Helm
cyfrowy ster

Na horyzoncie:

Wybory parlamentarne 2027

Helm cyfrowy ster

Helm digital democracy

Helm Cyfrowy Ster | Helm Digital Democracy

Mokotowska 43/104

00-551 Warszawa

KRS: 0001090110

EU Transparency Register: 114284692189-05

www.cyfrowy-ster.org | helm-democracy.org

Na horyzoncie:

Wybory parlamentarne 2027

Redakcja: Konrad Kiljan

Autorzy/rki: Sonia Horonziak, Konrad Kiljan, Michał Klincewicz, Jakub Kubś, Mateusz Łabuz, Jakub Szymik, Dobromił Wereszczyński, Aleksandra Wójtowicz

Opracowanie graficzne, skład publikacji i ilustracje: NIKIFOR.studio

Made possible by the support of Civitates.

The sole responsibility for the content lies with the author(s) and the content may not necessarily reflect the positions of NEF or the Partner Foundations.

Przedsięwzięcie finansowane ze środków Instytutu Współpracy Polsko-Węgierskiej im. Wacława Felczaka.

Publikacja wyraża jedynie poglądy autora/ów i nie może być utożsamiana z oficjalnym stanowiskiem Instytutu Współpracy Polsko-Węgierskiej im. Wacława Felczaka.

Civitates
The European Democracy Fund

INSTYTUT WSPÓŁPRACY POLSKO-WĘGIERSKIEJ

Number ISBN 978-83-978606-2-9

ISBN 978-83-978606-2-9
9 788397 860629

© Copyright by Helm Digital Democracy, Warszawa 2026

Spis treści:

Wstęp 01 strona 6

Jakub Szymik

Jak chatboty mogą manipulować wyborcami? 02 strona 7

Dr Mateusz Łabuz

Slopaganda: trendy i wyzwania 03 strona 11

Dr Michał Klincewicz

Cyberbezpieczeństwo i sabotaże 04 strona 14

Dr Jakub Kubś

Finansowanie i przejrzystość reklamy oraz kampanii politycznej 05 strona 18

Dr Sonia Horonziak

Hiperpersonalizacja 06 strona 22

Aleksandra Wójtowicz

Influencerzy i polityka: nowa szara strefa 07 strona 25

Dobromił Wereszczyński

Wstęp

01

Jakub Szymik

Prezes zarządu, Helm – Cyfrowy Ster

Od 2023 roku, kiedy po raz ostatni odbyła się kampania parlamentarna w Polsce, tempo zmian technologicznych wyraźnie przyspieszyło, a w każdym zakątku świata wzrosły obawy o ich konsekwencje. Obywatelki i obywatele niemal codziennie pytają o wpływ nowych technologii na młodych ludzi, a także o skutki rozwoju sztucznej inteligencji dla środowiska, rynku pracy i debaty publicznej. Coraz trudniej przy tym oddzielić obietnice nowych możliwości od ryzyk, które im towarzyszą, zaś propozycje odpowiedzi coraz częściej biorą na cel ograniczenie strat oraz wypracowanie międzynarodowych współprac na rzecz ucywilizowania nowych rozwiązań.

W kontekście wyborów te napięcia nabierają szczególnej wagi. Jak zapewnić równe warunki demokratycznej rywalizacji, gdy obieg informacji przekracza granice państw, a odpowiedzialność za organizację wyborów pozostaje krajowa? Czy Polska ma realne narzędzia działania wobec globalnych platform i ich zautomatyzowanych systemów dystrybucji treści? Kto powinien rozstrzygać o granicach wolności słowa w spolaryzowanej debacie politycznej i kto powinien kontrolować decyzje podejmowane w ramach masowej, automatycznej moderacji?

Pytania z codziennych dyskusji krok po kroku przekładają się nie tylko na dokumenty strategiczne, ale również faktyczną odpowiedzialność. Dotyczą tego, na jakiej podstawie podejmujemy decyzje wyborcze, komu ufamy i jakie mamy możliwości obrony przed manipulacją. Debata o odpowiedziach zbyt często zatrzymuje się jednak na nagłówkach artykułów czy nawoływaniach do wzmocnienia edukacji medialnej. Kolejne kontrowersje przyciągają uwagę, ale pozostawiają niewiele miejsca na spokojną ocenę mechanizmów i skuteczności proponowanych rozwiązań.

W tym raporcie przyglądamy się temu, co dzieje się pod powierzchnią bieżących sporów, jak dokładnie technologie są wykorzystywane do manipulowania wyborcami i jak można temu przeciwdziałać. Część odpowiedzi wymaga zmian w sposobie działania instytucji państwowych. Część – skutecznego stosowania europejskich przepisów. W innych obszarach odpowiednie narzędzia dopiero trzeba wypracować. Publikując ten raport na rok przed wyborami mówimy: czas na przegląd sprawności państwa. Na sprawdzenie, które rozwiązania działają, gdzie pozostają luki i kto powinien je wypełnić. Nie wystarczy pytać, co technologia robi z wyborami. Trzeba również ustalić, co instytucje publiczne mogą i powinny zrobić, by je mądrze chronić.

Po raz pierwszy jako Cyfrowy Ster – zapraszamy do lektury i wspólnego działania.

Jak chatboty mogą manipulować wyborcami?

02

Dr **Mateusz Łabuz**

Doktor w Instytucie Badań nad Pokojem i Polityką Bezpieczeństwa przy Uniwersytecie w Hamburgu (IFSH), kieruje projektem „International Cybersecurity". Jest autorem książki „Przed waszą erą. Człowiek w syntetycznym świecie deepfake'ów, ChataGPT i sztucznej inteligencji" (2026), która opowiada o różnych wymiarach relacji człowieka ze sztuczną inteligencją. W przeszłości pracował jako zawodowy dyplomata.

Wyobraźmy sobie wyborcę, który na kilka dni przed wyborami jeszcze nie zdecydował, na kogo zagłosuje. Nie sięga jednak po tradycyjne środki przekazu i komunikacji politycznej, by dowiedzieć się więcej na temat alternatyw. Nie porównuje programów partii politycznych. Nie wypełnia ankiety Latarnika Wyborczego. Zamiast tego zwraca się do chatbota. Wpisuje kilka słów kluczowych i zadaje pytanie: „Który kandydat byłby najlepszy dla kogoś takiego jak ja?".

Odpowiedź, jak to w przypadku dużego modelu językowego (LLM od *large language model*), jest natychmiastowa i niezwykle sugestywna. Chatbot wyjaśnia, że jeden z kandydatów wydaje się „bardziej wiarygodny", ma „silniejszy program gospodarczy" lub „lepsze zrozumienie kwestii bezpieczeństwa narodowego". Ostatecznie sugeruje jedną partię jako „najbardziej odpowiadającą" wartościom wyborcy.

Tyle że chatbot wcale nie jest obiektywną wyrocznią, w której została zgromadzona wiedza na temat wszystkich kandydatów i partii politycznych. Może się mylić. Co jeśli będzie stronniczy wobec jednego z podmiotów politycznych? Taka stronniczość może wynikać wprost z danych, na których model został wytrenowany. Także danych, które zostały celowo zmanipulowane, właśnie po to, by ostatecznie wpłynąć na wyborców. Ba, może być nawet pochodną bezpośrednich instrukcji przekazanych modelowi przez dostawcę infrastruktury.

W praktyce wyborca może nigdy nie zobaczyć reklamy, nie natknąć się na hasło wyborcze, ani nawet nie spojrzeć na polityka, a jego wybór zostanie ukształtowany w toku interakcji z systemem. Decyzja zostanie podjęta w zupełnie innym miejscu, choć może się wydawać „własna".

Algorytmy rekomendacyjne, głównie w mediach społecznościowych, od wielu lat przejmują rolę klasycznych *gatekeeperów* (jak media czy władze), swego rodzaju „strażników dostępu do informacji". Generatywna sztuczna inteligencja o charakterze konwersacyjnym tworzy nowy wymiar *gatekeepingu*. Według Digital News Report Instytutu Reutersa z LLM-ów do pozyskiwania informacji korzysta już 10% użytkowników, a w najmłodszej grupie wiekowej aż 17%[1]. W praktyce perswazja o charakterze politycznym jeszcze nigdy nie była tak groźna, bo oparta na daleko idącej personalizacji. Dziś może być sprowadzona do poziomu prywatnej, wręcz intymnej rozmowy z systemami, które jawią się jako bezstronni doradcy i wywierają coraz większy wpływ na decyzje podejmowane przez ludzi.

Od dezinformacji do perswazji za pośrednictwem maszyn

Zaufanie do LLM-ów to nowe ryzyko wyborcze o nieco innym charakterze od baniek informacyjnych, które znamy głównie z mediów społecznościowych, przy czym u jego podstaw wciąż leży łatwość, z jaką może zostać zmanipulowany nowoczesny ekosystem wymiany informacji i komunikacji. Już sam fakt, że jako ludzie coraz częściej kierujemy się wskazaniami LLM-ów powinien napawać niepokojem – pozbawia nas bowiem znacznej części podmiotowości i sprawczości, które stanowią przecież jeden z fundamentów demokracji[2]. Co gorsza, LLM-y są także podatne na wpływy z zewnątrz i mogą być celem strategii obliczonych na manipulację wynikami.

Pierwsza z nich jest stosunkowo prosta: chatbotowi można po prostu polecić, by faworyzował konkretnego kandydata lub stanowisko polityczne, a przez to bezpośrednio oddziaływać na preferencje osób wchodzących w interakcje z modelem. W eksperymentach przeprowadzonych przez naukowców z MIT w 2025 r. rozmowy z modelami sztucznej inteligencji opowiadającymi się za konkretnymi kandydatami spowodowały znaczące zmiany w preferencjach wyborczych. Efekty (zmiana deklaracji) były większe niż te, które obserwujemy za-

1 J. Egan i in., *Reuters Institute Digital News Report 2026*, Reuters Institute for the Study of Journalism, 2026.

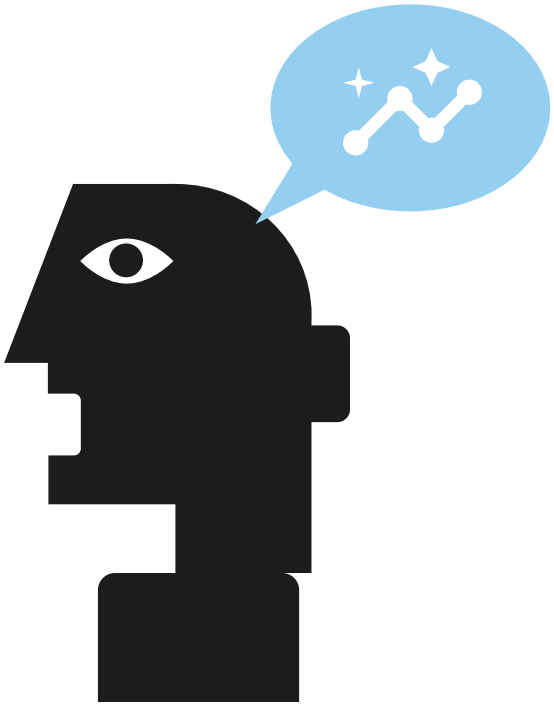
2 M. Łabuz, *Przed waszą erą. Człowiek w syntetycznym świecie deepfake'ów, ChataGPT i sztucznej inteligencji*, Filia, Poznań 2026.

zwyczaj w przypadku reakcji na tradycyjne reklamy w formie wideo[3]. Co ważne, modele nie potrzebowały wyrafinowanych sztuczek psychologicznych, by przekonać użytkowników. Były skuteczne, ponieważ dostarczały pozornie istotnych faktówi argumentów w trakcie spersonalizowanej rozmowy. Badanie dowiodło tym samym, jak wielki potencjał perswazji drzemie w generatywnej sztucznej inteligencji o charakterze konwersacyjnym.

Tego typu system nawiązuje z użytkownikiem zasadniczo odmienną relację niż klasyczne reklamy w postaci billboardów czy spotów wyborczych. Chatbot potrafi odpowiadać na pytania użytkownika, posługuje się przekonującą argumentacją, a przede wszystkim doskonale „czyta" odbiorcę, rozpoznając to, co dla użytkownika jest istotne i angażujące. Wszystkie te czynniki pozwalają dobrać odpowiednią argumentację. W praktyce taka konwersacja może przybrać formę spersonalizowanej, a przy tym adaptacyjnej perswazji politycznej. Problemem jest w tym wypadku możliwość uprzedniego zaprogramowania, w którym kierunku powinna zmierzać dyskusja sterowana przez chatbota, np. argumentowania w celu obrony określonych przekonań. Ten mechanizm sprawdza się w przypadku demistyfikowania pseudonaukowych teorii spiskowych, ale równie dobrze może być wykorzystywany do siania politycznej propagandy lub wprost promowania określonej partii lub kandydatów.

Druga forma manipulacji jest znacznie mniej widoczna, a przez to trudniejsza do wychwycenia. Stronniczość, także ta o charakterze politycznym czy ideologicznym, może być zakodowana w danych, na jakich model był trenowany, a następnie wzmacniana w toku dostrajania, a nawet komponowania procedur bezpieczeństwa. Dwa systemy szkolone w innych warunkach mogą udzielić skrajnie różnych odpowiedzi na to samo pytanie. Dlaczego tak się dzieje? Z badań wynika, iż duże modele językowe powielają poglądy charakterystyczne dla obszaru kulturowego, z którego zaczerpnięto dane treningowe. I tak, modele „zachodnie" będą zdecydowanie bardziej prodemokratyczne od swoich chińskich odpowiedników. Te ostatnie chętniej będą promować wartości charakterystyczne dla systemu komunistycznego[4]. Nawet bez wyraźnej decyzji o faworyzowaniu konkretnego kandydata czy partii politycznej różnice w informacjach dostępnych dla modelu mogą powodować różnice w odpowiedziach otrzymywanych przez wyborców, a przez to wpływać na postrzeganie świata.

