

ADAC-Genealogy 1.0 — Genealogy Profile Format Specification

Version: 1.0

Status: Stable

Copyright: © 2026 InnoVadens, LLC. All rights reserved.

Date: 2026

Base Format: [ADAC 1.0 — Archival Digital Asset Container](#)

Abstract

The **ADAC-Genealogy** profile defines a domain-specific metadata extension for the [ADAC 1.0](#) container format, designed for the long-term preservation and interchange of digitized genealogical records — census pages, vital records, church registers, land deeds, immigration manifests, and other primary sources used in family history research. The profile adds structured source citations, multi-page record linking, Genealogical Proof Standard evidence analysis, and per-region annotations for persons, transcriptions, and document condition — all without modifying the core ADAC specification.

Table of Contents

1. [Introduction](#)
2. [Conformance](#)
3. [Terminology](#)
4. [Relationship to ADAC](#)
5. [Container Format](#)
6. [Genealogy Profile — `metadata/profiles/genealogy.json`](#)
7. [Source Citation](#)

- 8. [Multi-Page Record Links](#)
 - 9. [Evidence Analysis](#)
 - 10. [Region Linked Entities](#)
 - 11. [Person Annotation](#) — `genealogy:person`
 - 12. [Transcription](#) — `genealogy:transcription`
 - 13. [Record Condition](#) — `genealogy:condition`
 - 14. [Well-Known Value Sets](#)
 - 15. [JSON Conventions](#)
 - 16. [Media Types & File Extensions](#)
 - 17. [Interoperability](#)
 - 18. [Complete Container Example](#)
 - 19. [Validation](#)
 - 20. [References](#)
 - 21. [Version History](#)
-

1. Introduction

1.1 Problem Statement

Genealogical research depends on the careful analysis of primary source documents — census enumerations, birth certificates, parish registers, probate records, and immigration manifests. When these documents are digitized, the resulting scans capture the visual content but lose the structured context that makes them useful for research: which repository holds the original, what jurisdiction it covers, who the people named on the page are, what the handwriting says, and how this source relates to other evidence in a proof argument.

Existing approaches store this context in proprietary databases, genealogy software files, or unstructured notes — creating fragile dependencies that break when systems change.

1.2 Solution

ADAC-Genealogy solves this by embedding genealogical context directly into the archival container alongside the scanned image:

- **Source citation** — archival repository, collection hierarchy, microfilm references, jurisdiction, date range, and language.
- **Person annotations** — per-region identification of individuals and their vital data as recorded in the source.
- **Transcriptions** — per-region text extracted from handwriting or print, with confidence scores for AI-assisted work.
- **Record condition** — per-region legibility and damage assessments, documenting why certain facts could not be confirmed.
- **Multi-page record links** — structural connections between masters when a single logical record (e.g., a household enumeration) spans multiple pages.
- **Evidence analysis** — Genealogical Proof Standard classification, correlation notes linking to other sources, and documented conclusions.

1.3 Design Philosophy

Principle	Description
ADAC-native	ADAC-Genealogy is a pure ADAC profile extension. No core specification modifications are required. Any ADAC reader can open an ADAC-Genealogy container.
Genealogy-first	Every schema element maps to a concept central to genealogical research and the Genealogical Proof Standard.
Non-intrusive	Non-genealogy ADAC consumers ignore all genealogy data. Linked entity values survive round-trip serialization unchanged.
Source-faithful	Date fields, names, and ages use free-text strings to faithfully represent the source record — genealogical documents routinely contain approximate, partial, or non-standard dates that structured types cannot capture.
Ecosystem-interoperable	Supports GEDCOM XREF cross-references, FamilySearch FHL film numbers, and external system identifiers for integration with existing genealogy platforms.
Forward-compatible	All JSON structures tolerate unknown properties. Future profile versions can add fields without breaking existing readers.

1.4 Scope

This specification defines:

- The genealogy profile file schema (`metadata/profiles/genealogy.json`)
- Source citation, collection hierarchy, film reference, and jurisdiction schemas
- Multi-page record linking
- Evidence analysis, classification, and correlation note schemas
- Per-region linked entity schemas: person annotations, transcriptions, and record conditions
- Well-known value sets for record types, evidence classifications, legibility values, and damage types

This specification does **not** define:

- The ADAC core container format (see the [ADAC 1.0 Format Specification](#))

- Application-level APIs, programming interfaces, or dependency injection
- User interface requirements for displaying genealogical data

1.5 Audience

This specification is intended for:

- Software developers implementing genealogy features in ADAC readers and writers
 - Genealogists and archivists evaluating the format for digitized record preservation
 - Developers of genealogy software (GEDCOM editors, family tree applications) considering ADAC integration
 - Standards bodies evaluating the format for adoption
-

2. Conformance

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#).

