dr inż. Michał Malinowski

https://www.drmalinowski.edu.pl

Temat 10.

Czy AI jest etyczne? Prawo, moralność i przyszłość technologii

Wprowadzenie do Sztucznej Inteligencji



- Global AI Index
- AI Compliance
- Prawo a Sztuczna Inteligencja w Polsce
- Wspólna Wizja Bezpiecznego Rozwoju AI
- AI Act, Projekt ustawy o systemach AI (PL)
- ISO/IEC 42001:2023
- RODO a AI, AI w MON
- AI a prawa autorskie
- Data Governance
- Responsible AI i Alignment AI
- Agenci AI – nowe wydanie
- Rynek pracy a AI

Global AI Index

to ranking, który ocenia i porównuje kraje pod względem ich zaawansowania w dziedzinie sztucznej inteligencji (AI). Tworzony jest przez firmę **Tortoise Media** uwzględnia różne aspekty rozwoju AI, takie jak:

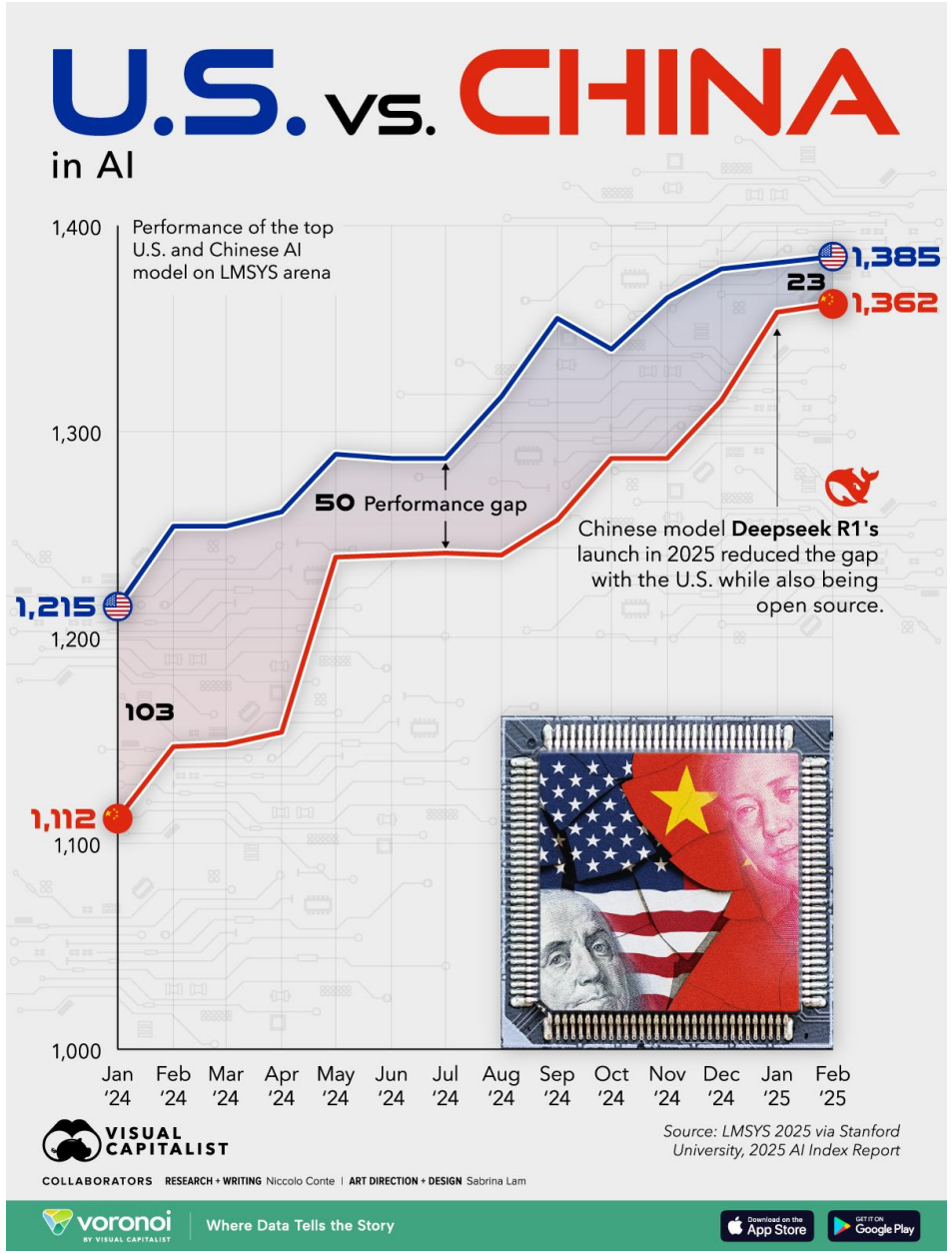
1. **Inwestycje w sztuczną inteligencję** (publiczne i prywatne).
2. **Potencjał badawczo-rozwojowy** (liczba publikacji naukowych, patenty, innowacje).
3. **Zdolność do wdrażania AI** w gospodarce i sektorze publicznym.
4. **Kapitał ludzki** (liczba specjalistów i talentów w AI).
5. **Infrastruktura AI** (dostępność technologii, superkomputerów i zasobów).



Global AI Index
https://www.tortoisemedia.com/intelligence/global-ai

Liderzy AI

- W styczniu 2024 roku najlepszy model amerykański miał przewagę 9,26% nad najlepszym chińskim modelem na LMSYS Chatbot Arena
- Do lutego 2025 roku ta różnica zmniejszyła się do 1,70%
- Przyczyną tego zbliżenia jest m.in. wprowadzenie chińskiego modelu DeepSeek R1, który osiągnął bardzo wysokie wyniki, będąc dodatkowo modelem otwartym
- Model DeepSeek R1 wykorzystuje znacznie mniej zasobów obliczeniowych niż wcześniejsze modele tego poziomu



Artificial Intelligence Index Report 2025
https://hai-production.s3.amazonaws.com/files/hai_ai_index_report_2025.pdf

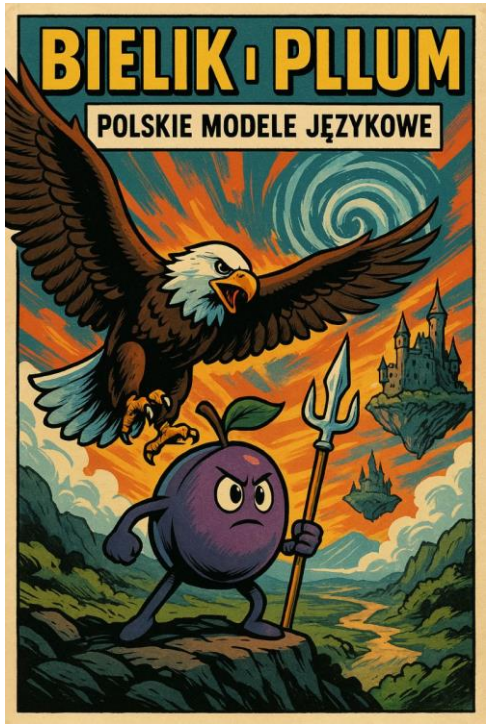
Polska

Polskie modele NLP

https://bielik.ai

https://pllum.clarin-pl.eu

Kategoria	Bielik	PLLuM 8x7B
Architektura	Decoder-only (transformer), oparty na Mistral-7B	Mixture of Experts: 8x7B, aktywowane 2 eksperci per token (efektywnie ~13B)
Parametry	7B (v0.1), 11B (v2.0)	8 x 7B (56B łącznie), aktywne 12.9B
Język	Polski	Polski + inne języki słowiańskie i bałtyckie + angielski
Dane treningowe	70+ miliardów tokenów tekstu w języku polskim	150 miliardów tokenów (PL + inne języki)
Zastosowania	ASR, klasyfikacja tekstu, QA, analiza sentymentu	Chatboty, QA, generacja tekstu, podsumowanie, asystenci specjalistyczni
Do pobrania	Hugging Face (Bielik-7B, Bielik-11B)	Hugging Face (PLLuM-8x7B)
Licencja	Open (Hugging Face, do zastosowań badawczych i wdrożeń)	Open (Hugging Face, zgodna z polityką CLARIN-PL i Cyfronetu)
Dodatkowe atuty	Specjalizacja w języku polskim	Obsługa wielu języków regionu Europy Środkowo-Wschodniej
Twórca modelu	Fundacja SpeakLeash i partnerzy (np. NASK, IPI PAN)	Cyfronet AGH, ICM UW, CLARIN-PL, NCBR, MEiN



AI Compliance



Nie mogę zaprojektować sprawiedliwego systemu AI…

Bo nie ma jednej definicji sprawiedliwości.

???

Dlaczego?

To prawda. A która definicja jest ważna dla twoich interesariuszy?



AI Compliance

AI Compliance

to zestaw zasad, regulacji i praktyk zapewniających, że systemy sztucznej inteligencji działają zgodnie z prawem, normami etycznymi i bezpieczeństwa.

Wyzwania AI Compliance

⚠ **Brak jednolitych standardów** – Różne kraje mają odmienne przepisy.
⚠ **Trudność w monitorowaniu AI** – Modele uczą się dynamicznie, co utrudnia kontrolę.
⚠ **Odpowiedzialność prawna** – Kto odpowiada za błędy AI?

Jak zapewnić AI Compliance?

✅ **Audyt AI** – Regularne przeglądy działania modeli.
✅ **Explainability (przejrzystość)** – AI powinno wyjaśniać swoje decyzje.
✅ **Etyczne programowanie** – Unikanie stronniczości i manipulacji danymi.
✅ **Bezpieczeństwo danych** – AI musi przestrzegać norm ochrony prywatności.

Dokumenty normujące Sztuczną Inteligencję w Polsce

Dokument	Rok	Zakres	Wpływ na Polskę	Link
Zasady OECD dotyczące sztucznej inteligencji (OECD AI Principles)	2019	Globalne ramy promujące sprawiedliwy i zrównoważony rozwój AI z poszanowaniem praw człowieka i wartości demokratycznych	Polska jako członek OECD zobowiązana jest do przestrzegania tych zasad	https://www.oecd.org/en/topics/ai-principles.html
Wytyczne UE dotyczące wiarygodnej sztucznej inteligencji (AI Ethics Guidelines for Trustworthy AI)	2019	Zasady dotyczące legalności, etyki i niezawodności AI	Polska zobowiązana do ich przestrzegania w ramach UE	https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
Polityka rozwoju sztucznej inteligencji w Polsce	2020	Strategia rozwoju AI w Polsce, promowanie transparentności i odpowiedzialności	Krajowa strategia rozwoju AI zgodna z regulacjami unijnymi	https://www.gov.pl/attachment/0a7fa2b7-236c-4081-8b03-f43a769fbee5
Rekomendacje UNESCO dotyczące etyki sztucznej inteligencji	2022	Globalne zasady etyki AI, promowanie sprawiedliwości społecznej i przejrzystości	Polska jako członek UNESCO popiera te zalecenia	https://unesdoc.unesco.org/ark:/48223/pf0000381137
EU Coordinated Plan on AI	2021 (aktualizowany)	Plan koordynacji polityk AI w UE, obejmujący rozwój innowacji, inwestycje w AI oraz harmonizację regulacyjną w krajach członkowskich.	Polska jako kraj UE wdraża elementy strategii w ramach polityki cyfryzacji i regulacji AI.	https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review
Deklaracja z Bletchley	2023	Międzynarodowe porozumienie dotyczące bezpieczeństwa AI, zobowiązujące do opracowania wspólnych standardów i badań nad ryzykiem AI	Polska jako sygnatariusz deklaracji zobowiązała się do wspierania globalnej współpracy w zakresie bezpiecznego rozwoju AI	https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration
G7 Hiroshima Process on Generative AI	2023	Międzynarodowe ramy regulacyjne dotyczące generatywnej sztucznej inteligencji, w tym transparentność, bezpieczeństwo i odpowiedzialność modeli AI.	Polska jako kraj UE może implementować wytyczne w ramach własnych strategii AI, szczególnie w kontekście zgodności z AI Act.	https://www.oecd.org/en/publications/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en.html

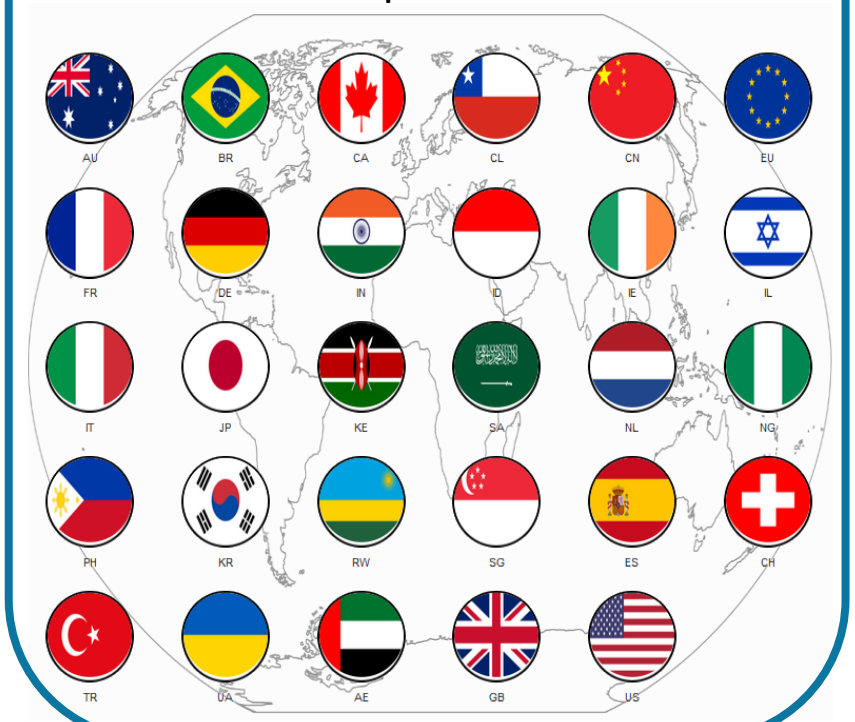
Dokumenty normujące Sztuczną Inteligencję w Polsce

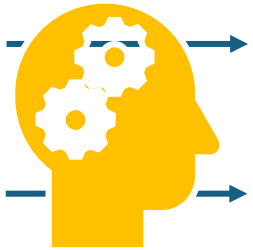
Dokument	Rok	Zakres	Wpływ na Polskę	Link
NIST AI Risk Management Framework (NIST AI 600-1)	2023	Ramy zarządzania ryzykiem AI opracowane przez NIST, obejmujące funkcje: zarządzanie, mapowanie, pomiar i zarządzanie ryzykiem	Choć opracowane przez NIST w USA, mogą stanowić cenne źródło wiedzy dla polskich organizacji wdrażających systemy AI	https://www.nist.gov/itl/ai-risk-management-framework
ISO/IEC 42001:2023 norma określająca system zarządzania sztuczną inteligencją (AI Management System, AIMS)	2023	Międzynarodowa norma określająca wymagania i wytyczne dotyczące ustanawiania, wdrażania, utrzymywania i ciągłego doskonalenia systemu zarządzania AI	Polska, jako członek ISO, może wdrażać tę normę w celu zapewnienia odpowiedzialnego rozwoju i wdrażania AI	https://www.iso.org/standard/81230.html
Unijne rozporządzenie dotyczące sztucznej inteligencji (AI Act)	2024	Pierwsze kompleksowe prawo regulujące AI w UE, klasyfikacja systemów AI według ryzyka	Obowiązuje w Polsce od 1 sierpnia 2024 r., z pełnym zastosowaniem od 2 sierpnia 2026 r.	https://eur-lex.europa.eu/legal-content/PL/TXT/?uri=CELEX:32024R1689
Konwencja ramowa Rady Europy o sztucznej inteligencji, prawach człowieka, demokracji i rządach prawa	2024	Międzynarodowa konwencja regulująca rozwój i wykorzystanie AI zgodnie z prawami człowieka, zasadami demokracji i rządami prawa. Obejmuje m.in. odpowiedzialność, przejrzystość i mechanizmy nadzoru nad AI.	Polska jako członek Rady Europy będzie zobowiązana do implementacji zapisów konwencji, co wpłynie na przyszłe krajowe regulacje dotyczące AI.	https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence
Projekt ustawy o systemach sztucznej inteligencji w Polsce (UC71)	2024 (w przygotowaniu)	Krajowe ramy prawne dla rozwoju i wdrażania AI w Polsce, zgodne z AI Act	Ustawa określi definicje systemów AI, obowiązki dostawców i użytkowników oraz mechanizmy nadzoru i egzekwowania przepisów	https://legislacja.rcl.gov.pl/projekt/12390551/katalog/13087901
OECD AI Policy Observatory	-	Narzędzie monitorujące polityki AI na świecie, prezentujące trendy regulacyjne, wytyczne etyczne i inicjatywy krajowe.	Polska jako członek OECD korzysta z analiz i benchmarków dla rozwoju własnej polityki AI.	https://oecd.ai/
Resortowa strategia sztucznej inteligencji do 2039 roku	2024	Określa strategie i ramy funkcjonowania AI w RON do 2039 roku		https://www.gov.pl/web/obrona-narodowa/resortowa-strategia-sztucznej-inteligencji-do-roku-2039