Stwarza to poważny problem. Jeśli wyborca zapyta: „Jakie są najsilniejsze argumenty za i przeciw kandydatowi lub postawie X?", odpowiedź nie zależy wyłącznie od tego, co obiektywnie jest słuszne moralnie czy korzystne z politycznego punktu widzenia. Zależy ona również od tego, które źródła model uznaje za istotne i które wartości kształtowały ekosystem informacyjny, w którym model został wytrenowany. W tym kontekście pominięcie informacji, stworzenie pewnych heurystyk czy też zmiana w obrębie sentymentu mogą mieć równie poważne konsekwencje polityczne jak fałsz podany wprost. Chatbot nie musi kłamać na temat kandydata, aby wpłynąć na wynik wyborów. Może po prostu uwypuklać określone wartości jako istotniejsze, a inne zwyczajnie przemilczeć.



Manipulowanie modelami LLM: zatruwanie ekosystemu informacji

Trzeci mechanizm działa na wcześniejszym etapie. Zamiast manipulować bezpośrednio chatbotem, podmiot może próbować manipulować informacjami, na podstawie których uczą się systemy sztucznej inteligencji. Taka jest logika stojąca za tzw. „groomingiem" dużych modeli językowych (LLM grooming). Chodzi o celowe zasilanie sieci masowo powielanymi treściami, które mają trafić do danych treningowych i źródeł wyszukiwania modeli językowych, by te następnie powtarzały zawarte w nich narracje. Badania naukowe wykazały podatność dużych modeli językowych na zaciąganie spreparowanych danych zaburzających ekosystem informacji. Aktorzy dezinformacji zdają sobie sprawę z tej podatności i już sięgają po agresywne operacje informacyjne[5].

Dobitnym przykładem jest aktywność rosyjskiej sieci „Pravda", która w samym tylko 2024 roku wyprodukowała 3,6 miliona artykułów opublikowanych następnie w różnych językach (najczęściej z wykorzystaniem automatycznego tłumaczenia, do czego prawdopodobnie także wykorzystywano LLM-y)[6]. W artykułach kolportowano prokremlowskie narracje kwalifikowane jako dezinformujące. W szeroko komentowanych dochodzeniach NewsGuard oraz American Sunlight Project,

3 H. Lin i in., *Persuading voters using human–artificial intelligence dialogues*, „Nature" 2025, nr 648.

4 M. Buyl i in., *Large Language Models Reflect The Ideology Of Their Creators*, „npj Artificial Intelligence" 2026, nr 2.

5 American Sunlight Project, *A Pro-Russia Content Network Foreshadows the Automated Future of Info Ops*, American Sunlight Project, Waszyngton 2025

6 Viginum, *PORTAL KOMBAT A structured and coordinated pro-Russian propaganda network*, Viginum, Paryż 2025.

badacze porównali piętnaście narracji promowanych przez tę sieć z odpowiedziami dziesięciu głównych modeli LLM i stwierdzili zbieżność w 33 procentach przypadków; w niektórych sytuacjach modele cytowały nawet materiały powiązane z „Pravdą" jako źródło[7]. Powtórzone rok później badanie wykazało zdecydowanie mniejszą podatność na manipulacje „Pravdy", ale i tutaj naukowcy wykazali konkretne podatności, szczególnie tam, gdzie wyraźny jest deficyt informacyjny, a propaganda może eksplorować konkretne luki[8].

Mechanizm ten jest pozornie prosty. Jeśli w sieci opublikowane zostaną duże ilości tekstów powielających identyczne lub zbliżone powiązania między słowami (na poziomie semantycznym), część z nich może ostatecznie zostać włączona do zbiorów danych wykorzystywanych do tworzenia modeli i stać się dla nich punktem odniesienia. To, co zaczęło się jako celowo wykreowana narracja, może zatem zyskać drugie życie w systemie sztucznej inteligencji, a stamtąd trafić z powrotem do ludzkiego ekosystemu wymiany informacji. To mechanizm swego rodzaju *„prania dezinformacji"*[9]. W przypadku *LLM groomingu* nie chodzi zatem o bezpośrednie dotarcie do odbiorców i podsuwanie im spreparowanego artykułu. Ten artykuł ma do nich dotrzeć już w formie przetworzonej i podanej jako wynik analizy modelu, niesłusznie uznawanej za obiektywny i autorytatywny. Celem jest zatem mechanizm dystrybucji informacji.

To nowy wymiar możliwej ingerencji w wybory, która zazwyczaj jest rozłożona w czasie. Kampania dezinformacyjna rozpoczęta miesiące lub lata przed wyborami może potencjalnie wpłynąć na to, w jaki sposób system sztucznej inteligencji opisuje kandydata, partię polityczną lub jakieś zjawisko, gdy wyborcy zaczną zadawać pytania. Atak nie jest zatem skierowany wprost przeciwko wyborcom, choć to oni znajdują się na końcu łańcucha wymiany informacji. Jest on wymierzony w infrastrukturę informacyjną, za pośrednictwem której wyborcy będą pozyskiwać wiedzę. Jednocześnie horyzont czasowy powodzenia operacji jest trudny do przewidzenia.

Duże modele językowe nie są ekwiwalentem wyszukiwarki internetowej, ale mogą być w ten sposób traktowane przez samych użytkowników. Istotna jest różnica w samym mechanizmie dochodzenia do wniosków. Nie tylko podczas wyborów. Wyborca, który wyszukuje hasło „program podatkowy kandydata A", może natrafić na dziesiątki konkurujących ze sobą źródeł. Wyborca, który zapyta wprost o to, który kandydat ma najlepszy program podatkowy, może otrzymać jedną, zsyntetyzowaną odpowiedź, która właściwie przesądzi o decyzji.

Należy jednak podkreślić, że systemy konwersacyjne wykazują przynajmniej bazową odporność zarówno na treści dezinformujące (w połowie przypadków narracje sieci „Pravda" był odrzucane wprost), jak i na próby opowiedzenia się za konkretnym kandydatem lub partią polityczną (modele mają zazwyczaj wbudowane odpowiednie ograniczenia, które uniemożliwiają im tak jednoznaczne rozstrzygnięcia)[10]. Nie zawsze jednak te bufory bezpieczeństwa działają, z czego aktorzy dezinformacji i aktorzy polityczni zdają sobie sprawę.

Nowy wymiar optymalizacji

Trudno się zatem dziwić, że na znaczeniu zyskują techniki tzw. optymalizacji wyszukiwania generatywnego (GEO). W praktyce są one próbami wywierania wpływu na to, jakie informacje systemy sztucznej inteligencji wyszukują, traktują priorytetowo i ostatecznie przekazują na temat konkretnych osób, organizacji, firm czy też kwestii politycznych. Tradycyjna optymalizacja wyszukiwarek miała na celu umieszczenie strony internetowej jak najwyżej na liście wyników zwracanych użytkownikom (tzw. pozycjonowanie). Celem GEO jest wywarcie wpływu na sam model i jego odpowiedź. Dla procesów demokratycznych może to oznaczać jeszcze silniejsze powiązanie ekspozycji kandydata, a tym samym możliwości dotarcia do odbiorców, ze zdolnością do wpływania na algorytmy i promowane przez nie treści.

Sztaby wyborcze już traktują „widoczność" w LLM-ach i GEO jako nową dziedzinę marketingu politycznego, co może skutkować próbami nowego pozycjonowania, swego rodzaju *„LLM groomingiem* w skali mikro". Nie oznacza to, że każda próba poprawy reprezentacji kandydata w odpowiedziach zwracanych przez system sztucznej inteligencji powinna być oceniona negatywnie. Dbanie o to, by chatbot wiernie odzwierciedlał oficjalny program partii ma demokratyczne uzasadnienie. Ryzyko dla demokracji pojawia się wtedy, gdy GEO przeradza się w próby manipulacji: fałszowania źródeł, czy też zalewania sieci treściami bez większej wartości, ale zaprojektowanymi tak, by przenikały do danych treningowych systemów sztucznej inteligencji.

Rekomendacje

Wartość LLM-ów zależy w dużej mierze od tego, jak są projektowane, zarządzane i kontrolowane. Obecnie demokratyczne społeczeństwa mają jedynie iluzoryczną kontrolę nad tymi procesami. Kluczowe jest sprawowanie adekwatnego nadzoru nad rynkiem LLM-ów, a tam, gdzie jest to niezbędne, odpowiednie regulowanie i zmuszenie firm technologicznych do lepszej moderacji danych treningowych oraz wbudowania skuteczniejszych buforów bezpieczeństwa na poziomie konwersacji.

1. Kraje przygotowujące się do wyborów mogą opracować ramy regulacyjne dotyczące **oznaczania informacji politycznych** przekazywanych za pośrednictwem sztucznej inteligencji. Należy wprowadzić obowiązkową przejrzystość w odniesieniu do materiałów publikowanych przez sztaby wyborcze, także tych obliczonych na GEO lub wprost *LLM grooming*. W polskich systemie prawnym wymagałoby to koordynacji na

7 American Sunlight Project, *op. cit*.

8 M. Alyukov i in., LLMs grooming or data voids? *LLM-powered chatbot references to Kremlin disinformation reflect information gaps, not manipulation*, „Harvard Kennedy School Misinformation Review", 2025.

9 M. Łabuz, *op. cit*.

10 American Sunlight Project, *op. cit*.

linii Państwowa Komisja Wyborcza - Komisja Rozwoju i Bezpieczeństwa Sztucznej Inteligencji.

---

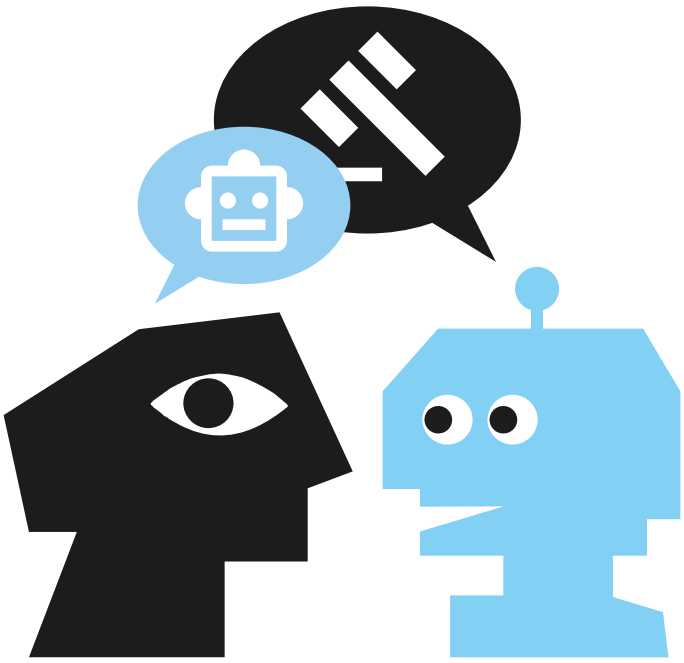
2. Niezbędne jest **konsekwentne monitorowanie** wpływu LLM-ów na środowisko informacyjne oraz podejmowanie dalszych kroków w celu rozpoznawania ryzyk systemowych stanowiących zagrożenie dla procesów demokratycznych, w tym wyborów. Zakwalifikowanie przez Komisję Europejską ChataGPT jako very *large online search engine* w myśl przepisów Aktu o Usługach Cyfrowych (Digital Services Act) należy ocenić jako słuszny krok wychodzący naprzeciw praktyce wykorzystywania tego narzędzia przez użytkowników[11]. Wyspecjalizowane instytucje administracji państwowej (w polskim systemie przede wszystkim Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy) powinny prowadzić monitoring sieci pod kątem prób tworzenia skoordynowanych sieci dezinformujących lub stosujących strategię *LLM groomingu*.

---

3. Zasadne wydaje się wprowadzenie **systematycznych audytów LLM-ów**. Organy wyborcze, niezależni badacze i NGOs powinny przeprowadzać okresowe testy chatbotów przy użyciu standardowych zestawów pytań. Audyty powinny mierzyć zgodność z faktami, jakość źródeł, widoczność kandydatów, asymetrie ideologiczne, pominięcia oraz różnice w odpowiedziach w zależności od języka i kontekstu demograficznego.

---

4. Konieczne jest także **podnoszenie świadomości i edukowanie obywateli** w zakresie ograniczeń i zagrożeń LLM-ów. Można to osiągnąć poprzez realizację kierunkowych programów edukacyjnych i kampanii społecznych, w szczególności obliczonych na podnoszenie kompetencji cyfrowych w zakresie korzystania z generatywnej sztucznej inteligencji.



---

11 Komisja Europejska, *Commission designates ChatGPT, Reddit, Roblox under Digital Services Act*, https://digital-strategy.ec.europa.eu/en/news/commission-designates-chatgpt-reddit-roblox-under-digital-services-act

Slopaganda: trendy i wyzwania

03

Dr **Michał Klincewicz**

Adiunkt w Katedrze Kognitywistyki Obliczeniowej Uniwersytetu w Tilburgu. Prowadzi badania na styku sztucznej inteligencji, kognitywistyki i filozofii, w tym etyki. Jego prace oscylują między technicznymi właściwościami systemów algorytmicznych a pytaniami poznawczymi i normatywnymi, jakie rodzą się przy ich wdrażaniu.

Od premiery ChataGPT mijają właśnie 4 lata. Od tego momentu generatywna sztuczna inteligencja zdobywała popularność i rozwija się w coraz szybszym tempie, umożliwiając szybkie i tanie wytwarzanie treści na masową skalę oraz ich precyzyjne dopasowywanie do poszczególnych grup odbiorców. Technologie te trafiły na podatny grunt: środowisko internetowe zdominowane przez platformy społecznościowe, nieustanny obieg wiadomości i konkurencję o uwagę użytkowników. Nadmiar dostępnych treści sprzyja przeciążeniu informacyjnemu, utrudnia ocenę wiarygodności przekazów i pogłębia obawy o ich wpływ na dobrostan odbiorców, zwłaszcza młodzieży. Z perspektywy polityki i wyborów szczególne znaczenie ma pojawienie się w tych warunkach „slopagandy", masowo wytwarzanych, często personalizowanych i zazwyczaj niskiej jakości treści politycznych tworzonych przy użyciu sztucznej inteligencji. Niski koszt ich produkcji i łatwość rozpowszechniania otwierają zarazem pole działania „przedsiębiorcom slopagandowym", którzy wykorzystują te możliwości do zdobywania zasięgów, osiągania korzyści finansowych lub wywierania wpływu politycznego. Dokładne zrozumienie mechanizmów działania slopagandy i analiza jej dotychczasowych przykładów jest pierwszym krokiem w kierunku przewidzenia konsekwencji dla debaty publicznej i przyszłych kampanii wyborczych.

