [Record Inventory, Classification and Dependencies](#)
 4. [Preservation Mode 1: Retain the Native System](#)
 5. [Preservation Mode 2: Controlled Viewer or Validated Render](#)
 6. [Preservation Mode 3: Migrate to a Dynamic Archive](#)
 7. [Preservation Mode 4: Certified Static Copy](#)
 8. [Feature Comparison](#)
 9. [Performance and Benchmarks](#)
 10. [Data Analysis and Evidence](#)
 11. [Access, Periodic Review and Decommissioning](#)
 12. [Implications and Future Directions](#)
 13. [Frequently Asked Questions \(FAQs\)](#)
 14. [Conclusion](#)
-

Executive Summary

GxP data archiving is best treated as a record-class decision, not a storage purchase. The question for each class of record is whether, after the source system is retired, the preserved representation is still authentic, complete, intelligible, retrievable and reconstructable. The UK Medicines and Healthcare products Regulatory Agency (MHRA) defines the data lifecycle as running through retention, archive/retrieval and destruction (assets.publishing.service.gov.uk), and both MHRA and the OECD (Organisation for Economic Co-operation and Development) state that backups kept for recovery do not replace long-term retention of data and metadata (assets.publishing.service.gov.uk) (www.oecd.org).

The regulator texts reviewed here converge on four expectations. Metadata are part of the record (assets.publishing.service.gov.uk). Electronic archives should maintain audit trails and metadata and keep records readable over the long term (www.gov.uk). Data that must be dynamic stay dynamic (picscheme.org), so a static render is not a safe default: the European Medicines Agency (EMA) says static formats of dynamic data are not adequate when a clinical database is decommissioned (www.ema.europa.eu). Migration to a new format must be validated and documented (www.gov.uk).

Retention periods differ by rule and by date, so this report uses none as a default. UK trial master file documents are kept at least 25 years (www.gov.uk) for applications on or after 28 April 2026, while earlier trials keep the previous five-year period for those documents (www.gov.uk). US batch records under 21 CFR 211.180 run at least one year after batch expiry (www.ecfr.gov), and investigational new drug sponsor records run two years after marketing application approval (www.ecfr.gov). A consultation on a revised EU Annex 11 closed on 7 October 2025 (health.ec.europa.eu), and the Volume 4 catalogue still listed Annex 11 as the January 2011 revision when checked on 27 October 2026 (health.ec.europa.eu). The draft is therefore not a requirement.

The report supplies four preservation modes (native retention, controlled viewer, migration to a dynamic archive, certified static copy), a ten-row record-class matrix, eight calculable verification measures and a restoration protocol. It does not rank products, set universal thresholds or give legal advice. Destruction needs documented quality and legal approval, and MHRA states that documents must not be destroyed before the end of the retention period (www.gov.uk).

Introduction and Background

This report is written for clinical, laboratory, manufacturing and quality-system owners, records managers, archivists and validation teams who must decide, for each legacy system, whether to keep it running, preserve a validated render, migrate dynamic data to an archive, or maintain a controlled viewer for the whole retention period. No vendor or product appears as an option in any comparison here; the analysis is a reading of regulator and standards texts, not a product evaluation.

Three terms are often conflated:

- **Backup** is a copy of current, editable data, metadata and system configuration kept for recovery (assets.publishing.service.gov.uk) (www.oecd.org).
- **Archive** is a designated secure area or facility, which can be a computerised system, for long-term retention (assets.publishing.service.gov.uk) (www.oecd.org), and it must be protected so records cannot be altered or deleted without detection (assets.publishing.service.gov.uk).
- **Long-term digital preservation** is a wider discipline. The Open Archival Information System (OAIS) reference model defines it as keeping information independently understandable by a Designated Community, with evidence supporting its authenticity (public.ccsds.org), and it defines "long term" not as a number of years but as a period long enough for changing technology to matter (public.ccsds.org).

Terminology is not uniform across regulators. The US Food and Drug Administration (FDA) uses "backup" in 21 CFR 211.68(b) to mean a true copy of the original record (www.fda.gov) and says that usage is consistent with "archive" (www.fda.gov). A project that borrows another jurisdiction's vocabulary without checking this can pass an inspection checklist on paper while lacking an archive in the MHRA or OECD sense.