Wspólna Wizja Bezpiecznego Rozwoju AI

Deklaracja z Bletchley

1 listopada 2023



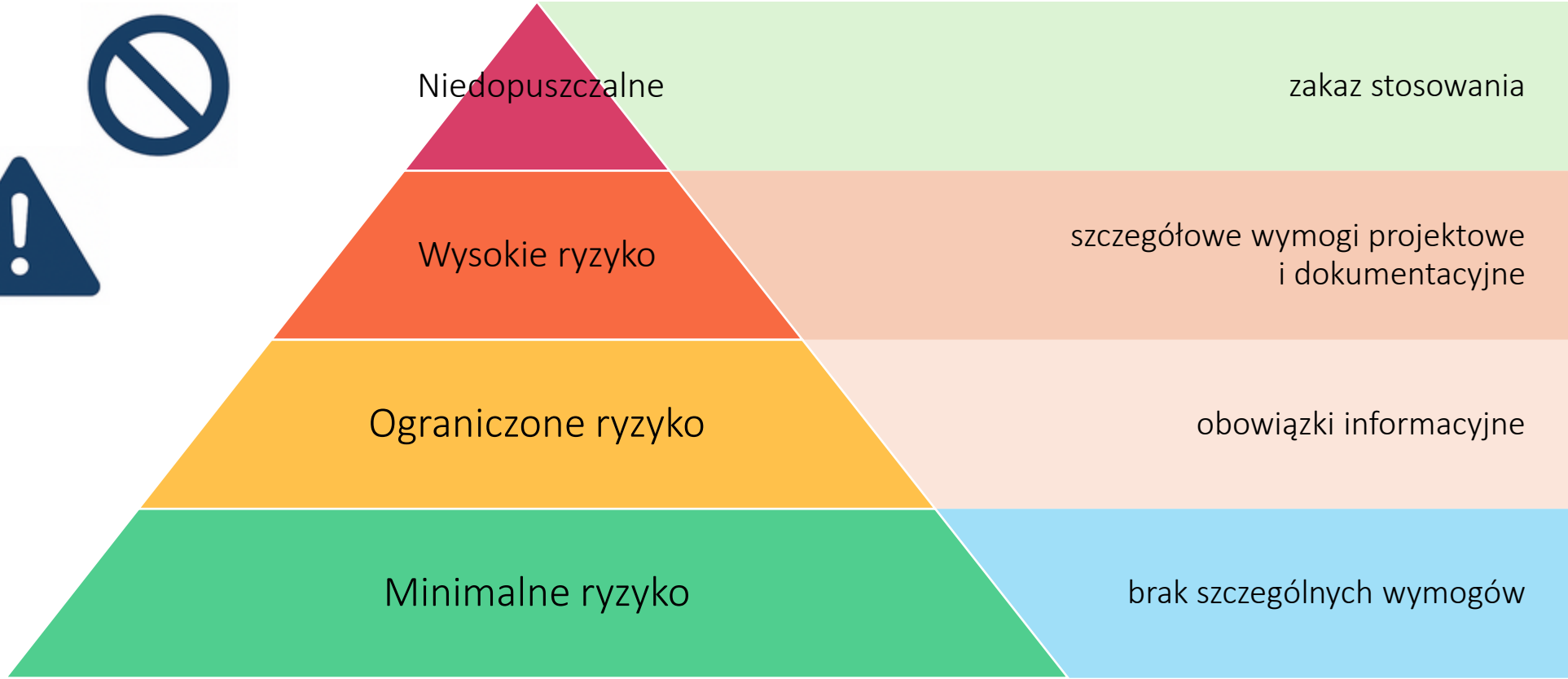


- Potencjał AI
- Ryzyka związane z AI
- Międzynarodowa współpraca
- Transparentność i odpowiedzialność sektora prywatnego
- Wzmocnienie zdolności technologicznych krajów rozwijających się
- Rozwój badań nad bezpieczeństwem AI
- Harmonizacja regulacji
- Inkluzja i etyka w AI

Poziomy Ryzyka Systemów AI według AI Act

AI Act klasyfikuje systemy na podstawie ryzyka, jakie stwarzają dla zdrowia, bezpieczeństwa i praw podstawowych obywateli.





Niedopuszczalne ryzyko (Unacceptable Risk AI Systems)



Systemy AI, których stosowanie jest sprzeczne z wartościami UE – naruszają prawa podstawowe lub narażają użytkowników na nieproporcjonalne ryzyko. Zgodnie z AI Act są one całkowicie zakazane.

Przykłady:

- Systemy social scoringu obywateli przez władze publiczne.
- Manipulacja poznawcza oparta na wrażliwych danych.
- Zdalna identyfikacja biometryczna w czasie rzeczywistym w przestrzeni publicznej (poza wyjątkami).

Implikacje techniczne:

- Wymóg wykrywania i blokowania zakazanych funkcji.
- Obowiązek audytu systemów pod kątem obecności ryzykownych komponentów.
- Eliminacja lub wycofanie architektur umożliwiających niedozwolone zastosowania.

Wysokie ryzyko (High-Risk AI Systems)



Systemy, które mogą znacząco wpłynąć na zdrowie, bezpieczeństwo lub prawa jednostki – dozwolone pod warunkiem spełnienia szeregu ścisłych wymogów technicznych i organizacyjnych.

Przykłady:

- Biometria – np. systemy rozpoznawania twarzy w dostępie do usług.
- Transport autonomiczny – np. pojazdy bezzałogowe.
- Systemy rekrutacyjne – wspierające decyzje kadrowe.

Implikacje techniczne:

- Architektura wspierająca kontrolowalność i przewidywalność działania.
- Obowiązkowe logowanie zdarzeń i dokumentacja techniczna.
- Wbudowane mechanizmy nadzoru człowieka i zarządzania ryzykiem.

Ograniczone ryzyko (Limited Risk AI Systems)



Systemy oddziałujące na użytkownika głównie poznawczo lub informacyjnie – dozwolone, ale wymagają zapewnienia transparentności i świadomego użycia przez odbiorcę.

Przykłady:

- Chatboty symulujące ludzką konwersację.
- Systemy rekomendacyjne – np. w e-commerce czy streamingu.
- Asystenci głosowi – np. w smartfonach i smart home.

Implikacje techniczne:

- Implementacja komunikatów o interakcji z AI („Jesteś obsługiwany przez system AI”).
- Zapewnienie opcji rezygnacji lub zmiany trybu działania.
- Przejrzysta architektura z uwzględnieniem funkcji wyjaśniania działania (explainability).

Minimalne ryzyko (Minimal Risk AI Systems)



Systemy, które nie wpływają istotnie na prawa, bezpieczeństwo ani dobrostan użytkownika – najmniej regulowane, choć wciąż zaleca się stosowanie dobrych praktyk.

Przykłady:

- Filtry antyspamowe w poczcie e-mail.
- Systemy predykcji zużycia energii.
- AI w grach komputerowych lub aplikacjach fitness.

Implikacje techniczne:

- Brak formalnych obowiązków, ale zalecenie stosowania zasad „cybersecurity by default”.
- Monitorowanie potencjalnego wzrostu ryzyka przy zmianie zastosowania systemu.
- Dokumentacja wewnętrzna wspierająca dobrowolną zgodność ze standardami.

Działania do podjęcia w ramach wdrożenia AI Act

Kary za naruszenie regulacji:
Do 35 mln euro lub 7% rocznego światowego obrotu firmy, w zależności od tego, która wartość jest wyższa.

- **2 sierpnia 2024** – Wejście w życie rozporządzenia.
- **2 lutego 2025** – Wdrożenie przepisów dotyczących podstawowych definicji, zakresu stosowania, klasyfikacji i oceny zgodności systemów AI wysokiego ryzyka.
- **2 maja 2025** – Zastosowanie kodeksów postępowania.
- **2 sierpnia 2025** – Obowiązywanie zasad zarządzania i obowiązków dla ogólnego zastosowania AI (GPAI).
- **2 sierpnia 2026** – Rozpoczęcie stosowania przepisów EU AI Act dla systemów AI .
- **2 sierpnia 2027** – Stosowanie całego EU AI Act dla wszystkich kategorii ryzyka .
- **Do 2 sierpnia 2031 r.** – Komisja Europejska przeprowadzi ocenę egzekwowania rozporządzenia i przedłoży sprawozdanie Parlamentowi Europejskiemu, Radzie oraz Europejskiemu Komitetowi Ekonomiczno-Społecznemu.

Podmioty objęte AI Act

Dostawcy systemów AI

- podmioty, które wprowadzają do obrotu lub oddają do użytku systemy AI w Unii, niezależnie od tego, czy mają siedzibę w UE, czy poza nią. (Art. 2(1)(a))

Podmioty stosujące systemy AI w UE

- organizacje, firmy i inne instytucje, które korzystają z systemów AI i mają siedzibę lub działalność w Unii. (Art. 2(1)(b))

Dostawcy i użytkownicy AI spoza UE, jeśli wyniki systemu AI są wykorzystywane w Unii

- oznacza to, że firmy spoza UE podlegają AI Act, jeśli ich systemy mają wpływ na podmioty w UE. (Art. 2(1)(c))

Importerzy i dystrybutorzy systemów AI

- podmioty zajmujące się sprowadzaniem i rozpowszechnianiem systemów AI na rynku unijnym. (Art. 2(1)(d))

Producenci produktów zintegrowanych z systemem AI

- jeśli producent wprowadza na rynek produkt, w którym AI jest zintegrowane, ponosi odpowiedzialność zgodnie z AI Act. (Art. 2(1)(e))

Upoważnieni przedstawiciele dostawców spoza UE

- firmy spoza UE muszą wyznaczyć przedstawiciela w UE, który odpowiada za zgodność z regulacjami. (Art. 2(1)(f))

Osoby, na które AI ma wpływ, znajdujące się w Unii

- AI Act obejmuje również ochronę osób, na które wpływa działanie systemów AI, jeśli znajdują się w UE. (Art. 2(1)(g))

Podmioty nie objęte AI Act

Podmioty wykorzystujące AI do celów wojskowych, obronnych i bezpieczeństwa narodowego

- systemy używane wyłącznie w tych celach nie podlegają regulacjom AI Act. (Art. 2(3))

Podmioty prowadzące badania naukowe i rozwój

- jeśli systemy AI i ich wyniki są wykorzystywane wyłącznie w ramach badań naukowych i rozwoju, nie podlegają AI Act. (Art. 2(6))

Podmioty spoza UE

- systemy AI, które nie są wprowadzone do obrotu ani oddane do użytku w UE oraz których wyniki nie są wykorzystywane w Unii, nie są objęte regulacją. (Art. 2(1)(a))

Osoby fizyczne korzystające z AI w celach osobistych

- użytkownicy indywidualni stosujący AI poza działalnością zawodową są zwolnieni z większości obowiązków wynikających z AI Act, z wyjątkiem niektórych wymogów dotyczących przejrzystości. (Art. 2(4))

Organy ścigania i wymiar sprawiedliwości spoza UE

- systemy AI stosowane przez organy publiczne państw trzecich lub organizacje międzynarodowe we współpracy z organami ścigania i wymiarem sprawiedliwości UE są wyłączone z regulacji, o ile spełniają odpowiednie zabezpieczenia. (Art. 2(8))

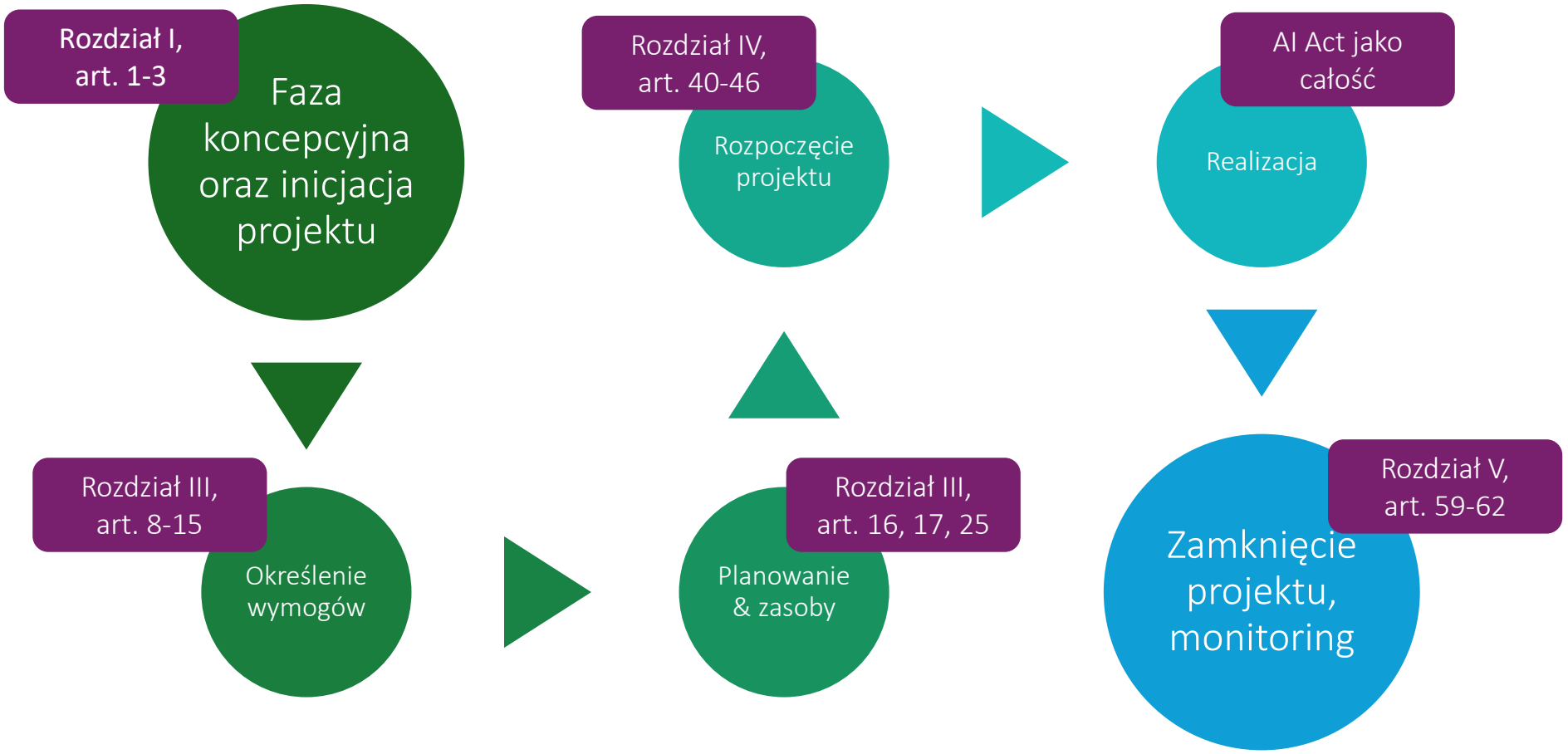
Twórcy i użytkownicy otwartego oprogramowania AI

- modele AI udostępniane na licencji open source są zwolnione z większości obowiązków, z wyjątkiem systemów o wysokim ryzyku oraz modeli AI ogólnego przeznaczenia o systemowym ryzyku. (Art. 2(10))

Role w AI Act

Rola	Opis	Przykłady obowiązków
Dostawca (Provider)	Podmiot (osoba fizyczna lub prawna), który rozwija system AI i wprowadza go na rynek lub oddaje do użytku.	Certyfikacja, dokumentacja, ocena ryzyka
Importer (Importer)	Podmiot wprowadzający system AI na rynek UE, jeśli jego dostawca znajduje się poza Unią Europejską.	Zapewnienie zgodności, rejestracja w UE
Dystrybutor (Distributor)	Podmiot, który sprzedaje, udostępnia lub wdraża system AI, ale nie jest jego twórcą ani importerem.	Weryfikacja certyfikatów, zgłaszanie naruszeń
Użytkownik (User)	Każda osoba fizyczna lub prawna korzystająca z systemu AI do swoich celów, ale nie rozwijająca go ani nie wprowadzająca na rynek.	Przestrzeganie zasad użytkowania, przejrzystość
Operator (Deploying Entity)	Podmiot odpowiedzialny za wdrożenie i eksploatację systemu AI w organizacji.	Monitorowanie działania, nadzór nad eksploatacją
Organ nadzorczy (Market Authority)	Instytucja odpowiedzialna za egzekwowanie przepisów AI Act i monitorowanie rynku AI w UE.	Kontrola rynku, nakładanie kar

Zależność budowy systemu AI i AI ACT



- **Rozdział III (art. 8-25)** jest kluczowy dla projektowania i planowania AI, ponieważ określa wymogi dla systemów wysokiego ryzyka.
- **Rozdział IV** dotyczy **rozpoczęcia projektu i nadzoru** nad zgodnością AI na etapie wdrażania.
- **Rozdział V** odpowiada za **monitorowanie i egzekwowanie przepisów**, co ma znaczenie po zamknięciu projektu.

