

Below are precise, copy-ready instructions for preparing **bible/metaphors.yml**. They mirror the rigor used in P1 (schemas) and add the fields/guidelines required for evaluation, retrieval, and controlled generation.

Purpose

Define **primary metaphors** (SOURCE→TARGET) and **bipolar axes** (two opposed poles on one concept) with affective priors, lexical cues (EN/FA), and compatibility rules. These entries drive: (1) retrieval (RAG), (2) expectation/"explosion" timing, (3) frame/Schema coherence checks, and (4) optional shot-planner biases.

What to Prepare (Creator-Curated)

A. Canonical metaphor list (IDs + definitions)

Prepare 12-24 items mixing *primary* and *bipolar*: - **Primary** (examples): `life_is_travel`, `knowledge_is_light`, `ignorance_is_darkness`, `emotion_is_temperature`, `intimacy_is_closeness`, `argument_is_war` (use carefully), `time_is_motion`, `ideas_are_food`, `importance_is_center`, `control_is_holding`, `difficulty_is_weight`. - **Bipolar axes** (examples): `raw_cooked`, `light_dark`, `near_far`, `open_closed`, `up_down`, `center_periphery`, `clean_dirty` (use ethically), `silent_loud`. For each, write a 1-2 sentence **definition** in your own words.

B. Source/Target anchoring

- **Primary**: link `source_schema` to an existing schema id (`path`, `container`, etc.).
- **Bipolar**: specify `axis` name and two **poles** (left/right). Optionally map each pole to schema tendencies (e.g., `raw → container+pressure(low)`; `cooked → balance`).

C. Affect & salience (0..1 unless noted)

- `valence` in `-1..+1` (neg→pos), `arousal` in `0..1`.
- Optional tags (all 0..1): `threat`, `comfort`, `tension`, `warmth`, `novelty`.
- Optional **pole_affect** for bipolar (per-pole overrides).

D. Lexical triggers (bilingual, lemmas w∈0..1)

- 5-20 lemmas each in **EN** and **FA**; weights in `[0, 1]`.
- Keep lemmas lowercase; no inflections. You may include short multi-word lemmas ("toward", "bridge").

E. Compatibility & control

- `coactivate_schemas` : schema ids that naturally co-activate.
- `preferred_frames` : frame ids where this metaphor is most at home.
- `banned_with` : metaphor ids disallowed together.
- `gating` : knobs for generation control:
- `beats_bias` : suggested beat for turn/reveal (e.g., `turn`, `reveal`).
- `expectation_increment` : 0..1 additive hint to the curve when active.

F. Examples (2–3 each)

- 1 short **text** example (1–2 lines) and optional **visual** cue list.

G. Governance metadata

- `provenance.source` (e.g., "SV_Extended v0.1"), `curator`, `license` (e.g., CC-BY), `confidence` (0..1).

File Structure & Validation Rules

- File header includes `version`.
- `id` must be **lower_snake_case**, globally unique.
- `type` \in {primary, bipolar}.
- **Primary** requires: `source_schema`, `target` (free text), optional `coactivate_schemas`, `preferred_frames`.
- **Bipolar** requires: `axis`, `bipolar`: [left, right]; optional `pole_affect` and `pole_bias`.
- All weights in [0,1]; `valence` \in [-1,1].
- Lexicon lemmas non-empty, lowercase; duplicates collapse; sort by weight desc during normalization.
- Cross-refs must exist: `source_schema` must be in `schemas.yml`; all `coactivate_schemas` and `preferred_frames` must exist.
- Keep examples short; no style words that contradict world rules.