Dlaczego to działa?

Kluczowe dla oceny zagrożeń związanych ze slopagandą jest zrozumienie, dlaczego tego rodzaju treści mogą skutecznie oddziaływać na odbiorców. Większość z nich wydaje się absurdalna lub łatwa do rozpoznania jako fałszywa, co może sprzyjać jej bagatelizowaniu jako kolejnym "głupotkom z Internetu". Jeśli jednak skorzystamy z teorii i narzędzi kognitywistyki oraz teorii przetwarzania informacji, zrozumiemy, że powszechna masowa konsumpcja takich treści wcale nie jest nieszkodliwa[1].

Generatywna sztuczna inteligencja pozwala szybko i niewielkim kosztem wytwarzać ogromne ilości materiałów, a następnie dopasowywać ich przekaz do zainteresowań, przekonań czy lęków poszczególnych grup i użytkowników. W rezultacie odbiorca styka się z nieustannym strumieniem pozornie wiarygodnych informacji, które przedstawiają określone zagrożenia i dostarczają kolejnych uzasadnień dla obaw i uprzedzeń. Nadmiar treści **zwiększa obciążenie poznawcze**, czyli angażuje ograniczone zasoby uwagi potrzebne do przetwarzania informacji, rozumienia ich i oceny. W stanie przeciążenia uwagi trudniej jest rozróżnić co ważne jest uwagi a co nie. O tym, co przyciągnie uwagę, decydują wtedy cechy niezwiązane z prawdziwością przekazu, na przykład jego emocjonalny charakter, powtarzalność lub zgodność z wcześniejszymi przekonaniami. Nasza uwaga pełni funkcję swoistego filtra: współdecyduje o tym, które informacje zostaną przyswojone i trafią do pamięci, a więc zasobu wiedzy i skojarzeń wykorzystywanego później do interpretowania wydarzeń oraz podejmowania decyzji. Jak wiemy z marketingu i reklam, asocjacje tworzą się pod progiem świadomości, ale potrzebują wielokrotnego przedstawienia do utrwalenia się. A nawet powierzchowne zaakceptowanie fałszywego twierdzenia jako wartego obejrzenia do końca lub dla rozrywki zwiększa naszą podatność na kolejne przekazy o podobnej treści. Dzieje się to również wtedy, gdy odbiorca później przyjmie do wiadomości jego sprostowanie. Skuteczności tego mechanizmu nie należy przy tym oceniać wyłącznie na poziomie pojedynczej osoby. Niewielki wpływ na przekonania lub decyzje, powielony wśród milionów odbiorców, może przełożyć się na istotne zmiany opinii publicznej czy zachowań wyborczych. Dlatego slopagandy nie można sprowadzać do zabawnych, niegroźnych filmików. Jej znaczenie wynika także z kumulacyjnego oddziaływania na sposób, w jaki grupy selekcjonują informacje, zapamiętują je i wykorzystują do rozumienia rzeczywistości.

Dzięki slopagandzie wyjątkowo skuteczna staje się także znana i bardzo niebezpieczna taktyka **flash flood**, polegająca na przeprowadzeniu nagłego zalewu informacyjnego. Atakujący wykorzystuje w niej próżnię informacyjną powstałą po ważnym wydarzeniu, na przykład po upadku rakiety w lesie. Zanim władze zdążą zareagować w odpowiedni sposób i zbadać in-

1 Klincewicz, M., Alfano, M., & Fard, A. E. (2025). Slopaganda: The interaction between propaganda and generative AI. *arXiv preprint arXiv:2503.01560*, opracowanie popularyzatorskie: Slopaganda — Propaganda in the age of generative AI https://slopaganda-two.vercel.app/

cydent, media zostają zalane sprzecznymi komunikatami na temat przebiegu i miejsca zdarzenia. Mieszkańcy i ogół społeczeństwa mogą czuć się zdezorientowani, a osoby publiczne – w tym urzędnicy – mogą nieświadomie stać się nośnikami dezinformacji. Nagłe powodzie osłabiają skuteczność instytucjonalnych działań na rzecz skoordynowanej reakcji i podważają zaufanie do instytucji. „Slopaganda" znacząco zwiększa skuteczność taktyk wykorzystujących **flash flood**, zwłaszcza że pozwala ona na personalizację przekazu pod kątem konkretnych grup, a nawet pojedynczych osób.

Astroturfing, czyli pozorowanie akcji oddolnych, to kolejna taktyka, której stosowanie umożliwia technologia generatywnej sztucznej inteligencji (AI). Jak niedawno donosił „The New York Times", wielu twórców treści wspieranych przez AI otrzymuje wynagrodzenie za tworzenie i rozpowszechnianie materiałów o charakterze politycznym[2]. Kreuje to wrażenie, że dany temat lub pogląd jest przedmiotem szerokiej debaty publicznej. Przykładem może być kampania popierająca republikańskiego kandydata na burmistrza Los Angeles, prowadzona przez osoby sprawiające wrażenie zwykłych obywateli, które publikowały nagrania w serwisie TikTok. Nie były to deklaracje poparcia ze strony znanych celebrytów, lecz skoordynowane działania mające na celu pozyskanie mikroinfluencerów za stawki zaczynające się już od 1500 USD za film. Oczywiście źródło finansowania takich działań nie musi być ujawniane i przy tej skali łatwo je ukryć, dlatego w niedalekiej przyszłości można spodziewać się prób prowadzenia kampanii typu astroturfing również w Polsce. Według „The Guardian" już w 2023 roku News Corp Australia produkowała przy użyciu generatywnej sztucznej inteligencji 3000 lokalnych materiałów informacyjnych tygodniowo[3].

Mikrotargetowanie jest już możliwe dzięki takim narzędziom jak reklamy kierowane do konkretnych odbiorców czy personalizowane kanały treści w mediach społecznościowych. Ze względu na niski koszt i łatwość tworzenia, slopaganda może teraz ułatwić prowadzenie mikrotargetowania na masową skalę. Jak zauważył Scott Goodstein, dyrektor ds. działań online w kampanii prezydenckiej Baracka Obamy z 2008 roku oraz główny strateg cyfrowy w kampanii Berniego Sandersa z 2016 roku: „Sztaby polityczne mogą teraz uruchamiać to, co nazywam »pociskami w mediach społecznościowych«: krótkie filmy, reklamy kierowane do konkretnych grup i wiadomości tekstowe, mające dotrzeć do niewielkich, precyzyjnie wyselekcjonowanych grup wyborców – często zanim dziennikarze, konkurencyjne sztaby czy weryfikatorzy faktów (fact-checkerzy) w ogóle dowiedzą się o ich istnieniu. Zanim je wykryjemy, szkody są już wyrządzone"[4]. Innymi słowy, większość materiałów typu slopaganda skierowanych do precyzyjnie określonych grup odbiorców, które realnie wpływają na sytuację, może nawet nie zostać zauważona ani zidentyfikowana jako takie zjawisko.

Różne źródła tego samego problemu

Federacja Rosyjska i Iran już teraz aktywnie włączyły slopagandę do swojego arsenału działań w ramach wojny informacyjnej i poznawczej. Do niedawnych przykładów należą operacje ingerencji informacyjnej koordynowane przez GRU[5] podczas wyborów w Mołdawii w 2025 roku oraz wyborów prezydenckich na Węgrzech w 2026 roku. Te ostatnie obejmowały materiały wideo typu slopaganda, przedstawiające realistycznie wyglądające aresztowania ukraińskich przestępców na Węgrzech, egzekucje węgierskich żołnierzy na froncie ukraińskim oraz deepfake'i z udziałem kandydata opozycji, wygłaszającego kontrowersyjne stwierdzenia. Można spodziewać się podobnych działań koordynowanych przez podmioty takie jak Social Design Agency[6] czy inne rosyjskie struktury mające na celu ingerencję w wybory w Polsce. Tematyka slopagandy będzie prawdopodobnie nawiązywać do kwestii budzących obecnie w Polsce największe emocje polityczne: wojny na Ukrainie, uchodźców z tego kraju, zaangażowania USA w regionie, kryzysu migracyjnego na wschodniej granicy z Białorusią oraz skandali z udziałem dowolnych polskich sił politycznych. Co istotne, treści te nie będą wymierzone wyłącznie w obecny rząd koalicyjny, choć siły polityczne sprzyjające celom Kremla mogą liczyć na większe wsparcie na wzór Orbána na Węgrzech. Biorąc pod uwagę wiedzę na temat kluczowych wskaźników skuteczności rosyjskich strategii manipulacji informacyjnej, ich nadrzędne cele mają charakter społeczny: chodzi o zakłócenie debaty publicznej, polaryzację społeczeństwa, wywołanie dezorientacji poznawczej oraz zwiększenie obecności tematyki dotyczącej Rosji w przestrzeni publicznej.

Instytucjonalna słopaganda w sferze politycznej stała się obecnie zjawiskiem powszechnym i w porównaniu do nieinstytucjonalnej slopagandy relatywnie nieszkodliwą. Głównym zagrożeniem związanym z tego typu treściami jest **normalizacja slopagandy**, a nie jej siła perswazji czy zdolność do wprowadzania w błąd. Same materiały są oczywiście stronnicze politycznie i ewidentnie stworzone przy użyciu generatywnej sztucznej inteligencji. Do kwestii tego zagrożenia powrócimy w sekcji poświęconej rekomendacjom.

Obecność i skuteczność slopagadny w komunikacji globalnej potwierdziły w 2026 roku animacje z klockami Lego tworzone i dystrybuowane na cały świat przez firmę Exploding Media dla

2 An Army of Paid Creators Is Carpeting the Web With Political Content, The New York Times, 22 września 2026 r. https://www.nytimes.com/2026/09/22/us/politics/influencers-midterms-echo-chamber.html

3 A. Meade, *News Corp using AI to produce 3,000 Australian local news stories a week*, „The Guardian", 31 lipca 2023 r. https://www.theguardian.com/media/2023/aug/01/news-corp-ai-chat-gpt-stories

4 S. Goodstein, *I helped build Obama's 2008 campaign. Democrats aren't ready for 2026.*, USA Today, 1 września 2026 r. https://eu.usatoday.com/story/opinion/2026/09/01/campaign-ads-midterms-misinformation-social-media-influencers/91547477007/

5 S. Panyi, *Putin's GRU-linked Election Fixers Are Already in Budapest to Help Orbán*, VSquare, 6 marca 2026 r. https://vsquare.org/putins-gru-linked-election-fixers-are-already-in-budapest-to-help-orban/

6 Psychological Defence Research Institute, *Beyond Operation Doppelgänger: A Capability Assessment of the Social Design Agency*, Lund University, 15 maja 2025 r. https://www.psychologicaldefence.lu.se/article/beyond-operation-doppelganger-capability-assessment-social-design-agency

Korpusu Strażników Rewolucji Islamskiej[7]. Według kanadyjskiego zespołu analizującego zagraniczne operacje informacyjne materiały te dotarły w marcu 2026 roku do co najmniej 83 milionów użytkowników na świecie[8]. W Stanach Zjednoczonych Donald Trump udostępnił w lutym 2025 roku film przedstawiający zniszczoną Strefę Gazy jako przyszły luksusowy kurort, z hotelem sygnowanym jego nazwiskiem i złotym pomnikiem prezydenta[9]. Fantastyczna wizja dobrobytu nadawała atrakcyjną wizualnie formę jego propozycji przejęcia tego terytorium przez USA. Z kolei w Indonezji podczas kampanii prezydenckiej już w 2024 roku sztab Prabowo Subianto wykorzystywał generowane przez AI kreskówkowe wizerunki kandydata, przedstawiające byłego generała jako sympatycznego, nieszkodliwego dziadka[10]. Były one elementem strategii „gemoy", czyli budowania „uroczego" i „rozczulającego" wizerunku, aby odsuwać uwagę od kontrowersyjnej przeszłości polityka. Do tej kategorii można również zaliczyć materiały emitowane przez TV Republika w programach Miłosza Kłeczka, takie jak te dotyczące „kryptodyktatury"[11]. Przykłady te pokazują, że polityczne oddziaływanie syntetycznych treści nie wymaga przekonania odbiorców, iż oglądają autentyczne wydarzenia: może opierać się również na humorze, widowiskowości i emocjonalnych skojarzeniach, które kształtują sposób postrzegania polityków i ich działań. Obecnie większość instytucjonalnej slopagandy rozpowszechniana jest za pośrednictwem prywatnych i instytucjonalnych kont na platformach X, Facebook oraz TikTok.

Równie szeroko działa jednak dziś slopaganda nieinstytucjonalna. W obecnym krajobrazie medialnym odbiorcy nie są już jedynie biernymi konsumentami mediów i informacji; aktywnie tworzą oni i udostępniają treści na własną rękę. Protokoły internetowe ułatwiają ukrywanie pochodzenia materiałów; Fałszywe konta, maskowane adresy IP, kopie stron internetowych itp. mogą służyć do ukrycia tożsamości twórcy treści w internecie. Stwarza to szansę na bezpieczniejsze wyrażanie opinii – zwłaszcza dla grup marginalizowanych – ale jednocześnie dostarcza wachlarza metod podmiotom państwowym i niepaństwowym zainteresowanym manipulacją informacyjną oraz ingerencją w sprawy wewnętrzne i międzynarodowe. Zagrożenia tutaj są bezpośrednie: chaos informacyjny, dezinformacje, manipulacje i inne powszechnie znane efekty kampanii FIMI. Slopaganda instytucjonalna, która normalizuje polityczne treści generowane przez AI, potęguje efektywność slopagandy nieinstytucjonalnej.