Wdrażanie AI Act – obowiązki i konsekwencje

Podmioty dostarczające systemy AI mają obowiązek zgłoszenia, dokumentowania, testowania i nadzoru.

Audyty wewnętrzne i zewnętrzne – także przez organy nadzorujące.

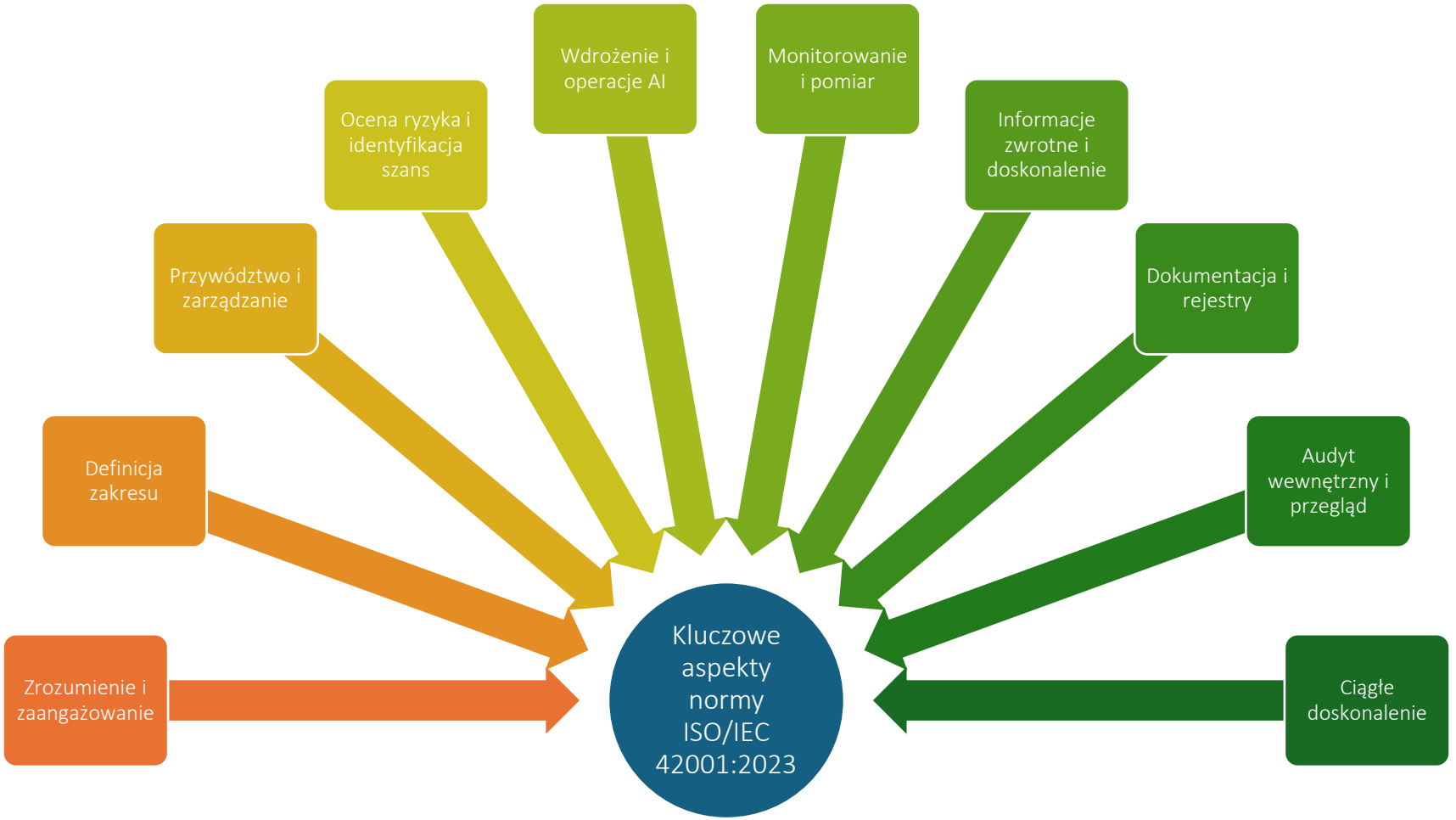
Sankcje: do 35 mln euro lub 7% rocznego światowego obrotu w przypadku poważnych naruszeń.

AI compliance staje się nową domeną zarządzania ryzykiem technologicznym.

Zalety i Wady polskiego projektu ustawy o systemach AI (PL)

✅ Zalety	❌ Wady
🏛️ **Harmonizacja z prawem unijnym:** Spójność z rozporządzeniem UE 2024/1689 ułatwia współpracę międzynarodową i integrację Polski z rynkiem cyfrowym UE.	💼 **Bariery dla małych firm i startupów:** Skomplikowane wymogi akredytacyjne mogą być trudne do spełnienia dla mniejszych podmiotów.
🛡️ **Kompleksowy nadzór i bezpieczeństwo:** Mechanizmy nadzoru nad AI zwiększają bezpieczeństwo technologii i zapobiegają incydentom.	💸 **Koszty administracyjne:** Tworzenie nowej struktury nadzorczej oraz wymnetania związane z oceną zgodności mogą generować dodatkowe koszty.
👥 **Ochrona konsumentów i obywateli:** Ustawa kładzie nacisk na ochronę praw użytkowników i minimalizację ryzyk związanych z AI.	🕒 **Opóźnienia we wdrożeniu innowacji:** Regulacje mogą spowalniać wprowadzenie nowych rozwiązań AI na rynek.
🔍 **Zwiększenie transparentności i odpowiedzialności:** Obowiązek raportowania incydentów i kary za naruszenia sprzyjają większej odpowiedzialności firm.	📏 **Ryzyko nadmiernej regulacji:** Zbyt restrykcyjne przepisy mogą zniechęcać do innowacji i ograniczać inwestycje w sektorze AI.
🚀 **Tworzenie piaskownic regulacyjnych:** Możliwość testowania nowych technologii AI w kontrolowanym środowisku sprzyja innowacjom.	❓ **Brak precyzyjnych definicji:** Niektóre kluczowe pojęcia związane z AI mogą być zbyt ogólne, co prowadzi do niejasności w interpretacji przepisów.

ISO/IEC 42001:2023
Zarządzanie Sztuczną Inteligencją w Organizacji



RODO a AI

RODO (Rozporządzenie o Ochronie Danych Osobowych)

to rozporządzenie o ochronie danych osobowych, obowiązujące w UE od 2018 r.
Reguluje sposób przetwarzania danych osobowych przez organizacje, w tym systemy AI.

Jakie wyzwania RODO stawia przed AI?

- **Brak przejrzystości modeli („czarna skrzynka”)** – AI podejmuje decyzje, których nie da się łatwo wyjaśnić. RODO wymaga wyjaśnialności.
- **Prawo do bycia zapomnianym** – użytkownicy mogą żądać usunięcia swoich danych, ale w modelach AI (np. deep learning) nie zawsze jest to możliwe.
- **Profilowanie i dyskryminacja** – AI nie może podejmować automatycznych decyzji (np. odrzucenie kredytu) bez wyjaśnienia i możliwości odwołania.
- **Przenoszenie danych** – AI zbierające dane użytkowników musi umożliwiać ich łatwy eksport i przekazanie do innego systemu.

RODO a AI

6 Kluczowych zasad RODO w kontekście AI

Zasada	Znaczenie dla AI
Legalność, rzetelność, przejrzystość	AI musi przetwarzać dane zgodnie z prawem i jasno informować użytkowników, jak dane są wykorzystywane.
Ograniczenie celu	Dane mogą być zbierane tylko w określonym celu, np. model AI do rozpoznawania twarzy nie może być używany do monitorowania bez zgody użytkownika.
Minimalizacja danych	AI nie powinno przetwarzać więcej danych, niż jest to konieczne – np. system scoringowy kredytowy nie musi analizować kont na mediach społecznościowych.
Dokładność	Modele AI muszą bazować na aktualnych i precyzyjnych danych – jeśli AI używa błędnych informacji, może to prowadzić do dyskryminacji.
Ograniczenie przechowywania	Dane osobowe nie mogą być przechowywane w nieskończoność – firmy muszą określić czas retencji danych.
Bezpieczeństwo danych	AI musi zapewniać ochronę danych (szyfrowanie, anonimizacja), aby uniknąć wycieków i ataków hakerskich.

AI w MON

Strategia Sztucznej Inteligencji w Resorcie Obrony Narodowej do 2039 r.

przedstawia plan wdrożenia i wykorzystania systemów AI w Resorcie Obrony Narodowej, uwzględniając zarówno szanse, jak i zagrożenia wynikające z tej technologii. Opisuje cele strategiczne, działania operacyjne oraz wyzwania związane z implementacją AI w kontekście obronności i bezpieczeństwa państwa.

Szanse i zagrożenia AI w wojsku

✅ Szanse:

- Szybka analiza danych i podejmowanie decyzji.
- Autonomiczne systemy operacyjne.
- Cyberbezpieczeństwo i rozpoznanie.

❌ Zagrożenia:

- Możliwość ataków AI na infrastrukturę.
- Dezinformacja i manipulacja informacyjna.
- Ryzyko broni autonomicznej.

AI w MON

Cele do osiągnięcia

1 Przygotowanie do implementacji sztucznej inteligencji

- Zapewnienie odpowiedniej infrastruktury i kompetencji.
- Stworzenie ram prawnych i etycznych dla AI.
- Wsparcie dla kadr w zakresie edukacji i adaptacji do AI.

2 Skalowalne wdrożenie sztucznej inteligencji

- Stopniowe wprowadzanie AI do działań operacyjnych.
- Integracja AI z systemami obronnymi i decyzyjnymi.
- Promowanie współpracy krajowej i międzynarodowej w zakresie AI.

3 Ochrona strategicznych zasobów i rozwiązań

- Zapewnienie cyberbezpieczeństwa systemów AI.
- Ochrona własności intelektualnej i danych kluczowych dla bezpieczeństwa państwa.
- Monitorowanie i kontrola wpływu AI na operacje militarne.

AI a prawa autorskie

Prawa autorskie a AI – kluczowe kwestie

✔ **Kto jest autorem?** – AI czy człowiek?
✔ **Prawa do danych treningowych** – Czy można trenować modele AI na chronionych treściach?
✔ **AI generujące treści** – Czy wygenerowany przez AI kod, tekst, obraz może być objęty prawem autorskim?
✔ **Licencje na modele AI** – Open-source vs. własność prywatna

Prawo autorskie – podstawowe zasady

📌 **Autorstwo** – Prawo autorskie przysługuje twórcy (człowiekowi), nie AI
📌 **Ochrona twórczości** – Dotyczy oryginalnych utworów (tekstów, grafik, kodów)
📌 **Dozwolony użytek** – Prawo do korzystania z cudzych utworów w określonych przypadkach (np. edukacja, cytowanie)
📌 **Plagiat i naruszenia** – Wykorzystanie treści generowanych przez AI może naruszać prawa osób trzecich

AI a prawa autorskie

AI a ochrona własności intelektualnej

- 🔷 **Problem „treningu” na cudzych danych** – Modele AI często uczą się na treściach chronionych prawem autorskim
- 🔷 **Odpowiedzialność prawna** – Kto odpowiada za naruszenia? Twórca modelu, użytkownik czy firma?
- 🔷 **Rozwój regulacji** – Różne podejścia w UE, USA, Chinach do regulacji AI i ochrony własności intelektualnej

Przykłady i kontrowersje

- 📌 **Kod generowany przez AI** – GitHub Copilot i problem praw autorskich
- 📌 **Sztuka AI (obrazy, muzyka)** – MidJourney, DALL·E – czy obrazy generowane przez AI są chronione?
- 📌 **Deepfake i prawo** – Manipulacje treściami a naruszenie praw do wizerunku

AI a prawa autorskie

Czym jest utwór według polskiego prawa?

Art. 1 ust. 1 ustawy:
„Przedmiotem prawa autorskiego jest każdy przejaw działalności twórczej o indywidualnym charakterze, ustalony w jakiejkolwiek postaci, niezależnie od wartości, przeznaczenia i sposobu wyrażenia."

Kluczowe elementy:

- **Działalność twórcza** – musi pochodzić od człowieka.
- **Indywidualny charakter** – unikalność i oryginalność.

AI jako twórca – czy to możliwe?

Art. 8 ust. 1 ustawy:
„Prawo autorskie przysługuje twórcy, o ile ustawa nie stanowi inaczej."

Kluczowe elementy:

- Polskie prawo **nie uznaje AI za podmiot prawa autorskiego.**
- Twórcą może być tylko **osoba fizyczna.**
- Utwory stworzone autonomicznie przez AI generalnie nie są chronione prawem autorskim.

Ustawa z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych.
https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=wdu19940240083

AI a prawa autorskie

Kiedy utwór z udziałem AI podlega ochronie?

- Jeśli człowiek wnosi istotny wkład twórczy (np. edycja, adaptacja), utwór może być chroniony.
- Decydujące znaczenie ma stopień zaangażowania człowieka w proces twórczy

Obowiązek oznaczania treści AI

- Obecnie brak ogólnego obowiązku prawnego w Polsce.
- Zalecane jest oznaczanie treści AI w celach transparentności.
- Przyszłe regulacje (np. AI Act) mogą wprowadzić takie obowiązki .

Data Governance

Data Governance

to zbiór zasad, procesów i standardów, które regulują sposób zarządzania danymi w organizacji.

Kluczowe elementy Data Governance

🛡 **Bezpieczeństwo danych** – Ochrona przed cyberzagrożeniami.
📜 **Zgodność z przepisami** – RODO, ISO 27001.
📊 **Jakość danych** – Zapewnienie dokładności i kompletności informacji.
🔍 **Dostępność danych** – Ułatwienie dostępu dla uprawnionych użytkowników.
⚙ **Zarządzanie metadanymi** – Organizacja i klasyfikacja danych.

