

STRATEGIA WYKORZYSTANIA AI

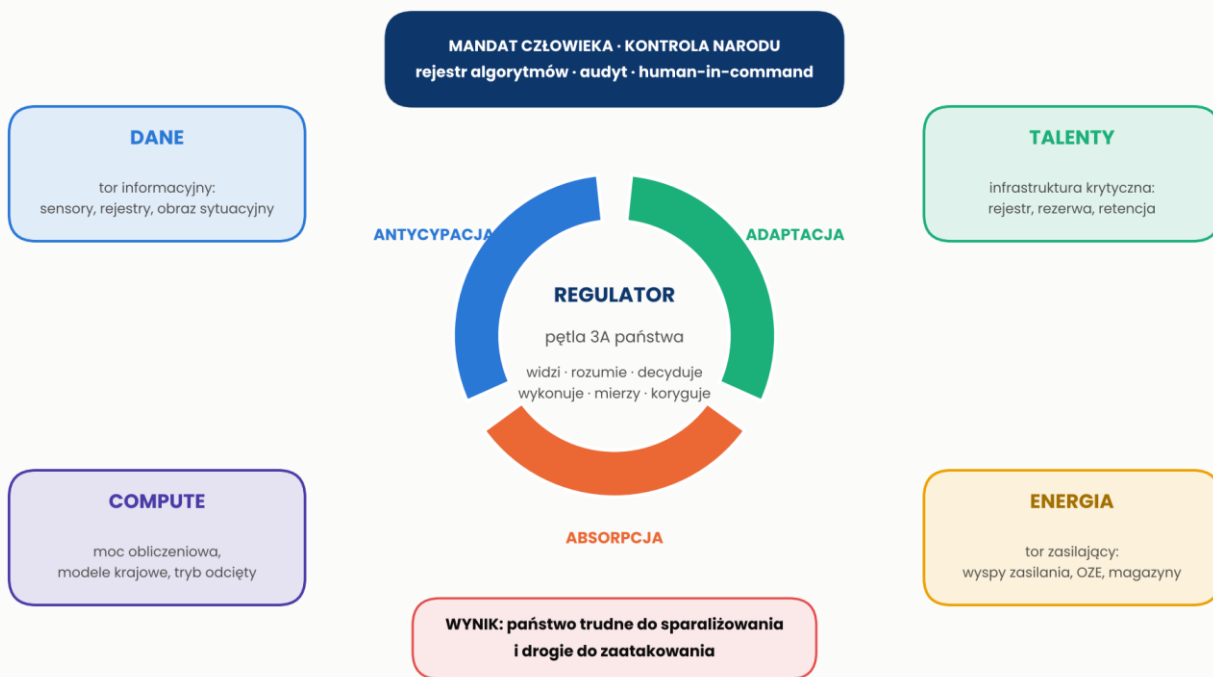
dla zapewnienia bezpieczeństwa Polski

Od państwa reagującego do państwa przewidującego: AI, cyberodporność, energia i talenty jako infrastruktura krytyczna Rzeczypospolitej.

FORMUŁA NACZELNA

Bezpieczeństwo Rzeczypospolitej w erze sztucznej inteligencji, robotów autonomicznych i biotechnologii jest funkcją szybkości i jakości pętli sterowania państwa. Dlatego Polska musi zbudować suwerenny, sprzężony i kontrolowany system AI bezpieczeństwa — oparty na własnych danych, mocy obliczeniowej, energii i talentach zarządzanych jak infrastruktura krytyczna — w którym maszyna przyspiesza poznanie, a człowiek zachowuje mandat, podpis i odpowiedzialność.

OTOCZENIE: cyberataki · drony i autonomia · operacje wpływu (FIMI) · sabotaż · bio-AI · przymus technologiczny



Bezpieczeństwo = szybkość pętli 3A × suwerenność torów (dane · compute · energia · talenty) × kontrola regulatora

Rysunek 1. Formuła naczelna jako pętla sterowania państwa: cztery tory suwerenności zasilają regulator, którego działanie pozostaje pod mandatem człowieka i kontrolą Narodu. Opracowanie własne TechnoSanacji.



Skrót strategii

Pole	Treść strategii
Najkrótsza formuła	Od państwa, które reaguje na incydenty, do państwa, które przewiduje, wykrywa i koryguje szybciej niż przeciwnik — dzięki suwerennej AI, odpornym torom danych i energii oraz talentom zarządzanym jak infrastruktura krytyczna.
Problem	Polska jest jednym z najczęściej atakowanych państw Europy (260 783 incydentów obsłużonych przez CERT Polska w 2025 r., wzrost o 152% r/r), a jednocześnie należy do ostatnich w UE pod względem wykorzystania AI w przedsiębiorstwach (8,4% wobec 20,0% średniej UE). Państwo ma nowe ustawy i strategie, ale nie ma regulatora, który łączyłby AI, cyberbezpieczeństwo, energię i kadry w jedną pętlę sterowania.
Rozstrzygnięcie	AI zostaje uznana za warstwę sterowniczą bezpieczeństwa narodowego, a talenty AI i cyber — za infrastrukturę krytyczną z rejestrem, wskaźnikiem stanu, planem ciągłości, rezerwą i właścicielem odpowiedzialności.
Cel główny	Narodowy System AI Bezpieczeństwa (NSAIB): suwerenny, sieciowy i odporny układ antycypacji, wykrywania, decyzji i korekty, obejmujący cyberprzestrzeń, infrastrukturę krytyczną, pole walki, przestrzeń informacyjną i bezpieczeństwo biologiczne.
Instrumenty	Rada Bezpieczeństwa AI przy Prezydenckim Centrum Strategicznym, Narodowy Obraz Sytuacyjny, Tarcza AI-Cyber dla infrastruktury krytycznej, Suwerenny Stos AI, Energia dla AI, Rejestr Zdolności Krytycznych AI-Cyber, Cyber-Rezerwa Narodowa, Akademia Obrony Cyfrowej, Sandbox Obronny, rejestr algorytmów publicznych i stały red teaming.
Zasada AI	AI wykrywa, modeluje i przyspiesza — człowiek decyduje, podpisuje i odpowiada. Użycie siły, decyzje wobec obywatela i zmiany ustrojowe pozostają w mandacie człowieka (human-in-command).
Wynik	Krótszy czas wykrycia i reakcji, rosnący koszt ataku na Polskę, malejąca podatność torów sterowania, dodatnie saldo talentów krytycznych i zdolność do przełączania się między wariantami strategicznymi bez utraty celów państwa.

FORMUŁA NACZELNA

Bezpieczeństwo Rzeczypospolitej w erze sztucznej inteligencji, robotów autonomicznych i biotechnologii jest funkcją szybkości i jakości pętli sterowania państwa. Wygrywa ten, kto szybciej widzi, trafniej rozumie, wcześniej decyduje i skuteczniej koryguje.

Dlatego Polska musi zbudować suwerenny, sprzężony i kontrolowany system AI bezpieczeństwa — oparty na własnych danych, mocy obliczeniowej, energii i talentach zarządzanych jak infrastruktura krytyczna — w którym maszyna przyspiesza poznanie, a człowiek zachowuje mandat, podpis i odpowiedzialność.

FORMUŁA W ZAPISIE SKRÓCONYM

Bezpieczeństwo = szybkość pętli 3A × suwerenność torów (dane · compute · energia · talenty) × kontrola regulatora przez Naród.

Iloczyn, a nie suma: zero w którymkolwiek członie zeruje wynik. Najszybsza pętla na cudzych danych jest pętlą przeciwnika; najlepsza technologia bez ludzi jest martwym obiektem; najsprawniejszy regulator bez kontroli staje się państwem nadzoru.

Pozycja	Wartość
Tytuł	Strategia wykorzystania AI dla zapewnienia bezpieczeństwa i cyberbezpieczeństwa Polski
Charakter	dokument programowy TechnoSanacji do debaty publicznej, eksperckiej i parlamentarnej
Miejsce w architekturze	rozwińnięcie Narodowej Strategii Bezpieczeństwa (Tom 1-jawny) w domenie AI i cyberprzestrzeni; dokument sprzężony z programami: Obrona Narodowa, Talenty – infrastruktura krytyczna, Narzędzia AI, Rezyliencja, Doktryna Energetyczna
Podstawa programowa	Manifest, Diagnoza Rzeczypospolitej 2026, Narodowa Strategia Bezpieczeństwa, Obrona Narodowa, Talenty – infrastruktura krytyczna państwa, Narzędzia AI, Controlling cybernetyczny, Infrastruktura odporności, Analityczny proces wyboru technologii militarnych, Graf multi-agentów.
Podstawa badawcza	destylaty TechnoSanacji: Obrona (611 PDF), Geopolityka (540 PDF), Odporność RES (1725 PDF), Cybernetyka państwa CybControl (623 PDF), Analiza systemowa SYS (3231 PDF), Analiza sieciowa GR (1667 PDF), AHP/MCDM, Innowacje, Edukacja; dokumenty urzędowe RP i UE; raporty CERT Polska i Eurostatu. Publikacja powstała na bazie danych portalu TechnoSanacja z wykorzystaniem narzędzi AI przy udziale autora w pętli kreatywnej i weryfikacyjnej.
Uzasadnienia	cybernetyczne, systemowe, sieciowe, technologiczne, ustrojowe i społeczno-gospodarcze (rozdział 11)
Status	wydanie pierwsze; propozycje ustrojowe wymagają legalnej ścieżki zmiany prawa i Konstytucji
Data	3 października 2026 r.
Licencja	Creative Commons — Uznanie autorstwa (CC BY 4.0)

Spis treści

0. Streszczenie strategiczne	6
1. Formuła naczelna i jej rozwinięcie	9
1.1. Zapis formuły i pięć rozstrzygnięć	9
1.2. Dlaczego teraz: potrójna konwergencja	10
1.3. Model 3A w bezpieczeństwie AI	11
1.4. Granice formuły: bezpieczeństwo bez państwa nadzoru	12
2. Tezy strategii i ich uzasadnienie	12
2.1. T1 — AI jako warstwa sterownicza	13
2.2. T2 — Własny model rzeczywistości	14
2.3. T3 — Szybkość pętli zamiast wielkości budżetu	14
2.4. T4 — Suwerenność jako koszt wyjścia	14
2.5. T5 — Talenty jako infrastruktura krytyczna	14
2.6. T6 — Energia jako tor zasilający	14
2.7. T7 — Węzły krytyczne i kaskady	15
2.8. T8 — Ekonomia autonomii	15
2.9. T9 — Odporność kognitywna bez propagandy	15
2.10. T10 — Bio-AI jako nowa klasa ryzyka	15
2.11. T11 — Kontrola regulatora	15
2.12. T12 — Pomiar efektu i domknięcie pętli	15
3. Środowisko zagrożeń i diagnoza stanu wyjściowego	16
3.1. Polska jako cel: fakty lat 2025–2026	16
3.2. Taksonomia zagrożeń ery AI	17
3.3. AI jako mnożnik ataku i obrony	19
3.4. Stan wyjściowy: ramy prawne i instytucje	19
3.5. Patologie systemowe: dlaczego dokumenty nie wystarczą	21
4. Analiza geopolityczna	22
4.1. Technologia jako oś potęgi	22
4.2. Położenie Polski: od flanki do węzła	22
4.3. Mapa oddziaływań geopolitycznych w domenę AI	23
4.4. Instrumenty siły DIME+T+L w domenę AI	24
4.5. Warianty strategiczne zależne od otoczenia	25
5. Analiza zasobowa	26
6. Analiza kadrowa: talenty jako infrastruktura krytyczna	28
6.1. Dlaczego talent AI i cyber jest infrastrukturą krytyczną	28
6.2. Bilans kadrowy Polski	29
6.3. Zasoby i przepływy talentu	30
6.4. Narodowy System Talentów AI-Cyber	31
6.5. Progi alarmowe dla infrastruktury talentu	32
7. Analiza kulturowa	33
8. Analiza technologiczna	34
8.1. Stos technologiczny bezpieczeństwa AI	34
8.2. Technologie przetomowe i ich znaczenie	35
9. Analiza strategiczna SWOT i TOWS	37
9.1. Macierz SWOT	37
9.2. Macierz TOWS — strategię wynikowe	38
9.3. Analiza PESTEL w skrócie	39
10. Scenariusze 2026–2035	40
11. Uzasadnienia tezy naczelnej: sześć dróg dowodowych	41
11.1. Uzasadnienie cybernetyczne	41

11.2. Uzasadnienie systemowe	43
11.3. Uzasadnienie sieciowe	45
11.4. Uzasadnienie technologiczne	46
11.5. Uzasadnienie ustrojowe	47
11.6. Uzasadnienie społeczno-gospodarcze	47
11.7. Test sześciu uzasadnień	48
12. Architektura Narodowego Systemu AI Bezpieczeństwa	49
12.1. Osiem warstw systemu	49
12.2. Instytucje systemu	50
12.3. Poziomy autonomii i stopniowany nadzór	51
13. Programy wykonawcze	52
13.1. P1 — Tarcza AI-Cyber dla infrastruktury krytycznej	53
13.2. P2 — Suwerenny Stos AI	53
13.3. P3 — Energia dla AI — AI dla energii	53
13.4. P4 — Narodowy Obraz Sytuacyjny	53
13.5. P5 — Autonomia z kontrolą i counter-UAS	54
13.6. P6 — Odporność Kognitywna	54
13.7. P7 — Bio-AI	54
13.8. P8 — Narodowy System Talentów AI-Cyber.....	54
13.9. P9 — Zamówienia Innowacyjne i Sandbox Obronny	54
13.10. P10 — Kontrola Regulatora	55
14. Mapa drogowa 2026–2035.....	55
15. Monitoring i ewaluacja: Indeks Gotowości AI-Cyber	57
15.1. Architektura pomiaru	57
15.2. KPI i KRI	57
15.3. Mechanizmy kontroli i korekty	58
16. Ryzyka wdrożenia i bezpieczniki	59
17. Spójność z programami TechnoRepubliki	60
18. Hasła programowe	61
19. Konkluzja programowa	62
Aneks A. Słownik pojęć, terminów i skrótów	63
A.1. Pojęcia strategiczne i programowe TechnoSanacji	63
A.2. Pojęcia cybernetyczne, systemowe i sieciowe	64
A.3. Pojęcia z zakresu AI, technologii i cyberbezpieczeństwa.....	65
A.4. Skróty.....	67
Aneks B. Bibliografia i korpus źródłowy	69
B.1. Cybernetyka i teoria systemów	69
B.2. Nauka o sieciach	69
B.3. Sztuczna inteligencja, technologia i bezpieczeństwo	70
B.4. Cyberbezpieczeństwo i odporność — standardy i raporty.....	71
B.5. Talenty, kadry i państwo	72
B.6. Akty prawne i dokumenty urzędowe	72
B.7. Źródła programowe TechnoSanacji (dr Andrzej Węgrzyn)	73
B.8. Destylaty i korpusy badawcze TechnoSanacji	73
B.9. Źródła danych bieżących (dostęp: październik 2026 r.)	74
Aneks C. Nota metodologiczna	75
C.1. Metoda opracowania	75
C.2. Zasady interpretacji	75
C.3. Zastrzeżenia do danych	75
C.4. Status i licencja	75

0. Streszczenie strategiczne

Sztuczna inteligencja zmieniła naturę bezpieczeństwa szybciej niż zmieniły się instytucje, które mają je zapewniać. W ciągu trzech lat AI przeszła od narzędzia, które odpowiada na pytania, do agenta, który samodzielnie planuje i wykonuje zadania, pisze i testuje kod, prowadzi rozpoznawanie, generuje treści perswazyjne na skalę przemysłową i steruje rojami dronów. Równolegle roboty autonomiczne stały się codziennością pola walki w Ukrainie, a biotechnologia — wspierana przez modele uczenia maszynowego — obniżyła próg wejścia zarówno do medycyny przyszłości, jak i do nowych klas ryzyka. Ta potrójna konwergencja nie jest kolejną falą cyfryzacji. Jest zmianą warunków brzegowych suwerenności.

Polska odczuwa tę zmianę bezpośrednio. W 2025 r. CERT Polska zarejestrował 260 783 incydenty — o 152% więcej niż rok wcześniej — przy 658 320 zgłoszeniach, czyli niemal dwóch tysiącach dziennie. Pod koniec grudnia 2025 r. przeprowadzono skoordynowany atak na systemy sterowania dwóch elektrociepłowni oraz system zarządzania farmami fotowoltaicznymi i wiatrowymi; według informacji rządowych w razie powodzenia nawet 500 tys. osób mogło zostać pozbawionych ciepła. We wrześniu 2025 r. polską przestrzeń powietrzną naruszyły rosyjskie drony, co doprowadziło do zestrzeleń nad terytorium państwa NATO z udziałem lotnictwa sojuszniczego i uruchomienia konsultacji z art. 4 Traktatu Północnoatlantyckiego. Ataki na stacje uzdatniania wody, oczyszczalnie i instalacje gazowe przestały być scenariuszem ćwiczeń, a stały się treścią raportów rocznych.

Państwo odpowiedziało aktywnie: nowelizacja ustawy o krajowym systemie cyberbezpieczeństwa wdrażająca dyrektywę NIS2 (podpisana 19 lutego 2026 r.), Strategia Cyberbezpieczeństwa RP na lata 2025–2029 (przyjęta 10 marca 2026 r.), ustawa o systemach sztucznej inteligencji powołująca Komisję Rozwoju i Bezpieczeństwa Sztucznej Inteligencji (KRiBSI), Polityka rozwoju sztucznej inteligencji do 2030 r., resortowa strategia AI Ministerstwa Obrony Narodowej do 2039 r., Centrum Implementacji Sztucznej Inteligencji w Wojskach Obrony Cyberprzestrzeni, fabryki AI w Poznaniu i Krakowie oraz udział w europejskim przetargu na gigafabryki AI. To realny postęp. Nie zmienia jednak tezy niniejszej strategii.

WNIOSEK NACZELNY

Polska nie cierpi na brak dokumentów o AI i cyberbezpieczeństwie. Cierpi na brak regulatora, który łączyłby AI, cyberprzestrzeń, energię, przemysł i kadry w jedną pętlę sterowania z pomiarem, progami alarmowymi i właścicielem korekty. Tempo zmian technologii i zagrożeń przewyższa tempo adaptacji państwa — a żadna instytucja nie ma obowiązku mierzyć, jak szybko ta różnica rośnie.

Strategia przyjmuje perspektywę programową TechnoSanacji: państwo jest regulatorem w sensie cybernetycznym — układem, który utrzymuje swoje zmienne krytyczne w granicach bezpieczeństwa mimo zakłóceń. Z tej perspektywy AI nie jest „technologią do wdrożenia”, lecz warunkiem zachowania wymaganej różnorodności regulatora wobec przeciwnika, który już używa AI do skalowania ataków.

Z prawa wymaganej różnorodności Ashby’ego wynika, że państwo bez AI przegra ze złożonością zagrożeń; z twierdzenia Conanta–Ashby’ego — że państwo bez własnego modelu rzeczywistości (danych, sensorów, obrazu sytuacji) steruje cudzym modelem; z teorii układów samodzielnych Mazura i torów sterowania Kosseckiego — że suwerenność jest pochodną kontroli jednocześnie toru informacyjnego (dane, modele, decyzje) i toru zasilającego (energia, compute, ludzie).

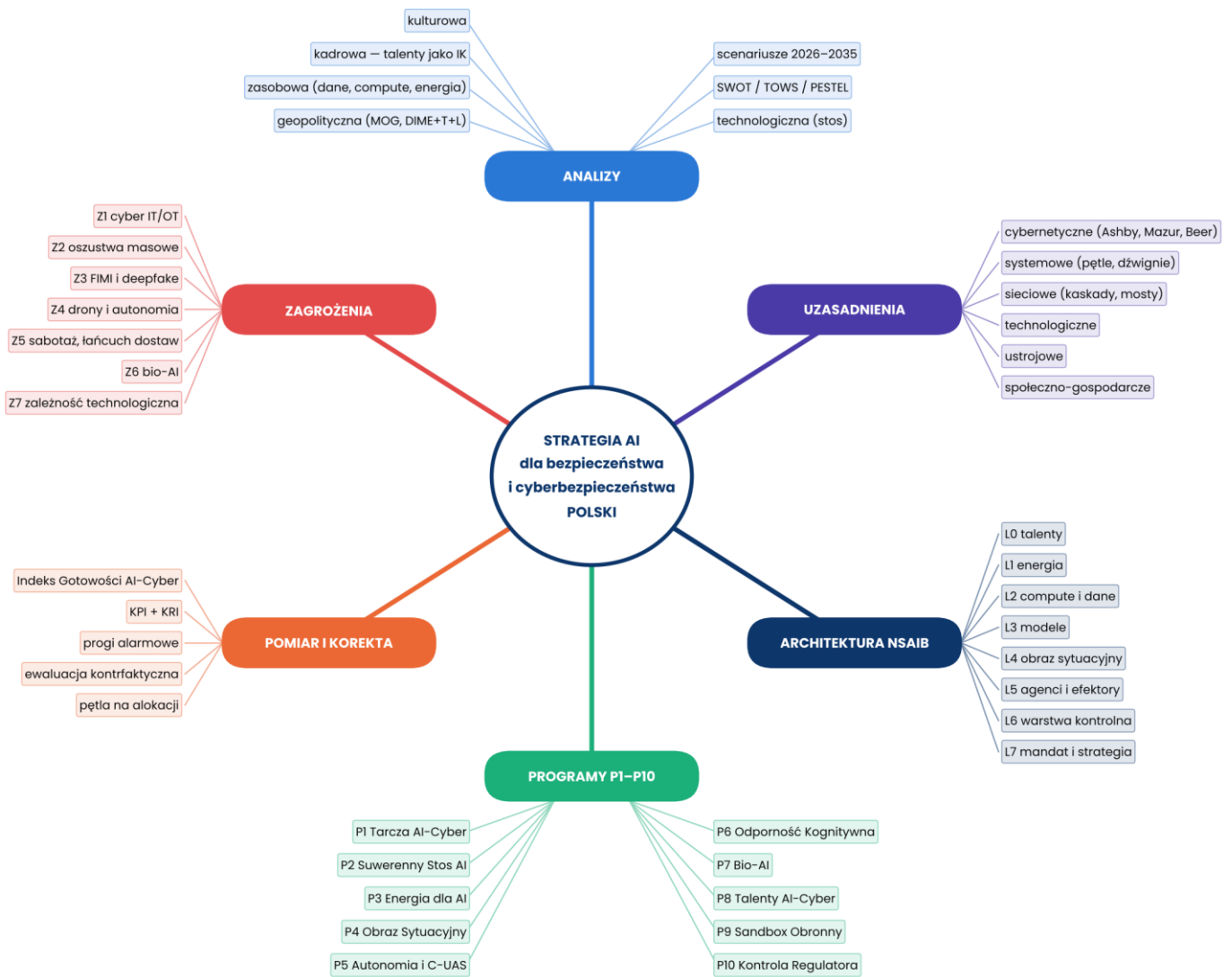
Najważniejszym i najściślej chronionym elementem tego układu są ludzie. Każda infrastruktura krytyczna ma warstwę ludzką: centrum operacji bezpieczeństwa bez analityków, model językowy bez badaczy, sieć elektroenergetyczna bez automatyków znających systemy OT, wojsko bez operatorów systemów bezzałogowych i specjalistów walki radioelektronicznej — to obiekty formalnie istniejące, lecz operacyjnie martwe. Dlatego strategia przenosi politykę talentów AI i cyber z obszaru rynku pracy do obszaru bezpieczeństwa narodowego: talent jest inwentaryzowany, monitorowany, chroniony, utrzymywany w cyrkulacji i odtwarzany jak sieć przesyłowa.

DECYZJA PROGRAMOWA

Rzeczpospolita uznaje sztuczną inteligencję za warstwę sterowniczą bezpieczeństwa narodowego, a kadry AI i cyberbezpieczeństwa — za infrastrukturę krytyczną państwa. Powstaje Narodowy System AI Bezpieczeństwa z jednym właścicielem strategicznym (Rada Bezpieczeństwa AI przy Prezydenckim Centrum Strategicznym), Narodowym Obrazem Sytuacyjnym, rejestrem zdolności krytycznych, progami alarmowymi KPI/KRI i obowiązkiem corocznej korekty. Spadek zdolności poniżej progu jest traktowany jak incydent bezpieczeństwa, a nie jak statystyka.

Strategia formułuje dwanaście tez (rozdział 2), przeprowadza siedem analiz wspierających — zagrożeń, geopolityczną, zasobową, kadrową, kulturową, technologiczną oraz SWOT/TOWS — i dla każdej tezy podaje sześć niezależnych dróg uzasadnienia: cybernetyczną, systemową, sieciową, technologiczną, ustrojową i społeczno-gospodarczą.

Zbieżność tych dróg jest najmocniejszym argumentem za tezą główną TechnoSanacji: w erze AI, robotów autonomicznych i biotechnologii modernizacja państwa przestaje być wyborem politycznym, a staje się warunkiem jego sterowności. Sanacja państwa — rozumiana jako przywrócenie zdolności widzenia, decydowania, wykonania i korekty — jest w tej perspektywie najtańszą formą obrony.



Rysunek 2. Mapa myśli strategii: zagrożenia, analizy, uzasadnienia, architektura i programy wykonawcze.
Opracowanie własne.

Rysunek 2 prezentuje rozbudowaną mapę myśli, która w ujęciu graficznym porządkuje całościową strukturę Strategii AI dla bezpieczeństwa Polski. Centralnym punktem diagramu jest hasło „STRATEGIA AI BEZPIECZEŃSTWA”, od którego odchodzą główne gałęzie tematyczne dokumentu.

Pierwszy obszar obejmuje diagnozę środowiska zagrożeń, wskazując na rosnącą liczbę incydentów cybernetycznych oraz ataki na infrastrukturę krytyczną. Kolejna sekcja przedstawia analizy strategiczne, w tym uwarunkowania geopolityczne, zasobowe, kadrowe, kulturowe i technologiczne.

Diagram wyróżnia także sześć niezależnych dróg uzasadnienia tezy naczelnej, w tym między innymi ścieżkę cybernetyczną, systemową i społeczno-gospodarczą. Na mapie ukazano również architekturę Narodowego Systemu AI Bezpieczeństwa oraz jego ośmiowarstwowy stos technologiczny. Całość schematu wieńczy zestawienie dziesięciu programów wykonawczych, które przekładają założenia strategiczne na konkretne działania wdrożeniowe.

1. Formuła naczelna i jej rozwinięcie

1.1. Zapis formuły i pięć rozstrzygnięć

Formuła naczelna ma trzy warstwy. Warstwa ontologiczna mówi, czym jest bezpieczeństwo w erze AI: nie stanem posiadania zasobów, lecz zdolnością układu do utrzymania celów mimo zakłóceń — mierzoną szybkością i jakością pętli sterowania. Warstwa operacyjna mówi, co państwo ma robić: widzieć, rozumieć, decydować, wykonywać, mierzyć i korygować szybciej niż przeciwnik, opierając się na własnych torach danych, mocy obliczeniowej, energii i talentów. Warstwa normatywna wyznacza granice: maszyna przyspiesza poznanie, lecz nie przejmuje mandatu; państwo broni przestrzeni informacyjnej Narodu, lecz nie staje się aparatem nadzoru.

Z formuły wynika pięć rozstrzygnięć, które odróżniają tę strategię od typowych dokumentów „cyfryzacji bezpieczeństwa”.

Rozstrzygnięcie	Treść	Wyklucza błąd
Rozstrzygnięcie sterownicze	AI jest warstwą sterowniczą bezpieczeństwa (rozpoznanie, fuzja danych, wsparcie decyzji, automatyzacja obrony), a nie projektem informatycznym jednego resortu.	traktowania AI jako kolejnego systemu IT dołączonego do niezmienionego procesu
Rozstrzygnięcie suwerennościowe	Zdolność, która działa tylko za cudzą zgodą — na cudzych danych, cudzym compute, cudzym oprogramowaniu i cudzej energii — jest zdolnością pozorną.	utożsamiania zakupu licencji z posiadaniem zdolności
Rozstrzygnięcie kadrowe	Talenty AI i cyber są infrastrukturą krytyczną: mają rejestr, wskaźnik stanu, plan ciągłości, rezerwę i właściciela.	traktowania niedoboru specjalistów jako problemu rynku pracy
Rozstrzygnięcie energetyczne	Energia jest torem zasilającym AI i cyberbezpieczeństwa; bez odpornej energii nie ma danych, a bez danych nie ma dowodzenia.	planowania centrów danych i SOC bez planu zasilania w kryzysie
Rozstrzygnięcie ustrojowe	Każda zdolność AI państwa ma przypisany mandat, właściciela, ślad decyzji i niezależną pętlę kontroli.	abdykacji odpowiedzialności na algorytm lub dostawcę i dryfu w stronę państwa nadzoru

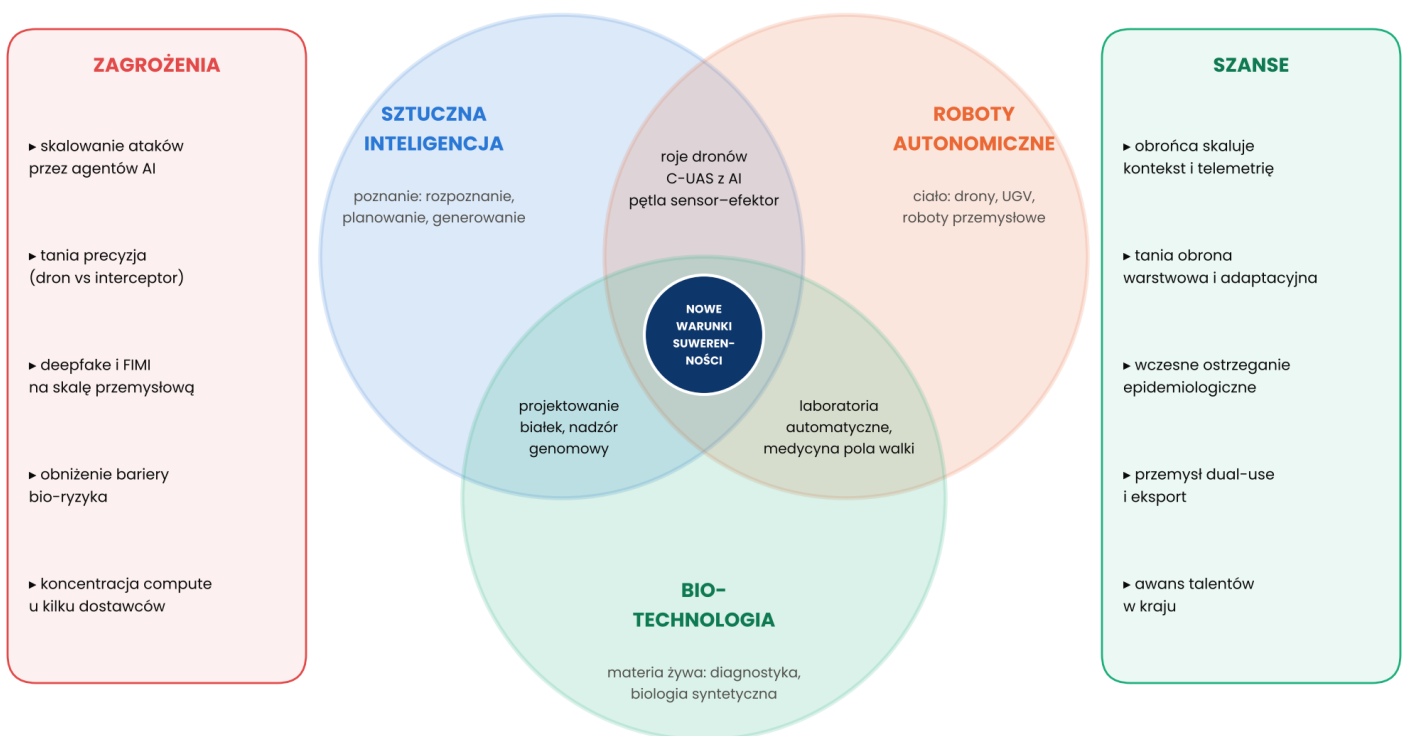
REGUŁA WARUNKÓW ŁĄCZNYCH

Pięć rozstrzygnięć nie jest menu do wyboru. Sterowanie bez suwerenności oddaje pętlę decyzyjną dostawcy. Suwerenność bez talentów buduje martwe obiekty. Talenty bez energii i compute emigrują tam, gdzie mogą pracować. Energia i compute bez kontroli ustrojowej tworzą aparat, który może zostać przechwycony. Kontrola bez sterowania daje bezsilną procedurę.