7 M. Alfano, M. Klincewicz, *Slopaganda wars: how (and why) the US and Iran are flooding the zone with viral AI-generated noise*, „The Conversation", 8 kwietnia 2026 r. https://theconversation.com/slopaganda-wars-how-and-why-the-us-and-iran-are-flooding-the-zone-with-viral-ai-generated-noise-280024

8 Rapid Response Mechanism Canada, Global Affairs Canada, 10 lipca 2026 r, https://international.canada.ca/en/global-affairs/corporate/reports/rapid-response-mechanism/2026-explosive-media-iran

9 Susan Carpenter, *Trump posts AI-generated Gaza video that shows it as proposed 'Riviera of the Middle East'*, Spectrum News, 26 lutego 2025 r. https://spectrumlocalnews.com/us/snplus/politics/2025/02/26/trump-ai-generated-gaza-video

10 Ross Tapsell, *The great rebrand, Inside Indonesia*, 24 lutego 2025 r. https://www.insideindonesia.org/editions/edition-158-oct-dec-2024/the-great-rebrand

11 Kanał Telewizji Republika na Youtube, Wysokie napięcie Miłosz Kłeczek https://youtu.be/_dJLROkId78?si=tMxhWgWYNyOfZSJx&t=8823

Rekomendacje

1. **Kompetencje cyfrowe** stanowią pierwszą linię obrony przed „slopagandą". Po pierwsze, jak potwierdza panujący obecnie konsensus naukowy, odbiorcy informacji są bardziej odporni na dezinformację i błędne informacje, jeśli mają świadomość mechanizmów ich wykorzystywania, potencjalnej tematyki oraz tego, kto może być ich twórcą. Można oczekiwać, że podobny mechanizm zadziała również w przypadku taktyk opartych na „slopagandzie". Gdy ludzie wiedzą, czym jest dane zjawisko, jaki ma na nich wpływ i kto za nim stoi, są lepiej przygotowani, by przeciwstawić się jego skutkom. Dodatkową korzyścią płynącą z kampanii na rzecz kompetencji cyfrowych jest budowanie zaufania do wiarygodnych źródeł wiadomości i informacji opartych na dowodach naukowych, a także kształtowanie przekonania, że władze interesują się codziennym życiem obywateli. W związku z tym rekomendujemy opracowanie kampanii edukacyjnych dotyczących sztucznej inteligencji, które powinny być realizowane na szczeblu krajowym i lokalnym przez organy wyborcze, samorządy, a nawet resort edukacji. Telewizja publiczna może na przykład emitować regularne materiały informacyjne na temat tzw. slopagandy (niskiej jakości treści propagandowych generowanych masowo przez AI) oraz treści wytwarzanych przez sztuczną inteligencję, kierując jednocześnie odbiorców do specjalnie przygotowanych serwisów internetowych zawierających bardziej szczegółowe informacje i przykłady.

2. Systemy informacyjne muszą zostać przystosowane do **standardów watermarkingu**[12]. Przedstawiciele branży oraz organy regulacyjne wprowadzą już rozwiązania techniczne umożliwiające – zarówno w sposób dyskretny, jak i jawny – oznaczanie treści generowanych przez AI (tzw. znakowanie wodne, czy 'watermarking'). Firma Anthropic zaczęła już stosować takie rozwiązanie w odniesieniu do wszystkich treści tworzonych przez model Claude[13]. Umożliwia to oznaczanie lub usuwanie nieoznaczonych materiałów typu "slopaganda", a także wyraźne wskazywanie pochodzenia treści oznaczonych jako wytworzone przez AI na platformach pełniących funkcję głównych kanałów dystrybucji wiadomości i innych istotnych informacji. Choć rozwiązanie to wymaga działań na szczeblu regulacyjnym, infrastruktura wyborcza może już teraz ułatwiać jego wdrożenie i przygotowywać się do niego. Wymaga to przede wszystkim zaangażowania personelu technicznego oraz współpracy z chętnymi partnerami z sektora przemysłowego, zwłaszcza z podmiotami funkcjonującymi w polskim ekosystemie informacyjnym.

12 N. Al-Khatib, AI Watermarking: *How It Works, Applications, Challenges*, DataCamp, 20 lutego 2025 r. https://www.datacamp.com/blog/ai-watermarking

13 Anthropic, *How Claude's text watermark works*, Anthropic, 14 sierpnia 2026 r. https://www.anthropic.com/news/claude-text-watermark

Cyberbezpieczeństwo i sabotaże

04

Dr **Jakub Kubś**

Doktor nauk humanistycznych, od 2021 roku pracuje w obszarze analizy dezinformacji. Jest autorem wielu raportów dotyczących rosyjskiej propagandy, skoordynowanego nieautentycznego zachowania w mediach społecznościowych (CIB) oraz pogłębionych analiz finansowania rosyjskich mediów. W pracy wykorzystuje analizę danych, wizualizację sieci powiązań, techniki i narzędzia OSINT oraz programowanie. Współzałożyciel fundacji Emerging Threats Observatory.

Niedziela, jesień 2027 roku, godzina 6:40. Za dwadzieścia minut otwierają się lokale wyborcze. Tymczasem w kraju zaczyna się paraliż komunikacyjny. Pociągi na głównych trasach nie mogą opuścić stacji. Nie doszło do wybuchu ani pożaru, nikogo nie ma na torach. W kilkudziesięciu miejscach ktoś przywiązał do szyn metalowe rury, które zwierają obwód torowy. System sterowania ruchem automatycznie uznaje tor za zajęty i zatrzymuje pociągi. W czasie kiedy służby próbują opanować kryzys, o 8:15 do kilkuset szkół w pięciu województwach trafiają maile o podłożonych ładunkach. Zwykle to w szkołach mieszczą się lokale wyborcze, więc komisje zarządzają przerwy w głosowaniu, a wyborcy czekają pod budynkami, aż pirotechnicy skończą sprawdzanie. Alarmy okazują się fałszywe, ale jeszcze przed południem kanały na Telegramie powtarzają, że głosowanie zostało zakłócone, a wybory są nieważne. Wieczorem przekaz dominuje w mediach społecznościowych, a w poniedziałek do Sądu Najwyższego trafiają pierwsze protesty wyborcze.

Ile kosztowała taka operacja? Zaledwie kilka tysięcy złotych. Wykonawcy działali w sposób skoordynowany, ale nigdy się nie spotkali. Zostali zwerbowani przez Telegram poprzez ogłoszenie o „prostym zleceniu", za które otrzymali przelew w kryptowalutach. Żaden nie wiedział, że pracował dla obcego wywiadu.

Ten scenariusz jest zmyślony, ale każdy jego element już się wydarzył, tylko nie tego samego dnia i nie w jednym kraju. 15 września 2026 roku w Holandii rury przywiązane do torów[1] w ponad trzydziestu miejscach sparaliżowały kolej w centrum i na wschodzie kraju, i to kilka godzin przed uroczystym otwarciem sesji parlamentu[2]. Sprawców na razie nie ustalono. W listopadzie 2025 roku pod Miką wysadzono tor na linii Warszawa–Lublin, a premier poinformował, że sprawcami byli Ukraińcy zwerbowani przez rosyjski wywiad[3]. Fale fałszywych alarmów bombowych[4] w polskich szkołach powtarzają się od lat. Za podobną kampanią w Czechach i na Słowacji w 2024 roku czeski wywiad widział rosyjski ślad[5]. Podczas wyborów w Mołdawii w 2025 roku fałszywe alarmy dotarły do lokali wyborczych w czterech niemieckich miastach[6], a niemieckie MSZ uznało je za skoordynowaną próbę powstrzymania diaspory przed głosowaniem.

Wszystkie te zdarzenia łączy jedno: zleceniodawca pozostaje poza zasięgiem, a ryzyko ponosi ktoś zwerbowany przez telefon. Dla bezpieczeństwa wyborów 2027 roku ten mechanizm jest groźniejszy niż jakakolwiek wyrafinowana technologia.

Diabeł nie mieszka w darknecie. Werbuje na Telegramie za pięćset złotych

W wyobraźni wielu decydentów to darknet jest miejscem, gdzie rodzą się takie operacje. Takie wyobrażenie jest nie tyle fałszywe, co mylące, i może prowadzić do złych decyzji.

Warto rozdzielić dwa pojęcia, które w debacie często się zle-

1 *Rail sabotage: Tubes tied to tracks in various locations; Collision near Steenwijk*, NL Times, 15 września 2026, https://nltimes.nl/2026/09/15/rail-sabotage-tubes-tied-tracks-various-locations-collision-near-steenwijk

2 *Sabotage disrupts Dutch train services ahead of state opening of parliament*, NBC News, 15 września 2026, https://www.nbcnews.com/world/europe/sabotage-disrupts-dutch-train-services-ahead-state-opening-parliament-rc-na597867

3 *Tusk: eksplozja miała najprawdopodobniej na celu wysadzenie pociągu na trasie Warszawa–Dęblin*, TVP Warszawa, 17 listopada 2025, https://warszawa.tvp.pl/90056257/tusk-eksplozja-miala-najprawdopodobniej-na-celu-wysadzenie-pociagu-na-trasie-warszawa-deblin

4 O. Klimczuk, *Fałszywe alarmy bombowe w polskich szkołach. Sprawcy nie mogą spać spokojnie*, CyberDefence24, 18 marca 2026, https://cyberdefence24.pl/armia-i-sluzby/policja/falszywe-alarmy-bombowe-w-polskich-szkolach-sprawcy-nie-moga-spac-spokojnie

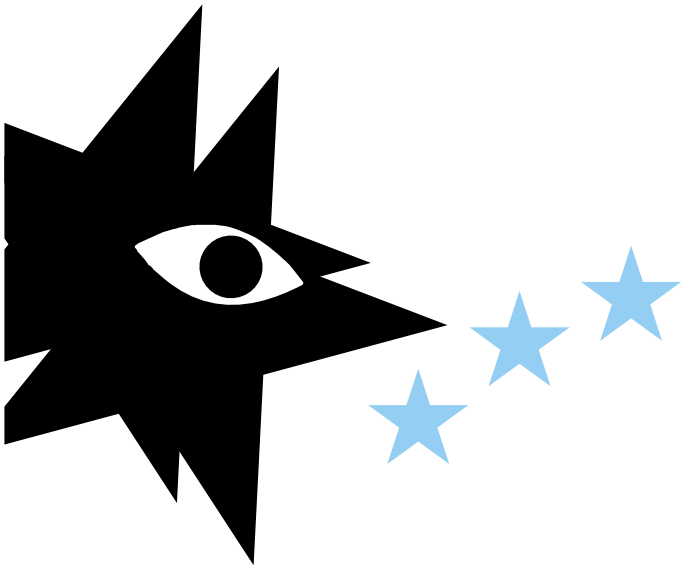
5 D. Basmat, *Russia behind recent bomb threats at Czech, Slovak schools, Czech intelligence chief says*, The Kyiv Independent, 20 listopada 2024, https://kyivindependent.com/russia-behind-recent-bomb-threats-at-czech-slovak-schools-czech-intelligence-chief-says/

6 Ch. Edwards, *Moldova warns Russia unleashing huge interference campaign to sway crucial election*, CNN, 26 września 2025, https://edition.cnn.com/2025/09/26/europe/moldova-election-russia-sandu-intl

wają. Darknet to część internetu dostępna wyłącznie przez specjalne oprogramowanie, takie jak sieć Tor, które ukrywa adres użytkownika i lokalizację serwera. Komunikator, taki jak Telegram, to z kolei zwykła aplikacja do wysyłania wiadomości, którą każdy pobiera ze sklepu w telefonie. W jego przypadku niczego się nie ukrywa, a szyfrowanie nie jest nawet domyślnie typu *end-to-end*: trzeba je włączyć ręcznie w „czacie tajnym", a kanały i grupy, w których odbywa się werbunek, nie mają go wcale. Żeby dotrzeć do rekrutera w darknecie, trzeba wiedzieć, gdzie szukać i mieć minimalną wiedzę techniczną. Żeby trafić na rekrutera na Telegramie, wystarczy aplikacja, którą większość mieszkańców Europy Wschodniej i tak ma już zainstalowaną w telefonie.

Najbardziej znanym przykładem darknetu jest sieć Tor, która powstała[7] w latach 90. w laboratorium badawczym amerykańskiej marynarki wojennej jako środek anonimowej komunikacji dla ludzi działających w dyskretny sposób w nieprzyjaznych państwach, czyli szpiegów. Kod udostępniono publicznie, bo anonimowość działa tylko w tłumie: gdyby z sieci korzystali wyłącznie agenci, sam fakt jej użycia by ich zdradzał. Im więcej osób korzysta z sieci Tor, tym bardziej jest ona anonimowa.

Polski darknet ma swoje fora, na których obok dyskusji o cyberbezpieczeństwie można znaleźć handel nielegalnymi substancjami i poradniki oszustw na BLIK-a. Dla służb takie miejsca bywają jednak bardziej wartościowe jako otwarte niż zamknięte. W 2017 roku holenderska policja przejęła darknetowy rynek Hansa[8] i przez miesiąc prowadziła go potajemnie. Gdy Amerykanie zamknęli konkurencyjny AlphaBay, jego użytkownicy przenieśli się wprost pod nadzór holenderskich śledczych, a liczba użytkowników Hansy wzrosła ośmiokrotnie. Wniosek jest prosty: obserwowana sieć to źródło informacji. Rozbita sieć rozprasza się na kilka mniejszych, których nikt jeszcze nie zmapował.



7 *History*, The Tor Project, 2019, https://www.torproject.org/about/history/

8 B. Krebs, *After AlphaBay's Demise, Customers Flocked to Dark Market Run by Dutch Police*, Krebs on Security, 20 lipca 2017, https://krebsonsecurity.com/2017/07/after-alphabays-demise-customers-flocked-to-dark-market-run-by-dutch-police/

Gdzie naprawdę toczy się gra

Werbunek dywersantów w Europie odbywa się przede wszystkim na Telegramie[9], czyli w aplikacji ani szczególnie bezpiecznej, ani ukrytej, za to popularnej na wschód od Bugu i praktycznie niemoderowanej.

