ChatGPT

Asymetria Twojej platformy wobec ryzyk SaaS+AI na tle ocen ekspertów i mediów

Zakres, definicje i metoda dowodowa

Potrzeby informacyjne (3–6): - Jakie **konkretne mechanizmy w repo** (gating, delta, fail-closed, kontrola pamięci, formalne osie 9D) tworzą *asymetrię* i gdzie są ich „twarde" punkty zaczepienia (artefakty wymuszające decyzję).
- Jak te mechanizmy mapują się na **najbardziej aktualne ryzyka** SaaS+AI wskazywane przez instytucje, standardy i rynek: pricing/telemetria, kaskady awarii i kosztów, bezpieczeństwo LLM (OWASP), governance (NIST), zgodność (AI Act), constraint energetyczny (IEA) i repricing rynkowy (Reuters).
- Czy Twoje repozytoria i badania **faktycznie ujmują** te problemy w sposób falsyfikowalny (metryki, testy, warunki obalenia), czy raczej tylko „pasują narracyjnie".
- Jakie są **luki/niezgodności** (bramka na brzegu vs brak hamulców w środku; kontrola supply chain vs brak sterowania ekonomią inferencji; model osi vs brak egzekucji).
- Jak przeprowadzić **statystyczną falsyfikację hipotez asymetrii**: metryki, testy, panel/effect-study, oraz Monte Carlo przy braku danych.

Zakres źródeł i kolejność (wymóg użytkownika).
1) Najpierw: analiza Twoich repozytoriów na GitHub [1] (sbom, writeups, chunk-chunk, swarm, ai_platform, HA2D). W tym raporcie dowód repo opieram na dostarczonych przez Ciebie dokumentach-syntezach z cytowanymi artefaktami (pipeline/polityki/protokoły), bo są one zebranym „materiałem dowodowym" dla kodu i specyfikacji.
2) Następnie: źródła zewnętrzne o najwyższej wadze dowodowej i świeżości: NIST [2] (AI RMF, profil dla GenAI), European Commission [3] / AI Act (art. 12/72), OWASP Foundation [4] (Top 10 LLM Apps), FinOps Foundation [5] (Framework 2025 + FOCUS), International Energy Agency [6] (Energy and AI), Reuters [7] (kaskady rynkowe i CAPEX anxieties), entity["organization","Bain & Company","management consulting firm"] (przejście seat→hybrydy i brak telemetryki jako bariera). [8]
Parametry nieokreślone (czas, budżet, zasoby) traktuję jako **„brak ograniczeń"**, a w testach podaję również warianty minimalne (MVP), bo brak ograniczeń rzadko jest prawdziwy operacyjnie.

Definicja operacyjna „asymetrii" (falsyfikowalna).
W reżimie naukowym (w sensie: „da się obalić") asymetria oznacza: **istotną redukcję ogonowego ryzyka** (kosztowego, awaryjnego, bezpieczeństwa, zgodności) **przy relatywnie małym koszcie wdrożenia** i przy zachowaniu funkcjonalności. Najprostsze wskaźniki to: spadek CVaR/p99 kosztu na jednostkę pracy, spadek kaskadowości awarii, skrócenie MTTR i wzrost auditability/log completeness.

Artefakty asymetrii w Twoich repozytoriach

Rdzeń wzorca: „pomiar → próg → akcja"

Twoje repozytoria (wspólny motyw) traktują złożoność SaaS+AI jako układ sprzężeń i progów, który trzeba **sterować**, a nie tylko „monitorować". Dokumenty nazywają to wprost anty-wzorcem wobec SaaS,

gdzie brak progów/metrowania powoduje akumulację długu aż do „dziwnego dnia” (heavy-tail) i przejścia fazowego w repricing/kaskady.

W ujęciu implementacyjnym masz trzy klasy stabilizatorów: - **Stabilizator supply-chain (CI/CD):** snapshot → SBOM → scan → delta → gate.
- **Stabilizator runtime (mesh):** rate limit + circuit breaker/outlier ejection (anti-cascade).
- **Stabilizator epistemiczno-audytowy:** protokół kontekstu + integralność pamięci + formalizacja osi decyzji (replikowalność, „delta kontekstu”).

sbom: asymetria przez wymuszalny dowód pochodzenia i zmiany

W modelu anty-SaaS najważniejsza przewaga sbom jest taka, że nie budujesz „ładnego raportu”, tylko **pętlę sterowania**: delta i gate są mechanizmem, który potrafi powiedzieć „STOP” i przerwać kaskadę kosztów/ryzyk zanim dojdzie do wdrożenia.