2.1 Profile Conformance

An ADAC container conforms to this profile specification when:

1. It is a valid ADAC 1.0 container (Minimal or Archival conformance).
2. It contains a file at a path listed in the manifest's `metadata.profiles` array whose filename portion matches `genealogy.json` (case-insensitive).
3. That file is valid JSON containing at least the `profileVersion` and `profileType` fields, with `profileType` equal to `"genealogy"`.

2.2 Forward Compatibility

Readers MUST tolerate unknown JSON properties in all genealogy profile structures. Unknown properties MUST be preserved during round-trip serialization.

3. Terminology

Term	Definition
ADAC- Genealogy Container	An ADAC container that includes a genealogy profile at <code>metadata/profiles/genealogy.json</code> . Uses the standard <code>.adac</code> extension.
Source Citation	Structured description of the archival repository, collection hierarchy, film reference, jurisdiction, and date range of the source record.
Person Annotation	Per-region metadata identifying a person and their vital data as recorded in the source document. Linked entity key: <code>genealogy:person</code> .
Transcription	Per-region text transcribed from the original handwriting or print. Linked entity key: <code>genealogy:transcription</code> .
Record Condition	Per-region assessment of legibility and physical damage. Linked entity key: <code>genealogy:condition</code> .
Page Link	A structural tie between multiple master files within a container, grouping them into a single logical multi-page record.
Evidence Analysis	GPS-compliant classification, correlation, and conclusion metadata for the source.
Correlation Note	A note linking the current source to another piece of evidence, optionally referencing another ADAC container by its manifest <code>id</code> .
GPS	Genealogical Proof Standard — the five-element framework for establishing genealogical conclusions: reasonably exhaustive search, complete and accurate citations, analysis and correlation, resolution of conflicting evidence, and a soundly reasoned written conclusion.
Collection Hierarchy	The standard archival arrangement: collection → series → box → folder → item.
Film Reference	Microfilm roll, frame, and FamilySearch Family History Library (FHL) film number.

4. Relationship to ADAC

4.1 Extension Mechanisms

ADAC-Genealogy extends ADAC through two mechanisms defined by the ADAC 1.0 specification:

1. **Profile file** — A self-describing JSON file placed at `metadata/profiles/genealogy.json` and referenced in the manifest's `metadata.profiles` array. Non-genealogy readers ignore this file entirely.
2. **Linked entities** — Per-region JSON objects stored in the `linkedEntities` dictionary on each region, keyed by namespaced strings (`genealogy:person` , `genealogy:transcription` , `genealogy:condition`). Non-genealogy readers preserve these values unchanged during round-trip serialization.

4.2 What This Profile Does NOT Modify

- ❌ The ADAC container format (ZIP structure)
- ❌ The manifest schema (`manifest.json`)
- ❌ The core metadata schema (`metadata/core.json`)
- ❌ The region annotation schema (structure of regions or bounds)
- ❌ The edit pipeline schema
- ❌ The provenance log or checksum manifest schemas
- ❌ Any ADAC validation rules or error codes

5. Container Format

5.1 File Extension

Property	Value
File extension	.adac
MIME type	application/vnd.adac.container

ADAC-Genealogy containers use the standard `.adac` extension. The presence of the genealogy profile is identified by the `profileType` field inside `metadata/profiles/genealogy.json`, not by the file extension. This avoids extension proliferation and keeps the format unified.

5.2 Directory Structure

```
<container>.adac
├─ manifest.json          (REQUIRED – ADAC)
├─ master/               (REQUIRED – ADAC)
│   └─ master_0001.tif
│   └─ ...
├─ metadata/            (REQUIRED – ADAC)
│   └─ core.json         (REQUIRED – ADAC)
│   └─ xmp/              (OPTIONAL – ADAC)
│       └─ ...
│   └─ profiles/        (REQUIRED for this profile)
│       └─ genealogy.json (REQUIRED for this profile)
├─ derivatives/         (OPTIONAL – ADAC)
│   └─ ...
├─ regions/             (RECOMMENDED for this profile)
│   └─ master-001.regions.json
│   └─ ...
├─ edits/               (OPTIONAL – ADAC)
│   └─ ...
└─ provenance/         (RECOMMENDED – ADAC)
    └─ log.json
    └─ checksums.json
```