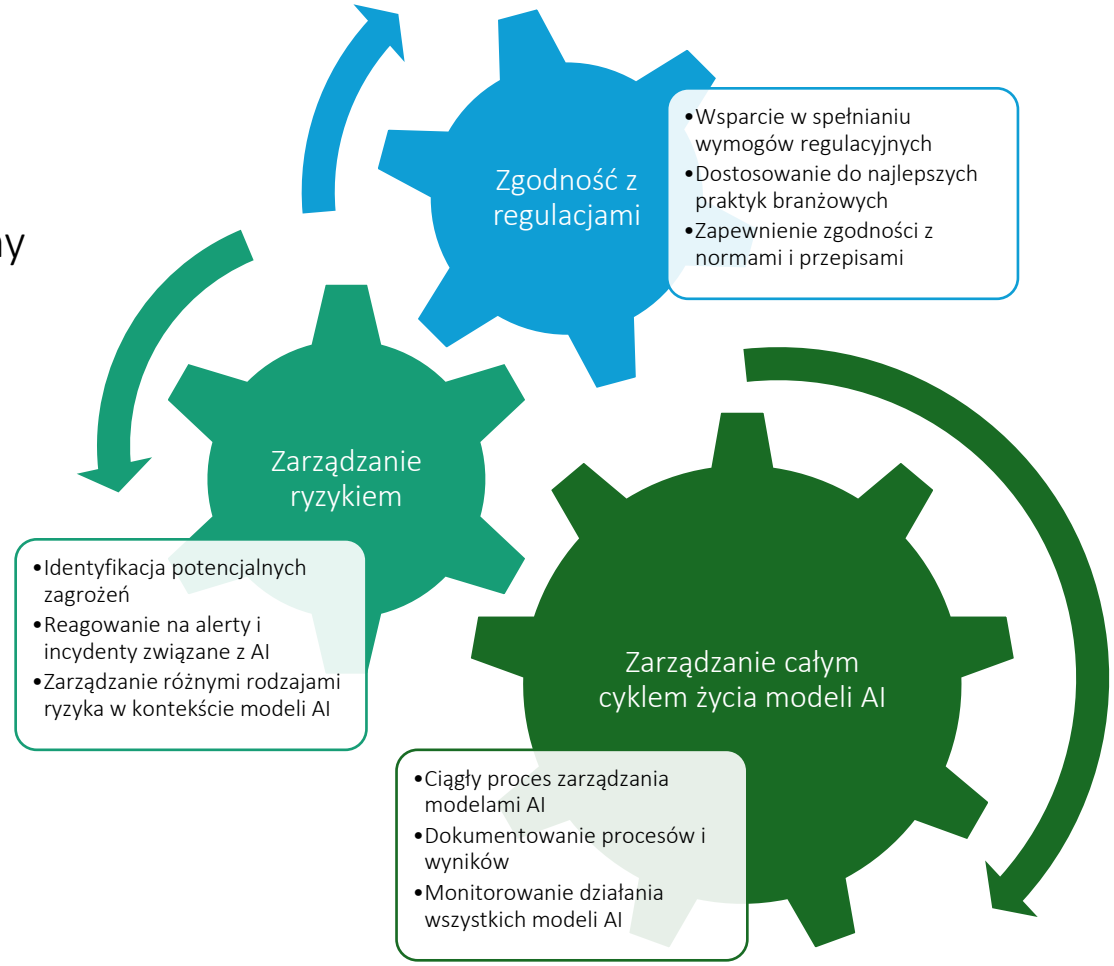
Data Governance - Zarządzanie Danymi w Organizacji



Data Governance

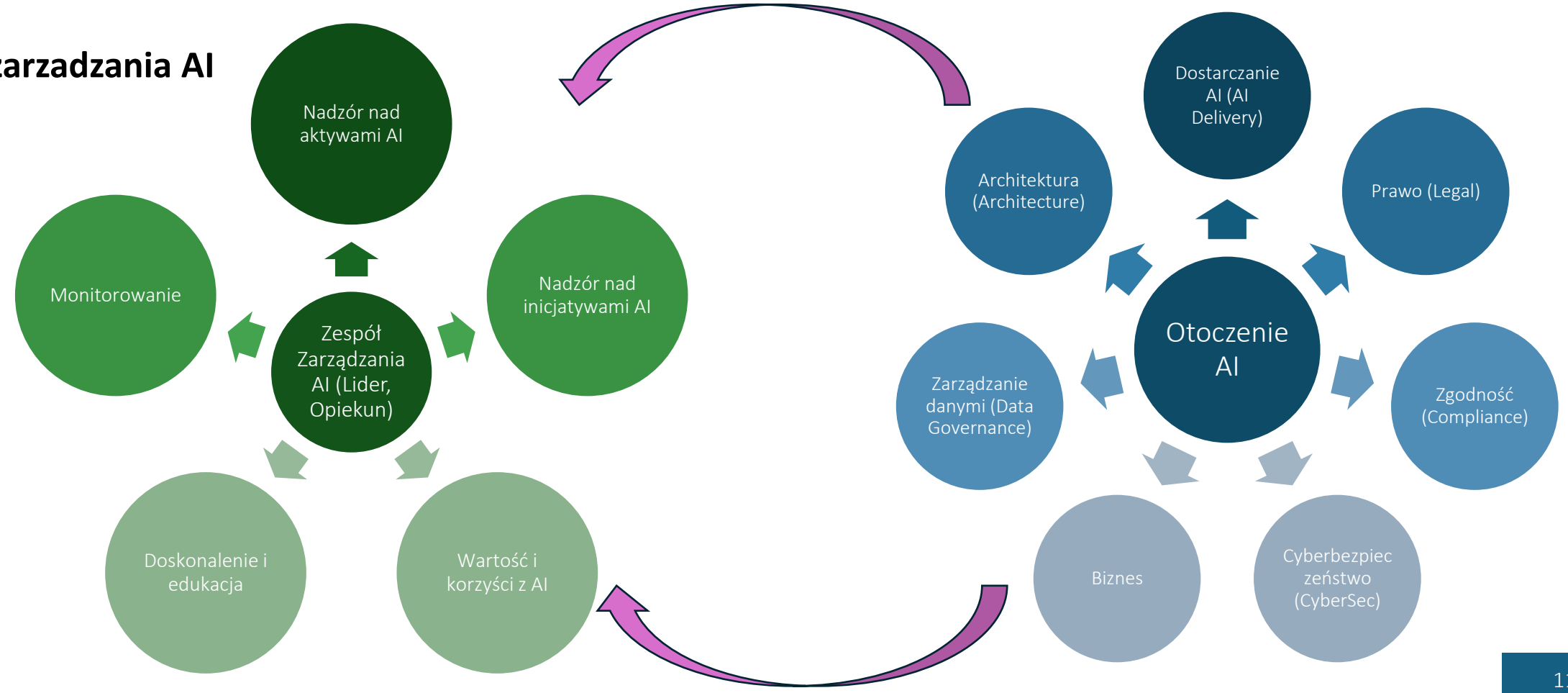
AI governance (zarządzanie AI)

to zestaw procesów organizacyjnych, ról i produktów, które pomagają zapewnić, że systemy AI są godne zaufania i korzystne dla organizacji.



Data Governance

Model zarzadzania AI



Responsible AI

Responsible AI

to podejście do tworzenia sztucznej inteligencji, które zapewnia jej etyczność, przejrzystość i zgodność z wartościami społecznymi.

Przykłady Responsible AI w praktyce
✓ AI wykorzystywana w medycynie – analiza diagnoz bez stronniczości.
✓ AI w rekrutacji – eliminacja uprzedzeń płciowych i rasowych.
✓ AI w finansach – zapobieganie dyskryminacji kredytowej.

Kluczowe zasady Responsible AI

🛡 **Etyczność** – AI powinna działać zgodnie z wartościami moralnymi.
⚖ **Sprawiedliwość** – Unikanie uprzedzeń i dyskryminacji w algorytmach.
🔍 **Przejrzystość** – AI powinna wyjaśniać swoje decyzje.

Bezpieczeństwo – Ochrona przed nadużyciami i cyberzagrożeniami.

Odpowiedzialność – Twórcy AI powinni ponosić odpowiedzialność za jej skutki.

Responsible AI

Zaufanie do AI

Co trzeba zrobić, aby zaufać decyzji podjętej przez AI?

WYJAŚNIALNOŚĆ

- Czy AI jest łatwy do zrozumienia?
- Dlaczego AI dostarcza takie wyniki?
- W jakich warunkach wyniki byłyby inne?

STRONNICZOŚĆ

- Czy AI jest sprawiedliwy?
- Czy AI faworyzuje określone grupy?
- Czy uwzględnia uprzedzenia w danych?

ODPORNOŚĆ

- Czy ktoś mógł wpłynąć na wynik?
- Czy system AI jest odporny na manipulacje?
- Jakie zabezpieczenia chronią przed zagrożeniami?

TRANSPARENTNOŚĆ

- Czy AI jest odpowiedzialny?
- Czy możemy zweryfikować rzetelność wyników AI?
- Jakie mechanizmy kontrolne są wdrożone?

PRYWATNOŚĆ

- Czy AI zabezpiecza dane?
- Jak zapewnić użytkownikom kontrolę nad swoimi danymi?
- Czy AI działa zgodnie z regulacjami dotyczącymi prywatności?

Alignment AI

Alignment AI

to dziedzina badań zajmująca się dostosowaniem działania sztucznej inteligencji do ludzkich wartości, norm etycznych i intencji użytkownika. Celem jest zapobieganie sytuacjom, w których AI podejmuje decyzje sprzeczne z oczekiwaniami społecznymi.

Główne aspekty Alignment AI

🤖 **Celowość działania** – AI powinna realizować ludzkie cele, unikając dosłownej interpretacji poleceń.

🛡️ **Bezpieczeństwo** – AI musi działać stabilnie i nie podejmować ryzykownych decyzji.

🔮 **Przewidywalność** – Ludzie powinni móc zrozumieć i przewidzieć zachowanie AI.

💡 **Zrozumiałość decyzji** – AI powinna wyjaśniać swoje działania, by budować zaufanie użytkowników.

Agenci AI – nowe wydanie

Modelowanie interakcji społecznych

- Eksperyment "Generative Agents" (2023) prowadzony przez Stanford University i Google Research.
- Cel: stworzenie agentów AI zdolnych do realistycznych interakcji społecznych.
- Generatywni agenci oparte na GPT-3.5: posiadają pamięć planują działania, prowadzą refleksję i komunikują się w języku naturalnym.
- Środowisko symulacyjne: Smallville – 25 agentów w wirtualnym miasteczku z domami, kawiarnią i miejscami pracy.

https://arxiv.org/abs/2304.03442
https://github.com/joonspk-research/generative_agents



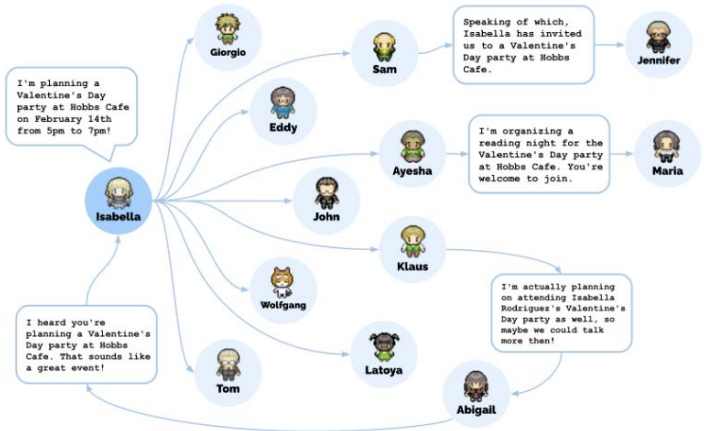
Agenci AI – nowe wydanie

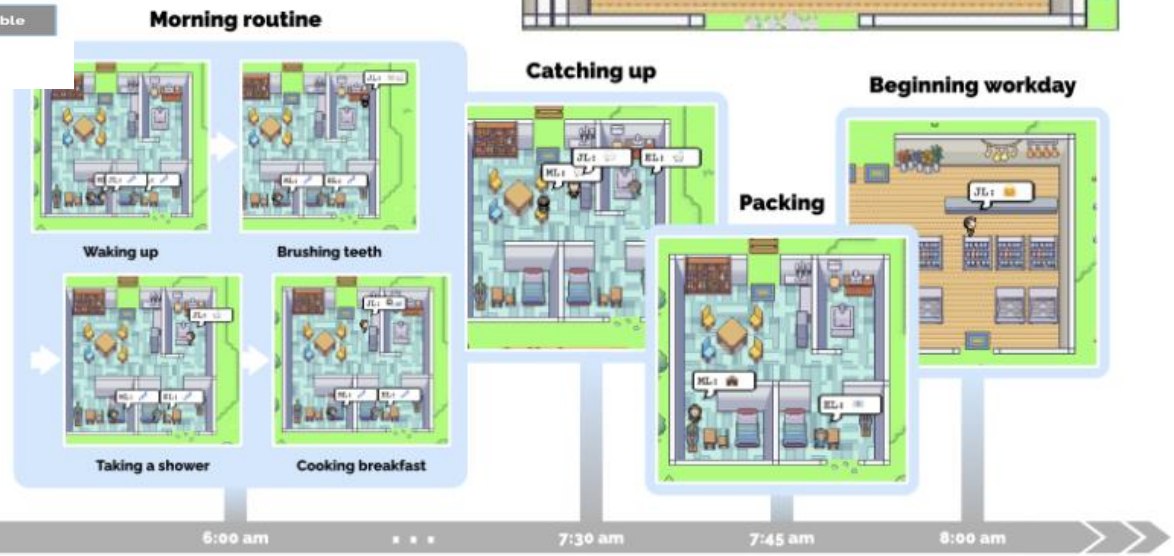
The Smallville sandbox world





Organizacja walentynek





Agenci AI – nowe wydanie

Wnioski

✅ Dobre wnioski	❌ Złe wnioski / ograniczenia
Agenci generatywni potrafią realistycznie symulować codzienne ludzkie zachowania.	Brakuje długoterminowej spójności w zachowaniach agentów – mogą "zapominać" wcześniejsze decyzje lub konteksty.
W środowisku multi-agentowym pojawiają się spontaniczne interakcje społeczne (np. organizacja wydarzeń, inicjatywy grupowe).	Modele są kosztowne obliczeniowo – każda decyzja agenta wymaga zapytań do dużego modelu językowego.
Architektura oparta na obserwacji, planowaniu i refleksji poprawia jakość symulacji.	Wysoka zależność od jakości i aktualności modelu językowego – błędy modelu przekładają się bezpośrednio na zachowania agentów.
Możliwość tworzenia środowisk testowych do badania teorii społecznych lub UX.	Brak rzeczywistej świadomości lub zrozumienia – działania agentów są tylko statystycznie zgodne z ludzkim zachowaniem.
Agenci są w stanie prowadzić wieloetapowe interakcje interpersonalne (np. przyjaźnie, romanse, współpraca).	Ograniczona możliwość kontroli nad emergentnymi zjawiskami – symulacja może przyjąć nieprzewidywalny kierunek.