1.2. Dlaczego teraz: potrójna konwergencja

Trzy technologie, rozwijane dotąd w odrębnych środowiskach, zaczęły się wzajemnie wzmacniać. Sztuczna inteligencja dostarcza zdolności poznawczych: rozpoznania wzorców, planowania, generowania kodu, języka i obrazu. Robotyka i systemy autonomiczne dostarczają ciała: drony powietrzne, nawodne i lądowe, roboty przemysłowe i logistyczne, platformy rozpoznawcze. Biotechnologia dostarcza zdolności przekształcania materii żywej: od diagnostyki i szczepionek po biologię syntetyczną. Każda z tych dziedzin osobno jest dual-use. Ich połączenie tworzy nową jakość zagrożeń i szans, w której czas od pomysłu do skutku skraca się z lat do tygodni.

Czas od pomysłu do skutku skraca się z lat do tygodni



Rysunek 3. Potrójna konwergencja AI, robotów autonomicznych i biotechnologii: obszary nakładania się zagrożeń i szans dla bezpieczeństwa Polski. Opracowanie własne.

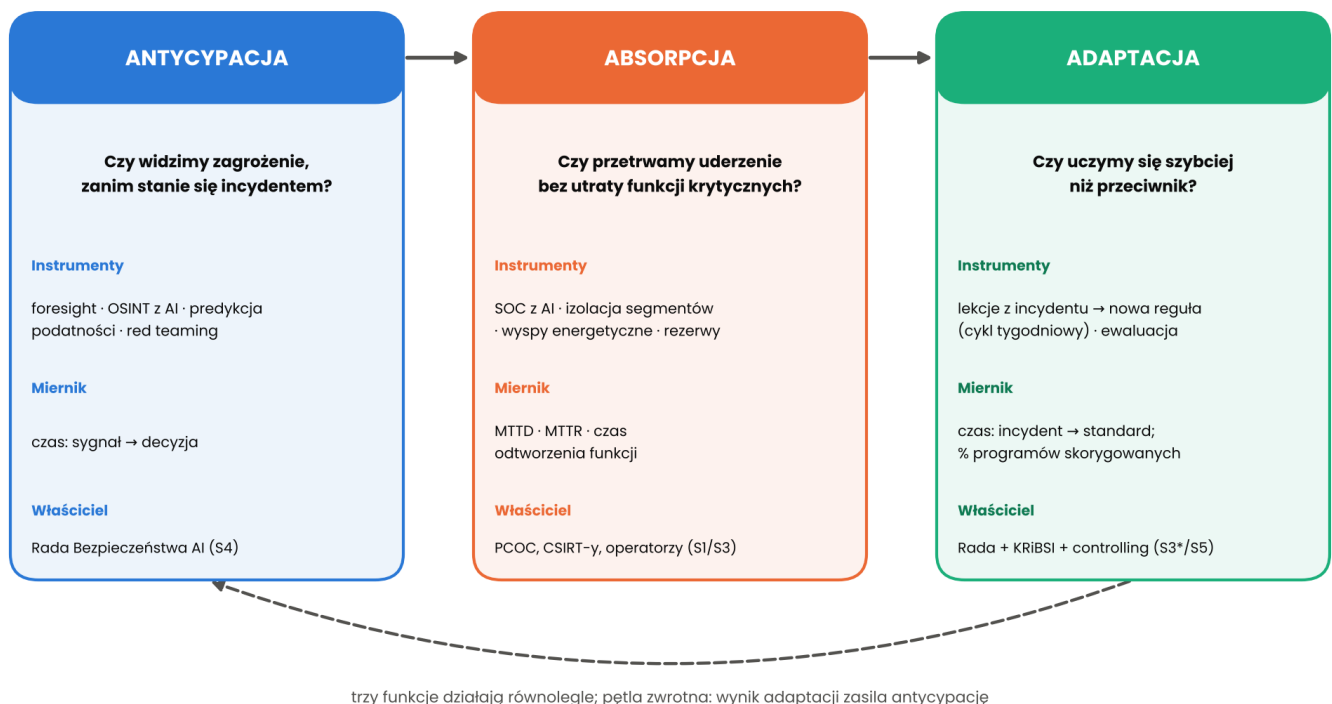
Dla bezpieczeństwa państwa konwergencja ma cztery konsekwencje. Po pierwsze, skaluje atak: jeden operator wyposażony w agentów AI prowadzi rozpoznawanie podatności, phishing i kampanię dezinformacyjną w skali, która wcześniej wymagała zespołu. Po drugie, demokratyzuje precyzję: tanie drony z modułami rozpoznawania obrazu zmieniają ekonomię wojny — dron za kilkadziesiąt tysięcy dolarów wymusza interceptor za miliony. Po trzecie, skraca pętlę decyzyjną: przewagę zdobywa ten, kto domyka cykl obserwacji, orientacji, decyzji i działania szybciej, a różnica minut przekłada się na różnicę między trafieniem a stratą. Po czwarte, przesuwą granicę suwerenności: państwo, które nie kontroluje danych, modeli, mocy obliczeniowej i energii, nie kontroluje własnej pętli decyzyjnej.

1.3. Model 3A w bezpieczeństwie AI

Model 3A TechnoSanacji — antycypacja, absorpcja, adaptacja — stosuje się do bezpieczeństwa AI wprost. Nie są to fazy następujące po sobie, lecz trzy równoległe utrzymywane funkcje, z których każda ma inny przyrząd pomiaru, innego właściciela i inny czas reakcji. Element strategii, który nie przechodzi wszystkich trzech testów, nie jest zdolnością, lecz kosztem.

Funkcja 3A	Pytanie sterujące	Instrumenty AI	Miernik
Antycypacja	Czy widzimy zagrożenie, zanim stanie się incydem?	foresight technologiczny, wywiad z otwartych źródeł wspierany AI, modele predykcji podatności, horizon scanning, red teaming	średni czas od sygnału do decyzji programowej; odsetek zagrożeń przewidzianych w rejestrze ryzyk
Absorpcja	Czy przetrwamy uderzenie bez utraty funkcji krytycznych?	SOC wspierane AI, automatyczna izolacja segmentów sieci, wyspy energetyczne, procedury ręczne, rezerwy kadrowe	średni czas wykrycia (MTTD) i reakcji (MTTR); czas odtworzenia funkcji krytycznej
Adaptacja	Czy uczymy się szybciej niż przeciwnik?	pętla „lekcje z incydentu → zmiana reguły” w cyklu tygodniowym, aktualizacja modeli, przebudowa programów według efektu	czas od incydentu do zmiany standardu; odsetek programów skorygowanych po ewaluacji

Element, który nie przechodzi trzech testów 3A, nie jest zdolnością, lecz kosztem



Rysunek 4. Model 3A w bezpieczeństwie AI: trzy równoległe funkcje sterowania z miernikami i właścicielami. Opracowanie własne.

1.4. Granice formuły: bezpieczeństwo bez państwa nadzoru

Strategia AI bezpieczeństwa niesie ryzyko, którego nie wolno przemilczeć: te same narzędzia, które pozwalają wykrywać wrogie operacje, pozwalają nadzorować obywateli. Cybernetyka drugiego rzędu przypomina, że regulator jest częścią systemu, który reguluje — ma własne interesy i skłonność do reprodukcji siebie. Dlatego formuła naczelna zawiera warunek kontroli regulatora przez Naród jako człon iloczynu, a nie przypis.

ZASADA GRANICZNA

Państwo broni toru poznawczego Narodu przed wrogimi operacjami wpływu, lecz nie sięga po manipulację wobec własnych obywateli. AI bezpieczeństwa wykrywa zagrożenia, a nie poglądy. Każdy system AI o charakterze władczym ma podstawę prawną, rejestr, ocenę skutków dla praw podstawowych, ślad decyzji, prawo do zakwestionowania i niezależnego audytora. Granica między obroną informacyjną a propagandą państwową jest granicą między TechnoRepubliką a państwem nadzoru.

2. Tezy strategii i ich uzasadnienie

Tezy sformułowano na podstawie destylatów korpusu TechnoSanacji (obrona, geopolityka, odporność, cybernetyka państwa, analiza systemowa i sieciowa) oraz aktualnych danych o stanie bezpieczeństwa cyfrowego Polski. Każda teza ma wskazaną podstawę i konsekwencję programową; uzasadnienia rozwinięto w rozdziale 11.

Nr	Teza	Podstawa	Konsekwencja
T1	AI jest warstwą sterowniczą bezpieczeństwa narodowego, a nie projektem IT.	prawo wymaganej różnorodności (Ashby); tezy Karpa-Zamiski o software jako twardej sile; doświadczenia wojny w Ukrainie	Narodowy System AI Bezpieczeństwa z jednym właścicielem strategicznym
T2	Dobry regulator musi być modelem systemu: bez własnego obrazu sytuacji państwo steruje cudzym modelem.	twierdzenie Conanta-Ashby'ego; destylat CybControl	Narodowy Obraz Sytuacyjny i suwerenna fuzja danych
T3	O wyniku decyduje szybkość pętli 3A, nie wielkość budżetu.	pętla OODA (Boyd); czas decyzyjny (Kossecki); destylat Obrona	mierniki MTDD/MTTR i czasu decyzji jako KPI strategiczne
T4	Suwerenność cyfrowa to koszt wyjścia i czas migracji, a nie autarkia.	Diagnoza Rzeczypospolitej 2026; destylat SYS	rejestr zależności krytycznych, plany wyjścia, zamówienia z kryterium przenaszalności

Nr	Teza	Podstawa	Konsekwencja
T5	Talenty AI i cyber są infrastrukturą krytyczną.	publikacja „Talenty – infrastruktura krytyczna państwa”; dyrektywy CER i NIS2	Rejestr Zdolności Krytycznych AI-Cyber i Cyber-Rezerwa Narodowa
T6	Energia jest torem zasilającym AI i cyberbezpieczeństwa.	tor zasilający (Kossecki); analiza AI Data Center łkowa; atak na energetykę XII 2025	program „Energia dla AI — AI dla energii” i wyspy zasilania dla węzłów krytycznych
T7	Ryzyko systemowe siedzi w węzłach o wysokiej centralności przepływu i w sprzężeniach międzywarstwowych.	perkolacja w sieciach współzależnych; destylaty SYS i GR	mapa 20–50 węzłów krytycznych z planem ostony
T8	Tania autonomia wymaga taniej obrony: cost-exchange decyduje o wyniku.	destylat Obrona (cost-exchange, magazine depth)	program autonomii z kontrolą i counter-UAS o koszcie porównywalnym z zagrożeniem
T9	Odporność kognitywna jest infrastrukturą krytyczną, ale budowaną wzmacnianiem odbiorcy, nie kontrolą treści.	teoria inokulacji; destylaty RES i CybControl	Program Odporności Kognitywnej z bezpiecznikami demokratycznymi
T10	Konwergencja AI–bio tworzy nową klasę ryzyka wymagającą nadzoru epidemiologicznego wspieranego AI.	podejście One Health; raporty o bezpieczeństwie AI	Program Bio-AI: wczesne ostrzeganie i bezpieczeństwo badań
T11	AI w rękach państwa wymaga kontroli drugiego rzędu: rejestru, audytu i zaskarżalności.	cybernetyka drugiego rzędu; AI Act; ustawa o systemach AI	warstwa kontrolna NSAIB, KRiBSI, red teaming instytucji
T12	Strategia bez pomiaru i prognozy alarmowego jest deklaracją.	standard KPI/KRI TechnoSanacji; prawo Goodharta	Indeks Gotowości AI-Cyber publikowany corocznie i domknięty na alokacji

2.1. T1 — AI jako warstwa sterownicza

Prawo wymaganej różnorodności mówi, że regulator opanuje tylko tyle klas zaburzeń, ile ma stanów odpowiedzi. Przeciwnik używający agentów AI generuje różnorodność ataku - wariantów phishingu, mutacji złośliwego oprogramowania, narracji dezinformacyjnych, trajektorii dronów — w tempie, którego nie pochłonie żadna kadra ludzka. Jedynym sposobem zrównania różnorodności jest automatyzacja po stronie obrony: AI do triażu alertów, korelacji zdarzeń, wykrywania anomalii w sieciach przemysłowych, analizy obrazu i klasyfikacji treści. AI nie jest więc dodatkiem do bezpieczeństwa, lecz warunkiem zachowania sterowności.

2.2. T2 — Własny model rzeczywistości

Twierdzenie Conanta–Ashby’ego głosi, że każdy dobry regulator systemu musi być modelem tego systemu. Państwo, które nie ma aktualnego obrazu stanu własnej infrastruktury krytycznej, łańcuchów dostaw, przestrzeni informacyjnej i pola walki, nie steruje — reaguje na doniesienia. Model zbudowany na cudzych sensorach, cudzej chmurze i cudzych modelach jest modelem, którego właścicielem jest ktoś inny. Stąd Narodowy Obraz Sytuacyjny jako wspólna, wielowarstwowa reprezentacja stanu państwa, zasilana danymi operatorów usług kluczowych, CSIRT-ów, wojska, służb, samorządów i obywateli.

2.3. T3 — Szybkość pętli zamiast wielkości budżetu

Nie wygrywa państwo, które wydaje najwięcej, lecz państwo, które najlepiej zamienia budżet w krótką pętlę decyzyjną, odporność i efekt. Polska wydaje na obronę rekordowe ok. 4,8% PKB, ale miarą strategii jest czas: od wykrycia anomalii do izolacji segmentu sieci, od zauważenia drona do jego neutralizacji, od incydentu do zmiany standardu. Mierniki czasu muszą stać się wskaźnikami strategicznymi, raportowanymi tak samo jak wydatki.

2.4. T4 — Suwerenność jako koszt wyjścia

Państwo nie musi produkować wszystkiego; musi umieć zmienić dostawcę bez utraty ciągłości działania. Zależność ekonomiczna to nie to samo co podatność polityczna: instytucja, której usługa może zostać zawieszona decyzją obcego państwa lub korporacji, traci zdolność działania. Regulacja zmienia to, co dostawca może zrobić z danymi; nie zmienia tego, czy instytucja może działać bez niego. Dlatego miarą suwerenności jest koszt wyjścia i czas migracji dla każdej warstwy stosu technologicznego.

2.5. T5 — Talenty jako infrastruktura krytyczna

Polska ma utalentowanych specjalistów, ale nie ma regulatora talentów. Nie mierzy zintegrowanie, ilu analityków SOC, inżynierów bezpieczeństwa OT, badaczy AI, kryptologów i operatorów systemów bezzatłogowych ma w dziedzinach krytycznych, z jaką rotacją, zastępowalnością i odpływem. Odejście jednego zespołu, który jako jedyny w kraju utrzymuje daną zdolność, jest ryzykiem bezpieczeństwa państwa, a nie sprawą wewnętrzną instytucji. Teza ta jest rozwinięta w rozdziale 6.

2.6. T6 — Energia jako tor zasilający

Bez danych nie ma dowodzenia; bez energii nie ma danych. Centra danych AI zużywają rosnącą część energii elektrycznej, a ich niezawodność zależy od sieci, która sama jest celem ataków. Atak z grudnia 2025 r. na systemy sterowania ciepłowni i farm OZE pokazał, że cyberprzestrzeń i energetyka tworzą jeden układ ściśle sprzężony. Odpowiedzią jest planowanie infrastruktury AI razem z energetyką: lokalne bilansowanie, magazyny, wodór, wyspy zasilania dla węzłów krytycznych i AI zastosowana do sterowania samą siecią.

2.7. T7 — Węzły krytyczne i kaskady

Infrastruktura państwa jest siecią sieci: energetyka, telekomunikacja, dane, finanse, transport, woda, administracja i obrona są wzajemnie zależne. Badania perkolacji w sieciach współzależnych pokazują, że takie układy załamują się skokowo, a nie łagodnie. Węzeł krytyczny nie musi być węzłem dużym — decyduje pozycja w przepływie. Strategia wymaga więc mapy 20–50 węzłów, których utrata kaskaduje na co najmniej dwie warstwy, i to one, a nie „ważne instytucje”, są przedmiotem priorytetowej osłony.

2.8. T8 — Ekonomia autonomii

Każdy tani, masowy środek ataku musi być zwalczany środkiem o porównywalnym lub niższym koszcie. Interceptor za miliony wobec drona za dziesiątki tysięcy to strategia przegranej na wyczerpanie. AI umożliwia taną obronę: drony przechwytyjące z autonomicznym naprowadzaniem, fuzję danych z tanich sensorów akustycznych i radiowych, automatyczną klasyfikację celów. Warunkiem jest krajowa produkcja i cykl adaptacji liczony w tygodniach, wzorowany na doświadczeniach ukraińskich.

2.9. T9 — Odporność kognitywna bez propagandy

Generatywna AI przeniosła wojnę informacyjną na poziom maszynowy: automatyczne operacje wpływu, deepfake’i, sieci nieautentycznych kont. Obrona kognitywna staje się infrastrukturą krytyczną, ale jej skuteczność zależy od zaufania. Odporność buduje się prebunkingiem, edukacją medialną, przejrzystością i szybką, wiarygodną komunikacją kryzysową — wzmacnianiem odbiorcy, a nie kontrolą treści.

2.10. T10 — Bio-AI jako nowa klasa ryzyka

Modele uczenia maszynowego przyspieszają projektowanie białek, analizę genomów i automatyzację laboratoriów. Te same zdolności, które skracają drogę do nowych leków, mogą obniżać barierę dla nadużyć. Strategia nie zajmuje się techniką — zajmuje się zdolnością państwa: wczesnym ostrzeganiem epidemiologicznym wspieranym AI (nadzór ścieków, syndromiczny, genomowy), bezpieczeństwem badań, kontrolą dostępu do infrastruktury biotechnologicznej i odpornością systemu ochrony zdrowia.

2.11. T11 — Kontrola regulatora

Państwo zdolne bez kontroli regulatora produkuje państwo nadzoru; kontrola bez zdolności produkuje bezsilną procedurę. Ustrój musi domknąć obie pętle jednocześnie: rejestr algorytmów publicznych, ocena skutków dla praw podstawowych, ślad decyzji, prawo do zaskarżenia, niezależny audyt, rotacja elit i referendum jako bezpiecznik ostateczny.

2.12. T12 — Pomiar efektu i domknięcie pętli

Bez pomiaru efektu każdego programu nie wiadomo, co działa. Każdemu wskaźnikowi wyniku (KPI) towarzyszy wskaźnik ryzyka (KRI) wykrywający jego obchodzenie, a wynik ewaluacji wraca do decyzji o środkach w kolejnym cyklu. Miernik bez progu i właściciela jest informacją opisową, a nie instrumentem sterowania.

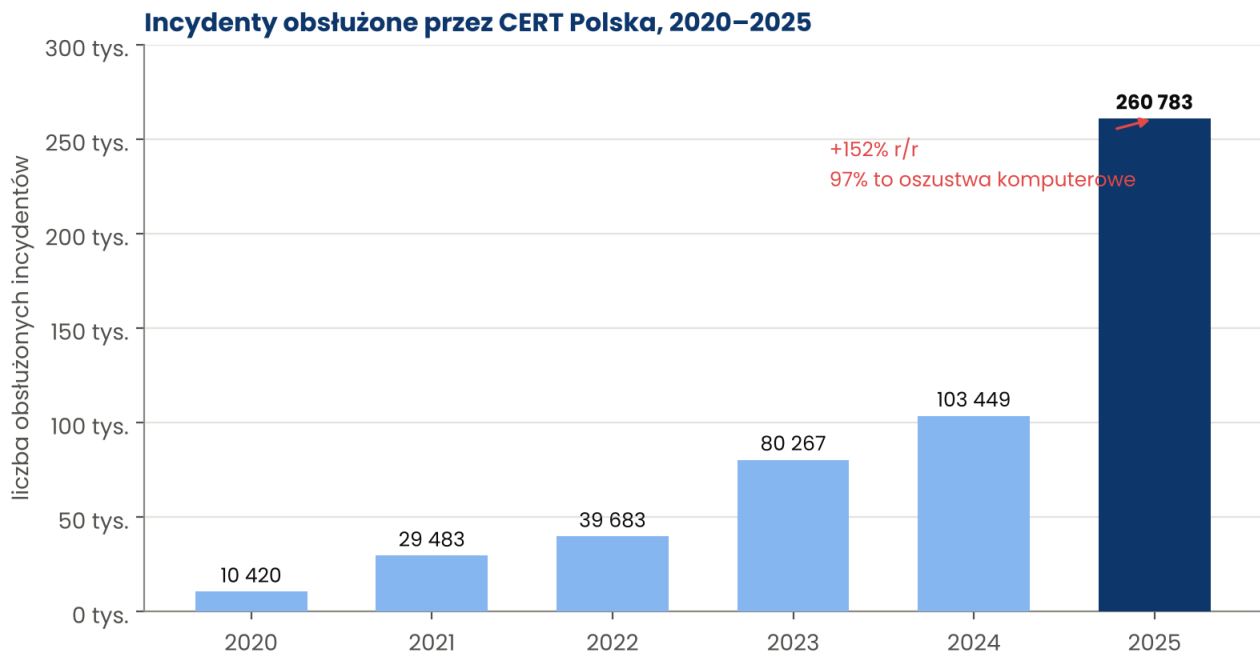
3. Środowisko zagrożeń i diagnoza stanu wyjściowego

3.1. Polska jako cel: fakty lat 2025–2026

Polska jest państwem frontowym nie tylko w sensie militarnym. Jest jednym z głównych celów rosyjskich operacji hybrydowych w Europie, a jednocześnie dużym, szybko cyfrującym się rynkiem z rozbudowanym sektorem usług informatycznych i niską — jak na skalę gospodarki — zdolnością absorpcji zaawansowanych technologii. Ten układ przyciąga zarówno przestępców, jak i państwowych przeciwników.

Zdarzenie lub wskaźnik	Fakty	Znaczenie strategiczne
Incydenty obsłużone przez CERT Polska w 2025 r.	260 783 incydenty (+152% r/r), 658 320 zgłoszeń; 97% incydentów to oszustwa komputerowe, ok. 78 tys. przypadków phishingu; 179 ataków ransomware	skala przekroczyła możliwości obsługi ręcznej — automatyzacja triażu przestaje być opcją
Atak na energetykę (koniec grudnia 2025 r.)	skoordynowane działania wobec dwóch elektrociepłowni i systemu zarządzania farmami PV i wiatrowymi; potencjalnie do 500 tys. osób bez ciepła	dowód ścisłego sprzężenia cyberprzestrzeni i energetyki; podatność rozproszonej generacji OZE
Ataki na wodociągi i instalacje gazowe (2025)	incydenty w stacjach uzdatniania wody, oczyszczalniach i stacji regazyfikacji LNG; w jednym przypadku ok. 6 godzin przerwy w dostawie wody dla ok. 2,5 tys. mieszkańców	małe podmioty komunalne są najłabszym ogniwem infrastruktury krytycznej
Główne wektory ataku	wykradzione poświadczenia VPN, usługi RDP dostępne z internetu, podatności urządzeń brzegowych	dominują błędy podstawowej higieny — dźwignia leży w skali i automatyzacji, nie w egzotycznych technologiach
Naruszenie przestrzeni powietrznej przez drony (9–10 września 2025 r.)	kilkanaście rosyjskich dronów nad Polską, zestrzelenia części obiektów przez lotnictwo polskie i sojusznicze, konsultacje z art. 4 NATO, operacja „Eastern Sentry”	test ekonomii obrony powietrznej: tanie drony wobec drogich środków przechwytywania
Dywersja i sabotaż infrastruktury (2024–2025)	akty sabotażu na kolei i w logistyce, podpalenia, werbunek wykonawców przez komunikatory	przeciwnik łączy domenę cyfrową, informacyjną i kinetyczną w jedną operację

Źródła: raport roczny CERT Polska za 2025 r.; komunikaty Rady Ministrów i Ministerstwa Cyfryzacji; komunikaty Dowództwa Operacyjnego Rodzajów Sił Zbrojnych; zestawienie w Aneksie B.



Źródło: raporty roczne CERT Polska / NASK. Wzrost odzwierciedla także poprawę wykrywania i zgłaszania.

Rysunek 5. Liczba incydentów obsługiwanych przez CERT Polska w latach 2020–2025. Źródło: raporty roczne CERT Polska (NASK).

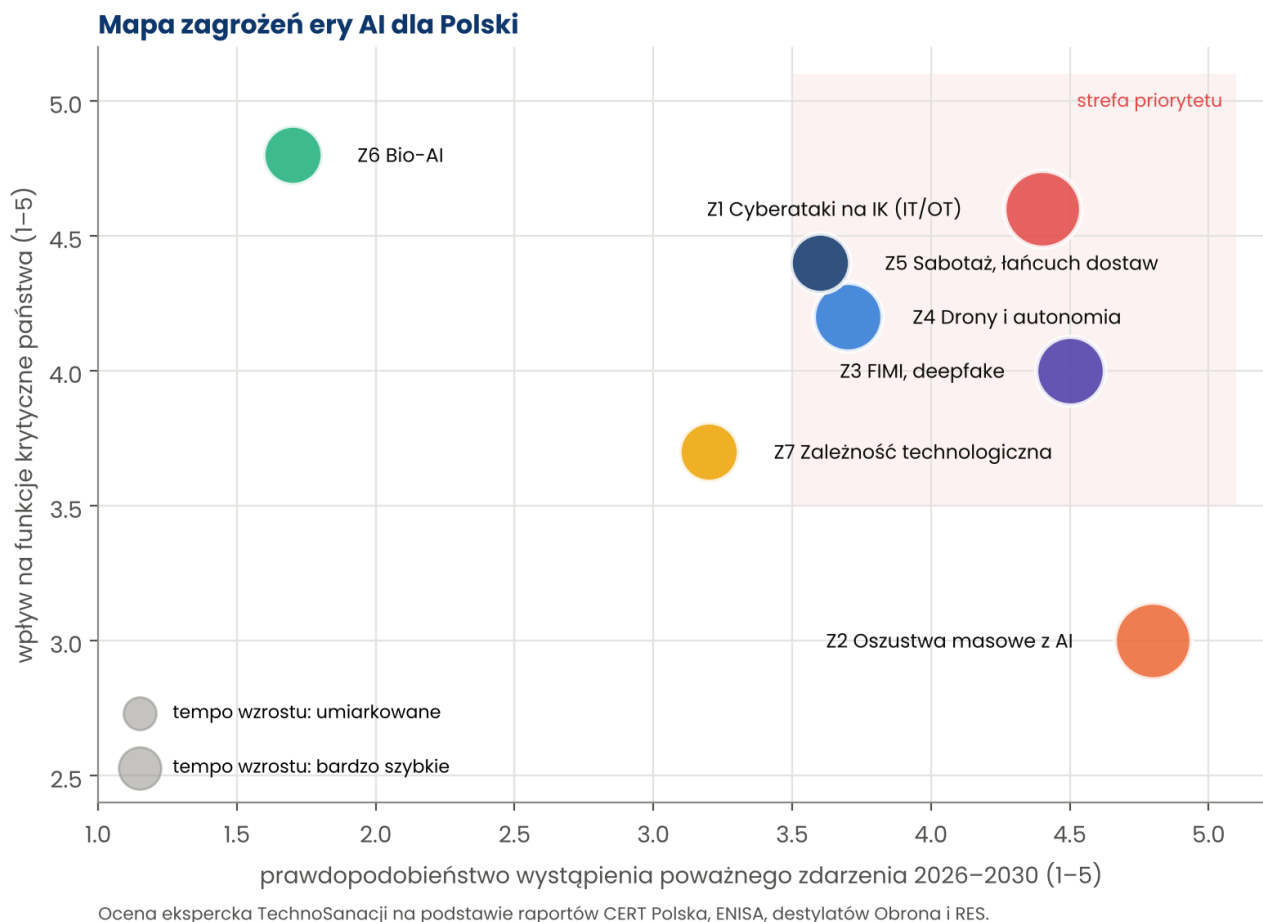
Krzywa incydentów ma kształt wykładniczy. Między 2020 a 2025 r. liczba obsługiwanych incydentów wzrosła ponad dwudziestokrotnie. Część wzrostu wynika z lepszego wykrywania i rosnącej świadomości zgłaszających, ale kierunek jest jednoznaczny: wolumen zagrożeń rośnie szybciej niż kadra zdolna je obsłużyć. To jest liczbowy obraz „dryfu strategicznego” opisanego w Diagnostyce Rzeczypospolitej 2026 — tempo zmian otoczenia przewyższa tempo adaptacji państwa.

3.2. Taksonomia zagrożeń ery AI

Zagrożenia nie występują osobno. Przeciwnik łączy je w kampanie, w których cyberatak przygotowuje sabotaż fizyczny, a operacja informacyjna wzmacnia jego efekt psychologiczny. Strategia porządkuje je w siedem klas, z których każda ma własny wskaźnik wczesnego ostrzegania (KRI).

