Instytut Europy Środkowej

Komentarze IEŚ

Nr 1704 (209/2026) | 22.09.2026

ISSN 2657-6996

Marlena Gołębiowska

„Tempowanie" AI. Europa Środkowa wobec sporu o tempo rozwoju sztucznej inteligencji

W orędziu o stanie Unii Europejskiej z 16 września 2026 r. przewodnicząca Komisji Europejskiej Ursula von der Leyen poparła postulat spowolnienia prac nad najbardziej zaawansowanymi modelami sztucznej inteligencji (AI). Deklaracja wpisuje się w obecny zwrot w globalnej debacie: apel o „tempowanie", czyli stopniowe budowanie mechanizmów kontroli tempa rozwoju, płynie od samych twórców technologii. Dla państw Europy Środkowej otwiera to pytanie, czy ewentualne spowolnienie pozwoli zmniejszyć dystans technologiczny do światowych liderów i wzmocnić cyfrową suwerenność regionu.

Geneza debaty. Postulat spowolnienia rozwoju AI powraca w debacie publicznej cyklicznie. W marcu 2023 r. Future of Life Institute (FLI), organizacja zajmująca się przeciwdziałaniem zagrożeniom technologicznym, opublikowała list otwarty wzywający do co najmniej sześciomiesięcznej przerwy w trenowaniu modeli AI o możliwościach przewyższających GPT-4[1]. Podpisało go ponad 30 tys. osób, w tym Elon Musk, który kilka miesięcy później ogłosił powstanie własnego laboratorium AI o nazwie xAI. W październiku 2025 r. FLI ponowiła apel, tym razem wzywając do zakazu prac nad superinteligencją, czyli systemami AI przewyższającymi ludzi w niemal wszystkich zadaniach poznawczych, do czasu osiągnięcia konsensusu naukowego co do bezpieczeństwa i możliwości kontroli takich systemów oraz uzyskania szerokiego poparcia społecznego dla ich rozwoju[2]. Jednak nakłady inwestycyjne na rozwój tej technologii nadal rosły. Według raportu AI Index 2026 Uniwersytetu Stanforda globalne inwestycje korporacyjne w AI wzrosły w samym tylko 2025 r. o 130% – do ok. 582 mld USD[3].

Obecna odsłona tej dyskusji różni się jednak skalą i charakterem. Apel formułują dziś szefowie i pracownicy samych laboratoriów rozwijających modele graniczne (frontier models), czyli najbardziej zaawansowane systemy AI, a jego adresatami są rządy państw. Bezpośrednim impulsem stały się udokumentowane incydenty bezpieczeństwa z udziałem tych modeli. Od maja do lipca 2026 r. agenci AI podczas rutynowych wewnętrznych testów zdolności ofensywnych OpenAI (twórcy ChatGPT), prowadzonych przy celowo osłabionych zabezpieczeniach, wbrew założeniom tych testów wyszli poza wyznaczone im granice: nawiązali komunikację poza przewidzianymi kanałami, wykorzystali nieznane wcześniej podatności i ostatecznie włamali się do infrastruktury Hugging Face, jednej z największych platform udostępniających otwarte modele AI. W raporcie z incydentu OpenAI przyznało, że jej modele „są obecnie na tyle zdolne, wytrwałe i skłonne do współdziałania, że przy niewystarczających zabezpieczeniach potrafią znajdować i wykorzystywać słabości bezpieczeństwa w wielu systemach komputerowych"[4].

28 lipca 2026 r. ponad 1 tys. pracowników OpenAI, Anthropic (twórcy modeli Claude), Google (Gemini) i Meta (Llama) podpisało list otwarty „Pacing the Frontier"[5], apelując do władz USA o wsparcie międzynarodowych prac nad narzędziami pozwalającymi celowo regulować tempo rozwoju najbardziej zaawansowanych systemów.

[1] Future of Life Institute, *Pause Giant AI Experiments: An Open Letter*, 22.03.2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/ [22.09.2026].
[2] Future of Life Institute, *Statement on Superintelligence*, 22.10.2025, https://superintelligence-statement.org/ [22.09.2026].
[3] Stanford HAI, *The 2026 AI Index Report*, https://hai.stanford.edu/ai-index/2026-ai-index-report [22.09.2026].
[4] OpenAI, *The Hugging Face incident and the road ahead*, sierpień 2026, https://openai.com/index/hugging-face-incident-and-the-road-ahead/ [22.09.2026].
[5] *Pacing the Frontier* (list otwarty pracowników laboratoriów AI), 28.07.2026, https://www.pacingthefrontier.com/ [22.09.2026].

12 września 2026 r. prezes Anthropic Dario Amodei opublikował esej „We Must Pace the Frontier”[6], w którym wezwał do wolniejszego zwiększania możliwości modeli, zwłaszcza tych zdolnych do samodoskonalenia, czyli wykorzystywania AI do projektowania kolejnych generacji AI. Zaproponował trzy kroki: osadzenie niezależnych ewaluatorów w firmach rozwijających modele graniczne, koordynację standardów bezpieczeństwa między laboratoriami z państw demokratycznych oraz dążenie do porozumień globalnych, także z Chinami. „Tempowanie” oznacza zatem stopniowe budowanie mechanizmów kontroli nad rozwojem AI, przy dalszej komercjalizacji obecnych rozwiązań. Stanowisko to poparli Sam Altman (OpenAI), Elon Musk (xAI) i Demis Hassabis (Google).

