ChatGPT

Obieg wartości intelektualnej w topologiach LLM, sieci agentowej i BTC z uwzględnieniem złożoności obliczeniowej oraz sieci energetycznej

Streszczenie wykonawcze

Wartość intelektualna w systemach LLM jest najtrafniej analizowana jako **aktywo niematerialne**: nie ma postaci fizycznej, ma generować przyszłe korzyści, a jej „posiadanie" zależy od mechanizmów kontroli (prawnych, kontraktowych i technicznych). To podejście jest spójne z opisem aktywów intelektualnych/„intangibles" oraz z ramami pomiaru „intellectual property products" jako wytworów R&D/innowacji, których użycie jest ograniczane ochroną prawną lub inną. [1]

W zcentralizowanej sieci LLM przepływ wartości jest łatwiejszy do **metryzacji i monetyzacji** (np. rozliczenia per token, platformy dystrybucji agentów/promptów), ale pojawia się silna **koncentracja ryzyk i praw** w operatorze usługi: regulaminy, polityki API, logowanie danych, dostęp do pamięci i narzędzi. Jednocześnie jest to środowisko szczególnie narażone na ataki typu black-box (model extraction), prompt injection i podatności łańcucha dostaw aplikacji LLM. [2]

W zdecentralizowanej sieci agentowej wartość intelektualna (dane, heurystyki, lokalne know-how) częściej pozostaje **blisko źródeł**, a transfer zachodzi poprzez protokoły i artefakty (np. federated learning). Zyskujemy odporność i potencjalnie lepszą prywatność „przez lokalność", ale rośnie trudność koordynacji praw/licencji i bezpieczeństwa oraz koszty komunikacji (w literaturze federated learning wskazywane jako główne ograniczenie). Dodatkowo aktualizacje/gradienty nie są automatycznie „bezpiecznym nośnikiem wartości" – możliwa jest rekonstrukcja danych z gradientów. [3]

Porównanie z BTC (Bitcoin) pokazuje skrajny przypadek decentralizacji: **wbudowaną trudność obliczeniową (PoW)** i **energetyczny koszt bezpieczeństwa**, jawny system zachęt (nagroda bloku i opłaty) oraz otwarty reżim własności IP implementacji (licencja MIT), przy jednoczesnej publiczności rejestru i probabilistycznej finalności. W przeciwieństwie do tego, w LLM „tokeny" są zwykle miarą tekstowo-kosztową (billing), a nie protokołową jednostką wartości i bezpieczeństwa. [4]

„Sieć energetyczna" jest wspólnym, często pomijanym, szkieletem wszystkich trzech topologii: przepływ wartości zależy od przepływu energii do data center / infrastruktury edge / kopalń oraz od lokalnych ograniczeń sieci elektroenergetycznych. International Energy Agency [5] szacuje, że zużycie energii przez data centres w 2024 wynosi ok. 415 TWh (~1,5% globalnej konsumpcji) i rosło ok. 12% rocznie w ostatnich latach; prognozy wskazują na szybkie przyrosty wraz z rozwojem AI. [6]
Dla BTC Cambridge Centre for Alternative Finance [7] przez CBECI publikuje bieżące estymacje zapotrzebowania na moc i rocznego zużycia energii, oparte m.in. na założeniu racjonalności ekonomicznej górników; analizy rządowe w USA przytaczają dla 2023 rzędy wielkości od dziesiątek do kilkuset TWh (z estymatą punktową ok. 120 TWh). [8]

Wreszcie, fakty z ostatniego tygodnia (zakres: 8–15 lutego 2026) wskazują na przyspieszenie przejścia od „modeli" do **zarządzania agentami**: ogłoszono i relacjonowano platformy oraz modele kładące nacisk

na agentowe działanie, integracje z narzędziami i granice uprawnień; równolegle rosną napięcia wokół distillation/„free-riding" i zastosowań w obszarach wysokiego ryzyka. [9]

Definicja wartości intelektualnej oraz prawa własności tej wartości

Definicja robocza i relacja do IP

W raporcie **wartość intelektualna** oznacza: *zdolność zasobu niematerialnego (wiedzy, danych, modeli, promptów, heurystyk, metadanych) do generowania przyszłej użyteczności i/lub korzyści ekonomicznej w procesach wspieranych przez LLM/agentów, przy jednoczesnym istnieniu mechanizmów kontroli umożliwiających legalny transfer albo zatrzymanie tej wartości*. Takie ujęcie jest zgodne z definicyjnymi cechami aktywów niematerialnych jako źródeł przyszłych zysków oraz z podejściem do produktów własności intelektualnej jako wytworzonych aktywów wynikających z badań/innowacji. [1]