Ta logika jest spójna z ramami ryzyka, które wymagają nie tylko pomiaru, ale *akcji zarządczej* (Measure→Manage). W praktyce: Twoje repo „widzi” supply chain jako źródło ryzyk heavy-tail i wprowadza progi zamiast liczenia, że „regulamin i on-call wystarczą”.

Asymetria: mała reguła decyzyjna (próg) może blokować duże koszty (incydent, podatność, regresja).
Ryzyko: brak domknięcia dowodu pochodzenia end-to-end (np. attestations) może spowodować, że SBOM jest poprawny, ale słabo „wiąże się” z łańcuchem zaufania w czasie.

swarm: asymetria przez fail-closed i anty-kaskady — ale też punkt „zniwelowania asymetrii”

W `swarm` masz klasyczne elementy anty-kaskadowe: fail-closed rate limiting i circuit breaking/outlier detection. To jest materializacja „kompromisu pokory”: nie skaluj bez granic, odcinaj outliery, tnij propagację błędu.

Jednocześnie te same materiały wskazują krytyczny kontrapunkt: **bramka na brzegu vs brak backpressure w środku**. W kodzie ingestu (wątek per UDP packet + brak timeout/backpressure) jest wzorzec, który może generować przeciążenie wewnętrzne i niwelować przewagę mesh-gatingu.

To jest szczególnie istotne dla SaaS+AI, bo agentowość jest „wbudowanym burstem”: jeden request potrafi uruchomić wiele kroków (tool-use, retry, retrieval), czyli kosztowe i awaryjne sprzężenie dodatnie.

chunk-chunk: asymetria przez formalny kontrakt decyzyjny i „energetykę” kroku

`hmk9d_protocol.yaml` wprowadza formalizm, który jest rzadki w praktycznych projektach SaaS: zachowanie jako złożenie kompresji i polityki (`H(s)=g(F(s))`), ryzyko jako spodziewana strata (`R(F,g)`) oraz energia kroku (`E(Δ)`, `E_total`) i bramki bezpieczeństwa (progi).

Asymetria: przesuwasz dyskusję z „AI to feature” na „AI to system sterowania i rachunku pracy” — to jest zgodne z tezą o renormalizacji: przejściu z seats do work units.
Ryzyko: bez operacjonalizacji (telemetria → policy engine → enforcement) formalizm pozostaje semantyką, a nie mechanizmem.

ai_platform i HA2D: asymetria przez mapowanie złożoności i pamięć dowodową

`ai_platform` próbuje rozwiązać problem „mnożenia logik" (wiki/folderów/metryk) przez wspólny układ współrzędnych `/QV9D`, mapujący warstwy/mosty/typy artefaktów na repo i procesy. To jest potencjalnie asymetryczne organizacyjnie: zamiast rosnącej entropii dokumentacji tworzysz rejestr sterowania.

Jednocześnie wprost nazwany jest brak deterministycznych identyfikatorów (`id_latarni` „do zaprojektowania"), co w reżimie dowodowym jest krytyczne: bez stabilnego ID nie ma porządnej korelacji, a bez korelacji nie ma statystyki i audytu.

`HA2D` wnosi protokół pamięci kontekstu (UUID, timestamp, payload, SHA256; STORE/RETRIEVE/LATEST), czyli budulec replikowalności i integralności „kontekstu jako obiektu".
To jest rdzeń Twojej tezy „wioski kosmicznej": produkcja AI jest zasobowo zamknięta, więc kontekst musi być zasobem sterowniczym, a nie „wiedzą w głowie".

writeups: asymetria epistemiczna i falsyfikacja protokołu

`writeups` dostarcza najważniejszego elementu dla rygoru: **warunku falsyfikacji** („protokół częściowo poznany, jeśli acc($\hat{H}$) > acc_bazowa na danych odłożonych") oraz opis ryzyk przeuczenia/dryfu/złej bazy i „kryterium śmierci". To spina Twoją architekturę z nauką: nie tylko „mamy progi", ale „umiemy powiedzieć, kiedy teza pada".