Najnowsze dane ICCT i GLOBSEC[10] z lutego 2026 roku mówią o 151 incydentach dywersyjnych powiązanych z Rosją w Europie od lutego 2022 roku. Z tego 31 miało miejsce w Polsce, najwięcej ze wszystkich państw. Wśród 172 ustalonych sprawców około 95 procent to zwykli cywile bez formalnych związków z rosyjskimi służbami, tak zwani agenci jednorazowi werbowani właśnie przez komunikatory. Z tego wynika wniosek praktyczny: sam kontrwywiad nie wystarczy. Pulę potencjalnych wykonawców trzeba ostrzec, zanim ktoś złoży im ofertę.

Mechanika przypomina rynek pracy. Zaczyna się od zadania banalnego i taniego, jak zdjęcie obiektu, naklejka czy graffiti za kilkadziesiąt euro. Stawka i ryzyko rosną stopniowo, aż do podpalenia za kilka tysięcy. W Wielkiej Brytanii powiązane z Rosją kanały na Telegramie oferowały[11] 2500 funtów za spalenie radiowozu i 100 funtów za nagranie dewastacji meczetu. Zwerbowany często nie wie, kto płaci, a gdy się domyśli, bywa szantażowany zebranym wcześniej materiałem.

Ten sam rynek pracuje już na rzecz kampanii wyborczych. Przed wyborami w Niemczech w 2025 roku ktoś zatykał pianką montażową rury wydechowe setek aut i naklejał na szyby wizerunek kandydata Zielonych Roberta Habecka. Media najpierw obwiniały klimatycznych radykałów. Według „Spiegla" wykonawcy zostali zwerbowani przez komunikatory, za 100 euro od samochodu[12], a niemieckie służby podejrzewają Moskwę. W Mołdawii w 2024 roku sieć oligarchy Ilana Shora zapisywała chętnych przez czatbota na Telegramie[13], przydzielała im „zadania" i płaciła za głosy. Policja ustaliła[14] około 138 tysięcy odbiorców pieniędzy. Werbunek dywersanta, opłaconego wyborcy i opłaconego komentatora działa w oparciu o tę samą infrastrukturę. Przed wyborami parlamentarnymi w Mołdawii w 2025 roku zarejestrowana w Polsce agencja reklamowa rekrutowała na Facebooku „obserwatorów wyborczych" wśród mołdawskiej diaspory. Obiecywała 500 euro gwarantowanej zapłaty i do 30 tysięcy euro premii za zgłoszone „nieprawidło-

9 F. Bryjka, A.M. Dyner, A. Kozioł, *White Paper on Russian Acts of Sabotage and Subversion against Members of the Council of the Baltic Sea States*, PISM, 29 maja 2026, https://pism.pl/publications/white-paper-on-russian-acts-of-sabotage-and-subversion-against-members-of-the-council-of-the-baltic-sea-states

10 J. Lanchès, K. Rekawek, *More of the Same? Russia's Crime-Terror Nexus: Criminality as a Tool of Hybrid Warfare Revisited*, ICCT / GLOBSEC, 23 lutego 2026, https://icct.nl/publication/more-same-russias-crime-terror-nexus-criminality-tool-hybrid-warfare-revisited

11 G. Davis, EXCLUSIVE: *A Russia-linked Telegram network is inciting terrorism and is behind hate crimes in the UK*, HOPE not hate, 10 lutego 2025, https://hopenothate.org.uk/2025/02/10/exclusive-a-russia-linked-telegram-network-is-inciting-terrorism-and-is-behind-hate-crimes-in-the-uk/

12 *"Be Greener!" Climate sabotage in Berlin... by Russian agents?*, The Berliner, 6 lutego 2025, https://www.the-berliner.com/english-news-berlin/be-greener-russia-secret-service-robert-habeck-climate-sabotage/

13 E. Nilsson-Julien, *Moldovan officials accuse Kremlin of bribing voters ahead of EU membership referendum*, Euronews, 4 października 2024, https://www.euronews.com/2024/10/04/moldovan-officials-accuse-kremlin-of-bribing-voters-ahead-of-eu-membership-referendum

14 *Moldova police accuse Shor, Russia oligarch of $39M vote buying*, RFE/RL, 25 października 2024, https://www.rferl.org/a/moldova-police-accuse-shor-russia-oligarch-39m-vote-buying/33172951.html

wości". Zwerbowani trafiali do grup na Telegramie i WhatsAppie, a według raportu FIMI-ISAC[15] infrastruktura operacji była powiązana z objętą unijnymi sankcjami firmą Stark Industries. Taki model nagradza fałszywe i przesadzone zgłoszenia, które potem łatwo nagłośnić jako dowód sfałszowania wyborów.

Powtarzalność schematu i jego niski koszt pokazują, że nie mamy do czynienia z pojedynczymi incydentami, lecz ze stałym mechanizmem. Ze skali i charakteru działań wynika, że powinniśmy myśleć o werbunku jak o oddzielnym rynku pracy, który przyciąga osoby zdesperowane lub zradykalizowane. Kluczowe jest to, że istnieje pula ludzi, dla których 500 złotych za „prostą robotę" to sensowna oferta, a państwo nie ma z nimi żadnego kanału kontaktu.

Co robią inni

Ostatnie lata to nie tylko rosnąca skala problemów, ale też coraz więcej inicjatyw podejmowanych, aby im przeciwdziałać. Ukraina traktuje werbunek jak kampanię zdrowia publicznego. SBU prowadzi spotkania w szkołach[16], wysyła masowe SMS-y[17] i udostępnia czatbota[18] do zgłaszania prób werbunku. Przekaz jest prosty: jeśli nieznajomy obiecuje łatwe pieniądze za „proste zadanie", zgłoś to. Od wiosny 2024 roku zatrzymano[19] tam ponad 700 osób zwerbowanych do aktów sabotażu, w tym około 175 nieletnich.

Szwecja i Europol biorą na cel pośredników. Od lipca 2023 roku Szwecja karze[20] samo wciąganie osoby poniżej 18 lat w przestępstwo, niezależnie od jej zgody. W ramach Europolu działa grupa zadaniowa GRIMM, wymierzona w model „przemocy na zlecenie" i która zatrzymała[21] w ciągu roku 280 osób, w tym dziesiątki rekruterów.

Wielka Brytania zbudowała narzędzie prawne. National Security Act z 2023 roku objął działania pośredników niemających bezpośredniego związku z obcym państwem. Organizator podpalenia magazynu ze sprzętem dla Ukrainy został skazany na 17 lat pozbawienia wolności[22]. Taki wyrok jest jednocześnie komunikatem dla kolejnych potencjalnych pośredników.

Mołdawia działała przed wyborami, a nie po nich. Władze zablokowały kanały Szora na Telegramie[23], a przed wyborami parlamentarnymi w 2025 roku rozbiły sieć przekupionych wyborców[24].

Na te wyzwania brakuje natomiast rozwiązań na poziomie europejskim. Telegram nie jest[25] bardzo dużą platformą w rozumieniu aktu o usługach cyfrowych (DSA), bo deklaruje mniej niż 45 milionów użytkowników w UE. Dzięki temu nie musi oceniać i ograniczać ryzyk systemowych, w tym dla wyborów. Parlament Europejski zapytał[26] Komisję, kiedy ta ocena zostanie zrewidowana. Od jej odpowiedzi zależy, czy Telegram zostanie objęty działaniami systemowymi.

Polska powinna domagać się od Telegrama obowiązków na miarę zasięgu: popierać w UE rewizję jego statusu w DSA, a przed wyborami ustalić z platformą szybką ścieżkę zgłaszania kanałów werbunkowych.

Rekomendacje

Przed wyborami w 2027 roku warto nie mylić narzędzia z problemem. Pomysły zakazu anonimizacji czy szyfrowania brzmią politycznie atrakcyjnie, ale nie trafiłyby w cel, bo werbunek odbywa się w zwykłej, powszechnie dostępnej aplikacji. Uderzyłyby natomiast w dziennikarzy, sygnalistów i badaczy bezpieczeństwa, a przeciwnika przeniosłyby tam, gdzie go nie widać. Lepszą inwestycją jest zdolność do obserwacji tych środowisk i wykorzystanie ludzi, którzy potrafią się w nich poruszać. Państwo musi umieć ich zarówno zatrudnić, jak i zatrzymać. Dlatego poniższe propozycje koncentrują się na ludziach i kanałach komunikacji, a nie na zakazach technologii.

1. **Reakcje w dniu wyborów** muszą być wielokrotnie przećwiczone. PKW, Krajowe Biuro Wyborcze, MSWiA i policja powinny przed 2027 rokiem zasymulować działania w scenariuszach takich jak te dotyczące masowych alarmów, tak aby jeden mail rozesłany do setek lokali nie zamykał ich wszystkich. Potrzeb-

15 *Moldova Election Country Report 2025*, FIMI-ISAC, 5 grudnia 2025, https://www.fimi-isac.org/reports/Moldova-Election-Country-Report-2025.pdf

16 A. Barkhush, T. Frolova, *Ukraine launches „Expose the FSB agent" campaign in schools to counter Russian recruitment tactics*, United24 Media, 1 maja 2025, https://united24media.com/latest-news/ukraine-launches-expose-the-fsb-agent-campaign-in-schools-to-counter-russian-recruitment-tactics-8033

17 *The SBU is issuing warnings about enemy recruitment following the terrorist attack in Bucha*, UA.News, 23 marca 2026, https://ua.news/en/ukraine/sbu-rozsilaie-poperedzhennia-pro-verbuvannia-vorogom-pislia-teraktu-v-buchi

18 V. Mereshchuk, *Olena Semykina: "The SBU chat-bot on recruitment attempts has already received over 18,000 reports."*, LB.ua, 25 lutego 2026, https://en.lb.ua/news/2026/02/25/37548_olena_semykina_sbu_chatbot.html

19 Ch. Miller, *Russia grooms Ukrainian teens as spies and saboteurs*, Financial Times, 29 czerwca 2025, https://www.ft.com/content/5cb5ebf9-bfb3-4a50-bc2f-79540baf8f87

20 Library of Congress, *Sweden: Law Criminalizing Recruitment of Children into Criminal Activity Enters into Force*, Global Legal Monitor, 11 września 2023, https://www.loc.gov/item/global-legal-monitor/2023-09-11/sweden-law-criminalizing-recruitment-of-children-into-criminal-activity-enters-into-force/

21 G. Blackburn, *Europol GRIMM taskforce nets 280 arrests as violence-for-hire spreads across Europe*, Euronews, 29 kwietnia 2026, https://www.euronews.com/2026/04/29/europol-grimm-taskforce-nets-280-arrests-as-violence-for-hire-spreads-across-europe

22 Crown Prosecution Service, *Updated sentence: How CPS used new National Security Act legislation to prosecute plot*, cps.gov.uk, 24 października 2025, https://www.cps.gov.uk/cps/news/updated-sentence-how-cps-used-new-national-security-act-legislation-prosecute-plot

23 M. Necsutu, *Moldova takes down pro-Kremlin oligarch Shor's Telegram accounts*, Balkan Insight, 11 października 2024, https://balkaninsight.com/2024/10/11/moldova-takes-down-pro-kremlin-oligarch-shors-telegram-accounts/bi/

24 G. Chizhov, *The $39 Million Election*, Free Russia Foundation (Substack), 30 czerwca 2026, https://freerussiafoundation.substack.com/p/the-39-million-election

25 H. Virkkunen, *Answer given by Executive Vice-President Virkkunen on behalf of the European Commission*, Parlament Europejski, 19 maja 2025, https://www.europarl.europa.eu/doceo/document/E-10-2025-001293-ASW_EN.html

26 U. Paet, *Curbing the activities of Telegram*, Parlament Europejski, 12 lutego 2026, https://www.europarl.europa.eu/doceo/document/P-10-2026-000628_EN.html

na jest szybka weryfikacja zagrożeń i gotowy komunikat PKW, który trafi do wyborców przed fałszywą narracją.

---

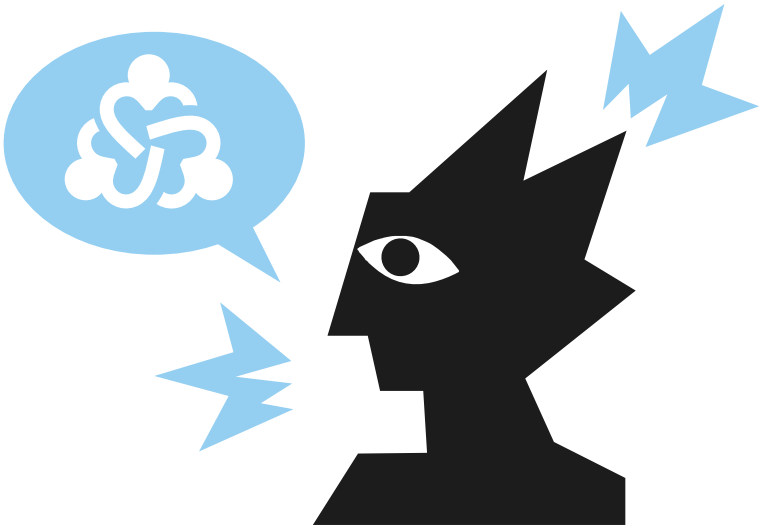
2. Celem uderzeń musi być **warstwa pośredników**. Warto rozważyć odrębną odpowiedzialność karną za wciąganie w przestępstwo, zarówno w odniesieniu do nieletnich (rozwiązanie szwedzkie), jak i do pośredników, którzy organizują sabotaż na zlecenie, nie mając bezpośredniego związku z obcym państwem (rozwiązanie brytyjskie). Rekruterzy działają ponad granicami, dlatego potrzebna jest także współpraca z partnerami, na przykład w ramach grupy GRIMM w Europolu. Surowe i nagłośnione wyroki dla organizatorów budują odstraszanie wśród kolejnych pośredników.