WIPO [10] definiuje IP jako „creations of the mind" (m.in. wynalazki, utwory, wzory, oznaczenia), natomiast WTO [11] opisuje prawa własności intelektualnej jako prawa przyznawane osobom do wytworów umysłu, zwykle dające wyłączne uprawnienia w zakresie użycia. [12]

Formy wartości intelektualnej w LLM i agentach

Wartość intelektualna w praktyce LLM najczęściej „żyje" w kombinacji kilku artefaktów:

- **Wiedza i know-how** (jawne i ukryte). Mechanizmy retrieval-augmented generation (RAG) formalizują przeniesienie części wartości z wag modelu do pamięci nieparametrycznej, co daje szybkie aktualizacje i potencjalnie lepszą atrybucję źródeł. [13]
- **Dane** (surowe, etykiety, logi, embeddingi). Jeśli zawierają dane osobowe, wchodzą w reżim definicyjny GDPR (dane osobowe jako informacje o identyfikowalnej osobie). [14]
- **Modele** (wagi, adaptery, konfiguracje) oraz **techniki transferu** (fine-tuning, distillation). Distillation jest klasycznie opisywana jako metoda kompresji zachowania/„wiedzy" do mniejszego modelu. [15]
- **Prompty i workflow** (instrukcje, łańcuchy promptów, polityki narzędzi, zestawy testowe). Przykładem komercjalizacji jest GPT Store jako marketplace dla „custom versions" asystenta. [16]
- **Heurystyki** i polityki decyzji – w agentach, w modelach użytkowników oraz w warstwie bezpieczeństwa LLM. Wzorzec Reason+Act (ReAct) opisuje interleaving rozumowania z akcjami/narzędziami dla zwiększenia skuteczności i ograniczania halucynacji. [17]

Prawa własności „wartości" i ich egzekucja

W praktyce LLM prawa do wartości intelektualnej są „złożone warstwowo": część chroni prawo IP, część chronią licencje i kontrakty, a część – tajemnica przedsiębiorstwa (know-how). Dyrektywa UE o tajemnicach przedsiębiorstwa podkreśla, że ochrona dotyczy informacji nieujawnionych, o wartości handlowej wynikającej z poufności i objętych rozsądnymi działaniami ochronnymi. [18]

W topologii usługowej istotne są regulaminy dostawców. Warunki usług wskazują, że użytkownik zachowuje własność Input i posiada Output (w relacji użytkownik–dostawca, w zakresie dopuszczalnym prawem). To wpływa na to, czy wynik może być traktowany jako „aktywo" użytkownika. [19]

W BTC kontrola wartości tokenowej jest wykonywana protokołowo: transakcje wydają środki według reguł walidacji, a outputy transakcji są „pod kontrolą" tego, kto spełni warunek skryptu (w praktyce: kto posiada odpowiednie klucze). [20]

Topologie przepływu wartości oraz teza „agent silniejszy od modelu"

Trzy topologie: zcentralizowana sieć LLM, zdecentralizowana sieć agentowa, BTC

Poniżej trzy topologie jako trzy różne odpowiedzi na pytanie: *gdzie „mieszka" wartość i jak się ją transportuje?*

Zcentralizowana sieć LLM: wartość jest w dużej mierze skupiona w modelu i warstwie usługowej (API, polityki, marketplace, logi). [21]

Zdecentralizowana sieć agentowa: wartość jest bardziej rozproszona i często lokalna (dane, pamięci, heurystyki), a transfer odbywa się przez wymianę wyników, komponentów lub aktualizacji (federated learning). [22]

BTC: wartość tokenowa i jej transfer są natywne w protokole; bezpieczeństwo oparte jest o pracę obliczeniową i reguły walidacji pełnych węzłów. [23]

