

Copyright and License

© 2026 InnoVadens, LLC. All rights reserved.

This specification is licensed under the Creative Commons Attribution–NoDerivatives 4.0 International License (CC BY-ND 4.0).

You are free to copy and redistribute this document in any medium or format, provided that you:

- Attribute the work to InnoVadens, LLC, and
- Do not remix, transform, or build upon the material.

No modifications of this specification may be distributed. Implementations may reference this specification, but derivative specifications or altered versions may not be published under the ADAC name or imply ADAC endorsement.

A copy of the license is available at: <https://creativecommons.org/licenses/by-nd/4.0/>

Status of This Document

This document defines the authoritative ADAC Specification for version 1.0.

InnoVadens, LLC is the steward of this specification. Only documents published at <https://adac-standard.com/specs/> are considered official and normative.

This specification may be referenced by third-party software, libraries, and services. However, no external party may publish modified versions of this specification or claim compatibility beyond what is explicitly defined in the normative text.

Normative requirements in this document use the keywords MUST, MUST NOT, SHOULD, SHOULD NOT, and MAY as defined in RFC 2119.

Trademark Notice

"ADAC" and the ADAC logo are trademarks of InnoVadens, LLC. Use of the ADAC name in derivative specifications, competing standards, or modified versions of this document is prohibited.

Software implementations may state that they "implement the ADAC Specification" provided such statements are accurate and not misleading.

Revision History

Version 1.0.0 — Initial publication.

ADAC-Genealogy v1.0.0 — Genealogy Profile Specification

Editor & Steward

John Vaden
Genealogist and Technical Workflow Contributor
InnoVadens, LLC — Cedar Lake, Michigan, USA

Contributors

AI-assisted drafting and refinement
Community feedback (ongoing)

Status

This document is maintained by InnoVadens, LLC as the official steward of the ADAC Standard.

John Vaden has decades of experience in genealogical research and in contributing to standards, workflows, and data practices across analyst and software development teams. He serves as the editor and steward of the ADAC Standard. InnoVadens, LLC is the legal steward of ADAC.

Version: v1.0.0 **Status:** Stable **Copyright:** © 2026 InnoVadens, LLC. All rights reserved. **Date:** 2026 **Base Format:** [ADAC 1.0]

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, per-region annotations for persons, transcriptions, and document condition, and explicit handling of living-person privacy concerns — 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. [Privacy and Living Persons](#)
 15. [Well-Known Value Sets](#)
 16. [JSON Conventions](#)
 17. [Media Types & File Extensions](#)
 18. [Interoperability](#)
 19. [Complete Container Example](#)
 20. [Validation](#)
 21. [References](#)
 22. [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. Researchers also face a recurring challenge with living-person privacy: census records become public after 72 years in the United States, but people named on more recent records may still be living and their personal information may be subject to privacy law (GDPR, CCPA, and state-level statutes).

1.2 Solution

ADAC-Genealogy solves these problems 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, with optional information-quality classification per person.
- **Transcriptions** — per-region text extracted from handwriting or print, with confidence scores for AI-assisted work and optional revision history.
- **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.
- **Privacy controls** — explicit guidance and schema support for living-person data, including integration with ADAC's `accessControl` and `EncryptionDescriptor` mechanisms.

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, ages, and historical demographic descriptors use free-text strings to faithfully represent the source record. Genealogical documents routinely contain approximate, partial, or non-standard values that structured types cannot capture, and historical demographic categories (race as recorded, sex as recorded) reflect the language of the original document, not modern classification.
Privacy-aware	The profile provides explicit guidance for handling living-person data and integrates with ADAC's access control and encryption mechanisms.
Ecosystem-interoperable	Supports GEDCOM 7 cross-references, GEDCOM 5.5.1 XREFs (legacy), 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
- Privacy guidance for living-person data
- Well-known value sets for record types, evidence classifications, source classifications, information classifications, legibility values, and damage types
- Profile-specific validation rules

This specification does **not** define:

- The ADAC core container format (see the [ADAC 1.0 Format Specification](#))
- DNA evidence or genetic genealogy data structures (deferred to a future companion profile)
- 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 at any conformance level (Minimal, Archival, or Signed Archival).
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 `profileId`, `profileType`, and `profileVersion` fields, with `profileType` equal to "genealogy" and `profileId` equal to "urn:adac:profile:genealogy:v1".