Konfrontacja z aktualną oceną ekspertów i mediów

Governance ryzyka i cykl życia

NIST [2] AI RMF 1.0 jest ramą „Govern/Map/Measure/Manage" jako językiem sterowania ryzykiem AI w cyklu życia (dobór, wdrożenie, użytkowanie, monitoring). [9]
Profil dla GenAI doprecyzowuje ryzyka generatywne i praktyki w tym samym paradygmacie (cykl życia, aktorzy, ryzyka specyficzne dla generacji). [10]

Zgodność z repo: Twoje „pętle sterowania" (SBOM gate, progi mesh, pamięć kontekstu) są zgodne z namiarami na Measure/Manage; Twoje writeups dostarczają formalnej falsyfikacji (co jest spójne z „Measure").

Luka: brakuje w repozytoriach „artefaktu planu" (konkret: metryki→progi→właściciele→eskalacje), czyli mechanizmu, który z obserwacji robi *zarządzanie ryzykiem w operacji*.

Zgodność i logowanie według AI Act

AI Act art. 12 wymaga, aby systemy wysokiego ryzyka umożliwiały automatyczne logowanie zdarzeń przez cały cykl życia; logi mają wspierać m.in. wykrywanie ryzyk i „substantial modification", post-market monitoring (art. 72) i monitoring operacji. [11]
Art. 72 wymaga post-market monitoring system i planu; Komisja ma przyjąć akt wykonawczy ustanawiający template planu do 2 lutego 2026. [12]

Zgodność z repo: AID/event-sourcing (sbom) i CMM (HA2D) są dobrym budulcem „dowodowego śladu".

Luka ryzyka regulacyjnego: brak jawnego planu post-market monitoring jako artefaktu repo/procesu oznacza ryzyko „compliance-by-logs”: logi są, ale nie sterują.

W polskim kontekście Ministerstwo Cyfryzacji [13] podkreśla etapowe wejście w życie AI Act (m.in. zakazane praktyki, AI literacy), co wzmacnia wagę budowania governance jako procesu, nie tylko jako kodu. [14]

Bezpieczeństwo LLM (OWASP) i przesunięcie ryzyk z „transportu” do „semantyki”

OWASP Foundation [4] Top 10 dla aplikacji LLM (v1.1) obejmuje m.in. prompt injection (LLM01), insecure output handling (LLM02), model DoS (LLM04), supply chain vulnerabilities (LLM05) i sensitive info disclosure (LLM06). [15]

Zgodność z repo:
- LLM05 (supply chain): silne wsparcie przez sbom i bramki CI/CD.
- LLM04 (DoS/koszty): częściowe wsparcie przez rate limiting i circuit breaker w swarm.

Luka krytyczna: brak „twardych bramek semantycznych” na wyjściu (LLM02) i brak mechaniki anty-prompt-injection jako zestawu polityk testowalnych. To jest miejsce, w którym AI-SaaS najczęściej „przenosi” ryzyko: progi transportowe nie chronią, gdy model generuje złośliwy/niebezpieczny output, który downstream potraktuje jak komendę.

Ekonomia SaaS+AI: telemetria jako warunek pricingu oraz „inference whales”

entity["organization","Bain & Company","management consulting firm"] wskazuje, że AI zmienia strukturę kosztu i wartości, a odejście od per-seat jest trudne m.in. dlatego, że firmy często nie mają telemetryki produktu i infrastruktury IT/billing/finance do pricingu usage/outcome; hybrydy (seat + metrum AI) dominują jako stan przejściowy. [16]
Media branżowe opisują zjawisko „inference whales”: ciężki ogon zużycia potrafi generować koszty radykalnie większe od abonamentu, wymuszając limity i capy. [17] (kontekst ogonów kosztu)

Zgodność z repo: Twoje badania formalizują próg ekonomiczny i „iskrę” wprost przez równanie marży (P – q·u) i próg u* = P/q: kiedy przychód jest „płaski”, a koszt zmienny i ciężkoogonowy, system staje się strukturalnie kruchy.

Luka krytyczna: repo (w warstwie implementacyjnej) ma świetny wzorzec telemetryki dla supply chain (sbom_snapshot/scan/delta/gate), ale nie ma jeszcze analogicznego, ustandaryzowanego strumienia telemetryki AI-runtime: tokeny/czas/koszt/tenant/workflow i budżety, które da się egzekwować.