Diagramy architektury przepływu wartości

Zcentralizowana sieć LLM (wartość przez API, polityki i integracje)

```
flowchart LR
  U[Użytkownik] -->|dane+prompty| TOK[Tokenizacja]
  TOK --> API[API/Gateway]
  API --> GOV[Heurystyki LLM: safety,routing,cache]
  GOV --> LLM[Centralny LLM]
  LLM --> OUT[Output]
  OUT --> SYS[Systemy: aplikacje, RPA, BI]
  SYS --> VAL[Wartość: decyzje, przychód, oszczędność]

  subgraph AG[Agenci aplikacyjni]
    P[Planowanie]
    T[Wywołania narzędzi]
    E[Walidacja]
    P --> T --> E
  end
  API --> AG --> SYS
```

Zdecentralizowana sieć agentowa (wartość przez lokalność i federację)

```
flowchart TB
  subgraph A[Węzeł A]
```

```
    UMA[Model użytkownika + heurystyki prywatności]
    AGA[Agenci lokalni]
    DDA[Dane lokalne]
    UMA --> AGA <--> DDA
  end
  subgraph B[Węzeł B]
    UMB[Model użytkownika]
    AGB[Agenci lokalni]
    DDB[Dane lokalne]
    UMB --> AGB <--> DDB
  end
  subgraph FL[Federacja]
    SA[Secure Aggregation]
    AGG[Agregator aktualizacji]
    GM[Model globalny]
    SA --> AGG --> GM
  end
  A -->|aktualizacje| SA
  B -->|aktualizacje| SA
  GM -->|wersje| A
  GM -->|wersje| B
```

BTC (wartość i bezpieczeństwo w protokole + energia jako koszt)

```
flowchart LR
  W[Portfele / użytkownicy] -->|transakcje| P2P[Sieć P2P]
  P2P --> N[Pełne węzły: walidacja reguł]
  P2P --> M[Mempool]
  M --> MIN[Miners: Proof-of-Work]
  MIN -->|blok + coinbase| N
  N -->|akceptacja/odrzucenie| CHAIN[Łańcuch bloków]
  CHAIN -->|saldo/UTXO| W
  MIN -->|koszt energii| GRID[Sieć elektroenergetyczna]
  CHAIN -->|opłaty+subsidy| MIN
```

Reguły coinbase w dokumentacji Bitcoin Developer Guide wiążą block reward z block subsidy i opłatami, a emisja i difficulty adjustment są w praktyce mechanizmami kontrolującymi rzadkość i tempo bloków. [24]

Teza użytkownika: „agent jest silniejszy od modelu, bo człowiek chce mieć wartości w kieszeni”

Tę tezę można ująć precyzyjnie jako twierdzenie o **miejscu zakotwiczenia wartości**:

- „Model” jest źródłem zdolności generatywnej/poznawczej, ale w praktyce jest wymienialny (modele się starzeją, zmieniają wersje, zmienia się dostawca lub licencja).

- „Agent” jest **opakowaniem normatywno-własnościowym**: przenosi i egzekwuje wartości użytkownika/organizacji poprzez (a) pamięć i kontekst, (b) polityki uprawnień, (c) narzędzia, (d) walidację oraz (e) reguły retencji danych.

Właśnie tę warstwę „kieszeni na wartości” widać w języku produktów agentowych z ostatniego tygodnia: ogłoszona platforma klasy enterprise do budowy i zarządzania agentami podkreśla „shared context”, „permissions”, „boundaries” i procesy wdrażania agentów do organizacji. [25]
Analogicznie, komunikaty o modelach „agentic tasks” i „operate with tools” podkreślają zdolność utrzymywania dłuższych zadań, pracy z narzędziami i granicami wykonania. [26]

W skrócie: agent staje się silniejszy jako *nośnik kontroli i własności wartości* (w tym prywatności i zgodności), podczas gdy model pozostaje *silnikiem obliczeniowym* – bardzo ważnym, ale częściej wymienialnym.

Trudność i złożoność obliczeniowa: podobieństwa i różnice

Co znaczy „trudność” w trzech topologiach?

W BTC trudność jest zdefiniowana protokołowo: difficulty adjustment (co 2016 bloków) ma utrzymywać średni czas bloku ~10 minut, a bezpieczeństwo wynika z kosztu pracy obliczeniowej (PoW) i reguł walidacji pełnych węzłów. [27]

W LLM „trudność” ma dwa oblicza:

1) **Złożoność obliczeniowa inferencji i treningu** (ile operacji i pamięci wymaga model),
2) **Złożoność systemowa agentów** (wiele pętli: planowanie → narzędzia → walidacja → pamięć → polityki).

