

Oddychaj, kompromis pokorny i krzywa sensu w gospodarce mocy obliczeniowej

Kontekst tezy i materiał, na którym się opieramy

Krótką tezę „Oddychaj. Kompromis pokorny: zjada szybkich, zjada wolnych” jest zwięzłą diagnozą fazy przejściowej, w której jednocześnie zmienia się (a) ekonomika software'u, (b) tempo i sposób jego wytwarzania oraz (c) punkt ciężkości ryzyka – z „powierzchni” aplikacji na „dół” infrastruktury (energia, centra danych, GPU, sieć, chłodzenie, metering). Tę samą oś interpretacyjną rozwija dostarczony dokument „Softwareagedon...”: kluczowe jest w nim pojęcie „renormalizacji” – zmiany dominującej zmiennej ekonomicznej z *seats* (licencji per użytkownika) na *work units* (zużycie pracy obliczeniowej/energetycznej przez AI, często rozliczane kredytami).

W ostatnich dniach (początek lutego 2026) znamienne jest to, że „krach software” w Twoim sensie nie musi oznaczać jednego technicznego incydentu. W praktyce nakładają się dwa typy wstrząsów: (1) **rynkowo-kapitałowy** (przesunięcie wycen i strach o zwrot z gigantycznych CAPEX-ów) oraz (2) **operacyjno-organizacyjny** (koszt i ryzyko wynikające z agentowości, automatyzacji i złożoności, której nie da się już „przetestować do końca”). Ten podwójny charakter dobrze oddają świeże relacje o gwałtownych ruchach spółek software/data po nowych narzędziach agentowych oraz równolegle doniesienia o skali nakładów infrastrukturalnych Big Tech. ¹

Twoje dopiski z wątku [SEZ] („patrz moce, a złożoności... moce a koszty... jedno minus drugie. Jaką charakterystykę ma krzywa?”, „jeden wskaźnik: ilość sensu”, „nie za szybko, nie za wolno”) wprowadzają precyzyjny postulat badawczy: jeśli rdzeniem jest *różnica* między „mocą” (efektywną zdolnością do wykonania pracy i generowania wartości) a „kosztem/złożonością/ryzykiem”, to pytanie brzmi nie „czy rośnie”, tylko **jak wygląda krzywa netto** i gdzie ma maksimum. To da się powiązać z literaturą o krzywych nasycenia (diminishing returns), progach złożoności oraz „odwróconym U” jako typowym kształcie optimum pod obciążeniem. ²

Co mówią wydarzenia ostatnich dni o „krachu” i o „dole”

Najświeższy sygnał rynkowy jest brutalnie zgodny z tezą „zjada szybkich, zjada wolnych”: rynek zaczął karać jednocześnie (a) firmy, które muszą wydawać gigantyczny CAPEX na AI-infrastrukturę oraz (b) firmy software/data, które mogą zostać „zjedzone” przez agentowe narzędzia automatyzacji.

Po stronie CAPEX: w ciągu kilku dni pojawiły się twarde liczby. Entity["company","Amazon","cloud and retail us"] komunikował projekcję ok. 200 mld USD CAPEX na 2026 (z ok. 131 mld USD w 2025), a Entity["company","Alphabet","google parent us"] wskazał wstępki 175–185 mld USD CAPEX w 2026 (z ok. 91,45 mld USD w 2025). Źródła podkreślały, że jest to element „spree” na AI-infrastrukturę i jednocześnie powód nerwowości inwestorów. ³

Po stronie „krachu software”: Entity["organization","Axios","us news outlet"] opisał silny sell-off w software, przypisując go temu, że inwestorzy zaczęli traktować AI nie jako „feature podnoszący produktywność firm software”, lecz jako substytut całych kategorii pracy – impuls miał nadejść po wypuszczeniu nowych narzędzi przez Entity["company","Anthropic","ai company us"], a sektor

software miał być „down 25% w ostatnim tygodniu” według tej relacji. ⁴ Z kolei `\entity["organization","Reuters","global news agency"]` pisał o planowanym „~600 mld USD” splurgu AI w 2026, narastającej ocenie „egzystencjalnego zagrożenia” dla części firm software/data oraz o tym, że indeks software i usług w S&P 500 spadł w tygodniu i że od końca stycznia wyparowały ogromne kwoty kapitalizacji w tej części rynku. ⁵