Link do symulacji https://reverie.herokuapp.com/UIST_Demo/



Każdy może mieć agenta na miarę swoich możliwości



Korzyści dla organizacji

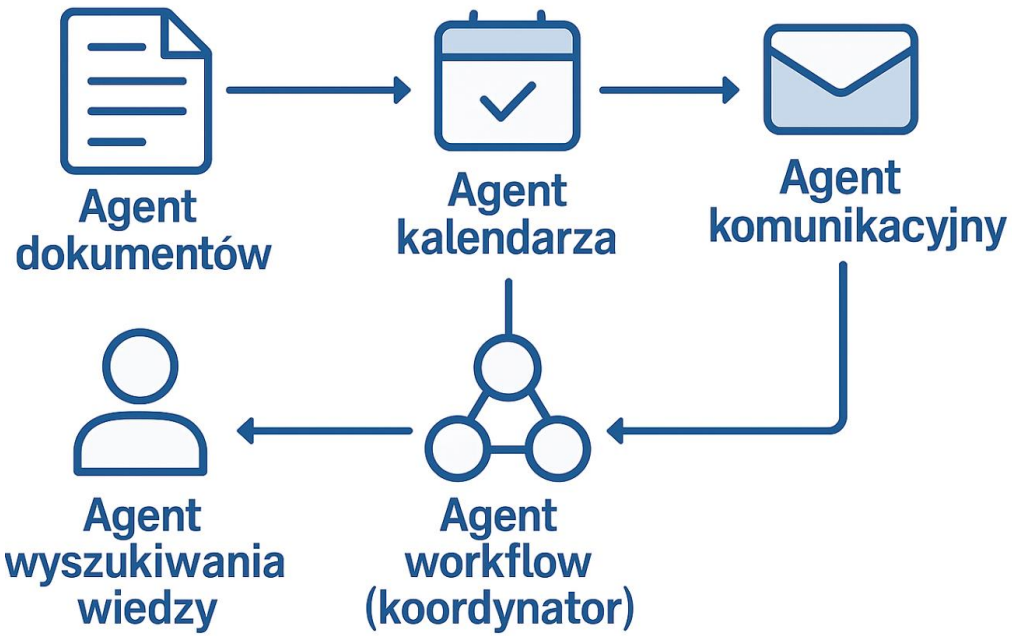
BOT pracuje **7** razy szybciej niż **CZŁOWIEK**

BOT	KOSZT 1 godz. pracy	CZŁOWIEK
n*		4,5n

- Kilkukrotne przyspieszenie pracy oraz wzrost elastyczności w zarządzaniu procesami
- Wyższa jakość usług
- Obniżenie kosztów operacyjnych
- Wzrost satysfakcji pracowników i klientów

Agenci AI – nowe wydanie

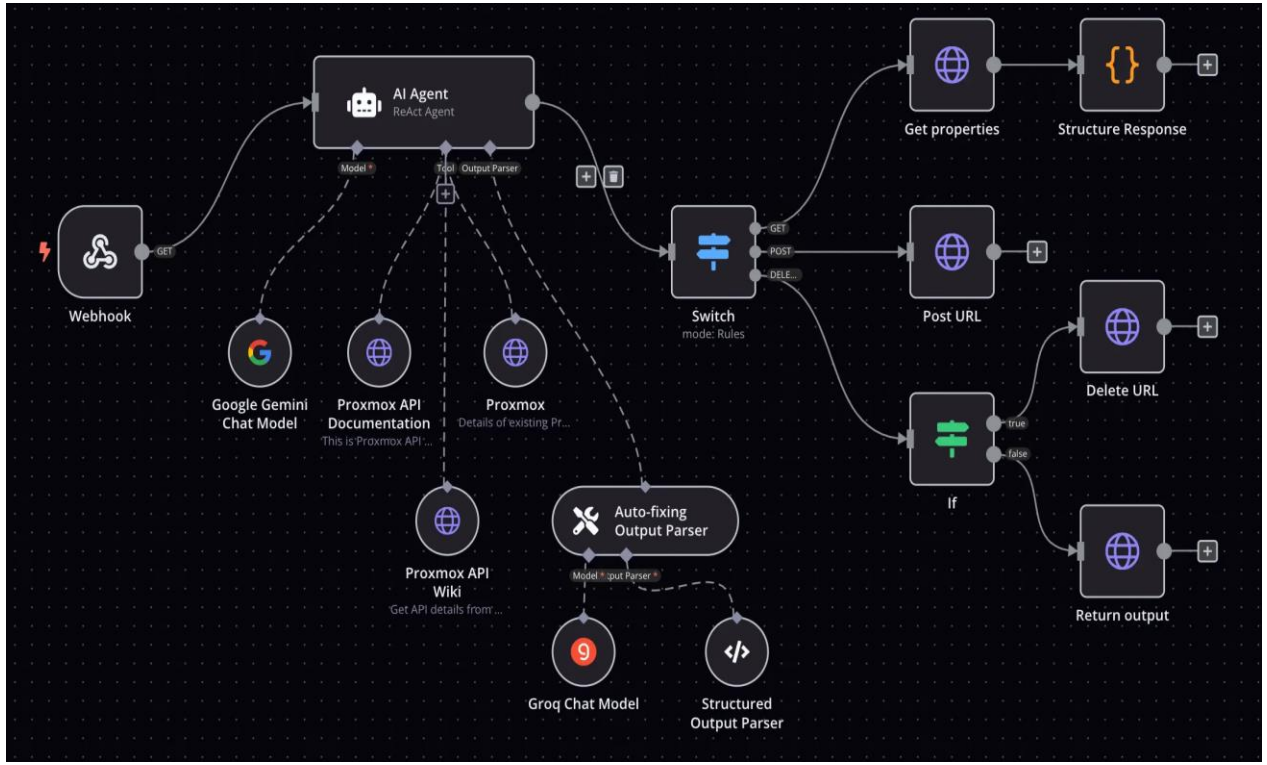
Typ agenta AI	Przykładowa rola
📄 **Agent dokumentów**	Klasyfikuje dokumenty, wyciąga dane, rozpoznaje wzorce faktur, umów.
📧 **Agent komunikacyjny**	Odpowiada na maile, podsumowuje wątki, przypomina o zadaniach.
📅 **Agent kalendarza**	Organizuje spotkania, sprawdza dostępność uczestników, proponuje terminy.
🔍 **Agent wyszukiwania wiedzy**	Przeszukuje bazy danych, intranet, repozytoria, odpowiada na pytania.
✅ **Agent workflow (koordynator)**	Nadzoruje realizację zadań między pracownikami i agentami.



Agenci AI – nowe wydanie

Narzędzia no-code / low-code do orkiestracji AI

Nazwa	Opis	Link
Make.com	Automatyzacja z LLM i setkami aplikacji SaaS	https://www.make.com/en
n8n.io	Open source, workflow automation z integracją OpenAI, Hugging Face, API, baz danych	https://n8n.io
FlowiseAI	Interfejs do tworzenia agentów i chatbotów z wykorzystaniem LangChain, LLM i narzędzi (open source)	https://flowiseai.com/



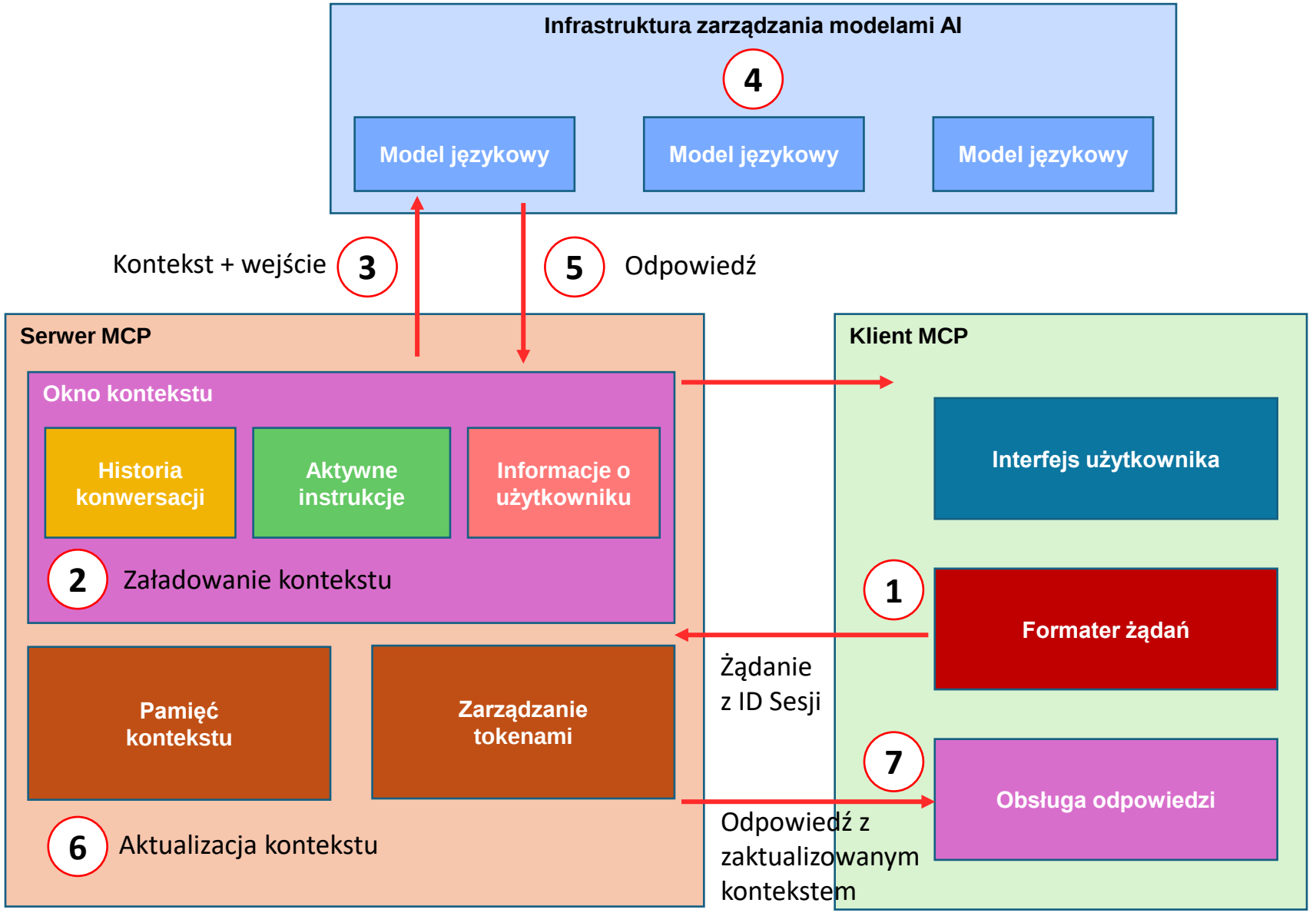
Agenci AI – nowe wydanie

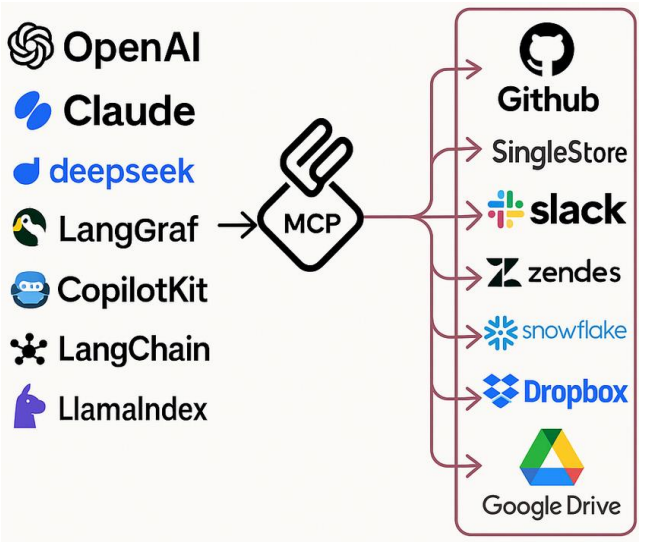
Model Context Protocol

to **architektoniczny wzorzec komunikacji** stosowany w systemach wieloagentowych, systemach opartych na wiedzy (Knowledge-Based Systems) oraz w kontekście **model-driven engineering (MDE)**, służący do **zarządzania kontekstem działania modelu lub komponentu w systemie złożonym**.

Element	Opis	Przykład
Model	Reprezentacja wiedzy, logiki, reguł lub zachowania (np. model decyzyjny, predykcyjny, behawioralny).	*AI przewiduje ryzyko choroby na podstawie danych pacjenta*
Context	Informacje z otoczenia, które wpływają na działanie modelu (np. czas, lokalizacja, stan użytkownika, urządzenia).	*Wiek, płeć, pora dnia, lokalizacja (np. dom vs szpital)*
Protocol	Zbiór reguł lub interfejsów komunikacyjnych, które określają, jak modele dostosowują się lub komunikują w zależności od kontekstu.	*Określa, kiedy i jak dane z czujników mają być interpretowane przez model (np. różne zakresy norm zależnie od sytuacji)*

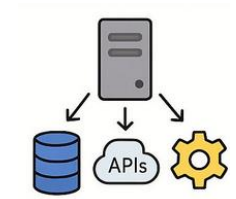
Model Context Protocol

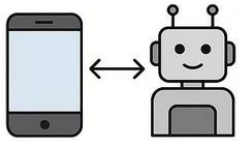


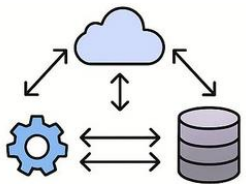


Agenci AI – nowe wydanie

Inne protokoły







Protokół	Nazwa (PL)	Opis	Charakterystyka techniczna
MCP (Model Context Protocol)	Protokół Kontekstu Modelu	Dostarcza zewnętrzny kontekst (np. dane, narzędzia) do dużych modeli językowych (LLM).	- Przydatny do pobierania danych z API, baz danych, narzędzi w czasie rzeczywistym. - Scentralizowany: serwer zarządza wszystkimi informacjami o narzędziach.
ACP (Agent Communication Protocol)	Protokół Komunikacji Agentów	Umożliwia komunikację między lokalnymi, osadzonymi agentami (np. w telefonach, robotach).	- Nie wymaga zewnętrznego serwera – agenci komunikują się bezpośrednio. - Wspiera różne protokoły: gRPC, IPC, ZeroMQ. - Działa w środowiskach offline.
A2A (Agent-to-Agent Protocol)	Protokół Agent-do-Agenta	Przeznaczony do rozproszonej, skalowalnej współpracy między agentami.	- Agenci mogą działać na różnych urządzeniach, w chmurach, organizacjach. - Bezpieczna komunikacja międzyplatformowa. - Wspiera horyzontalne współdziałanie agentów (LLM, bazy danych, planery itd.).

Pętla OODA

Pętla OODA (Observe, Orient, Decide, Act)

to koncepcja decyzyjna opracowana przez Johna Boyda, pierwotnie w kontekście strategii wojskowej.
Opisuje cztery etapy podejmowania decyzji w dynamicznym i szybko zmieniającym się środowisku.

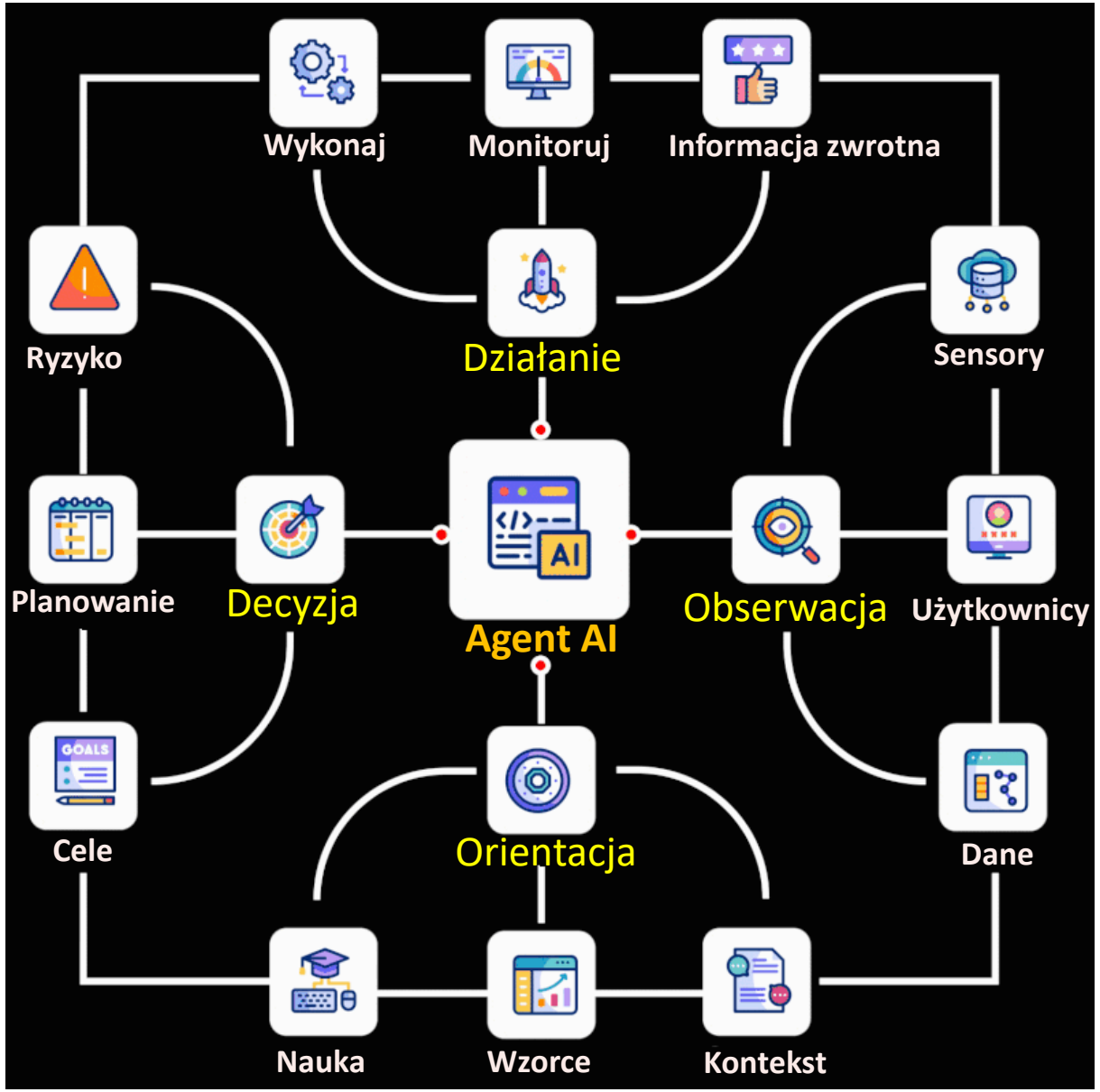
Dlaczego OODA pasuje do projektów IT?