The status of sources matters. Requirements here are those of a cited regulation or guidance. The revised Annex 11 and Chapter 4 texts, prepared with the Pharmaceutical Inspection Co-operation Scheme (PIC/S), are consultation drafts whose response period has ended (health.ec.europa.eu). The Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM) is described by its publisher as a format for exchanging and archiving clinical research data (www.cdisc.org), which is not the same as an archiving mandate. Matrix choices, test designs and owners are editorial implementation choices and are labelled as such. Governing domain, jurisdiction and predicate rule must be confirmed by the reader's legal and quality functions before any decision is acted on.

Record Inventory, Classification and Dependencies

Inventory and the predicate retention rule

An archive decision starts from an inventory. The EMA guideline on computerised systems in clinical trials calls for an inventory of all essential data and documents with their retention periods (www.ema.europa.eu), and the ICH E6(R3) guideline (International Council for Harmonisation, Good Clinical Practice) asks sponsors and investigators to record where essential records are located (database.ich.org). For GMP records, EU Chapter 4 requires regulated users to define which data are raw data (health.ec.europa.eu).

In the US, the predicate rule decides what must be kept. FDA's 2003 scope guidance says records not required to be retained under predicate rules are not within Part 11 merely because they are held electronically (www.fda.gov), and it suggests basing record-maintenance decisions on predicate rule requirements and a documented risk assessment (www.fda.gov). Part 11 defines an electronic record to include information that is archived (www.ecfr.gov), so archiving is within the definition, not outside it.

Static, dynamic and hybrid records

MHRA defines a static record as fixed, with little or no interaction with content (assets.publishing.service.gov.uk), and a dynamic record as one that allows an interactive relationship between user and content (assets.publishing.service.gov.uk). FDA's data integrity guidance gives the same distinction (www.fda.gov), and the OECD says information captured in a dynamic state should remain available in that state (www.oecd.org). FDA states that a fixed printout of a reprocessible record would not satisfy requirements to retain original records or true copies (www.fda.gov).

Hybrid records combine paper and electronic parts. MHRA requires references between the physical and electronic parts to be maintained for the whole retention period (assets.publishing.service.gov.uk), which makes the link itself a record that needs a preservation owner.

Dependency classes

A raw data file is rarely intelligible alone. The sources reviewed support the following dependency classes, which drive the matrix in the Feature Comparison section:

- **Metadata.** FDA defines metadata as the contextual information required to understand data (www.fda.gov) and expects the relationships between data and metadata to be preserved securely and traceably (www.fda.gov).
- **Audit trails.** An audit trail allows reconstruction of the course of events relating to the creation, modification or deletion of a record (www.fda.gov). Part 11 requires audit trail documentation to be retained at least as long as the subject records (www.ecfr.gov), and the EMA says data, contextual information and audit trail should not be separated in migration (www.ema.europa.eu).
- **Signatures.** Signed records must show the signer, date and time, and meaning (www.ecfr.gov), and signatures must be linked to their records so they cannot be excised (www.ecfr.gov). Annex 11 expects permanent linkage (health.ec.europa.eu).

- **Methods, configurations and validation evidence.** The OECD GLP Principles list validation documentation for computerised systems among archived materials (www.oecd.org). Where methods and code lists are not named in a source, their inclusion in this report is an editorial inference from the "complete information package" principle (www.oecd.org).
- **Attachments, links and relationships.** The EMA expects archived data to be retrievable as usable datasets with their metadata (www.ema.europa.eu).
- **Rendering dependencies.** The WHO notes electronic data are normally legible only through the original software application (cdn.who.int), and the Library of Congress lists external dependencies among its format sustainability factors (www.loc.gov).

Preservation Mode 1: Retain the Native System

Capabilities

The native system keeps the dynamic record, its metadata and audit trail in the environment that created them. MHRA allows data to be retained either in the original system, subject to suitable controls, or in an appropriate archive (assets.publishing.service.gov.uk). Archive arrangements must permit recovery and readability of data and metadata throughout the retention period (assets.publishing.service.gov.uk).

Adoption and regulatory position

No adoption data are available in the sources reviewed, and none are inferred. For legacy systems, MHRA says the ability to review data should be periodically verified (assets.publishing.service.gov.uk). For unsupported software, MHRA, the OECD and PIC/S all point to maintaining the software for data accessibility (assets.publishing.service.gov.uk) (www.oecd.org) (picscheme.org), with a virtual environment given as one route.