2.2 Recommended Conformance for Genealogical Archives

For long-term preservation of genealogical records, **Archival** conformance is RECOMMENDED. For records intended to serve as legal evidence (probate, land transfer, naturalization records used in citizenship claims), **Signed Archival** conformance is RECOMMENDED so that chain of custody can be cryptographically verified.

2.3 Forward Compatibility

Readers MUST tolerate unknown JSON properties in all genealogy profile structures. Unknown properties MUST be preserved during round-trip serialization. Writers that cannot preserve unknown fields MUST set the manifest's `roundTripFidelity` to "degraded" and append a `fidelityLoss` provenance event, in accordance with ADAC 1.0 Section 6.2.

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 (such as a household enumeration spanning two census sheets).
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.
Source Classification	Whether the source itself is original (created at the time of the event) or derivative (a later copy or compilation).
Information Classification	Whether the information within the source is primary (recorded by an eyewitness or participant) or secondary (recorded by someone with second-hand knowledge).
Evidence Classification	Whether information from the source is direct , indirect , or negative evidence for a specific genealogical claim.
Collection Hierarchy	The standard archival arrangement: collection → series → box → folder → item.
Film Reference	Microfilm roll, frame, and FamilySearch Family History Library (FHL) film number.
Living Person	A person who is, or may reasonably be presumed to be, alive at the time of container creation or distribution. See Section 14 for handling guidance.

4. Relationship to ADAC

4.1 Extension Mechanisms

ADAC-Genealogy extends ADAC through three primary 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.
3. **Provenance events** — Genealogy applications SHOULD append events to ADAC's hash-chained provenance log (`provenance/log.json`) when significant genealogical analysis actions occur (transcription added, evidence analyzed, person annotation created or modified, container correlated to another source). Recommended event types include `genealogyTranscription`, `genealogyAnalysis`, and `genealogyCorrelation`. These appear in the same chronological hash chain as ADAC's built-in events and inherit the same tamper-evident properties.

4.2 Optional ADAC Features Used by This Profile

ADAC-Genealogy MAY make use of the following optional ADAC 1.0 features:

ADAC Feature	Genealogy Use Case
<code>metadata/structure.json</code>	Generic structural relationships between masters (paged-document sequences). Genealogy-specific multi-page record links live in the genealogy profile (Section 8) and add semantic grouping on top of the generic structure.
<code>signatures/</code>	Cryptographic chain of custody for records used as legal evidence (probate, naturalization, land transfer).
<code>accessControl</code> (per-master, per-derivative, or container-level)	Tagging records or specific derivatives that contain living-person data with confidentiality levels and embargo dates. See Section 14.
<code>EncryptionDescriptor</code>	Storing recent records with living-person data as encrypted ciphertext when the container is shared outside controlled environments. See Section 14.
Hash-chained provenance log	Tamper-evident audit trail of transcription work, evidence analysis revisions, and cross-container correlation.

4.3 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 schema or hash-chain mechanism
- The checksum manifest schema or fixity computation
- The signature mechanism or signature index schema
- 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)
│   ├── structure.json         (OPTIONAL - 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/                 (REQUIRED for Archival ADAC)
│   ├── log.json              (hash-chained events)
│   └── checksums.json
└── signatures/                 (OPTIONAL - REQUIRED for Signed Archival)
    ├── index.json
    └── *.p7s

```

5.3 Manifest Integration

```
{
  "metadata": {
    "core": "metadata/core.json",
    "structure": "metadata/structure.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 uses the standard ADAC 1.0 profile wrapper (Section 13.2 of the ADAC core specification) to declare its identity, then carries genealogy-specific content under the data field.

6.1 Schema

```
{
  "profileId": "urn:adac:profile:genealogy:v1",
  "profileType": "genealogy",
  "profileVersion": "1.0.0",
  "title": "ADAC Genealogy Profile",
  "schema": "https://adac.example.org/schemas/genealogy-v1.json",
  "data": {
    "sourceCitation": { ... },
    "pageLinks": [ ... ],
    "evidence": { ... }
  }
}
```

6.2 Profile Wrapper Fields

Field	Type	Required	Value	Description
profileId	string	REQUIRED	"urn:adac:profile:genealogy:v1"	URN identifying this specific profile and version.

Field	Type	Required	Value	Description
profileType	string	REQUIRED	"genealogy"	The profile type identifier. MUST be "genealogy".
profileVersion	string	REQUIRED	"1.0.0" for this spec	Semantic version (MAJOR.MINOR.PATCH).
title	string	OPTIONAL	—	Human-readable title for display.
schema	string	OPTIONAL	—	URI of a JSON Schema document that validates the data subobject.
data	GenealogyData	OPTIONAL	—	The genealogy-specific payload. See Section 6.3.