✅ Szybkie reagowanie na zmiany w wymaganiach
✅ Minimalizacja ryzyka poprzez iteracyjne podejmowanie decyzji
✅ Wsparcie dla metodyk zwinnych (Scrum, Kanban)
✅ Efektywne zarządzanie kryzysowe i eskalacja problemów



Jak działają agenci AI (Pętla OODA)

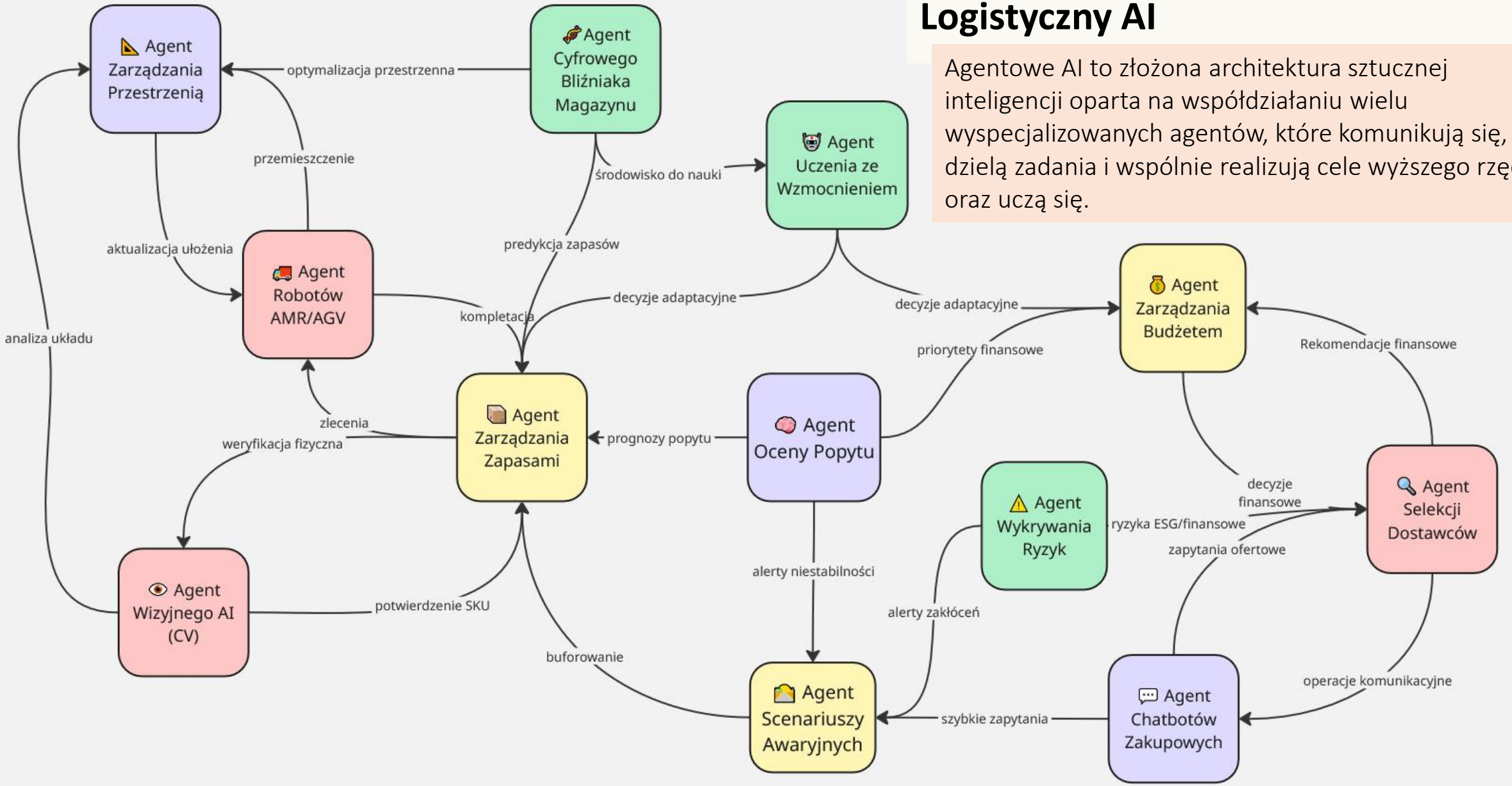
Obserwacja	
Sensory	Zbierają dane z otoczenia fizycznego (np. czujniki IoT, kamery, mikrofony).
Użytkownicy	Dane wejściowe generowane przez użytkowników (interakcje, kliknięcia, zapytania).
Dane	Strumienie danych zewnętrznych (np. API, bazy danych, źródła w czasie rzeczywistym).
Orientacja	
Nauka	Mechanizmy uczenia maszynowego analizujące dane historyczne i bieżące.
Wzorce	Identyfikacja powtarzających się struktur i schematów w danych.
Kontekst	Interpretacja danych z uwzględnieniem warunków otoczenia i bieżącej sytuacji.
Decyzja	
Cele	Definiują, co system ma osiągnąć; służą jako kryterium oceny decyzji.
Planowanie	Tworzenie sekwencji działań prowadzących do realizacji celów.
Ryzyko	Analiza zagrożeń, niepewności i kosztów potencjalnych decyzji.
Działanie	
Wykonaj	Realizacja zaplanowanego działania w środowisku fizycznym lub cyfrowym.
Monitoruj	Śledzenie skutków działania w czasie rzeczywistym.
Informacja zwrotna	Dane służące do oceny skuteczności działania i dalszego uczenia się.



Agentowe AI =N x Agent AI

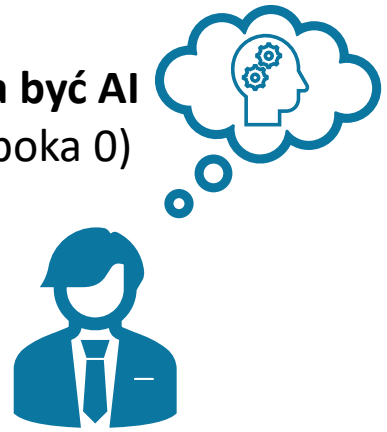
Agentowy System Zakupowo-Logistyczny AI

Agentowe AI to złożona architektura sztucznej inteligencji oparta na współdziałaniu wielu wyspecjalizowanych agentów, które komunikują się, dzielą zadania i wspólnie realizują cele wyższego rzędu oraz uczą się.



Model AI vs Agent AI vs Agentowe AI

Ma być AI
(Epoka 0)



Prosty model AI
(Epoka 1)

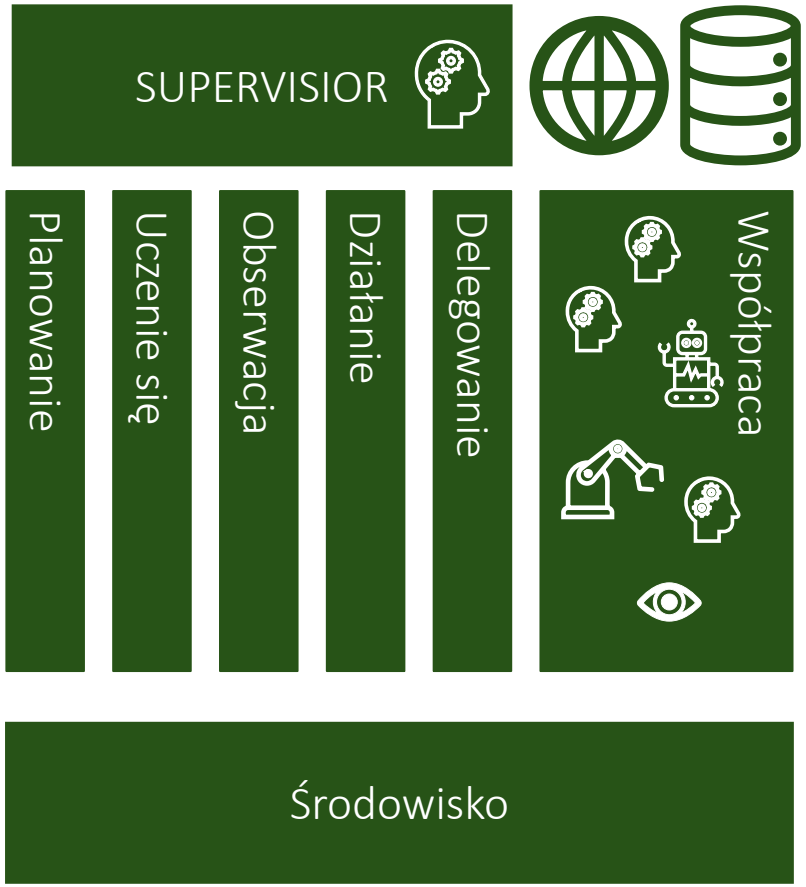




Agent AI
(Epoka 2)

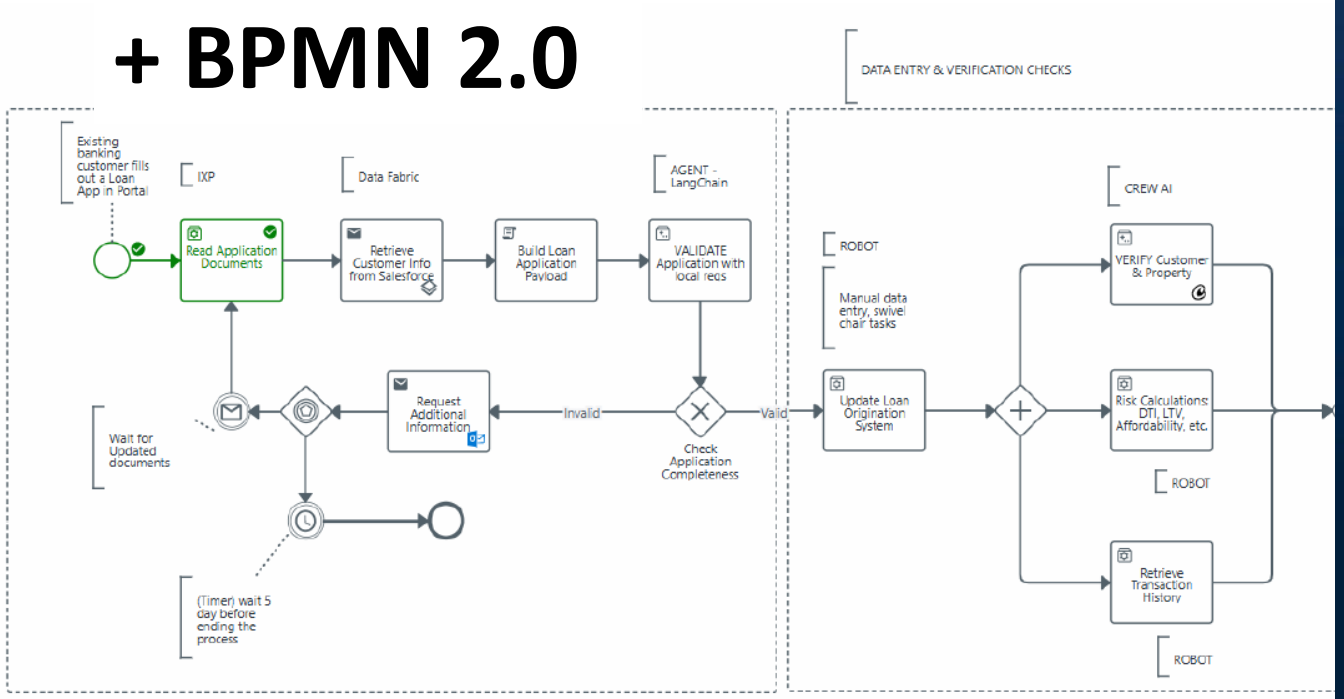


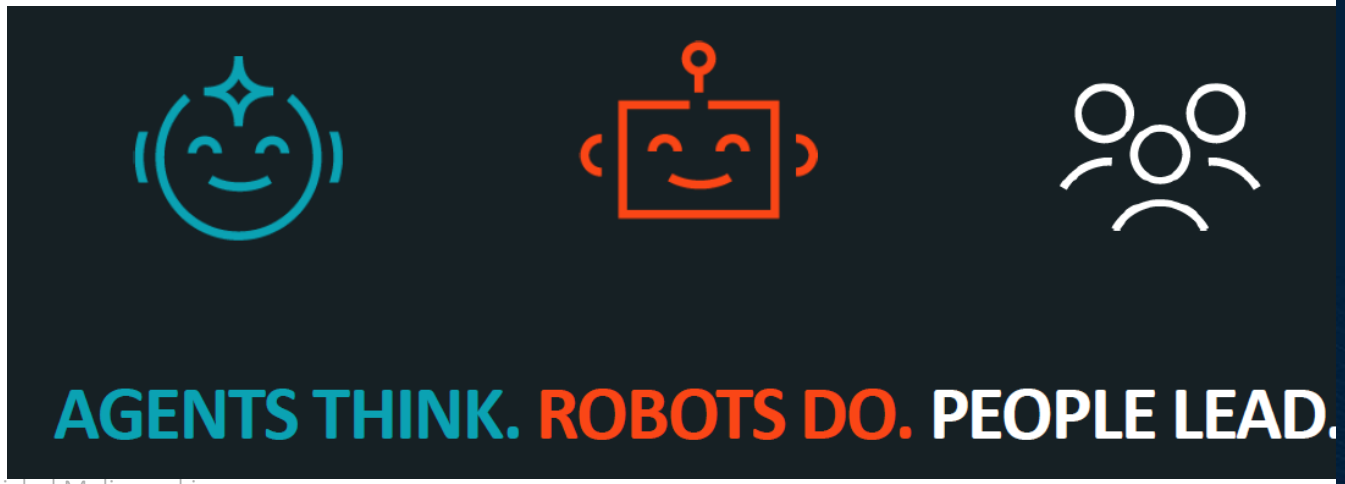
Agentowe AI
(Epoka 3)



Agentowe AI

+ BPMN 2.0







Model AI vs Agent AI vs Agentowe AI

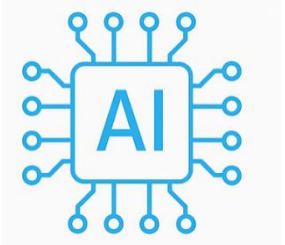
Kryterium	Prosty model AI (Epoka 1)	Agent AI (Epoka 2)	Agentowe AI (Epoka 3)
Opis	Modele statyczne uczące się na danych historycznych	Modułowe systemy podejmujące działania autonomiczne	Ekosystem inteligentnych agentów współpracujących w środowisku złożonym
Forma działania	Statyczna predykcja / klasyfikacja	Autonomiczne zadania o jasno określonym celu	Współpraca wielu agentów w dynamicznym środowisku
Przykłady technologii	Regresja liniowa, drzewo decyzyjne, SVM	LangChain Agents, ReAct, AutoGPT	CrewAI, OpenDevin, AgentVerse, ChatDev
Zakres autonomii	Brak lub minimalna autonomia	Ograniczona autonomia w obrębie zadania	Wysoka autonomia z rozpoznaniem celów, planowaniem i współpracą
Źródła danych	Zestawy treningowe	Dostęp do API, baz danych, plików	Wielomodalne źródła: sensory, API, przeglądarki, narzędzia operacyjne
Reaktywność	Brak reakcji na zmieniające się środowisko	Reaktywność względem promptów lub triggerów	Reaktywność i proaktywność w odpowiedzi na zmienne konteksty i cele
Koordynacja	Brak	Samodzielna realizacja zadań	Współdziałanie wielu agentów z podziałem ról (planer, wykonawca, analityk)
Zastosowania	Detekcja fraudów, klasyfikacja obrazów, predykcja sprzedaży	Automatyzacja procesów, zarządzanie dokumentami	Zarządzanie złożonymi procesami, projektowanie systemów wieloagentowych
Tryb działania	Offline	Online, trigger-based	Online, ciągła pętla percepcji, planowania i działania
Możliwość planowania	Brak	Tak, w ograniczonym zakresie	Tak, z dynamiczną adaptacją do zmiennych celów
Możliwość współpracy między jednostkami	Brak	Ograniczona	Wysoka – wieloagentowa współpraca w systemach agentów
Typowa architektura	Input → Model → Output	Obserwacja → Agent → Działanie	Cel → Planer → Agenci wykonawczy → Koordynator → Feedback Loop
Typ interakcji	Jednostronna (input → output)	Dwustronna – środowisko ↔ agent	Wielostronna – agent ↔ środowisko ↔ inne agenty ↔ narzędzia (API, bazy, LLM)
Integracja z narzędziami i pamięcią	Brak lub statyczna integracja	Ograniczona integracja (np. lokalna pamięć stanu)	Dynamiczne użycie zewnętrznych narzędzi, kontekstu, pamięci długoterminowej i LLM

Organizacja 2.0

Organizacja 2.0 to firma, która łączy możliwości sztucznej inteligencji, automatyzacji i pracy cyfrowej, aby tworzyć inteligentne środowisko pracy oparte na danych, zwinności i ciągłym uczeniu się.