Strengths and limitations

The strength is fidelity: nothing is converted. The limitations are operational. The EMA expects contracts so that vendor documentation stays accessible even if the vendor stops supporting the system or ceases activity (www.ema.europa.eu). Annex 11 requires periodic evaluation to confirm a system remains valid (health.ec.europa.eu), so a retained system carries ongoing validation, security and infrastructure duties. PIC/S adds that after software installation or update, users should ensure existing and archived data can still be read, converting archived data where necessary (picscheme.org). The OECD requires records of a system that cannot be maintained to be archived under a documented strategy before decommissioning (www.oecd.org). Keeping a system is therefore a commitment to fund its lifecycle for the full retention period: patching, hardware refresh, access control and periodic review all continue. No source reviewed quantifies that cost, and no figures are offered here. Native retention suits record classes with high reprocessing value and a retention period short enough that vendor support is realistic; for longer periods it is best treated as a bridge to another mode.

Preservation Mode 2: Controlled Viewer or Validated Render

Capabilities

A controlled viewer is a validated tool that opens preserved data read-only, either through emulation of the original software or through a purpose-built reader. The EMA lists using a new system to emulate the old software as one way to keep archived data accessible (www.ema.europa.eu), and its computerised systems guideline says data should be archived in a read-only state (www.ema.europa.eu). The OECD says software should be retained in the archive if needed to read or reconstruct data (www.oecd.org).

Adoption and regulatory position

FDA states that true copies of dynamic records may be kept in the original format or one that preserves content and meaning, provided a suitable reader and copying equipment are readily available (www.fda.gov). For inspections, the MHRA says sponsors must provide the equipment and software needed to access electronic trial master file records (www.gov.uk).

Strengths and limitations

A viewer can lower the cost of keeping a full system alive, but it adds its own dependency. The Library of Congress notes that encryption is incompatible with format transparency (www.loc.gov), and a viewer that depends on an unsupported runtime merely relocates the technology-change problem. Validation scope is also open: no source reviewed prescribes viewer test content, so the render-equivalence sampling later in this report is an editorial design. A viewer used for dynamic reprocessing must demonstrate that reprocessing is possible, because a read-only display of a result is functionally a static view. Part 11 also requires the ability to generate accurate and complete copies in human readable and electronic form (www.ecfr.gov), so a viewer needs an export path as well as a display. Where an external party operates the viewer or archive, PIC/S expects the provider to be assessed and responsibilities recorded in a quality technical agreement (picscheme.org).

Preservation Mode 3: Migrate to a Dynamic Archive

Capabilities

Migration moves stored data from one durable storage location to another, possibly changing format but not content or meaning (assets.publishing.service.gov.uk). Annex 11 requires migration validation to check that data are not altered in value or meaning (health.ec.europa.eu). The EMA says validating a data migration should have no less focus than validating the system itself (www.ema.europa.eu), and ICH E6(R3) requires migration to be documented for traceability (database.ich.org).

Adoption and regulatory position

For clinical trials, the EMA expects a dated, certified copy of a decommissioned database to be archived (www.ema.europa.eu). Draft Annex 11 (not in force) would expect integrity checks, for example by checksum, before any data are deleted after a move (health.ec.europa.eu). OECD Advisory Document No. 17 says conversion to a different format, for example a proprietary format to PDF, should be treated as data migration

(www.oecd.org), and that electronic archiving is an independent procedure needing its own validation (www.oecd.org). CDISC ODM carries clinical data with metadata, administrative data, reference data and audit information (www.cdisc.org), which makes it a candidate exchange representation, not a mandated one.

Strengths and limitations

Migration can decouple records from vendor software. The OECD, however, warns that the difficulty of migrating data is often underestimated (www.oecd.org). MHRA accepts that preserving accessibility may require migration to a format that loses some attributes or dynamic functionality, and expects the choice to be risk assessed (assets.publishing.service.gov.uk). The Digital Preservation Coalition (DPC) cites an aerospace example where migrating older design files required extensive validation (dpconline.org). The EMA allows a certified copy in a different file format only when the equivalent dynamic nature, including metadata, is retained (www.ema.europa.eu). No public benchmarks of conversion fidelity for GxP formats were found.

Preservation Mode 4: Certified Static Copy

Capabilities

A certified copy is verified to contain the same information, including context, content and structure, as the original. MHRA's true-copy definition allows verification by dated signature or by a validated generation process (assets.publishing.service.gov.uk), and ICH E6(R3) adds that the verified information includes relevant metadata where applicable (database.ich.org). FDA's 2024 clinical investigations guidance says a copy retained in place of an original should be a certified copy that includes the date and time of copying (www.fda.gov), and that an original can be discarded after a certified copy is created (www.fda.gov).