Energia jako constraint i „logika mocy”

International Energy Agency [6] szacuje zużycie energii przez data centers na ok. 415 TWh w 2024 i projekcję ~945 TWh do 2030 w scenariuszu bazowym; raport podkreśla niepewności i potrzebę scenariuszy oraz fakt, że data center można postawić w 2–3 lata, podczas gdy infrastruktura energetyczna ma dłuższe lead-time’y. [18]

To wzmacnia Twoją oś „Semantyka–Energia”: jeśli energia jest bottleneckiem, wtedy „koszt kroku” nie jest metaforą, tylko twardym parametrem sterowania (w ekonomice i w ryzyku reputacyjnym/ESG).

Kaskady rynkowe i dynamika „AI scare trade”

W lutym 2026 relacje rynkowe opisują rozlewanie się paniki wokół AI poza sektor software (brokerzy, dane, real estate, insurance), z istotnymi spadkami w wielu branżach; narracja brzmi jak kaskada progowa: „sell first, think later”. [19]
W Indiach Reuters opisał silne spadki w IT services i jednocześnie kontrargument analityków, że integracja AI w enterprise nadal wymaga tych firm (czyli: rynek może przeszacowywać bezpośrednią dysintermediację). [20]

Twoje tezy „pętla w pętli” i „mnożenie logik” są spójne z mechaniką obserwowaną przez rynek: katalizator agentowy + presja na pricing + CAPEX/energia + napięcia w łańcuchu vendor–client.

Ocena zgodności, ryzyk drugiego rzędu i „czy to falsyfikuje krytykę”

Czy Twoje repo wskazują te problemy jasno?

Tak — ale asymetrycznie (w sensie: „w pewnych klasach bardzo mocno, w innych jeszcze nie”).
Najmocniejsze dowody implementacyjne dotyczą mechanizmów „pokory” jako progów wymuszanych: sbom-owy gate/delta oraz swarm-owe polityki sieciowe. To są artefakty, które da się testować i obalać (false positive/negative, burst/sustained, MTTR).

Formalizacje osi (HMK9D/QV9D) i protokołów kontekstu (HA2D/writeups) są metodologicznie mocne, bo budują język falsyfikacji (replikowalność, delta-kontekstu), ale ich „siła dowodowa” zależy od tego, czy zostaną spięte z enforcement (CI/runtime).

Czy repo falsyfikują lub zgadzają się z krytyką mediów/ekspertów?

Zgadzają się z krytyką w diagnozie i w kierunku mechaniki, ale **nie falsyfikują jej w pełni** bez telemetryki AI-runtime i bez testów obciążeniowych/ekonomicznych na danych rzeczywistych:
- Krytyka „pricing/telemetria” (Bain/FinOps) nie jest jeszcze obalona, bo nie ma produkcyjnego odpowiednika “sbom-telemetry” dla AI zużycia i kosztu. [21]
- Krytyka OWASP LLM01/LLM02 nie jest obalona, bo w repo brak zestawu polityk i bramek semantycznych równoważnych gate’owi supply chain. [15]
- Krytyka AI Act (plan monitoringu) nie jest obalona, bo brakuje planu jako artefaktu procesu (metryki/progi/właściciele). [22]

Luki i ryzyka drugiego rzędu

Najbardziej krytyczne „niwelatory asymetrii” (czyli miejsca, gdzie przewaga może zniknąć):
- **Bramka na brzegu, brak hamulców w środku:** mesh ciętnie ruch, ale aplikacja wewnątrz może przeciążać się poprzez brak backpressure (wzorzec thread-per-packet).
- **Koszt/energia w modelu, nie w telemetryce:** HMK9D ma oś energii, ale bez uziemienia w kosztach (tokens/GPU-seconds/energia/PLN) progi nie odpowiadają realnej ekonomice.
- **QV9D bez deterministycznych ID:** brak stabilnych identyfikatorów osłabia audyt i statystykę (brak porównywalności w czasie).

