

< GOVTECH_ FESTIVAL />



OŚRODEK
PRZETWARZANIA
INFORMACJI
PAŃSTWOWY INSTYTUT BADAWCZY



GOVTECH_
FESTIVAL />

AUTONOMIA KONTRA AUTOMATYZACJA – ETYKA SZTUCZNEJ INTELIGENCJI

mec. Robert Kroplewski,



Rewolucja 4.0 v. (+)

„Bardziej niż zapoczątkowanie czwartej rewolucji przemysłowej, sztuczna inteligencja prowokuje rewolucję kulturową.

potrafi (?) lub ma (?) na celu przekształcenie naszej przyszłości, ale nie wiemy dokładnie, **jak i w co ...**

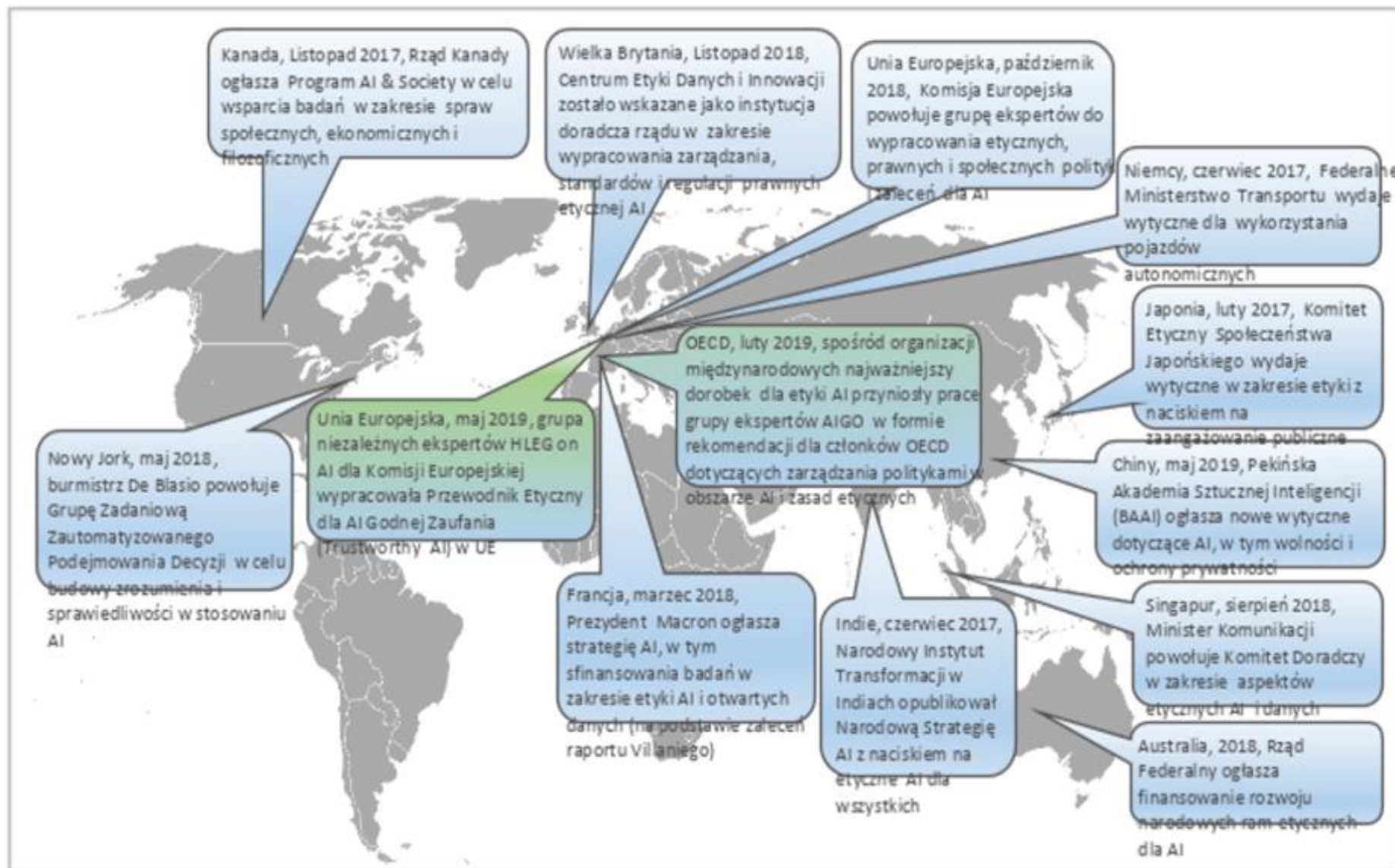
Dlatego AI pobudza zarówno fascynację, jak i strach.”

dyrektor generalny Audrey Azoulay (UNESCO)

Systemy etyczne

- Cywilizacje – osie organizacji społeczeństw
- Quincunx, PIĘCIOMIAN Konecznego (Dobro. Prawda, Zdrowie, Dobrobyt, Piękno) – stosunek do nich definiuje cywilizację (też i jej etykę)
- Europa = Humanizm, Indywidualizm, Wolność, Różnorodność, (r)Ewolucje
- Rosja = zależność – totalizacja - osamotnienie
- Azja = zbiorowość – całość – holistyczność
- Indie = hybrydowość adaptacyjna
- Chiny = sprawność monowładzy – „2030 lider wszechczasów”
- Arabia = feudalizm – organizacja
- Ameryka Pn = Utylitaryzm – ekonomizacja
- Ameryka Pld = środowisko – kruchość - odnawialność
- Afryka = wykluczenie – więzy plemienne – asymilacja
- **KORPORACJE = e-commerce (usługi a nie wartości)**