Minimal YAML Template

```
version: "0.1.0"
metaphors:
  # === PRIMARY ===
  - id: life_is_travel
    type: primary
    definition: "We understand life in terms of a journey with
```

```

source→path→goal."
  source_schema: path          # must exist in schemas.yml
  target: life
  affect: { valence: 0.1, arousal: 0.6, tension: 0.3, novelty: 0.4 }
  lexicon:
    en: [{ lemma: "journey", w: 0.8 }, { lemma: "road", w: 0.6 }, { lemma:
"cross", w: 0.6 }]
    fa: [{ lemma: "سفر", w: 0.8 }, { lemma: "olj", w: 0.6 }, { lemma: "عبور",
w: 0.6 }]
  coactivate_schemas: [path, boundary, container]
  preferred_frames: [journey]
  gating:
    beats_bias: development
    expectation_increment: 0.15
  examples:
    text: ["we cross the years like bridges"]
    visual: ["bridge", "tunnel", "stairs"]
  provenance: { source: "SV_Extended v0.1", curator: "Mahyar", license: "CC-
BY", confidence: 0.85 }

# === BIPOLAR ===
- id: raw_cooked
  type: bipolar
  definition: "Meaning unfolds along the RAW→COOKED axis (readiness/
pressure)."
  axis: readiness
  bipolar: [raw, cooked]
  affect: { valence: -0.1, arousal: 0.7, threat: 0.6, warmth: 0.2 }
  pole_affect:
    raw: { valence: -0.2, arousal: 0.8, threat: 0.7 }
    cooked: { valence: 0.2, arousal: 0.5, warmth: 0.4 }
  lexicon:
    en: [{ lemma: "raw", w: 0.9 }, { lemma: "uncooked", w: 0.7 }, { lemma:
"cooked", w: 0.7 }]
    fa: [{ lemma: "خام", w: 0.9 }, { lemma: "نیپخته", w: 0.7 }, { lemma: "پخته",
w: 0.7 }]
  coactivate_schemas: [container, balance]
  preferred_frames: [union_separation]
  banned_with: []
  gating:
    beats_bias: turn
    expectation_increment: 0.25
  examples:
    text: ["her voice is still raw inside the sealed room"]
  provenance: { source: "SV_Extended v0.1", curator: "Mahyar", license: "CC-
BY", confidence: 0.9 }

```

```
# Optional global constraints
constraints:
  max_active_bipolar_axes: 2
  disallowed_pairs:
    - [life_is_travel, raw_cooked]
```

Spreadsheet (Optional) → YAML Columns

If you prefer to curate in Sheets/CSV first, use these columns (one row per metaphor):

```
id,type,definition,source_schema,axis,pole_left,pole_right,valence,arousal,threat,comfort,tension
w,lex_fa|
w,coactivate_schemas,preferred_frames,banned_with,beats_bias,expectation_increment,examples_text,
```

- `lex_en|w` / `lex_fa|w`: semicolon-separated pairs like `journey|0.8;road|0.6`.

Curation Workflow

1) Draft entries in YAML or Sheet; keep IDs stable and lowercase snake_case. 2) Check cross-refs: `source_schema`, `coactivate_schemas`, `preferred_frames`. 3) Keep weights within bounds; ensure bilingual lexicon has ≥ 5 lemmas per language when possible. 4) Run local validation (engineers provide `sv metaphors validate`). 5) Iterate examples and affect until evaluator scores look sensible in dev runs. 6) Commit with message `data(metaphors): seed v0.1`.

Quality Checklist (Before Handoff)

- [] Each entry has `type`, `definition`, and **either** `source_schema` **or** `axis+bipolar`.
- [] Affect numbers within ranges; bilingual lexicon present and weighted.
- [] `coactivate_schemas` exist in `schemas.yml`; `preferred_frames` will exist in `frames.yml`.
- [] No disallowed pairings contradict global constraints.
- [] At least one short text example per entry.
- [] Provenance filled (source/curator/license/confidence).

Common Pitfalls (Avoid)

- Mixing style adjectives with metaphors (e.g., “cinematic” is not a metaphor).

- Poles that duplicate frames/schemas instead of mapping to them.
 - Lemma lists with inflected forms or uppercase; keep lemmas as lowercase base forms.
 - Overlapping IDs with schemas or frames.
-

Hand-Off Note to Engineering

- Treat this file as **read-only source of truth**.
- Validators should enforce ranges, cross-refs, and unique IDs; normalization sorts lexicon by weight and clamps to `[0, 1]`.
- Retrieval indexes both `definition` and `lexicon` fields; generation uses `gating.beats_bias` and `expectation_increment` as soft controls.