Klasa zagrożenia	Rola AI po stronie atakującego	Przykładowy KRI	Właściciel ryzyka
Z1. Cyberataki na infrastrukturę krytyczną (IT i OT)	automatyczne rozpoznanie podatności, generowanie exploitów, adaptacyjne złośliwe oprogramowanie	odsetek operatorów usług kluczowych z podatnościami krytycznymi niezatartymi > 30 dni	Minister Cyfryzacji, CSIRT sektorowe
Z2. Przestępczość i oszustwa masowe	phishing generowany przez modele językowe, klonowanie głosu, fałszywe sklepy i inwestycje	wolumen oszustw na 100 tys. mieszkańców; czas blokady domeny	CERT Polska, Policja, sektor finansowy

Klasa zagrożenia	Rola AI po stronie atakującego	Przykładowy KRI	Właściciel ryzyka
Z3. Operacje wpływu i wojna kognitywna (FIMI)	farmy botów, deepfake'i, mikrotargetowanie narracji, „fabryki propagandy”	zasięg skoordynowanych kampanii nieautentycznych; spadek zaufania do instytucji	Rada Bezpieczeństwa AI, media publiczne, NASK
Z4. Systemy autonomiczne i drony	nawigacja i naprowadzanie bez łączności, roje, rozpoznanie obrazu	stosunek kosztu środka obrony do kosztu celu; czas wykrycia	MON, DORSZ
Z5. Sabotaż hybrydowy i ataki na łańcuch dostaw	analiza sieci logistycznych, werbunek przez platformy, ataki na dostawców oprogramowania	udział dostawców wysokiego ryzyka w systemach kluczowych	ABW, Kolegium ds. Cyberbezpieczeństwa
Z6. Bio-AI i zagrożenia biologiczne	obniżenie bariery wiedzy, automatyzacja laboratoriów	sygnały anomalii w nadzorze ścieków i syndromicznym	Minister Zdrowia, GIS
Z7. Zależność technologiczna i przymus gospodarczy	koncentracja compute, modeli i chmury u kilku dostawców pozaeuropejskich	koszt wyjścia i czas migracji dla warstw krytycznych	Rada Bezpieczeństwa AI, Minister Cyfryzacji



Rysunek 6. Mapa zagrożeń ery AI dla Polski: prawdopodobieństwo w horyzoncie 2026–2030 × wpływ na funkcje państwa; wielkość punktu oznacza tempo wzrostu. Ocena ekspercka TechnoSanacji.

3.3. AI jako mnożnik ataku i obrony

Sztuczna inteligencja jest technologią podwójnego zastosowania w najczystszej postaci. Te same modele, które wykrywają anomalie w ruchu sieciowym, pozwalają je ukrywać; te same narzędzia, które generują raporty analityczne, generują dezinformację. Przewaga nie wynika więc z samego dostępu do AI, który staje się powszechny, lecz z trzech czynników: jakości danych własnych, szybkości integracji AI z procesami operacyjnymi i jakości ludzi, którzy te procesy projektują i nadzorują.

W krótkim okresie AI daje przewagę atakującemu, ponieważ obniża koszt prób, a obrońca musi mieć rację za każdym razem. W średnim okresie przewaga może się odwrócić: obrońca ma uprzywilejowany dostęp do własnych danych telemetrycznych, architektury sieci i kontekstu, a AI pozwala ten kontekst wykorzystać w skali. Warunkiem jest jednak to, by obrońca te dane zbierał, łączył i udostępniał modelom — czego dziś w Polsce w większości sektorów nie robi.

ASYMETRIA DO ODWRÓCENIA

Atakujący wykorzystuje AI do skalowania prób; obrońca może wykorzystać AI do skalowania kontekstu. Państwo, które połączy telemetrię operatorów, CSIRT-ów, wojska i samorządów w jeden obraz sytuacyjny, zamienia własną złożoność z obciążenia w przewagę informacyjną.

3.4. Stan wyjściowy: ramy prawne i instytucje

Lata 2024–2026 przyniosły największą od dekady przebudowę ram prawnych i instytucjonalnych bezpieczeństwa cyfrowego. Strategia nie zaczyna od zera — zaczyna od zestawu narzędzi, które trzeba połączyć w system.

Dołączona tabela przedstawia szczegółowe zestawienie kluczowych instrumentów prawnych, strategicznych oraz infrastrukturalnych wraz z ich stanem na październik 2026 roku i oceną z perspektywy przyjętej strategii.

Analiza obejmuje m.in. nowelizację ustawy o krajowym systemie cyberbezpieczeństwa (NIS2), Strategię Cyberbezpieczeństwa RP 2025–2029 oraz polską Ustawę o systemach sztucznej inteligencji. Wśród regulacji unijnych uwzględniono unijny AI Act po pakiecie „Digital Omnibus”, wskazując przesunięte terminy obowiązków dla systemów wysokiego ryzyka oraz fakt, że systemy wojskowe pozostają poza jego zakresem.

Dokumenty programowe, takie jak Polityka rozwoju sztucznej inteligencji w Polsce do 2030 r. i Resortowa strategia AI MON do 2039 r., zostały ocenione pod kątem obecności twardych komponentów bezpieczeństwa oraz spójności z sektorem cywilnym. Ostatnie wiersze tabeli poświęcono zapleczu technologicznemu, w tym infrastrukturze obliczeniowej (np. PLGrid, fabryki AI w Poznaniu i Krakowie) oraz krajowym modelom językowym takim jak PLLuM i Bielik.

Instrument	Stan na październik 2026 r.	Ocena z perspektywy strategii
Nowelizacja ustawy o krajowym systemie cyberbezpieczeństwa (NIS2)	podpisana 19 lutego 2026 r.; podmioty kluczowe i ważne — zakres rozszerzony z ok. 400 do ponad 10 tys. organizacji w 18 sektorach; procedura dostawcy wysokiego ryzyka (7 lat na wycofanie); skierowana do TK w trybie kontroli następczej	fundament prawny jest; luka leży w zdolności nadzoru, kadrach i wsparciu małych podmiotów
Strategia Cyberbezpieczeństwa RP 2025–2029	przyjęta 10 marca 2026 r.; filary: suwerenność technologiczna, walka z przestępczością, jednolite działanie (Połączone Centrum Operacyjne Cyberbezpieczeństwa — PCOC, CSIRT sektorowe), wsparcie regionów (Cyberbezpieczny Samorząd, Cyberbezpieczne Wodociągi, ochrona anty-DDoS)	dobra diagnoza; brak spięcia z AI, energetyką i polityką talentów w jedną pętlę
Ustawa o systemach sztucznej inteligencji	uchwalona 3 lipca 2026 r.; KRiBSI jako krajowy organ nadzoru; opinie indywidualne i piaskownice regulacyjne; przepisy o karach od 28 października 2026 r.	szansa na regulatora AI; ryzyko zawężenia roli do nadzoru zgodności zamiast budowy zdolności
AI Act (rozporządzenie UE 2024/1689) po pakiecie „Digital Omnibus”	obowiązki dla systemów wysokiego ryzyka z Załącznika III przesunięte na 2 grudnia 2027 r., z Załącznika I na 2 sierpnia 2028 r.; obowiązki przejrzystości bez zmian	czas na przygotowanie administracji; systemy wojskowe poza zakresem AI Act wymagają własnych reguł
Polityka rozwoju sztucznej inteligencji w Polsce do 2030 r.	zaktualizowana w listopadzie 2025 r.; filary „Sprawne państwo”, „Serce kontynentu AI”, AI zorientowana na człowieka	brak twardego komponentu bezpieczeństwa i obronności; brak powiązania z rejestrem kadr krytycznych
Resortowa strategia AI MON do 2039 r. i Centrum Implementacji Sztucznej Inteligencji	strategia przedstawiona w 2024 r.; CISI w strukturze Dowództwa Komponentu Wojsk Obrony Cyberprzestrzeni	właściwy kierunek; potrzebne sprzężenie z sektorem cywilnym i przemysłem dual-use
Infrastruktura obliczeniowa	PLGrid, fabryki AI w Poznaniu (PCSS) i Krakowie (Cyfronet); udział w przetargu EuroHPC na gigafabryki AI (termin ofert: 12 listopada 2026 r.), konsorcjum z Chorwacją, Węgrami, Litwą i Słowacją	szansa na skok zdolności; ryzyko, że compute powstanie bez energii, kadr i zastosowań bezpieczeństwa
Krajowe modele językowe	PLLuM (konsorcjum naukowe, wdrożenia w administracji), Bielik (inicjatywa otwarta)	załączek suwerennego stosu; potrzebne wersje do zastosowań bezpieczeństwa z audytem i testami odporności

3.5. Patologie systemowe: dlaczego dokumenty nie wystarczą

Diagnoza Rzeczypospolitej 2026 opisała mechanizm, który w domenie AI i cyberbezpieczeństwa działa ze szczególną siłą: autopojezę instytucjonalną — systemy utrzymują własne procedury, budżety i kadry nawet wtedy, gdy wynik publiczny pozostaje słaby. W bezpieczeństwie cyfrowym przyjmuje ona pięć postaci.

1. **Silosowość.** Cyberbezpieczeństwo cywilne, wojskowe, służby specjalne, energetyka, ochrona zdrowia i samorząd mają odrębne systemy, dane, budżety i kultury. Przeciwnik nie zna podziału kompetencji; atakuje tam, gdzie styk jest najślabszy.
2. **Brak regulatora.** Wiele instytucji odpowiada za fragmenty, żadna za wynik całości. Nie istnieje organ zobowiązany do mierzenia luki między tempem zagrożeń a tempem adaptacji.
3. **Zgodność zamiast zdolności.** Nowe obowiązki NIS2 i AI Act mogą zostać zrealizowane jako dokumentacja, a nie jako realna zdolność wykrywania i reakcji. Audyt zgodności nie mierzy czasu reakcji.
4. **Zależność od dostawców (vendor lock-in).** Zależność licencyjna administracji od oprogramowania własnościowego jest niemal całkowita w systemach operacyjnych i pakietach biurowych; chmura i modele AI pochodzą głównie spoza Unii. Brakuje planów wyjścia.
5. **Kadry bez rejestru.** Państwo wie, ile wydaje na programy kadrowe, ale nie wie, ile ma zdolności w dziedzinach krytycznych. Odptyw specjalistów do sektora prywatnego i za granicę nie jest monitorowany jako ryzyko bezpieczeństwa.

OTWARTA PĘTLA

Instytucja bez sprzężenia zwrotnego to fasada, nie regulator. Dopóki wynik - czas wykrycia, czas reakcji, liczba skutecznych ataków, stan kadr - nie wraca do decyzji o zasobach, kolejne strategie będą zwiększały liczbę dokumentów, a nie zdolność państwa.

4. Analiza geopolityczna

4.1. Technologia jako oś potęgi

Świat wszedł w erę wielobiegunowej rywalizacji mocarstw, w której sztuczna inteligencja, moc obliczeniowa, półprzewodniki, kosmos i cyberprzestrzeń wyznaczają hierarchię. Stany Zjednoczone koncentrują większość światowej mocy obliczeniowej dla AI i kontrolują kluczowe ogniwa łańcucha półprzewodników; Chiny budują równoległy stos technologiczny i stosują kontrolę eksportu surowców krytycznych; Unia Europejska próbuje łączyć regulację z inwestycjami (AI Act, fabryki i gigafabryki AI, InvestAI), lecz pozostaje zależna od zewnętrznych dostawców chipów, chmury i modeli granicznych. Rosja nie jest liderem technologicznym, ale jest liderem w stosowaniu technologii w operacjach hybrydowych i informacyjnych.

Z perspektywy Teorii Oddziaływań Geopolitycznych stosowanej w Narodowej Strategii Bezpieczeństwa ryzyko jest iloczynem mocy przeciwnika, jego wrogości i własnej podatności. AI zwiększa moc przeciwnika (skalowanie operacji), a zależności technologiczne zwiększają podatność. Jedynym członem, na który Polska ma bezpośredni wpływ, jest podatność — stąd priorytet redukcji zależności jednostronnych w torach sterowania.

4.2. Położenie Polski: od flanki do węzła

Polska leży na styku głównych osi strategicznych kontynentu: Bałtyku, wschodniej flanki NATO, przesmyku suwalskiego, korytarzy północ-południe i wschód-zachód. W domenie cyfrowej położenie to ma trzy wymiary. Po pierwsze, Polska jest zapleczem logistycznym i informacyjnym dla Ukrainy, a zatem celem rozpoznania i sabotażu. Po drugie, przez Polskę i Bałtyk biegą kable telekomunikacyjne i energetyczne, gazociągi i połączenia z offshore, wrażliwe na działania podmorskie. Po trzecie, Polska może stać się węzłem infrastruktury obliczeniowej Europy Środkowo-Wschodniej — jeżeli połączy tanią, czystą i odporną energię z kadrami i stabilnym prawem.

Rysunek 7 ilustruje geopolityczne i strategiczne położenie Polski, wykazując jej dwoistą rolę jako państwa frontowego oraz potencjalnego węzła cyfrowo-energetycznego Europy Środkowej. Od strony wschodniej i północno-wschodniej wykres zaznacza teatry presji hybrydowej ze strony Rosji i Białorusi, w tym zagrożenia dronowe, sabotażowe, migracyjne i cybernetyczne, a także strefy wrażliwe, takie jak Przesmyk Suwalski i Obwód kaliningradzki. Na obszarze Morza Bałtyckiego wskazano kluczową infrastrukturę krytyczną oraz połączenia podmorskie, do których należą m.in. gazociąg Baltic Pipe, kable telekomunikacyjne i energetyczne oraz linie przesyłowe SwePol Link i NordBalt.

Równolegle na mapie zaznaczono strategiczne węzły energetyczne i logistyczne, w tym terminal LNG w Świnoujściu, farmy offshore wind koło Łeby oraz Rzeszów pełniący funkcję huba logistycznego dla Ukrainy. Istotnym elementem schematu jest wymiar technologiczny, reprezentowany przez istniejące fabryki AI w Poznaniu (PCSS) i Krakowie (Cyfronet) działające w ramach EuroHPC. Mapę dopełniają rozważane lokalizacje pod budowę gigafabryki AI (m.in. Warszawa, Wrocław, Skawina) oraz koncepcja AI Data Center w Łowej oparta na odnawialnych źródłach energii i wodrze.

Polska: państwo frontowe i potencjalny węzeł cyfrowy Europy Środkowej

Rysunek 7. Polska jako państwo frontowe i potencjalny węzeł cyfrowy Europy Środkowej: teatry presji, infrastruktura krytyczna i ośrodki mocy obliczeniowej. Opracowanie własne na podstawie danych publicznych.

4.3. Mapa oddziaływań geopolitycznych w domenie AI

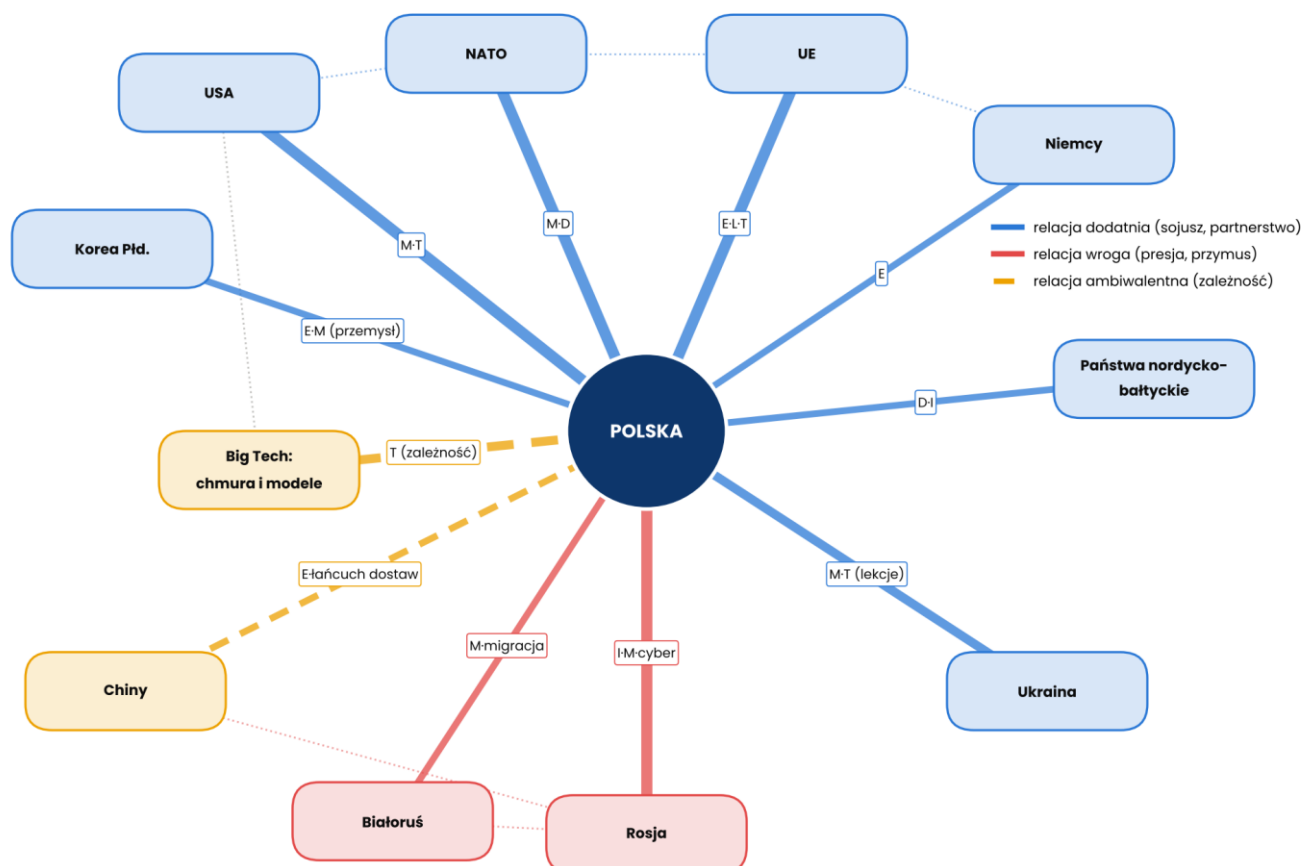
Relacje Polski z kluczowymi aktorami w domenie AI i cyberbezpieczeństwa mają różny znak i różną moc. Stany Zjednoczone są głównym dostawcą technologii i gwarantem bezpieczeństwa, ale ich polityka staje się bardziej transakcyjna, a kontrola eksportu mocy obliczeniowej może obejmować także sojuszników.

Unia Europejska jest źródłem regulacji, finansowania i rynku. Ukraina jest najważniejszym laboratorium wojny dronowej, walki radioelektronicznej i cyberobrony w warunkach wojny — partnerem, od którego Polska może uczyć się szybciej niż ktokolwiek w NATO. Rosja i Białoruś są źródłem presji hybrydowej.

Chiny są jednocześnie dostawcą komponentów i źródłem ryzyka w łańcuchu dostaw (procedura dostawcy wysokiego ryzyka). Państwa nordyckie i bałtyckie są wzorcem obrony totalnej i odporności społecznej; Korea Południowa — partnerem przemysłowym.

Rysunek 8 przedstawia graf oddziaływań geopolitycznych Polski w domenie sztucznej inteligencji i cyberbezpieczeństwa, opracowany na podstawie Teorii Oddziaływań Geopolitycznych. Schemat ukazuje sieć relacji łączących Polskę z kluczowymi aktorami międzynarodowymi, odróżniając stosunki sojusznicze, wrogie oraz ambiwalentne. Grubość poszczególnych linii odzwierciedla moc relacji, stanowiącą iloczyn jej siły oraz częstotliwości kontaktów. Etykiety przypisane do krawędzi grafu definiują dominujące instrumenty oddziaływania z katalogu DIME+T+L, takie jak wojsko, technologia, gospodarka, prawo czy informacja. Graf w przejrzysty sposób ilustruje partnerstwa ze Stanami Zjednoczonymi, Unią Europejską, NATO czy Ukrainą, jednocześnie wskazując obszary presji ze strony Rosji i Białorusi oraz zależności od Chin i globalnych korporacji technologicznych.

Grubość linii = moc relacji (siła × częstotliwość); etykieta = dominujący instrument DIME+T+L



Rysunek 8. Graf oddziaływań geopolitycznych Polski w domenę AI i cyberbezpieczeństwa: znak relacji, jej moc i dominujący instrument. Opracowanie własne według Teorii Oddziaływań Geopolitycznych.

4.4. Instrumenty siły DIME+T+L w domenę AI

Narodowa Strategia Bezpieczeństwa rozszerza klasyczne instrumenty siły — dyplomację, informację, wojsko i gospodarkę — o technologię i prawo. W domenę AI każdy z tych instrumentów ma konkretną treść.

Instrument	Zastosowanie w strategii AI bezpieczeństwa
D — dyplomacja	koalicje compute i danych w Europie Środkowej; udział w standardach NATO i UE dla AI wojskowej; partnerstwo technologiczne z Ukrainą, państwami nordyckimi i Koreą Płd.
I — informacja	Narodowy Obraz Sytuacyjny; Program Odporności Kognitywnej; wspólne z sojusznikami wykrywanie operacji FIMI
M — wojsko	AI w C4ISR, obronie powietrznej i counter-UAS; autonomia z człowiekiem w dowodzeniu; Wojska Obrony Cyberprzestrzeni jako rdzeń kompetencji
E — gospodarka	zamówienia innowacyjne i pierwszy klient dla krajowych firm; przemysł dual-use; data center przy źródłach energii
T — technologia	suwerenny stos AI: compute, modele krajowe, dane, kryptografia postkwantowa, bezpieczeństwo łańcucha dostaw
L — prawo	spójne wdrożenie NIS2, AI Act i ustawy o systemach AI; reguły dla AI w obronności; rejestr algorytmów i kontrola drugiego rzędu

4.5. Warianty strategiczne zależne od otoczenia

Strategia nie stawia na jeden scenariusz. Definiuje warianty odpowiedzi zależne od dwóch zmiennych krytycznych: wiarygodności gwaranta technologicznego i bezpieczeństwa (USA, NATO, UE) oraz agresywności przeciwnika. W wariantcie „Kotwica” (gwarant wiarygodny, przeciwnik powściągliwy) Polska pogłębia integrację z sojuszniczymi systemami AI i buduje hub obliczeniowy regionu. W wariantcie „Tarcza frontowa” (gwarant wiarygodny, przeciwnik agresywny) priorytetem jest maksymalna zdolność obrony cyfrowej i counter-UAS pod parasolem sojuszu. W wariantcie „Autonomia zdolności” (gwarant osłabiony, przeciwnik powściągliwy) przyspiesza budowa europejskiego i krajowego stosu AI. W wariantcie „Samodzielna odporność” (gwarant osłabiony, przeciwnik agresywny) decydujące stają się krajowe compute, modele, energia i kadry, zdolne działać bez dostępu do zewnętrznych usług.

REGUŁA PRZEŁĄCZANIA

Państwo utrzymuje wymaganą różnorodność, jeżeli potrafi przełączyć się między wariantami bez utraty celów nadrzędnych. Dlatego każda krytyczna zdolność AI musi mieć wersję działającą w trybie „odciętym” — na krajowym compute, z krajowym modelem i lokalnym zasilaniem — nawet jeśli w czasie pokoju korzysta z usług sojuszniczych lub komercyjnych.

Rozdział czwarty przedstawia analizę geopolityczną, ukazując technologię jako nową oś potęgi w erze wielobiegunowej rywalizacji mocarstw. W tym układzie Stany Zjednoczone kontrolują rynek mocy obliczeniowej, Chiny budują własne łańcuchy dostaw, a Unia Europejska stara się łączyć regulacje z inwestycjami.

W oparciu o Teorię Oddziaływań Geopolitycznych wskazano, że Polska ma bezpośredni wpływ wyłącznie na zmniejszanie własnej podatności na zagrożenia. Położenie Polski definiuje ją nie tylko jako państwo frontowe i zaplecze dla Ukrainy, ale także jako potencjalny węzeł cyfrowo-energetyczny Europy Środkowej.

Kluczową rolę w tej domenie odgrywa współpraca techniczna i wymiana doświadczeń wojennych z Ukrainą oraz państwami nordyckimi i bałtyckimi. Jednocześnie zidentyfikowano ryzyka wynikające ze zbyt dużej zależności od globalnych gigantów technologicznych oraz dostawców z Chin.

W analizie posłużono się rozszerzonym katalogiem instrumentów siły państwa DIME+T+L, dodającym wymiar technologii oraz prawa. Każdy z tych instrumentów otrzymał konkretną rolę operacyjną, od budowania koalicji obliczeniowych po suwerenny stos technologiczny. Strategia nie opiera się na jednym scenariuszu, lecz definiuje cztery warianty odpowiedzi zależne od wiarygodności sojuszników oraz presji ze strony przeciwnika. Niezależnie od przyjętego wariantu, kluczowa pozostaje zasada utrzymania wersji krytycznych systemów AI zdolnych do działania w trybie całkowicie odciętym.

5. Analiza zasobowa

Analiza zasobowa odpowiada na pytanie, czym Polska dysponuje, aby zbudować Narodowy System AI Bezpieczeństwa i gdzie leżą wąskie gardła. W logice cybernetycznej zasoby dzielą się na tor informacyjny (dane, modele, wiedza), tor zasilający (energia, compute, kapitał) i tor ludzki (talenty), który spina oba pozostałe.

Zasób	Stan 2026	Kluczowa luka	Dźwignia strategiczna
Dane	rozbudowane rejestry państwowe, mObywatel, telemetria operatorów i CSIRT-ów, dane wojskowe	silosy, brak interoperacyjności semantycznej, brak wspólnego obrazu sytuacyjnego, ograniczone udostępnianie między sektorami	Narodowy Obraz Sytuacyjny; federacyjne uczenie i wymiana wskaźników zagrożeń (IoC) w czasie rzeczywistym
Moc obliczeniowa	PLGrid, fabryki AI w Poznaniu i Krakowie, inwestycje komercyjnych centrów danych	brak zasobów dedykowanych bezpieczeństwu i obronności w trybie odciętym; zagraniczna własność dużych centrów danych	rezerwa compute bezpieczeństwa; udział w gigafabryce AI z klauzulą dostępu dla obronności
Modele	PLLuM, Bielik, kompetencje naukowe w NLP i wizji komputerowej	brak modeli wyspecjalizowanych w cyber i analizie wywiadowczej; brak testów odporności i krajowego centrum ewaluacji	Narodowe Laboratorium Ewaluacji i Bezpieczeństwa Modeli; specjalizacja zamiast wyścigu skali
Energia	dywersyfikacja źródeł, rozwój OZE, plany jądrowe, magazyny	koszty energii, ograniczenia przyłączeniowe, podatność systemów sterowania rozproszonej generacji	data center przy źródłach i magazynach; wyspy zasilania dla węzłów krytycznych
Kapitał	rekordowe wydatki obronne (ok. 4,8% PKB w 2026 r.), pożyczki SAFE (43,7 mld euro), fundusze UE	mała część budżetów trafia do software, AI i krajowych dostawców; brak kapitału cierpliwego dla deep tech	budżet wyniku z jawną kategorią „system nerwowy” (C4ISR, cyber, AI, dane)
Przemysł	rosnąca baza przemysłu obronnego, sektor IT i gier, producenci dronów	rozproszenie, brak pierwszego klienta, długie zamówienia	Sandbox Obronny i zamówienia innowacyjne z cyklem 90–180 dni
Instytucje	KRiBSI, PCOC, CSIRT GOV/MON/NASK, WOC, CISI, RCB	rozproszenie kompetencji, brak właściciela wyniku całości	Rada Bezpieczeństwa AI jako regulator S4/S5
Talenty	duży sektor IT, silne olimpiady informatyczne i matematyczne, uznane zespoły badawcze	brak rejestru, odpływ do sektora prywatnego i za granicę, niedobór kilkunastu tysięcy specjalistów cyber	Narodowy System Talentów AI-Cyber (rozdział 6)

Wniosek z analizy zasobowej jest jednoznaczny: Polska nie jest państwem ubogim w zasoby potrzebne do bezpieczeństwa AI. Jest państwem, w którym zasoby nie są połączone. Dane nie spotykają się z modelami, modele z compute, compute z energią, a wszystkie trzy - z ludźmi zdolnymi je zintegrować. Strategia zasobowa polega więc mniej na dokupowaniu, a bardziej na łączeniu - co w języku sieci oznacza budowę mostów między klastrami, a w języku cybernetyki - domknięcie torów sterowania.



Miara suwerenności = koszt wyjścia i czas migracji dla każdej warstwy. Ocena ekspercka TechnoSanacji.

Rysunek 9. Stos suwerenności technologicznej: warstwy od energii i półprzewodników po agentów AI i warstwę kontrolną, z oceną stopnia zależności Polski. Opracowanie własne.

Rysunek 9 przedstawia stos suwerenności technologicznej Polski, podzielony na osiem poziomów od L0 (Energia) do L7 (Warstwa kontrolna). Dla każdej warstwy wskazano obecny stan jej rozwoju w kraju oraz oceniono poziom zewnętrznej zależności Polski. Najwyższy, krytyczny stopień zależności występuje w warstwie L1 (Półprzewodniki i sprzęt), co wynika z całkowitego opierania się na importowanych akceleratorach i sprzęcie sieciowym. Wysoki poziom zależności odnotowano przy mocy obliczeniowej (L2 Compute i chmura) oraz zaawansowanych modelach AI (L4), gdzie dominują podmioty spoza Unii Europejskiej. Niską zależnością i większą suwerennością charakteryzują się natomiast warstwy danych (L3), kryptografii (L6) oraz warstwa kontrolna (L7).

Rozdział piąty dokonuje analizy zasobów niezbędnych do zbudowania Narodowego Systemu AI Bezpieczeństwa, dzieląc je na tor informacyjny, zasilający i ludzki. Wykazano, że Polska nie jest państwem ubogim w zasoby, lecz cierpi na brak ich wzajemnego połączenia i koordynacji. Głównym wąskim gardłem pozostaje niedobór specjalistów i brak rejestru kadr krytycznych, co prowadzi do drenażu talentów z sektora publicznego.