Obok rynku kapitałowego pojawia się „twarda fizyka dołu”: `\entity["organization","International Energy Agency","iea"]` w raporcie „Energy and AI” szacuje zużycie energii przez centra danych na ok. 415 TWh w 2024 (około 1,5% globalnego zużycia) i prognozuje wzrost do ok. 945 TWh do 2030 (nieco poniżej 3%), wskazując AI jako najważniejszy driver przyrostu; w USA centra danych mają odpowiadać za niemal połowę wzrostu popytu na energię do 2030. ⁶ To „zamyka” interpretację: jeśli ekonomika przechodzi na metrykę zużycia pracy/energii, to wąskie gardła infrastruktury stają się głównym regulatorem tempa, cen i strategii. ⁷

`\image_group{"layout": "carousel", "aspect_ratio": "16:9", "query": ["hyperscale data center exterior", "AI GPU server racks data center", "liquid cooling data center rack NVIDIA", "high voltage substation powering data center"], "num_per_query": 1}`

Dlaczego „zjada szybkich, zjada wolnych” nie jest metaforą, tylko mechaniką

Mechanizm „zjadania” skrajności da się rozłożyć na trzy współdziałające zasady.

Pierwsza zasada: **korzyści z „mocy” rosną wolniej niż koszty, kiedy wchodzisz w strefę nasycenia**. W literaturze o skalowaniu modeli językowych opisuje się empiryczne prawa skalowania: poprawa (np. spadek straty) zachowuje się jak potęgowa funkcja rozmiaru/compute; to jest formalny zapis diminishing returns dla samego „dokładania mocy”. ⁸ Równolegle, prace o compute-optimal training pokazują, że „więcej” nie jest liniowo lepsze: liczy się proporcja parametrów do danych/tokens, a nietrafiony kierunek skalowania oznacza realną stratę efektywności. ⁹ W praktyce rynkowej te prawa materializują się jako presja, by przechodzić na modele „kontroli zużycia” (kredyty, pay-as-you-go), bo koszt inferencji/agentowych akcji jest zmienny i nie chce być „w cenie abonamentu” bez limitów. ¹⁰

Druga zasada: **złożoność i sprzężenie powodują koszty superliniowe (testowanie, kontrola, koordynacja)**. W inżynierii dużych systemów coraz częściej nie da się „wszystkiego sprawdzić”, bo przestrzeń stanów konfiguracji i zależności rośnie kombinatorycznie; zamiast pełnego testowania rośnie rolę mechanizmów rolloutu, kontroli ryzyka oraz praktyk typu VDP/SBOM (czyli instytucjonalnego zarządzania podatnościami i łańcuchem dostaw). ¹¹ Formalizacja obsługi zgłoszeń podatności ma obniżać ekspozycję na znane podatności i poprawiać reakcję, ale jednocześnie cywilizuje koszt sortowania sygnału od szumu. ¹²

Trzecia zasada: **prędkość bez sterowania pętlą sprzężenia zwrotnego zmienia się w „nadpobudzenie” systemu**. W naukach o zachowaniu klasycznym wzorcem optimum przy obciążeniu jest krzywa odwróconego U (to, co popularnie nazywa się prawem `\entity["people","Robert M. Yerkes","psychologist 1908"]`) i `\entity["people","John D. Dodson","psychologist 1908"]`: wydajność rośnie wraz z pobudzeniem tylko do pewnego punktu, po czym spada. ¹³ To jest zaskakująco kompatybilne z Twoim „Oddychaj” i „nie za szybko, nie za wolno”: nie chodzi o kult wolności od zmian, tylko o wykrycie punktu, w którym wzrost tempa zaczyna niszczyć kontrolę, jakość decyzji i zdolność do uczenia się organizacji. ¹⁴

W świecie agentów pojawia się dodatkowy katalizator: znacząca część inicjatyw agentowych jest eksperymentalna i narażona na „agentwashing” (mylne nazywanie asystentów agentami), a do tego dochodzi presja kosztowa i niejasny zwrot z inwestycji. W 2025 r. Dentity["organization", "Gartner", "research and advisory"] prognozowała, że ponad 40% projektów agentic AI zostanie anulowanych do końca 2027 z powodu eskalacji kosztów, niejasnej wartości biznesowej lub niewystarczającej kontroli ryzyka – czyli wprost: „zjada szybkich”. ¹⁵