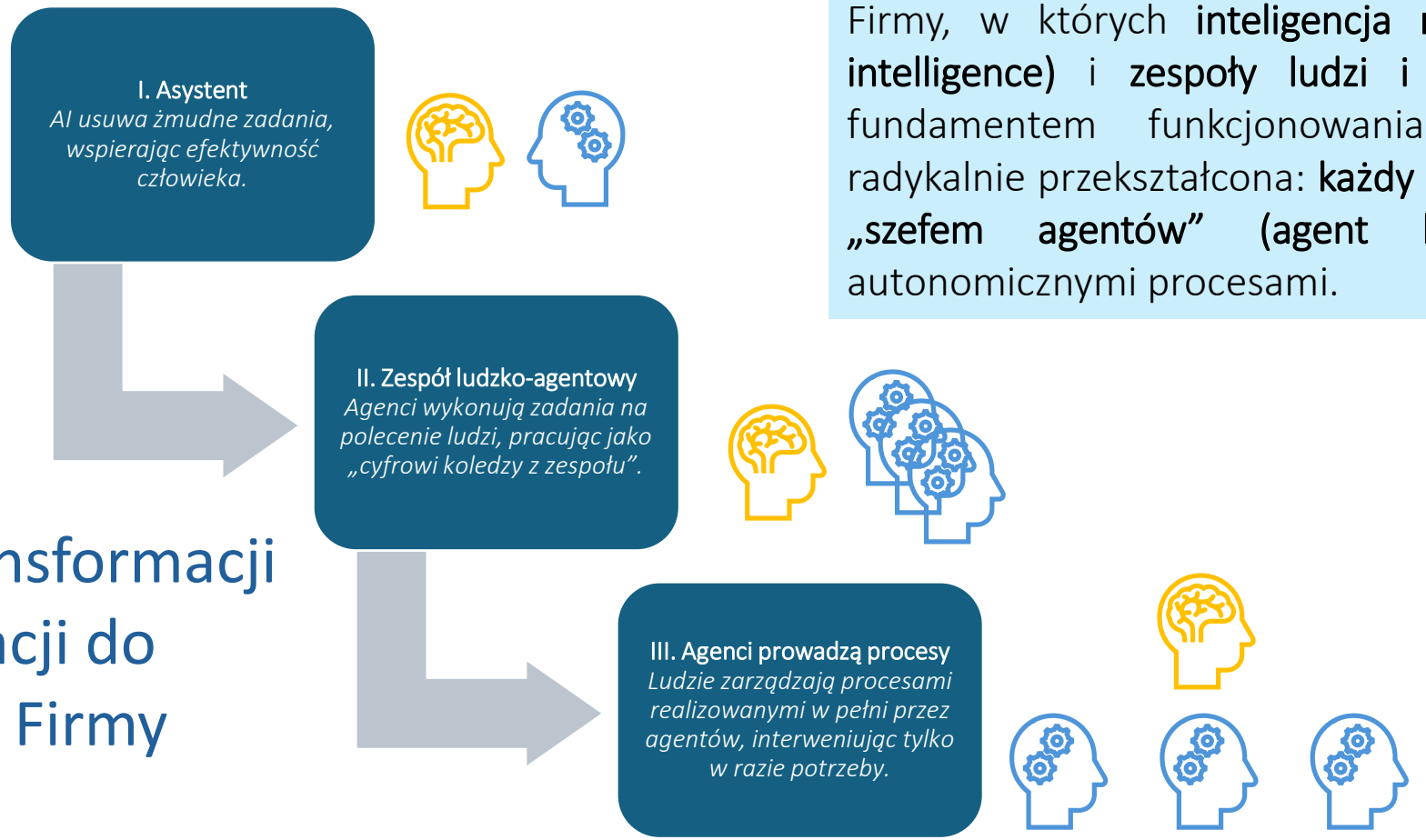
Organizacja 1.0	Organizacja 2.0
Tradycyjna praca biurowa	Praca cyfrowa
Ręczne procesy	Automatyzacja i AI
Decyzje oparte na intuicji	Decyzje oparte na danych
Struktura hierarchiczna	Struktura zwinna (agile)





Frontier Firms

Fazy transformacji organizacji do Frontier Firmy



Firmy, w których **inteligencja na żądanie (on-tap intelligence)** i **zespoły ludzi i agentów AI** będą fundamentem funkcjonowania. Praca zostanie radykalnie przekształcona: **każdy pracownik stanie się „szefem agentów” (agent boss)**, zarządzając autonomicznymi procesami.

2025 Work Trend Index Annual Report
https://assets-c4akfrf5b4d3f4b7.z01.azurefd.net/assets/2025/04/2025_Work_Trend_Index_Annual_Report_680aaa7fe52dd.pdf

+AI vs AI+





Element	+AI Tradycyjne wykorzystanie AI	AI+ Nowoczesne wykorzystanie AI
Rola AI	Wspiera istniejące procesy	Tworzy nowe modele biznesowe
Pozycja w strategii	Narzędzie pomocnicze	Centralny element strategii
Zakres zastosowań	Automatyzacja, optymalizacja	Innowacje, redefinicja biznesu
Wpływ na organizację	Wzrost wydajności	Transformacja kultury organizacyjnej

AI wspiera analizę danych klientów i optymalizację receptur produktów

+AI
poprawianie istniejącego

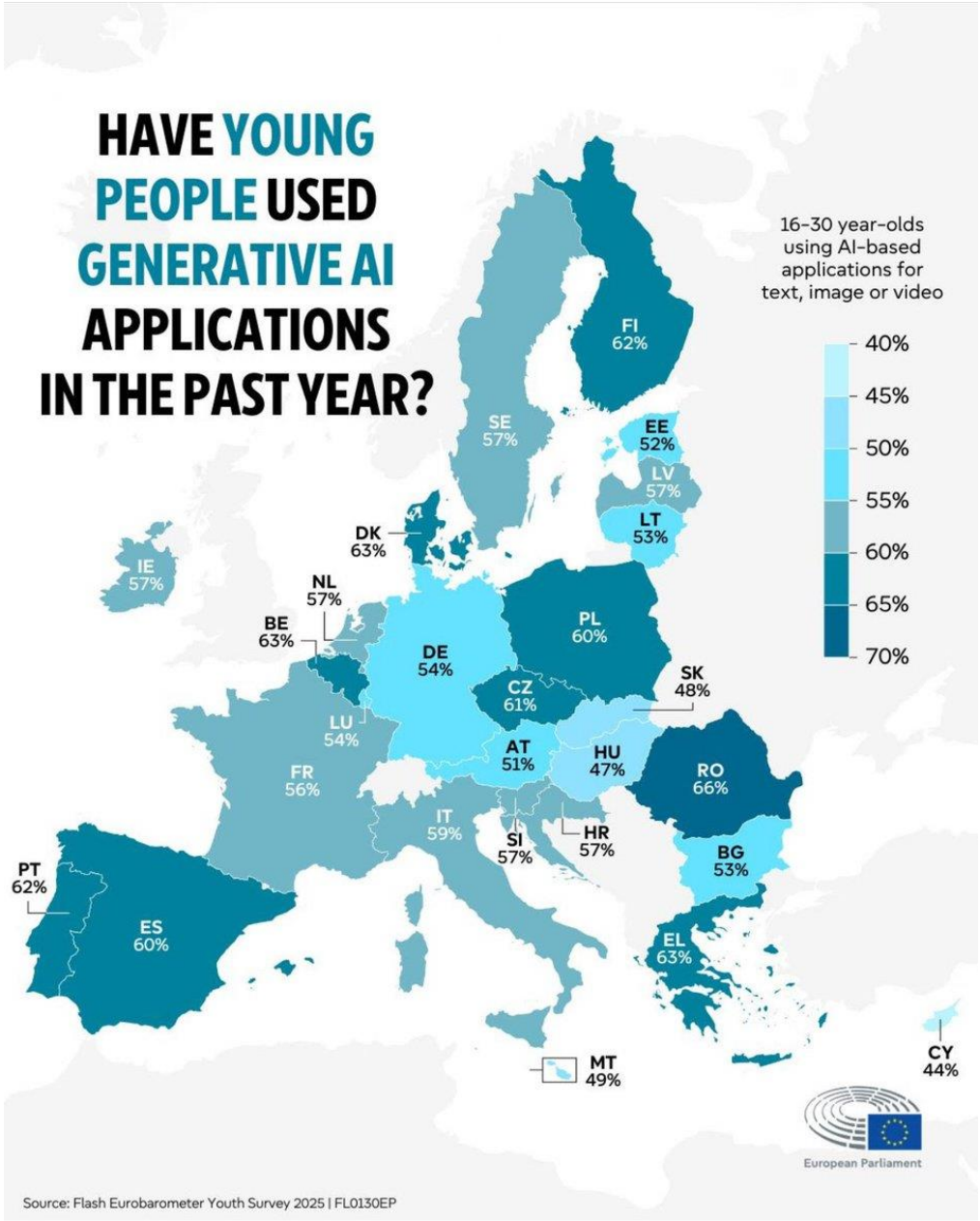
AI+
budowanie nowego

Kompleksowe zarządzanie ruchem i nawigacją w pojazdach autonomicznych

Tempo zmian

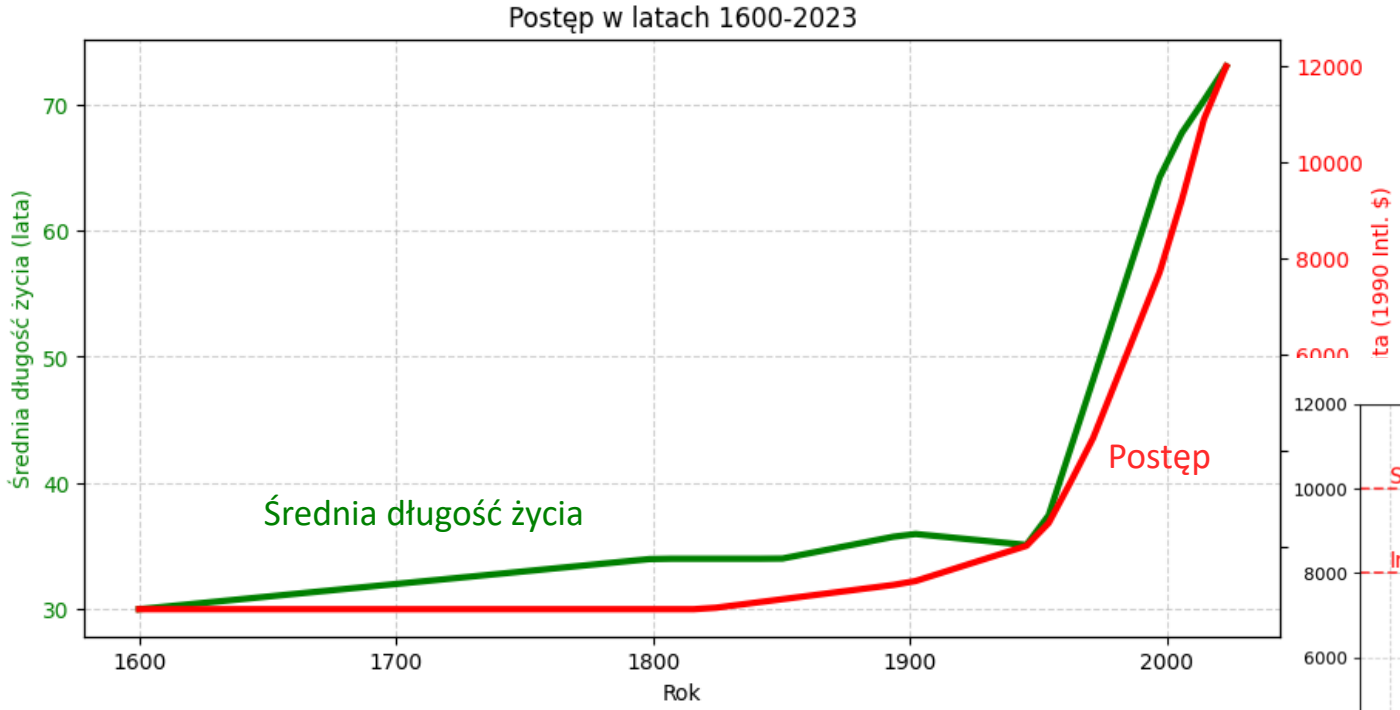
Czy młodzi ludzie korzystali z aplikacji generatywnej sztucznej inteligencji w ciągu ostatniego roku?

	Kraj	Pełna nazwa	Udział (%)
1	RO	Rumunia	66
2	EL	Grecja	63
3	DK	Dania	63
4	ES	Hiszpania	62
5	PT	Portugalia	62
6	PL	Polska	60
7	DE	Niemcy	60
8	IT	Włochy	59
9	LV	Łotwa	57
10	HR	Chorwacja	57
11	IE	Irlandia	57
12	SE	Szwecja	57
13	BE	Belgia	57
14	SK	Słowacja	56
15	NL	Niderlandy	56
16	FR	Francja	56
17	LU	Luksemburg	54
18	CZ	Czechy	54
19	BG	Bułgaria	53
20	EE	Estonia	52
21	FI	Finlandia	52
22	AT	Austria	51
23	SI	Słowenia	51
24	LT	Litwa	51
25	MT	Malta	49
26	HU	Węgry	47
27	CY	Cypr	44



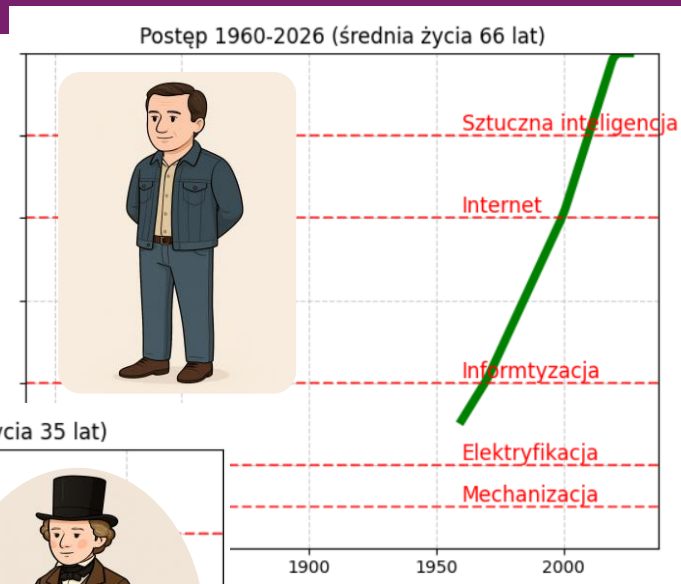
Osoby w wieku 16–30 lat korzystające z aplikacji opartych na sztucznej inteligencji do **tekstu, obrazu lub wideo**

Tempo zmian

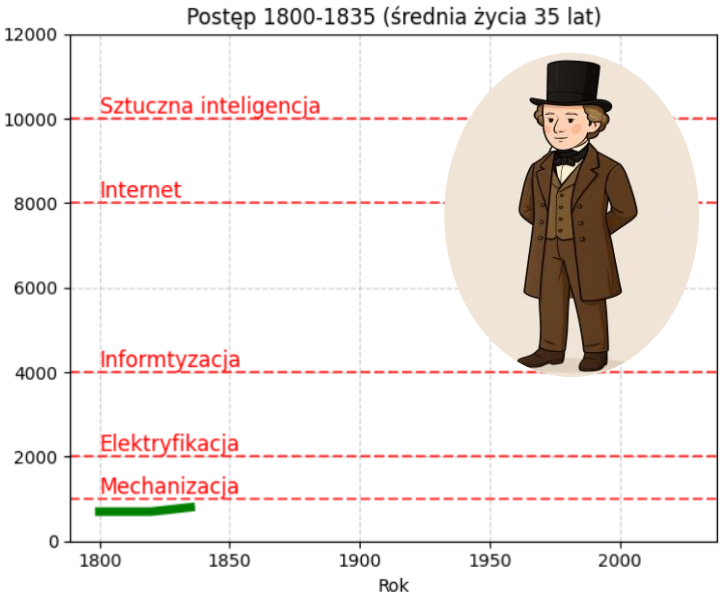


„Dłuższe życie i szybki postęp technologiczny wymagają ciągłego rozwoju kompetencji i zmiany zawodów.”