6. Analiza kadrowa: talenty jako infrastruktura krytyczna

6.1. Dlaczego talent AI i cyber jest infrastrukturą krytyczną

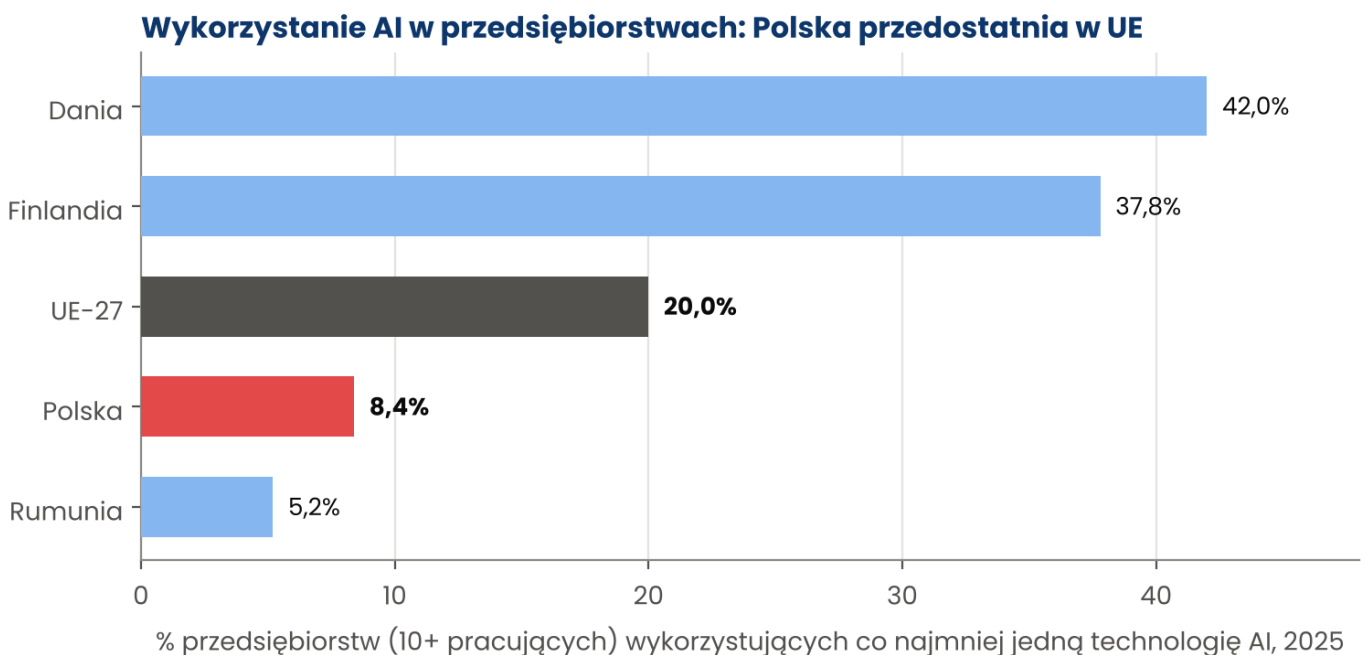
W polskim porządku prawnym infrastruktura krytyczna to systemy oraz powiązane z nimi obiekty, urządzenia, instalacje i usługi kluczowe dla bezpieczeństwa państwa i obywateli. Prawo Unii Europejskiej — dyrektywa CER o odporności podmiotów krytycznych i dyrektywa NIS2 — przesunęło akcent z ochrony obiektów na odporność usług kluczowych. To przesunięcie otwiera drogę do formuły TechnoSanacji: usługa kluczowa działa tylko wtedy, gdy są ludzie zdolni ją zaprojektować, utrzymać, zabezpieczyć i odbudować.

W domenie AI i cyberbezpieczeństwa ta prawda jest szczególnie ostra. Centrum operacji bezpieczeństwa (SOC) jest tak dobre, jak jego analitycy. System sterowania siecią elektroenergetyczną jest tak bezpieczny, jak wiedza automatyków o jego podatnościach. Model językowy jest tak godny zaufania, jak zespół, który go trenuje, testuje i aktualizuje. Dron jest tak skuteczny, jak operator i inżynier, który w ciągu tygodnia przeprogramuje go po zmianie taktyki walki radioelektronicznej przeciwnika. Każda z tych zdolności ma czas odtworzenia liczony w latach — a jej utrata może nastąpić w ciągu jednej oferty pracy.

Cecha infrastruktury krytycznej	Odpowiednik w zasobie talentów AI-cyber	Konsekwencja dla polityki
Niezbędność dla funkcji państwa	bez analityków, inżynierów OT, kryptologów, badaczy AI i operatorów BSP nie działają cyberobrona, energetyka, łączność i wojsko	talent w katalogu zasobów krytycznych; właściciel odpowiedzialności w każdej dziedzinie
Długi czas odtworzenia	8–15 lat od identyfikacji uzdolnień do samodzielnego eksperta; 2–4 lata do zgrania zespołu SOC lub zespołu badawczego	inwestycja ciągła i antycykliczna, chroniona przed kadencyjnością
Efekty kaskadowe	odejście kluczowego zespołu przerywa program, łańcuch wiedzy i kształcenie następców	mapa zależności kompetencyjnych i plan ciągłości
Podatność na przechwycenie	rekrutacja przez globalne korporacje, obce programy talentowe, operacje wywiadowcze	bezpieczeństwo badań, konkurencyjna oferta, ochrona przed wrogim transferem wiedzy — bez izolacji
Efekt sieciowy	produktywność rośnie z gęstością otoczenia: zespołu, danych, compute, misji	klastry, centra doskonałości, suwerenny compute dla badaczy i obrońców
Wymóg utrzymania	półokres życia wiedzy w cyberbezpieczeństwie i AI liczony jest w miesiącach	uczenie się przez całe życie jako „serwis” infrastruktury, a nie przywilej
Redundancja	funkcja krytyczna oparta na jednej osobie jest pojedynczym punktem awarii	minimalny współczynnik zastępowalności dla każdej zdolności krytycznej

6.2. Bilans kadrowy Polski

Bilans kadrowy jest paradoksalny. Polska ma jeden z największych w Europie Środkowej sektorów usług informatycznych, wysokie pozycje w międzynarodowych olimpiadach i konkursach programistycznych, uznane zespoły kryptologiczne i matematyczne - a jednocześnie należy do ostatnich w UE pod względem wykorzystania AI w przedsiębiorstwach (8,4% firm wobec 20,0% średniej unijnej w 2025 r.) i ma udział specjalistów ICT w zatrudnieniu wyraźnie niższy od średniej UE (5,0% w 2025 r.). Analizy rynku pracy szacują niedobór specjalistów cyberbezpieczeństwa na kilkanaście tysięcy osób, przy skokowo rosnącym popycie wynikającym z rozszerzenia obowiązków NIS2 z ok. 400 do ponad 10 tys. podmiotów.



Źródło: Eurostat, komunikat z 11 grudnia 2025 r. (najwyższe i najniższe wartości oraz średnia UE).

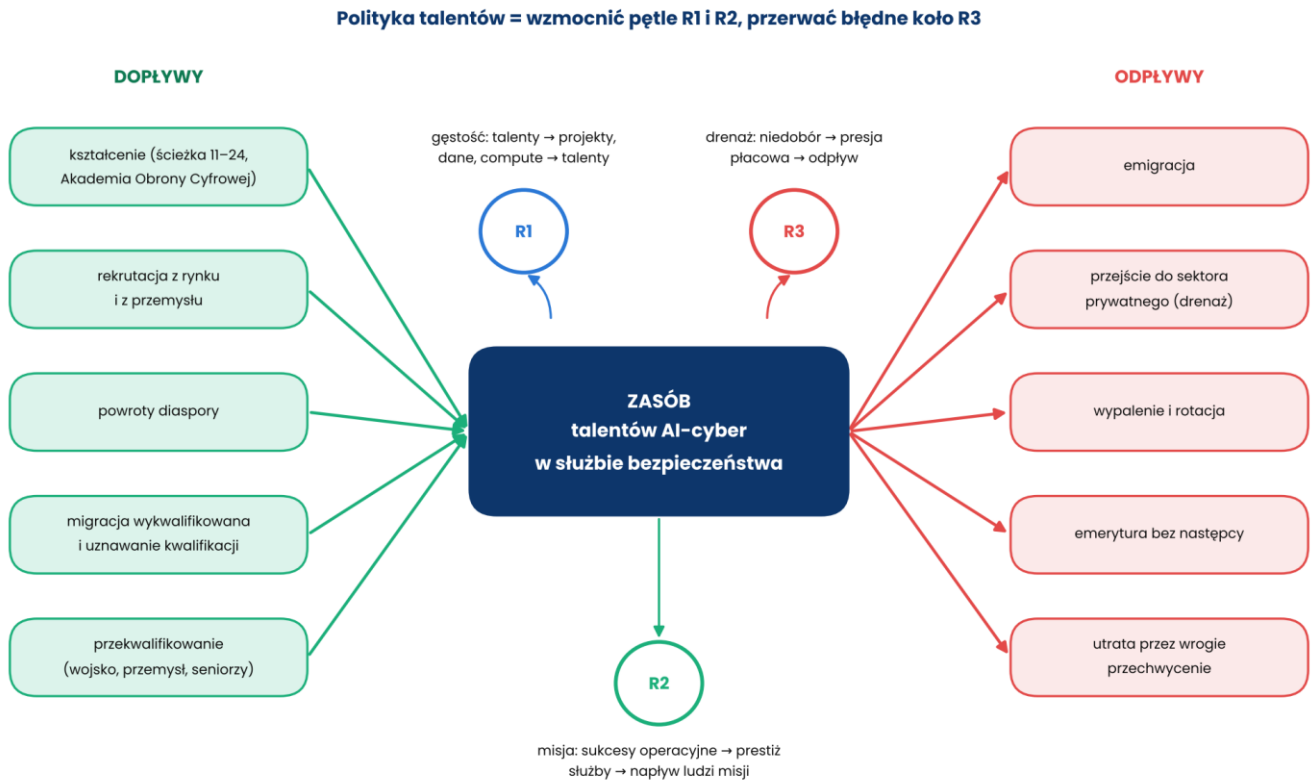
Rysunek 10. Odsetek przedsiębiorstw (10+ pracujących) wykorzystujących technologie AI w 2025 r. — Polska na tle UE.

Źródło: Eurostat, grudzień 2025 r.

Paradoks ma wyjaśnienie systemowe. Talent wytwarzany w Polsce jest w dużej mierze eksportowany — fizycznie lub w modelu pracy zdalnej dla zagranicznych firm — ponieważ krajowe otoczenie nie oferuje porównywalnej gęstości misji, danych, compute i kapitału. Sektor publiczny konkuruje o tych samych ludzi z siatką płac niezdolną do konkurencji, a sektor obronny — z procedurami niezdolnymi do szybkiego zatrudnienia cywilnego eksperta. W efekcie państwo jest zależne od kilku węzłów kompetencyjnych (NASK, WOC, wybrane uczelnie i firmy), których utrata miałaby skutki kaskadowe.

6.3. Zasoby i przepływy talentu

W ujęciu dynamiki systemów talent jest zasobem (stock), który zmienia się przez przepływy: dopływ (kształcenie, rekrutacja, powroty, migracja), odpływ (emigracja, przejście do sektora prywatnego, wypalenie, emerytura) i konwersję (przekwalifikowanie, awans kompetencyjny). Polityka kadrowa państwa przez lata zajmowała się prawie wyłącznie dopływem — liczbą miejsc na studiach i kursach — ignorując odpływ i konwersję. Tymczasem w dziedzinach o wysokim popycie globalnym to odpływ i retencja decydują o stanie zasobu.



Rysunek 11. Talenty jako infrastruktura krytyczna: zasoby, przepływy i pętle sprzężeń zwrotnych w systemie kadr AI-cyber. Opracowanie własne.

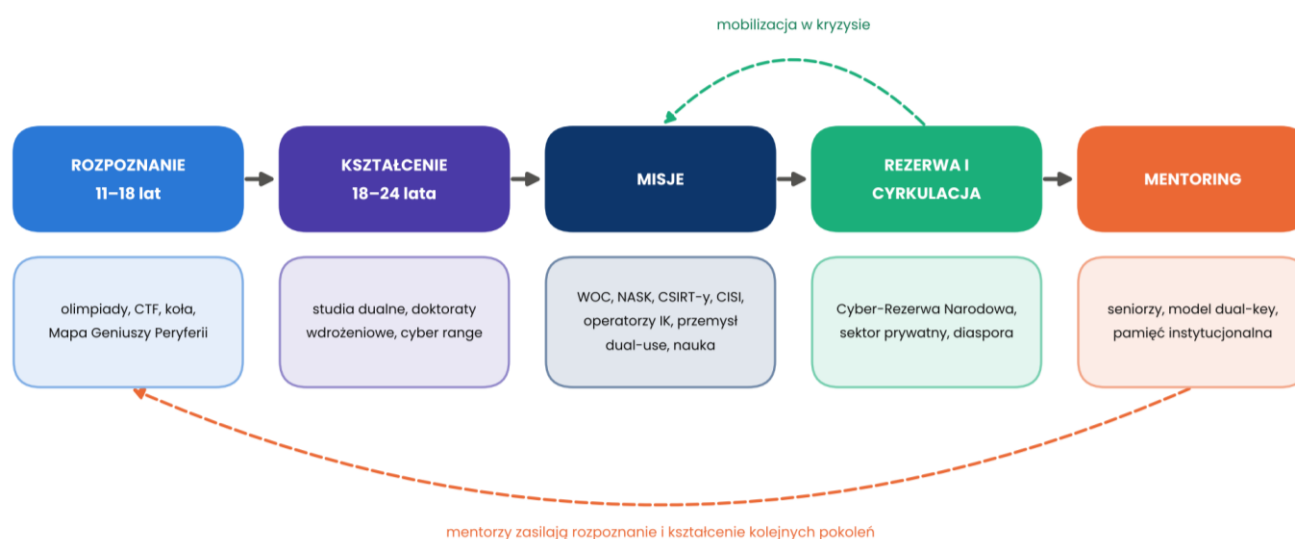
Model wskazuje trzy pętle wzmacniające, z których dwie działają na korzyść państwa, a jedna przeciw niemu. Pętla R1 (gęstość): więcej talentów przyciąga więcej projektów, danych i compute, co przyciąga kolejne talenty — to mechanizm klastrów takich jak Tel Awiw, Tallin czy Cambridge. Pętla R2 (misja): sukcesy w realnych operacjach obronnych zwiększają prestiż służby i przyciągają ludzi, dla których misja jest ważniejsza niż płaca.

Pętla R3 (drenaż) jest błędnym kołem: rosnący niedobór podnosi presję płacową sektora prywatnego i zagranicy, co wysysa specjalistów z sektora publicznego i pogłębia jego niedobór. Strategia polega na wzmocnieniu R1 i R2 oraz przerwaniu R3 — pakietem retencyjnym, misją, dostępem do compute i modelem mieszanym (rezerwa, kontrakty eksperckie).

6.4. Narodowy System Talentów AI-Cyber

Strategia przenosi do domeny AI i cyberbezpieczeństwa instytucje zaprojektowane w publikacji „Talenty – infrastruktura krytyczna państwa” i dodaje elementy specyficzne dla bezpieczeństwa.

Instrument	Funkcja	Analogia infrastrukturalna
Rejestr Zdolności Krytycznych AI-Cyber	inwentaryzacja specjalistów w 12 rolach krytycznych (analityk SOC, inżynier bezpieczeństwa OT/ICS, kryptolog i specjalista PQC, badacz AI, inżynier MLOps i bezpieczeństwa modeli, red teamer, analityk FIMI, operator i inżynier BSP, specjalista walki radioelektronicznej, analityk wywiadu z otwartych źródeł, audytor algorytmów, architekt systemów danych)	rejestr obiektów infrastruktury krytycznej
Wskaźnik Talentu Krytycznego (WTK)	pomiar stanu: liczebność, struktura wieku i kompetencji, rotacja, współczynnik zastępowalności, odpływ	wskaźnik stanu technicznego sieci
Plan ciągłości kompetencyjnej	dla każdej zdolności krytycznej — następcy, dokumentacja wiedzy, minimalny współczynnik zastępowalności 1:3	plan ciągłości działania operatora
Cyber-Rezerwa Narodowa	przeszkoleni specjaliści cywilni z sektora prywatnego, nauki i diaspory, mobilizowani w kryzysie na podstawie umów i ćwiczeń; wzorce estońskiej Cyber Defence Unit i rezerw cyber państw nordyckich	rezerwa strategiczna
Akademia Obrony Cyfrowej	sieć kształcenia łącząca uczelnie, WOC, NASK i przemysł: ścieżka 11–24 dla uzdolnionych, studia dualne, doktoraty wdrożeniowe, poligony cyber (cyber range)	modernizacja i rozbudowa sieci
Pakiet retencyjny służby cyfrowej	odrębna siatka płac dla ról krytycznych powiązana z rynkiem, autonomia zadaniowa, dostęp do compute i danych, praca na misjach o znaczeniu państwowym, szybka ścieżka poświadczeń bezpieczeństwa	utrzymanie i serwis
Program Cyrkulacji Diaspory	afiliacje wirtualne, granty powrotowe z gwarancją infrastruktury, wspólne zespoły z polskimi specjalistami za granicą	połączenia międzysystemowe
Szybka ścieżka kwalifikacji	uznawanie kompetencji migrantów i osób przekwalifikowujących się (np. z wojska i przemysłu) w miesiącach, nie latach	podłączenie istniejącej infrastruktury do sieci
Model dual-key Gen Z + seniorzy	młody specjalista wnosi szybkość i narzędzia AI, doświadczony mentor — kontekst, pamięć błędów i odpowiedzialność; rozbieżność jest dokumentowana	redundancja i pamięć instytucjonalna

Ścieżka talentu AI-cyber: rejestr, wskaźnik stanu i plan ciągłości na każdym etapie

Rysunek 12. Ścieżka talentu AI-cyber: od wczesnego rozpoznania (11–24) przez kształcenie i misje do rezerwy i mentoringu — przepływy między sektorami. Opracowanie własne.

6.5. Progi alarmowe dla infrastruktury talentu

Talent jako infrastruktura krytyczna wymaga takich samych progów alarmowych jak sieć energetyczna. Przekroczenie progu uruchamia procedurę, a nie komentarz.

Wskaźnik	Próg ostrzegawczy	Próg krytyczny	Reakcja
Współczynnik zastępowalności w roli krytycznej	< 1:3	< 1:2	program następców i rekrutacja priorytetowa w 90 dni
Roczna rotacja w zespołach SOC sektora publicznego	> 15%	> 25%	przegląd pakietu retencyjnego, rozmowy wyjściowe, korekta płac
Wakaty w rolach krytycznych WOC/NASK/CSIRT	> 10%	> 20%	uruchomienie Cyber-Rezerwy, kontrakty eksperckie
Odsetek absolwentów kierunków cyber i AI pracujących dla podmiotów krajowych po 3 latach	< 60%	< 45%	program misji, staży i grantów
Czas poświadczenia bezpieczeństwa dla eksperta cywilnego	> 90 dni	> 180 dni	ścieżka priorytetowa, dodatkowe zasoby ABW/SKW

ZASADA GRANICZNA POLITYKI TALENTÓW

Infrastrukturę państwo posiada; talentów nie posiada. Formuła „talent jako infrastruktura krytyczna” nie uzasadnia zakazów wyjazdu, zobowiązań lojalnościowych ani profilowania bez zgody. Uzasadnia coś odwrotnego: obowiązek państwa, by tworzyć warunki, w których najlepsi chcą pracować dla bezpieczeństwa Polski — z dowolnego miejsca świata.

7. Analiza kulturowa

Strategie bezpieczeństwa zawodzą najczęściej nie z powodu technologii, lecz kultury: sposobu, w jaki ludzie i instytucje postrzegają ryzyko, błąd, autorytet i współpracę. Analiza kulturowa wskazuje cztery napięcia, które strategia musi świadomie zarządzać.

Kultura procedury a kultura wyniku. Polska administracja jest wychowana w logice zgodności: sukcesem jest poprawnie przeprowadzona procedura, a nie osiągnięty efekt. W cyberbezpieczeństwie prowadzi to do „bezpieczeństwa na papierze” — polityk, rejestrów i audytów, które nie skracają czasu reakcji. Strategia wymaga przesunięcia akcentu na mierzalny wynik i nagradzania instytucji za szybkie wykrycie i zgłoszenie incydentu, a nie za jego brak w statystyce.

Kultura ukrywania błędu a kultura zgłaszania. Organizacje wysokiej niezawodności — lotnictwo, energetyka jądrowa, kontrola ruchu lotniczego — łączy zaabsorbowanie możliwością porażki i bezpieczne zgłaszanie błędów. W polskiej administracji zgłoszenie incydentu bywa traktowane jako przyznanie się do winy. Bez kultury „just culture” AI nie dostanie danych, na których mogłaby się uczyć.

Nieufność wobec państwa a potrzeba współpracy. Niski poziom zaufania instytucjonalnego utrudnia budowę wspólnego obrazu sytuacyjnego i programów odporności kognitywnej. Jednocześnie badania odporności pokazują, że zaufanie trzeba budować warstwowo: zaufanie do instytucji i nauki chroni, a bezkrytyczne zaufanie uogólnione może zwiększać podatność. Państwo zdobywa zaufanie przejrzystością — rejestrem algorytmów, jawnymi kryteriami, rozliczalnością — a nie kampaniami.

Cyfrowa odwaga młodych a pamięć instytucjonalna starszych. Generacja Z wnosi intuicję technologiczną i szybkość prototypowania; seniorzy — pamięć błędów, znajomość procedur kryzysowych i etos odpowiedzialności. Osobno obie grupy są podatne: młodzi na manipulację algorytmiczną, starsi na oszustwa z klonowaniem głosu. Razem tworzą regulator, który TechnoSanacja opisuje jako sojusz pokoleń. Strategia przekłada go na model dual-key w zespołach bezpieczeństwa i na wspólne programy edukacji cyfrowej.

TEZA KULTUROWA

Technologia AI jest dziś dostępna niemal każdemu. Przewagę daje kultura, która pozwala ją szybko i odpowiedzialnie wdrożyć: kultura wyniku, zgłaszania błędów, przejrzystości i współpracy pokoleń. Kultury nie da się kupić w przetargu — trzeba ją zbudować jako zdolność państwa.

8. Analiza technologiczna

8.1. Stos technologiczny bezpieczeństwa AI

Bezpieczeństwo AI jest stosem warstw, z których każda może stać się pojedynczym punktem awarii. Analiza obejmuje osiem warstw, od fizycznej po instytucjonalną.

Warstwa	Kluczowe technologie	Stan Polski	Kierunek
0. Energia	sieć, magazyny, wodór, SMR, wyspy zasilania	dywersyfikacja w toku, podatne systemy sterowania OZE	zasilanie węzłów krytycznych w trybie wyspowym, cyberbezpieczeństwo OT
1. Półprzewodniki i sprzęt	GPU/akceleratory, sprzęt sieciowy, sensory	pełna zależność od importu	dywersyfikacja dostaw, zapasy strategiczne, udział w europejskich projektach
2. Compute i chmura	centra danych, HPC, chmura suwerenna	fabryki AI, inwestycje komercyjne	rezerwa compute bezpieczeństwa, gigafabryka z klauzulą obronną
3. Dane	rejstry, telemetria, dane wywiadowcze, dane OSINT	silosy	Narodowy Obraz Sytuacyjny, standardy wymiany, federacyjne uczenie
4. Modele	modele językowe, wizja komputerowa, modele detekcji anomalii	PLLuM, Bielik, zespoły akademickie	modele specjalistyczne dla cyber i analizy, ewaluacja i red teaming
5. Agenci i automatyzacja	agenci SOC, automatyczna reakcja (SOAR), autonomia BSP	pojedyncze wdrożenia	agenci z tożsamością maszynową i warstwą kontrolną
6. Kryptografia i tożsamość	kryptografia postkwantowa (PQC), zero trust, tożsamość cyfrowa	mObywatel, kompetencje kryptologiczne	plan migracji do PQC, zero trust w administracji
7. Warstwa kontrolna	rejestr algorytmów, audyt, ślad decyzji, human-in-command	załączek w ustawie o systemach AI	warstwa kontrolna dla każdego systemu władczego

8.2. Technologie przełomowe i ich znaczenie

Agentowa AI w cyberbronie. Agenci zdolni do samodzielnego triażu alertów, korelacji zdarzeń i proponowania działań naprawczych mogą wielokrotnie zwiększyć wydajność analityków. Warunkiem jest warstwa kontrolna: tożsamość maszynowa agenta, ograniczone uprawnienia, rejestr działań i stopniowany nadzór człowieka zależny od odwracalności działania.

Modele otwarte a zamknięte. Modele o otwartych wagach dają możliwość uruchomienia, audytu i dostrojenia na krajowej infrastrukturze, co jest warunkiem trybu „odciętego”. Modele zamknięte oferują zwykle wyższą wydajność graniczną. Strategia przyjmuje zasadę portfelową: zdolności krytyczne muszą mieć wersję działającą na modelach, które państwo może uruchomić samodzielnie; modele komercyjne są dopuszczalne tam, gdzie koszt wyjścia jest znany i akceptowalny.

Kryptografia postkwantowa. Dane wykradane dziś mogą zostać odszyfrowane w przyszłości („zbiierz teraz, odszyfruj później”). Migracja do algorytmów postkwantowych w systemach państwowych i infrastrukturze krytycznej jest procesem wieloletnim i powinna rozpocząć się od inwentaryzacji kryptografii.

Autonomia i counter-UAS. Wojna w Ukrainie pokazała, że tanie drony z elementami autonomii — nawigacją wizyjną odporną na zakłócenia, rozpoznaniem celów — zmieniają pole walki. Obrona wymaga fuzji danych z tanich sensorów, automatycznej klasyfikacji i tanich efektorów. Cykl adaptacji liczony jest w tygodniach.

AI w energetyce. Algorytmy prognozowania, sterowania magazynami i zarządzania obciążeniem (np. sieci LSTM i uczenie ze wzmocnieniem opisane w studium AI Data Center w Łłowej) zwiększają odporność systemu, ale jednocześnie poszerzają powierzchnię ataku. Każdy algorytm sterujący infrastrukturą fizyczną musi przejść testy odporności na manipulację danymi wejściowymi.

Bio-AI. Modele projektowania białek i analiza genomowa przyspieszają medycynę, a jednocześnie wymagają nadzoru nad dostępem do usług syntezy i automatyzacji laboratoriów. Dla państwa kluczowe jest wczesne ostrzeżenie — nadzór ścieków, nadzór syndromiczny i genomowy — wspierane AI w podejściu One Health.

Przełomowe technologie, takie jak agentowa AI, drony autonomiczne czy modele Bio-AI, diametralnie zwiększają efektywność działań operacyjnych, lecz stwarzają nowe wyzwania w zakresie kontroli i bezpieczeństwa. Aby zapewnić odporność państwa w trybie odciętych, kluczowe jest stosowanie modeli o otwartych wagach na własnej infrastrukturze oraz wieloletnia migracja na kryptografię postkwantową.

W obszarze obronności i energetyki decydujące znaczenie mają szybkie cykle adaptacji oraz stałe testowanie algorytmów sterujących pod kątem podatności na manipulację. Całość strategii technologicznej opiera się na zasadzie portfelowej, łączeniu tanich sensorów z automatyzacją oraz wdrażaniu nadzoru w podejściu One Health.

Radar zdolności bezpieczeństwa AI: stan 2026 i cel 2030 (0–5)



Ocena ekspercka TechnoSanacji.

Rysunek 13. Radar zdolności Polski w dziesięciu wymiarach bezpieczeństwa AI: ocena stanu 2026 i cel 2030 (skala 0–5).
Ocena ekspercka TechnoSanacji.

Rysunek 13 przedstawia wykres radarowy oceniający potencjał Polski w dziesięciu kluczowych wymiarach bezpieczeństwa AI na pięciostopniowej skali. Diagram w przejrzysty sposób zestawia wyjściowy stan z 2026 roku, oznaczony czerwoną linią, z docelowym poziomem planowanym do osiągnięcia w 2030 roku, zaznaczonym niebieską linią przerywaną.

Ocena stanu wyjściowego ujawnia istotne luki m.in. w obszarze mocy obliczeniowej (Compute bezpieczeństwa), danych i obrazu sytuacyjnego oraz systemów Counter-UAS i autonomii. Wyznaczone na rok 2030 cele zakładają znaczne skoki zdolności we wszystkich wymiarach, ze szczególnym uwzględnieniem rozwoju kadr AI-cyber, cyberobrony infrastruktury krytycznej oraz ram prawnych

9. Analiza strategiczna SWOT i TOWS

9.1. Macierz SWOT

Analiza SWOT zestawia czynniki wewnętrzne (mocne i słabe strony państwa) z czynnikami zewnętrznymi (szanse i zagrożenia). W logice TechnoSanacji nie jest ona listą obserwacji, lecz wejściem do macierzy TOWS, z której wyprowadza się strategię.

Mocne strony (S)	Słabe strony (W)
S1. Silna baza talentów informatycznych i matematycznych, sukcesy olimpijskie, duży sektor IT i gier	W1. Brak regulatora łączącego AI, cyber, energię i kadry; rozproszona odpowiedzialność
S2. Doświadczone instytucje: CERT Polska/NASK, Wojska Obrony Cyberprzestrzeni, CSIRT GOV i MON	W2. Niska absorpcja AI w gospodarce (8,4% firm) i administracji
S3. Nowe ramy prawne: KSC/NIS2, ustawa o systemach AI, Strategia Cyberbezpieczeństwa 2025–2029	W3. Silosy danych i brak wspólnego obrazu sytuacyjnego
S4. Rekordowe wydatki obronne i dostęp do pożyczek SAFE	W4. Zależność od zagranicznych dostawców chmury, modeli, sprzętu i oprogramowania
S5. Krajowe modele językowe (PLLuM, Bielik) i fabryki AI w Poznaniu i Krakowie	W5. Niedobór kilkunastu tysięcy specjalistów cyber; brak rejestru talentów i pakietu retencyjnego w sektorze publicznym
S6. Bezpośredni dostęp do doświadczeń wojny w Ukrainie (drony, EW, cyberobrona)	W6. Kultura procedury zamiast wyniku; długie zamówienia publiczne
S7. Sprawna cyfryzacja usług publicznych (mObywatel, e-administracja)	W7. Podatne małe podmioty infrastruktury krytycznej (wodociągi, ciepłownie, szpitale, gminy)
S8. Silny etos obronny społeczeństwa i rosnące zainteresowanie obronnością młodych	W8. Wysokie koszty energii i ograniczenia przyłączeniowe dla centrów danych

Szanse (O)	Zagrożenia (T)
O1. Europejskie inwestycje w AI: gigafabryki, InvestAI, EuroHPC; konsorcjum z Chorwacją, Węgrami, Litwą i Słowacją	T1. Wykładniczy wzrost cyberataków i oszustw wspieranych AI
O2. Przesunięcie obowiązków AI Act dla systemów wysokiego ryzyka do 2027–2028 r. — czas na przygotowanie	T2. Ataki państwowe na energetykę, wodę i transport; sabotaż hybrydowy
O3. Popyt NATO i UE na zdolności dual-use, counter-UAS i cyberobronę; programy SAFE i EDIP	T3. Drony i systemy autonomiczne zmieniające ekonomię obrony powietrznej
O4. Powrót części diaspory i przyciąganie talentów z regionu	T4. Operacje wpływu i deepfake'i podważające zaufanie i procesy wyborcze
O5. Polska jako węzeł energetyczno-cyfrowy Europy Środkowej (OZE, atom, wodór, data center)	T5. Koncentracja compute i modeli u kilku dostawców; kontrola eksportu
O6. Partnerstwo technologiczne z Ukrainą i państwami nordyckimi	T6. Globalna wojna o talenty AI i cyber; drenaż do korporacji i za granicę
O7. Agentowa AI zwiększająca wydajność obrońców wielokrotnie	T7. Konwergencja AI–bio obniżająca bariery dla zagrożeń biologicznych
O8. Kryptografia postkwantowa i zero trust jako okazja do modernizacji architektury państwa	T8. Osłabienie wiarygodności gwaranta i transakcyjność sojuszy



Rysunek 14. Macierz SWOT strategii AI bezpieczeństwa Polski. Opracowanie własne.