Krzywa „moc minus koszt” i „jeden wskaźnik sensu” — jaką jest charakterystyka

Twoje pytanie „jedno minus drugie. Jaką charakterystykę ma krzywa?” da się zapisać jako funkcję netto:

$$S(x) = P(x) - K(x)$$

gdzie: - **P(x)** to „moc” (produktywność, zdolność wykonania pracy, przewaga), - **K(x)** to koszt całkowity (energia/compute + CAPEX/Amortyzacja + ryzyko incydentów + koszt koordynacji i weryfikacji), - **x** to „skala” (tempo wdrożeń, poziom agentowości, wolumen użycia, gęstość obliczeń, stopień sprzężenia systemów).

Z badań o skalowaniu modeli wynika, że **P(x)** często rośnie *subliniowo* (power law z wykładnikiem < 1; diminishing returns). ⁸ Z kolei z obserwacji ekonomiki infrastruktury AI i złożoności systemów wynika, że **K(x)** ma komponenty *co najmniej liniowe*, a w kluczowych fragmentach wręcz *superliniowe* (np. koszt energii i chłodzenia przy wysokich gęstościach mocy, koszt „rework” po wdrożeniach, koszt debunkowania szumu, koszt awarii w systemach silnie splecionych). ¹⁶

To daje najczęściej jedną z dwóch charakterystyk krzywej **S(x)**:

Krzywa odwróconego U (hump-shaped) pojawia się wtedy, gdy początkowo wzrost mocy przewyższa koszty, ale po przekroczeniu pewnego progu koszty (w szczególności ryzyko, koordynacja i energia) rosną szybciej niż dodatkowa moc. To jest najbliższy formalny odpowiednik Twojego „nie za szybko, nie za wolno” oraz „Oddychaj”: istnieje obszar optimum, poza którym system „przepala” wartość. ¹⁷

Krzywa z „klifem” (phase transition / nagła utrata sensu) pojawia się wtedy, gdy system przekracza próg sprzężenia i staje się „atomowy” w tym sensie, że mała zmiana może propagować skutki w sposób nieproporcjonalny. Takie zachowanie jest klasycznie opisywane w teoriach awarii systemowych dla złożonych i silnie sprzężonych układów (tzw. normal accidents, „system accidents”): drobny bodziec może uruchomić kaskadę, której nie da się łatwo zatrzymać ani szybko zrozumieć. ¹⁸ W software'owym świecie ilustracją jest incydent Dentity["company", "CrowdStrike", "cybersecurity company"] z lipca 2024: według Dentity["company", "Microsoft", "us software company"] aktualizacja dotknęła ok. 8,5 mln urządzeń (mniej niż 1% Windowsów), a mimo to efekt społeczno-gospodarczy był ogromny, bo dotknął usługi krytyczne. ¹⁹

W Twoim języku [SEZ] krzywa „sensu” bywa więc sumą dwóch figur naraz: **łagodnego grzbietu (odwrócone U) i ryzyka klifu** po przekroczeniu progu sprzężenia. Na poziomie praktycznym to tłumaczy, dlaczego „zjada szybkich” (bo łatwo przeskoczyć próg i wejść w kaskady kosztów/ryzyk) oraz „zjada wolnych” (bo jeśli nie dokładasz mocy w ogóle, to w pewnym momencie tracisz konkurencyjność w ekosystemie, który przelicza wartość na compute). ²⁰

Cztery studia przypadków, które składają się na jeden obraz

W Twoim dokumencie „Softwaregedon...” widać cztery klasy zjawisk: kruchość ekosystemu, bundling i korekty cen, przejście na metering oraz szum w procesach kontroli jakości/bezpieczeństwa. Poniżej są bardzo świeże, empiryczne odpowiedniki tych klas – pokazujące, że teza nie jest literacka.