5.3 Manifest Integration

```
{
  "metadata": {
    "core": "metadata/core.json",
    "profiles": [
      "metadata/profiles/genealogy.json"
    ],
    "provenanceLog": "provenance/log.json",
    "checksums": "provenance/checksums.json"
  }
}
```

6. Genealogy Profile

Path: metadata/profiles/genealogy.json

Status: REQUIRED for ADAC-Genealogy containers

The genealogy profile is the root metadata file for this extension. It aggregates source citation, multi-page linking, and evidence analysis. Per-region data (person annotations, transcriptions, record conditions) is stored separately in each region's `linkedEntities` dictionary.

6.1 Schema

```
{
  "profileVersion": "1.0",
  "profileType": "genealogy",
  "sourceCitation": { ... },
  "pageLinks": [ ... ],
  "evidence": { ... }
}
```

6.2 Fields

Field	Type	Required	Default	Description
profileVersion	string	REQUIRED	"1.0"	The profile schema version.
profileType	string	REQUIRED	"genealogy"	The profile type identifier. MUST be "genealogy".
sourceCitation	SourceCitation	OPTIONAL	null	Archival source information (see Section 7).
pageLinks	PageLink[]	OPTIONAL	null	Multi-page record links (see Section 8).
evidence	EvidenceAnalysis	OPTIONAL	null	GPS evidence classification and analysis (see Section 9).

6.3 Forward Compatibility

Unknown JSON properties MUST be preserved during round-trip serialization. This ensures that future profile versions can add fields without breaking existing readers.

7. Source Citation

The source citation describes where the original record is held, what collection it belongs to, how to locate it on microfilm, what jurisdiction and date range it covers, and what language(s) it is written in.

7.1 Schema

```
{
  "sourceCitation": {
    "recordType": "census",
    "repository": "National Archives, Washington, D.C.",
    "collection": {
      "collection": "T9",
      "series": "1870 Census",
      "box": null,
      "folder": null,
      "item": "Roll 1042"
    },
    "film": {
      "roll": "1042",
      "frame": "0142",
      "fhlFilmNumber": "0553042"
    },
    "url": "https://www.familysearch.org/ark:/61903/3:1:S3HY-X9BS-F6R",
    "jurisdiction": {
      "country": "United States",
      "state": "Ohio",
      "county": "Licking",
      "town": "Newark"
    },
    "dateRangeStart": "1870-06-01",
    "dateRangeEnd": "1870-08-31",
    "languages": ["en"]
  }
}
```

7.2 Source Citation Fields

Field	Type	Required	Description
<code>recordType</code>	<code>string</code>	OPTIONAL	The genealogical record type. See Section 14.1 for well-known values.
<code>repository</code>	<code>string</code>	OPTIONAL	The archival repository or holding institution.
<code>collection</code>	<code>CollectionHierarchy</code>	OPTIONAL	Collection / series / box / folder / item hierarchy (see Section 7.3).
<code>film</code>	<code>FilmReference</code>	OPTIONAL	Microfilm or FHL film reference (see Section 7.4).
<code>url</code>	<code>string</code>	OPTIONAL	URL to an online source (FamilySearch, Ancestry, FindMyPast, etc.).
<code>jurisdiction</code>	<code>Jurisdiction</code>	OPTIONAL	Geographic jurisdiction the record covers (see Section 7.5).
<code>dateRangeStart</code>	<code>string</code>	OPTIONAL	Start of the date range covered by the source (ISO-8601 date).
<code>dateRangeEnd</code>	<code>string</code>	OPTIONAL	End of the date range covered by the source (ISO-8601 date).
<code>languages</code>	<code>string[]</code>	OPTIONAL	Language(s) of the document. IETF BCP 47 tags RECOMMENDED (e.g., <code>"en"</code> , <code>"de"</code> , <code>"la"</code>).

7.3 Collection Hierarchy

Describes the archival collection hierarchy following the standard arrangement:
collection → series → box → folder → item.

```
{
  "collection": "T9",
  "series": "1870 Census",
  "box": null,
  "folder": null,
  "item": "Roll 1042"
}
```

Field	Type	Description
collection	string	The collection name or number.
series	string	The series identifier within the collection.
box	string	The box number or identifier.
folder	string	The folder number or identifier.
item	string	The item number or identifier.