9.2. Macierz TOWS — strategie wynikowe

Macierz TOWS przekształca SWOT w cztery typy strategii. Strategia SO wykorzystuje mocne strony do chwytania szans; ST — mocne strony do neutralizacji zagrożeń; WO — szanse do usuwania słabości; WT — minimalizuje słabości, by nie spotkały się z zagrożeniami.

Typ	Strategie wynikowe
SO — ofensywna	SO1. Talenty + gigafabryka: oferta „compute + misja” dla badaczy i obrońców (S1, S5 × O1, O7). SO2. Laboratorium ukraińskie: wspólne centrum adaptacji dronów, EW i cyberobrony (S6 × O3, O6). SO3. Eksport zdolności dual-use z polskich firm przez zamówienia NATO i UE (S4 × O3).
ST — konfrontacyjna	ST1. Tarcza AI-Cyber: SOC wspierane AI dla operatorów kluczowych z wykorzystaniem kompetencji CERT i WOC (S2 × T1, T2). ST2. Counter-UAS o korzystnym cost-exchange z krajową produkcją (S4, S6 × T3). ST3. Program Odporności Kognitywnej oparty na etosie obronnym i współpracy pokoleń (S8 × T4).
WO — przekształcająca	WO1. Rada Bezpieczeństwa AI jako regulator wykorzystujący okno czasowe AI Act (W1 × O2). WO2. Narodowy Obraz Sytuacyjny jako warunek udziału w europejskich systemach wymiany danych (W3 × O1). WO3. Data center przy źródłach OZE i magazynach — energia jako przewaga, nie bariera (W8 × O5). WO4. Program Cyrkulacji Diaspory i szybka ścieżka kwalifikacji (W5 × O4).
WT — defensywna	WT1. Plany wyjścia i rezerwa compute w trybie odciętych (W4 × T5, T8). WT2. Cyber-Rezerwa Narodowa i pakiet retencyjny (W5 × T6). WT3. Pakiet minimalnej higieny dla małych podmiotów IK: zarządzane SOC, MFA, likwidacja RDP z internetu (W7 × T1, T2). WT4. Bio-AI: nadzór wczesnego ostrzegania i bezpieczeństwo badań (W6 × T7).



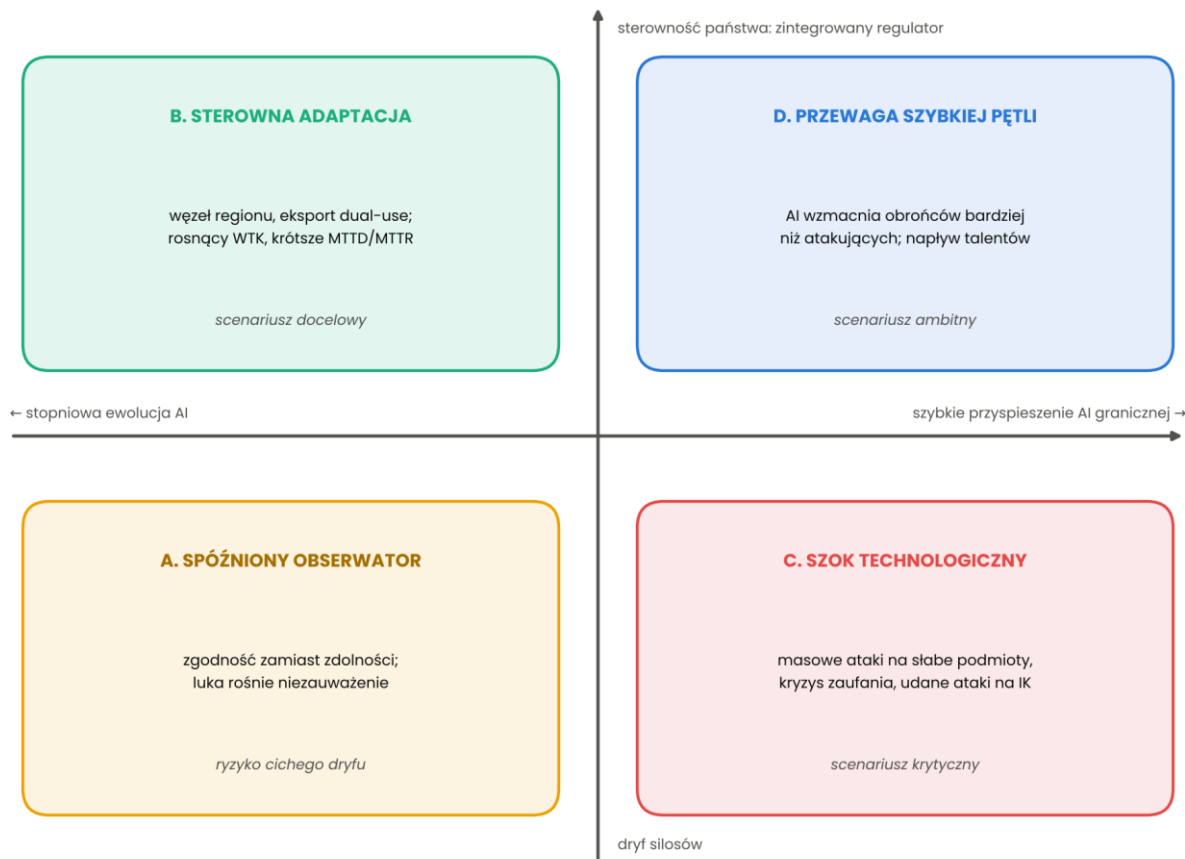
Rysunek 15. Macierz TOWS: od czynników SWOT do czterech typów strategii i programów wykonawczych. Opracowanie własne.

9.3. Analiza PESTEL w skrócie

Wymiar	Czynniki kluczowe dla strategii
Polityczny	polaryzacja i spór kompetencyjny między ośrodkami władzy; rosnąca rola bezpieczeństwa w debacie publicznej; transakcyjność sojuszy
Ekonomiczny	deficyt finansów publicznych przy rekordowych wydatkach obronnych; koszt energii; niska produktywność i absorpcja technologii
Spółeczny	starzenie się społeczeństwa; wysoka podatność na oszustwa cyfrowe; niskie zaufanie instytucjonalne; silny etos obronny
Technologiczny	agentowa AI, autonomia, kryptografia postkwantowa, konwergencja AI-bio; koncentracja compute
Ekologiczny	presja dekarbonizacji; zapotrzebowanie centrów danych na energię i wodę; zjawiska ekstremalne zagrażające infrastrukturze
Prawny	NIS2/KSC, AI Act i ustawa o systemach AI, CER, DSA, Cyber Resilience Act; kontrola TK nad nowelizacją KSC; brak reguł dla AI w obronności

10. Scenariusze 2026–2035

Scenariusze budowane są na dwóch osiach niepewności, które najsilniej różnicują przyszłość bezpieczeństwa AI w Polsce: tempo i dostępność AI granicznej (od stopniowej ewolucji do gwałtownego przyspieszenia zdolności agentowych) oraz spójność i sterowność państwa (od dryfu silosów do zintegrowanego regulatora).



Rysunek 16. Cztery scenariusze bezpieczeństwa AI Polski 2026–2035 wyznaczone przez tempo AI i sterowność państwa. Opracowanie własne.

Scenariusz	Opis	Sygnaty wczesnego ostrzeżenia	Odpowiedź strategii
A. „Spóźniony obserwator” (wolne AI, dryf państwa)	AI rozwija się stopniowo, Polska wdraża przepisy jako zgodność, luka kompetencyjna rośnie powoli i niezauważenie	stagnacja wskaźnika absorpcji AI; brak wzrostu kadr krytycznych	wykorzystać czas na budowę regulatora i rejestru talentów, zanim przyspieszenie nastąpi
B. „Sterowna adaptacja” (wolne AI, spójne państwo)	Polska buduje NSAIB metodycznie, staje się węzłem regionu i eksporterem zdolności dual-use	rosnący WTK, skracające się MTDD/MTTR	scenariusz docelowy; skalować to, co działa
C. „Szok technologiczny” (szybkie AI, dryf państwa)	gwałtowny skok zdolności agentowych po stronie przeciwnika i przestępców; masowe ataki na słabo chronione podmioty; kryzys zaufania	skokowy wzrost incydentów, udane ataki na IK, kampanie deepfake w wyborach	tryb kryzysowy: Cyber-Rezerwa, zarządzane SOC dla wszystkich podmiotów IK, priorytet trybu odciętego
D. „Przewaga szybkiej pętli” (szybkie AI, spójne państwo)	Polska wykorzystuje przyspieszenie do skoku zdolności obronnych i gospodarczych; AI wzmacnia obrońców bardziej niż atakujących	dynamiczny wzrost absorpcji AI, napływ talentów, eksport technologii	maksymalizować SO-strategie, rozbudować compute i misje

WNIOSEK ZE SCENARIUSZY

O tym, czy Polska znajdzie się w scenariuszu B lub D, a nie A lub C, nie decyduje tempo rozwoju AI, na które Polska nie ma wpływu. Decyduje oś sterowności państwa — w całości w polskich rękach. Modernizacja i sanacja państwa są jedyną zmienną strategiczną, którą możemy kontrolować.

11. Uzasadnienia tezy naczelnej: sześć dróg dowodowych

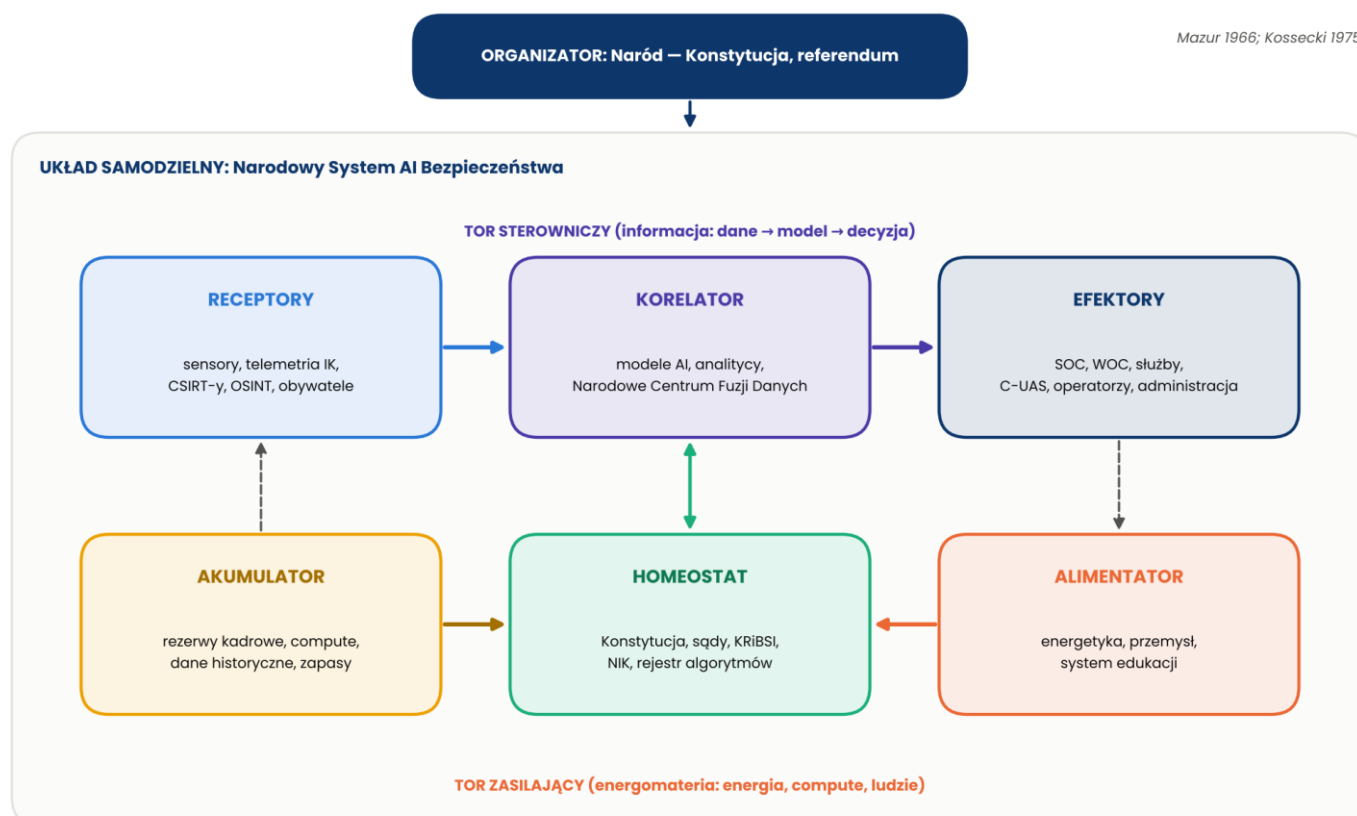
Teza naczelna strategii brzmi: w erze AI, robotów autonomicznych i biotechnologii bezpieczeństwo Polski wymaga głębokiej modernizacji i sanacji państwa, a nie dopisania nowych zadań do istniejących struktur. Teza ta jest uzasadniana sześcioma niezależnymi drogami. Każda wychodzi z innych przesłanek i posługuje się innym aparatem pojęciowym. To, co wynika z jednej teorii, bywa artefaktem tej teorii; to, co wynika z sześciu teorii naraz, jest własnością rzeczywistości.

11.1. Uzasadnienie cybernetyczne

Prawo wymaganej różnorodności. Ashby wykazał, że tylko różnorodność może pochłoniąć różnorodność: regulator musi dysponować repertuarem odpowiedzi nie mniejszym niż repertuar zakłóceń. AI dramatycznie zwiększyła różnorodność ataków — liczba wariantów phishingu, mutacji złośliwego kodu, narracji i trajektorii dronów rośnie szybciej niż liczba ludzi zdolnych na nie odpowiedzieć. Państwo ma dwie drogi: zwiększyć własną różnorodność (AI po stronie obrony, Cyber-Rezerwa, współpraca z sektorem prywatnym i obywatelami) albo zmniejszyć różnorodność zakłóceń (standaryzacja, eliminacja podatności podstawowych, zero trust). Strategia wymaga obu jednocześnie. Instytucje o stałym repertuarze, projektowane dla świata sprzed AI, z definicji nie spełniają prawa Ashby’ego — dlatego potrzebna jest zmiana strukturalna, a nie przyrostowa.

Twierdzenie o dobrym regulatorze. Conant i Ashby dowiedli, że każdy dobry regulator systemu musi być modelem tego systemu. Państwo bez wspólnego obrazu sytuacyjnego, bez mapy zależności krytycznych i bez rejestru własnych zdolności kadrowych nie jest modelem rzeczywistości, którą ma regulować. Narodowy Obraz Sytuacyjny, Rejestr Zdolności Krytycznych i mapa węzłów krytycznych nie są więc projektami informatycznymi, lecz technicznym wykonaniem warunku sterowalności.

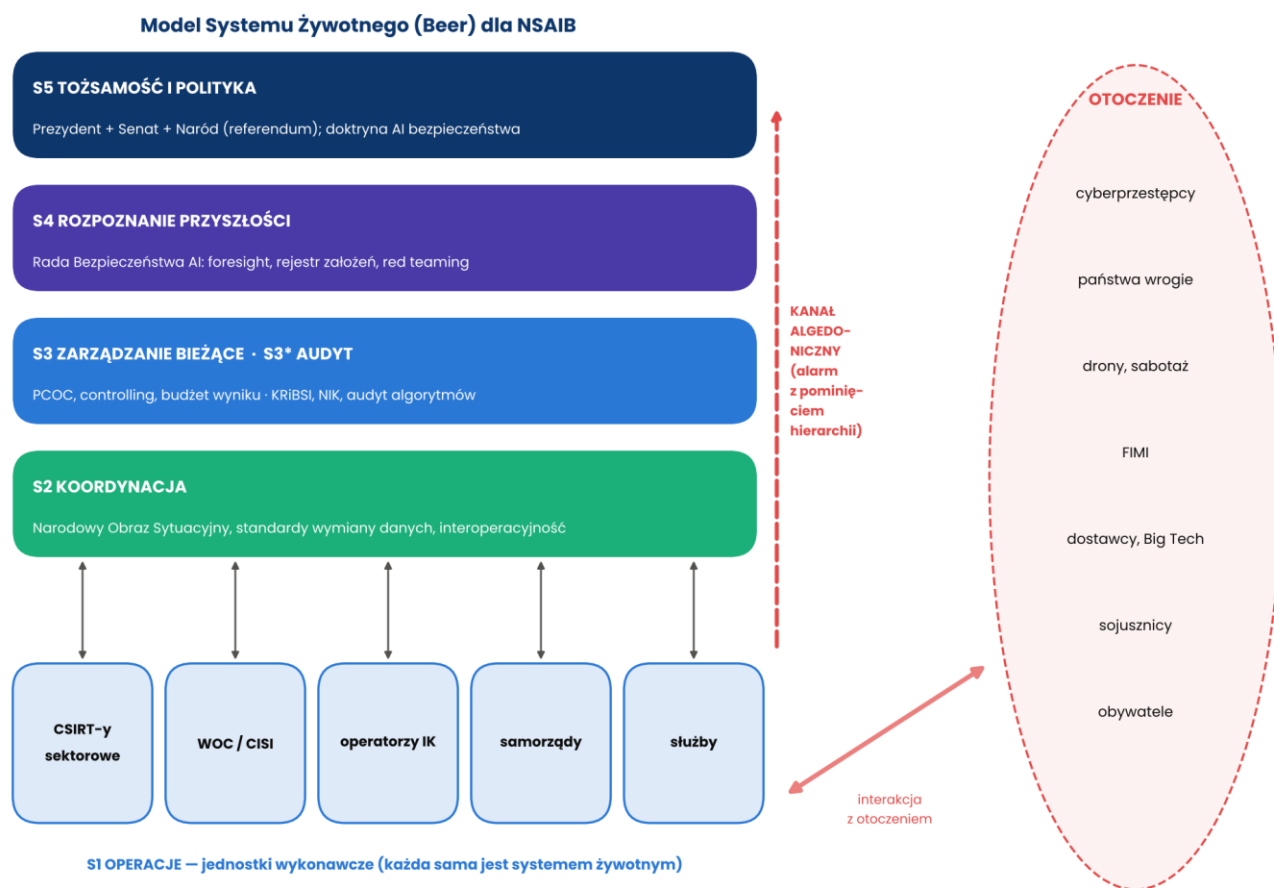
Układ samodzielny Mazura. Marian Mazur opisał układ samodzielny jako zdolny do przeciwdziałania utracie zdolności sterowania, złożony z receptorów, korelatora, efektorów, homeostatu, akumulatora i alimentatora. W domenie bezpieczeństwa AI receptory to sensory, telemetria i wywiad; korelator — modele AI, analitycy i centra fuzji; efekторы — SOC, wojsko, służby i operatorzy; homeostat — kontrola konstytucyjna, audyt, KRiBSI i sądy; akumulator — rezerwy kadrowe, compute, zapasy sprzętu i dane historyczne; alimentator — energetyka, przemysł i system edukacji. Brak któregośkolwiek bloku niszczy samodzielność układu, a nie tylko obniża jego sprawność.



Rysunek 17. Układ samodzielny Mazura w domenie bezpieczeństwa AI: bloki funkcjonalne, tor sterowniczy i tor zasilający Kosseckiego. Opracowanie własne.

Tor sterowniczy i tor zasilający. Józef Kossecki rozróżnił tor sterowniczy (informacja) i tor zasilający (energomateria). Państwo, które utraciło kontrolę nad torem sterowniczym, wykonuje cudze decyzje; państwo, które utraciło kontrolę nad torem zasilającym, wykonuje cudze warunki. W erze AI tor sterowniczy to dane, modele i algorytmy decyzyjne; tor zasilający — energia, compute i ludzie. Suwerenność jest pochodną kontroli obu torów jednocześnie. Kossecki wprowadził też miarę czasu decyzyjnego: struktura organizacyjna powinna minimalizować czas od bodźca do odpowiedzi. Obecna struktura bezpieczeństwa cyfrowego Polski, z wieloma niezależnymi ośrodkami, maksymalizuje ten czas na stykach.

Model Systemu Żywnotnego Beera. Każdy układ zdolny do przetrwania ma pięć podsystemów: operacje (S1), koordynację (S2), zarządzanie bieżące z audytem (S3, S3*), rozpoznanie przyszłości (S4) i tożsamość (S5). Diagnoza bezpieczeństwa cyfrowego Polski wskazuje rozwinięte S1 (liczne CSIRT-y, WOC, służby), słabe S2 (interoperacyjność), fragmentaryczne S3 (brak właściciela wyniku całości), bardzo słabe S4 (brak zinstytucjonalizowanego foresightu technologicznego i rejestru założeń krytycznych) i niezdefiniowane S5 dla domeny AI. Beer opisał też kanał algedoniczny — tor alarmowy omijający hierarchię. W cyberbezpieczeństwie pierwsza informacja o ataku przychodzi zwykle od operatora, administratora lub obywatela; kanał ten musi być projektowany, a nie pozostawiony przypadkowi.



Rysunek 18. Model Systemu Żywnotnego Beera zastosowany do Narodowego Systemu AI Bezpieczeństwa: S1–S5, kanał algedoniczny i otoczenie. Opracowanie własne.

Pętla OODA i przewaga decyzyjna. Boyd opisał walkę jako rywalizację cykli obserwacji, orientacji, decyzji i działania. AI skraca każdy etap cyklu — ale tylko w państwie, które ma dane, modele, uprawnienia decyzyjne na niskim szczeblu i zdolność wykonania. Szybki model w wolnej strukturze nie skraca pętli; przenosi wąskie gardło z analizy do zatwierdzenia.

TEZA CYBERNETYCZNA

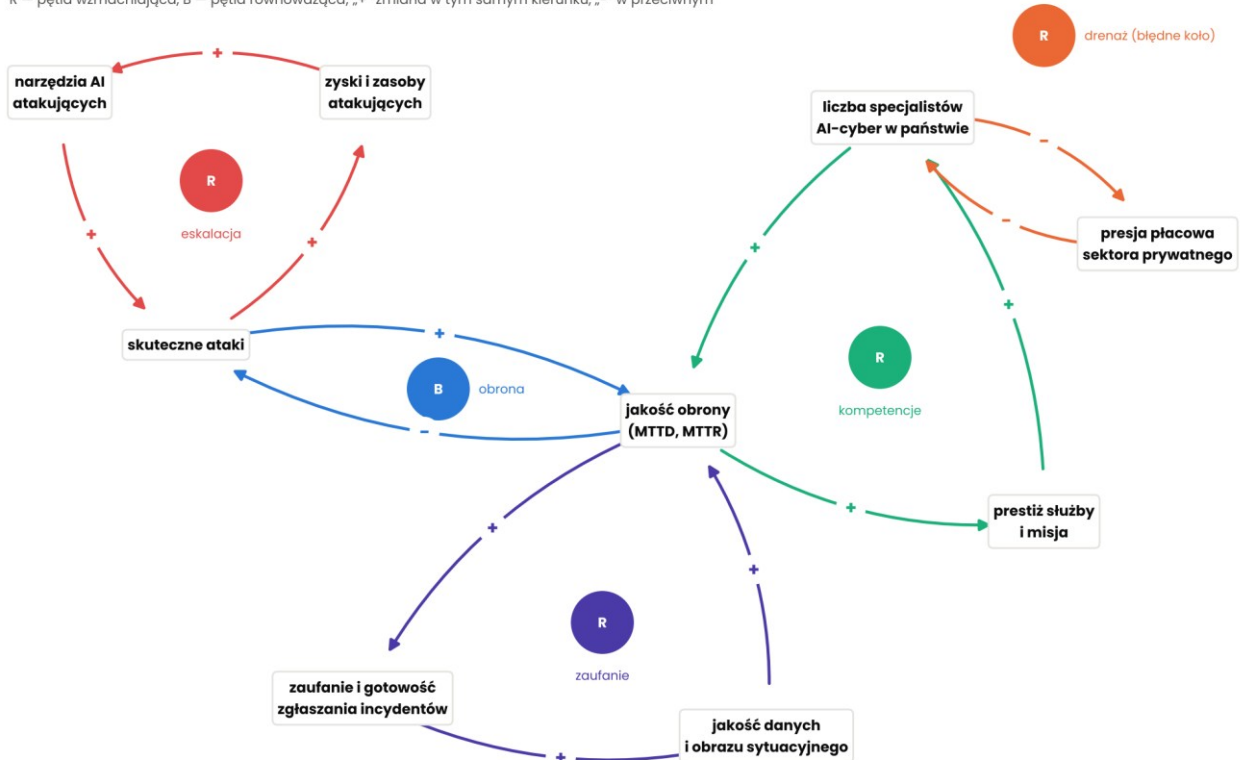
Bezpieczeństwo w erze AI jest sterowaniem, a sterowanie wymaga modelu, wymaganej różnorodności, kontroli obu torów i krótkiej pętli. Żadnego z tych warunków nie da się spełnić przez dopisanie zadań do istniejących struktur — wymagają one przebudowy architektury państwa.

11.2. Uzasadnienie systemowe

Pętle sprzężeń zwrotnych. System bezpieczeństwa cyfrowego zawiera pętle wzmacniające i równoważące. Pętla eskalacji (R) łączy skuteczne ataki ze wzrostem zysków przestępców, ich inwestycjami w narzędzia AI i kolejnymi atakami. Pętla kompetencji (R) łączy liczbę specjalistów z jakością obrony, prestiżem służby i napływem talentów. Pętla drenażu (R — błędne koło) łączy niedobór kadr z presją płacową i odpływem z sektora publicznego. Pętla obrony (B) jest jedyną pętlą równoważącą: skuteczne ataki mobilizują nakłady i poprawiają jakość obrony, co ogranicza kolejne ataki — ale działa z opóźnieniem, którego przeciwnik

używa. Pętla zaufania (R) łączy przejrzystość państwa z gotowością obywateli i firm do zgłaszania incydentów, co poprawia dane i skuteczność obrony.

R – pętla wzmacniająca, B – pętla równoważąca; „+” zmiana w tym samym kierunku, „-” w przeciwnym



Punkty dźwigni (Meadows): struktura informacji (dane, obraz sytuacyjny), reguły i cele (jakość obrony, zaufanie), paradygmat – talent jako infrastruktura krytyczna (specjaliści, misja); przerwanie błędnego koła drenażu

Rysunek 19. Diagram pętli przyczynowych systemu bezpieczeństwa AI: pętle eskalacji, kompetencji, drenażu i zaufania oraz punkty dźwigni. Opracowanie własne.

Punkty dźwigni. Donella Meadows wskazała, że najślabszymi punktami dźwigni są parametry (np. kwoty budżetu), a najsilniejszymi — struktura informacyjna, reguły systemu, jego cele i paradygmat. Strategia celowo uderza w punkty wysokiej dźwigni: zmienia strukturę przepływu informacji (Narodowy Obraz Sytuacyjny), reguły (obowiązek mierzenia i progów alarmowych), cel (od zgodności do zdolności) i paradygmat (talent jako infrastruktura krytyczna). Zwiększanie budżetów bez tych zmian działa na najślabszej dźwigni.

Awarie normalne i ściśle sprzężenie. Charles Perrow wykazał, że w układach złożonych i ściśle sprzężonych awarie są własnością konstrukcji. Cyfryzacja energetyki, wody, transportu i administracji zwiększa sprzężenie. Odpowiedź systemowa to rozluźnienie sprzężeń w punktach krytycznych: segmentacja sieci, tryby wyspowe, procedury ręczne, redundancja kadrowa i zapasowe centra decyzyjne.

Organizacje wysokiej niezawodności. Weick i Sutcliffe opisali organizacje utrzymujące zdatność mimo ciągłego ryzyka: zaabsorbowanie możliwością porażki, niechęć do upraszczania, wrażliwość na operacje, zaangażowanie w odporność i szacunek dla wiedzy eksperckiej ponad rangą. Centra cyberbezpieczeństwa państwa muszą działać według tych zasad — co wymaga zmian w kulturze, awansach i odpowiedzialności, a nie tylko w technologii.

Prawo Goodharta. Gdy miernik staje się celem, przestaje być dobrym miernikiem. Dlatego każdy KPI ma towarzyszący KRI wykrywający obchodzenie — np. spadek liczby zgłoszonych incydentów może oznaczać poprawę bezpieczeństwa albo ukrywanie zdarzeń.