6.3 GenealogyData Fields

The data subobject carries the genealogy-specific content:

Field	Type	Required	Description
sourceCitation	SourceCitation	OPTIONAL	Archival source information (see Section 7).
pageLinks	PageLink[]	OPTIONAL	Multi-page record links (see Section 8).
evidence	EvidenceAnalysis	OPTIONAL	GPS evidence classification and analysis (see Section 9).

6.4 Forward Compatibility

Unknown JSON properties at any level (within the wrapper or within data) MUST be preserved during round-trip serialization.

7. Source Citation

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

7.1 Schema

```
{
  "sourceCitation": {
```

```

"recordType": "census",
"sourceClassification": "original",
"repository": "National Archives, Washington, D.C.",
"collection": {
  "collection": "T9",
  "series": "1870 Federal Census",
  "box": null,
  "folder": null,
  "item": "Roll 1042"
},
"film": {
  "roll": "1042",
  "frame": "0142",
  "fhlFilmNumber": "0553042"
},
"pageNumber": "42",
"folio": null,
"entryNumber": "household 42",
"url": "https://www.familysearch.org/ark:/61903/3:1:S3HY-X9BS-F6R",
"jurisdiction": {
  "country": "United States",
  "state": "Ohio",
  "county": "Licking",
  "town": "Newark",
  "historicalContext": "Licking County formed 1808 from Fairfield County."
},
"dateRangeStart": "1870-06-01",
"dateRangeEnd": "1870-08-31",
"languages": ["en"]
}
}

```

7.2 Source Citation Fields

Field	Type	Required	Description
recordType	string	OPTIONAL	The genealogical record type. See Section 15.1 for well-known values.
sourceClassification	string	OPTIONAL	GPS source classification: "original" (created at the time of the event) or "derivative" (a later copy, transcript, or compilation). See Section 15.2 .
repository	string	OPTIONAL	The archival repository or holding institution.
collection	CollectionHierarchy	OPTIONAL	Collection / series / box / folder / item hierarchy (see Section 7.3).
film	FilmReference	OPTIONAL	Microfilm or FHL film reference (see Section 7.4).
pageNumber	string	OPTIONAL	Page number as printed or stamped on the source document.

Field	Type	Required	Description
folio	string	OPTIONAL	Folio number for early registers and bound volumes.
entryNumber	string	OPTIONAL	Entry, household, or item number within the page (e.g., "household 42", "entry 17").
url	string	OPTIONAL	URL to an online source (FamilySearch, Ancestry, FindMyPast, etc.).
jurisdiction	Jurisdiction	OPTIONAL	Geographic jurisdiction the record covers (see Section 7.5).
dateRangeStart	string	OPTIONAL	Start of the date range covered by the source (ISO-8601 date).
dateRangeEnd	string	OPTIONAL	End of the date range covered by the source (ISO-8601 date).
languages	string[]	OPTIONAL	Language(s) of the document. IETF BCP 47 tags RECOMMENDED (e.g., "en", "de", "la").

7.3 Collection Hierarchy

Describes the archival collection hierarchy: collection → series → box → folder → item. All fields are OPTIONAL; populate only the levels applicable to the record.

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

7.4 Film Reference

Describes microfilm, roll, and FHL (Family History Library) film references. All fields are OPTIONAL.

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

7.5 Jurisdiction

Describes the geographic jurisdiction the record covers, from broadest to most specific. The optional `historicalContext` field captures notes about the jurisdiction as it existed at the time the record was created. All fields are OPTIONAL.

```
{
  "country": "United States",
  "state": "Ohio",
  "county": "Licking",
  "town": "Newark",
  "historicalContext": "Licking County formed 1808 from Fairfield County."
}
```

Field	Type	Description
<code>country</code>	<code>string</code>	The country.
<code>state</code>	<code>string</code>	The state, province, or equivalent administrative division.
<code>county</code>	<code>string</code>	The county, parish, or equivalent subdivision.
<code>town</code>	<code>string</code>	The town, city, or municipality.
<code>historicalContext</code>	<code>string</code>	Free-text notes about the jurisdiction as it existed at the time of the record (boundary changes, parent jurisdictions, name changes).