Adoption and regulatory position

Static copies are well accepted for records that are static by nature. 21 CFR 211.180 allows records to be retained as originals or as true copies (www.ecfr.gov), and FDA states it does not object to archiving required records in standard electronic file formats or non-electronic media where predicate rule requirements are met (www.fda.gov). The limit is dynamic data. A copy may be in a different electronic format but must retain the metadata and audit trail needed to preserve full meaning (assets.publishing.service.gov.uk).

Strengths and limitations

PDF/A is a constrained form of PDF intended to be suitable for long-term preservation of page-oriented documents (www.loc.gov). The Library of Congress also says PDF/A does not permit encryption (www.loc.gov), that rendering support is not guaranteed (www.loc.gov), and that logical structure appears only if the creating process adds it (www.loc.gov). No source reviewed states that PDF/A preserves audit trails, signatures or reprocessing capability. A scanned or rendered image therefore preserves appearance, not behaviour. The EMA states that a conversion failing the certified-copy criteria cannot replace the original (www.ema.europa.eu).

Feature Comparison

Table 1 below compares the four modes on the dimensions the cited sources raise. The cells summarise points cited in the mode sections above; the final column is an editorial residual-risk judgement.

Mode	Dynamic behaviour	Metadata and audit trail	Technology dependency	Source support	Editorial residual risk
Native system retention	Preserved where the system is maintained (picscheme.org)	Remain in place; the system must be kept valid	Vendor and platform support; restoration to validated state	Retention in the original system is permitted with suitable controls	Support ends before retention does
Controlled viewer	Depends on emulation or reader; read-only access	Must be displayed and linked to the data	Viewer or emulator lifecycle	Reader must be readily available (www.fda.gov)	Display without reprocessing
Migration to dynamic archive	Must be shown retained, not assumed	Migrated together	Archive platform and schema	Validation no less rigorous than for the system	Silent loss of attributes
Certified static copy	Not retained for dynamic records	Only if included in the copy	Low (open, documented format)	Accepted as true copy when criteria are met (www.ecfr.gov)	Wrong choice for reprocessible data (www.fda.gov)

The table shows no mode dominating. The decision turns on whether later reprocessing, query or audit-trail interrogation is a realistic need for the record class, which is why a record-class matrix follows.

Table 2 is the reference asset: a record-class, dependency, preservation-mode, risk, test, retention-driver and owner matrix. Dependency citations show the source basis; the "candidate mode", "test" and "owner" columns are editorial implementation choices, not requirements, and must be adapted to the governing predicate rule.

Record class	Dependencies	Candidate mode (editorial)	Principal risk	Verification test (editorial)	Retention driver	Owner (editorial)
Raw data (dynamic)	Metadata, method, software to reprocess (picscheme.org)	Native retention or dynamic archive	Static copy cannot be reprocessed (www.fda.gov)	Reprocess a sample and compare results	Predicate rule or trial regulation	Data owner with QA
Raw data (static)	Identity and date metadata	Certified static copy	Appearance loss, format obsolescence	Fixity and render check	Predicate rule	Records manager
Metadata	Link to parent record (www.fda.gov)	Archived with parent	Orphaned context	Dependency-coverage count	Same as parent	Data owner

Record class	Dependencies	Candidate mode (editorial)	Principal risk	Verification test (editorial)	Retention driver	Owner (editorial)
Audit trails	Parent record, time base (www.ecfr.gov)	Dynamic archive; searchable export where practical (www.fda.gov)	Separation from data	Reconstruct a sample edit history	At least as long as parent	System owner
Electronic signatures	Signer, time, meaning, link (www.ecfr.gov)	Native, or archive that preserves manifestation	Link lost on conversion (www.ecfr.gov)	Verify manifestation on sample	Same as signed record	QA
Methods and processing parameters	Version, audit trail	Archived with each result	Wrong method version	Re-trace a result to method version	Same as results	Lab or process owner
System configurations	Validated state	Archived with validation evidence (www.oecd.org)	Cannot restore state	Document-level completeness check	Life of dependent records	Validation lead
Attachments and links	Parent-child references	Archive with reference map	Broken references	Link and attachment integrity rate	Same as parent	Records manager
Code lists and reference data	Version effective at the time (www.cdisc.org)	Archived as versioned reference data	Terms reinterpreted	Version-resolution check	Same as coded data	Data standards owner
Rendering dependencies	Viewer, fonts, runtime (www.loc.gov)	Controlled viewer inventory	End of vendor support	Approaching-end-of-support count	Life of rendered records	IT with QA

The matrix is a decision aid, not a standard. Teams that disagree with a candidate mode should record the rationale. Where two modes are combined, for example a dynamic archive plus certified static copies for human reading, the dependency column should show which copy is the system of record.