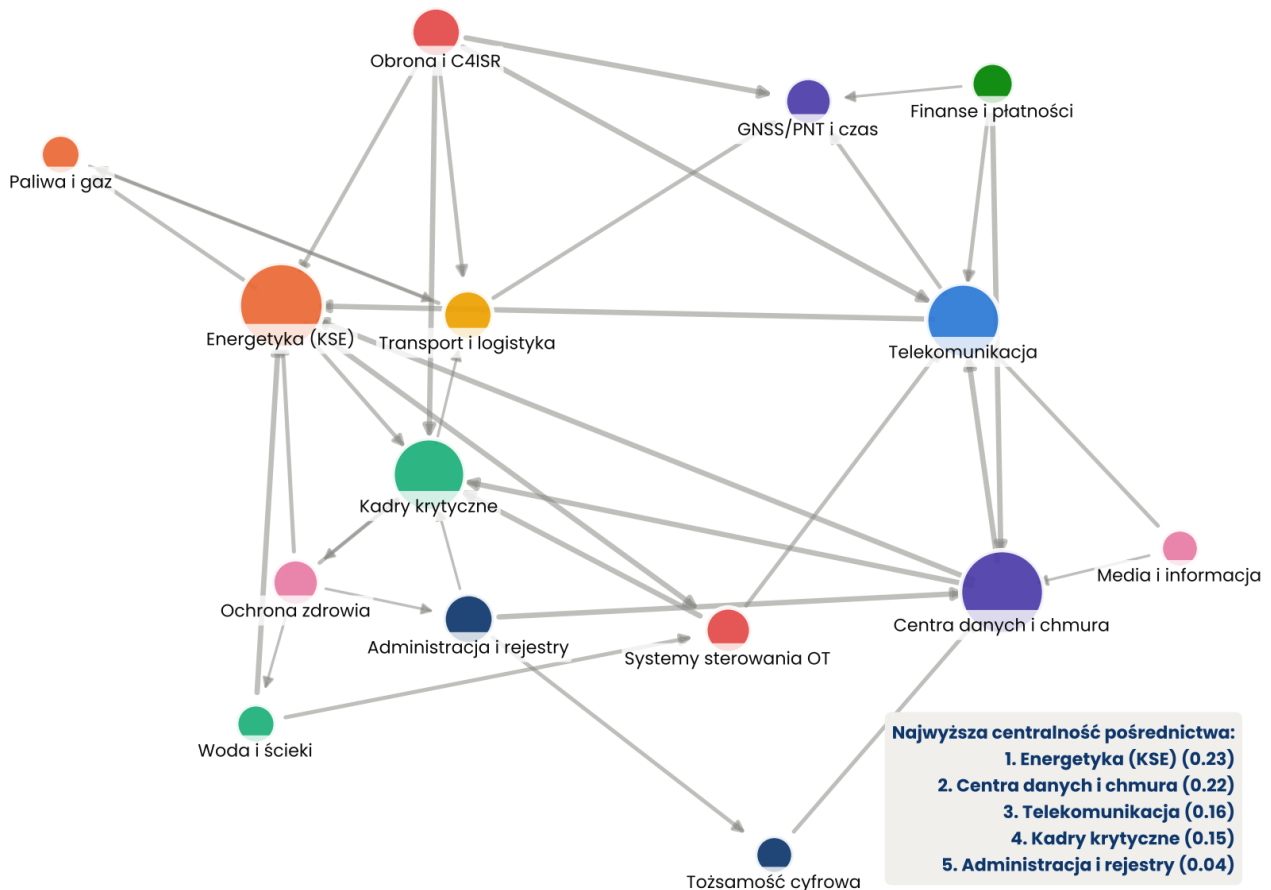
11.3. Uzasadnienie sieciowe

Sieci współzależne załamują się skokowo. Badania perkolacji w sieciach współzależnych (Buldyrev i in., 2010) pokazały, że awaria w jednej sieci może wywołać kaskadę prowadzącą do nagłego, nieciągłego załamania obu sieci. Energetyka i telekomunikacja, dane i finanse, woda i sterowanie — to klasyczne pary współzależne. Atak z grudnia 2025 r. na systemy sterowania ciepłowni i OZE jest empiryczną ilustracją tego mechanizmu.

Odporność na przypadek, kruchość na atak. Albert, Jeong i Barabási wykazali, że sieci zdominowane przez węzły centralne są odporne na losowe awarie i bardzo kruche na atak celowany. Przeciwnik państwowy nie atakuje losowo — prowadzi rozpoznanie i uderza w węzły o najwyższej centralności. Dlatego ochrona musi być celowana na węzły krytyczne, a nie rozłożona równomiernie.

Kadry jako węzeł krytyczny. Ilustracyjny model zależności piętnastu sektorów infrastruktury krytycznej (rysunek 20), policzony narzędziami analizy sieci, wskazuje cztery węzły o najwyższej centralności pośrednictwa: energetykę (0,23), centra danych i chmurę (0,22), telekomunikację (0,16) oraz kadry krytyczne (0,15). Zasób ludzki, którego żaden katalog infrastruktury krytycznej formalnie nie obejmuje, okazuje się węzłem o randze porównywalnej z siecią przesyłową — to sieciowe potwierdzenie tezy T5.

Mosty i dziury strukturalne. Ronald Burt pokazał, że węzły łączące odrębne klastry (mosty) mają nieproporcjonalny wpływ i są pojedynczymi punktami awarii. W systemie bezpieczeństwa cyfrowego mostami są: kilka zespołów eksperckich łączących sektor cywilny i wojskowy, kilku dostawców usług zarządzanych, kilka węzłów wymiany ruchu internetowego. Granovetter wskazał siłę słabych więzi: nowa informacja przychodzi przez kontakty spoza własnego klastra. Program Cyrkulacji Diaspory i współpraca z sektorem prywatnym to inwestycja w słabe więzi.

Graf zależności infrastruktury krytycznej (strzałka: „zależy od”)

Model ilustracyjny TechnoSanacji; wielkość węzła ~ betweenness centrality (NetworkX), grubość krawędzi ~ siła zależności.

Rysunek 20. Graf zależności infrastruktury krytycznej Polski: wielkość węzła odpowiada centralności pośrednictwa (betweenness), grubość krawędzi — siła zależności. Model ilustracyjny TechnoSanacji.

Odporność (2f+1). W systemach wieloagentowych sieć musi być odporna typu (2f+1), by przetrwać f nielojalnych lub przejętych węzłów bez utraty zdolności decyzyjnej całości. To policzalny wymóg redundancji, a nie postulat: każda funkcja krytyczna musi mieć co najmniej trzy niezależne ścieżki realizacji, jeżeli ma przetrwać przejęcie jednej.

TEZA SIECIOWA

Bezpieczeństwo państwa w erze AI jest własnością topologii sieci, a nie sumą zabezpieczeń poszczególnych instytucji. Topologii nie zmienia się przetargiem — zmienia się ją architekturą: mapą węzłów krytycznych, celowaną osłoną, redundancją ścieżek i mostami między klastrami.

11.4. Uzasadnienie technologiczne

Komplementarność talentu, danych i compute. Wydajność AI wynika z iloczynu trzech czynników: jakości danych, mocy obliczeniowej i kompetencji ludzi. Inwestycja w jeden czynnik bez pozostałych daje znikomy przyrost zdolności. Gigafabryka AI bez zastosowań bezpieczeństwa i bez kadr jest kosztownym obiektem; model językowy bez danych domenowych — ciekawostką; analitycy bez compute — kadrą, która wyjedzie tam, gdzie będzie mogła pracować.

Dual-use i wyścig adaptacji. W domenie, w której atakujący i obrońca używają tych samych narzędzi, przewagę daje szybkość adaptacji, a nie posiadanie technologii. Doświadczenie ukraińskie pokazuje cykle adaptacji dronów i walki radioelektronicznej liczone w tygodniach; zamówienia trwające latami zapewniają sprzęt przestarzały w dniu dostawy.

Zero trust i kryptografia postkwantowa. Architektura „nie ufaj, weryfikuj zawsze” (NIST SP 800-207) i migracja do kryptografii postkwantowej wymagają przebudowy systemów tożsamości, dostępu i komunikacji w całej administracji. Są to zmiany architektoniczne o horyzoncie wieloletnim — możliwe wyłącznie przy jednym właścicielu programu.

Ryzyka samej AI. Modele podlegają atakom specyficznym: zatrutowaniu danych, wstrzykiwaniu poleceń (prompt injection), ekstrakcji danych, halucynacjom i obchodzeniu zabezpieczeń. Systemy AI w bezpieczeństwie wymagają własnej warstwy bezpieczeństwa: testów odporności, red teamingu, monitoringu dryfu i możliwości natychmiastowego wyłączenia.

11.5. Uzasadnienie ustrojowe

Rozproszona egzekutywa. Diagnoza Rzeczypospolitej 2026 wskazała rozproszenie odpowiedzialności między ośrodkami władzy jako źródło dryfu strategicznego. W bezpieczeństwie cyfrowym rozproszenie jest szczególnie kosztowne, bo przeciwnik operuje w czasie minut, a spory kompetencyjne trwają miesiące. Projekt ustrojowy TechnoSanacji — zintegrowana egzekutywa z Prezydentem odpowiedzialnym za wykonanie, Prezydenckim Centrum Strategicznym jako ośrodkiem S4/S5 i controllingiem cybernetycznym — daje Narodowemu Systemowi AI Bezpieczeństwa jednego właściciela strategicznego.

Kontrola drugiego rzędu. Państwo zdolne bez kontroli regulatora produkuje państwo nadzoru. Dlatego projekt ustrojowy przewiduje równolegle: Senat stanowiący reguły i budżet, sąd chroniący legalność, prawa obywatelskie do wyjaśnienia decyzji i ludzkiej kontroli (human-in-command), rejestr użycia AI, jednokadencyjność i rotację elit ograniczające oligarchizację oraz referendum jako bezpiecznik ostateczny. Narzędzia AI bezpieczeństwa są dopuszczalne tylko w tej podwójnej architekturze.

Legalna transformacja. W obecnym porządku prawnym Rada Bezpieczeństwa AI może działać jako ciało analityczno-koordynacyjne przy legalnych organach państwa, opierając się na istniejących ustawach (KSC, ustawa o systemach AI, ustawa o ochronie ludności, ustawa o obronie Ojczyzny). Docelowe zmiany ustrojowe wymagają legalnej procedury i zatwierdzenia przez Naród.

11.6. Uzasadnienie społeczno-gospodarcze

Koszt braku działania. Oszustwa komputerowe stanowią 97% incydentów obsługiwanych przez CERT Polska; ich koszt ponoszą obywatele, w szczególności seniorzy i osoby mniej biegłe cyfrowo. Ataki na wodociągi, ciepłownie i szpitale uderzają w podstawowe usługi. Każdy złoty zainwestowany w higienę podstawową małych podmiotów infrastruktury krytycznej chroni usługi, od których zależy codzienne życie milionów ludzi.

Bezpieczeństwo jako warunek konkurencyjności. Inwestorzy w centra danych, przemysł zaawansowany i usługi cyfrowe oceniają nie tylko koszt pracy i energii, ale też bezpieczeństwo i przewidywalność. Państwo z wiarygodną cyberodpornością przyciąga kapitał; państwo kojarzone z udanymi atakami na infrastrukturę — odstrasza go. Strategia AI bezpieczeństwa jest więc także strategią gospodarczą.

Przemysł dual-use jako motor awansu. Zamówienia na counter-UAS, oprogramowanie obronne, systemy fuzji danych i cyberobronę mogą stać się pierwszym klientem dla polskich firm technologicznych i ścieżką awansu dla młodych inżynierów — zgodnie z Manifestem TechnoSanacji: Polska ma być miejscem, w którym młodzi nie muszą wyjeżdżać, by pracować przy technologiach przyszłości.

Sprawiedliwość cyfrowa. Automatyzacja bezpieczeństwa nie może obrócić się przeciw najsłabszym: osobom starszym, wykluczonym cyfrowo, mieszkańcom małych gmin. Strategia przewiduje symetrię narzędzi — każdej zdolności AI państwa odpowiada narzędzie ochrony obywatela (ostrzeżenia o oszustwach, weryfikacja tożsamości rozmówcy, prosty kanał zgłoszeń, kanał analogowy).

11.7. Test sześciu uzasadnień

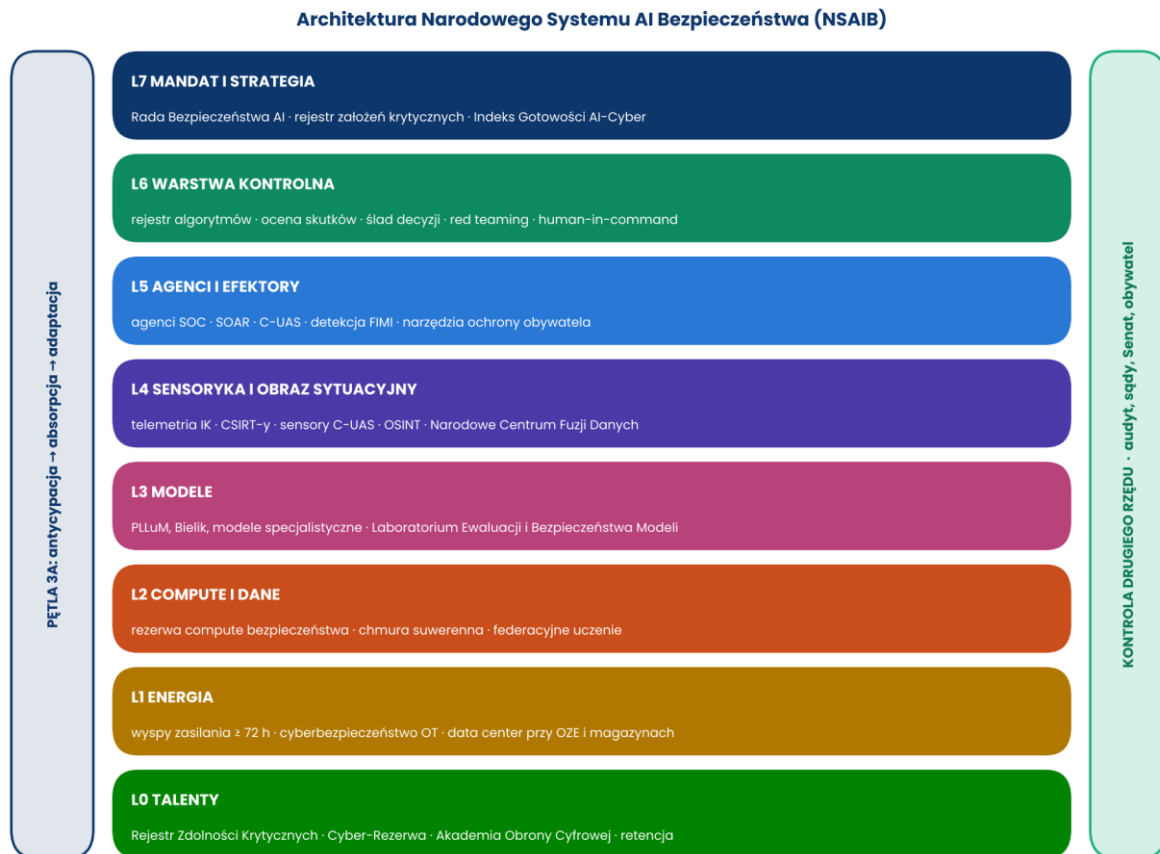
Element strategii	Cybern.	System.	Sieciowe	Technol.	Ustroj.	Spot.-gosp.
Rada Bezpieczeństwa AI (regulator S4/S5)	● dobry regulator, VSM	● punkt dźwigni: cel i reguły	● most między klastrami	○	● jeden właściciel wyniku	○
Narodowy Obraz Sytuacyjny	● model systemu	● struktura informacji	● fuzja warstw	● dane dla AI	○ wymaga kontroli	○
Talenty jako infrastruktura krytyczna	● różnorodność regulatora	● pętla R1/R2 vs R3	● węzły i mosty	● kompletność	● merytokracja	● awans i retencja
Tarcza AI-Cyber dla IK	● amplifikacja różnorodności	● rozluźnienie sprzężeń	● ostona węzłów	● agenci SOC	○	● ochrona usług
Energia dla AI	● tor zasilający	● redundancja	● sieci współzależne	● compute	○	● inwestycje
Program Odporności Kognitywnej	● ochrona toru poznawczego	● pętla zaufania	● bez progu epidemicznego w sieciach bezskalowych	● detekcja FIMI	● granica propagandy	● ochrona demokracji
Kontrola regulatora	● drugi rząd	● Goodhart	● odporność (2f+1)	● rejestr i audyt	● bezpieczniki	● zaufanie

Legenda: ● — uzasadnienie bezpośrednie; ○ — uzasadnienie pośrednie lub warunkowe. Element, który nie uzyskuje co najmniej czterech uzasadnień bezpośrednich, nie wchodzi do rdzenia strategii.

12. Architektura Narodowego Systemu AI Bezpieczeństwa

12.1. Osiem warstw systemu

Narodowy System AI Bezpieczeństwa (NSAIB) nie jest nową instytucją ani jednym systemem informatycznym. Jest architekturą łączącą istniejące instytucje, dane, infrastrukturę i ludzi w jedną pętlę sterowania. Składa się z ośmiu warstw — od fundamentu ludzkiego i energetycznego po mandat Narodu.



Rysunek 21. Architektura Narodowego Systemu AI Bezpieczeństwa: osiem warstw od talentów i energii po mandat Narodu, z warstwą kontrolną przenikającą całość. Opracowanie własne.

Warstwa	Zawartość	Właściciel	Funkcja w VSM
L0. Talenty	Rejestr Zdolności Krytycznych AI-Cyber, Akademia Obrony Cyfrowej, Cyber-Rezerwa Narodowa, pakiet retencyjny	Agencja Talentów / Rada Bezpieczeństwa AI	akumulator i alimentator
L1. Energia	zasilanie węzłów krytycznych, wyspy energetyczne, cyberbezpieczeństwo OT w energetyce	Minister Energii, operatorzy systemów	alimentator

Warstwa	Zawartość	Właściciel	Funkcja w VSM
L2. Compute i dane	rezerwa compute bezpieczeństwa, chmura suwerenna, standardy wymiany danych, federacyjne uczenie	Minister Cyfryzacji, NASK, PCSS, Cyfronet	akumulator, S2
L3. Modele	modele krajowe i specjalistyczne, Narodowe Laboratorium Ewaluacji i Bezpieczeństwa Modeli	KRiBSI, konsorcja naukowe	korelator
L4. Sensoryka i obraz sytuacyjny	telemetria operatorów, CSIRT-ów, wojska; sensory counter-UAS; OSINT; Narodowy Obraz Sytuacyjny	Narodowe Centrum Fuzji Danych	receptory, S4
L5. Agenci i efektory	agenci SOC, automatyczna reakcja, systemy counter-UAS, narzędzia wykrywania FIMI, narzędzia obywatela	PCOC, WOC, CSIRT-y, operatorzy	efektory, S1
L6. Warstwa kontrolna	rejestr algorytmów, ocena skutków, ślad decyzji, red teaming, stopniowany nadzór człowieka	KRiBSI, NIK, sądy, Rzecznik Praw Obywatelskich	homeostat, S3*
L7. Mandat i strategia	Rada Bezpieczeństwa AI przy Prezydenckim Centrum Strategicznym, rejestr założeń krytycznych, Indeks Gotowości AI-Cyber	Prezydent, Senat, Naród (referendum)	S3, S4, S5

12.2. Instytucje systemu

Rada Bezpieczeństwa AI. Ciało strategiczne przy Prezydenckim Centrum Strategicznym, w obecnym porządku prawnym — przy legalnych organach państwa jako zespół analityczno-koordynacyjny. Skład: przedstawiciele bezpieczeństwa narodowego, obrony, cyfryzacji, energetyki, zdrowia, służb, nauki i przemysłu, z niezależnymi ekspertami i obserwatorem z Senatu. Zadania: właściciel wyniku całości NSAIB, rejestr założeń krytycznych, Indeks Gotowości AI-Cyber, progi alarmowe, coroczny raport korekty, rozstrzyganie sporów kompetencyjnych w czasie godzin, nie miesięcy.

Narodowe Centrum Fuzji Danych. Operator Narodowego Obrazu Sytuacyjnego — wielowarstwowej reprezentacji stanu infrastruktury krytycznej, cyberprzestrzeni, przestrzeni informacyjnej i zagrożeń fizycznych. Działa w modelu federacyjnym: dane pozostają u właścicieli, a centrum łączy wskaźniki, sygnały i modele. Zasila Radę, PCOC, WOC, RCB i operatorów.

Narodowe Laboratorium Ewaluacji i Bezpieczeństwa Modeli. Krajowe centrum testowania modeli AI przed użyciem w zastosowaniach bezpieczeństwa i administracji: odporność na ataki, halucynacje, stronnictwo, wycieki danych. Współpracuje z europejskimi i sojuszniczymi instytutami bezpieczeństwa AI. Publikuje karty modeli dla systemów publicznych.

Agencja Talentów (komponent AI-Cyber). Operator Rejestru Zdolności Krytycznych, Wskaźnika Talentu Krytycznego i Cyber-Rezerwy Narodowej; koordynator Akademii Obrony Cyfrowej i Programu Cyrkulacji Diaspory.

Sandbox Obronny. Przestrzeń testowania innowacji dual-use w kontrolowanych warunkach, z uproszczoną ścieżką zamówień dla zwycięzców testów i cyklem decyzyjnym 90–180 dni, wzorowana na ukraińskiej platformie Brave1 i natowskiej DIANA, osadzona w ramach publikacji TechnoSanacji o sandboxie regulacyjnym.

12.3. Poziomy autonomii i stopniowany nadzór

Każdy system AI w NSAIB ma przypisany poziom autonomii i odpowiadający mu tryb nadzoru człowieka. Poziom zależy od odwracalności działania i jego skutków dla praw człowieka.

Poziom	Opis	Przykład w bezpieczeństwie	Tryb nadzoru
A0	brak AI	ręczna analiza incydentu	—
A1	AI informuje	podsumowanie alertów, tłumaczenie, wyszukiwanie	człowiek decyduje o wszystkim
A2	AI rekomenduje	priorytetyzacja incydentów, propozycja reguły blokady	człowiek zatwierdza (human-in-the-loop)
A3	AI działa odwracalnie, człowiek nadzoruje	automatyczna izolacja stacji roboczej, blokada domeny phishingowej	człowiek monitoruje i może cofnąć (human-on-the-loop)
A4	AI działa szybciej niż percepcja człowieka w granicach dyrektywy	obrona przed atakiem DDoS, zakłócanie drona w strefie chronionej	dyrektywa wyprzedzająca i audyt ex post (human-in-command)
A5	pełna autonomia bez nadzoru	—	niedopuszczalna wobec ludzi i decyzji nieodwracalnych

ZASADA HUMAN-IN-COMMAND

Decyzja o użyciu siły wobec człowieka, decyzja prawnie doniosła wobec obywatela i decyzja nieodwracalna pozostają w mandacie człowieka zdolnego zrozumieć, zatrzymać i skorygować system. Nadzór jest realny tylko wtedy, gdy człowiek ma czas, kompetencje, informację i prawo zatrzymania — obecność człowieka „w pętli” bez tych warunków jest fikcją odpowiedzialności.

13. Programy wykonawcze

Strategia realizowana jest przez dziesięć programów wykonawczych. Każdy ma właściciela, cel mierzalny, główne instrumenty i powiązanie ze strategiami TOWS.

Program	Cel mierzalny do 2030 r.	Główne instrumenty	TOWS
P1. Tarcza AI-Cyber dla infrastruktury krytycznej	100% podmiotów kluczowych objętych SOC z AI; MTDD < 1 h, MTTR < 24 h dla incydentów poważnych	zarządzane SOC dla małych podmiotów (wodociągi, ciepłownie, szpitale, gminy), agenci triażu, pakiet higieny podstawowej	ST1, WT3
P2. Suwerenny Stos AI	każda zdolność krytyczna ma wersję działającą w trybie odciętym; znany koszt wyjścia dla 100% systemów krytycznych	rezerwa compute bezpieczeństwa, modele krajowe i specjalistyczne, laboratorium ewaluacji, plany wyjścia	SO1, WT1
P3. Energia dla AI — AI dla energii	100% węzłów krytycznych z zasilaniem wyspowym ≥ 72 h; data center przy źródłach i magazynach	lokalne bilansowanie, magazyny BESS i wodór, cyberbezpieczeństwo OT, AI w sterowaniu siecią	WO3
P4. Narodowy Obraz Sytuacyjny	jeden federacyjny obraz stanu IK, cyberprzestrzeni i przestrzeni informacyjnej dla decydentów w czasie bliskim rzeczywistemu	Narodowe Centrum Fuzji Danych, standardy wymiany, kanał algedoniczny	WO2
P5. Autonomia z kontrolą i counter-UAS	koszt neutralizacji drona ≤ koszt drona; cykl adaptacji ≤ 4 tygodnie	drony przechwytyjące, fuzja tanich sensorów, AI klasyfikacji celów, produkcja krajowa, centrum adaptacji z Ukrainą	SO2, ST2
P6. Odporność Kognitywna	wykrywanie skoordynowanych kampanii w < 24 h; prebunking w programach szkolnych i mediach publicznych	detekcja FIMI i deepfake'ów, edukacja medialna, komunikacja kryzysowa, współpraca z platformami w ramach DSA	ST3
P7. Bio-AI	wczesne ostrzeganie epidemiologiczne z AI w każdym województwie; standard bezpieczeństwa badań dla laboratoriów	nadzór ścieków, syndromiczny i genomowy; kontrola dostępu do usług syntezy; ćwiczenia	WT4
P8. Narodowy System Talentów AI-Cyber	dotądnie saldo talentów krytycznych; współczynnik zastępowalności ≥ 1:3; Cyber-Rezerwa ≥ 5 tys. osób	Rejestr, WTK, Akademia Obrony Cyfrowej, pakiet retencyjny, Program Cyrkulacji Diaspory, dual-key	SO1, WO4, WT2
P9. Zamówienia Innowacyjne i Sandbox Obrony	≥ 30% budżetu „systemu nerwowego” dla krajowych dostawców; decyzja zakupowa ≤ 180 dni od testu	pierwszy klient, kontrakty za wynik, testy w sandboxie	SO3
P10. Kontrola Regulatora	100% systemów władczych w rejestrze algorytmów; roczny raport red teamingu instytucji	rejestr algorytmów, ocena skutków dla praw podstawowych, ślad decyzji, audyt, prawo do zakwestionowania	wszystkie

13.1. P1 — Tarcza AI-Cyber dla infrastruktury krytycznej

Najślabszym ogniwem infrastruktury krytycznej nie są duże koncerny energetyczne, lecz tysiące małych podmiotów: wodociągów gminnych, ciepłowni miejskich, szpitali powiatowych i urzędów. Nie mają one i nie będą miały własnych zespołów bezpieczeństwa. Program tworzy sieć zarządzanych centrów operacji bezpieczeństwa (SOC jako usługa publiczna), wykorzystujących agentów AI do triażu i korelacji, z analitykami w modelu regionalnym. Pierwszym krokiem jest pakiet higieny podstawowej, odpowiadający na dominujące wektory ataku: uwierzytelnianie wieloskładnikowe dla zdalnego dostępu, likwidacja usług RDP dostępnych z internetu, aktualizacja urządzeń brzegowych, kopie zapasowe offline. Program rozwija istniejące inicjatywy „Cyberbezpieczny Samorząd” i „Cyberbezpieczne Wodociągi”, zmieniając ich logikę z grantu na usługę ciągłą z mierzonym wynikiem.

13.2. P2 — Suwerenny Stos AI

Program zapewnia, że każda zdolność krytyczna ma wersję działającą bez dostępu do zewnętrznych usług. Obejmuje rezerwę mocy obliczeniowej dedykowaną bezpieczeństwu (w fabrykach AI i — jeżeli Polska ją uzyska — w gigafabryce AI, z klauzulą dostępu priorytetowego w kryzysie), rozwój modeli krajowych w wersjach specjalistycznych (analiza zagrożeń, analiza dokumentów, tłumaczenie z języków przeciwnika), Narodowe Laboratorium Ewaluacji i Bezpieczeństwa Modeli oraz obowiązkowe plany wyjścia dla systemów krytycznych. Zasada: nie autarkia, lecz znany koszt wyjścia i przetestowany tryb odcięty.

13.3. P3 — Energia dla AI — AI dla energii

Program łączy politykę infrastruktury AI z Doktryną Energetyczną TechnoRepubliki. Centra danych lokalizowane są przy źródłach OZE, magazynach i — w przyszłości — małych reaktorach modułowych, z udziałem w lokalnym bilansowaniu sieci (model analizowany w koncepcji AI Data Center w Łowej: integracja PV, wiatru, elektrolizerów i magazynów wodoru oraz baterii, z inteligentnym zarządzaniem obciążeniem zapewniającym ciągłość pracy przy ograniczeniu mocy obliczeniowej o nie więcej niż kilkanaście procent w sytuacji kryzysowej). Równolegle AI wspiera sterowanie siecią, prognozowanie i wykrywanie anomalii w systemach OT — pod warunkiem testów odporności na manipulację danymi. Węzły krytyczne NSAIB otrzymując zasilanie wyspowe na co najmniej 72 godziny.

13.4. P4 — Narodowy Obraz Sytuacyjny

Program buduje federacyjny obraz stanu państwa: dane pozostają u właścicieli, a centrum łączy wskaźniki zagrożeń, stan usług kluczowych, sygnały z przestrzeni informacyjnej i sensorów fizycznych. Obraz ma warstwy dostępu odpowiadające poziomom decyzyjnym i obywatelską warstwę jawną (stan zagrożeń, ostrzeżenia, zalecenia). Projektowany jest kanał algedoniczny — prosty, szybki tor alarmowy dla operatorów, administratorów i obywateli, omijający hierarchię.

13.5. P5 — Autonomia z kontrolą i counter-UAS

Program przenosi do polskiej obrony lekcje wojny w Ukrainie: obrona przed tanimi dronami musi być tania, warstwowa i szybko adaptowana. Obejmuje sieć tanich sensorów (akustycznych, radiowych, optycznych) z fuzją danych wspieraną AI, drony przechwytyjące, środki walki radioelektronicznej i — w perspektywie — broń energii skierowanej. Kluczowe jest centrum adaptacji, w którym informacja zwrotna z ćwiczeń i od partnerów ukraińskich zamienia się w aktualizację oprogramowania w cyklu tygodniowym. Wszystkie systemy działają w reżimie human-in-command zgodnie z poziomami autonomii.

13.6. P6 — Odporność Kognitywna

Program łączy wykrywanie operacji wpływu (detekcja skoordynowanych zachowań nieautentycznych, deepfake'ów, sieci botów) z budową odporności społecznej: prebunkingiem z regularnym „przypominaniem”, edukacją medialną w szkołach i dla seniorów, wiarygodną komunikacją kryzysową. Działa według zasady granicznej: państwo ujawnia i wyjaśnia wrogie operacje, ale nie ocenia poglądów i nie prowadzi własnej manipulacji. Szczególną ochroną objęte są procesy wyborcze.

13.7. P7 — Bio-AI

Program buduje zdolność wczesnego ostrzegania przed zagrożeniami biologicznymi w podejściu One Health: nadzór ścieków, nadzór syndromiczny i genomowy wspierany AI, wymiana danych z sieciami europejskimi. Obejmuje standard bezpieczeństwa badań dla laboratoriów i usług syntezy, zgodny z zaleceniem Rady UE w sprawie bezpieczeństwa badań, oraz ćwiczenia międzyresortowe. Strategia nie zajmuje się szczegółami technicznymi zagrożeń — zajmuje się zdolnością państwa do ich wczesnego wykrycia i odpowiedzi.