Konkurencja na ramy etyki jako standardów techniczno-organizacyjnych



GOVTECH FESTIVAL

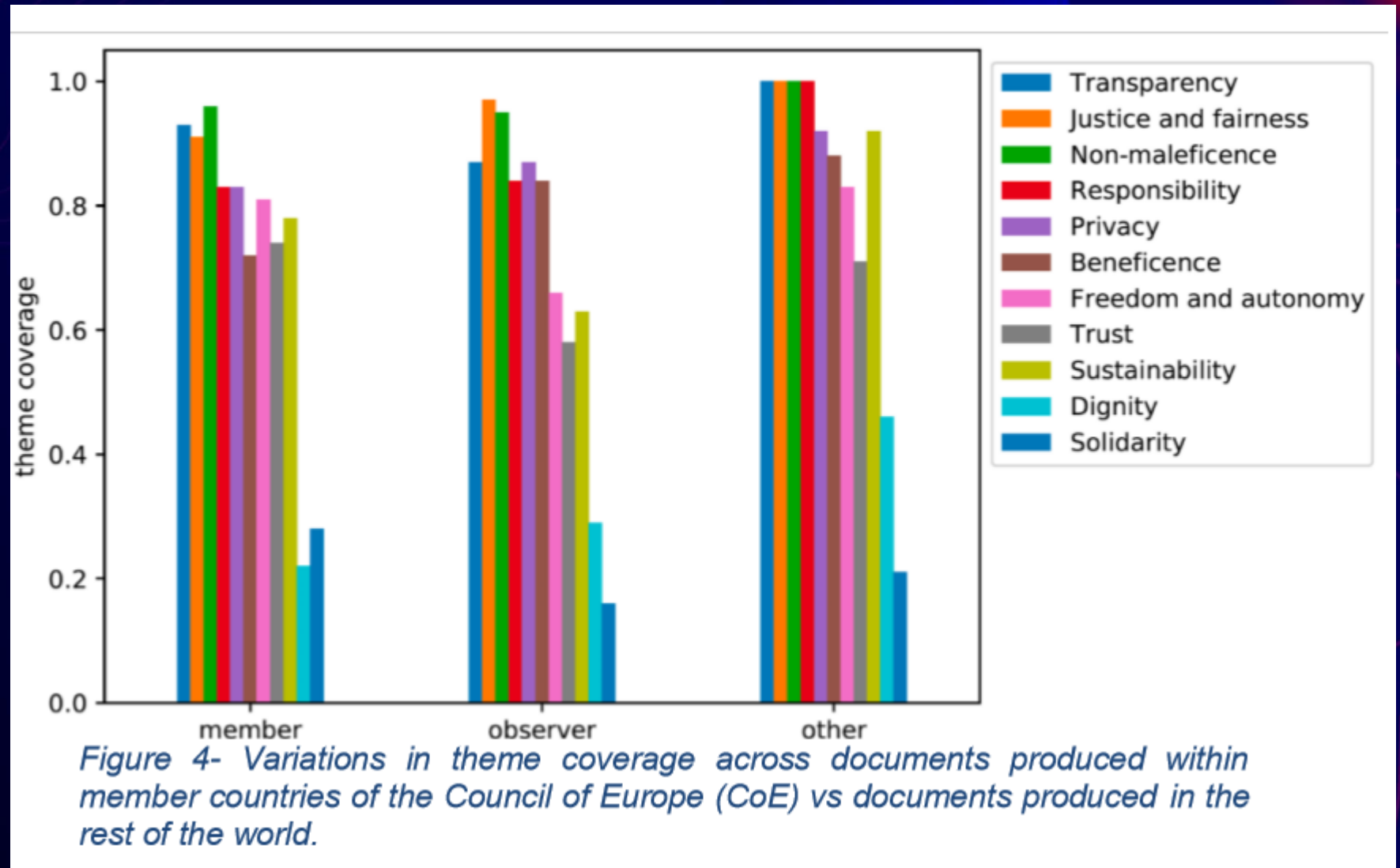


Zasady OECD dotyczące SI dla odpowiedzialnego zarządzania godną zaufania sztuczną inteligencją:

- SI powinna przynosić korzyści ludziom i planecie, stymulując wzrost sprzyjający włączeniu społecznemu, zrównoważony rozwój i dobrobyt.
- Systemy SI powinny być projektowane z poszanowaniem praworządności, praw człowieka, wartości demokratycznych i różnorodności, a także powinny obejmować odpowiednie zabezpieczenia - na przykład umożliwiające w razie potrzeby interwencję człowieka - w celu zapewnienia uczciwego i sprawiedliwego społeczeństwa.
- Powinna istnieć przejrzystość i odpowiedzialne ujawnianie informacji na temat systemów SI, aby ludzie rozumieli wyniki oparte na sztucznej inteligencji i mogli je kwestionować.
- Systemy SI muszą działać solidnie, pewnie i bezpiecznie przez cały cykl życia, a potencjalne zagrożenia powinny być stale oceniane i zarządzane.
- Organizacje i osoby opracowujące, wdrażające lub obsługujące systemy SI powinny zostać pociągnięte do odpowiedzialności za ich prawidłowe funkcjonowanie

< GOVTECH_ FESTIVAL / >

Statystyki
priorytetyzacji
zasad etycznych
w dokumentach
strategicznych i
badaniach
społecznych





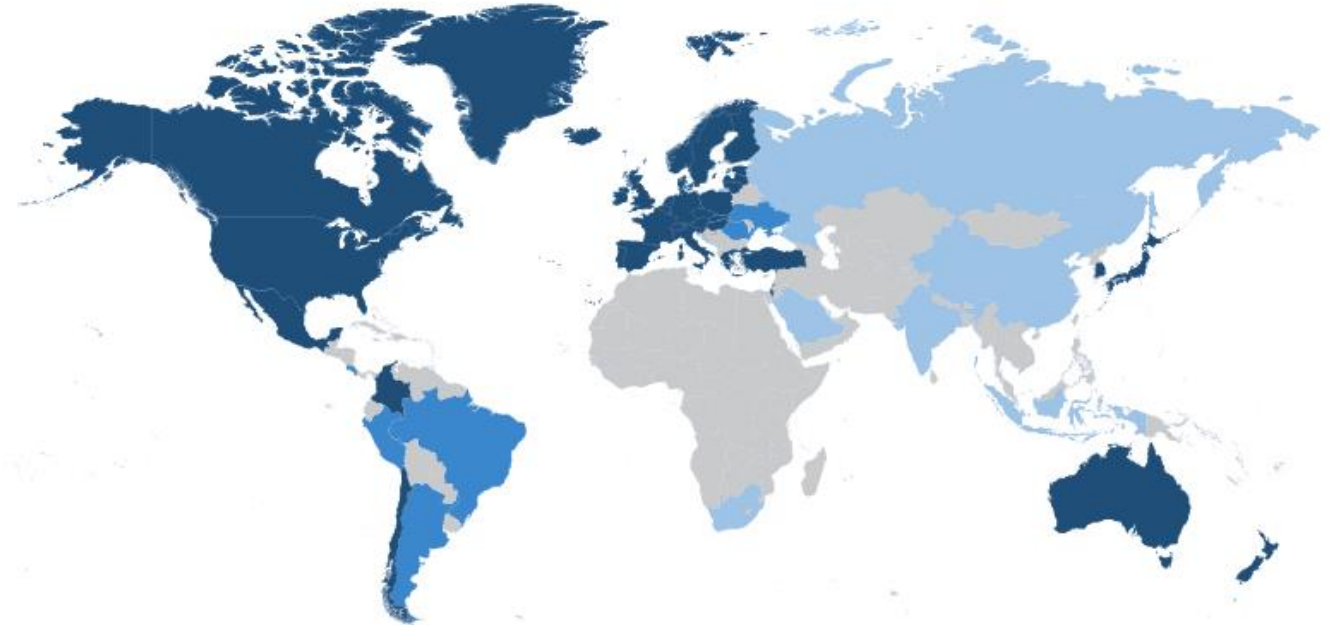
Polityka Rozwoju Sztucznej
Inteligencji w Polsce
na lata 2020 – 2027