Tabela porównawcza: artefakty repo vs oceny mediów/ekspertów

Repo / artefakt	Mechanizm asymetrii	Ryzyko zewnętrzne (najnowsze)	Ocena zgodności	Największa luka/ ryzyko
sbom: pipeline SBOM→scan→delta→gate	Twarde bramki CI/CD i „dowód zmiany”	Supply chain jako ryzyko heavy-tail (NIST/SSDF trend) [23]	Wysoka	Nie obejmuje AI-runtime (prompty, modele, dane)
swarm: fail-closed rate limit	Anty-kaskada kosztowa/ awaryjna	OWASP LLM04 (DoS i koszty) [15]	Wysoka	Rate-limit service jako element krytyczny + brak kalibracji do kosztu/WU
swarm: circuit breaker/ outlier	Odcinanie outlierów, redukcja propagacji błędów	SRE logika kaskad (overload) — zgodne z praktykami branży	Wysoka	Agresywność ustawień + brak „hamulca w środku” aplikacji
chunk-chunk: HMK9D	Formalizacja ryzyka i energii kroku, progi jako część kontraktu	IEA: energia jako constraint; OWASP: kosztowe DoS [24]	Kierunkowo bardzo wysoka	Brak operacjonalizacji w telemetry/billing/ policy enforcement
ai_platform: QV9D	Redukcja „mnożenia logik” przez wspólną ontologię artefaktów	Bain/ FinOps: telemetria jako warunek pricingu [25]	Średnia-wysoka	Brak deterministycznych ID → słaba korelacja i audyt
HA2D: protokół kontekstu	Integralność pamięci kontekstu (hash), replikowalność	AI Act art. 12/72: logi i monitoring [26]	Wysoka jako budulec	Brak spięcia w plan monitoringu i polityki retencji/ danych

Repo / artefakt	Mechanizm asymetrii	Ryzyko zewnętrzne (najnowsze)	Ocena zgodności	Największa luka/ ryzyko
writeups: falsyfikacja protokołu	Kryteria obalenia i ryzyka metodologiczne (drift, baseline)	NIST: Measure/ Manage, TEVV; OWASP: output handling jako ryzyko	Wysoka metodologicznie	Brak danych produkcyjnych i prerejestracji testów → ryzyko „zwycięstwa narracyjnego”

Statystyczna falsyfikacja hipotez asymetrii

W reżimie „absolutnego rygoru” (tj. minimalizacja wniosków nieobalalnych) potrzebujesz: definicji jednostki pracy, telemetryki, oraz testów odpornych na ogony (percentyle, CVaR, bootstrap), bo kaskady są rzadkie i duże.

Hipotezy i kryteria obalenia

Wersja robocza hipotez (kompatybilna z Twoimi repo i z krytyką rynku):
- **H1-S (supply chain):** SBOM-gating redukuje ekspozycję na krytyczne podatności w wydaniach oraz skraca „dwell time” ryzyka w CI/CD. Obalenie: brak spadku lub wysoki koszt fałszywych alarmów, który zabija przepływ.
- **H1-R (reliability / anti-cascade):** progi runtime zmniejszają kaskadowość awarii (mniej 5xx, więcej kontrolowanego 429, krótszy MTTR) w testach burst i sustained. Obalenie: brak różnic lub pogorszenie p99/MTTR.
- **H1-E (ekonomia):** budżety i bramki AI-runtime redukują ogon kosztu na jednostkę pracy (CVaR/p99) i obniżają P(loss) w modelach flat/hybrid. Obalenie: brak spadku ogona kosztu mimo progów, lub spadek kosztu okupiony nieakceptowalną degradacją jakości produktu.
- **H1-G (governance/compliance):** protokół kontekstu + eventy + plan monitoringu zwiększają audytowalność i skracają „czas do wyjaśnienia” incydentu. Obalenie: brak poprawy kompletności logów i brak redukcji czasu analizy.

Wymagane dane wejściowe (minimum)

Aby rzeczywiście testować H1-E i H1-G musisz mieć telemetrykę na „work units”, nie na „seats”: to jest centralny punkt Twojej ekonomicznej tezy o u* = P/q.

Minimalny event dla AI-runtime powinien zawierać: tenant/workflow_id, model, tokeny in/out, czas, liczba kroków/iteracji, narzędzia, koszt szacowany, decyzja bramki (degrade/deny/human review). To jest analog Twojego sbom-owego zestawu eventów (sbom/scan/delta/gate).

Technicznie, spójne przenoszenie kontekstu można realizować przez standardy obserwowalności: OpenTelemetry [27] opisuje mechanizm „Baggage” do propagacji par nazwa/wartość w ramach transakcji (pomaga indeksować zdarzenia w usługach downstream). [28]

Testy statystyczne i projekty przyczynowe

W praktyce (dla metryk heavy-tail) rekomendowane są: bootstrap różnic percentyli (p95/p99), testy nieparametryczne dla rozkładów, oraz projekty quasi-eksperymentalne (ITS/DiD) zamiast naiwnego before/after. Te zasady są spójne z Twoimi materiałami o „rygorze falsyfikacji" i ostrzeżeniem przed myleniem poprawności mechanizmu z dowodem systemowym.