13.8. P8 — Narodowy System Talentów AI-Cyber

Program wdraża instrumenty opisane w rozdziale 6: Rejestr Zdolności Krytycznych AI-Cyber, Wskaźnik Talentu Krytycznego, plan ciągłości kompetencyjnej, Cyber-Rezerwę Narodową, Akademię Obrony Cyfrowej, pakiet retencyjny służby cyfrowej, Program Cyrkulacji Diaspory, szybką ścieżkę kwalifikacji i model dual-key. Jest programem o najwyższym priorytecie, bo wszystkie pozostałe od niego zależą.

13.9. P9 — Zamówienia Innowacyjne i Sandbox Obrony

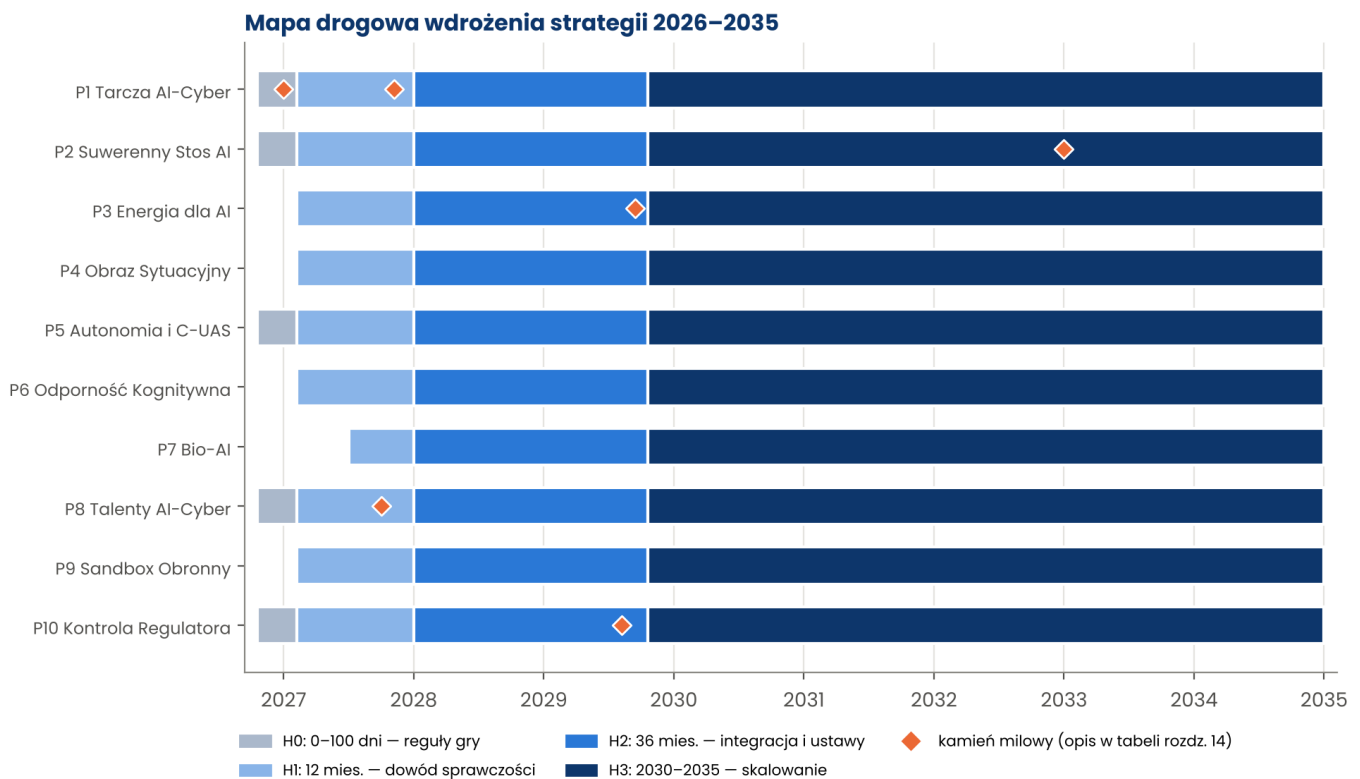
Program skraca drogę od prototypu do zdolności. Wprowadza ścieżkę „test — decyzja — kontrakt” z decyzją zakupową nie później niż 180 dni od pozytywnego testu, kontrakty za wynik (płatność za osiągnięty efekt, nie za dostarczony sprzęt) i jawną kategorię budżetową „system nerwowy” obejmującą C4ISR, cyber, AI i dane. Państwo staje się pierwszym klientem dla krajowych firm technologicznych, co jednocześnie buduje suwerenność i tworzy miejsca pracy dla talentów.

13.10. P10 — Kontrola Regulatora

Program zapewnia, że NSAIB pozostaje pod kontrolą Narodu. Obejmuje obowiązkowy rejestr algorytmów publicznych (dostępny także przez interfejs maszynowy), ocenę skutków dla praw podstawowych przed wdrożeniem systemów władczych, ślad decyzji, prawo do wyjaśnienia i zakwestionowania, niezależny audyt (NIK, KRiBSI) oraz coroczny red teaming instytucji — ćwiczenie, w którym niezależny zespół próbuje wykazać, jak system mógłby zostać nadużyty lub przechwycony.

14. Mapa drogowa 2026–2035

Wdrożenie przebiega w czterech horyzontach, zgodnie z logiką wdrożeniową TechnoSanacji: najpierw dowód sprawczości w obecnym prawie, następnie integracja, przygotowanie legalnych zmian i skalowanie tego, co działa.



Rysunek 22. Mapa drogowa wdrożenia strategii 2026–2035: dziesięć programów w czterech horyzontach. Opracowanie własne.

Horyzont	Cel	Działania kluczowe	Dowód dla obywatela
H0: 0–100 dni	ustawić reguły gry	powołanie Rady Bezpieczeństwa AI jako zespołu koordynacyjnego; inwentaryzacja zdolności i baseline KPI/KRI; mapa 20–50 węzłów krytycznych; pakiet higieny podstawowej dla małych podmiotów IK; pilotaż Rejestru Zdolności Krytycznych	jawne: kto odpowiada, za co, według jakiego miernika
H1: 12 miesięcy	udowodnić sprawczość w obecnym prawie	pierwsze regionalne SOC jako usługa; pilotaż Narodowego Obrazu Sytuacyjnego; Cyber-Rezerwa (pierwsze 1000 osób); pakiet retencyjny; Sandbox Obronny; laboratorium ewaluacji modeli; ćwiczenie blackout + cyber	krótszy czas reakcji na incydenty, pierwsze publiczne raporty Indeksu Gotowości
H2: 36 miesięcy	połączyć system i przygotować legalne zmiany	federacyjny obraz sytuacyjny dla wszystkich sektorów; rezerwa compute bezpieczeństwa; zasilanie wysp węzłów krytycznych; centrum adaptacji counter-UAS; pakiet ustawowy NSAIB; plan migracji do PQC	jedna sprawa obywatela przechodzi przez system bez przerzucania między silosami; mniej skutecznych oszustw
H3: 2030–2035	skalować przewagę	pełna integracja NSAIB z ustrojem TechnoRepubliki; eksport zdolności dual-use; Polska jako węzeł obliczeniowy regionu; migracja PQC zakończona w systemach krytycznych	państwo trudne do sparaliżowania i drogie do zaatakowania; dodatnie saldo talentów

KRYTERIUM SKALOWANIA

Rozwiązanie przechodzi do kolejnego horyzontu tylko wtedy, gdy jednocześnie poprawia wynik, chroni prawa obywatela, ogranicza ryzyko, zachowuje interoperacyjność i posiada plan wyłączenia. Pilotaż bez ewaluacji kontrfaktycznej nie jest dowodem.

15. Monitoring i ewaluacja: Indeks Gotowości AI-Cyber

15.1. Architektura pomiaru

Strategia bez pomiaru jest deklaracją. Indeks Gotowości AI-Cyber (IGAC) jest publikowanym corocznie, nierankingującym pulpitem stanu Narodowego Systemu AI Bezpieczeństwa. Składa się z sześciu wymiarów odpowiadających torom formuły naczelnej i funkcjom Modelu 3A: szybkość pętli, suwerenność torów, talenty, odporność infrastruktury, odporność społeczna i kontrola regulatora. Każdy miernik ma definicję, źródło danych, właściciela, częstotliwość aktualizacji, próg ostrzegawczy, próg krytyczny i przypisany wariant awaryjny.

15.2. KPI i KRI

Wymiar	Wskaźnik wyniku (KPI)	Wskaźnik ryzyka (KRI)	Próg ostrzegawczy / krytyczny	Właściciel
Szybkość pętli	mediana czasu wykrycia (MTTD) i reakcji (MTTR) incydentów poważnych w podmiotach kluczowych	odsetek incydentów wykrytych przez stronę trzecią, a nie przez podmiot	> 30% / > 50%	PCOC, CSIRT-y
Szybkość pętli	czas od incyduentu o nowym wzorcu do aktualizacji standardu	liczba powtórzeń tego samego wektora ataku	> 3 / > 10 w roku	Rada Bezpieczeństwa AI
Suwerenność	odsetek zdolności krytycznych z przetestowanym trybem odciętym	udział systemów krytycznych bez planu wyjścia	> 20% / > 40%	Minister Cyfryzacji
Suwerenność	dostępna rezerwa compute bezpieczeństwa (PFLOPS)	czas uruchomienia rezerwy w kryzysie	> 24 h / > 72 h	NASK, PCSS, Cyfronet
Talenty	Wskaźnik Talentu Krytycznego; saldo talentów w rolach krytycznych	współczynnik zastępowalności; rotacja w SOC publicznych	< 1:3 i > 15% / < 1:2 i > 25%	Agencja Talentów
Talenty	liczebność i gotowość Cyber-Rezerwy (odsetek uczestników ćwiczeń)	odsetek rezerwistów nieaktywnych > 12 miesięcy	> 30% / > 50%	WOC, Agencja Talentów
Odporność IK	odsetek małych podmiotów IK objętych SOC jako usługą	odsetek podmiotów z RDP dostępnym z internetu lub bez MFA	> 5% / > 15%	Minister Cyfryzacji

Wymiar	Wskaźnik wyniku (KPI)	Wskaźnik ryzyka (KRI)	Próg ostrzegawczy / krytyczny	Właściciel
Odporność IK	czas pracy węzłów krytycznych na zasilaniu wyspowym	węzły bez zasilania awaryjnego ≥ 72 h	$> 10\%$ / $> 25\%$	Minister Energii
Odporność IK	koszt neutralizacji drona względem kosztu drona	głębokość magazynka środków counter-UAS	< 30 dni / < 10 dni intensywnych działań	MON
Odporność społeczna	odsetek obywateli rozpoznających typowe techniki oszustw i manipulacji	wolumen skutecznych oszustw na 100 tys. mieszkańców	wzrost $> 10\%$ / $> 25\%$ r/r	NASK, MEN
Odporność społeczna	czas wykrycia i ujawnienia skoordynowanej kampanii FIMI	zasięg kampanii przed ujawnieniem	> 48 h / > 7 dni	Narodowe Centrum Fuzji Danych
Kontrola regulatora	odsetek systemów władczych w rejestrze algorytmów z oceną skutków	skargi uwzględnione na decyzje wspierane AI	wzrost $> 20\%$ / $> 50\%$	KRiBSI, RPO

15.3. Mechanizmy kontroli i korekty

Każdy próg krytyczny uruchamia procedurę: raport do Rady Bezpieczeństwa AI w ciągu 72 godzin, plan korekty w ciągu 30 dni, weryfikację po 90 dniach. Programy, które przez dwa kolejne cykle nie poprawiają swoich wskaźników, są przebudowywane lub wygaszane, a ich środki realokowane. Ewaluacja kontrfaktyczna — porównanie z grupą kontrolną lub scenariuszem bazowym — jest obowiązkowa dla wszystkich programów o budżecie powyżej progu ustalonego przez Radę. Wyniki ewaluacji wracają do decyzji budżetowej kolejnego roku, domykając pętlę na alokacji zasobów zgodnie z zasadami controllingu cybernetycznego TechnoRepubliki.

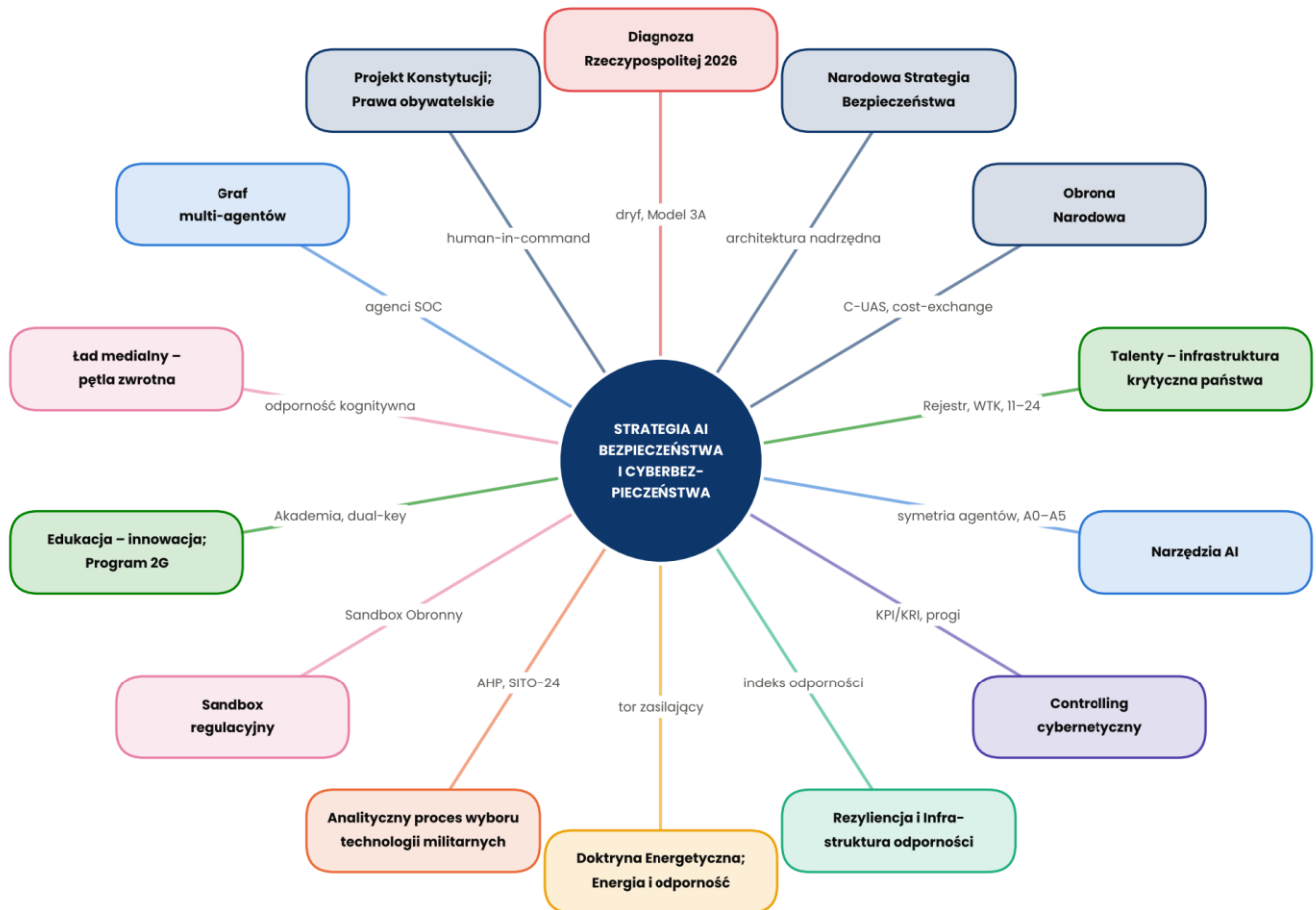
16. Ryzyka wdrożenia i bezpieczniki

Ryzyko	Mechanizm	Bezpiecznik
Państwo nadzoru	narzędzia wykrywania zagrożeń używane wobec obywateli i opozycji	podstawa ustawowa, rejestr algorytmów, ocena skutków, sądowa kontrola, rozdział wykrywania zagrożeń od oceny poglądów, coroczny red teaming instytucji
Przechwycenie regulatora (capture)	dostawcy, lobby lub sieci towarzyskie przejmują Radę i programy	jawność beneficjentów, rotacja, zakaz „drzwi obrotowych” w okresie karencji, otwarte standardy, kryterium przenaszalności w zamówieniach
Zgodność zamiast zdolności	wdrożenie przepisów jako dokumentacji	KPI czasu reakcji i wyników ćwiczeń zamiast liczby polityk; audyt efektu
Nowa zależność od jednego dostawcy	szybkie wdrożenia oparte na jednej platformie	plany wyjścia, portfel dostawców, modele możliwe do uruchomienia krajowo
Automatyzacja błędu	AI zwiększa przepustowość decyzji, ale także błędów	poziomy autonomii A0–A5, testy, monitoring dryfu, natychmiastowe wyłączenie
Drenaż talentów z programów publicznych	sektor prywatny i zagranica oferują więcej	pakiet retencyjny, misja, compute, model mieszany (rezerwa, kontrakty eksperckie)
Fragmentacja budżetowa	środki rozproszone między resortami bez wspólnego celu	kategoria budżetowa „system nerwowy”, jeden pulpit wyniku
Kryzys zaufania	incydent nadużycia AI przez państwo	przejrzystość, szybkie ujawnienie, niezależne dochodzenie, prawo do zadośćuczynienia
Spór prawny i konstytucyjny	kontrola TK, spory kompetencyjne	rozdzielenie działań w obecnym prawie, wymagających ustaw i wymagających zmian konstytucyjnych
Optymizm technologiczny	przecenianie AI i niedocenywanie podstaw	zasada „najpierw higiena podstawowa”: MFA, aktualizacje, kopie zapasowe i segmentacja przed zaawansowanymi projektami

17. Spójność z programami TechnoRepubliki

Strategia AI bezpieczeństwa nie jest dokumentem samodzielny. Jest węzłem łączącym wiele publikacji programowych TechnoSanacji, z których każda dostarcza jej elementu architektury.

Strategia jako węzeł architektury programowej TechnoSanacji



Rysunek 23. Powiązania strategii AI bezpieczeństwa z dokumentami programowymi TechnoSanacji: wkłady i zależności. Opracowanie własne.

Dokument TechnoSanacji	Wkład do strategii	Wkład strategii do dokumentu
Diagnoza Rzeczypospolitej 2026	teza o dryfie strategicznym, Model 3A, patologie systemowe	domena AI-cyber jako najszybciej rosnąca luka dryfu
Narodowa Strategia Bezpieczeństwa	architektura nadrzędna, Prezydenckie Centrum Strategiczne, DIME+T+L, warianty W1–W4	rozwińnięcie wykonawcze w domenie AI i cyberprzestrzeni
Obrona Narodowa	wojsko nowego wzoru, cost-exchange, produkcja jako odstraszenie, AI z człowiekiem w dowodzeniu	program counter-UAS, autonomia z kontrolą, AI w C4ISR
Talenty – infrastruktura krytyczna państwa	formuła talentu jako infrastruktury, Rejestr, WTK, ścieżka 11–24	komponent AI-Cyber Narodowego Systemu Talentów, Cyber-Rezerwa

Dokument TechnoSanacji	Wkład do strategii	Wkład strategii do dokumentu
Narzędzia AI	symetria agentów, poziomy autonomii, warstwa kontrolna, rejestr algorytmów	zastosowanie w bezpieczeństwie; narzędzia ochrony obywatela przed oszustwami
Controlling cybernetyczny	KPI/KRI, progi, pętla na alokacji	Indeks Gotowości AI-Cyber
Rezyliencja i Infrastruktura odporności	Narodowy Indeks Odporności, 7 wymagań NATO, ochrona ludności	warstwa cyfrowa odporności, zasilanie wyspowe węzłów
Doktryna Energetyczna; Strategia Energii i Odporności	tor zasilający, lokalne bilansowanie, wodór	program Energia dla AI — AI dla energii
Analityczny proces wyboru technologii militarnych	AHP/MCDM wspierane AI, metoda SITO-24	kryteria wyboru technologii dla Sandboxu Obronnego
Sandbox regulacyjny	testowanie innowacji w bezpiecznej przestrzeni	Sandbox Obronny
Edukacja – innowacja; Program 2G	przedsiębiorcza edukacja, sojusz pokoleń	Akademia Obrony Cyfrowej, model dual-key
Ład medialny – pętla zwrotna	jakość informacji publicznej	Program Odporności Kognitywnej
Graf multi-agentów	architektura systemów wieloagentowych państwa	agenci SOC i warstwa kontrolna
Projekt Konstytucji; Prawa obywatelskie	mandat, granice władzy, human-in-command jako prawo	kontrola regulatora AI bezpieczeństwa

18. Hasła programowe

1. Nie wygrywa ten, kto ma więcej AI, lecz ten, kto szybciej domyka pętlę.
2. Talent to infrastruktura krytyczna — mierzymy go, chronimy i odtwarzamy jak sieć przesyłową.
3. Bez danych nie ma dowodzenia; bez energii nie ma danych; bez ludzi nie ma niczego.
4. Suwerenność to koszt wyjścia, nie autarkia.
5. Tani atak wymaga taniej obrony.
6. AI wykrywa i przyspiesza — człowiek decyduje i odpowiada.
7. Najpierw higiena, potem rakiety: MFA, aktualizacje, kopie i segmentacja ratują więcej niż najnowszy model.
8. Bronimy toru poznawczego Narodu — nie sterujemy nim.
9. Strategia bez progu alarmowego jest deklaracją.
10. Państwo trudne do sparaliżowania i drogie do zaatakowania.

19. Konkluzja programowa

Era sztucznej inteligencji, robotów autonomicznych i biotechnologii nie daje państwom średniej wielkości czasu na stopniową adaptację. Zmienia warunki brzegowe suwerenności szybciej niż zmieniają się instytucje. Polska ma zasoby, talenty, nowe ramy prawne i doświadczenia z pierwszej linii wojny hybrydowej. Brakuje jej regulatora, który połączyłby te elementy w jedną pętlę sterowania — z modelem rzeczywistości, wymaganą różnorodnością, kontrolą torów informacyjnego i zasilającego, krótkim czasem decyzyjnym i kontrolą drugiego rzędu.

Sześć niezależnych dróg uzasadnienia — cybernetyczna, systemowa, sieciowa, technologiczna, ustrojowa i społeczno-gospodarcza — prowadzi do tego samego wniosku: bezpieczeństwa Polski w erze AI nie da się zapewnić przez dopisanie nowych zadań do starych struktur. Wymaga ono modernizacji i sanacji państwa — przywrócenia mu zdolności widzenia, rozumienia, decydowania, wykonania i korekty. Ta sama przebudowa, która chroni przed cyberatakiem, chroni przed dryfem strategicznym w każdej innej domenie.

Centrum tej przebudowy są ludzie. Najdroższy system AI, najnowocześniejsze centrum danych i najsurowsza ustawa pozostaną martwe bez analityków, inżynierów, badaczy, operatorów i obywateli przygotowanych do działania. Dlatego strategia zaczyna się i kończy na talentach traktowanych jak infrastruktura krytyczna — rozpoznawanych wcześniej niż przez rynek globalny, rozwijanych krótszą ścieżką, chronionych przed zmarnowaniem i przechwyceniem, utrzymywanych w cyrkulacji z diasporą i światem, zasilanych compute, danymi i misjami, i mierzonych tak rygorystycznie, jak mierzy się stan każdego systemu krytycznego.

KONKLUZJA

TechnoRepublika nie konkuruje liczbą deklaracji, lecz szybkością pętli. Polska, która szybciej widzi, trafniej rozumie, wcześniej decyduje i skuteczniej koryguje - na własnych danych, własnej energii, własnym compute i z własnymi talentami, pod kontrolą Narodu - będzie państwem trudnym do sparaliżowania, drogim do zaatakowania i atrakcyjnym dla ludzi, którzy budują przyszłość.

Dołącz do rozmowy na [TechnoSanacja – państwo zdolne do działania](#) . Sprawdź założenia. Zgłoś ryzyko. Zaproponuj rozwiązanie. Dołącz kompetencje.

Aneks A. Słownik pojęć, terminów i skrótów

A.1. Pojęcia strategiczne i programowe TechnoSanacji

Pojęcie	Definicja i znaczenie w strategii
TechnoSanacja	Droga przemiany państwa: przywrócenie mu zdolności widzenia, rozumienia, decydowania, wykonania, pomiaru i korekty z wykorzystaniem technologii, pod kontrolą Narodu. Ruch programowy i portal wiedzy TechnoSanacja.pl.
TechnoRepublika	Cel przemiany: państwo zdolne do działania, które przewiduje, działa i odpowiada przed obywatelem; ustrój łączący sprawczość egzekutywy, demokrację bezpośrednią i kontrolę regulatora.
Formuła naczelna	Jednozdaniowe twierdzenie otwierające każdą publikację programową TechnoSanacji, wobec którego weryfikowana jest cała treść dokumentu.
Model 3A	Antycypacja — absorpcja — adaptacja: trzy równolegle utrzymywane funkcje odporności i sterowania, z odrębnymi miernikami, właścicielami i czasem reakcji.
Dryf strategiczny	Narastająca luka między tempem zmian otoczenia a tempem adaptacji strategii państwa (w ujęciu Ansoffa — strategiczne niedopasowanie).
Narodowy System AI Bezpieczeństwa (NSAIB)	Architektura ośmiu warstw (od talentów i energii po mandat Narodu) łącząca instytucje, dane, infrastrukturę i ludzi w jedną pętlę sterowania bezpieczeństwem w erze AI.
Rada Bezpieczeństwa AI	Proponowane ciało strategiczne przy Prezydenckim Centrum Strategicznym, właściciel wyniku całości NSAIB, rejestru założeń krytycznych i Indeksu Gotowości AI-Cyber.
Narodowy Obraz Sytuacyjny	Federacyjna, wielowarstwowa reprezentacja stanu infrastruktury krytycznej, cyberprzestrzeni, przestrzeni informacyjnej i zagrożeń fizycznych, dostępna decydentom w czasie bliskim rzeczywistemu.
Talent jako infrastruktura krytyczna	Zasada, zgodnie z którą kadry w dziedzinach krytycznych są inwentaryzowane, monitorowane, chronione, utrzymywane i odtwarzane jak infrastruktura krytyczna państwa — z rejestrem, wskaźnikiem stanu, planem ciągłości i właścicielem.
Rejestr Zdolności Krytycznych	Inwentarz specjalistów w rolach krytycznych (liczebność, kompetencje, wiek, rotacja, zastępowalność) — model stanu zasobu talentów.
Wskaźnik Talentu Krytycznego (WTK)	Złożony miernik stanu infrastruktury talentu w danej dziedzinie, z progami ostrzegawczymi i krytycznymi.
Współczynnik zastępowalności	Liczba osób zdolnych przejąć funkcję krytyczną w przeliczeniu na jedną osobę ją pełniącą; minimalny standard strategii 1:3.
Cyber-Rezerwa Narodowa	Przeszkoleni specjaliści cywilni (sektor prywatny, nauka, diaspora) mobilizowani w kryzysie na podstawie umów i regularnych ćwiczeń.
Akademia Obrony Cyfrowej	Sieć kształcenia łącząca uczelnie, WOC, NASK i przemysł: ścieżka 11–24, studia dualne, doktoraty wdrożeniowe, poligony cyber.
Ścieżka 11–24	Model wczesnego rozpoznania i rozwijania uzdolnień od 11. do 24. roku życia, zaprojektowany w programie talentów TechnoSanacji.
Model dual-key	Zasada pracy zespołowej łącząca młodego specjalistę (szybkość, narzędzia AI) z doświadczonym mentorem (kontekst, pamięć błędów, odpowiedzialność); rozbieżności są dokumentowane.
Program 2G	Program współpracy Generacji Z i seniorów jako regulatora nowej Rzeczypospolitej.

Pojęcie	Definicja i znaczenie w strategii
Tryb odcięty	Zdolność systemu krytycznego do działania bez dostępu do zewnętrznych usług (chmury, modeli, aktualizacji) — na krajowym compute, z krajowym modelem i lokalnym zasilaniem.
Koszt wyjścia	Koszt i czas migracji od danego dostawcy technologii bez utraty ciągłości działania; podstawowa miara suwerenności cyfrowej w strategii.
Indeks Gotowości AI-Cyber (IGAC)	Coroczny, nierankingujący pulpit stanu NSAIB w sześciu wymiarach: szybkość pętli, suwerenność, talenty, odporność IK, odporność społeczna, kontrola regulatora.
Sandbox Obronny	Przestrzeń testowania innowacji dual-use w kontrolowanych warunkach, z uproszczoną ścieżką zamówień dla zwycięzców testów.
System nerwowy (kategoria budżetowa)	Proponowana jawna kategoria wydatków obejmująca C4ISR, cyberbezpieczeństwo, AI i dane — odróżniona od wydatków na platformy.
Prezydenckie Centrum Strategiczne (PCS)	Zaprojektowane w Narodowej Strategii Bezpieczeństwa TechnoSanacji centrum foresightu, mapy oddziaływań, pulpitu KPI i red teamingu.
Teoria Oddziaływań Geopolitycznych	Autorski aparat dr. Andrzeja Węgrzyna: moc relacji = siła × częstotliwość; ryzyko = moc × wrogość × podatność; modelowanie grafowe relacji państw.
DIME+T+L	Instrumenty siły państwa: dyplomacja, informacja, wojsko, gospodarka, rozszerzone o technologię i prawo.

A.2. Pojęcia cybernetyczne, systemowe i sieciowe

Pojęcie	Definicja i znaczenie w strategii
Regulator	Układ utrzymujący wybrane zmienne w wyznaczonych granicach mimo zakłóceń; w strategii — państwo jako całość i jego podsystemy bezpieczeństwa.
Prawo wymaganej różnorodności (Ashby)	Tylko różnorodność może pochłoniąć różnorodność: regulator opanuje tyle klas zaburzeń, ile ma stanów odpowiedzi.
Twierdzenie o dobrym regulatorze (Conant–Ashby)	Każdy dobry regulator systemu musi być modelem tego systemu.
Układ samodzielny (Mazur)	Układ zdolny do przeciwdziałania utracie zdolności sterowania, złożony z receptorów, korelatora, efektorów, homeostatu, akumulatora i alimentatora.
Receptor / korelator / efektor	W ujęciu Mazura: odbiór informacji / przetwarzanie informacji w decyzje / wykonanie oddziaływań.
Homeostat	Blok utrzymujący równowagę układu; w państwie — kontrola konstytucyjna, sądy, audyt.
Akumulator / alimentator	Blok gromadzący zasoby / blok zasilający układ energią i materią.
Tor sterowniczy i tor zasilający (Kossecki)	Kanał informacji sterującej oraz kanał energomaterii; suwerenność wymaga kontroli obu.
Moc swobodna	Zasób, którym układ dysponuje na prowadzenie samodzielnej polityki (Kossecki).
Czas decyzyjny	Czas od zarejestrowania bodźca do wykonania odpowiedzi; parametr ustroju, który należy minimalizować.
Model Systemu Żywnotnego (VSM, Beer)	Model pięciu podsystemów koniecznych do przetrwania organizacji: S1 operacje, S2 koordynacja, S3 zarządzanie bieżące (S3* audyt), S4 rozpoznanie przyszłości, S5 tożsamość i polityka.