W federated learning dochodzi trzeci wymiar: **złożoność komunikacyjna** (ile i jak często przesyłamy aktualizacje), wskazywana jako główne ograniczenie praktyczne. [28]

Złożoność obliczeniowa „silnika” LLM i konsekwencje dla przepływu wartości

Transformery są projektowane tak, by zredukować sekwencyjność i zwiększyć paralelizację; autorzy porównują typy warstw pod kątem złożoności per warstwa, liczby operacji sekwencyjnych i długości ścieżek zależności. W szczególności, w sekcji „Why Self-Attention” i tabeli porównawczej wskazują, że samo-uwaga łączy pozycje stałą liczbą operacji sekwencyjnych i omawiają relacje złożoności obliczeniowej względem długości sekwencji i wymiaru reprezentacji. [29]

Dla obiegu wartości ma to praktyczne skutki: - w centralizacji koszt rośnie wraz z wolumenem tokenów i długością kontekstu (bo rośnie praca obliczeniowa i pamięć), co bezpośrednio mapuje się na billing per-token i na decyzje o RAG/kompresji kontekstu; [30]
- w agentach rośnie koszt „kolekcji kroków” (więcej wywołań, walidacji, narzędzi), ale jednocześnie rośnie kontrolowalność wartości (granice uprawnień i weryfikowalność). [31]

Złożoność obliczeniowa walidacji vs generacji: BTC kontra LLM

Jedna z kluczowych asymetrii (ważna dla praw i własności wartości) to relacja **„łatwo zweryfikować vs trudno wytworzyć"**:

- BTC: wytworzenie bloku wymaga kosztownej pracy (PoW), ale walidacja przez pełne węzły jest projektowana jako wykonalna i deterministyczna według reguł konsensusu. [32]
- LLM: generacja jest „łatwa" po stronie użytkownika (wywołanie API), ale walidacja semantyczna (czy output jest prawdziwy i zgodny z wartościami/polityką) jest trudniejsza i często wymaga dodatkowych agentów-weryfikatorów, RAG i mechanizmów provenance. [33]

To wspiera tezę „agent > model" w sensie kontroli wartości: agent buduje warstwę weryfikacji i egzekucji polityk.

Porównawcza tabela złożoności i „trudności"

Wymiar trudności/ złożoności	Zcentralizowana sieć LLM	Zdecentralizowana sieć agentowa	BTC
Dominująca „twarda" trudność	Koszt obliczeń inferencji/treningu; ekonomizacja w tokenach	Koszt inferencji + koszt koordynacji i komunikacji; bezpieczeństwo aktualizacji	Trudność PoW regulowana protokołowo (difficulty); koszt ataku rośnie z hashrate/energią
Złożoność obliczeniowa rdzenia	Samo-uwaga: złożoność zależna od długości sekwencji i wymiaru reprezentacji; nacisk na paralelizację	Jak w LLM + koszty pętli agentowych (wiele kroków)	Generacja bloku kosztowna, walidacja reguł przez pełne węzły projektowana jako wykonalna
Złożoność komunikacyjna	Głównie ruch API i retrieval	Kluczowa w FL (komunikacja jako główne ograniczenie)	P2P propagacja transakcji/bloków; spójność przez reguły konsensusu
„Walidacja wartości"	Polityki operatora i ewaluatory; trudna walidacja semantyczna	Częściej lokalna walidacja i wieloetapowe sprawdzanie	Deterministyczna walidacja reguł + probabilistyczna finalność zależna od potwierdzeń

Uzasadnienia elementów: złożoności self-attention i porównania warstw są omawiane w pracy o Transformerach; komunikacja jako główne ograniczenie federated learning jest opisane w klasycznej pracy FedAvg; mechanizmy difficulty adjustment i docelowy czas bloku wynikają z praktyki protokołu i są szeroko dokumentowane. [34]

Sieć energetyczna jako „ukryta topologia” obiegu wartości

Wspólny mianownik: energia jako warunek istnienia wartości

W każdej z trzech topologii przepływ wartości jest ograniczony przepływem energii:

- Centralne LLM: energia zasila centra danych (GPU, chłodzenie), a ograniczenia sieciowe mogą stać się wąskim gardłem dla rozwoju infrastruktury. [35]
- Sieci agentowe: energia rozkłada się na węzły edge/on-prem; może zmniejszać obciążenie centrów, ale może duplikować obliczenia i przenosić koszty do uczestników. [36]
- BTC: energia jest „skonsumowana” w PoW jako mechanizm bezpieczeństwa i emisji/nagród. [37]

Fakty energetyczne: AI/data centres i BTC

IEA szacuje, że zużycie energii przez data centres wynosi ok. 415 TWh w 2024 (~1,5% globalnej energii elektrycznej) i rosło ok. 12% rocznie w ostatnich pięciu latach; prognozy wskazują, że globalna generacja energii na potrzeby data centres może przekroczyć 1 000 TWh w 2030 (w scenariuszu bazowym). [38]

CBECI publikuje bieżące estymacje mocy i zużycia energii BTC, a metodologia opiera się m.in. na koszyku sprzętu i założeniu racjonalności ekonomicznej górników (uruchamiają sprzęt, gdy jest opłacalny). [39]
U.S. Energy Information Administration [40] przytacza dla 2023 r. estymacje CBECI w zakresie 67–240 TWh (punktowo ok. 120 TWh) i odnosi je do globalnego zużycia energii. [41]

Energetyka jako sieć zależności wartości i praw własności

W topologiach AI energia staje się częścią „praw do wartości” w dwóch sensach:

1) **Prawa do danych i outputu** (kontraktowe/prawne) są bezprzedmiotowe, jeśli infrastruktura energetyczna i obliczeniowa jest niedostępna lub zbyt droga – w praktyce rosną koszty tokenowe/compute, a więc zmienia się ekonomika transferu wartości. [42]
2) **Lokalność energii** wpływa na lokalność danych: jeśli przetwarzanie przenosi się do edge (bo energia/połączenia są tańsze lokalnie albo dlatego, że prawo/ryzyko wymusza lokalność danych), rośnie znaczenie agentów jako warstwy, która pozwala „przenieść wartości w kieszeni” (polityki, pamięć, uprawnienia), niezależnie od tego, jaki model jest używany w danym miejscu. [43]

W BTC zależność jest jeszcze ostrzejsza: energia jest bezpośrednio przekształcana w bezpieczeństwo i nagrody; stąd stabilność i geografia sieci energetycznej stają się elementem systemu wartości (i sporów regulacyjnych). [44]

Ekonomia i mechanizmy transferu wartości: LLM/agent vs BTC

LLM: tokenizacja kosztu i platformizacja wartości

Rozliczenia per token sprawiają, że tokenizacja jest jednocześnie jednostką obliczeniową i jednostką kosztu (transfer wartości jest ekonomicznie „skwantowany” w tokenach). [45]
Mechanika marketplace (np. GPT Store) jest zgodna z teorią platform wielostronnych: platforma umożliwia interakcje grup użytkowników, a model biznesowy zależy od przyciągnięcia i skoordynowania obu stron rynku. [46]

OECD [47] zwraca uwagę na rolę aktywów intelektualnych w modelach biznesowych (zwłaszcza firm opartych o kapitał intelektualny); w LLM praktycznym odpowiednikiem jest fakt, że prompty, modele wyspecjalizowane i workflow stają się aktywami, które można dystrybuować i monetyzować. [48]

Zdecentralizowana sieć agentowa: zachęty, wycena wkładów i koszty komunikacji

W federated learning komunikacja jest wskazywana jako główne ograniczenie, a jednocześnie istnieje intensywna literatura o mechanizmach zachęt do udziału i o wycenie wkładów (dane/obliczenia). [49]
Secure aggregation zwiększa prywatność wkładów, ale nie eliminuje całej klasy ryzyk inferencji (gradient leakage), co komplikuje zarówno bezpieczeństwo, jak i ekonomię „rynku wkładów". [50]

BTC: opłaty, subsidy, kontrolowana emisja i fee market

W dokumentacji Bitcoin Developer Guide coinbase transaction ma zbierać block reward, tj. subsidy + opłaty transakcyjne. [51]
Kontrolowana emisja opisuje halving co 210 000 bloków, a difficulty adjustment co 2016 bloków ma utrzymywać średni czas bloku. [52]
To daje BTC bardzo silną cechę: prawa do wartości tokenowej są egzekwowane „mechanicznie" przez reguły protokołu, a nie przez regulaminy pojedynczego operatora.