Godna Zaufania Sztuczna Inteligencja
autonomia i konkurencja
+PL

projekt dla konsultacji społecznych

Raczejpolska Polska, Warszawa, 20 sierpnia 2019

Governments that have committed to the AI Principles



OECD members

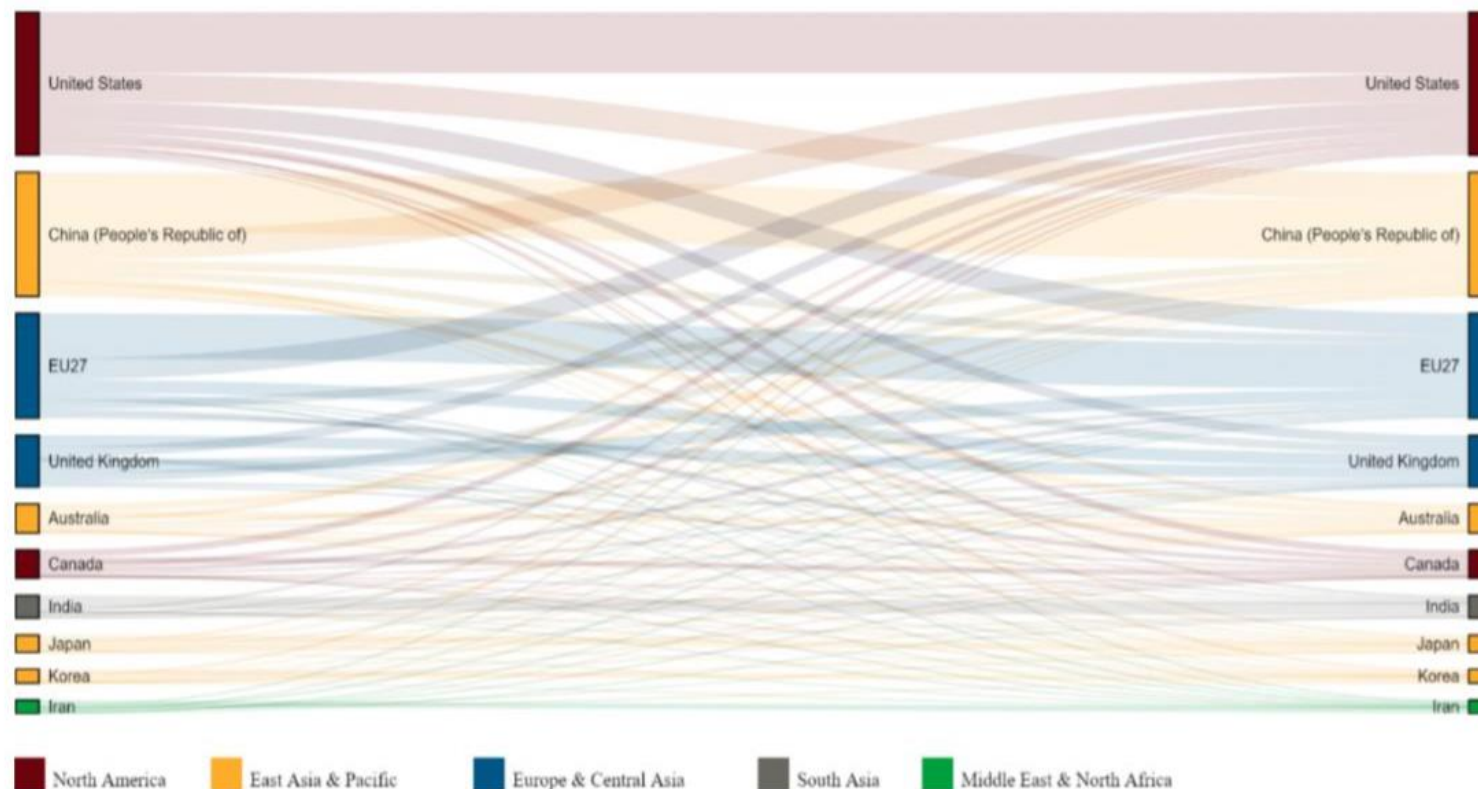
G20 principles, based on OECD

Adherents

Find detailed information about hundreds of policy initiatives in over 60 countries and territories, and the EU.