8. Multi-Page Record Links

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

8.1 Relationship to ADAC `structure.json`

ADAC 1.0 defines a generic mechanism (`metadata/structure.json`) for expressing structural relationships between masters such as paged documents. ADAC-Genealogy `pageLinks` are complementary — they add genealogy-specific semantics (the concept of a record group, such as a single household within a multi-household census page) on top of the generic page-ordering relationships in `structure.json`. A container MAY use both mechanisms.

8.2 Schema

```
{
  "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 – continuation"
  }
]
}

```

8.3 Fields

Field	Type	Required	Description
recordGroupId	string	REQUIRED	Groups pages of the same logical record. All masters belonging to the same record MUST share this value. MUST be non-empty.
pageSequence	integer	REQUIRED	1-based reading order within the record group. MUST start at 1 with consecutive integers.
masterId	string	REQUIRED	The master identifier within this container. MUST match a master id in the manifest.
description	string	OPTIONAL	Describes this page's role in the record.

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. The evidence analysis at the profile level captures the overall conclusion drawn from the entire source. Per-claim or per-person evidence classification is captured at the region level via `genealogy:person` (see Section 11.4).

9.1 Schema

```

{
  "evidence": {
    "claim": "John Smith born in Ohio circa 1828",
    "classification": "direct",
    "informationClassification": "primary",
    "correlationNotes": [
      {
        "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
        "containerId": "550e8400-e29b-41d4-a716-446655440099",
        "note": "Same household head listed 10 years earlier confirms continuity of

```

```

residence."
    }
  ],
  "conclusion": "John Smith, age 42, farmer, born Ohio circa 1828.",
  "analysisText": "Direct evidence from the 1870 census enumeration lists John Smith as
head of household, age 42, farmer, born in Ohio. Corroborated by the 1860 census showing the
same individual at age 32 in the same township.",
  "analyst": "Jane Researcher",
  "analyzedOn": "2026-02-01T14:00:00Z"
}
}

```

9.2 Evidence Analysis Fields

Field	Type	Required	Description
claim	string	OPTIONAL	The specific genealogical claim this analysis addresses. RECOMMENDED to make classification meaningful.
classification	string	OPTIONAL	The evidence classification (direct/indirect/negative). See Section 15.3 .
informationClassification	string	OPTIONAL	Whether the information within the source is "primary" or "secondary". See Section 15.4 .
correlationNotes	CorrelationNote[]	OPTIONAL	Notes correlating this source with other evidence.
conclusion	string	OPTIONAL	The researcher's conclusion drawn from this source.
analysisText	string	OPTIONAL	The full analysis text explaining the reasoning.
analyst	string	OPTIONAL	The researcher who performed the evaluation.
analyzedOn	string	OPTIONAL	ISO-8601 timestamp of when the analysis was performed.

9.3 Correlation Note

```

{
  "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
  "containerId": "550e8400-e29b-41d4-a716-446655440099",
  "note": "Same household head listed 10 years earlier confirms continuity."
}

```

Field	Type	Required	Description
sourceDescription	string	REQUIRED	A human-readable description of the correlated source.
containerId	string	OPTIONAL	The ADAC container id (UUID from the manifest's id field) for the correlated source. When present, MUST be a valid UUID (RFC 4122) matching the manifest id of the referenced container. UUIDs are required so that cross-container references remain reliable across renames and repository transfers.
note	string	REQUIRED	Text explaining how the two sources relate.

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, with optional revision history (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 unchanged.

