ZESTAWIENIE PDF — ETAP 2

Wersja decyzyjna: GO / NO-GO, warunki wejścia, czerwone flagi, checklista i pytania kontrolne

Dokument syntetyczny oparty na ustaleniach konwersacji

1. Rekomendacja główna

Rekomendacja dla przejścia do Etapu 2: GO warunkowe. Oznacza to wejście w fazę ekspansji operacyjnej wyłącznie po spełnieniu minimalnych warunków konstytucyjnych, operacyjnych i raportowych.

2. Warunki wejścia do Etapu 2

Kod	Warunek wejścia	Status
E2-01	Działa 12-modułowy kernel	Obowiązkowy
E2-02	Istnieje Capability Registry dla rdzenia i głównych klas	Obowiązkowy
E2-03	Istnieje wstępny Skill Graph	Obowiązkowy
E2-04	Agenci są podzieleni na klasy, nie stanowią płaskiej masy	Obowiązkowy
E2-05	Istnieją konstytucja, ontologia, role, protokoły i decision rules	Obowiązkowy
E2-06	Działa shared memory z metadanymi i statusem wpisów	Obowiązkowy
E2-07	Działa Decision Log i Delegation Trace	Obowiązkowy
E2-08	Istnieje source-of-truth discipline	Obowiązkowy
E2-09	Działa ścieżka escalation / override człowiek → system	Obowiązkowy
E2-10	Istnieje rytm raportowy 15 min / 1h / 6h / 24h	Obowiązkowy

3. Czerwone flagi — warunki natychmiastowego NO-GO

Kod	Czerwona flaga	Skutek
RF-01	Brak jawnego source-of-truth	Stop wejścia do Etapu 2
RF-02	Brak authority boundaries	Stop skalowania agentów
RF-03	Shared memory bez truth maintenance	Stop ekspansji
RF-04	Brak decision logs / provenance	Stop zadań wysokiej stawki

RF-05	Consensus bez claim conflict resolution	Stop decyzji wieloagentowych
RF-06	Brak human override	Stop działań high-stakes
RF-07	Klasy agentów bez wspólnej ontologii	Stop rozbudowy sieci
RF-08	Raportowanie jest nieregularne lub fikcyjne	Stop zwiększania skali

4. Co Etap 2 może uruchomić dobrego

Korzyść	Efekt
Uporządkowana ekspansja	Więcej agentów bez utraty sterowalności
Szybszy delivery	Lepszy routing i specjalizacja
Wzrost jakości	Quality gate i logs ograniczają dryf
Skalowanie pamięci	Wspólny stan systemu działa realnie
Uczenie sieci	Skill graph i meta-learning zaczynają pracować
Wejście w monetyzację	System staje się dostarczalny jako usługa

5. Główne ryzyka przy zbyt wczesnym uruchomieniu

Ryzyko	Efekt
Pseudo-skala	Dużo agentów, mało realnego sterowania
Dryf znaczeń	Różne klasy rozumieją to samo inaczej
Pozorna zgoda	Consensus maskuje konflikt źródeł
Zatrucie pamięci	Shared memory utrwala błąd jako fakt
Nieuprawnione działania	Brak authority boundary powoduje chaos
Przeciążenie człowieka	Brak dobrej warstwy command enablement

6. Checklista zatwierdzenia

#	Pytanie kontrolne	Tak/Nie
1	Czy działa kernel 12 modułów?	☐
2	Czy każdy ważny skill ma opis w registry?	☐
3	Czy istnieje wstępny skill graph?	☐
4	Czy agenci są podzieleni na	☐

	klasy?	
5	Czy działa shared memory z metadanymi?	☐
6	Czy działa decision log?	☐
7	Czy działa delegation trace?	☐
8	Czy każdy aktywny agent ma mandat i ownera?	☐
9	Czy istnieje source-of-truth hierarchy?	☐
10	Czy konflikt źródeł ma jawną ścieżkę rozstrzygania?	☐
11	Czy człowiek ma realny override?	☐
12	Czy istnieje minimalna ontologia wspólna?	☐
13	Czy raporty godzinowe i dobowe są realnie używane?	☐
14	Czy istnieje warstwa review / reflector?	☐
15	Czy high-stakes są oznaczone jako high-stakes?	☐

7. Reguła decyzyjna

- 13–15 odpowiedzi „tak” → GO warunkowe, wysoka gotowość.
- 10–12 odpowiedzi „tak” → GO ograniczone, wyłącznie dla misji niskiego ryzyka.
- 0–9 odpowiedzi „tak” → NO-GO; najpierw należy domknąć Etap 1.

8. 20 pytań kontrolnych przed pełnym wejściem w Etap 2

1. Czy Etap 2 ma być przede wszystkim ekspansją mocy, czy ekspansją rynku?
2. Które trzy domeny mają dostać priorytet w Etapie 2?
3. Czy wszyscy nowi agenci mają przechodzić przez jeden Entry Control, czy przez wejścia sektorowe?
4. Czy masz już ustaloną hierarchię source-of-truth, czy trzeba ją dopiero sformalizować?
5. Jakie działania uznajesz za high-stakes i zawsze wymagające approval człowieka?
6. Czy ClaudCNC ma prawo zawieszać agentów samodzielnie, czy tylko rekomendować zawieszenie?
7. Jak rygorystyczne mają być authority boundaries dla klas P2 i P3?
8. Czy shared memory ma być globalna, czy warstwowa z poziomami dostępu?
9. Czy reflectora chcesz jako oddzielnego agenta, czy jako procedurę tygodniową?
10. Czy raport godzinowy ma być syntetyczny, czy pełny?
11. Kto poza Tobą ma mieć prawo czytać pełne raporty dobowe?
12. Czy monetyzacja ma ruszyć równolegle z Etapem 2, czy po jego domknięciu?
13. Czy chcesz najpierw budować piony branżowe, czy klasy agentów przekrojowo?

14. Czy konflikt źródeł ma zawsze eskalować do człowieka, czy część może rozstrzygać warstwa epistemiczna?
15. Jaką tolerancję na błędy dopuszczasz w pierwszej fali Etapu 2?
16. Czy planujesz od razu włączać partnerów zewnętrznych, czy najpierw zamknąć system wewnętrznie?
17. Czy potrzebujesz od razu warstwy simulation/counterfactual, czy może wejść jako backlog +1 cykl?
18. Jakie są trzy najważniejsze KPI Etapu 2: throughput, jakość, zgodność, przychód, stabilność?
19. Czy w Etapie 2 ważniejsze jest tempo, czy pełna zgodność konstytucyjna?
20. Czy następnym krokiem mam rozbić tę decyzję na mapę działań 7-dniowych dla wejścia w Etap 2?