Performance and Benchmarks

No public benchmark of GxP archive restoration time, render equivalence or reconstruction success was found, and none is implied. What exists is a set of regulator expectations to test and a set of measures a team can compute. Targets are set by local risk assessment; the OECD, for instance, expects the periodicity of evaluating archived signed records to be justified by management (www.oecd.org).

Authenticity and fixity

OAIS defines Fixity Information as documenting the mechanisms that ensure content has not been altered in an undocumented manner (public.ccsds.org), and the DPC defines fixity as assurance that a file has remained unchanged (dpconline.org). The DPC suggests, as a general guideline, checks every six months for hard-drive

based systems (dpconline.org), while noting that MD5 suffices for accidental damage but not where malicious alteration matters (dpconline.org). Checksum algorithms have a lifecycle: NIST announced a transition away from SHA-1 for applying cryptographic protection (csrc.nist.gov), and recorded that SHA-1 may still be needed to handle information protected earlier (csrc.nist.gov). Fixity alone does not prove authenticity; the OECD also expects records moved into an archive to be locked against undetected alteration or deletion (www.oecd.org).

Encryption and key longevity

NIST defines a key archive as a repository of keys and associated information (nvlpubs.nist.gov), says some keys may need to be archived beyond their usage period (nvlpubs.nist.gov), and requires archived key information to be stored separately from operational data (nvlpubs.nist.gov). For signed records, the verification period can outlast the signing period (nvlpubs.nist.gov). An encrypted archive is only as durable as its key archive, a point the regulator texts reviewed do not address directly.

Eight verification measures (editorial calculations)

- **Inventory coverage** = records classified with a retention rule and owner / records identified.
- **Dependency coverage** = records whose required metadata, audit trail, signature manifestation, method and configuration are verified present / records requiring them.
- **Fixity-check success** = files whose recomputed digest matches the stored digest / files checked.
- **Link and attachment integrity** = references that resolve to the intended target / references checked.
- **Render equivalence exceptions** = sampled records with a defined attribute differing from source / records sampled.
- **Retrieval success and elapsed time** = successful retrievals / requested, with median and maximum elapsed time against a locally set target.
- **Reconstruction success** = test scenarios an independent reviewer rebuilt from the archive alone / scenarios attempted.
- **End-of-support exposure** = formats and viewers whose vendor support ends within the review horizon / formats and viewers inventoried.

(Hypothetical Example) If 400 records are sampled and 8 show a differing signature manifestation, render equivalence exceptions are 2.0 percent. The figure has no meaning until the team has defined the attributes compared and the acceptable response, which this report does not set.

Archive restoration protocol (editorial, built from cited expectations)

1. **Select** a risk-weighted sample across record classes, including at least one record that depends on each dependency class.
2. **Retrieve** from the archive, not from the source, and record elapsed time. Annex 11 expects the ability to retrieve data to be ensured and tested when systems change (health.ec.europa.eu).
3. **Check** fixity and accessibility, readability and integrity, as Annex 11 clause 17 expects (health.ec.europa.eu).
4. **Compare** the rendered record with the source or certified copy on pre-defined attributes.

5. **Reconstruct** one event history using audit trail, signature and method, by a reviewer who did not operate the system.
6. **Record** results and deviations; the EMA expects periodic test retrieval or restores to confirm availability and integrity (www.ema.europa.eu), and the OECD expects stored data to be verified for restorability, accessibility, readability and accuracy (www.oecd.org).
7. **Repeat** after any major platform change. A clinical data management trade-association chapter recommends testing archived records after every major upgrade of the active system (scdm.org); that source is an industry association text, not a regulator.

Data Analysis and Evidence

Table 3 lists examples of retention periods stated in the sources, to show variation. It is not a retention schedule and must not be used as one; each entry carries its own conditions.