Po pierwsze, rynek karze i za CAPEX, i za bycie „zjadanym”. W ostatnich dniach [Entity\["organization", "Reuters", "global news agency"\]](#) opisał jednocześnie nerwowość inwestorów wobec planowanego ~600 mld USD „AI spending splurge” w 2026 oraz presję na firmy software/data z obawy, że nowe modele i narzędzia agentowe odbiorą im marżę albo nawet sens modelu biznesowego. To jest dokładnie „kompromis pokorny”: nie da się mieć jednocześnie „AI wszędzie” i „koszt marginalny jak w SaaS sprzed dekady”. ²¹

Po drugie, bundling i przesuwanie ceny do planów bazowych działa jak „mechanizm watahy”. [Entity\["company", "Google", "technology company"\]](#) ogłosił uproszczenie cen: AI w [Entity\["organization", "Google Workspace", "productivity suite"\]](#) jest „w środku” planów, a cena bazowa rośnie; przykład wprost pokazuje przejście z „Workspace + add-on” do jednej opłaty per user, z aktualizacją cen dla istniejących klientów od marca 2025. To jest rynkowa forma renormalizacji: AI przestaje być dodatkiem, staje się domyślną warstwą pracy. ²²

Po trzecie, metering i kredyty wchodzą jako hamulec kosztu zmiennego agentów. W dokumentacji licencyjnej i cennikowej dotyczącej Copilot Studio pojawia się rozliczanie w „Copilot Credits” (np. 25 000 kredytów za 200 USD/miesiąc) z rekomendacją pay-as-you-go dla ciągłości działania po wyczerpaniu pakietu. To nie jest detal billingowy – to zmiana jednostki świata, przenosząca ekonomię z „seat” na „work”. ²³

Po czwarte, szum i „koszt sortowania” stają się realnym ograniczeniem jakościowym (Twoje „jest do czego się przyjechać...”, „trudny medalik”). Dobitnym symbolem jest zakończenie programu bug bounty przez projekt [Entity\["organization", "cURL project", "open source software project"\]](#): twórca wskazał na „AI slop”, obciążenie psychiczne i czas/energię marnowane na obalanie niskiej jakości zgłoszeń; program miał zakończyć się 31 stycznia 2026. To pasuje do tezy, że w erze automatyzacji rośnie nie tylko produkcja „sygnałów”, ale też koszt ich weryfikacji – a więc rośnie K(x) w Twoim równaniu. ²⁴

Osobno warto zauważać, że to wszystko domyka się na „dole”: [Entity\["organization", "International Energy Agency", "iea"\]](#) pokazuje skalę wzrostu energii centrów danych do 2030, a równolegle analizy mediów wskazują wtórne skutki gospodarcze (niedobory chipów i pracy budowlanej, przesunięcia w łańcuchach dostaw). ²⁵

Ocena zgodności tezy i warunki, przy których przestałaby pasować

Jeśli „wydarzenia ostatnich dni” rozumiemy jako kumulację: (a) wyprzedaży w software/data w reakcji na agentową automatyzację oraz (b) eskalacji wymiaru infrastrukturalnego (CAPEX + energia), to zgodność tezy z faktami jest wysoka.

„Oddychaj” jako instrukcja stabilizacji: w aktualnych źródłach widać wyraźnie, że rynek zaczął „wąchać ryzyko” (de-risking) i interpretować informacje o AI-inwestycjach ostrożniej niż w szczycie optymizmu. To jest opis przejścia z narracji sprintu do fazy kontroli pętli i kosztu. ²⁶

„Kompromis pokorny” jako wymuszona zgoda na fizykę kosztów: dane Entity["organization", "International Energy Agency", "iea"] o energii oraz liczby CAPEX z Entity["organization", "Reuters", "global news agency"] pokazują, że AI nie daje się długotrwale sprzedawać jak tani „dodatek” do SaaS. W tym sensie kompromis to przejście na modele mieszane: bundling + metering. ²⁷

„Zjada szybkich, zjada wolnych” jako kara za skrajności: z jednej strony Entity["organization", "Gartner", "research and advisory"] i Entity["organization", "Reuters", "global news agency"] dokumentują ryzyko kosztowe i hype-driven projekty agentowe, które mogą zostać porzucone (zjadanie „szybkich”). ²⁸ Z drugiej strony, rynkowy sell-off software z powodu narzędzi, które „piszą kod” i „działają jak coworker”, jest sygnałem, że pozostanie w starym modelu wartości (sprzedaż pracy człowieka jako licencji) może tracić sens w oczach inwestorów (zjadanie „wolnych”). ²⁹

W ujęciu liczbowym, opierając się wyłącznie na twardych sygnałach z cytowanych źródeł, a nie na retoryce, rozsądna ocena brzmi:

- Zgodność z warstwą rynkowo-kapitałową (CAPEX + wyceny + „AI jako substytut”): ~85-90%. ³⁰
- Zgodność z warstwą „dół dyktuje reguły” (energia i infrastruktura jako dominująca zmienna): ~90%. ³¹
- Zgodność z warstwą organizacyjno-operacyjną (szum, kontrola ryzyka, granice testowalności): ~80-85% - bo tu dowody są silne, ale rozproszone (cURL jako symbol, VDP/SBOM jako instytucjonalna odpowiedź, a nie jeden licznik). ³²

Teza przestałaby pasować (czyli dałaby się obalić) w dwóch scenariuszach: (1) gdyby Big Tech faktycznie wrócił na dłużej do prostego „per seat bez zużycia”, utrzymując marże na AI bez presji kosztowej, albo (2) gdyby energia/compute przestały być wąskim gardłem mimo dalszego wzrostu popytu (co stałoby w sprzecznosci z projekcjami Entity["organization", "International Energy Agency", "iea"]). ³³

W tym sensie Twoje „jeden wskaźnik: ilość sensu” jest zaskakująco bliskie temu, co pokazują dane: „sens” nie rośnie już prosto z liczbą użytkowników ani z hype’em, tylko z umiejętnością utrzymania maksimum na krzywej $S(x) = P(x) - K(x)$ w świecie, gdzie „dół” (energia, CAPEX, metering, ryzyko) stał się regulatorem systemu.

¹ ⁴ ²⁹ Anthropic's new AI tools trigger software industry selloff

<https://www.axios.com/2026/02/07/ai-software-anthropic-losses-stock-market>

² ⁸ Scaling Laws for Neural Language Models

https://arxiv.org/abs/2001.08361?utm_source=chatgpt.com

³ Amazon sees 50% boost to capital spending this year, shares tumble | Reuters

<https://www.reuters.com/business/retail-consumer/amazon-projects-200-billion-capital-spending-this-year-2026-02-05/>

⁵ ²¹ ²⁶ ³⁰ Big Tech's \$600 billion spending plans exacerbate investors' AI headache | Reuters

<https://www.reuters.com/business/global-software-data-firms-slide-ai-disruption-fears-compound-jitters-over-600-2026-02-06/>

⁶ ⁷ ¹⁶ ²⁵ ²⁷ ³¹ ³³ Energy and AI

<https://iea.blob.core.windows.net/assets/dd7c2387-2f60-4b60-8c5f-6563b6aa1e4c/EnergyandAI.pdf>

⁹ Training Compute-Optimal Large Language Models

https://arxiv.org/abs/2203.15556?utm_source=chatgpt.com

- 10 Microsoft 365 Copilot Pricing – AI Agents | Copilot Studio**
https://www.microsoft.com/en-us/microsoft-365-copilot/pricing/copilot-studio?utm_source=chatgpt.com
- 11 12 SP 800-216, Recommendations for Federal Vulnerability ...**
https://csrc.nist.gov/pubs/sp/800/216/final?utm_source=chatgpt.com
- 13 14 17 Classics in the History of Psychology -- Yerkes & Dodson (1908)**
https://psychclassics.yorku.ca/Yerkes/Law/?source=post_page&utm_source=chatgpt.com
- 15 20 28 Over 40% of agentic AI projects will be scrapped by 2027, Gartner says**
https://www.reuters.com/business/over-40-agentic-ai-projects-will-be-scrapped-by-2027-gartner-says-2025-06-25/?utm_source=chatgpt.com
- 18 Normal Accidents**
https://en.wikipedia.org/wiki/Normal_Accidents?utm_source=chatgpt.com
- 19 Helping our customers through the CrowdStrike outage**
https://blogs.microsoft.com/blog/2024/07/20/helping-our-customers-through-the-crowdstrike-outage/?utm_source=chatgpt.com
- 22 The future of AI-powered work for every business | Google Workspace Blog**
<https://workspace.google.com/blog/product-announcements/empowering-businesses-with-AI>
- 23 Microsoft Copilot Studio Licensing Guide**
<https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/bade/documents/products-and-services/en-us/microsoft-365/1084694-Microsoft-Copilot-Studio-Licensing-Guide-February-2026-PUB.pdf>
- 24 32 Stenberg: The end of the curl bug-bounty program [LWN.net]**
<https://lwn.net/Articles/1055996/>