Symulacja Monte Carlo jako test mechanizmu (gdy brak danych)

Jeżeli nie masz jeszcze danych produkcyjnych, Monte Carlo może przetestować *konsekwencje założeń* (nie „dowód świata"). W Twoich materiałach jest to jawnie zaakceptowane jako „demonstracja mechanizmu" oraz plan kalibracji po zebraniu danych.

Poniższy wykres pokazuje syntetyczny, ciężkoogonowy rozkład kosztu AI na tenant w scenariuszu flat fee oraz variant z prostym capem (break-even). Taki ogon jest intuicyjnie zgodny z problemem „whales" i OWASP LLM04 (kosztowe przeciążenie). [15]

Symulacja Monte Carlo: ogon kosztu inferencji i efekt cap (break-even)

Interpretacja falsyfikacyjna (wprost): jeśli po wdrożeniu budżetów/capów nie obserwujesz spadku CVaR/p99 kosztu per WU ani spadku przypadków ujemnej marży per tenant, hipoteza H1-E upada.

Rekomendacje techniczne i komunikacyjne

Rekomendacje techniczne o najwyższej dźwigni

Najwyższa dźwignia (wprost wynikająca z konfrontacji repo ↔ Bain/FinOps/Reuters): **skopiować wzorzec SBOM-telemetrii na AI-runtime**. To jest pojedynczy ruch, który domyka lukę „pricing bez telemetryki". [21]

Konkretnie: - Zdefiniować family eventów: `ai_usage_snapshot`, `ai_cost`, `ai_delta`, `ai_budget_gate`, analogicznie do `sbom/scan/delta/gate`.
- Wprowadzić budżety per tenant/workflow (degrade, deny, require human) oraz limity iteracji agentów jako pętle równoważące „mnożenie logik".
- Ustandaryzować dane koszt/usage pod FOCUS (docelowy format), bo FOCUS 1.3 jest ratyfikowany i explicitnie dotyczy ujednolicenia danych billingowych; wspiera m.in. kompletność/recency. [29]
- Naprawić „wewnętrzny generator kaskady": backpressure w aggregatorze (kolejka + worker pool + limit inflight), jawne timeouty i retry budget; inaczej bramka mesh nie wystarczy.
- Zbudować OWASP-LLM guardrails: walidacja outputu schematem, allow-listy narzędzi, sandbox; to jest brakująca część „bramek semantycznych". [15]
- Wpisać w repo artefakt „post-market monitoring plan" zgodny z AI Act art. 72 (metryki, progi, właściciele, eskalacje, retencja), bo do 2 lutego 2026 jest wymagana template'owa forma. [22]

Rekomendacje komunikacyjne

- Zastąpić metafory publicznym językiem sterowania: „progi, liczniki, zasady degradacji, budżety, wyjątki z expiry" — to jest język, na który rynek reaguje w „AI scare trade" i w debacie o zwrocie z CAPEX w AI. [30]
- W polskim kontekście podpiąć narrację pod regulacje: AI Act (obowiązki i etapowanie) i Data Act (zmiana gry w lock-in, obniżenie/zniesienie opłat za switching od 12 stycznia 2027). [31]

- Ustawić „rdzeń obietnicy” jako: **redukcja ogonów ryzyka** (CVaR/p99 kosztu/UW, kaskadowość, MTTR) zamiast „AI robi więcej feature’ów”. To jest spójne z Twoim anty-SaaS: przewaga jest w sterowaniu.