All fields are OPTIONAL. Populate only the levels applicable to the record.

7.4 Film Reference

Describes microfilm, roll, and FHL (Family History Library) film references.

```
{
  "roll": "1042",
  "frame": "0142",
  "fhlFilmNumber": "0553042"
}
```

Field	Type	Description
roll	string	The microfilm roll number.
frame	string	The frame number on the roll.
fhlFilmNumber	string	The FamilySearch FHL film number.

All fields are OPTIONAL.

7.5 Jurisdiction

Describes the geographic jurisdiction that the record covers, from broadest to most specific.

```
{  
  "country": "United States",  
  "state": "Ohio",  
  "county": "Licking",  
  "town": "Newark"  
}
```

Field	Type	Description
country	string	The country.
state	string	The state, province, or equivalent administrative division.
county	string	The county, parish, or equivalent subdivision.
town	string	The town, city, or municipality.

All fields are OPTIONAL. Populate only the levels relevant to the record's jurisdiction.

8. Multi-Page Record Links

Page links tie multiple master files within a container into a single logical record. This is common in genealogical records where a census enumeration, church register entry, or probate document spans multiple pages.

8.1 Schema

```
{
  "pageLinks": [
    {
      "recordGroupId": "household-42",
      "pageSequence": 1,
      "masterId": "master-001",
      "description": "Census sheet A"
    },
    {
      "recordGroupId": "household-42",
      "pageSequence": 2,
      "masterId": "master-002",
      "description": "Census sheet B (continuation)"
    }
  ]
}
```

8.2 Fields

Field	Type	Required	Description
<code>recordGroupId</code>	<code>string</code>	REQUIRED	Groups pages of the same logical record. All masters belonging to the same record MUST share this value.
<code>pageSequence</code>	<code>integer</code>	REQUIRED	1-based reading order within the record group. MUST start at 1 with consecutive integers.
<code>masterId</code>	<code>string</code>	OPTIONAL	The master identifier within this container that this link refers to. SHOULD match a master <code>id</code> in the manifest.
<code>description</code>	<code>string</code>	OPTIONAL	Describes this page's role in the record (e.g., "Census sheet header", "Continuation of household 42").

8.3 Usage Guidelines

- **Record Group ID:** Use a descriptive, unique-within-container identifier (e.g., `"household-42"`, `"probate-entry-001"`).
- **Page Sequence:** Always start at **1**. Consecutive integers establish reading order.

- **Multiple Record Groups:** A container MAY have multiple record groups (e.g., a container with 10 census pages may have 5 households, each spanning 2 pages).

9. Evidence Analysis

The evidence analysis section supports the Genealogical Proof Standard (GPS) by providing structured fields for evidence classification, source correlation, and documented conclusions.

9.1 Schema

```
{
  "evidence": {
    "classification": "direct",
    "correlationNotes": [
      {
        "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
        "containerId": "adac-1860-census-042",
        "note": "Same household head listed 10 years earlier confirms continuity of re
      }
    ],
    "conclusion": "John Smith, age 42, farmer, born Ohio circa 1828.",
    "analysisText": "Direct evidence from the 1870 census enumeration lists John Smith
    "analyst": "Jane Researcher",
    "analyzedOn": "2025-02-01T14:00:00Z"
  }
}
```

9.2 Evidence Analysis Fields

Field	Type	Required	Description
<code>classification</code>	<code>string</code>	OPTIONAL	The evidence classification. See Section 14.2 for well-known values.
<code>correlationNotes</code>	<code>CorrelationNote[]</code>	OPTIONAL	Notes correlating this source with other evidence.
<code>conclusion</code>	<code>string</code>	OPTIONAL	The researcher's conclusion drawn from this source.
<code>analysisText</code>	<code>string</code>	OPTIONAL	The full analysis text explaining the reasoning.
<code>analyst</code>	<code>string</code>	OPTIONAL	The researcher who performed the evaluation.
<code>analyzedOn</code>	<code>string</code>	OPTIONAL	ISO-8601 timestamp of when the analysis was performed.

9.3 Correlation Note

A correlation note links the current source to another piece of evidence, optionally referencing another ADAC container by its manifest `id`.

```
{
  "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
  "containerId": "adac-1860-census-042",
  "note": "Same household head listed 10 years earlier confirms continuity."
}
```

Field	Type	Required	Description
<code>sourceDescription</code>	<code>string</code>	REQUIRED	A human-readable description of the correlated source.
<code>containerId</code>	<code>string</code>	OPTIONAL	The ADAC container <code>id</code> (from its manifest) for the correlated source, enabling cross-container linking.
<code>note</code>	<code>string</code>	REQUIRED	Text explaining how the two sources relate.