10.3 Multiple Entities per Region

A single region MAY have all three genealogy linked entities simultaneously. 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,
  "sexAsRecorded": "M",
  "raceAsRecorded": "W",
  "birthDate": "1828",
  "deathDate": null,
  "marriageDate": null,
  "age": "42",
  "occupation": "Farmer",
  "birthplace": "Ohio",
  "relationshipToHead": "Self",
  "isLiving": false,
  "evidenceClassification": "direct",
  "informationClassification": "primary",
  "gedcom7Id": "@I123@",
  "gedcomXref": "@I123@",
  "externalId": "FSFT-ABC-1234"
}
```

11.2 Fields

Field	Type	Required	Description
givenName	string	OPTIONAL	Given (first) name as recorded.
surname	string	OPTIONAL	Family name as recorded.
maidenName	string	OPTIONAL	Maiden name, if applicable and recorded.
sexAsRecorded	string	OPTIONAL	Sex as recorded in the source document (e.g., "M", "F", "male", "female"). This field captures the historical record verbatim and is NOT a statement of modern gender identity.
raceAsRecorded	string	OPTIONAL	Race or ethnicity as recorded in the source document (e.g., "W", "B", "Mu", "Indian", free text). This field captures the historical categorization used by the original document, including outdated or offensive terminology. It is preserved for source fidelity and is NOT an endorsement of those categories. Applications displaying this field SHOULD provide context indicating that this reflects historical record-keeping, not modern classification.
birthDate	string	OPTIONAL	Birth date as recorded. Free-text to accommodate partial dates (e.g., "1828", "circa 1828", "15 Mar 1828", "12/23 Feb 1731/2" for dual-calendar dates).

Field	Type	Required	Description
deathDate	string	OPTIONAL	Death date as recorded.
marriageDate	string	OPTIONAL	Marriage date as recorded.
age	string	OPTIONAL	Age at the time of the record. String to accommodate ranges and notations (e.g., "42", "about 40", "6/12" for 6 months, "under 5").
occupation	string	OPTIONAL	Occupation at the time of the record.
birthplace	string	OPTIONAL	Birthplace as recorded.
relationshipToHead	string	OPTIONAL	Relationship to the head of household (census records).
isLiving	boolean	OPTIONAL	Whether this person is, or may reasonably be presumed to be, alive. When <code>true</code> , see Section 14 for required handling. Default if absent: <code>false</code> for records older than 100 years; otherwise undefined.
evidenceClassification	string	OPTIONAL	The evidence classification specifically for this person's identification within this source. See Section 15.3 .
informationClassification	string	OPTIONAL	Whether the information about this person is "primary" or "secondary". See Section 15.4 .
gedcom7Id	string	OPTIONAL	GEDCOM 7 record identifier. New implementations SHOULD prefer this field.
gedcomXref	string	OPTIONAL	GEDCOM 5.5.1 XREF identifier. Retained for backward compatibility with legacy genealogy databases.
externalId	string	OPTIONAL	External system identifier (e.g., FamilySearch PID, Ancestry tree person ID).

11.3 Design Rationale — String Dates

Date fields use `string` rather than structured date types. 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.

11.4 Per-Person Evidence Classification

The GPS distinguishes evidence at the level of individual claims, not entire sources. A single census page may provide direct evidence for one person's age but only indirect evidence for another's place of birth. ADAC-Genealogy therefore supports evidence classification at two levels:

- **Source-level** classification (in `evidence.classification`, Section 9): the overall character of the source for the primary claim under analysis.
- **Person-level** classification (in `genealogy:person.evidenceClassification`): the specific classification for each person's identification within the source.

11.5 Source Fidelity for Demographic Fields

The `sexAsRecorded` and `raceAsRecorded` fields exist to faithfully preserve the source record. Historical documents used categories that are now considered inaccurate, offensive, or scientifically unfounded. ADAC-Genealogy does not sanitize or modernize these terms because doing so would falsify the historical record. Applications that display these fields are responsible for providing appropriate context for end users. If modern interpretations are needed, store them in separate fields using custom linked entities, not by overwriting the as-recorded values.

12. Transcription

Linked Entity Key: `genealogy:transcription`

Transcriptions capture the text read from a specific region, whether performed manually by a researcher or automatically by OCR/AI handwriting recognition. Transcriptions support optional revision history so successive corrections are preserved.

12.1 Schema

