



AN INDEPENDENT ARCHIVAL FIELD GUIDE / 25 OCT 2026

A Hundred-Year Archive Checklist.

Build an archive that outlasts its moment.

A usable guide to formats, context, provenance, backups, rights and stewardship. For a creator, researcher, independent publisher or future reader who wants more than a beautiful dead link.

NOT A 100-YEAR GUARANTEE. A PRACTICE OF MAKING WORK UNDERSTANDABLE.

FIELD NOTE / 01 ORIENTATION

Preserve an honest question.

An archive is an invitation to interpretation, not a claim to immortality.

A successful preservation plan does not begin with finding the most impressive tool. It begins by choosing what should remain intelligible: an essay, research project, source notebook, oral history, piece of code, or a group of related objects. Specify the intended audience, legitimate access, risks and what would count as a faithful reading.

The time horizon of a creative work is not the same thing as its public attention. A brief project can carry extraordinary meaning; a wildly shared work can be incomprehensible once its platform changes. The aim is to keep the evidence and context that allow later readers to assess a work, not to guarantee agreement or reverence.

A five-part preservation question

- **Object:** What files, originals, images, datasets or pages constitute the work?
- **Meaning:** What definitions, sources, dates and interpretive limits make it understandable?
- **Rights:** What may be copied, moved, published or restricted?
- **Technical independence:** What can be opened without today's account or code?
- **Care:** Who may inspect, test, repair or responsibly retire it?

Scope exercise: Write a single sentence: "I want to keep [collection] understandable for [legitimate reader] because [purpose]. I will not make public [sensitive or unlicensed material]."

This guide is educational. It does not create a repository, establish copyright, conduct a backup, certify preservation readiness or promise a particular lifespan.

FIELD NOTE / 02 THE OBSERVATORY

One, ten, one hundred.

Use horizons to change your questions. Do not use them as predictions.

01 year / The work is still near

Check that the published page and its references open. Save the original authoring files and a readable export. Record the date, author, intended audience, and notable errors already found. While memories are fresh, annotate the terms you assume everyone knows. Select a named custodian, even if that person is only you.

10 years / The interface has moved

A site may be redesigned, an account removed or a font unavailable. Keep stable bibliographic details, a revision history, descriptions of input data and instructions for opening the files without the original platform. Reassess whether personally identifying material should still be distributed. Make the handover plan understandable to someone with no access to your private notes.

100 years / The context becomes unfamiliar

The crucial challenge may not be byte-level readability. A stranger could open a document without knowing the city, social setting, original audience or ethical limits. Record why the work existed, who it affected, what was controversial, what is fictional and what cannot be confirmed. Explain the difference between preserving a historical record and endorsing it.

Worksheet A — three lines:

In one year, I should verify: _____

In ten years, I should explain: _____

For an unknown future reader, I should preserve: _____

There is no magic calendar. More copies, good documentation and human care can reduce particular risks; they cannot turn a hundred-year aspiration into a verified outcome.

FIELD NOTE / 03 FORMATS

A readable copy is not enough.

Prefer interoperable, documented exports alongside editable source and its context.

Store the authored original whenever lawful and practical, then produce a form that a different program can inspect. A simple HTML document is useful for prose and links; plain UTF-8 text is unusually adaptable; CSV is useful for simple rows if accompanied by column definitions; JSON is useful for structured records if its field names and rules are explained. PDF can preserve intended pagination but may not contain all editable sources. None of these formats is invulnerable.

A practical paired-format inventory

WORK	SOURCE / ORIGINAL	READABLE / EXPORT
Essay	Markdown, HTML or editor source	HTML, UTF-8 TXT, optionally PDF
Dataset	Original records plus codebook	CSV with encoding and data dictionary
Website	Source HTML/CSS/JS and local assets	Static HTML snapshot + README
Visual work	Editable layers / master original	PNG or PDF plus descriptive notes
Audio / oral history	Original audio with permissions	Transcript + format and speaker notes

The dependency receipt

- [] 01 **Software.** List applications, versions and proprietary files required to rebuild or open the work
- [] 02 **Web services.** Identify APIs, hosted scripts, external data and third-party accounts
- [] 03 **Fonts and media.** Record licences, missing originals and reasonable fallbacks
- [] 04 **Instructions.** Test on a clean machine or by a person who does not know your setup

A format label is a description, not a preservation policy. Re-open exports periodically, and keep the original when permissions permit.