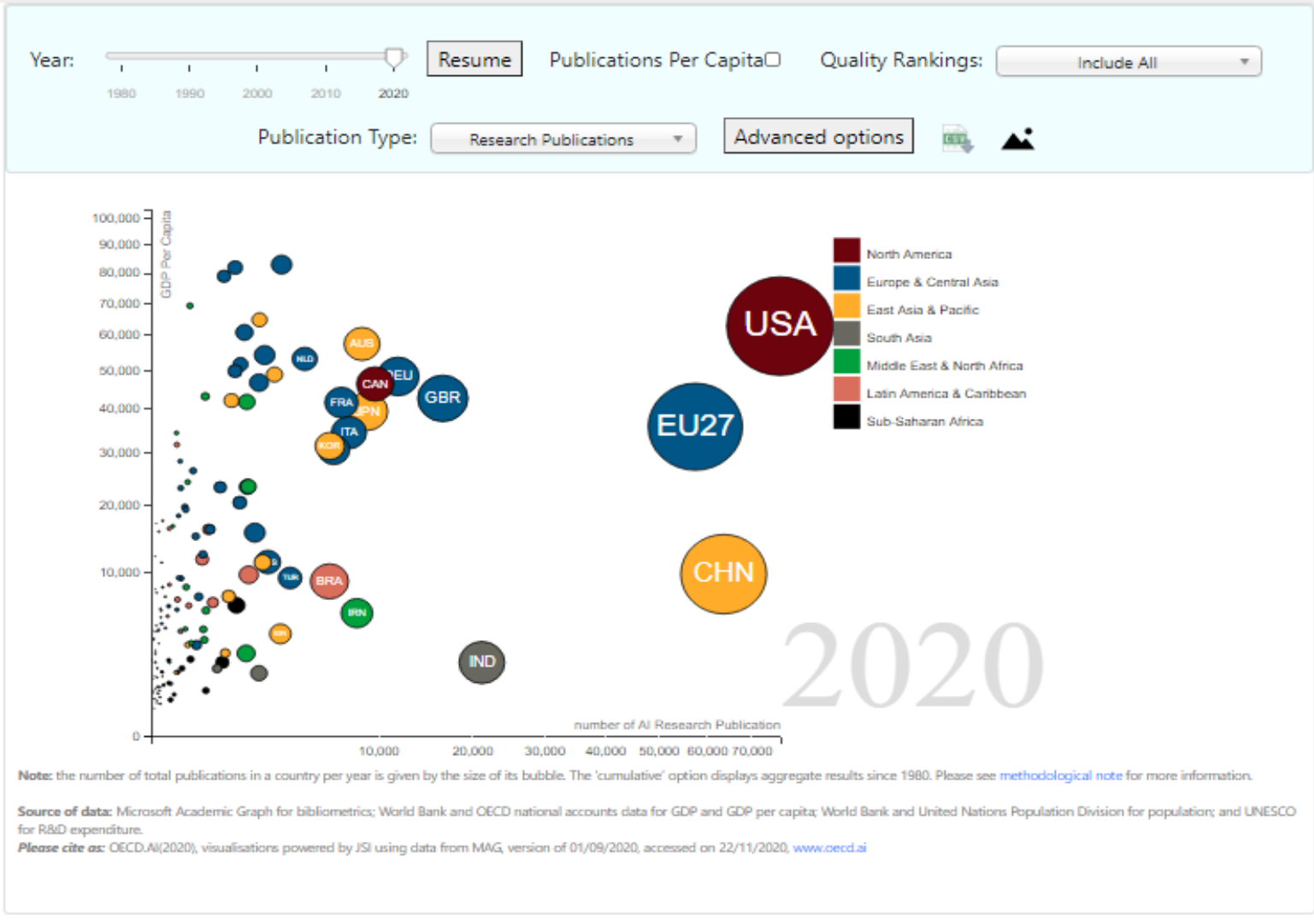


Figure 5. Domestic and international AI research collaboration

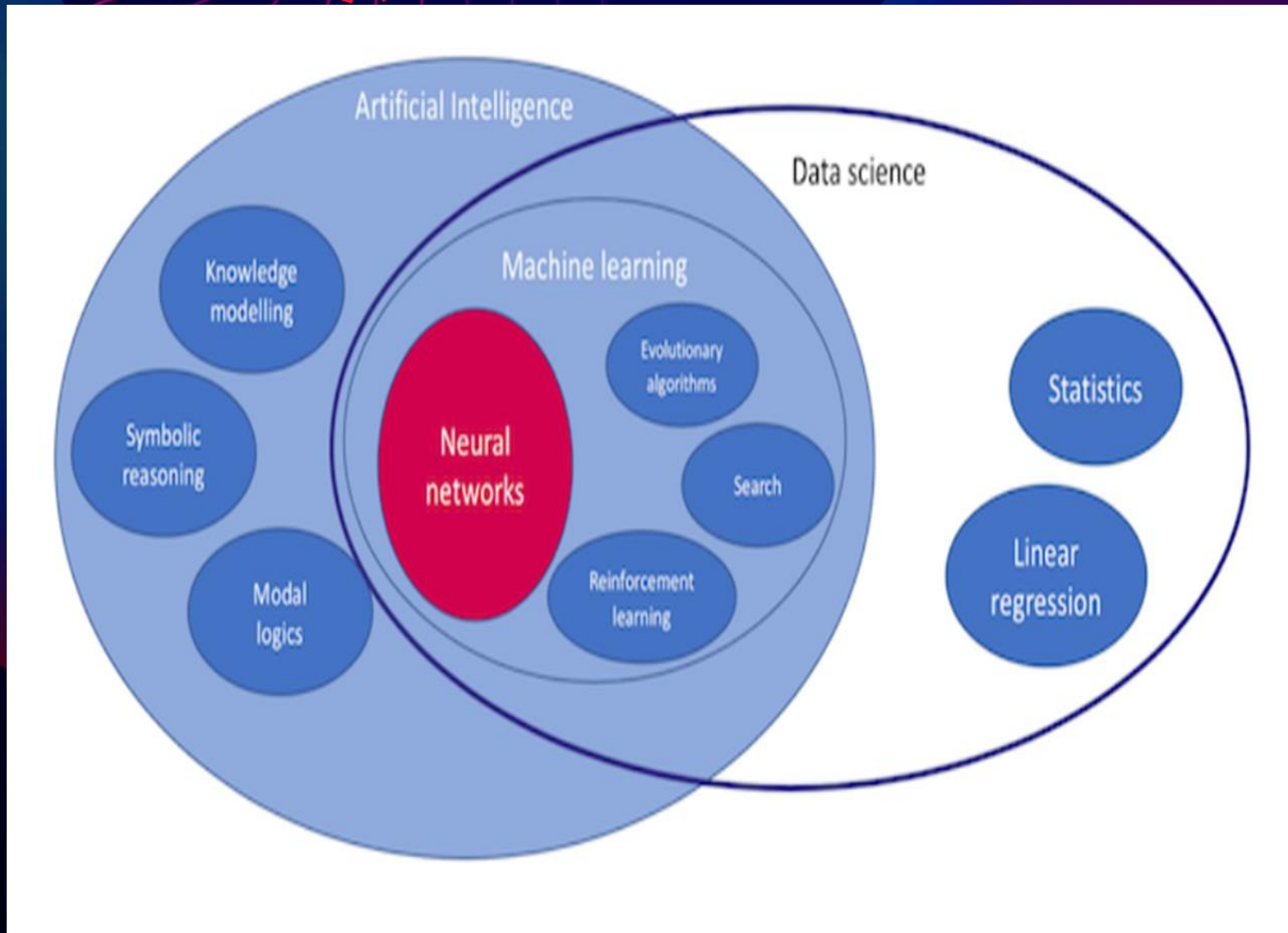


Notes: The thickness of a connection represents the number of joint AI publications between two countries since 1980. "Domestic collaboration" shows co-authorship involving different institutions within the same country. Please see methodological note for more information at www.oecd.ai. EU27 = the European Union minus the United Kingdom.

AI publications vs GDP per capita by country, region, in time



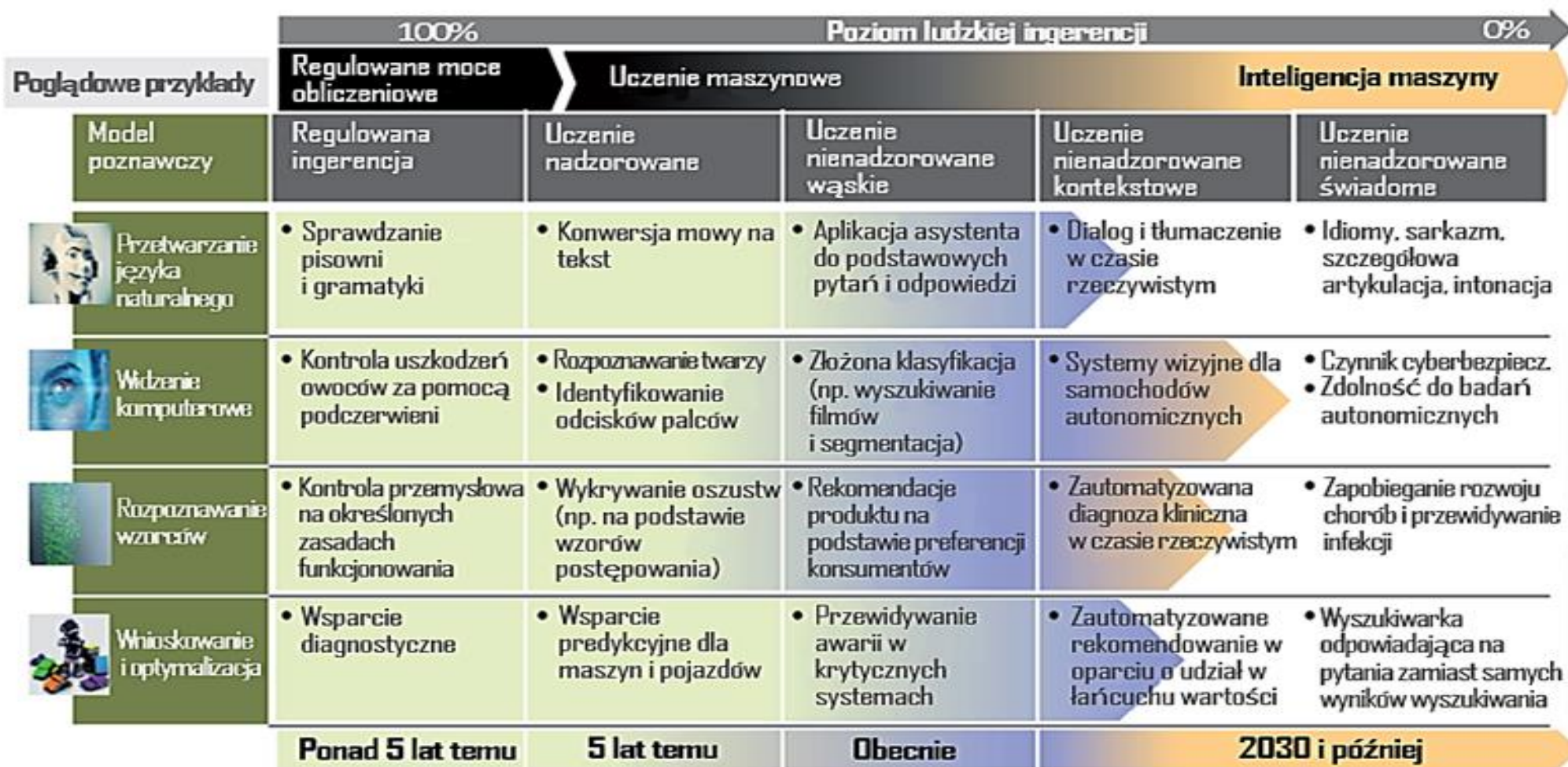
Stan rzemiosła AI



Dziedzina Sztucznej Inteligencji rozwija się jako multidyscyplinarny obszar badawczy, który powstał na przecięciu filozofii, matematyki, ekonomii, psychologii, neurobiologii, teorii kontroli, robotyki oraz inżynierii komputerowej, w ramach którego rozwijany jest szereg technologii obliczeniowych (informatycznych).