```
{
  "text": "Smith, John 42 M W Farmer 2000 Ohio",
  "confidence": 0.95,
  "transcriber": "GeneaScan AI v2.0",
  "transcribedOn": "2026-01-15T11:00:00Z",
  "originalLanguage": "en",
  "translatedText": null,
  "translatedLanguage": null,
  "revisions": [
    {
      "text": "Smith, John 42 M W Farmer ?000 Ohio",
      "confidence": 0.78,
      "transcriber": "GeneaScan AI v1.0",
      "transcribedOn": "2025-08-10T09:30:00Z",
      "supersededBy": "2026-01-15T11:00:00Z",
      "supersedeReason": "Re-transcribed with improved AI model."
    }
  ]
}
```

12.2 Fields

Field	Type	Required	Description
<code>text</code>	<code>string</code>	REQUIRED	The current (most recent) transcribed text from the region.

Field	Type	Required	Description
confidence	number	OPTIONAL	AI confidence score (0.0–1.0). Omit (or null) for manual transcription.
transcriber	string	OPTIONAL	The person or system that produced the current transcription.
transcribedOn	string	OPTIONAL	ISO-8601 timestamp of when the current transcription was performed.
originalLanguage	string	OPTIONAL	Language of the original text (IETF BCP 47 tag).
translatedText	string	OPTIONAL	Translated text, if the original is in a different language.
translatedLanguage	string	OPTIONAL	Language of the translation (IETF BCP 47 tag).
revisions	TranscriptionRevision[]	OPTIONAL	Ordered list of prior transcription versions, oldest first. See Section 12.3.

12.3 Transcription Revisions

When a transcription is corrected or re-performed, the prior version SHOULD be preserved in the `revisions` array.

Field	Type	Required	Description
text	string	REQUIRED	The prior transcription text.
confidence	number	OPTIONAL	The prior confidence score, if applicable.
transcriber	string	OPTIONAL	The person or system that produced the prior transcription.
transcribedOn	string	OPTIONAL	ISO-8601 timestamp of when the prior transcription was performed.
supersededBy	string	OPTIONAL	ISO-8601 timestamp when this revision was superseded.
supersedeReason	string	OPTIONAL	Free-text reason for the supersession.

13. Record Condition

Linked Entity Key: `genealogy:condition`

Record condition assessments document the physical state and legibility of a specific region. 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 15.5 .
damageTypes	string[]	OPTIONAL	null	List of observed damage types. See Section 15.6 .
notes	string	OPTIONAL	null	Free-text condition notes.

14. Privacy and Living Persons

Genealogical research routinely encounters records containing information about people who may still be alive. Such records may be subject to data protection laws (GDPR, CCPA, state-level statutes) or ethical considerations independent of legal requirements.

14.1 The `isLiving` Flag

The `isLiving` field on `genealogy:person` indicates whether the person is, or may reasonably be presumed to be, alive. When `true`, the consuming application SHOULD apply additional handling consistent with applicable privacy law and institutional policy. For records older than 100 years, all named individuals can typically be presumed deceased and `isLiving` defaults to `false`. For more recent records, `isLiving` SHOULD be set explicitly.

14.2 Recommended Container Patterns for Living-Person Records

Scenario	Recommended Pattern
Research-only, controlled environment	Set <code>isLiving: true</code> on relevant person annotations. No additional encryption required if access is controlled at the storage layer.
Container shared outside controlled environment	Apply <code>accessControl</code> (ADAC core Section 12) at the container, master, or derivative level with <code>confidentialityLevel: "restricted"</code> and an <code>embargoUntil</code> date if appropriate.
Container containing PII subject to GDPR/CCPA	Encrypt master files using ADAC's <code>EncryptionDescriptor</code> mechanism (ADAC core Section 20) with AES-256-GCM and KMS-managed keys.

Scenario	Recommended Pattern
Container intended for public release	Generate redacted derivatives masking living-person data; keep originals encrypted; reference both in the manifest.

14.3 Right to Erasure

GDPR Article 17 recognizes a right to erasure. ADAC's master immutability principle is in fundamental tension with this right. Implementations that need to support right-to-erasure requests SHOULD: (1) remove the affected container from active distribution; (2) generate new derivatives with the affected data redacted; (3) record the erasure request and response as `redaction` events in the provenance log; and (4) document the original master's continued existence and access controls in institutional records external to the container. Compliance is the responsibility of the institution holding the container.

15. 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.

15.1 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.
"newspaper"	Newspaper article, obituary, or notice.
"directory"	City directory, phone book, or business directory.

Value	Description
"slaveSchedule"	U.S. slave schedule (1850, 1860).

15.2 Source Classifications

Used in `sourceCitation.sourceClassification`.

Value	Description
"original"	The source was created at the time of the event being researched.
"derivative"	The source is a later copy, transcript, abstract, index, or compilation.

15.3 Evidence Classifications

Used in `evidence.classification` and `genealogy:person.evidenceClassification`.

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.

15.4 Information Classifications

Used in `evidence.informationClassification` and `genealogy:person.informationClassification`.

Value	Description
"primary"	The information was recorded by an eyewitness or participant in the event.
"secondary"	The information was recorded by someone with second-hand knowledge.
"undetermined"	It is not clear whether the informant had direct knowledge.

15.5 Legibility Values

Used in record condition `legibility`.

Value	Description
"clear"	Fully readable without difficulty.
"partial"	Partially readable with some effort required.
"poor"	Mostly unreadable but some text is discernible.

Value	Description
"illegible"	Completely unreadable.

15.6 Damage Types

Used in record condition `damageTypes`.

Value	Description
"torn"	Physical tear in the document.
"faded"	Ink or print has faded.
"waterDamage"	Water damage causing staining or warping.
"fireDamage"	Fire or heat damage.
"foxing"	Brown spots from oxidation.
"moLd"	Mold or mildew damage.
"insectDamage"	Insect damage (bookworm holes, etc.).
"bleedThrough"	Ink from the reverse side bleeding through.
"binding"	Text obscured by binding edge or gutter.

16. 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 (except free-text genealogical date fields, which preserve the source)
UUID format	RFC 4122, lowercase hexadecimal

17. 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.

18. Interoperability

18.1 GEDCOM Integration

The `gedcom7Id` field on person annotations links person regions to records in a GEDCOM 7 file (the current standard, released by FamilySearch in 2021). The legacy `gedcomXref` field provides backward compatibility with GEDCOM 5.5.1 databases.

18.2 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.

18.3 Profile Coexistence

ADAC-Genealogy containers MAY include additional profiles. A probate document could carry both a genealogy profile (source citation, person annotations, transcriptions) and a legal profile (case reference, custody chain, confidentiality level). The profiles coexist without conflict.

18.4 Cross-Container Evidence Networks

The `containerId` field in correlation notes enables researchers to build evidence networks across multiple containers. Because `containerId` is required to be a valid UUID matching a manifest `id`, these references remain reliable even when containers are renamed or moved between repositories.

18.5 Future DNA Profile

Genetic genealogy and DNA evidence are out of scope for ADAC-Genealogy 1.0. A future companion specification (such as ADAC-DNA) may define structures for storing DNA test results, match data, and centimorgan analyses. Until such a specification exists, implementations needing to attach DNA-related metadata SHOULD use custom linked entities under a reverse-domain namespace (e.g., `com.example.dna:match`).

19. Complete Container Example

19.1 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
│  ├─ structure.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
```