Fakty z ostatniego tygodnia istotne dla obiegu wartości

W okresie 8–15 lutego 2026 (kontekst: „ostatni tydzień" względem 15 lutego 2026) wydarzenia istotne dla przepływu wartości oraz praw do tej wartości obejmują:

- Ogłoszenie i relacje o platformie enterprise do budowy i zarządzania agentami (akcent na uprawnienia, granice i kontekst współdzielony), co wzmacnia przesunięcie wartości z „samego modelu" na „warstwę agenta i governance". [53]
- Wydania i komunikaty modeli podkreślających pracę „z narzędziami" i dłuższe zadania agentowe, co sugeruje, że przewaga konkurencyjna przesuwa się w stronę utrzymywania kontekstu, planowania, narzędzi i weryfikacji. [54]
- Spór narracyjno-polityczny wokół distillation i „free-riding": w relacjach Reuters pojawia się motyw wykorzystywania distillation przez podmioty trzecie do replikowania możliwości modeli, co jest kluczowe dla ryzyk IP i dla ekonomiki platform LLM. [55]
- Rosnące napięcia wokół wdrożeń AI w obszarach wysokiego ryzyka (np. użycie w sieciach klasyfikowanych i presja na redukcję ograniczeń), co bezpośrednio wpływa na polityki governance i odpowiedzialność prawno-organizacyjną. [56]

Ryzyka, KPI i trzy krótkie case studies przepływu wartości

Ryzyka utraty wartości intelektualnej i ich relacja do trudności obliczeniowej

W LLM i agentach kluczowe klasy ryzyk obejmują:

- **Model extraction** przez API (kradzież funkcji modelu) – udokumentowana jako wykonalna w scenariuszach usługowego ML. [57]
- **Ekstrakcja danych treningowych** z modeli językowych (odzyskiwanie przykładów treningowych) oraz **membership inference** (czy rekord był w zbiorze treningowym), co uderza w prywatność i wartość danych. [58]
- **Prompt injection i podatności łańcucha dostaw** – ujęte w taksonomii OWASP Top 10 dla aplikacji LLM. [59]

- **Gradient leakage** w uczeniu rozproszonym – pokazuje, że aktualizacje mogą nieść informację o danych. [60]

W BTC ryzyka mają inną naturę: atakujący musiałby kontrolować znaczną moc obliczeniową, bo bezpieczeństwo opiera się o pracę i reguły konsensusu – to jest wprost projektowane jako kosztowne. [61]

KPI oceny obiegu wartości z uwzględnieniem energii i złożoności

W praktyce warto monitorować KPI w tym układzie:

- **Wartość/produktywność**: value-per-token, ROI fine-tuning vs ROI prompt/workflow, czas aktualizacji wiedzy (RAG). [62]
- **Złożoność i koszt**: rozkłady opóźnień p50/p95/p99 dla zadań agentowych i całych „misji"; koszt komunikacji w FL na jednostkę poprawy jakości. [63]
- **Energia jako KPI infrastruktury**: energia na jednostkę pracy (np. na 1 mln tokenów / na zadanie agentowe), wykorzystanie mocy i „wąskie gardła" lokalnych sieci; sensowność tego KPI rośnie wraz z prognozami wzrostu zużycia energii przez data centres. [35]
- **Ochrona IP i prywatności**: wskaźniki prób model extraction, testy podatności na ekstrakcję danych treningowych, incydenty prompt injection. [64]
- **BTC-analogiczne metryki decentralizacji** (dla sieci agentowych): dystrybucja „mocy" (kto kontroluje kluczowe zasoby), odporność na awarie i zdolność do weryfikacji wkładów.