Rocznik 1800



Rocznik 1960



Wpływ AI na sektory gospodarki

Sektor gospodarki	AIIE	AIIE LLM	AIIE GenIm
K – Działalność finansowa i ubezpieczeniowa	2,05	1,97	1,60
M – Działalność profesjonalna, naukowa i techniczna	1,71	1,56	1,93
P – Edukacja	1,34	1,49	0,90
J – Informacja i komunikacja	1,20	1,13	1,38
Q – Opieka zdrowotna i pomoc społeczna	0,62	0,74	0,10
L – Działalność związana z obsługą rynku nieruchomości	0,36	0,34	0,37
D – Wytwarzanie i zaopatrywanie w energię elektryczną, gaz, parę wodną i gorącą wodę	0,20	0,05	0,65
O – Administracja publiczna i obrona narodowa; obowiązkowe zabezpieczenia społeczne	0,39	0,39	-0,33
G – Handel hurtowy i detaliczny	0,03	0,19	-0,55
S – Pozostała działalność usługowa	-0,25	-0,16	-0,38
C – Przetwórstwo przemysłowe	-0,40	-0,51	-0,06
B – Górnictwo i wydobywanie	-0,68	-0,85	-0,13
N – Działalność w zakresie usług administrowania i działalność wspierająca	-0,61	-0,47	-0,85
H – Transport i gospodarka magazynowa	-1,02	-1,09	-0,64
F – Budownictwo	-1,10	-1,11	-0,70
R – Działalność związana z kulturą, rozrywką i rekreacją	-1,60	-0,26	-1,40
I – Działalność związana z zakwaterowaniem i usługami gastronomicznymi	-1,12	-0,73	-2,07
A – Rolnictwo, leśnictwo, łowiectwo i rybactwo	-2,00	-1,85	-1,94

AIIE – sztuczna inteligencja ogólnie
AIIE LLM – duże modele językowe
AIIE GenIm – generowanie obrazów

Wskaźnika AIIE (Artificial Intelligence Industry Exposure) mierzy **stopień narażenia sektorów** na wpływ sztucznej inteligencji.

- Wartości > 0 → sektor bardziej narażony na AI niż średnia gospodarki.
- Wartości < 0 → sektor mniej narażony na AI niż średnia gospodarki.

Źródło:
AI na polskim rynku pracy. Polski Instytut Ekonomiczny *(październik 2024)*
https://pie.net.pl/wp-content/uploads/2024/10/AI-na-polskim-rynku-pracy.pdf

Nowe zawody

Nowe zawody

📱 Inżynier AI

- 🔷 Projektowanie i trenowanie modeli sztucznej inteligencji
- 🔷 Integracja systemów AI z infrastrukturą IT

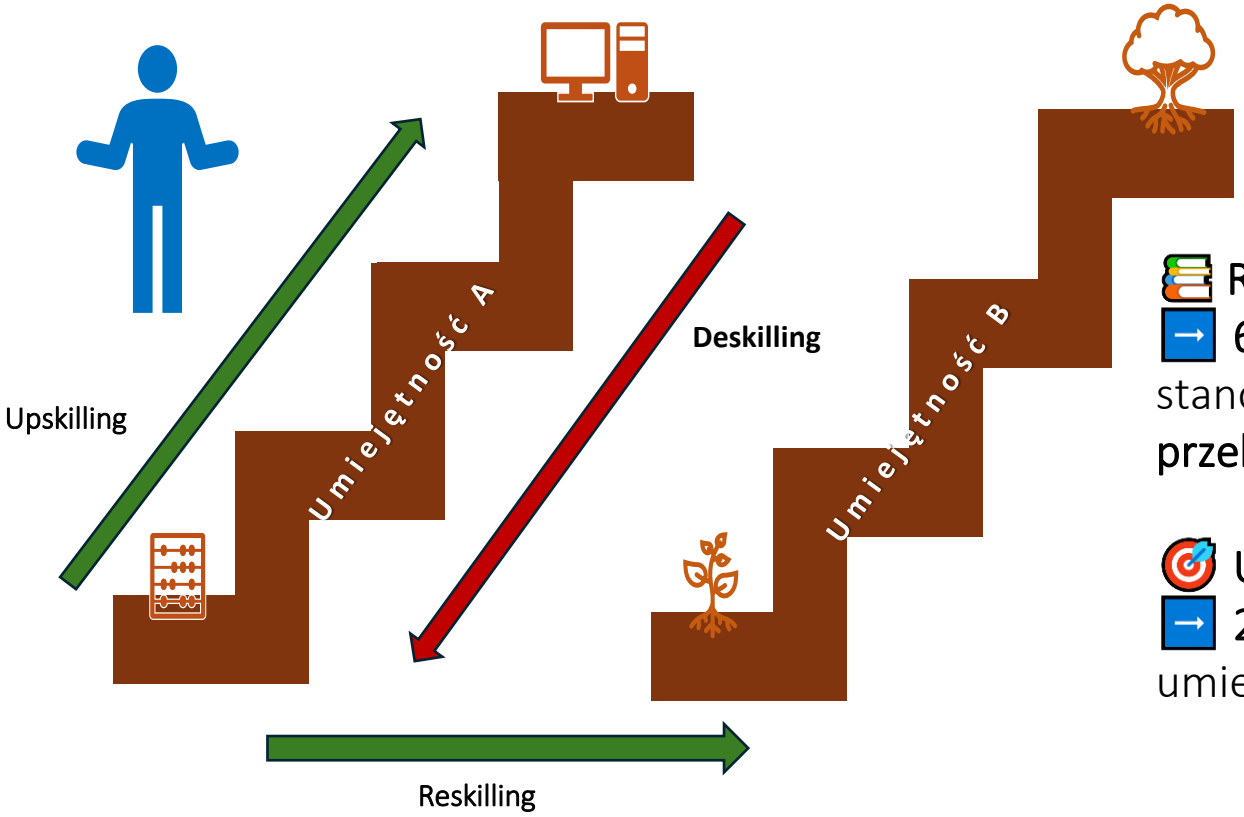
 Inżynierowie promptów

- 🔷 Optymalizacja zapytań do dużych modeli językowych (LLM)
- 🔷 Dostosowywanie AI do konkretnych zadań biznesowych

⚖️ Specjaliści ds. etyki AI

- 🔷 Tworzenie zasad odpowiedzialnego wdrażania AI
- 🔷 Kontrola systemów AI pod kątem zgodności z regulacjami

Rozwój umiejętności



📚 **Reskilling** – przekwalifikowanie się do nowego zawodu:
➡️ **60% pracowników** obawia się, że AI wpłynie na ich stanowisko pracy, ale **75% jest gotowych się przekwalifikować**

🎯 **Upskilling** – rozwijanie nowych kompetencji:
➡️ **20% wzrostu efektywności** dzięki AI w nauce nowych umiejętności (Harvard Business Review, 2024).

Nowe formy wspierania kwalifikacji

Gamifikacja



+ AI +

Storytelling



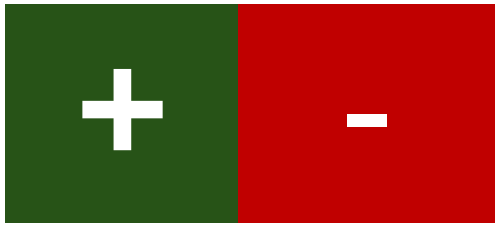
Ale to już inna historia …… ☺

Dyskusja ?

Czy AI powinno zastąpić człowieka w podejmowaniu decyzji?

https://docs.google.com/document/d/1ppugfJri962lcjaa9TgT6CvzvVRxDm1kyaa-mKq5evE/edit?usp=sharing

- **Ekonomia** (Economics) 🏢
- **Technologia** (Technology) 💻
- **Etyka** (Ethics) ⚖️
- **Społeczeństwo** (Society) 🌍
- **Prawo** (Law) ⚖️
- **Środowisko** (Environment) 🌱
- **Bezpieczeństwo i prywatność** (Security and Privacy) 🔒
- **Organizacja** (Organization) 🏢
- **Innowacyjność i kreatywność** (Innovation and Creativity) 💡
- **Dane i ich pozyskiwanie** (Data and Its Acquisition) 📊
- **Finanse** (Finance) 💰
- **Relacje biznesowe** (Business Relationships) 🤝
- **Regulacje i zgodność** (Regulations and Compliance) 📜
- **Bezpieczeństwo** (Security) 🛡️
- **Adaptacyjność i skalowalność** (Adaptability and Scalability) 📈
- **Kompetencje ludzkie** (Human Competencies) 👤
- **Biologia** (Biology) 🧬
- **Dominacja** (Dominance) 🏆
- **Edukacja i rozwój personalny** (Education and Personal Development) 🎓

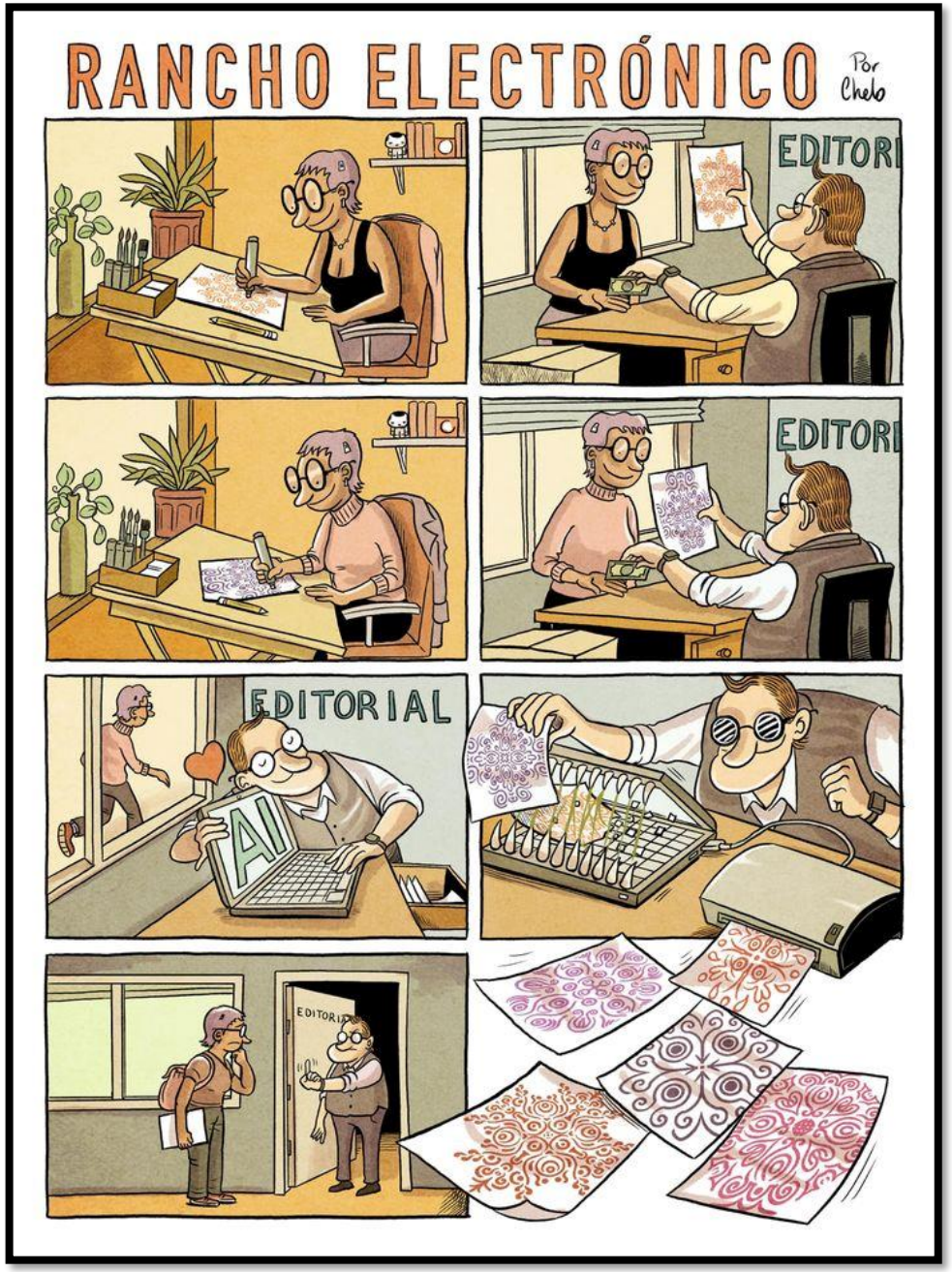


Dyskusja ?

Czy AI powinno zastąpić człowieka w podejmowaniu decyzji?



https://youtu.be/VsZpSzLeohY?si=09SAiiz3A9XtcF5S


RANCHO ELECTRÓNICO
Por Chelo
EDITORI
EDITORI
EDITORIAL
AI
EDITORIA

dr inż. Michał Malinowski

https://www.drmalinowski.edu.pl

Temat n.

Podsumowanie

Wprowadzenie do Sztucznej Inteligencji

„Po przeprowadzeniu tokenizacji tekstu, lematyzujemy każdy token, aby zminimalizować jego semantyczną złożoność, a następnie wrzucamy wyniki do embeddingu. W tym procesie musimy być świadomi potencjalnego biasu w korpusie uczącym, danych odstających oraz overfittingu, które mogą prowadzić do halucynacji podczas generowania wyników.
Kluczowe jest także, aby korpus uczący, na którym trenujemy model ML, był odpowiednio zrównoważony, by uniknąć zakłóceń w predykcjach, a także zapewnić modelowi możliwość generalizacji. Warto także zwrócić uwagę na interpretowalność i explainability modelu, szczególnie przy analizie wyników generowanych przez ANN.
Będzie istniało ryzyko Jailbreaking i Data Poisoning jeśli nie zadałbym o Data Governance, a … i jeszcze jedno zapewne ktoś zrobi Slop.”

Czy rozumiesz teraz ten tekst ?

Domeny Sztucznej Inteligencji

AI - Sztuczna Inteligencja
AIs - Autonomiczne Systemy Inteligentne
ANN - Sztuczne Sieci Neuronowe
cAI – Współpracująca Sztuczna Inteligencja
CC - Przetwarzanie kognitywne
CNN - Sieci Konwolucyjne
DL - Uczenie Głębokie
ES - Systemy Ekspertowe
HCI - Interakcja człowiek-komputer
IoT - Internet Rzeczy
KBS - Systemy Oparte na Wiedzy
LLMs - Duże Modele Językowe
MAS - Systemy Wieloagentowe
ML - Uczenie Maszynowe
NLP - Przetwarzanie Języka Naturalnego
RNN - Rekurencyjne Sieci Neuronowe
RPA - Zrobotyzowana Automatyzacja Procesów
XAI - Wytłumaczalna Sztuczna Inteligencja

