**EAR‑QA Windows‑First Roadmap (B.28–B.41)**

**B.28 – Provenance & delta loads**

- Modules: earCrawler/transforms, earCrawler/kg (delta), tests/, docs/.

- Deliverables: provenance triples, content hashes, idempotent upserts.

- Tests: repeat build → zero diff; provenance required in CI.

**B.29 – Deterministic IRIs & canonicalization**

- Modules: earCrawler/kg, earCrawler/ontology, earCrawler/utils.

- Deliverables: frozen IRI rules; alias registry; canonicalization pass.

- Tests: same input → same IRIs; variants collapse.

**B.30 – Entity resolution (dedup)**

- Modules: earCrawler/kg/reconcile, cli/reconcile (if present).

- Deliverables: deterministic merges; reversible with audit facts.

- Tests: synthetic collisions; no data loss.

**B.31 – FR text anchors & citations**

- Modules: api\_clients/federalregister\_client.py, earCrawler/ingestion.

- Deliverables: anchors linked to part: IRIs; indices for prompt assembly.

- Tests: each part has ≥1 anchor when applicable.

**B.32 – Mention extraction upgrade**

- Modules: earCrawler/transforms (mentions.py).

- Deliverables: token/pattern logic; graded strength; stoplists/context.

- Tests: precision/recall sample; quality floor in CI.

**B.33 – Policy‑aware linking rules**

- Modules: kg/assembler (rules.yml), earCrawler/kg.

- Deliverables: data‑driven hints; never override facts.

- Tests: deterministic outputs; schema/version checks in CI.

**B.34 – Windows batch jobs**

- Modules: scripts/, earCrawler/cli.

- Deliverables: scheduled jobs; structured logs; dry‑run mode.

- Tests: exit codes/logs; offline smoke in CI.

**B.35 – Caching & retries**

- Modules: api\_clients/\*, earCrawler/core (cache utils).

- Deliverables: normalized‑key cache; bounded retries; warm/invalidate.

- Tests: 429/5xx paths; call‑budget checks.

**B.36 – Observability**

- Modules: earCrawler/telemetry, earCrawler/monitor.

- Deliverables: run id, timings, JSON summaries.

- Tests: presence/format; artifacts published.

**B.37 – Access control & secrets**

- Modules: earCrawler/security, earCrawler/utils/secure\_store.py, cli/auth.

- Deliverables: centralized secrets; redaction; static secret scan.

- Tests: clear failures when missing; scan blocks violations.

**B.38 – RDF export profiles**

- Modules: earCrawler/kg, cli/bundle.

- Deliverables: TTL/NT/gz + checksums + manifests.

- Tests: round‑trip parse; checksum verification; artifact checks.

**B.39 – Graph integrity checks**

- Modules: earCrawler/kg/validate.py, sparql/ sanity queries.

- Deliverables: orphan/label/namespace checks; human‑readable report.

- Tests: seeded bad graphs detected; CI fails on violations.

**B.40 – Admin CLI (Windows)**

- Modules: earCrawler/cli/\*.

- Deliverables: single command surface; quiet/verbose; Windows help.

- Tests: arg‑parsing and chaining; happy‑path in CI.

**B.41 – Performance pass**

- Modules: perf/, earCrawler/transforms, kg serializers.

- Deliverables: batched writes; streaming; benchmarks/targets.

- Tests: moderate‑size perf assertions; regression warnings in CI.

Windows baseline: All milestones execute on Windows workstations; DAL/TACC onboarding stays gated for later phases.