Ich główną funkcją jest uczenie maszynowe, przetwarzanie języka naturalnego, automatyczne wnioskowanie, przewidywanie, optymalizowanie, nawigowanie i rekomendowanie, rozpoznawanie oraz generowanie mowy i obrazu.

Rozwój sztucznej inteligencji i jej przyszłych form



„AI będzie dobra w określonych zadaniach obliczeniowych, ale jesteśmy daleko od ogólnej inteligencji..”

„Każdy obiekt stanie się świadomy kontekstu, będzie ultra-personalizowany i dostosowany do potrzeb dzięki postępowi AI.”

„Inteligencja ogólna wymaga innego rodzaju niedeterministycznej architektury obliczeniowej, innej niż ta, która dziś istnieje.”

Źródła: Strony internetowe firm; A.T. Kearney; A.T. Kearney/warsztat Światowego Forum Ekonomicznego; listopad 2016; wywiady ekspertów



„Większość technologii ma świetlisty awers, ale życie dało im rewers – czarną rzeczywistość”

– Stanisław Lem

Konstytucja RP

„Art. 30 Przyrodzona i niezbywalna godność człowieka stanowi źródło wolności i praw człowieka i obywatela. Jest ona nienaruszalna, a jej poszanowanie i ochrona jest obowiązkiem władz publicznych.”

GOVTECH_ FESTIVAL

„Świadomość (...) nie jest po prostu sumą informacji z receptorów zmysłowych, jest rzeczywiście nową jakością. (...) Świadomość ma w sobie dynamikę układu o innej kombinacji niż ilościowe bogacenie homogennej masy informacyjnej.”

Włodzimierz Sedlak, w: Homo electronicus,
Syzyfa Wcielenie Drugie.

(Rozdział 3, Homo electronicus, PIW, 1980,

„Sztuczna Inteligencja jest martwa... Absolutnie nie ma postępu w zrozumieniu natury inteligencji ludzkiej” (nurt poznawczy Sztucznej Inteligencji)

Aktualne stwierdzenie twórcy nowoczesnej AI, Marvin'a Minskiego, z 2006 r.: wobec Manifestu z Darmouth, Warsztaty Darmouth College, USA, 1956r.

< GOVTECH_ FESTIVAL />

Stan techniki na dziś pozwala na ustalenie, że mimo pozorów, AI przetwarzając informację nie potrafi myśleć i daleko jeszcze (jeśli w ogóle jest to możliwe) do możliwości pobrania wszystkich komponentów człowieka do maszyny algorytmicznej. Robot przestrzega zestawu procedur, które pozwalają mu na interakcję z nami, ale poza bardzo precyzyjnymi ramami, w których ma oddziaływać, nie potrafi stworzyć prawdziwej relacji społecznej, nawet wzbogacony systemem oprogramowań samouczących się.

Jednakże jest też zupełnie odwrotnie. Nie ma zauważalnych znaków postępu po stronie człowieka w rozumieniu działania AI.

A właśnie relacja autonomii człowieka *versus* automatyka maszyn uczących się pozostaje centralnym wyzwaniem dla rozwoju AI dla samostanowienia człowieka i dobrobytu społeczeństw.

Ramy dla Godnej Zaufania AI

Wprowadzenie

Godna Zaufania AI

AI zgodna z prawem

(nie omówione w tym dokumencie)

Etyczna AI

Odporna AI

Rozdział I

Podstawy Godnej Zaufania AI

Dostosowane do zasad etycznych opartych na prawach podstawowych

4 zasady etyczne

Uznawanie i zmniejszanie napięć pomiędzy nimi

- poszanowanie ludzkiej autonomii
- przeciwdziałanie szkodom
- uczciwość
- zrozumiałość

Rozdział II

Realizacja Godnej Zaufania AI

Wdrożenie kluczowych wymagań

7 praktycznych zasad

Stała ocena i wdrażanie ich w ciągu cyklu życia Systemu AI za pomocą:

Metody techniczne

Metody nietechniczne

- przewodnia i nadzorczą rolę człowieka
- stabilność i bezpieczeństwo
- ochrona prywatności i danych
- przejrzystość
- różnorodność, niedyskryminacja i sprawiedliwość
- dobrostan społeczny i środowiskowy
- odpowiedzialność

Rozdział III

Ocena Godnej Zaufania AI

Zoperacjonalizowanie kluczowych wymagań

Lista kontrolna Godnej Zaufania AI

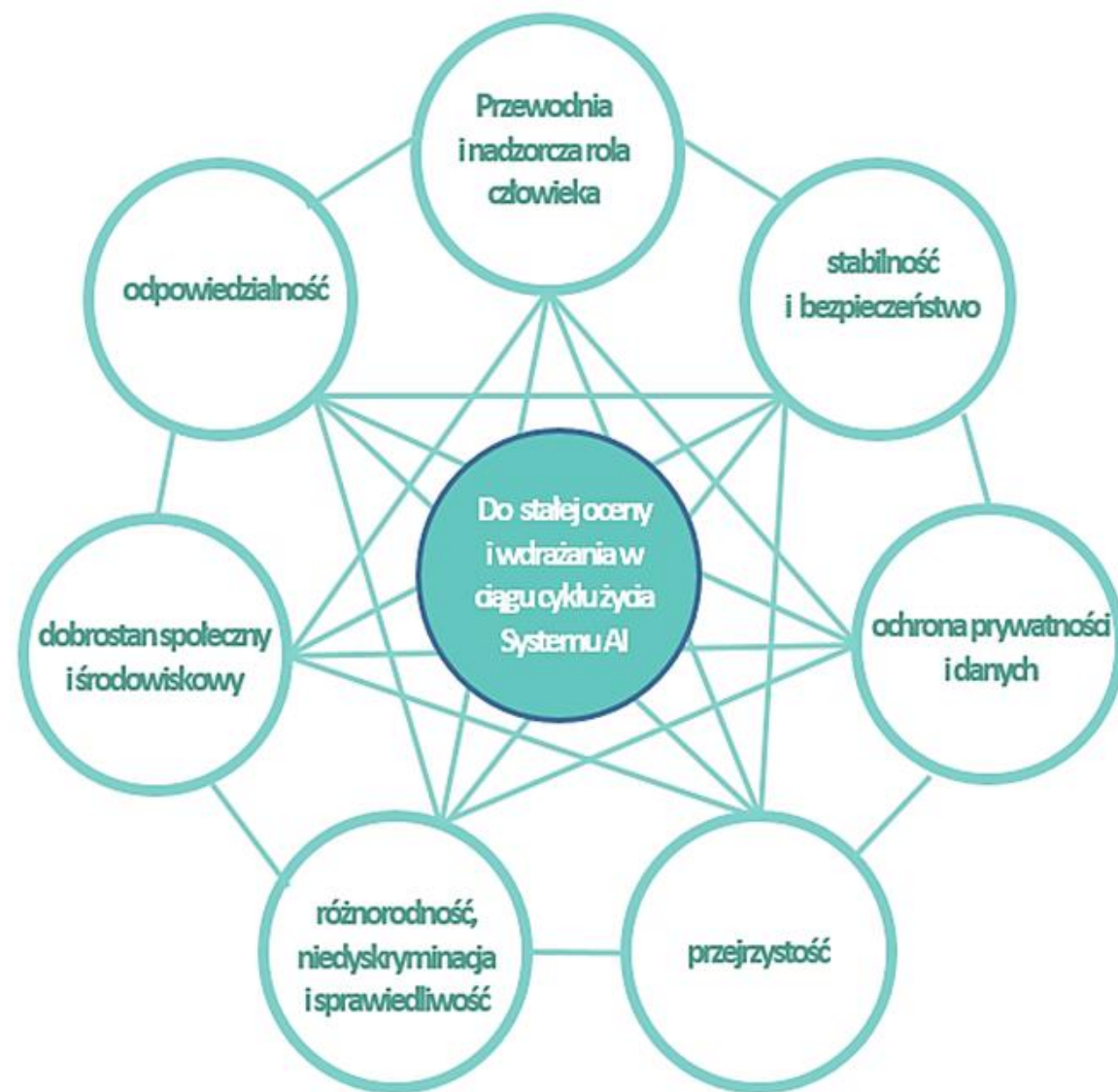
Dostosowanie do konkretnych wdrożeń AI

Źródło: Przewodnik Etyczny dla AI opracowany przez HLEG on AI dla KE.