9.4 Cross-Container Linking

The `containerId` field in correlation notes enables researchers to build a web of correlated evidence across multiple containers. When present, applications can navigate to the referenced container to display the correlated source alongside the current one. The `containerId` value SHOULD match the `id` field in the referenced container's manifest.

10. Region Linked Entities

ADAC-Genealogy stores per-region genealogical data in the ADAC region `linkedEntities` dictionary — a JSON object on each region where keys are namespaced strings and values are arbitrary JSON objects.

10.1 Linked Entity Keys

Key	Description
<code>genealogy:person</code>	Person identification and vital data (see Section 11).
<code>genealogy:transcription</code>	OCR/handwriting transcription (see Section 12).
<code>genealogy:condition</code>	Legibility and damage assessment (see Section 13).

10.2 Round-Trip Safety

Non-genealogy consumers that read a container and write it back MUST preserve all linked entity data. Because the ADAC core treats `linkedEntities` values as opaque JSON objects, no genealogy data is lost during round-trip serialization by tools that do not understand the genealogy namespace.

10.3 Multiple Entities per Region

A single region MAY have all three genealogy linked entities simultaneously. For example, a bounding box around a person's name in a census record could have:

- A `genealogy:person` entity identifying the individual
- A `genealogy:transcription` entity with the transcribed text
- A `genealogy:condition` entity noting partial legibility due to faded ink

Genealogy linked entities also coexist with linked entities from other profiles (e.g., `legal:exhibit`) without conflict.

11. Person Annotation

Linked Entity Key: `genealogy:person`

Person annotations link a bounding-box region on a scanned document to a specific person and their associated vital data as recorded in the source.

11.1 Schema

```
{
  "givenName": "John",
  "surname": "Smith",
  "maidenName": null,
  "birthDate": "1828",
  "deathDate": null,
  "marriageDate": null,
  "age": "42",
  "occupation": "Farmer",
  "birthplace": "Ohio",
  "relationshipToHead": "Self",
  "gedcomXref": "@I123@",
  "externalId": "FSFT-ABC-1234"
}
```

11.2 Fields

Field	Type	Required	Description
<code>givenName</code>	<code>string</code>	OPTIONAL	Given (first) name.
<code>surname</code>	<code>string</code>	OPTIONAL	Family name.
<code>maidenName</code>	<code>string</code>	OPTIONAL	Maiden name, if applicable.
<code>birthDate</code>	<code>string</code>	OPTIONAL	Birth date as recorded in the document. Free-text to accommodate partial dates (e.g., <code>"1828"</code> , <code>"circa 1828"</code> , <code>"15 Mar 1828"</code>).
<code>deathDate</code>	<code>string</code>	OPTIONAL	Death date as recorded.
<code>marriageDate</code>	<code>string</code>	OPTIONAL	Marriage date as recorded.
<code>age</code>	<code>string</code>	OPTIONAL	Age at the time of the record. String to accommodate ranges and notations (e.g., <code>"42"</code> , <code>"about 40"</code> , <code>"6/12"</code> for 6 months).
<code>occupation</code>	<code>string</code>	OPTIONAL	Occupation at the time of the record.
<code>birthplace</code>	<code>string</code>	OPTIONAL	Birthplace as recorded.
<code>relationshipToHead</code>	<code>string</code>	OPTIONAL	Relationship to the head of household (census records).
<code>gedcomXref</code>	<code>string</code>	OPTIONAL	GEDCOM XREF identifier linking to a person in a genealogy database (e.g., <code>"@I123@"</code>).
<code>externalId</code>	<code>string</code>	OPTIONAL	External system identifier for cross-referencing (e.g., FamilySearch PID, Ancestry tree person ID).

11.3 Design Rationale — String Dates

Date fields use `string` rather than structured date types. This is a deliberate design decision: genealogical records routinely contain approximate (`"circa 1828"`), partial (`"Mar 1828"`), dual-calendar (`"12/23 Feb 1731/2"`), or descriptive (`"about age 40"`) dates. Structured date types would lose this information. The ADAC-Genealogy profile faithfully preserves dates exactly as they appear in the source record.

12. Transcription

Linked Entity Key: `genealogy:transcription`

Transcriptions capture the text read from a specific region of a scanned document, whether performed manually by a researcher or automatically by OCR/AI handwriting recognition.