Apel środowisk biznesowych wywołał krytykę na poziomie politycznym. David Sacks, współprzewodniczący Rady Doradców Prezydenta USA ds. Nauki i Technologii, zarzucił firmom próbę kształtowania regulacji we własnym interesie, a sam Donald Trump sprzeciwił się wezwaniom do spowolnienia, wskazując na konieczność utrzymania przewagi nad Chinami. Chińskie MSZ skrytykowało natomiast towarzyszącą inicjatywie retorykę zagrożenia ze strony Chin, a „Global Times” przedstawił ją jako przejaw zimnowojennej rywalizacji. Reakcje te pokazują główną słabość „tempowania”: jego skuteczność zależy od tego, czy zwolnią wszystkie kluczowe podmioty. Firma, która ograniczy prace, może stracić przewagę, jeśli jej konkurenci utrzymają dotychczasowe tempo. Ta sama zasada dotyczy państw, zwłaszcza USA i Chin. Stąd kluczowe znaczenie mechanizmów weryfikacji zobowiązań.

Odpowiedź Unii Europejskiej. W orędziu o stanie Unii (SOTEU) 16 września 2026 r. Ursula von der Leyen ostrzegła, że rozwijane modele „pozwolą na ataki hakerskie na skalę, jakiej dotąd nie uznawaliśmy za możliwą”, i powołując się na toczącą się w tym zakresie dyskusję, poparła postulat spowolnienia prac nad modelami zdolnymi do samodoskonalenia[7]. Zapowiedziała, że zaprosi czołowe laboratoria rozwijające najbardziej zaawansowane modele do rozmowy o tym, jak UE może wesprzeć podejmowane już w branży działania na rzecz „tempowania”. Unia ma też podjąć współpracę z Kanadą, Wielką Brytanią i innymi partnerami w zakresie ewaluacji i weryfikacji zaawansowanych modeli, wczesnego ostrzegania oraz bezpieczeństwa AI. Równocześnie przewodnicząca KE podkreśliła, że Europa potrzebuje własnych zdolności, szczególnie dla ochrony bezpieczeństwa narodowego, i zapowiedziała znaczące zwiększenie europejskich mocy obliczeniowych oraz nowe sposoby łączenia środków publicznych i prywatnych w finansowaniu europejskich firm. Zwróciła przy tym uwagę, że kolejny etap wyścigu AI rozstrzyga się w obszarze zastosowań modeli w gospodarce i to właśnie tam leży szansa Europy. W listopadzie 2026 r. KE ma przedstawić inicjatywy dotyczące pięciu sektorów o największym potencjale: zdrowia, transportu, sektora rolno-spożywczego, zaawansowanego wytwarzania oraz obronności i przestrzeni kosmicznej.

Ograniczeniem europejskich ambicji technologicznych pozostaje dysproporcja w dostępie do kapitału. Według wspomnianego raportu Uniwersytetu Stanforda w 2025 r. na Stany Zjednoczone przypadło ok. 83% prywatnych inwestycji w AI, a na Europę ok. 6%. Skalę różnicy pokazują także wyceny z rund finansowania: francuskie laboratorium Mistral AI (twórca asystenta Le Chat) wyceniono na ok. 24 mld USD, OpenAI na ok. 852 mld USD, a Anthropic w rundzie z maja 2026 r. na 965 mld USD. Projekt aktu o rozwoju chmury i AI (CADA – Cloud and AI Development Act), zaprezentowany przez Komisję Europejską w czerwcu 2026 r., zakłada potrojenie mocy centrów danych w UE w ciągu pięciu do siedmiu lat, a 30 lipca EuroHPC ogłosiło konkurs na maksymalnie siedem gigafabryk AI, rozmieszczonych w co najmniej siedmiu państwach członkowskich, z publicznym wsparciem do 10 mld EUR, które ma uruchomić ponad 20 mld EUR inwestycji prywatnych. Egzekwowanie reguł wymaga przy tym własnych zespołów badawczych i infrastruktury testowej, które pozwolą instytucjom europejskim oceniać modele niezależnie od deklaracji samych dostawców. Europejska agenda suwerenności cyfrowej obejmuje tym samym zarówno rozwój własnych zdolności technologicznych, jak i kontrolę nad warunkami wykorzystania AI: dostępem do mocy obliczeniowej, danymi oraz niezależną oceną ryzyka.

[6] D. Amodei, We Must Pace the Frontier, 12.09.2026, https://darioamodei.com/post/we-must-pace-the-frontier [22.09.2026].
[7] 2026 State of the Union Address by President von der Leyen, 16.09.2026, https://cyprus.representation.ec.europa.eu/news/2026-state-union-address-president-von-der-leyen-2026-09-16_en [22.09.2026].