Źródła priorytetowe i mapa cytowań

Najwyższy priorytet (ramy normatywne i procesowe): - NIST AI RMF 1.0 + Playbook + profil GenAI. [32]
- AI Act (oficjalny tekst i artykuły 12/72) – logowanie i monitoring jako obowiązek procesu. [33]
- OWASP Top 10 LLM Apps v1.1. [15]

Wysoki priorytet (ekonomia i koszt jako rdzeń sporu): - FinOps Framework 2025 (Scopes: Public Cloud, SaaS, Data Center; możliwość definiowania scope’ów). [34]
- FOCUS 1.3 (ratyfikacja i sens: uniform billing datasets, recency/completeness). [35]
- Bain: hybrydy pricingu i telemetryka jako bariera transformacji. [16]

Wysoki priorytet (constraint energetyczny i ryzyko rynkowe): - IEA Energy and AI: 415 TWh (2024) → ~945 TWh (2030) i lead-time’y energii vs DC. [18]
- Reuters: kaskady strachu i „AI scare trade” rozszerzające się na sektory. [36]
- Reuters przez Investing.com: obawy o skalę CAPEX (~200 mld USD Amazon; ~630 mld USD Big Tech) i zwrot z inwestycji AI. [37]

Polskie źródła (preferowane tam, gdzie dostępne): - Ministerstwo Cyfryzacji: AI Act (etapowanie) i Data Act (ułatwienia switchingu, zniesienie opłat od 12.01.2027). [38]
- PARP o Data Act (to samo w skrócie praktycznym). [39]

Twoje dokumenty-dowody (repo + program badawczy):
- Anty-wzorzec wobec SaaS (mechanika progów, falsyfikacje).
- Konfrontacja repo z oceną ekspertów/mediów (tabela mapowania).
- Badania kontekstu („wioska kosmiczna”, kontekst jako zasób sterowniczy).
- Ekonomiczny dowód „pętli w pętli” (P, q, u, próg u*).
- Ocena naukowa testów amortyzacji kolapsu (ważność, ograniczenia i metryki).
- „Wybuch” cloud pricing jako efekt pętli (szkic mechanizmu).
- Synteza ryzyk i złożoności (mapa luk i rekomendacje).
- Superteza Manhattan (heurystyka wyścigu i Monte Carlo jako technika przeszukiwania).
- Teza o mnożeniu logik i kaskadach (hipotezy falsyfikowalne i kontrargumenty).
- Weryfikacja naukowa anty-wzorca na kodzie (procedury testów, kryteria obalenia).

---

[1] [19] [36] https://www.reuters.com/business/software-real-estate-us-sectors-under-grip-ai-scare-trade-2026-02-13/
https://www.reuters.com/business/software-real-estate-us-sectors-under-grip-ai-scare-trade-2026-02-13/

[2] [4] [8] [9] [23] [27] [32] https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10
https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

[3] [29] [35] https://focus.finops.org/focus-specification/
https://focus.finops.org/focus-specification/

[5] [12] [22] https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-72
https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-72

6 16 21 25 https://www.bain.com/insights/per-seat-software-pricing-isnt-dead-but-new-models-are-gaining-steam/
https://www.bain.com/insights/per-seat-software-pricing-isnt-dead-but-new-models-are-gaining-steam/

7 13 34 https://www.finops.org/insights/2025-finops-framework/
https://www.finops.org/insights/2025-finops-framework/

10 https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence
https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence

11 26 33 https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-12
https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-12

14 31 38 https://www.gov.pl/web/cyfryzacja/pierwsze-przepisy-rozporzadzenia-o-sztucznej-inteligencji-ai-act-zaczynaja-obowiazywac
https://www.gov.pl/web/cyfryzacja/pierwsze-przepisy-rozporzadzenia-o-sztucznej-inteligencji-ai-act-zaczynaja-obowiazywac

15 https://owasp.org/www-project-top-10-for-large-language-model-applications/
https://owasp.org/www-project-top-10-for-large-language-model-applications/

17 18 24 https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai
https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

20 https://www.reuters.com/world/india/indian-it-stocks-extend-slump-head-worst-week-nearly-six-years-2026-02-13/
https://www.reuters.com/world/india/indian-it-stocks-extend-slump-head-worst-week-nearly-six-years-2026-02-13/

28 https://opentelemetry.io/docs/reference/specification/overview/
https://opentelemetry.io/docs/reference/specification/overview/

30 37 https://www.investing.com/news/stock-market-news/amazon-shares-sink-as-big-techs-ai-spending-plans-worry-investors-4489723
https://www.investing.com/news/stock-market-news/amazon-shares-sink-as-big-techs-ai-spending-plans-worry-investors-4489723

39 https://www.parp.gov.pl/component/content/article/89279%3Aakt-w-sprawie-danych-nowe-regulacje-dot-wymiany-danych
https://www.parp.gov.pl/component/content/article/89279%3Aakt-w-sprawie-danych-nowe-regulacje-dot-wymiany-danych