12.1 Schema

```
{
  "text": "Smith, John    42    M    W    Farmer    2000    Ohio",
  "confidence": 0.95,
  "transcriber": "GeneaScan AI v2.0",
  "transcribedOn": "2025-01-15T11:00:00Z",
  "originalLanguage": "en",
  "translatedText": null,
  "translatedLanguage": null
}
```

12.2 Fields

Field	Type	Required	Description
<code>text</code>	<code>string</code>	REQUIRED	The transcribed text from the region.
<code>confidence</code>	<code>number</code>	OPTIONAL	AI confidence score (0.0–1.0). <code>null</code> for manual transcription.
<code>transcriber</code>	<code>string</code>	OPTIONAL	The person or system that produced the transcription.
<code>transcribedOn</code>	<code>string</code>	OPTIONAL	ISO-8601 timestamp of when the transcription was performed.
<code>originalLanguage</code>	<code>string</code>	OPTIONAL	Language of the original text (IETF BCP 47 tag, e.g., <code>"en"</code> , <code>"de"</code> , <code>"la"</code>).
<code>translatedText</code>	<code>string</code>	OPTIONAL	Translated text, if the original is in a different language.
<code>translatedLanguage</code>	<code>string</code>	OPTIONAL	Language of the translation (IETF BCP 47 tag).

12.3 Usage Guidelines

- Set `confidence` to `null` for manual transcriptions by human researchers. A value between `0.0` and `1.0` indicates AI-assisted transcription.
- The `transcriber` field SHOULD identify the software version for AI transcriptions (e.g., `"GeneaScan AI v2.0"`) or the researcher's name for manual work.
- When a document is in a non-English language, provide `originalLanguage` + `text` and optionally `translatedLanguage` + `translatedText`.

13. Record Condition

Linked Entity Key: `genealogy:condition`

Record condition assessments document the physical state and legibility of a specific region on a scanned document. This metadata supports GPS compliance —

documenting that a record was partially illegible strengthens the evidentiary chain when explaining why certain facts could not be confirmed.

13.1 Schema

```
{
  "legibility": "partial",
  "damageTypes": ["faded", "waterDamage"],
  "notes": "Bottom-right corner water-stained, last two digits of year unreadable."
}
```

13.2 Fields

Field	Type	Required	Default	Description
legibility	string	OPTIONAL	"clear"	Legibility assessment. See Section 14.3 .
damageTypes	string[]	OPTIONAL	null	List of observed damage types. See Section 14.4 .
notes	string	OPTIONAL	null	Free-text condition notes describing specific observations.

13.3 Usage Guidelines

- Default legibility is "clear". Override only when the region has reduced readability.
- Multiple damage types MAY be specified simultaneously (e.g., a region that is both faded and water-damaged).
- The notes field SHOULD describe location-specific observations (e.g., "bottom-right corner torn, last two digits of year missing").

14. Well-Known Value Sets

ADAC-Genealogy defines controlled vocabularies for several fields. Implementations SHOULD use these well-known values when applicable but MAY use custom values for

domain-specific needs. All values are lowercase strings.

14.1 Record Types

Used in `sourceCitation.recordType`.

Value	Description
"birth"	Birth record or certificate.
"death"	Death record or certificate.
"marriage"	Marriage record, license, or certificate.
"census"	Census enumeration.
"immigration"	Immigration, naturalization, or passenger list.
"military"	Military service, draft, or pension record.
"church"	Church record (baptism, burial, confirmation, etc.).
"probate"	Probate, will, or estate record.
"land"	Land, property, or deed record.
"tax"	Tax record or assessment.

14.2 Evidence Classifications

Used in `evidence.classification`.

Value	Description
"direct"	The source explicitly states the fact being researched.
"indirect"	The source requires inference or combination with other sources to establish the fact.
"negative"	The absence of expected information in the source is itself meaningful evidence.

14.3 Legibility Values

Used in record condition `legibility`.

Value	Description
<code>"clear"</code>	Fully readable without difficulty.
<code>"partial"</code>	Partially readable with some effort required.
<code>"poor"</code>	Mostly unreadable but some text is discernible.
<code>"illegible"</code>	Completely unreadable.

14.4 Damage Types

Used in record condition `damageTypes`.

