

Source Inventory Template

Purpose

The Source Inventory is the canonical register of every input a GROW-built system is allowed to read, retrieve, embed, summarize, or cite. Nothing enters lineage that is not first listed here. If an agent reaches for data not on the inventory, that is a provenance violation and a reliability incident, not a data-quality problem to solve at output time.

A complete inventory answers four questions for every byte the system touches:

- 1. Where did it come from and who owns it?
- 2. Is it authoritative for the claim we will use it to support?
- 3. How fresh is it, and how often can it move under us?
- 4. What source confidence do we attach to outputs derived from it?

The confidence_band field on every row is the value supplied to s2-scoring-system per locked contract C4.

Required Fields

Field	Type	Required	Notes					
source_id	string (kebab-case)	yes	Stable across versions; never renumbered.					
name	string	yes	Human-readable.					
type	enum	yes	dataset \	document \	api \	connector \	upload \	generated
owner	string	yes	Role or named individual accountable for the source.					
authoritative	bool	yes	True only if this is the source-of-record for the claim type.					
update_frequency	string	yes	E.g., realtime, hourly, daily, weekly, event-driven, static.					

Field	Type	Required	Notes					
last_validated	date (ISO 8601)	yes	Last time owner confirmed the source is fit-for-purpose.					
permissions	string	yes	Access scope; public \	internal \	restricted \	regulated .		
confidence_band	enum	yes	high \	medium \	low \	unknown		
band_rationale	string	yes	One sentence explaining why this source carries its assigned confidence_band. Per locked contract C4, this is a structured per-row field, not narrative prose.					

Optional but recommended: retention_class, pii_flag, jurisdiction, cost_per_call, rate_limit, change_log_url.

Worked Example — Doug's Operating Stack

The following is a real-shape inventory for a multi-surface workload spanning a Florida special district seat, a 30A STR campus, and a podcast production pipeline.

source_id	name	type	owner	authoritative	update_frequency	last_validated	permissions	confidence_band	band_rationale
fl-dep-lpa0381	FL DEP Grant LPA0381 disbursement ledger	dataset	District Treasurer	true	weekly	2026-05-12	regulated	high	Statutory system-of-record for grant disbursement; reconciled weekly by the treasurer against bank settlement.

source_id	name	type	owner	authoritative	update_frequency	last_validated	permissions	confidence_band	band_rationale
wcpa-parcels	Walton County Property Appraiser parcel records	api	County GIS	true	daily	2026-05-20	public	high	County-published parcel system-of-record under FS Ch. 193; refreshed daily with documented change log.
ownerrez-bookings	OwnerRez bookings + payouts feed (Point Preserve)	connector	STR Ops	true	hourly	2026-05-25	internal	high	Booking system-of-record for confirmed reservations; high for booking events themselves, downgraded for revenue claims until reconciled with Airbnb.
aifg-transcripts	AI for Good podcast Otter/De script transcripts	document	Producer	false	event-driven	2026-05-22	internal	medium	Machine-generated transcript with measurable named-entity and timestamp drift; safe for topic-level reuse but not verbatim quotation without review.

source_id	name	type	owner	authoritative	update_frequency	last_validated	permissions	confidence_band	band_rationale
slack-pp-archive	Point Preserve internal Slack workspace export	dataset	Founder	false	weekly	2026-04-30	restricted	medium	Conversational and informal; not authoritative for any external claim, acceptable only as supporting context behind a primary source.
ecoguardian-stream	EcoGuardian sensor telemetry (water quality, turbidity)	api	EcoGuardian.AI	true	realtime	2026-05-27	internal	medium	Authoritative sensor feed but calibration window is open; high only post-calibration, medium during the current verification cycle.
fasd-bulletins	FL Association of Special Districts bulletin archive	document	Commissioner's office	false	weekly	2026-05-15	public	medium	Reputable secondary commentary on statute and practice, but not a primary legal source; cite as context, not authority.

source_id	name	type	owner	authoritative	update_frequency	last_validated	permissions	confidence_band	band_rationale
airbnb-csv-dump	Manual Airbnb reservation CSV exports	upload	STR Ops	false	monthly	2026-04-01	internal	low	Manual export with monthly latency and known overlap with ownerrez-bookings; usable only as a cross-check, never standalone.

Why these confidence bands

The band_rationale column above is the canonical, per-row justification consumed by s2-scoring-system under contract C4. The notes below restate the operating intuition for builders reading the inventory the first time; they do not replace the structured field.

- fl-dep-lpa0381 and wcpa-parcels are statutory systems-of-record. High.
- ownerrez-bookings is the booking system-of-record, but reconciliation gaps with Airbnb mean medium-confidence for revenue claims — promoted to high only after reconciliation. Listed as high here for booking events themselves.
- aifg-transcripts are machine-generated with named-entity error; medium.
- slack-pp-archive is conversational, not authoritative for any external claim; medium at best, often low for factual reuse.
- airbnb-csv-dump is manual, late, and overlaps ownerrez-bookings for the same events; low when used standalone.

Operating Rules

1. New source → inventory row before first read. No backfilling after the fact.
2. last_validated older than the update_frequency window flips confidence_band down one notch automatically until re-validated.
3. authoritative=true is exclusive per claim type. Two authoritative sources for the same claim is a governance bug, not a redundancy.
4. Deprecation is a row update (status note), never a deletion — provenance records reference historical source_ids.
5. Quarterly review: owners re-attest, stale rows are downgraded, and the inventory diff is logged into the lineage map.