Source and rule	Stated period	Trigger or condition
UK clinical trial TMF (from 28 April 2026)	At least 25 years (www.gov.uk)	From the day after the trial concludes; transitional trials keep 5 years for TMF documents (www.gov.uk)
EU Clinical Trials Regulation Article 58 (wording checked against a UK-hosted retained copy; confirm against the current EUR-Lex text, CELEX 32014R0536)	At least 25 years (www.legislation.gov.uk , UK-hosted retained copy)	After trial end; subject medical files follow national law (www.legislation.gov.uk , UK-hosted retained copy)
UK advanced therapy traceability	30 years (www.gov.uk)	After product expiry, or longer if authorised
21 CFR 211.180	At least 1 year (www.ecfr.gov)	After batch expiry
EU GMP Chapter 4	1 year after expiry or 5 years after certification (health.ec.europa.eu)	Whichever is longer
21 CFR 58.195 (GLP)	At least 5 years (www.ecfr.gov)	After results submitted to FDA; shortest applicable period shown
21 CFR 312.57 (IND sponsor)	2 years (www.ecfr.gov)	After marketing application approval
21 CFR 314.80(j)	10 years (www.ecfr.gov)	Postmarketing adverse drug experience records

The spread, from two years to thirty, is the reason a single enterprise default is unsafe. The UK entry also shows that a change in law can create different periods for different trials within one organisation, and that the controlling date may be an application date, not a study date.

Quantitative evidence about storage reliability is sparse and dated. A peer-reviewed analysis of 1.53 million production disk drives found more than 400,000 checksum mismatches (static.usenix.org), with 3855 disks developing mismatches overall (static.usenix.org); the authors report that silent corruption could lead to data loss more often than latent sector errors (static.usenix.org). The study was presented in 2008

(static.usenix.org) and describes one vendor's population, so it supports the existence of silent corruption, not a rate for current systems. The DPC notes that media and systems typically last three to five years, requiring storage migration (www.dpconline.org), and the DPC fixity interval is a general archival guideline, not a GxP requirement (dpconline.org).

Regulatory text versions need date stamps. This report reads the MHRA clinical trial archiving page, which states its amendments come into force on 28 April 2026 (www.gov.uk); the UK amending regulations, in force from the same date (www.legislation.gov.uk); the January 2011 text of Annex 11 (health.ec.europa.eu); and the ICH E6(R3) Step 4 text of January 2025. Section numbers refer to those versions and should be rechecked before citation.

Access, Periodic Review and Decommissioning

Access, security and retrieval

Access to archives is restricted in each domain reviewed. MHRA restricts trial archive access to appointed individuals and auditors or inspectors (www.gov.uk); the OECD Principles restrict it to personnel authorised by management (www.oecd.org). The EMA requires continued timely inspector access to data after a system is decommissioned (www.ema.europa.eu). No source reviewed sets a retrieval time target, so retrieval service levels must be set locally. Legal hold is a legal process outside the regulator texts reviewed here; a hold should suspend the destruction step defined below.

Periodic technology and media review

PIC/S flags the risk that access and readability of archived data may be lost through software updates or superseded equipment (picscheme.org), and expects the restore procedure to be regularly tested (picscheme.org). The DPC warns that failing to run technology watch risks loss of access and higher costs (www.dpconline.org). The UK National Archives advises profiling file formats to manage risk (www.nationalarchives.gov.uk). The end-of-support exposure measure above operationalises this.

Decommissioning, destruction approval and the evidence pack

MHRA requires that, if a system cannot be maintained, records be archived under a documented strategy before decommissioning (assets.publishing.service.gov.uk). FDA's 2024 guidance expects metadata to be obtained, retained and linkable to each data element when systems are decommissioned (www.fda.gov). Draft Chapter 4 would add a documented process for disposal of records (health.ec.europa.eu), and ICH E6(R3) allows trial data and metadata to be destroyed only when no longer required by applicable regulatory requirements (database.ich.org).

An evidence pack for sign-off might contain the following (editorial):

- **Inventory:** the record inventory and the retention rule for each class.
- **Dependency map:** metadata, audit trail, signature, method and configuration links per class.
- **Validation report:** the migration or certification validation, including the migration plan (www.oecd.org).
- **Fixity manifests:** digests recorded before and after any move.
- **Restoration results:** the protocol outcomes and any deviations.

- **Contract terms:** provisions covering readability and access for the retention period.
- **Approvals:** the legal and quality approval record, and any legal-hold check.

Where an external provider holds the archive, the sponsor retains responsibility for confidentiality and retrieval (www.ema.europa.eu), and the agreement should cover the provider going out of business.