Kompas dla kompromisów w konfliktach wartości:

Godność ludzka, dobrostan, ochrona przed szkodą

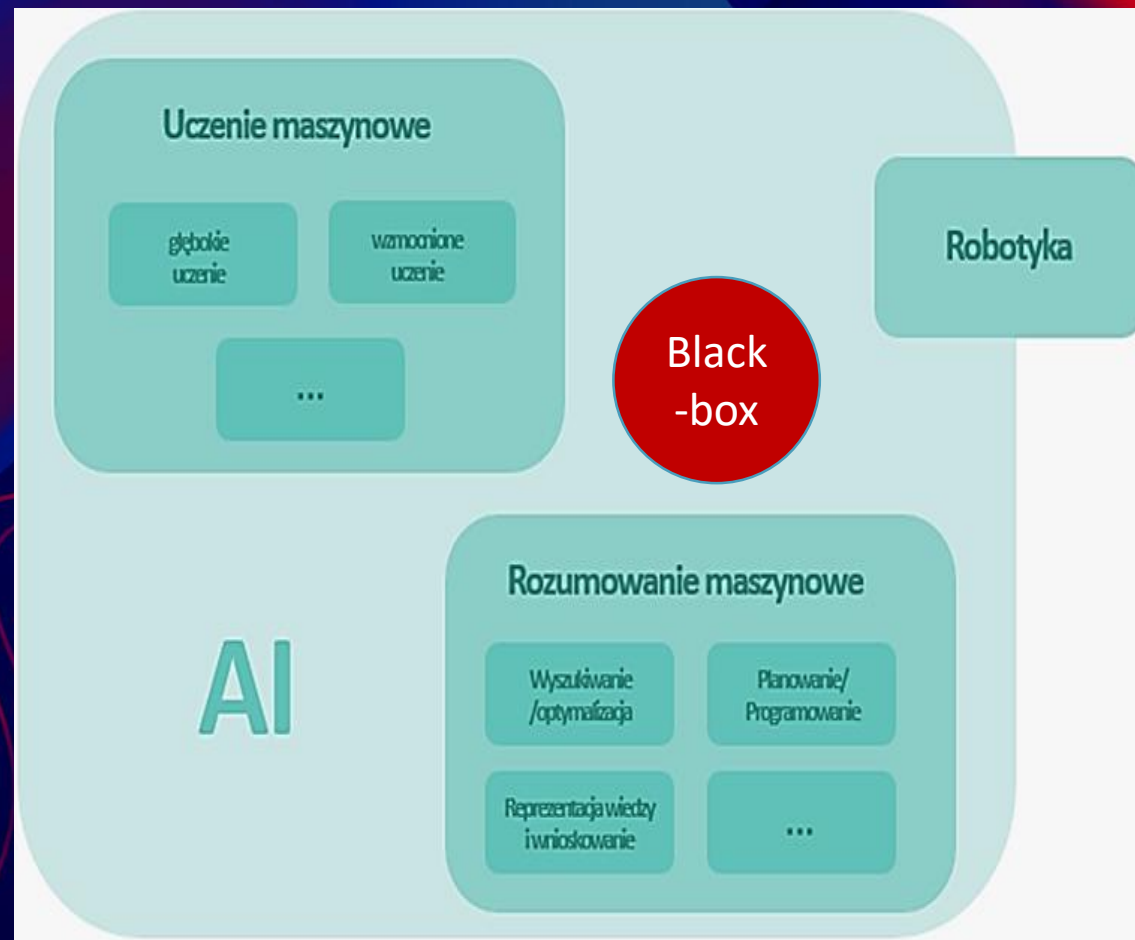
Projekt the Global Code of Ethics by AHEG – UNESCO



GOVTECH FESTIVAL

Obecny opis tego, czym jest system sztucznej inteligencji, oparty jest na koncepcyjnym spojrzeniu na sztuczną inteligencję, wyszczególnionym w „Artificial Intelligence: A Modern Approach” (Russel, S. i Norvig, P., 2009 [1]).

Pogląd ten jest zgodny z szeroko stosowaną definicją sztucznej inteligencji jako „studium obliczeń, które umożliwiają postrzeganie, rozumowanie i działanie” (Winston, 1992 [2]) oraz z podobnymi ogólnymi definicjami (Gringsjord, S. & Govindarajulu, NS, 2018 [3]).



Źródło: Przewodnik Etyczny dla Trustworthy AI opracowany przez HLEG AI dla Komisji Europejskiej.