FIELD NOTE / 04 LINKS AND CITATIONS

A link is a direction, not a source.

Write citations that remain understandable when the destination cannot be reached.

Link rot includes deleted pages, changed domains and pages whose substance changes while the URL stays stable. A naked “see here” gives a future reader almost nothing. Link preservation is partly a writing practice: pair URLs with author or organization, title, publication date, access date where useful, and stable identifiers such as DOIs when provided.

Before and after

Fragile: “Our methods follow the report here.”

More useful: “Our method draws on the [organization] report, [full title], [edition/year], section [name], accessed [date], URL or persistent identifier [value]. We changed the method in the following ways: [list].”

Where legally and ethically permissible, retain a source excerpt, web snapshot, or citation export with its origin documented. Do not assume a commercial archive will always retain material or grant reuse. For sources that cannot be redistributed, preserve descriptive metadata and a record of where the source was consulted.

Link repair worksheet

- [] **01 Locate.** What external services and URLs does the project rely on?
- [] **02 Describe.** Does every important link have a meaningful human-readable citation?
- [] **03 Mark.** Which references are missing, changed or uncertain?
- [] **04 Replace carefully.** Can a legitimate stable identifier or lawful copy be used?
- [] **05 Record edits.** When a link changes, retain a dated note of what was replaced and why

Broken links are an expected preservation risk, not proof that a creator was careless. Honest annotations are more valuable than silently substituting a different source.

FIELD NOTE / 05 PROVENANCE

Give the work a passport.

A metadata record answers who, what, when, how, from where and under which permissions.

Provenance is a trail of evidence and description about origins and changes. It helps a later reader decide whether a record belongs to a particular work, which version they are seeing and what its limitations are. Metadata does not, by itself, authenticate an object or grant permissions.

A compact human-readable specimen

title	When the city learned to listen [fictional example]
creator	Example Author [fictional]
created	2026-10-25 [demonstration date]
revision	v1.0 / educational specimen
source	Fictional teaching example; no empirical source
rights	Undetermined; seek permission
format	text/html; charset=utf-8
context	Invented claims; original data and consent missing

Fill in the gaps, never manufacture certainty

If you do not know the creator or licence, use “unknown” or “undetermined” and document why. Distinguish creation date, publication date and export date. Keep revision notes if material claims changed. Avoid publishing private contacts or unlicensed attachments merely to make a record feel complete.

Try it: Use the website's editable Provenance Drawer to export this record as JSON or CSV. Those are real local downloads, not uploads or registrations.

FIELD NOTE / 06 BACKUPS AND INTEGRITY

Test the restore, not the badge.

Copies matter when they are independent, accessible, lawful and recoverable.

The Library of Congress recommends making multiple copies in different locations and regularly checking that documents can be read. The widely used “3-2-1” mnemonic can be a planning starting point, but counts alone do not establish safety. Consider physical risks, shared credentials, encryption-key loss, provider dependence, corruption and accidental deletion. A backup is valuable because it can restore what you need.

A minimal restoration protocol

- Write an inventory of the collection and identify the authoritative original.
- Make independent copies with appropriate privacy and access controls. A synced folder is not automatically an isolated backup.
- Pick a sample that includes unusual formats, linked assets and important metadata.
- Restore that sample somewhere separate. Check it opens, the contents make sense and the permitted reader can access it.
- Record what worked, what failed, the test date and the person who performed it; adjust the plan.

Integrity in ordinary language

A checksum is a computed fingerprint of bytes. A changed checksum can warn of changed data, but the number alone does not say whether the change was authorized or harmful. Keep checksum manifests with clear generation dates and compare them according to a documented process. Preservation programs often consider integrity and information security together.

Restoration log: Tested date _____ Copy/location _____ Sample files _____
Result _____ Problems observed _____ Person responsible _____
Next review _____

The website's health checklist records user declarations only. Checking “independent copies exist” does not create or validate a single copy.

FIELD NOTE / 07 RIGHTS AND CARE

Some files should not be public.

Long-term access must coexist with consent, privacy and the ability to say no.

Before publishing an archive, identify who owns what. Text may be yours while a photograph, sound clip, chart or quoted interview is not. A licence affects potential reuse only if the person applying it has the necessary rights. Rights metadata should describe the actual situation, not the status you wish the work had.

An archival decision may reasonably involve restricted access, removal, redaction or disposal. An oral history may involve a promise of confidentiality; an early draft may include a private address; a dataset may contain identifiers that would be harmful if made public. Metadata can preserve the existence and ethical context of a restricted item without releasing it.