Implications and Future Directions

Three implications follow. First, the unit of decision is the record class, not the system: a single legacy platform may justify a dynamic archive for raw data, a certified static copy for reports and a controlled viewer for audit trails. Second, the regulatory text is moving toward explicit archive controls. The consultation draft of Annex 11 includes a dedicated archiving section requiring audit trails to be protected from deletion and change (health.ec.europa.eu) and addresses long-term archiving on media with limited durability (health.ec.europa.eu). These remain drafts and should be tracked, not applied.

Third, the ICH E6(R3) guideline requires archiving that allows retrieval and reading and protects against unauthorised alteration (database.ich.org), which regional authorities implement in their own texts.

Outside the GxP texts, the National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation reached version 2.1 in March 2026 (ndsa.org); the earlier matrix expected fixity of content to be checked at fixed intervals (digitalpreservation.gov). Teams may map their own controls to such a framework, as an editorial choice, without treating it as a GxP requirement.

A practical monitoring list is:

- Changes to the governing retention rule.
- Changes to the regulator catalogue entry for the relevant annex or guideline.
- Vendor end-of-support notices for every viewer and format in the inventory.
- The algorithm status of any checksum or encryption method used.

The practical bottleneck is likely the inventory and dependency work that precedes any tooling choice; this is the author's opinion, not evidence. Teams structuring a record-class inventory can start from the regulator texts cited above. Nothing in this report is legal advice; retention periods, acceptable representations and destruction approval depend on jurisdiction and predicate rule and need review by qualified legal and quality professionals.

Frequently Asked Questions (FAQs)

What are the main GxP archive requirements? Across the sources: archives protect records from undetected alteration (www.oecd.org), keep them retrievable and readable (www.oecd.org), restrict access (www.oecd.org), and name a responsible individual, as 21 CFR 58.190 does for GLP (www.ecfr.gov).

How should dynamic records be archived under GxP? Keep dynamic data in a form that retains interactivity where that is critical (picscheme.org); FDA says a printout of a dynamic record does not preserve the dynamic format (www.fda.gov).

Is metadata retention required for data integrity? The EMA treats data without metadata as meaningless (www.ema.europa.eu), and FDA expects copies provided at inspection to include metadata and audit trail information (www.fda.gov).

How is long-term readability ensured? By periodic accessibility checks irrespective of format (www.ema.europa.eu) and a documented migration or emulation plan.

What does 21 CFR Part 11 say about archiving? It requires protection of records for accurate and ready retrieval throughout the retention period (www.ecfr.gov) and the ability to generate accurate and complete copies (www.ecfr.gov); a trade-association chapter notes Part 11 itself sets no specific retention periods (scdm.org).

How does ALCOA+ apply to archives? MHRA's attributes include data being complete, consistent, enduring and available (assets.publishing.service.gov.uk), and "available" means readily available for review or inspection.

How should a validated archive migration be handled? Validate and document it (database.ich.org) (cdn.who.int), keep data, context and audit trail together (www.ema.europa.eu), and verify with checksums where appropriate.

What are GxP data retention and archiving best practices? Define retention periods in an authorised procedure (cdn.who.int), inventory the records (www.ema.europa.eu), keep multiple independent copies in different geographic locations (www.dpconline.org), and verify restorability periodically (www.oecd.org). These are practices drawn from the cited texts, not a prescribed sequence.

How do archive solutions compare? This report compares preservation modes, not products. The ISPE guide references retention, archiving and migration in an appendix (ispe.org); no product rankings are offered.

Conclusion

Archiving regulated records is a series of record-class decisions made against four tests: whether the preserved representation stays authentic, complete, intelligible and reconstructable after technology changes. The regulator texts reviewed agree on the direction: metadata and audit trails travel with the record, dynamic data are not reduced to static images without justification, migration is validated, archives are read-only and access-controlled, and readability is verified over time rather than assumed.

They differ on retention periods, terminology and acceptable representations, and parts of the most detailed EU archiving language remain consultation drafts. For that reason this report offers a matrix and a protocol, not a standard answer. The measures defined here (inventory coverage, dependency coverage, fixity success, link integrity, render equivalence exceptions, retrieval success, reconstruction success and end-of-support exposure) give a team evidence for its own thresholds. No benchmark values are claimed, and none should be borrowed from this report.