Trzy case studies ilustrujące przepływ wartości

Komercyjne wykorzystanie promptów i modeli

Firma tworzy kompozycję: prompty + narzędzia + walidator i dystrybuuje ją w marketplace. Wartość powstaje w workflow i heurystykach, a rozliczenie odbywa się per token / przez platformę. [65]
Ryzyko: prompt injection i wyciek tajemnicy przedsiębiorstwa (jeśli prompty/heurystyki są kluczowe). [66]

Współdzielenie wiedzy w federated learning

Konsorcjum trenuje model bez wynoszenia surowych danych. Wartość transferuje się w aktualizacjach; kosztem jest komunikacja, a ochroną – secure aggregation. [67]
Ryzyko: rekonstrukcja danych z gradientów → wymaga dodatkowych zabezpieczeń, bo „aktualizacja" może być nośnikiem niechcianej informacji. [60]

Atak na własność intelektualną

Konkurent przeprowadza model extraction przez API i/lub wykorzystuje distillation jako narzędzie odtworzenia funkcji modelu; równolegle testuje podatności na wyciek danych treningowych. Wykonalność model extraction i training data extraction jest udokumentowana w literaturze. [68]
Kontekst z ostatniego tygodnia pokazuje, że temat distillation/free-riding jest także elementem gry regulacyjno-biznesowej, a nie wyłącznie kwestią techniczną. [55]

Rekomendacje: jak optymalizować obieg wartości i praw do niej

Architektura i polityki: „wartości w kieszeni” jako projekt agenta

Jeśli celem jest, by człowiek/organizacja „miała swoje wartości w kieszeni”, praktyczna recepta brzmi: **agent jako warstwa własności i kontroli**, model jako wymienialny silnik.

- Buduj agentów jako kompozycje: pamięć + polityki uprawnień + narzędzia + walidacja + mechanizm atrybucji/provenance. To jest spójne zarówno z praktyką agentic workflows, jak i z trendem produktów enterprise akcentujących permissions i boundaries. 69
- Tam, gdzie kluczowe są prawa i dowodliwość pochodzenia, wdrażaj standardy provenance dla treści (np. C2PA) oraz dokumentowanie modeli (model cards) jako narzędzia separacji dozwolonych zastosowań od ryzyk i wzmocnienia zaufania. 70

Ochrona IP i prywatności: kontrola logów, tajemnica, licencje, compliance

- Traktuj prompty/heurystyki jako potencjalne tajemnice przedsiębiorstwa: kontrola dostępu, retencja logów, rozdział środowisk, polityki „need-to-know”, bo ochrona w UE zależy od poufności i rozsądnych działań ochronnych. 71
- Projektuj „minimalizację danych” w pamięci agentów i w logowaniu; definicja danych osobowych jest szeroka, więc brak dyscypliny logów może zjeść wartość (ryzyko prawne i reputacyjne). 72
- Utrzymuj metadane licencyjne dla modeli/danych/narzędzi w repozytoriach i pipeline’ach, bo brak jasności licencyjnej potrafi zatrzymać legalny transfer wartości. 73
- Wykorzystuj ramy zarządzania ryzykiem AI (np. AI RMF) do operacjonalizacji kontroli i audytu w całym cyklu życia systemu. 74

Bezpieczeństwo aplikacji agentowych: obrona przed kradzieżą funkcji i prompt injection

- Stosuj detekcję i ograniczanie model extraction (rate limiting, watermarkowanie interfejsów, red-teaming „fidelity stealing”), bo ataki tego typu mają silne podstawy badawcze. 57
- Traktuj prompt injection i insecure output handling jako ryzyko pierwszego rzędu; buduj separację zaufania między outputem modelu a akcjami narzędzi, zgodnie z taksonomią OWASP dla aplikacji LLM. 59

Artefakty operacyjne do utrzymania

Aby obieg wartości i praw do niej był mierzalny, utrzymuj (jako „żywe” artefakty):

- rejestr artefaktów wartości (dane, prompty, heurystyki, modele, indeksy RAG) z właścicielem, licencją, poufnością i ścieżkami dystrybucji; 75
- rejestr transferów (API, fine-tuning, distillation, FL) z kosztem tokenowym/energetycznym i oceną ryzyka; 76
- diagramy mermaid pokazujące: przepływy danych, granice uprawnień agentów oraz zależności energetyczne (gdzie „energia warunkuje wartość”). 77

navlistŹródła z ostatniego tygodnia o agentach i ryzykachturn14news39,turn14news40,turn14news38,turn13news32,turn13news31

---

1 40 48 Intellectual Assets and Innovation (EN)
https://www.oecd.org/content/dam/oecd/en/publications/reports/2011/12/intellectual-assets-and-innovation_g1g1458e/9789264118263-en.pdf?utm_source=chatgpt.com