---

3. Komunikacja państwa stale musi się dostosowywać się do **dynamiki środowiska informacyjnego**. W grudniu 2025 roku ABW uruchomiła[27] na Telegramie czatbota do zgłaszania dywersji i prób werbunku. To trafna decyzja, ale brakuje kampanii, dzięki której dowie się o nim magazynier z Rzeszowa i licealista z Lublina, a nie tylko czytelnicy portali branżowych. Profilaktykę trzeba prowadzić w tych samych kanałach, w których odbywa się werbunek: na Telegramie, TikToku i w grupach ogłoszeniowych, po polsku, ukraińsku, rosyjsku i białorusku. Przekaz powinien być arytmetyczny, a nie tylko moralalizujący: obiecali ci tysiąc euro, a dostaniesz co najmniej dziesięć lat[28] za dywersję na rzecz obcego wywiadu. Rolę mają tu też media publiczne i organizacje społeczne.

---

4. Czujność zachować muszą **komitety i organizacje obserwatorskie**. Komitety wyborcze i organizacje delegujące obserwatorów społecznych powinny weryfikować, kogo zgłaszają do komisji, i sygnalizować podejrzane oferty „pracy" przy wyborach. PKW powinna z kolei być przygotowana na falę zgłoszeń o rzekomych nieprawidłowościach, z których część może być zamówiona.



---

27 Sz. Palczewski, *ABW stworzyło chatbota do zgłaszania przypadków dywersji*, CyberDefence24, 18 grudnia 2025, https://cyberdefence24.pl/armia-i-sluzby/abw-stworzylo-chatbota-do-zglaszania-przypadkow-dywersji

28 Art. 130. *Udział w działalności obcego wywiadu; szpiegostwo*, w: Ustawa z dnia 6 czerwca 1997 r. Kodeks karny, Dz.U. 1997 nr 88 poz. 553.

Finansowanie i przejrzystość reklamy oraz kampanii politycznej

05

Dr **Sonia Horonziak**

Doktor nauk politycznych i administracji, dyrektorka programu Demokracji i Społeczeństwa Obywatelskiego w Instytucie Spraw Publicznych oraz wykładowczyni na Uniwersytecie Warszawskim. Absolwentka Uniwersytetu Jagiellońskiego w Krakowie oraz badawczego programu stypendialnego na Uniwersytecie Technicznym w Dreźnie. Autorka licznych publikacji w zakresie polaryzacji politycznej, teorii demokracji oraz procesów wyborczych.

Wyobraźmy sobie sytuację, w której tuż przed końcem kampanii wyborczej partia X decyduje się na wypuszczenie do sieci szeregu reklam oczerniających jej przeciwników politycznych. Głoszone w tych reklamach treści nie są prawdziwe, jednak partia X nie robi tego za pośrednictwem własnych kanałów tylko prosi o to zaprzyjaźnione jednostki i instytucje, które wykładają na to setki tysięcy złotych zdecydowanie przekraczając wszelkie limity wydatków na kampanię. Materiały nie są w żaden sposób oznaczone, a sposób ich skonstruowania (udawane wywiady, wyniki sondy, zainscenizowane sytuacje) sprawia, że platformy społecznościowe nie rozpoznają ich jako reklamy politycznej, a treść społeczną i dopuszczają do dalszego obrotu zyskując tysiące wyświetleń. Choć zakamuflowana kampania szybko zostaje zidentyfikowana przez pozostałe partie startujące w wyborach, to organy wyborcze nie mają narzędzi by zareagować natychmiastowo, a platformy społecznościowe usuwają jedynie część rozpowszechnionych treści. Partia X na ostatnim odcinku kampanii uzyskuje nieznaczną przewagę i ostatecznie wygrywa wybory, gdyż część wyborców do których dotarły treści kampanii, postanowiło nie oddać swojego głosu na inne partie. Choć wszyscy wiedzą, że to partia X stoi za tą kampanią, to jednocześnie jej sprawozdanie finansowe z kampanii wyborczej, złożone po wygranej w wyborach, nie wykazuje żadnych nieprawidłowości. Ostatecznie, dopiero po kilku tygodniach od zakończenia kampanii wyborczej, wszystkie nieprawidłowe reklamy zostają usunięte.

Scenariusz ten, choć ekstremalny, nie jest jednak całkowicie abstrakcyjny. Choć jest szansa, że osoby stojące za oczerniającymi treściami zostałyby wykryte, a platformy społecznościowe zareagowałyby na zgłoszenia natychmiast, to jednak cały system funkcjonuje w granicach niepewności. Kontrola kampanii i reklamy politycznej w Internecie to przede wszystkim kwestia określenia limitów i granic wydatków, bieżącego ich monitoringu oraz możliwości natychmiastowej reakcji na nieprawidłowości. Tymczasem kwestie te w polskim systemie ciągle pozostawiają wiele do życzenia.

Wyzwania polskiego prawa wyborczego

Jednym z największych wyzwań przed którymi stoi polskie prawo w kontekście wyborów jest brak **sprawnego systemu bieżącego monitoringu finansowania kampanii wyborczych**. Problem ten destabilizuje przejrzystość prowadzenia kampanii zarówno w zakresie ograniczeń dostępu do informacji o podmiotach, które angażują się finansowo w kampanię, jak i tego kiedy i w jaki sposób podmioty kontrolujące mogą analizować wydatki kampanijne komitetów wyborczych, a co za tym idzie reagować na potencjalne nieprawidłowości. Ponadto, przeniesienie w znacznej mierze kampanii do Internetu sprawia, że większość obecnych przepisów nie przystaje do rzeczywistości dynamiki kampanii w sieci. Tym samym, zwiększa się pole do nadużyć oraz obniża odpowiedzialność jaką prowadzący kampanię powinni ponosić za łamanie prawa.

Pierwsze z wyzwań wynika ze zmienionych w 2018 roku przepisów Kodeksu wyborczego, przede wszystkim brzmienia art. 106 §1, określającego kto może prowadzić agitację wyborczą oraz uchylenia art. 499 wprowadzającego sankcje za nieprawidłowe wydatkowanie środków na kampanię. Od tego momentu każdy obywatel zyskał prawo angażowania się w agitację wyborczą bez uzyskania zgody pełnomocnika wyborczego, co *de facto* doprowadziło do odsunięcia odpowiedzialności komitetu wyborczego za agitację określonych jednostek. Tym samym, przepisy kontrolujące i limitujące wydatki komitetów wyborczych na kampanię przestały obejmować nieograniczoną liczbę obywateli, którzy mogą włączać się swobodnie w agitację na rzecz danej partii. Otworzyło to ścieżkę do pozaformalnego wspierania kampanii wyborczych, gdzie osoby prywatne i stojące za nimi instytucje mogą wspierać kampanię bez żadnej odpowiedzialności danego komitetu wyborczego, nawet w przypadku omijania przepisów. Co więcej, obecne brzmienie przepisu sprawia, że nawet zarejestrowany kandydat finansując wydatki na swoją własną kampanię może ominąć przepisy dotyczące ewidencji, ze względu na fakt, że agituje jako „zwykły obywatel". Praktyki wykorzystywania tych przepisów można było obserwować w wyborach każde-

go szczebla: parlamentarnych w 2023 roku, do Parlamentu Europejskiego w 2024 czy w wyborach samorządowych w 2024 roku[1].

Drugim wyzwaniem, szczególnie widocznym w kontekście przeniesienia się w ostatnich latach kampanii wyborczych do sieci, dotyczy **przestarzałych przepisów odnośnie do sprawozdań finansowych** komitetów wyborczych z przeprowadzonej kampanii. Przepisy *Kodeksu wyborczego* określają[2], że komitet wyborczy ma obowiązek złożyć sprawozdanie w terminie trzech miesięcy po dniu głosowania, na dosyć ogólnych papierowych szablonach. Oznacza to, że większość wydatków[3], które ponoszą komitety wyborcze na kampanię pozostaje niedostępna nie tylko dla obywateli, ale również dla samej administracji wyborczej. Ponadto, nawet złożone sprawozdanie umożliwia kontrolę jedynie w ograniczonym stopniu, biorąc pod uwagę że zawiera się w setkach czy tysiącach stron dokumentacji papierowej, których analiza stanowi wyzwanie dla samego Krajowego Biura Wyborczego, nie mówiąc o indywidualnych obywatelach, którzy przecież mają prawo dostępu do tych informacji i mogliby chcieć przeprowadzić własną kontrolę. W tym wypadku, identyfikacja i reakcja na jakiekolwiek nieprawidłowości, nawet w przypadku reklam w przestrzeni fizycznej jest mocno ograniczona, nie mówiąc o specyfice reklam w Internecie.

Równie istotną kwestią są narzędzia i zasoby, którymi dysponują organy wyborcze mające prowadzić kontrolę prowadzenia kampanii. Wykrywalność naruszeń i nieprawidłowości, zwłaszcza w przypadku kampanii w sieci, pozostaje ograniczona przez **brak spójnego systemu współpracy** między organami wyborczymi, platformami internetowymi i innymi podmiotami monitorującymi kampanię np. społecznymi. Obecnie decyzyjność w przypadku naruszenia przepisów prowadzenia i wydatkowania na kampanię w dużej mierze opiera się na zasadach i reakcji prywatnych platform internetowych, które niekoniecznie działają w zgodzie z przepisami i standardami krajowymi. Po zapowiedziach i ostatecznym wdrożeniu przez Unię Europejską rozporządzenia w sprawie przejrzystości i targetowania reklamy politycznej tzw. TTPAR[4] wiele platform internetowych uznało, że zamiast dostosowywać się do surowych i obarczonych ryzykiem wysokich kar przepisów, wolą ograniczyć lub wręcz zakazać reklamy politycznej (m.in. przez Meta, X/Twitter czy TikTok), odsuwając tym samym swoją odpowiedzialność za kontrolę wyświetlanych treści. Nie przełożyło się to jednak na faktyczny brak tego typu treści na tych platformach, a co więcej umożliwiło niekontrolowane rozproszenie reklam omijających przepisy platformowe bez możliwości ich kontroli. Jednym z najbardziej wyraźnych przykładów tego typu działań jest kampania parlamentarna na Węgrzech w 2026 roku, gdzie niezależni badacze zidentyfikowali setki płatnych reklam politycznych emitowanych przez partię Fidesz i podmioty z nią związane na platformach firmy META, którym udało się ominąć zasady[5]. Ze względu na zakaz reklamy politycznej, faktyczne możliwości identyfikacji i kontroli takich reklam stały się realnym wyzwaniem, a reakcja samej platformy była na tyle opóźniona, że zakamuflowane reklamy mogły dotrzeć do setek tysięcy użytkowników przed ich usunięciem lub wstrzymaniem.

Rozwiązania istnieją, należy tylko po nie sięgnąć[6]

Choć finansowanie i kontrola prowadzenia kampanii wyborczej w sieci to problem systemowy i wymaga wdrożenia rozwiązań na różnych szczeblach zarządzania procesem wyborów, to możliwe do wprowadzenia zmiany, niekoniecznie muszą oznaczać przebudowę całego systemu. Istnieje również wiele rozwiązań w bliskich nam geograficznie krajach, gdzie od lat już funkcjonują spójne systemy monitoringu wydatków kampanii.

W przypadku Polski, przede wszystkim należałoby wprowadzić **wymóg uzyskania zgody pełnomocnika komitetu wyborczego na prowadzenie agitacji związanej z wydatkowaniem pieniędzy**. Przepływ finansów wokół kampanii byłby dzięki temu powiązany z danym komitetem, a co za tym idzie środki te mogłyby zostać rozliczone i sprawozdane bez ograniczania możliwości agitacji przez obywateli. Dzięki temu organy wyborcze mogłyby sprawniej wykrywać nieprawidłowości i ustalać czy dany komitet przekroczył ustawowe limity wydatków. W zdecydowany sposób pomogłoby to również kontrolować przebieg kampanii w sieci, gdzie w szczególności widoczne jest w ostatnich latach prowadzenie nieformalnej kampanii, która nie podlega kontroli. Przykładem może być kampania prezydencka 2025, gdzie dodatkowe działania reklamowe na rzecz kandydatów były prowadzone przez szereg innych kont wspierających, z których niektóre podejrzewano o bycie finansowanymi ze źródeł zagranicznych[7]. Wiele z tych reklam nie było w żaden sposób oznaczone, a osoby za nimi stojące oraz skala wydatków pozostawała jedynie częściowo możliwa do ustalenia. Z tego względu związanie wydatków na kampanię z danym komitetem oraz obowiązkowe oznaczanie źródeł finansowania reklam politycznych powinno być jedną

1 T. Gąsior, A. Kobylińska, P. Wachowiec, *Jak poprawić przejrzystość finansowania polityki. Propozycje zmian prawnych*, Instytut Spraw Publicznych, Warszawa 2025, https://www.isp.org.pl/pl/publikacje/jak-poprawic-przejrzystosc-finansowania-polityki-propozycje-zmian-prawnych

2 Ustawa z dnia 5 stycznia 2011 r. – Kodeks wyborczy (t.j. Dz. U. z 2025 r. poz. 365, z późn. zm.), art. 142.

3 Wyjątek stanowi rejestr wpłat od osób fizycznych oraz rejestr kredytów, który jest dostępny w trakcie trwania kampanii wyborczej.

4 Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2024/900 z dnia 13 marca 2024 r. w sprawie przejrzystości i targetowania reklamy politycznej, https://eur-lex.europa.eu/legal-content/PL/ALL/?uri=CELEX:32024R0900

5 E. Nilsson-Julien, T.Paternoster, *Orbán's Fidesz bypassing Meta's political advertising rules, researchers warn*, https://www.euronews.com/my-europe/2026/04/08/orbans-fidesz-bypassing-metas-political-advertising-rules-researchers-warn

6 Zaproponowane w tekście rozwiązania stanowią jedynie część szerszego katalogu propozycji zmian mających na celu usprawnienie procesu wyborczego w Polsce. W kontekście bardziej kompleksowych rozwiązań warto zapoznać się z publikacjami:

- Równe i powszechne v. 2.0: społeczne propozycje zmian w prawie wyborczym, Fundusz Civitates dla Grupy Wybory, Warszawa 2025, https://www.isp.org.pl/pl/publikacje/rowne-i-powszechne-v-2-0-spoleczne-propozycje-zmian-w-prawie-wyborczym.

