**AI-01 — AI Governance for Assistive Features (Advisory-Only, HIL)**

**Intent**  
Govern your assistive AI (validation hints, code suggestions, license picks) so it’s **advisory-only**, explainable, and opt-out per tenant — with bias/drift checks and human-in-the-loop.

**Actors**  
Model Owner • Product Owner • Compliance • DPO/Legal • Audit Vault

**Preconditions**

* AI used only for suggestions (never the final decider).
* Tokens/decisions already signed (CBT/LRT) to separate advice vs enforcement.

**Flow (happy path)**

1. **Model Inventory**
   * Register each assist (Classifier Assistant, Drift Mapper, License Picker Helper) with: inputs, outputs, storage, retention, and **tenant opt-out flag**.
2. **Guardrails**
   * Hard constraints: no auto-actions; confidence threshold; show **WHY snippet** for every suggestion; display “advice only” banner.
3. **Evaluation & Bias**
   * Build labeled eval sets (multi-script names, diverse geos).
   * Score precision/recall & false-positive rates; track by segment/locale; store reports.
4. **Drift & Monitoring**
   * Monitor live acceptance rates & error deltas; trigger re-eval when deltas exceed threshold.
5. **Human-in-the-Loop (HIL)**
   * Reviewer must accept/override; capture reason categories; feed back into evals.
6. **Change Management**
   * Any model/config update → governance review; publish change note; rollout behind a feature flag; rollback path.
7. **Tenant Controls**
   * Per-tenant toggle: **Enable AI assists** + per-assist granularity; log state in decisions.

**Edge cases**

* Low confidence → surface “What to add” hints; never silently choose.
* Tenant opt-out → hide assistants; keep deterministic flows intact.

**Done when**

* Model cards + eval reports published; assistive features show WHY and confidence; toggles work per tenant; post-change metrics stable.

**Deliverables**  
AI Governance Document, Model Cards, Eval reports, Flag config, Audit log of HIL decisions.