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

```
{
  "profileId": "urn:adac:profile:genealogy:v1",
  "profileType": "genealogy",
  "profileVersion": "1.0.0",
  "title": "ADAC Genealogy Profile",
  "data": {
    "sourceCitation": {
      "recordType": "census",
      "sourceClassification": "original",
      "repository": "National Archives, Washington, D.C.",
      "collection": {
        "collection": "T9",
        "series": "1870 Federal Census",
        "item": "Roll 1042"
      },
    },
    "film": {
      "roll": "1042",
      "frame": "0142",
      "fhlFilmNumber": "0553042"
    },
    "pageNumber": "42",
    "entryNumber": "household 42",
    "url": "https://www.familysearch.org/ark:/61903/3:1:S3HY-X9BS-F6R",
    "jurisdiction": {
      "country": "United States",
    },
  },
}
```

```

    "state": "Ohio",
    "county": "Licking",
    "town": "Newark",
    "historicalContext": "Licking County formed 1808 from Fairfield County."
  },
  "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": {
  "claim": "John Smith born in Ohio circa 1828",
  "classification": "direct",
  "informationClassification": "primary",
  "correlationNotes": [
    {
      "sourceDescription": "1860 U.S. Federal Census, Licking County, Ohio",
      "containerId": "550e8400-e29b-41d4-a716-446655440099",
      "note": "Same household head listed 10 years earlier confirms continuity of residence."
    }
  ],
  "conclusion": "John Smith, age 42, farmer, born Ohio circa 1828.",
  "analysisText": "Direct evidence from the 1870 census enumeration lists John Smith as head of household, age 42, farmer, born in Ohio. Corroborated by the 1860 census showing the same individual at age 32 in the same township.",
  "analyst": "Jane Researcher",
  "analyzedOn": "2026-02-01T14:00:00Z"
}
}
}

```

19.3 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",
        "sexAsRecorded": "M",
        "raceAsRecorded": "W",
        "birthDate": "1828",
        "age": "42",
        "occupation": "Farmer",
        "birthplace": "Ohio",
        "relationshipToHead": "Self",
        "isLiving": false,
        "evidenceClassification": "direct",
        "informationClassification": "primary",
        "gedcom7Id": "@I123@"
      },
      "genealogy:transcription": {
        "text": "Smith, John 42 M W Farmer 2000 Ohio",
        "confidence": 0.95,
        "transcriber": "GeneaScan AI v2.0",
        "transcribedOn": "2026-01-15T11:00:00Z",
        "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",
        "sexAsRecorded": "F",
        "raceAsRecorded": "W",
        "birthDate": "1832",

```