Pojęcie	Definicja i znaczenie w strategii
Kanał algedoniczny	Tor alarmowy omijający hierarchię, przenoszący informację o zdarzeniu krytycznym bezpośrednio do poziomu decyzyjnego (Beer).
Cybernetyka drugiego rzędu	Ujęcie, w którym obserwator i regulator są częścią obserwowanego systemu; uzasadnia kontrolę nad regulatorem.
Pętla OODA	Cykl obserwacja — orientacja — decyzja — działanie (Boyd); przewagę zdobywa ten, kto szybciej go domyka.
Sprzężenie zwrotne	Powrót informacji o wyniku do regulatora; pętla wzmacniająca (R) lub równoważąca (B).
Punkt dźwigni (Meadows)	Miejsce w systemie, w którym niewielka zmiana wywołuje dużą zmianę zachowania; najsilniejsze są struktura informacji, reguły, cele i paradygmat.
Awaria normalna (Perrow)	Awaria wynikająca z konstrukcji układu złożonego i ściśle sprzężonego, a nie z pojedynczego błędu.
Organizacja wysokiej niezawodności (HRO)	Organizacja utrzymująca zdatność mimo ciągłego ryzyka dzięki kulturze uważności (Weick, Sutcliffe).
Prawo Goodharta	Gdy miernik staje się celem, przestaje być dobrym miernikiem; uzasadnia parę KPI/KRI.
Autopojeza instytucjonalna	Samoreprodukcja instytucji (procedur, budżetów, kadr) niezależnie od wyniku publicznego.
Centralność pośrednictwa (betweenness)	Miara częstości, z jaką węzeł leży na najkrótszych ścieżkach między innymi węzłami; wskazuje węzły krytyczne dla przepływu.
Sieć bezskalowa	Sieć o potęgowym rozkładzie stopni, zdominowana przez nieliczne huby; odporna na losowe awarie, krucha na atak celowany (Barabási).
Perkolacja w sieciach współzależnych	Proces, w którym awaria w jednej sieci kaskaduje do sieci od niej zależnej, prowadząc do skokowego załamania (Buldyrev i in.).
Most i dziura strukturalna	Węzeł lub krawędź łącząca odrębne klastry; źródło wpływu i pojedynczy punkt awarii (Burt).
Odporność (2f+1)	Wymóg redundancji sieci, która ma przetrwać f niełojalnych lub przejętych węzłów bez utraty zdolności decyzyjnej.
Diagram pętli przyczynowych (CLD)	Graficzny model dynamiki systemu przedstawiający zmienne i znaki zależności tworzące pętle sprzężeń.

A.3. Pojęcia z zakresu AI, technologii i cyberbezpieczeństwa

Pojęcie	Definicja i znaczenie w strategii
Sztuczna inteligencja (AI)	Systemy maszynowe zdolne do wnioskowania na podstawie danych wejściowych, generowania prognoz, treści, rekomendacji lub decyzji; w rozumieniu AI Act — systemy działające z różnym poziomem autonomii.
AI agentowa (agent AI)	System AI samodzielnie planujący i wykonujący wieloetapowe zadania z użyciem narzędzi, systemów i interfejsów, w granicach nadanego mandatu.
Model ogólnego przeznaczenia (GPAI)	Model AI trenowany na dużych zbiorach danych, zdolny do wykonywania szerokiego zakresu zadań; podlega odrębnym obowiązkom AI Act.
Model o otwartych wagach	Model, którego parametry są udostępnione i mogą być uruchomione, audytowane i dostrojone na własnej infrastrukturze.

Pojęcie	Definicja i znaczenie w strategii
Duży model językowy (LLM)	Model AI przetwarzający i generujący język naturalny; krajowe przykłady: PLLuM, Bielik.
Compute	Moc obliczeniowa (akceleratorzy, centra danych, HPC) niezbędna do trenowania i uruchamiania modeli AI.
Fabryka AI / gigafabryka AI	Europejskie ośrodki mocy obliczeniowej dla AI finansowane w ramach EuroHPC; gigafabryki — największe instalacje (dziesiątki tysięcy akceleratorów).
Red teaming	Kontrolowane symulowanie działań przeciwnika w celu wykrycia słabości systemu, modelu lub instytucji.
Zatrutowanie danych	Atak polegający na wprowadzeniu zmanipulowanych danych do zbioru treningowego modelu w celu zmiany jego zachowania.
Wstrzykiwanie poleceń (prompt injection)	Atak polegający na ukryciu w danych wejściowych instrukcji, które przejmują kontrolę nad zachowaniem modelu lub agenta.
Deepfake	Syntetycznie wygenerowany lub zmanipulowany obraz, dźwięk lub wideo przedstawiający osoby lub zdarzenia w sposób fałszywy.
FIMI	Foreign Information Manipulation and Interference — zagraniczna manipulacja informacyjna i ingerencja; kategoria stosowana przez UE (EEAS).
Prebunking (inokulacja)	Upředzające uodpornianie odbiorców na techniki manipulacji przez ich wcześniejsze wyjaśnienie.
SOC	Security Operations Center — centrum operacji bezpieczeństwa monitorujące i reagujące na incydenty; w strategii także jako usługa publiczna dla małych podmiotów.
SOAR	Security Orchestration, Automation and Response — automatyzacja reakcji na incydenty.
MTTD / MTTR	Mean Time to Detect / Respond — średni czas wykrycia / reakcji na incydent.
OT / ICS	Operational Technology / Industrial Control Systems — technologie operacyjne i systemy sterowania przemysłowego (energetyka, woda, przemysł).
Zero trust	Architektura bezpieczeństwa, w której żaden użytkownik ani urządzenie nie jest domyślnie zaufany; każdy dostęp jest weryfikowany (NIST SP 800-207).
Kryptografia postkwantowa (PQC)	Algorytmy kryptograficzne odporne na ataki z użyciem komputerów kwantowych.
MFA	Uwierzytelnianie wieloskładnikowe.
RDP	Remote Desktop Protocol — protokół zdalnego pulpitu; udostępniony w internecie jest jednym z głównych wektorów ataku.
Ransomware	Złośliwe oprogramowanie szyfrujące dane w celu wymuszenia okupu.
Phishing	Wyłudzenie danych uwierzytelniających lub płatności przez podszywanie się pod zaufane podmioty.
Dostawca wysokiego ryzyka (DWR)	Dostawca sprzętu lub oprogramowania uznany w procedurze KSC za zagrożenie dla bezpieczeństwa; jego produkty podlegają wycofaniu z systemów kluczowych.
Counter-UAS (C-UAS)	Zdolność wykrywania, identyfikacji i neutralizacji bezzałogowych statków powietrznych.
BSP / UAV	Bezzałogowy statek powietrzny.

Pojęcie	Definicja i znaczenie w strategii
Walka radioelektroniczna (WRE / EW)	Działania w spektrum elektromagnetycznym: rozpoznanie, zakłócanie i ochrona własnych systemów.
C4ISR	Dowodzenie, kierowanie, łączność, informatyka, wywiad, obserwacja i rozpoznanie.
Cost-exchange ratio	Relacja kosztu środka obrony do kosztu środka ataku (lub odwrotnie); kluczowa miara ekonomii obrony przed dronami.
Głębokość magazynka	Liczba dostępnych środków obrony (np. interceptorów) względem przewidywanej skali ataku.
Human-in-the-loop / on-the-loop / in-command	Człowiek zatwierdza każdą decyzję / nadzoruje i może przerwać / zachowuje mandat i odpowiedzialność, wyznaczając granice działania systemu.
Rejestr algorytmów publicznych	Jawny, ustrukturyzowany rejestr systemów AI i algorytmów stosowanych przez administrację, z opisem celu, danych, nadzoru i właściciela.
Ocena skutków dla praw podstawowych (FRIA)	Wymagana przez AI Act ocena wpływu systemu AI wysokiego ryzyka na prawa podstawowe.
Dual-use	Technologia lub produkt podwójnego zastosowania — cywilnego i wojskowego.
One Health	Podejście łączące zdrowie ludzi, zwierząt i środowiska w nadzorze i odpowiedzi na zagrożenia biologiczne.
Nadzór ścieków	Monitoring patogenów w ściekach komunalnych jako wczesny sygnał zagrożeń epidemiologicznych.

A.4. Skróty

Skrót	Rozwinięcie
ABW	Agencja Bezpieczeństwa Wewnętrznego
AI Act	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 w sprawie sztucznej inteligencji
AHP	Analytic Hierarchy Process — analityczny proces hierarchiczny
BESS	Battery Energy Storage System — baterijny magazyn energii
CER	Dyrektywa (UE) 2022/2557 w sprawie odporności podmiotów krytycznych
CERT Polska	zespół reagowania na incydenty działający w NASK
CISI	Centrum Implementacji Sztucznej Inteligencji (Wojska Obrony Cyberprzestrzeni)
CRA	Cyber Resilience Act — rozporządzenie (UE) 2024/2847 o cyberodporności
CSIRT	Computer Security Incident Response Team (CSIRT GOV, CSIRT MON, CSIRT NASK)
DIANA	Defence Innovation Accelerator for the North Atlantic (NATO)
DKWOC	Dowództwo Komponentu Wojsk Obrony Cyberprzestrzeni
DORSZ	Dowództwo Operacyjne Rodzajów Sił Zbrojnych
DSA	Digital Services Act — akt o usługach cyfrowych
EDIP	European Defence Industry Programme
ENISA	Agencja Unii Europejskiej ds. Cyberbezpieczeństwa
EuroHPC	Wspólne Przedsięwzięcie w dziedzinie obliczeń wielkiej skali

Skrót	Rozwinięcie
HPC	High Performance Computing — obliczenia wielkiej skali
IGAC	Indeks Gotowości AI-Cyber
IK	infrastruktura krytyczna
IoC	Indicator of Compromise — wskaźnik naruszenia bezpieczeństwa
KPI / KRI	Key Performance Indicator / Key Risk Indicator — wskaźnik wyniku / wskaźnik ryzyka
KRiBSI	Komisja Rozwoju i Bezpieczeństwa Sztucznej Inteligencji
KSC	krajowy system cyberbezpieczeństwa (ustawa z 5 lipca 2018 r. z późn. zm.)
KSE	Krajowy System Elektroenergetyczny
MCDM	Multi-Criteria Decision Making — wielokryterialne podejmowanie decyzji
NASK	NASK – Państwowy Instytut Badawczy
NIK	Najwyższa Izba Kontroli
NIS2	Dyrektywa (UE) 2022/2555 w sprawie środków na rzecz wysokiego wspólnego poziomu cyberbezpieczeństwa
NIST	National Institute of Standards and Technology (USA)
NSAIB	Narodowy System AI Bezpieczeństwa
OSINT	Open Source Intelligence — wywiad z otwartych źródeł
PCOC	Połączone Centrum Operacyjne Cyberbezpieczeństwa
PCSS	Poznańskie Centrum Superkomputerowo-Sieciowe
PLLuM	Polish Large Language Model — polski model językowy rozwijany przez konsorcjum naukowe
PQC	Post-Quantum Cryptography — kryptografia postkwantowa
RCB	Rządowe Centrum Bezpieczeństwa
RPO	Rzecznik Praw Obywatelskich
SAFE	Security Action for Europe — unijny instrument pożyczek na zdolności obronne
SKW	Służba Kontrwywiadu Wojskowego
SMR	Small Modular Reactor — mały reaktor modułowy
TOWS	macierz strategii wyprowadzonych z analizy SWOT (SO, ST, WO, WT)
VSM	Viable System Model — Model Systemu Żywego
WOC	Wojska Obrony Cyberprzestrzeni
WTK	Wskaźnik Talentu Krytycznego

Aneks B. Bibliografia i korpus źródłowy

B.1. Cybernetyka i teoria systemów

1. Ashby W.R., *An Introduction to Cybernetics*, Chapman & Hall, London 1956.
2. Conant R.C., Ashby W.R., Every Good Regulator of a System Must Be a Model of That System, „International Journal of Systems Science” 1970, t. 1, nr 2, s. 89–97.
3. Beer S., *Brain of the Firm*, Allen Lane, London 1972; wyd. 2: Wiley, Chichester 1981.
4. Beer S., *The Heart of Enterprise*, Wiley, Chichester 1979.
5. Mazur M., *Cybernetyczna teoria układów samodzielnych*, PWN, Warszawa 1966.
6. Mazur M., *Cybernetyka i charakter*, PIW, Warszawa 1976.
7. Kossecki J., *Cybernetyka społeczna*, PWN, Warszawa 1975.
8. Kossecki J., *Elementy nowoczesnej wiedzy o wojnie i walce informacyjnej*, Wyższa Szkoła Pedagogiczna, Kielce 1996.
9. Wiener N., *Cybernetics: Or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge MA 1948.
10. von Foerster H., *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer, New York 2003.
11. Deutsch K.W., *The Nerves of Government: Models of Political Communication and Control*, Free Press, New York 1963.
12. Medina E., *Cybernetic Revolutionaries: Technology and Politics in Allende’s Chile*, MIT Press, Cambridge MA 2011.
13. Meadows D.H., *Leverage Points: Places to Intervene in a System*, The Sustainability Institute, Hartland 1999.
14. Meadows D.H., *Thinking in Systems: A Primer*, Chelsea Green, White River Junction 2008.
15. Perrow C., *Normal Accidents: Living with High-Risk Technologies*, Basic Books, New York 1984.
16. Weick K.E., Sutcliffe K.M., *Managing the Unexpected: Resilient Performance in an Age of Uncertainty*, Jossey-Bass, San Francisco 2007.
17. Boyd J.R., *Patterns of Conflict*, prezentacja niepublikowana, 1986; *Destruction and Creation*, 1976.
18. Ansoff H.I., *Strategic Management*, Macmillan, London 1979.
19. Holling C.S., Resilience and Stability of Ecological Systems, „Annual Review of Ecology and Systematics” 1973, t. 4, s. 1–23.

B.2. Nauka o sieciach

20. Barabási A.-L., *Linked: The New Science of Networks*, Perseus, Cambridge MA 2002.
21. Albert R., Jeong H., Barabási A.-L., Error and Attack Tolerance of Complex Networks, „Nature” 2000, t. 406, s. 378–382.

22. Buldyrev S.V., Parshani R., Paul G., Stanley H.E., Havlin S., Catastrophic Cascade of Failures in Interdependent Networks, „Nature” 2010, t. 464, s. 1025–1028.
23. Watts D.J., Strogatz S.H., Collective Dynamics of „Small-World” Networks, „Nature” 1998, t. 393, s. 440–442.
24. Granovetter M.S., The Strength of Weak Ties, „American Journal of Sociology” 1973, t. 78, nr 6, s. 1360–1380.
25. Burt R.S., *Structural Holes: The Social Structure of Competition*, Harvard University Press, Cambridge MA 1992.
26. Newman M.E.J., *Networks*, wyd. 2, Oxford University Press, Oxford 2018.
27. Rinaldi S.M., Peerenboom J.P., Kelly T.K., Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies, „IEEE Control Systems Magazine” 2001, t. 21, nr 6, s. 11–25.
28. Farrell H., Newman A.L., Weaponized Interdependence: How Global Economic Networks Shape State Coercion, „International Security” 2019, t. 44, nr 1, s. 42–79.

B.3. Sztuczna inteligencja, technologia i bezpieczeństwo

29. Karp A.C., Zamiska N.W., *The Technological Republic: Hard Power, Soft Belief, and the Future of the West*, Crown, New York 2025.
30. Scharre P., *Army of None: Autonomous Weapons and the Future of War*, W.W. Norton, New York 2018.
31. Scharre P., *Four Battlegrounds: Power in the Age of Artificial Intelligence*, W.W. Norton, New York 2023.
32. Kissinger H.A., Schmidt E., Huttenlocher D., *The Age of AI: And Our Human Future*, Little, Brown, New York 2021.
33. Suleyman M., Bhaskar M., *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma*, Crown, New York 2023.
34. Miller C., *Chip War: The Fight for the World's Most Critical Technology*, Scribner, New York 2022.
35. Russell S., *Human Compatible: Artificial Intelligence and the Problem of Control*, Viking, New York 2019.
36. Buchanan B., *The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics*, Harvard University Press, Cambridge MA 2020.
37. Rid T., *Active Measures: The Secret History of Disinformation and Political Warfare*, Farrar, Straus and Giroux, New York 2020.
38. Greenberg A., *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*, Doubleday, New York 2019.
39. Zuboff S., *The Age of Surveillance Capitalism*, PublicAffairs, New York 2019.
40. Mazzucato M., *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*, Anthem Press, London 2013.
41. National Security Commission on Artificial Intelligence, *Final Report*, Washington 2021.

42. Bengio Y. i in., *International AI Safety Report*, raport przygotowany na AI Action Summit, Paryż, styczeń 2025.
43. Roozenbeek J., van der Linden S., *The Psychology of Misinformation*, Cambridge University Press, Cambridge 2024.
44. van der Linden S., *Foolproof: Why Misinformation Infects Our Minds and How to Build Immunity*, W.W. Norton, New York 2023.

B.4. Cyberbezpieczeństwo i odporność — standardy i raporty

45. NIST, *Cybersecurity Framework (CSF) 2.0*, NIST CSWP 29, Gaithersburg 2024.
46. NIST, *Zero Trust Architecture*, SP 800-207, Gaithersburg 2020.
47. NIST, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, Gaithersburg 2023.
48. NIST, *Developing Cyber-Resilient Systems: A Systems Security Engineering Approach*, SP 800-160 Vol. 2 Rev. 1, Gaithersburg 2021.
49. NIST, *Post-Quantum Cryptography Standards: FIPS 203, FIPS 204, FIPS 205*, Gaithersburg 2024.
50. ENISA, *ENISA Threat Landscape 2025*, Ateny 2025.
51. ENISA, *Foresight Cybersecurity Threats for 2030*, Ateny 2023 (aktualizacja 2024).
52. CERT Polska / NASK, *Krajobraz bezpieczeństwa polskiego internetu. Raport roczny z działalności CERT Polska 2025*, Warszawa 2026.
53. CERT Polska / NASK, raporty roczne z działalności za lata 2020–2024, Warszawa 2021–2025.
54. NATO, *Revised NATO Artificial Intelligence Strategy*, Bruksela 2024.
55. NATO, *Strengthened Resilience Commitment oraz Seven Baseline Requirements for National Resilience*, Bruksela 2016–2023.
56. OECD, *OECD AI Principles*, Paryż 2019, aktualizacja 2024.
57. OECD, *Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions*, Paryż 2025.
58. Eurostat, *Use of Artificial Intelligence in Enterprises*, Statistics Explained, grudzień 2025.
59. Eurostat, *ICT Specialists in Employment*, Statistics Explained, 2026.
60. Komisja Europejska, *AI Continent Action Plan*, Bruksela, kwiecień 2025.
61. Komisja Europejska, *State of the Digital Decade 2025*, Bruksela 2025.
62. Rada UE, Zalecenie w sprawie zwiększenia bezpieczeństwa badań naukowych, Bruksela, maj 2024.
63. RAND Corporation, prace z zakresu odporności narodowej, odstraszania i wargamingu na wschodniej flance NATO (2022–2025), według korpusu destylatów TechnoSanacji.
64. CSIS, CNAS, RUSI, IISS — analizy wojny dronowej, cost-exchange, counter-UAS i bazy przemysłowo-obronnej (2024–2026), według destylatu Obrona.

B.5. Talenty, kadry i państwo

65. Saxenian A., *The New Argonauts: Regional Advantage in a Global Economy*, Harvard University Press, Cambridge MA 2006.
66. Lubinski D., Benbow C.P., Study of Mathematically Precocious Youth After 35 Years, „Perspectives on Psychological Science” 2006, t. 1, nr 4, s. 316–345.
67. Bell A., Chetty R., Jaravel X., Petkova N., Van Reenen J., Who Becomes an Inventor in America? The Importance of Exposure to Innovation, „Quarterly Journal of Economics” 2019, t. 134, nr 2, s. 647–713.
68. Heckman J.J., Skill Formation and the Economics of Investing in Disadvantaged Children, „Science” 2006, t. 312, s. 1900–1902.
69. Michels R., *Zur Soziologie des Parteiwesens in der modernen Demokratie*, Klinkhardt, Leipzig 1911.
70. Fukuyama F., What Is Governance, „Governance” 2013, t. 26, nr 3, s. 347–368.
71. Besley T., Persson T., *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton University Press, Princeton 2011.
72. (ISC)², *Cybersecurity Workforce Study*, edycje 2023–2024.

B.6. Akty prawne i dokumenty urzędowe

73. Ustawa z dnia 5 lipca 2018 r. o krajowym systemie cyberbezpieczeństwa (tekst jedn. z późn. zm.), w tym nowelizacja wdrażająca dyrektywę NIS2 podpisana 19 lutego 2026 r.
74. Ustawa z dnia 3 lipca 2026 r. o systemach sztucznej inteligencji.
75. Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym (tekst jedn. z późn. zm.).
76. Ustawa z dnia 11 marca 2022 r. o obronie Ojczyzny (tekst jedn. z późn. zm.).
77. Ustawa z dnia 5 grudnia 2024 r. o ochronie ludności i obronie cywilnej (Dz.U. 2024 poz. 1907 z późn. zm.).
78. Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie sztucznej inteligencji (AI Act), wraz ze zmianami pakietu „Digital Omnibus” (2026).
79. Dyrektywa Parlamentu Europejskiego i Rady (UE) 2022/2555 z dnia 14 grudnia 2022 r. (NIS2).
80. Dyrektywa Parlamentu Europejskiego i Rady (UE) 2022/2557 z dnia 14 grudnia 2022 r. w sprawie odporności podmiotów krytycznych (CER).
81. Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/2847 z dnia 23 października 2024 r. (Cyber Resilience Act).
82. Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. (Digital Services Act).
83. Rada Ministrów, *Strategia Cyberbezpieczeństwa Rzeczypospolitej Polskiej na lata 2025–2029*, przyjęta 10 marca 2026 r. (M.P. 2026 poz. 309).
84. Ministerstwo Cyfryzacji, *Polityka rozwoju sztucznej inteligencji w Polsce do 2030 roku*, aktualizacja listopad 2025 r.

85. Ministerstwo Cyfryzacji, *Przewodnik po sztucznej inteligencji dla administracji publicznej*, Warszawa 2026.
86. Ministerstwo Obrony Narodowej, *Resortowa strategia sztucznej inteligencji do roku 2039*, Warszawa 2024.
87. Rada Europy, *Ramowa konwencja o sztucznej inteligencji, prawach człowieka, demokracji i państwie prawa*, Wilno 2024.

B.7. Źródła programowe TechnoSanacji (dr Andrzej Węgrzyn)

88. *Manifest TechnoSanacji*, TechnoSanacja.pl, 14.07.2026.
89. *Diagnoza Rzeczypospolitej 2026 — państwo w dryfie strategicznym*, TechnoSanacja.pl, 2026.
90. *Narodowa Strategia Bezpieczeństwa. Tom 1. Ogólna strategia bezpieczeństwa*, TechnoSanacja.pl, 2026.
91. *Obrona Narodowa Rzeczypospolitej. Projekt obrony narodowej TechnoRepubliki opartej o technologię i obywateli*, TechnoSanacja.pl, 2026.
92. *Talenty – infrastruktura krytyczna państwa. Pozyskiwanie i zatrzymywanie talentów dla Rzeczypospolitej*, TechnoSanacja.pl, 2026.
93. *Narzędzia AI dla obywateli i administracji w epoce rewolucji agentowej*, wyd. 2, TechnoSanacja.pl, 16.09.2026.
94. *Controlling cybernetyczny TechnoRepubliki*, TechnoSanacja.pl, 2026.
95. *Rezyliencja — odporność państwa oraz Infrastruktura odporności*, TechnoSanacja.pl, 2026.
96. *Analityczny proces wyboru technologii militarnych*, TechnoSanacja.pl, wrzesień 2026.
97. *Doktryna Energetyczna TechnoRepubliki oraz Strategia energii i odporności*, TechnoSanacja.pl, 2026.
98. *Sandbox regulacyjny oraz Graf multi-agentów*, TechnoSanacja.pl, 2026.
99. *Mapa programowa i Mapa dziewięciu postulatów*, TechnoSanacja.pl, 2026.
100. Węgrzyn A., Energetyczne aspekty budowy AI Data Center, „Nowa Energia” 2025, nr 3(99), s. 28–30.
101. Węgrzyn A., Wodór infrastrukturą energetyczną serwerowni AI, „Nowa Energia” 2026, nr 1(102).
102. Węgrzyn A., Technologiczna rewolucja v. węglowy retrofit, „Nowa Energia” 2026, nr 3(104), s. 40–43.
103. Węgrzyn A., *TechnoRepublika. Strategia transformacji Polski*, TechnoRepublika.pl, 2026.

B.8. Destylaty i korpusy badawcze TechnoSanacji

104. *Narodowa Obrona TechnoRepubliki Polskiej i zarządzanie ryzykiem wojennym — destylat biblioteki Obrona* (611 PDF, 15 klastrów).

105. *Narodowa Strategia TechnoRepubliki Polskiej* — destylat biblioteki Geopolitics (540 PDF, 15 klastrów).
106. *Budowanie odporności społecznej, gospodarczej i technologicznej TechnoRepubliki* — destylat rundy RES (1725 PDF).
107. *Cybernetyka państwa, narodu i organizacji politycznej* — destylat biblioteki CybControl (623 PDF).
108. *Analiza systemowa państwa* — destylat biblioteki SYS (3231 PDF, 6 klastrów, w tym SYS-DIGITAL i SYS-DEF).
109. *Analiza sieciowa agend i struktur państwa* — destylat rundy GR (1667 PDF: teoria grafów, network science, GNN, grafy wiedzy, TDA).
110. *Wielokryterialny proces decyzyjny AHP* — destylat rundy AHP/MCDM.
111. *System multiagentowy TechnoRepubliki, Innowacje państwa TechnoRepublika, Od przedsiębiorczej edukacji do wdrożonej innowacji w RP* — destylaty projektu TechnoSanacja.

B.9. Źródła danych bieżących (dostęp: październik 2026 r.)

112. NASK, *Prawie 2 tys. zgłoszeń każdego dnia. Raport CERT Polska za 2025 rok*, nask.pl, 2026.
113. CyberDefence24, *CERT Polska podsumowuje rekordowy 2025 rok*, cyberdefence24.pl, 2026.
114. CyberDefence24, *Tusk o ataku na polską energetykę: nawet 500 tys. osób mogło zostać pozbawionych ciepła*, cyberdefence24.pl, 2026.
115. Portal Komunalny, *Ustawa o KSC podpisana: NIS2 wchodzi do polskiego prawa*, portalkomunalny.pl, luty 2026.
116. Portal Komunalny, *Rząd przyjął nową Strategię Cyberbezpieczeństwa RP*, portalkomunalny.pl, marzec 2026.
117. GDPR.pl, *Ustawa o systemach sztucznej inteligencji weszła w życie*, gdpr.pl, 2026.
118. Gibson Dunn, *EU AI Act Omnibus Agreement — Postponed High-Risk Deadlines and Other Key Changes*, 2026.
119. CyberDefence24, *Polska walczy o historyczną inwestycję w AI. Mocni konkurenci w przetargu UE*, cyberdefence24.pl, 2026.
120. Ministerstwo Cyfryzacji, *Polityka rozwoju sztucznej inteligencji w Polsce do 2030 roku — bezpieczne i odpowiedzialne wykorzystanie sztucznej inteligencji*, gov.pl, listopad 2025.
121. Ministerstwo Obrony Narodowej, *W Wojsku Polskim powstaje Centrum Implementacji Sztucznej Inteligencji*, gov.pl, 2024.
122. Eurostat, *20% of EU enterprises use AI technologies*, komunikat z 11 grudnia 2025 r.

Aneks C. Nota metodologiczna

C.1. Metoda opracowania

Strategię opracowano według standardu publikacji programowych TechnoSanacji: formuła naczelna → tezy → analizy wspierające → uzasadnienia wielodrogowe → architektura → programy → mierniki → słownik i bibliografia. Podstawą były destylaty korpusów badawczych TechnoSanacji (łącznie kilka tysięcy publikacji źródłowych w klastrach obrony, geopolityki, odporności, cybernetyki państwa, analizy systemowej i sieciowej), teksty programowe dr. Andrzeja Węgrzyna oraz aktualne dane urzędowe i statystyczne zweryfikowane według stanu na październik 2026 r. Materiał przygotowano w oparciu o bazę danych TechnoSanacji z wykorzystaniem narzędzi AI, pod redakcją i odpowiedzialnością autora.

C.2. Zasady interpretacji

Dokument rozdziela fakty (dane urzędowe i statystyczne, opatrzone źródłem), oceny eksperckie (oznaczone jako ocena TechnoSanacji — m.in. mapa zagrożeń, radar zdolności, graf zależności infrastruktury) i propozycje programowe. Rysunki o charakterze modelowym mają funkcję ilustracyjną i nie zastępują ilościowej analizy ryzyka na danych niejawnych. Propozycje ustrojowe opisują model docelowy, a nie obowiązujące prawo; ich wdrożenie wymaga legalnej procedury, kontroli zgodności i zatwierdzenia przez Naród. Dokument nie przypisuje odpowiedzialności konkretnym osobom i nie zastępuje opinii prawnej.

C.3. Zastrzeżenia do danych

Liczby incydentów CERT Polska odzwierciedlają zarówno zmianę skali zagrożeń, jak i zmianę zdolności wykrywania i zgłaszania; porównania wieloletnie należy interpretować ostrożnie. Szacunki niedoboru specjalistów cyberbezpieczeństwa pochodzą z analiz rynkowych i różnią się metodyką. Ocena skutków ataku na energetykę z grudnia 2025 r. różniła się w wypowiedziach przedstawicieli rządu; przyjęto ostrożną interpretację. Strategia celowo nie zawiera szczegółów technicznych zagrożeń biologicznych ani podatności — zajmuje się zdolnościami państwa, a nie technikami ataku.

C.4. Status i licencja

Wydanie pierwsze, 3 października 2026 r. Dokument programowy do debaty publicznej, eksperckiej i parlamentarnej. Materiał udostępniony na licencji Creative Commons — Uznanie autorstwa 4.0. Uwagi, korekty i propozycje: [TechnoSanacja.pl](https://technosanacja.pl).