The stewardship conversation

- [] 01 **Who.** Who is currently responsible and who may make decisions later?
- [] 02 **Access.** How does an authorized custodian find files, passwords or recovery instructions securely?
- [] 03 **Rights.** Which parts can be shared, migrated, quoted or withdrawn?
- [] 04 **Capacity.** What time, money and technical skills would continued care require?
- [] 05 **Exit.** When should the project be restricted, described as inactive or retired?

A successor can refuse

A plan is not a conscription notice for family, friends or collaborators. Ask a potential successor whether they want the responsibility. Preserve enough documentation to allow an informed decision, while accepting that a future person may legitimately choose to stop.

The future reader test.

Take any page out of context. Try to interpret it without asking its author.

Example: The fictional essay “When the city learned to listen” says: “Last summer, the experiment changed everything. We used the new dashboard and interviewed eleven people.” Which material would allow a fair interpretation?

- [] **01 Title and date.** Is the original publication date explicit?
- [] **02 Authorship.** Who created, edited and commissioned the work?
- [] **03 Place and audience.** Which city and social circumstances are involved?
- [] **04 Source citations.** Can claims be traced beyond bare hyperlinks?
- [] **05 Methods.** How were observations gathered and interpreted?
- [] **06 Data and units.** Are source data, labels and calculations accessible?
- [] **07 Consent.** May the affected people or records be republished?
- [] **08 Rights.** Do you have permissions for the included materials?
- [] **09 Revision trail.** What claims or files have changed since the first edition?
- [] **10 Exports.** Does a non-specialist readable version exist?
- [] **11 Dependencies.** What apps, data feeds or licences must exist to reconstruct it?
- [] **12 Disagreement.** Are uncertainty, alternatives and known failures named?

Your three most important gaps:

1. _____
2. _____
3. _____

A living archive exam.

Tick only verified activities. Unchecked items are research questions, not a failing grade.

- ☐ **01 Inventory.** I can list the files and their purposes
- ☐ **02 Originals.** Editable originals have been located and opened
- ☐ **03 Context.** A README explains audience, method and limitations
- ☐ **04 Attribution.** Creator, dates and revisions have been documented
- ☐ **05 Sources.** Important citations include titles and identifiers
- ☐ **06 Permissions.** Rights, restrictions and privacy have been reviewed
- ☐ **07 Exports.** An independent readable export was actually opened
- ☐ **08 Dependencies.** Required programs, data and services are described
- ☐ **09 Copies.** Independent copies exist in different places
- ☐ **10 Restoration.** A restore and read test actually succeeded
- ☐ **11 Integrity.** Checking of changes is considered and documented
- ☐ **12 Care plan.** A realistic custodian, handover or exit plan exists

Next real actions

Priority 1: _____ Owner: _____

Priority 2: _____ Owner: _____

Priority 3: _____ Owner: _____

Review / restoration date: _____ Source of evidence: _____

Honesty rule: A checkbox cannot inspect files, make a backup, restore a drive, verify rights, create a steward or guarantee future access. Verify every claim in the real world.

Pass the work forward.

A long game is a craft of transfer, not a promise of permanence.

A six-line handover letter

To a future reader: This collection was created in _____ for _____. It tried to answer _____. Its most important source files are _____. The uncertainties and restrictions are _____. If the interactive layer no longer works, start with _____. Permission to care for or retire this material is governed by _____.

What this guide does not prove

A working URL today is not evidence of century-scale durability. No single format prevents obsolescence. A checksum does not supply social context. An attribution field does not establish copyright. A named successor has not accepted responsibility unless they actually agree. No advice here replaces legal, records-management or professional archival review when such review is needed.

Further exploration / primary guidance

Library of Congress / Personal Digital Records — Practical advice on identifying, organizing, copying and checking personal digital materials.

NDSA / Levels of Digital Preservation — Structured areas for reviewing preservation practice; version 2.1 available in 2026.

Digital Preservation Coalition / Metadata and Documentation — Metadata, recording preservation events and documenting provenance.

W3C / Web Sustainability Guidelines — Progressive enhancement, accessibility, reuse, low dependencies and responsible lifecycle design.

Read this guide on-screen or print the worksheets. The companion interactive studio is at sahilverse.online/thelonggame/; the central Sahilverse entrance is sahiltayal.com. Contact: sahil@sahiltayal.com.