Value	Description
<code>"torn"</code>	Physical tear in the document.
<code>"faded"</code>	Ink or print has faded.
<code>"waterDamage"</code>	Water damage causing staining or warping.
<code>"fireDamage"</code>	Fire or heat damage.
<code>"foxing"</code>	Brown spots from oxidation.
<code>"mold"</code>	Mold or mildew damage.
<code>"insectDamage"</code>	Insect damage (bookworm holes, etc.).

15. JSON Conventions

All JSON within this profile follows the ADAC 1.0 conventions:

Convention	Requirement
Property naming	camelCase
Formatting	Indented (human-readable)
Null handling	Null-valued properties SHOULD be omitted
Encoding	UTF-8 without BOM
Unknown properties	MUST be preserved during round-trip serialization
Date/time format	ISO-8601

16. Media Types & File Extensions

Format	Extension	MIME Type
ADAC-Genealogy Container	.adac	application/vnd.adac.container

All ADAC containers — regardless of which domain profiles they include — use the `.adac` file extension. The profile type is determined by reading the container's metadata, not by the file extension.

17. Interoperability

17.1 GEDCOM Integration

The `gedcomXref` field on person annotations allows bidirectional linking between ADAC-Genealogy containers and GEDCOM databases. A genealogy application can:

- Export a person annotation with a `gedcomXref` matching an individual in a GEDCOM file.
- Import a GEDCOM file and create corresponding person annotations on scanned document regions.

17.2 FamilySearch Integration

The `fhLFilmNumber` field in film references and the `externalId` field on person annotations support integration with FamilySearch systems — FHL film catalogs, FamilySearch Family Tree PIDs, and Ark identifiers.

17.3 Profile Coexistence

ADAC-Genealogy containers MAY include additional profiles. A probate document could carry both:

- A genealogy profile (source citation, person annotations, transcriptions)
- A legal profile (case reference, custody chain, confidentiality level)

The profiles coexist without conflict. Linked entity keys use distinct namespaces (`genealogy:person` vs. `legal:exhibit`).

17.4 Cross-Container Evidence Networks

The `containerId` field in correlation notes enables researchers to build evidence networks across multiple containers. Applications can navigate between containers to display correlated sources side by side, supporting the GPS requirement for reasonably exhaustive search and source correlation.

18. Complete Container Example

18.1 Container Structure

```
1870-census-licking-oh-042.adac
├─ manifest.json
├─ master/
│  ├─ master_0001.tif          (Census sheet A – 600 DPI TIFF)
│  └─ master_0002.tif        (Census sheet B – continuation)
├─ metadata/
│  ├─ core.json
│  └─ profiles/
│     └─ genealogy.json
├─ derivatives/
│  └─ deriv_0001.jpg          (Web preview of sheet A)
├─ regions/
│  ├─ master-001.regions.json
│  └─ master-002.regions.json
└─ provenance/
   ├─ log.json
   └─ checksums.json
```

18.2 Genealogy Profile (metadata/profiles/genealogy.json)

```
{
  "profileVersion": "1.0",
  "profileType": "genealogy",
  "sourceCitation": {
    "recordType": "census",
    "repository": "National Archives, Washington, D.C.",
    "collection": {
      "collection": "T9",
      "series": "1870 Federal Census",
      "item": "Roll 1042"
    }
  },
  "film": {
    "roll": "1042",
    "frame": "0142",
    "fhlFilmNumber": "0553042"
  },
  "url": "https://www.familysearch.org/ark:/61903/3:1:S3HY-X9BS-F6R",
  "jurisdiction": {
    "country": "United States",
    "state": "Ohio",
    "county": "Licking",
    "town": "Newark"
  },
  "dateRangeStart": "1870-06-01",
  "dateRangeEnd": "1870-08-31",
  "languages": ["en"]
},
"pageLinks": [
  {
    "recordGroupId": "household-42",
    "pageSequence": 1,
    "masterId": "master-001",
    "description": "Census sheet A – household 42 begins"
  },
  {
    "recordGroupId": "household-42",
    "pageSequence": 2,
    "masterId": "master-002",
    "description": "Census sheet B – household 42 continuation"
  }
],
}
```

```
"evidence": {
  "classification": "direct",
  "correlationNotes": [
    {
      "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
      "containerId": "adac-1860-census-042",
      "note": "Same household head listed 10 years earlier confirms continuity of re
    }
  ],
  "conclusion": "John Smith, age 42, farmer, born Ohio circa 1828.",
  "analysisText": "Direct evidence from the 1870 census enumeration lists John Smith
  "analyst": "Jane Researcher",
  "analyzedOn": "2025-02-01T14:00:00Z"
}
```