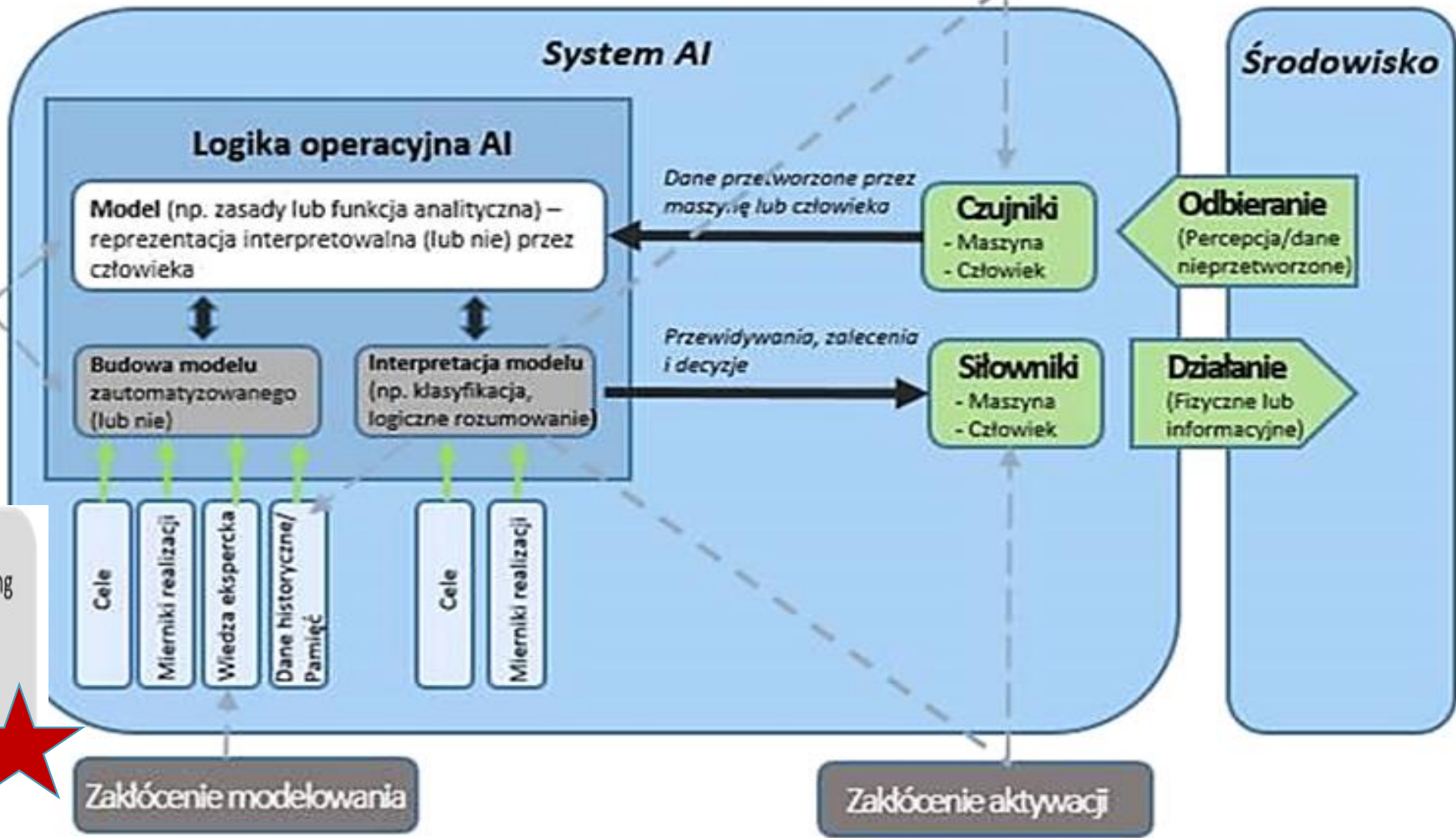
Zakłócenie techniczne

- General Principles**
1. Inclusive and sustainable growth and well-being
 2. Human-centred values and fairness
 3. Transparency and explainability
 4. Robustness and safety
 5. Accountability



Zakłócenie modelowania

Zakłócenie percepcyjne



Zakłócenie aktywacji

Cykl życia AI



Źródło: Przewodnik Etyczny dla Trustworthy AI, HLEG on AI dla KE.

Tabela nr 8 *Maszyna vs. Człowiek vs. Zespoły Kognitywne Ludzi i Maszyn*

Jednostka	Zalety	Wady
Komputer	<ul style="list-style-type: none">• szybkie generowanie i testowanie w dużej przestrzeni do poszukiwania rozwiązań• szybkie przetwarzanie dużej ilości danych	<ul style="list-style-type: none">• generator rozwiązań jest niedokładny w otwartym świecie• dane nie w pełni przedstawiają otwarty świat
Zespół ludzi	<ul style="list-style-type: none">• doświadczenie ludzkie w otwartym świecie• wszechstronne doświadczenie interdyscyplinarnych zespołów w wielu dziedzinach	<ul style="list-style-type: none">• występują koszty koordynacji
Zespół kognitywny Ludzi i Maszyn (najwyższa wydajność)	<ul style="list-style-type: none">• komplementarne i wzajemne uczenie się eliminuje sytuacje awaryjne obu stron i zwiększa ich efektywność	<ul style="list-style-type: none">• potrzebujemy lepszej teorii i praktyk w zakresie budowania zespołów ludzi i maszyn

Źródło: *Keflik, M. w PARC Blog, Half human, half computer? Meet the modern centaur.*

Tabela nr 9 Porównanie przewag ludzi i maszyn cyfrowych:

Przewaga maszyn nad ludźmi

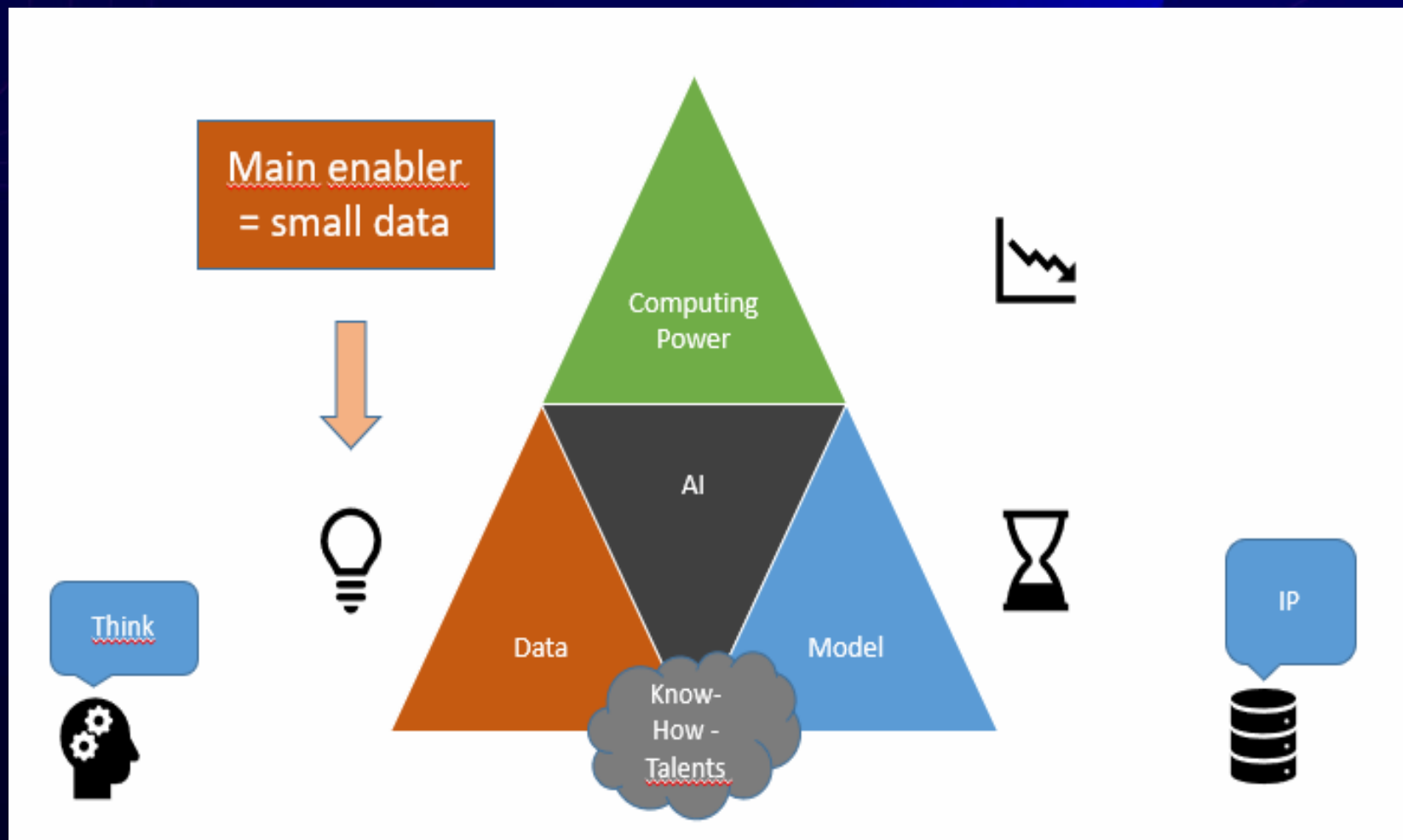
- powtarzalność przewidywalnych zadań;
- wykonywanie zadań, które zależą od mocy obliczeniowej;
- klasyfikacja wielkich ilości danych i wkładów pochodzących z percepcji maszyn i ludzi;
- podejmowanie decyzji na podstawie skonkretyzowanych reguł;

Przewaga ludzi nad maszynami:

- doświadczanie autentycznych emocji i budowanie relacji;
- formułowanie pytań i wyjaśnień w różnych skalach, poziomach i źródłach;
- decydowanie, jak użyć limitowanych zasobów w strategicznych wymiarach, w tym które z zadań maszyny powinny być wykonane i w jakie dane je wyposażyć;
- wytwarzanie produktów i rezultatów pracy użytecznych dla ludzi, komunikowanie o tym ludziom;
- podejmowanie decyzji w związku z abstrakcyjnymi pojęciami dotyczącymi wartości.