Before any source system is switched off or any record destroyed, the governing domain, jurisdiction, predicate rule and retention period should be confirmed, the evidence pack assembled, and legal and quality approval documented. Regulatory and standards statements here are dated as of 27 October 2026 and should be rechecked against the current regulator catalogue before reuse.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://assets.publishing.service.gov.uk/media/5aa2b9ede5274a3e391e37f3/MHRA_GxP_data_integrity_guide_March_edited_Final.pdf
2. https://www.oecd.org/environment/glp-data-integrity_45779212-en
3. <https://www.gov.uk/government/publications/clinical-trials-for-medicines-archiving-and-retention-of-clinical-trial-records/archiving-and-retention-of-clinical-trial-records>
4. <https://picscheme.org/docview/4234>
5. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf
6. <https://www.ecfr.gov/current/title-21/section-211.180>
7. <https://www.ecfr.gov/current/title-21/section-312.57>
8. https://health.ec.europa.eu/consultations/stakeholders-consultation-eudralex-volume-4-good-manufacturing-practice-guidelines-chapter-4-annex_en
9. https://health.ec.europa.eu/medicinal-products/eudralex/eudralex-volume-4_en
10. <https://public.ccsds.org/Pubs/650x0m3.pdf>
11. <https://www.fda.gov/media/119267/download>
12. <https://www.cdisc.org/standards/data-exchange/odm>
13. https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf
14. https://health.ec.europa.eu/document/download/104b3eb8-81a7-4858-9419-cb06562adb66_en?filename=chapter4_01-2011_en.pdf
15. <https://www.fda.gov/media/75414/download>
16. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-11>
17. https://health.ec.europa.eu/document/download/8d305550-dd22-4dad-8463-2ddb4a1345f1_en?filename=annex11_01-2011_en.pdf
18. https://www.oecd.org/content/dam/oecd/en/publications/reports/1998/01/oecd-principles-on-good-laboratory-practice_g1gh32e8/9789264078536-en.pdf
19. https://www.oecd.org/content/dam/oecd/en/publications/reports/2016/04/application-of-glp-principles-to-computerised-systems_7d05366e/e3c9cda5-en.pdf
20. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-content-management-and-archiving-clinical-trial-master-file-paper-and-or-electronic_en.pdf
21. <https://cdn.who.int/media/docs/default-source/medicines/norms-and-standards/guidelines/inspections/trs1033-annex4-guideline-on-data-integrity.pdf>
22. <https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml>
23. <https://www.gov.uk/guidance/good-clinical-practice-for-clinical-trials>
24. https://health.ec.europa.eu/document/download/40231f18-e564-4043-94de-c031f813d38b_en?filename=mp_vol4_chap4_annex11_consultation_guideline_en.pdf
25. <https://dpconline.org/handbook/technical-solutions-and-tools/file-formats-and-standards>
26. <https://www.fda.gov/media/166215/download>

27. <https://www.loc.gov/preservation/digital/formats/fdd/fdd000125.shtml>
28. <https://dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums>
29. <https://csrc.nist.gov/news/2022/nist-transitioning-away-from-sha-1-for-all-apps>
30. https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/establishment-and-control-of-archives-that-operate-in-compliance-with-the-principles-of-glp_g1ghcbc4/9789264084995-en.pdf
31. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-57pt1r5.pdf>
32. <https://scdm.org/wp-content/uploads/2024/07/Clinical-Data-Archiving.pdf>
33. <https://www.legislation.gov.uk/eur/2014/536/article/58>
34. <https://www.ecfr.gov/current/title-21/section-58.195>
35. <https://www.ecfr.gov/current/title-21/section-314.80>
36. https://static.usenix.org/events/fast08/tech/full_papers/bairavasundaram/bairavasundaram_html/index.html
37. <https://www.dpconline.org/handbook/organisational-activities/storage>
38. <https://www.legislation.gov.uk/ukxi/2025/538/regulation/22>
39. <https://www.dpconline.org/handbook/organisational-activities/preservation-planning>
40. <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records>
41. https://health.ec.europa.eu/document/download/fa336a6e-753b-46fc-b53b-9178bacd8878_en?filename=mp_vol4_chap4_consultation_guideline_en.pdf
42. <https://ndsa.org/publications/levels-of-digital-preservation>
43. <https://digitalpreservation.gov/meetings/documents/ndiipp13/NDSA-Levels-DigitalPreservation-NDSA-2013.pdf>
44. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-58/subpart-J/section-58.190>
45. <https://ispe.org/publications/guidance-documents/gamp-records-pharmaceutical-data-integrity>

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore [AI enablement](#), [pharma and biotech software development](#), [data engineering and BI](#), and [Veeva services](#). [Contact IntuitionLabs](#) to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: <https://intuitionlabs.ai>

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.