- T. Gąsior, A. Kobylińska, P. Wachowiec, *Jak poprawić przejrzystość finansowania polityki. Propozycje zmian prawnych*, Instytut Spraw Publicznych, Warszawa 2025, https://www.isp.org.pl/pl/publikacje/jak-poprawic-przejrzystosc-finansowania-polityki-propozycje-zmian-prawnych.

7 S. Horonziak, A. Mierzyńska, D. Owczarek, M. Pańków, R. Załęski, *Algorytmy i agendy polityczne: cyfrowa walka o prezydenturę Polski w 2025 r.*, https://www.isp.org.pl/pl/publikacje/algorytmy-i-agendy-polityczne-cyfrowa-walka-o-prezydenture-polski-w-2025-r

z pierwszych zmian przepisów.

Kolejną kwestią jest **wdrożenie systemu bieżącego monitoringu wydatków kampanijnych oraz cyfryzacja procesu sprawozdawczości wydatków za kampanię**. Polska, jako kraj przodujący w rozwiązaniach dotyczących systemów internetowych na poziomie rządowym, ma wszelkie narzędzia by wprowadzić system automatycznie pozyskujący informacje o wydatkach czy wpływach na kampanię danego komitetu wiążąc go np. z indywidualnym dla komitetu kontem bankowym czy też wprowadzanymi na bieżąco fakturami czy rachunkami. Nie jest to zresztą rozwiązanie innowacyjne. Litwa ma swój własny system ewidencji księgowej finansowania kampanii politycznej w postaci elektronicznego dziennika księgowego[8]. Skarbnik danego uczestnika kampanii politycznej ma obowiązek bieżącego wprowadzania danych o darowiznach, wydatkach i wszelkich zobowiązaniach finansowych. Co istotne, dane te są publikowane i dostępne na portalu, dzięki czemu każdy wyborca może zapoznać się z tymi informacjami praktycznie w czasie rzeczywistym[9]. Jest to niezwykłe ułatwienie dla samych uczestników kampanii, a system automatycznie generuje raport z takiej kampanii, oznacza błędy do poprawienia i niejako rozlicza komitet. Alternatywne rozwiązanie, pozwalające również na bieżącą kontrolę wydatkowania przedstawia Słowacja. Każdy podmiot prowadzący kampanię ma przypisane tzw. konto transparentne, czyli wyznaczony rachunek bankowy, na którym prowadzone są wszystkie obroty kampanijne, a którego historia transakcji jest publicznie dostępna w Internecie w trakcie trwania kampanii[10]. Choć system ten nie jest całkowicie odporny na nieprawidłowości, to jednak znacząco zwiększa poziom przejrzystości finansowania kampanii i możliwości bieżącej reakcji. Zarówno z perspektywy obywatela, jak i administracji wyborczej oba te rozwiązania dają większość jawność prowadzenia kampanii oraz pełniejszą kontrolę wydatków jej uczestników. W związku z dynamiką kampanią w sieci i rozprzestrzeniania się bezprawnych materiałów wyborczych, istotna w tym kontekście jest możliwość szybkiej korekty nieprawidłowości.

Finalną kwestią jest wprowadzenie w Polsce **spójnej polityki cyfrowej**, która wyposażyłaby odpowiednie organy w narzędzia i kompetencje do reakcji na dynamicznie zmieniający się rynek przestrzeni cyfrowej. Wdrożenie *Aktu o usługach cyfrowych*[11] wraz z wyznaczeniem Koordynatora ds. Usług Cyfrowych to niezbędne pierwsze kroki, ale ich znaczenie zależy od zapewnienia skutecznego stosowania nowych regulacji w praktyce.

Kluczowe są nie tylko odpowiednie zasoby kadrowe i narzędzia, lecz także sprawne mechanizmy reagowania, współpracy między instytucjami i dostosowywania działań do zmieniających się form aktywności w sieci. Dopiero takie podejście może zapewnić organom państwowym realną zdolność do reagowania na skoordynowane kampanie prowadzone w przestrzeni cyfrowej.

Jeszcze parę lat temu sztuczna inteligencja nie odgrywała istotnej roli w procesie wyborczym, a obecnie nawet najbardziej popularne chatboty oskarżane są o podawanie fałszywych informacji związanych z procesem wyborczym[12]. Monitoring zagrożeń dla procesów demokratycznych, nie powinien skupiać się więc jedynie na działaniach dezinformacyjnych określonych aktorów, ale na pojawiających się nowych sposobach manipulacji wyborczej utrudniających kontrolę nad wydatkami finansów publicznych. Tym samym polityka cyfrowa w Polsce powinna opierać się na przeprowadzaniu regularnej modernizacji swoich narzędzi, testowaniu określonych rozwiązań oraz aktualizowaniu katalogu zagrożeń wraz z propozycją innowacji. Obecna kontrola, następująca długo po zakończeniu kampanii wyborczej, jest niewystarczająca i nie chroni przed zagrożeniami i oszustwami kampanii w sieci. Niezbędne jest też uregulowanie kwestii reklamy politycznej z największymi platformami społecznościowymi oraz usprawnienie systemu reakcji na nieprawidłowe treści omijające przepisy.

Rekomendacje:

1. Konieczne jest powiązanie finansowania agitacji wyborczej z komitetami wyborczymi. **Należy wprowadzić obowiązek powiązania płatnej agitacji** z konkretnym komitetem wyborczym oraz obowiązkowego oznaczania źródła finansowania reklam politycznych. Pozwoliłoby to lepiej kontrolować wydatki, ograniczać obchodzenie limitów i zwiększyć przejrzystość nieformalnych działań w sieci.

2. **Cyfrowy monitoring finansowania kampanii** powinien być prowadzony na bieżąco. Polska, jako państwo o dobrze rozwiniętej infrastrukturze cyfrowej i doświadczeniu w cyfryzacji usług publicznych, ma odpowiednie zaplecze, aby wdrożyć taki system. Warto wykorzystać istniejące rozwiązania i połączyć je z dobrymi praktykami innych państw, np. litewskim elektronicznym dziennikiem księgowym czy słowackimi rachunkami transparentnymi. Pozwoliłoby to na bieżące rejestrowanie, udostępnianie i kontrolowanie przychodów oraz wydatków kampanijnych.

3. Wzmocnienie cyfrowego nadzoru nad kampanią. Potrzebny jest **elastyczny system monitorowania kampanii interneto-**

8 S. Horonziak, *Przejrzystość finansowania kampanii wyborczych – przegląd rozwiązań zagranicznych*, Instytut Spraw Publicznych, Warszawa 2025.

9 rinkejopuslapis.lt – litewska strona zaawansowanego świadczenia usług online dla wyborców.

10 transparentneucty.sk – słowacka strona z danymi ze wszystkich kont transparentnych na Słowacji.

11 Rozporządzenie (UE) 2022/2065 w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (akt o usługach cyfrowych), https://eur-lex.europa.eu/eli/reg/2022/2065/oj/pol

12 C. Weinmann, D. Allen, O. Denkovski *Inconsistent and Unreliable: Chatbots Provide Inaccurate Information on German Elections*, Democracy Reporting International 2025, https://democracy-reporting.org/en/office/global/publications/inconsistent-and-unreliable-chatbots-provide-inaccurate-information-on-german-elections

wych, który będzie mógł reagować na nowe formy agitacji, reklamy politycznej czy wykorzystania sztucznej inteligencji. Polska może rozwijać takie rozwiązania, opierając je na już funkcjonujących systemach cyfrowych i administracyjnych oraz regularnie aktualizując narzędzia i katalog zagrożeń.

Hiperpersonalizacja

06

Aleksandra Wójtowicz

Analityczka ds. nowych technologii i cyfryzacji. Jej zainteresowania badawcze obejmują cyfrowe regulacje, cyberbezpieczeństwo i dezinformację. Stypendystka Europejskiego Forum Alpbach, Transatlantic Security Initiative IRI i Humanity in Action. Wcześniej związana m.in. z NASK, WiseEuropa, Polityką Insight czy ECFR. Absolwentka Uniwersytetu Warszawskiego, studiowała też na George Mason University w Waszyngtonie i na Uniwersytecie Karola w Pradze.

Jest wrzesień 2027 roku, ostatni tydzień kampanii wyborczej. W jednym z okręgów sztab kandydata w ciągu jednego popołudnia generuje czterysta wariantów jednego spotu. Model predykcyjny podzielił wcześniej elektorat na kilkadziesiąt bardzo konkretnych grup. Każda z nich otrzymuje inny komunikat, jak najbardziej dostosowany do jej potrzeb.

Tydzień później dwoje znajomych, Gajusz i Penelopa, odkrywają, że głosowali na tego samego kandydata, ale zupełnie z innych względów. Gajusz zdecydował się oddać na niego głos, ponieważ ten obiecał, że obniży podatki. Za to Penelopa liczy na zdecydowane działania w ochronie zdrowia i przede wszystkim większe wydatki na szpitale powiatowe. Co do podatków, to uważa, że powinny być wyższe, szczególnie dla najbogatszych, ale o tym, że jej kandydat ma zdanie odmienne, nawet nie wie. Algorytm pokazywał jej bowiem tylko jedną wersję kandydata, ściśle dopasowaną do jej profilu.

Taki scenariusz jest wynikiem coraz szerszego wykorzystania sztuczna inteligencja wykorzystywana będzie do hiperpersonalizacji kampanii. Choć ryzyko, że partia przedstawi się różnym wyborcom jako inne ugrupowanie „jednej sprawy", literatura opisywała jeszcze przed erą generatywnej AI (Zuiderveen Borgesius i in. 2018)[1], to AI zmienia skalę tego zjawiska. Narzędzia AI mogą bowiem być łatwo wykorzystywane do produkcji i dystrybucji treści tak, by te jak najlepiej odpowiadały na potrzeby poszczególnych wyborców.

Czym jest hiperpersonalizacja

Jednym z największych wyzwań przed którymi stoi polskie Hiperpersonalizacja to dopasowywanie treści nie do szerokiej grupy odbiorców, tylko do jej wąskich segmentów, a potencjalnie także do pojedynczych osób. Samo dzielenie elektoratu na mniejsze grupy nie jest nowe: mikrotargetowanie stosowano w kampaniach od dawna. Ujawniona w 2018 r. afera Cambridge Analytica pokazała, że w kampaniach z 2016 r. do typowania wyborców według cech osobowości wykorzystano dane nawet 87 mln użytkowników Facebooka, pozyskane bez ich wiedzy. Sztuczna inteligencja wzmacnia jednak to zjawisko. Po pierwsze, radykalnie obniża koszt przygotowania osobnego komunikatu dla każdej grupy, co dotąd było głównym hamulcem mikrotargetowania. Po drugie, chatbot może prowadzić z wyborcą indywidualny dialog i na bieżąco dobierać argumenty, a więc reagować na potencjalne wątpliwości wyborców.

Gdy każda grupa widzi inną wersję kandydata, ogranicza się wyborcom możliwość oceny całego programu: kandydat może składać różnym grupom obietnice trudne do pogodzenia. Niesie to ryzyko fragmentacji debaty publicznej i selektywnego informowania wyborców. Taka kampania jest też dużo trudniejsza w kontekście kontroli i obserwacji. Czterystu wariantów spotu nie zobaczą dziennikarze, konkurenci ani organy wyborcze, szczególnie jeśli wszystkie te warianty nie trafią do publicznego repozytorium - a nawet wtedy liczba danych do analizy będzie po prostu znacznie większa niż w przypadku "tradycyjnych" kampanii. Sprzecznej lub fałszywej obietnicy nie da się sprostować, gdy nie wiadomo, komu ją złożono.

Oddzielnym problemem jest też kwestia bezpieczeństwa danych użytych do hiperpersonalizacji. Hiperpersonalizacja wymaga gromadzenia w jednym miejscu szczegółowych profili milionów wyborców. Taka baza jest celem dla przestępców i obcych służb, a jej wyciek szkodzi bardziej niż utrata zwykłych danych kontaktowych - a jednocześnie warto podkreślić, że w trakcie kampanii wyborczej ryzyka związane z cyberbezpieczeństwem są szczególnie wysokie. Profilowanie oznacza też wykorzystywanie danych wrażliwych dotyczących kwestii takich jak np. czyjeś zdrowie, na co odbiorcy skierowanych do nich komunikatów wcześniej świadomie się nie zgodzili. Co więcej, tego typu spersonalizowane kampanie mogą być de facto sposobem na obchodzenie lokalnych praw wyborczych. Chodzi tu na przykład o treści przygotowane przez sztab, a opublikowane przez sympatyka lub influencera „od siebie", która nie jest oznaczona jako materiał komitetu. Wszystko to negatywnie wpływa na transparentność kampanii wyborczej.

1 F.J. Zuiderveen Borgesius, J. Möller, S. Kruikemeier, R. Ó Fathaigh, K. Irion, T. Dobber, B. Bodó, C. de Vreese, *Online Political Microtargeting: Promises and Threats for Democracy*, „Utrecht Law Review" 2018, t. 14, nr 1, s. 82–96

Poszczególne fazy a odpowiedzi regulacyjne

Gdy każda grupa widzi inną wersję kandydata, ogranicza się wyborcom możliwość oceny całego programu: kandydat może składać różnym grupom obietnice trudne do pogodzenia. Niesie to ryzyko fragmentacji debaty publicznej i selektywnego informowania wyborców. Taka kampania jest też dużo trudniejsza w kontekście kontroli i obserwacji. Czterystu wariantów spotu nie zobaczą dziennikarze, konkurenci ani organy wyborcze, szczególnie jeśli wszystkie te warianty nie trafią do publicznego repozytorium - a nawet wtedy liczba danych do analizy będzie po prostu znacznie większa niż w przypadku "tradycyjnych" kampanii. Sprzecznej lub fałszywej obietnicy nie da się sprostować, gdy nie wiadomo, komu ją złożono.

Zjawisko to składa się z trzech istotnych elementów, a mianowicie profilowania psychologicznego, produkcji, a także dystrybucji.