Źródło: Opracowanie własne MC na podstawie [Artificial Intelligence in Education, promises and Implication for teaching & Learning](#)

Wnętrze AI na rozdrożach



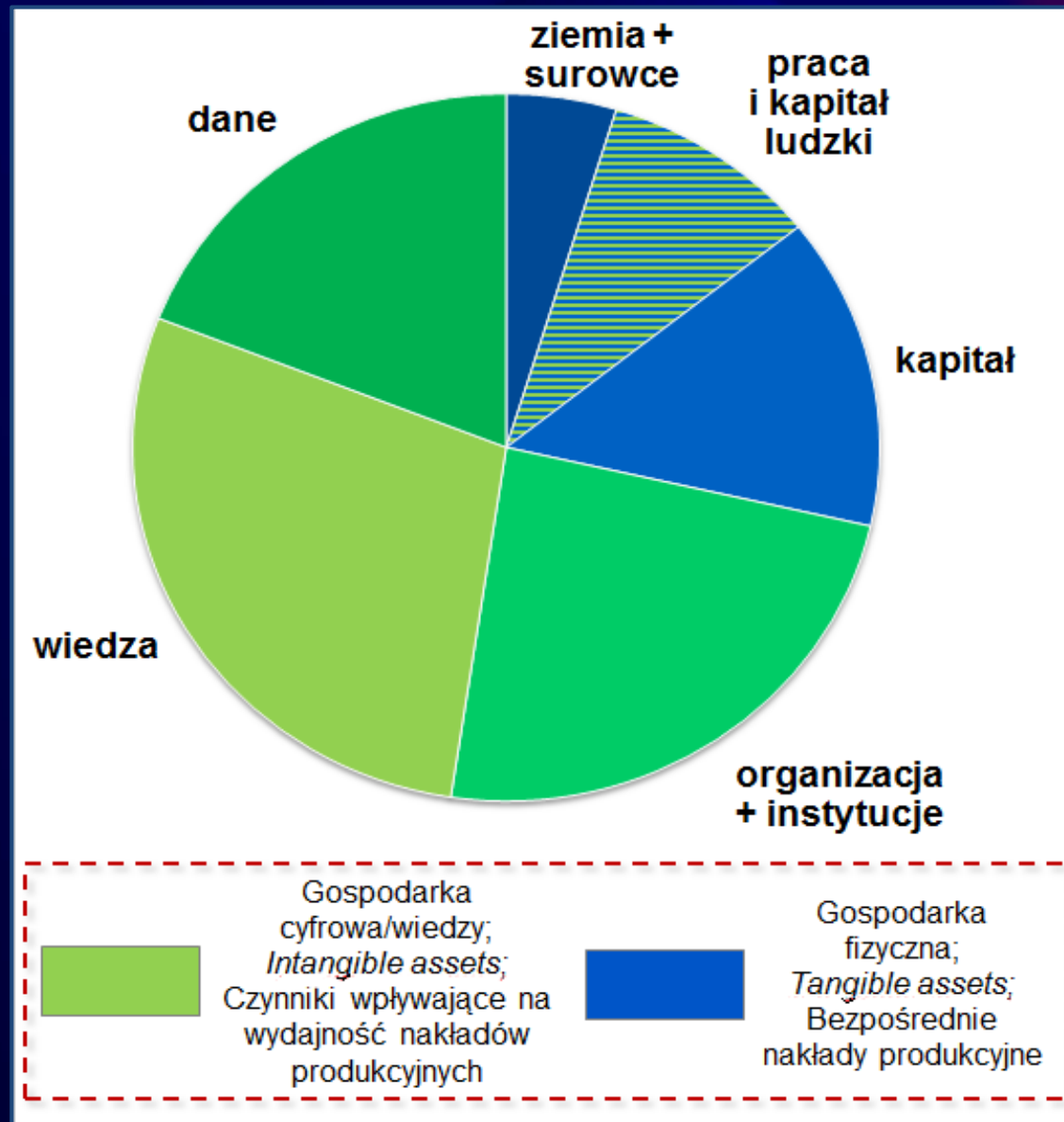
- Konkurencja, kooperacja
- Zarządzanie danymi, przenaszalność danych
- Privacy, Security, Consumer Protection
- Handel cyfrowy, FTA
- Cyfrowe zarządzanie

- Zgodność z prawem
- Etyka
- Odporność
- Wolność innowacji
- Autonomia i kreatywność ludzi



AI Otwarte Ramy Innowacji

Schematyczny podział wpływu czynników produkcji, MR, projekt Strategia Produktyności



< GOVTECH_ FESTIVAL />

„Człowiek jest bezsprzecznie najinteligentniejszym zwierzęciem i podobno z tego powodu niezbyt pasuje do biologicznego królestwa zwierząt.

Nałożył na siebie pojęciową siatkę, od góry nakrywając się logiką, którą wymyślił jako konieczną wymierność wszelkich poczynań oddzielających go od zwierząt. Przez 23 wieki sądzono, że istnieje tylko jedna logika sformułowana przez Arystotelesa, jak niemal przez tyle samo stuleci istniało przekonanie o jednej geometrii wymyślonej przez Euklidesa. Wiek XIX obalił oba intelektualne mity. **Istnieje tyle logik i geometrii, ile ich potrafi wymyślić człowiek.** Podobnie, jak współrzędne kartezjańskie nie stanowią jedynej "siatki geograficznej" dla naszych pojęć.

W biologii jest inaczej. W logice nieobserwowanie zasad prowadzi do nieciągłości, czyli pomyłki w rozumowaniu, a tym samym - pomyłki w poznaniu rzeczywistości. Wydaje się, że w biologii odstępstwo od przyjętych pojęć jest przeciwstawnością życia określamy ją jako śmierć. W logice biologicznej człowiek obawia się nie tyle pomyłki w poznaniu, ile samej natury przedmiotu badań. Być może jest to lęk podświadomy." Homo electronicus, Włodzimierz Sedlak, 1980.

< GOVTECH_ FESTIVAL />

<https://ec.europa.eu/futurium/en/ai-alliance-consultation>

<https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence>

https://ec.europa.eu/knowledge4policy/ai-watch_en

<https://oecd.ai/>

<https://en.unesco.org/artificial-intelligence>

https://www.wipo.int/about-ip/en/artificial_intelligence/

<https://www.gov.pl/web/cyfryzacja/ai>



MEC. ROBERT KROPLEWSKI

solicitor,
plenipotentiary of Minister of Digital Affairs in
Poland resp. for Information Society
CEO of Future Industry Digital Platform
expert #HLEGonAI (EC)
expert #AIGO (OECD)
expert #CAHAI (COF)

robert@kroplewski.com

www.kroplewski.com

TT@bobo.kroplewski



DZIĘKUJEMY

...