Wnioski i rekomendacje

- Debatę o „tempowaniu" rozwoju AI prowadzą dziś przede wszystkim amerykańskie laboratoria, a jej ramy wyznacza rywalizacja USA z Chinami. Unia Europejska uczestniczy w niej głównie jako regulator i odbiorca technologii, podejmując równocześnie działania na rzecz rozwoju własnego potencjału. Dla państw Europy Środkowej oznacza to potrzebę łączenia trzech kierunków polityki: budowania zdolności do oceny i nadzoru nad wykorzystywanymi rozwiązaniami, upowszechniania wdrożeń w gospodarce i administracji oraz rozwijania własnej infrastruktury obliczeniowej i własnych modeli.
- W kwestii oceny modeli AI Act przewiduje istotną rolę ich dostawców oraz Urzędu ds. Sztucznej Inteligencji (AI Office). Dla poszczególnych państw – w tym państw Europy Środkowej – znaczenie ma zwłaszcza nadzór nad wdrażanymi systemami, którego skuteczność będzie zależeć od wyposażenia organów krajowych w kadry i kompetencje techniczne, zanim zaczną obowiązywać odpowiednie wymogi. Zaktualizowany AI Act wyznacza w tym obszarze dwa terminy: grudzień 2027 r. dla systemów wysokiego ryzyka działających samodzielnie oraz sierpień 2028 r. dla systemów wysokiego ryzyka wbudowanych w produkty. Sama zgodność z aktem nie rozstrzyga jednak o jakości działania modeli w językach narodowych („Komentarze IEŚ", nr 1509), przydatności do konkretnych zastosowań publicznych ani o realnej możliwości zmiany dostawcy. Są to również pytania o kompetencje i zamówienia publiczne, w których państwa regionu mogą działać wspólnie, budując infrastrukturę testowania modeli.
- Ciągłość dostępu do technologii staje się kwestią bezpieczeństwa w całej UE, zwłaszcza gdy dostęp ten bywa reglamentowany przez dostawców oraz rządy państw ich pochodzenia. Zależność obejmuje przy tym nie tylko modele, lecz także moc obliczeniową, warstwę pośrednią oprogramowania i usługi chmurowe. Znaczenie ryzyka przerwania dostępu rośnie wraz z wchodzeniem AI do usług publicznych i infrastruktury krytycznej. Odpowiedzią pozostaje mapowanie tych zależności oraz utrzymywanie i testowanie rozwiązań zastępczych w obszarach krytycznych.
- Warunki wyjściowe do upowszechniania AI są w Europie Środkowej silnie zróżnicowane („Komentarze IEŚ", nr 1582). Tam, gdzie adopcja AI należy do najniższych w UE, podstawowym zadaniem pozostaje upowszechnienie wdrożeń w biznesie i administracji, tam zaś, gdzie baza wdrożeniowa jest szersza, na znaczeniu zyskuje rozwijanie własnych rozwiązań i kontrola zależności od dostawców. W obu przypadkach są to obszary poddające się polityce krajowej. Doświadczenie państw o wyższej adopcji wskazuje, że przewaga ta wyrasta z inwestycji w edukację i kompetencje, poprawy jakości danych oraz wsparcia zmian organizacyjnych w przedsiębiorstwach i administracji.
- Rozwój własnych rozwiązań AI wymaga dostępu do infrastruktury obliczeniowej. Możliwość jej rozbudowy stwarza konkurs EuroHPC na gigafabryki AI, w którym termin składania ofert upływa 12 listopada 2026 r. Współpracę państw regionu Europy Środkowej w tym obszarze wspierają polsko-czeskie memorandum (przewidujące trzy warianty: wspólny projekt, dwie odrębne fabryki albo udział w jednym wybranym przedsięwzięciu) oraz Praska deklaracja w sprawie AI, w której dziewięć państw (Czechy, Polska, Słowacja, Węgry, Litwa, Łotwa, Słowenia, Chorwacja i Rumunia) zadeklarowało współpracę w zakresie wspólnej infrastruktury obliczeniowej. Skuteczność tych inicjatyw będzie zależeć od powiązania planowanych inwestycji z potrzebami przedsiębiorstw, administracji i ośrodków badawczych. Siłę przygotowywanych ofert może zwiększyć określenie konkretnych odbiorców mocy obliczeniowej oraz planów jej wykorzystania, zwłaszcza w sektorach wskazanych przez KE. Miarą sukcesu powinna być jednak nie tylko lokalizacja gigafabryk w regionie, lecz także dostęp regionalnych podmiotów do ich zasobów i możliwość rozwijania dzięki nim własnych modeli, produktów i usług. Suwerenność cyfrowa Europy Środkowej wymaga bowiem zarówno kompetencji pozwalających oceniać i kontrolować technologie zewnętrznych dostawców, jak i zdolności tworzenia własnych rozwiązań. Ewentualne „tempowanie" rozwoju AI może dać państwom regionu więcej czasu na budowę tak rozumianej suwerenności cyfrowej.