2 7 21 30 42 45 62 76 What are tokens and how to count them?
https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them?utm_source=chatgpt.com

3 22 28 36 47 Communication-Efficient Learning of Deep Networks from Decentralized Data
https://arxiv.org/abs/1602.05629?utm_source=chatgpt.com

4 23 37 61 https://bitcoin.org/bitcoin.pdf
https://bitcoin.org/bitcoin.pdf

5 29 34 https://arxiv.org/html/1706.03762v7
https://arxiv.org/html/1706.03762v7

6 35 38 77 https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai
https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

8 39 44 https://ccaf.io/cbnsi/cbeci/methodology
https://ccaf.io/cbnsi/cbeci/methodology

9 25 43 53 69 Introducing OpenAI Frontier
https://openai.com/index/introducing-openai-frontier/?utm_source=chatgpt.com

10 70 https://c2pa.org/specifications/specifications/2.3/specs/C2PA_Specification.html
https://c2pa.org/specifications/specifications/2.3/specs/C2PA_Specification.html

11 17 31 https://arxiv.org/abs/2210.03629
https://arxiv.org/abs/2210.03629

12 What is Intellectual Property?
https://www.wipo.int/en/web/about-ip?utm_source=chatgpt.com

13 33 Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks
https://arxiv.org/abs/2005.11401?utm_source=chatgpt.com

14 https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng

15 Distilling the Knowledge in a Neural Network
https://arxiv.org/abs/1503.02531?utm_source=chatgpt.com

16 46 65 https://openai.com/index/introducing-the-gpt-store/
https://openai.com/index/introducing-the-gpt-store/

18 DIRECTIVE (EU) 2016/ 943 OF THE EUROPEAN ... - EUR-Lex
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX%3A32016L0943&utm_source=chatgpt.com

19 Terms of Use
https://openai.com/policies/row-terms-of-use/?utm_source=chatgpt.com

20 https://developer.bitcoin.org/reference/transactions.html
https://developer.bitcoin.org/reference/transactions.html

24 32 51 https://developer.bitcoin.org/devguide/block_chain.html
https://developer.bitcoin.org/devguide/block_chain.html

26 54 Introducing Claude Opus 4.6
https://www.anthropic.com/news/claude-opus-4-6?utm_source=chatgpt.com

27 https://www.blockchain.com/charts/difficulty
https://www.blockchain.com/charts/difficulty

41 https://www.eia.gov/todayinenergy/detail.php?id=61364
https://www.eia.gov/todayinenergy/detail.php?id=61364

49 63 67 Communication-Efficient Learning of Deep Networks from ...
https://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf?utm_source=chatgpt.com

50 https://arxiv.org/abs/1611.04482
https://arxiv.org/abs/1611.04482

52 https://en.bitcoin.it/wiki/Controlled_supply
https://en.bitcoin.it/wiki/Controlled_supply

55 OpenAI says China's DeepSeek trained its AI by distilling ...
https://www.reuters.com/world/china/openai-accuses-deepseek-distilling-us-models-gain-advantage-bloomberg-news-2026-02-12/?utm_source=chatgpt.com

56 Pentagon pushing AI companies to expand on classified networks, sources say
https://www.reuters.com/business/pentagon-pushing-ai-companies-expand-classified-networks-sources-say-2026-02-12/?utm_source=chatgpt.com

57 64 https://arxiv.org/abs/1609.02943
https://arxiv.org/abs/1609.02943

58 https://arxiv.org/abs/2012.07805
https://arxiv.org/abs/2012.07805

59 66 https://owasp.org/www-project-top-10-for-large-language-model-applications/
https://owasp.org/www-project-top-10-for-large-language-model-applications/

60 https://arxiv.org/abs/1906.08935
https://arxiv.org/abs/1906.08935

68 https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_tramer.pdf
https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_tramer.pdf

71 Directive - 2016/943 - EN - EUR-Lex - European Union
https://eur-lex.europa.eu/eli/dir/2016/943/oj/eng?utm_source=chatgpt.com

72 https://gdpr-info.eu/art-4-gdpr/
https://gdpr-info.eu/art-4-gdpr/

73 75 https://huggingface.co/docs/hub/en/repositories-licenses
https://huggingface.co/docs/hub/en/repositories-licenses

74 https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf
https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf