You’re right — this *is* the merged master plan. 🎉  
Everything we discussed has been folded in:

| **Area** | **Earlier tweaks** | **Present in your consolidated doc?** |
| --- | --- | --- |
| Nightly E2E + chaos / perf harness | ✓ | ✓ “Automated Nightly E2E Pipeline” & “Resilience Drills” |
| Feature-freeze end of Week 2 | ✓ | ✓ in Release-Management section |
| phase3\_flags.yaml with risky-feature toggles | ✓ | ✓ plus example flags |
| Axe-core gate → Playwright keyboard pass | ✓ | ✓ under Accessibility |
| Secret-flow regression & Vault fuzzing | ✓ | ✓ “Secret Detection Flow” + “Vault Security Hardening” |
| Prometheus metrics endpoint & Grafana JSON | ✓ | ✓ Observability section |
| Plugin sandbox validation | ✓ | ✓ “Plugin Sandbox Validation” |
| In-app feedback hook | ✓ | ✓ “Instrumentation & Feedback” |
| Offline demo + captioned assets | ✓ | ✓ Demo resilience |

**Gaps?** Only two ultra-minor items aren’t explicit:

1. **Load-testing dataset license check** – just verify the sample repo is MIT/BSD or public-domain.
2. **OpenAPI auto-publish CI step** – you mention Redoc, but add a failing gate if schema generation breaks.

If you’re happy with the scope, the next practical move is to **break Week 1 into tickets**:

bash

CopyEdit

github issue create --title "CI: axe-core gate"

github issue create --title "Test: ExecutionTracer ↔ FieldMeditationMode harness"

github issue create --title "Infra: nightly E2E pipeline"

...

I can draft those ticket checklists or write the initial config/phase3\_flags.yaml + loader snippet—just say the word.

**You said:**

provide an updated list please

**ChatGPT said:**

Thought for 11 seconds

**Updated Phase-3 “To-Do” Master List**

*(includes the two tiny gaps we spotted + granular Week-1 tickets ready for GitHub)*

**A. New Items Added**

| **ID** | **Item** | **Why** | **Owner** | **When** |
| --- | --- | --- | --- | --- |
| **A-1** | **Dataset-license verifier** for demo repos | Avoid shipping non-redistributable code/data in sample project | Dev Rel | Wk 1 |
| **A-2** | **OpenAPI auto-publish CI job** + failing gate | Keep docs + SDKs in sync; block drift | Dev Infra | Wk 1 |

**B. Week-1 Ticket Breakdown *(create as GitHub Issues)***

| **#** | **Title** | **Labels** | **Acceptance Criteria** |
| --- | --- | --- | --- |
| **1** | CI · **axe-core accessibility gate** | ci, accessibility | Jest-axe runs on Storybook & editor pages ⇒ build fails on any rule ≥ serious |
| **2** | Test · **ExecutionTracer ↔ FieldMeditation harness** | test, visualization | RTL test mounts tracer mock, fires 10 events, asserts SVG node count == 10 |
| **3** | Test · **RefactorService ↔ EditorSync harness** | test, code-sync | Apply concept morph, expect AST diff + concept-graph diff both truthy |
| **4** | CI · **Nightly E2E pipeline** | ci, e2e | GH Action runs full matrix @ 02:00 UTC; green badge surfaces in README |
| **5** | Perf · **Load harness (5 k files, 50 agents)** | perf | p95 SSE latency < 500 ms, FPS drop < 30 % |
| **6** | Sec · **Secret-detection regression test** | security | Hard-coded AWS key triggers vault flow; test passes when blocked |
| **7** | Proto · **MCP compliance suite in CI** | ci, mcp | FastMCP mcp validate passes; schema version pinned |
| **8** | Infra · **phase3\_flags.yaml loader util** | feature-flag | useFeatureFlag('telemetryOptIn') returns boolean from YAML |
| **9** | Infra · **Commit initial config/phase3\_flags.yaml** | feature-flag | File merged with flags listed below |
| **10** | Test · **Plugin sandbox validation harness** | plugin, security | Third-party plugin can’t access FS outside allowed path; test fails if breach |
| **11** | Tool · **Dataset license verifier script** | demo, legal | CLI returns non-zero for GPL or unknown licenses |
| **12** | Docs · **OpenAPI build & Redoc publish job** | docs, ci | PR blocked if docs/api/openapi.yaml changes but Redoc HTML not regenerated |
| **13** | Resilience · **Chaos-test script (kill MCP, corrupt Vault)** | reliability | Recovery within 5 s; no unhandled promise rejections |
| **14** | Observability · **Prometheus metrics endpoint** | ops, metrics | /metrics exposes default + custom tool\_latency\_seconds histogram |
| **15** | Observability · **Grafana dashboard JSON stub** | ops, metrics | Dashboard shows RPS, latency p95, agent\_error\_rate; JSON committed |
| **16** | Sec · **Dependabot + bandit + npm audit** | security, ci | All three scanners wired; PR fails on high-severity CVEs |

**C. Feature-Flag File (config/phase3\_flags.yaml)**

yaml

CopyEdit

# Phase-3 risky/experimental toggles

telemetryOptIn: false # GDPR toggle

secretDetection: true

newAgentPanels: true

offlineDemoMode: false

multiAgentParallel: false # turn on after perf tuning

autoWriteFile: false # safety valve

newGraphVirtualization:true

**D. Remaining Weeks (unchanged)**

* Week-2: **Bug fixes · Security pipeline · Feature freeze**
* Week-3: **Docs · Dev-experience · Demo polish**
* Week-4: **Final testing · Observability · Release packaging**

Everything else in your consolidated implementation plan stands; the table above slots the missing bits and gives you ready-made tickets for the kickoff board.