18.3 Region Annotations (`regions/master-001.regions.json`)

```
{
  "mediaId": "master-001",
  "coordinateSystem": "pixel",
  "width": 4800,
  "height": 6400,
  "regions": [
    {
      "id": "region-001",
      "type": "boundingBox",
      "label": "John Smith – Head of household",
      "bounds": {
        "x": 100.0,
        "y": 200.0,
        "width": 4600.0,
        "height": 80.0
      },
      "linkedEntities": {
        "genealogy:person": {
          "givenName": "John",
          "surname": "Smith",
          "birthDate": "1828",
          "age": "42",
          "occupation": "Farmer",
          "birthplace": "Ohio",
          "relationshipToHead": "Self",
          "gedcomXref": "@I123@"
        },
        "genealogy:transcription": {
          "text": "Smith, John    42    M    W    Farmer    2000    Ohio",
          "confidence": 0.95,
          "transcriber": "GeneaScan AI v2.0",
          "transcribedOn": "2025-01-15T11:00:00Z",
          "originalLanguage": "en"
        },
        "genealogy:condition": {
          "legibility": "clear"
        }
      }
    },
    {
      "id": "region-002",
```

```

"type": "boundingBox",
"label": "Mary Smith – Wife",
"bounds": {
  "x": 100.0,
  "y": 280.0,
  "width": 4600.0,
  "height": 80.0
},
"linkedEntities": {
  "genealogy:person": {
    "givenName": "Mary",
    "surname": "Smith",
    "maidenName": "Jones",
    "birthDate": "1832",
    "age": "38",
    "occupation": "Keeping house",
    "birthplace": "Virginia",
    "relationshipToHead": "Wife"
  },
  "genealogy:transcription": {
    "text": "Smith, Mary 38 F W Keeping house Virginia",
    "confidence": 0.88,
    "transcriber": "GeneaScan AI v2.0",
    "transcribedOn": "2025-01-15T11:00:00Z",
    "originalLanguage": "en"
  },
  "genealogy:condition": {
    "legibility": "partial",
    "damageTypes": ["faded"],
    "notes": "Maiden name partially faded; confirmed by marriage record cross-re"
  }
}
]
}

```

19. Validation

ADAC-Genealogy containers are validated by the standard ADAC validator. The ADAC validator verifies:

- The manifest references the genealogy profile correctly (ADAC-050 error if the referenced profile file is missing).
- All region annotation files exist (ADAC-023 error if missing).
- SHA-256 checksums for all files including the genealogy profile (ADAC-082 on mismatch).

19.1 Profile-Specific Validation

This specification does **not** define additional validation error codes. Profile-level semantic validation is the responsibility of the consuming application.

19.2 Recommended Application-Level Checks

Check	Description
Page link master references	Verify that <code>pageLink.masterId</code> values exist in the manifest's master entries.
Record group completeness	Verify that all pages in a record group have contiguous, 1-based sequence numbers.
Linked entity consistency	Verify that regions with <code>genealogy:person</code> entities have non-empty <code>givenName</code> or <code>surname</code> .
Evidence classification validity	Verify that <code>classification</code> uses one of the well-known values (<code>direct</code> , <code>indirect</code> , <code>negative</code>).

20. References

20.1 Normative References

Reference	Description
ADAC 1.0 Format Specification	The base container format this profile extends.
RFC 2119	Key words for use in RFCs to Indicate Requirement Levels.

20.2 Informative References

Reference	Description
Genealogical Proof Standard	Board for Certification of Genealogists — the five-element framework for genealogical proof.
GEDCOM 5.5.1	The standard format for genealogical data interchange. ADAC-Genealogy supports GEDCOM XREF cross-references.
GEDCOM X	A modern, JSON-based genealogical data model. Informs several ADAC-Genealogy schema decisions.
IETF BCP 47	Tags for Identifying Languages. Recommended for language fields.
Dublin Core Metadata Element Set	Vocabulary for describing resources. ADAC core metadata aligns with Dublin Core.
ADAC-Legal 1.0 Format Specification	The legal profile extension for ADAC. Companion profile that can coexist with the genealogy profile.

21. Version History

Version	Date	Description
1.0	2026	Initial release. Source citation, person annotations, transcriptions, record conditions, multi-page record links, evidence analysis with GPS support, well-known value sets for record types, evidence classifications, legibility, and damage types.