```

    "age": "38",
    "occupation": "Keeping house",
    "birthplace": "Virginia",
    "relationshipToHead": "Wife",
    "isLiving": false,
    "evidenceClassification": "direct",
    "informationClassification": "secondary"
  },
  "genealogy:transcription": {
    "text": "Smith, Mary 38 F W Keeping house Virginia",
    "confidence": 0.88,
    "transcriber": "GeneaScan AI v2.0",
    "transcribedOn": "2026-01-15T11:00:00Z",
    "originalLanguage": "en"
  },
  "genealogy:condition": {
    "legibility": "partial",
    "damageTypes": ["faded"],
    "notes": "Maiden name partially faded; confirmed by marriage record cross-
reference."
  }
}
]
}

```

20. Validation

ADAC-Genealogy containers are validated by the standard ADAC validator, which 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), and SHA-256 checksums for all files including the genealogy profile (ADAC-082 on mismatch).

20.1 Profile-Specific Validation Rules

In addition to standard ADAC validation, validators MAY enforce the following ADAC-Genealogy-specific rules. These rules SHOULD be implemented as warnings rather than errors so that legitimate edge cases are not rejected.

Rule	Severity	Description
GENL-001	Error	The profile file's <code>profileType</code> is not "genealogy".
GENL-002	Error	The profile file's <code>profileId</code> is not "urn:adac:profile:genealogy:v1".

Rule	Severity	Description
GENL-010	Warning	A <code>pageLink</code> has an empty or missing <code>recordGroupId</code> .
GENL-011	Warning	A <code>pageLink</code> has a <code>pageSequence</code> value that is not a positive integer.
GENL-012	Warning	A <code>pageLink.masterId</code> does not match any master entry in the manifest.
GENL-013	Warning	A record group has non-contiguous or non-1-based <code>pageSequence</code> values.
GENL-020	Warning	A <code>genealogy:transcription.confidence</code> value is outside the range 0.0–1.0.
GENL-021	Warning	A <code>genealogy:person</code> entity has none of <code>givenName</code> , <code>surname</code> , or <code>relationshipToHead</code> set.
GENL-022	Warning	A <code>genealogy:person.evidenceClassification</code> value is not in the well-known set (<code>direct</code> , <code>indirect</code> , <code>negative</code>).
GENL-023	Warning	A <code>genealogy:person.informationClassification</code> value is not in the well-known set (<code>primary</code> , <code>secondary</code> , <code>undetermined</code>).
GENL-030	Warning	A <code>correlationNote.containerId</code> is present but is not a valid UUID.
GENL-040	Warning	A <code>genealogy:person</code> entity has <code>isLiving: true</code> but the container has no <code>accessControl</code> metadata and no encryption descriptor on the master file.
GENL-050	Info	The <code>sourceCitation.recordType</code> value is not in the well-known set. (Custom values are permitted.)

21. References

21.1 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.
RFC 4122	UUID URN Namespace. Required format for <code>containerId</code> values.
ISO 8601	Date and time representations. Used for all timestamp fields.
IETF BCP 47	Tags for Identifying Languages. Used for language fields.

21.2 Informative References

Reference	Description
Genealogical Proof Standard	Board for Certification of Genealogists — the five-element framework for genealogical proof.
Evidence Explained	Elizabeth Shown Mills, <i>Evidence Explained: Citing History Sources from Artifacts to Cyberspace</i> . The standard reference for genealogical citation.
GEDCOM 7	The current standard format for genealogical data interchange (FamilySearch, 2021). ADAC-Genealogy supports GEDCOM 7 cross-references via <code>gedcom7Id</code> .
GEDCOM 5.5.1	The legacy genealogical data interchange format. ADAC-Genealogy supports GEDCOM 5.5.1 XREF cross-references via <code>gedcomXref</code> for backward compatibility.
Dublin Core Metadata Element Set	Vocabulary for describing resources. ADAC core metadata aligns with Dublin Core.
GDPR Article 17	Right to erasure. Relevant to handling living-person data.
ADAC-Legal 1.0 Profile Specification	The legal profile extension for ADAC. Companion profile that can coexist with the genealogy profile.
ADAC-Preservation 1.0 Profile Specification	The Preservation profile extension for ADAC. Companion profile for institutional archival management.

About the Editor

ADAC is edited and maintained by **John Vaden**, a genealogist and technical workflow contributor with decades of experience in genealogical research and in shaping data practices and process standards across analyst and software development teams. His work on ADAC emphasizes clarity, interoperability, and practical implementation for genealogists, developers, and archival practitioners. InnoVadens, LLC serves as the legal steward of the ADAC Standard.

Document History

- v1.0 — Initial release by Editor (John Vaden); AI-assisted drafting and refinement