Podczas **profilowania** narzędzia analityczne budują profile wyborców obejmujące nie tylko dane demograficzne, lecz także wartości, lęki i wrażliwość na konkretne argumenty. Profil to zbiór informacji o danej osobie, z którego kampania korzysta, by zdecydować, co i jak jej powiedzieć. Profil demograficzny obejmuje wiek, płeć, miejsce zamieszkania czy wykształcenie. Profil psychograficzny opisuje wartości, styl życia, lęki i cechy osobowości, np. czy dana osoba silniej reaguje na argumenty odwołujące się do bezpieczeństwa czy do zmiany. Powstaje on ze śladów cyfrowych, np. polubień, zakupów online czy przeglądanych stron, które model statystyczny przekłada na prawdopodobieństwo określonych postaw.

Firma Resonate udostępniła w 2024 r. na platformie reklamowej Yahoo ponad tysiąc gotowych segmentów odbiorców, dzieląc ich według zachowań wyborczych, skłonności do głosowania czy sympatii partyjnych. Resonate łączy odpowiedzi z własnego, prowadzonego na bieżąco badania konsumenckiego (ponad 100 mln odpowiedzi) z miliardami obserwowanych codziennie zachowań online, a jej system rAI przypisuje na tej podstawie ponad 250 mln dorosłych Amerykanów ponad 15 tys. cech – od danych demograficznych, przez wartości, aż po sympatie polityczne – i wiąże je z identyfikatorami urządzeń, tak by reklamę można było skierować od razu[2].

Jeśli chodzi o profilowanie, to zjawisko to po części w Unii Europejskiej regulują przepisy o ochronie danych i o reklamie politycznej. RODO co do zasady zakazuje przetwarzania danych wrażliwych, w tym ujawniających poglądy polityczne, wyznanie czy stan zdrowia (art. 9), i daje każdemu prawo, by nie podlegać decyzjom opartym wyłącznie na zautomatyzowanym profilowaniu (art. 22). W ramach TTPA targetowanie reklamy politycznej online jest dopuszczalne tylko na podstawie danych zebranych od samej osoby, za jej wyraźną i odrębną zgodą, i bez profilowania z użyciem szczególnych kategorii danych, takich jak poglądy polityczne, zdrowie czy wyznanie. Akt o AI zalicza z kolei systemy przeznaczone do wpływania na wynik wyborów lub zachowania wyborcze do systemów wysokiego ryzyka (załącznik III pkt 8 lit. b). Obowiązki z tego tytułu zaczną jednak obowiązywać dopiero 2 grudnia 2027 r., czyli po wyborach parlamentarnych w Polsce. Z kategorii wyłączono ponadto narzędzia do organizacji i optymalizacji kampanii, z których wynikami wyborcy nie stykają się bezpośrednio, co może obejmować również część narzędzi profilujących.

Istotna w hiperpersonalizacji jest też faza **produkcji**, czyli po prostu tworzenia treści, tekstu, grafiki, nagrania czy spotu. W tradycyjnej kampanii nadawcą jest komitet wyborczy, a treść powstaje w niewielu wersjach; w hiperspersonalizowanej kampanii treść tworzony się w bardzo dużej liczbie wersji tak, by odpowiadały na potrzeby każdej z wielu licznych grup wyborców. Wyspecjalizowane narzędzia, np. BattlegroundAI czy Quiller, generują dla sztabów warianty reklam, postów i e-maili. BattlegroundAI, uruchomiona w 2024 r. z myślą o kampaniach lokalnych i stanowych, oferuje nielimitowane generowanie treści za 19 dol. miesięcznie, a jej założycielka jako typową potrzebę sztabu wskazuje „szybkie przygotowanie stu reklam w wyszukiwarce"; firmy konsultingowe tworzą już dla klientów automatycznie setki wariantów jednej reklamy[3]. Konsekwencje są dwie. Po pierwsze, odbiorcy są zalewani komunikatami. Po drugie, takiej kampanii nie da się rzetelnie zbadać ani opisać: nie istnieje rejestr wygenerowanych wariantów, sztaby nie mają interesu w ujawnianiu, których użyły, a nad całością nie ma żadnego nadzoru, ciężko więc ocenić rzeczywistą skalę zjawiska.

Dla tej fazy najistotniejszy jest akt o AI, a obecnie w szczególności jego art. 50, który stosuje się od 2 sierpnia 2026 r. Od tego czasu treści stworzone przy pomocy AI muszą być jako takie oznaczone. Przepis nie obejmuje też tekstów, które przeszły weryfikację człowieka i za które ktoś ponosi odpowiedzialność redakcyjną, co może dotyczyć dużej części tekstowych materiałów sztabowych.

Dystrybucja to za to sposób, w jaki treść trafia do odbiorcy. Może to być płatna reklama emitowana przez platformę albo publikacja „organiczna", czyli post udostępniony przez zwykłego użytkownika. To rozróżnienie jest tym istotniejsze, ponieważ prawo reguluje przede wszystkim reklamę płatną, a kampanie coraz częściej przenoszą się do dystrybucji organicznej. Co ważne, systemy reklamowe od dawna potrafią uzależniać emisję od kontekstu: pogody, pory dnia, lokalizacji czy urządzenia. W 2017 r. wyciekła prezentacja Facebooka dla australijskich reklamodawców, w której firma zapewniała, że potrafi wskazać momenty, gdy nastolatkowie czują się „bezwartościowi", „niepewni" lub „zestresowani" i „potrzebują zastrzyku pewności siebie"; Facebook przeprosił, ale zaprzeczył,

2 Resonate, *Resonate and Yahoo DSP Take Audience Targeting to the Next Level with the Launch of Syndicated Segments*, komunikat prasowy, 26.07.2024, https://www.resonate.com/newsroom/resonate-and-yahoo-dsp-take-audience-targeting-to-the-next-level-with-the-launch-of-syndicated-segments/; Resonate Launches rAI for Consumer Data at Scale, „Daily Research News Online", 5.06.2023, https://www.mrweb.com/drno/news35128.htm; Resonate, Methodology, https://www.resonate.com/reports/resonate-methodology/

3 M. Swant, *AI Briefing: How political startups are helping small political campaigns scale content and ads with AI*, „Digiday", 26.07.2024, https://digiday.com/media/ai-briefing-how-political-startups-are-helping-small-political-campaigns-scale-content-and-ads-with-ai/; N.E. Sanders, B. Schneier, AI Is Changing How Politics Is Practiced in America, „The American Prospect", 10.10.2025, https://prospect.org/2025/10/10/ai-artificial-intelligence-campaigns-midterms/

by kierował reklamy na podstawie stanu emocjonalnego[4]. W kampanii wyborczej ten sam mechanizm oznacza, że komunikat o bezpieczeństwie może trafić do odbiorcy późnym wieczorem, po lekturze alarmujących wiadomości, a obietnica obniżki cen energii – w dniu, w którym przyszedł rachunek. Odbiorca nie będzie wiedział, że czytany przez niego materiał danej partii widzi dlatego, że został wybrany właśnie ze względu na stan, w jakim się znajduje.

Kwestii tej dotyka po części akt o AI, który zakazuje stosowania technik celowo manipulacyjnych lub wprowadzających w błąd oraz wykorzystywania słabości osób wynikających z ich wieku, niepełnosprawności lub szczególnej sytuacji społecznej czy ekonomicznej (art. 5 ust. 1 lit. a i b) – ale tylko wtedy, gdy techniki te powodują lub mogą powodować poważną szkodę. Najszerzej na fazę dystrybucji odpowiada jednak akt o usługach cyfrowych (DSA), który jako jedyny pozwala spojrzeć na kampanię całościowo. Bardzo duże platformy muszą oceniać i ograniczać ryzyko systemowe dla procesów wyborczych (art. 34–35) oraz prowadzić repozytoria reklam (art. 39). Żadna platforma internetowa nie może wyświetlać reklam opartych na profilowaniu z użyciem danych wrażliwych (art. 26 ust. 3). Komisja wydała w 2024 r. wytyczne dotyczące ograniczania ryzyka dla wyborów. Skuteczność DSA zależy jednak od sprawności Komisji i krajowych koordynatorów; co więcej, sam Akt nie podnosi bezpośrednio tematu wykorzystywania AI do hiperpersonalizacji treści w trakcie kampanii wyborczej.

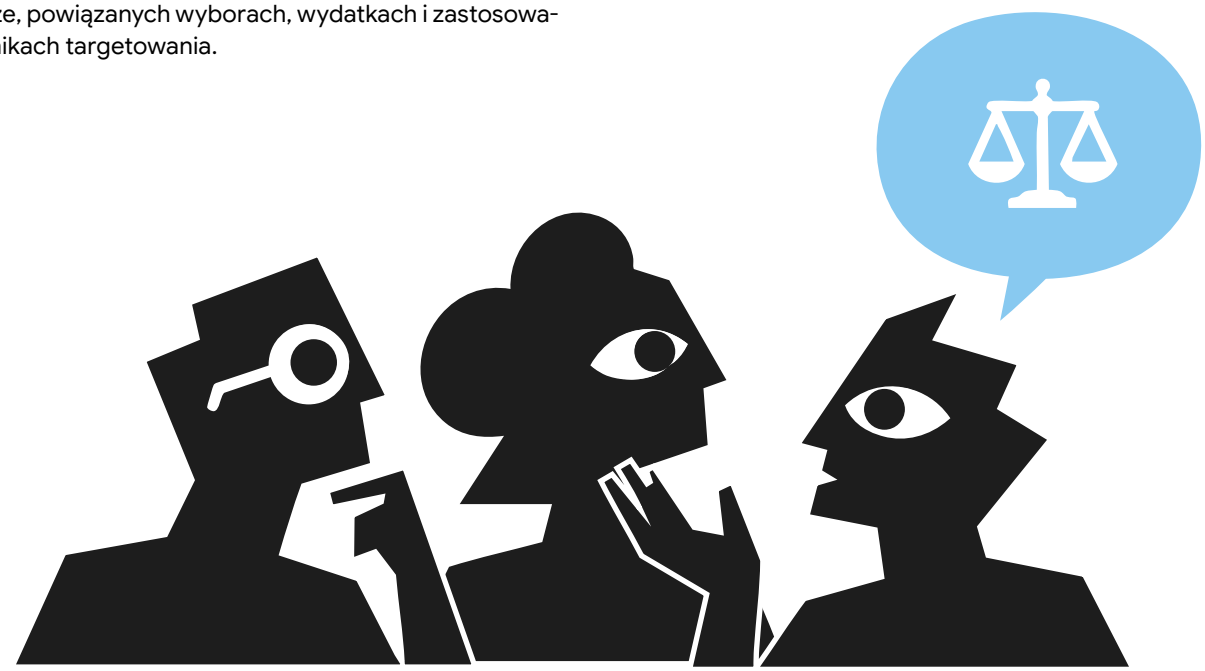
Oddzielną odpowiedzią jest też Rozporządzenie (UE) 2024/900 w sprawie przejrzystości i targetowania reklamy politycznej (TTPA), które pomimo zawetowania w Polsce implementującej ustawy, obowiązuje w całej Unii Europejskiej stosuje się w pełni od 10 października 2025 r. Wymaga ono, by każda reklama polityczna była oznaczona i opatrzona informacją o sponsorze, powiązanych wyborach, wydatkach i zastosowanych technikach targetowania.

Rekomendacje

1. Definicja materiału wyborczego powinna **objąć treści generowane przez AI**. Kodeks wyborczy nie zna pojęć reklamy targetowanej, treści generowanej przez AI ani skoordynowanej nieautentycznej aktywności w sieci. Wymaga jedynie, by materiały wyborcze zawierały wyraźne oznaczenie komitetu, od którego pochodzą. To istotne, by prawo wyborcze w krajach Unii Europejskiej, w tym oczywiście w Polsce, uwzględniało takie koncepty. Równolegle konieczne jest szybkie uchwalenie ustawy wykonującej TTPA.

---

2. Czas na zwiększenie zaangażowania Polski w tworzenie prawa unijnego. **Wszystkie wspomniane regulacje wdrażają prawo UE**, które - choć po części odpowiada na wyzwania związane z wykorzystaniem AI w kampanii wyborczej - nie odpowiada bezpośrednio na kwestie mikrotargetowania i hiperpersonalizacji. Polska powinna więc aktywnie uczestniczyć w tworzeniu kolejnych aktów unijnych, między innymi w pracach nad wytycznymi i kodeksem praktyk do art. 50 aktu o AI, nad wytycznymi EROD do art. 18–19 TTPA oraz w przeglądzie samego rozporządzenia, tak, by standardy odpowiadały realiom polskich kampanii, w których coraz większa część przekazu rozchodzi się organicznie, poza płatną reklamą.



4 J. Condliffe, *Is Facebook Targeting Ads at Sad Teens?*, „MIT Technology Review", 1.05.2017, https://www.technologyreview.com/2017/05/01/105987/is-facebook-targeting-ads-at-sad-teens/

Influencerzy i polityka: nowa szara strefa

07

Dobromił Wereszczyński

Analityk Cyfrowego Steru, zajmuje się wykorzystaniem danych z mediów społecznościowych w tropieniu i badaniu ingerencji w procesy demokratyczne oraz kwestiami dotyczącymi reklamy politycznej i polityki publicznej. Interesuje się tematyką wpływu narracji strategicznych na politykę międzynarodową, szczególnie w obszarze obronności i bezpieczeństwa. Ukończył studia na London School of Economics, studiuje Bezpieczeństwo Informacyjne i Cyberbezpieczeństwo na Akademii Sztuki Wojennej.

Polityków i influencerów łączy nie tylko to, że reprezentują zawody cieszące się najniższym poziomem zaufania społecznego. Przedstawiciele obu tych grup tworzą spektakl, zabiegają o wizerunek autentycznych reprezentantów bliskich swoim odbiorcom i zarabiają na przyciąganiu uwagi. Przeniesienie naszej uwagi do sieci sprawiło jednak, że obie te role coraz częściej się przenikają, a granica między nimi staje się trudna do uchwycenia.

43% Polaków obserwuje influencerów, a wśród osób w wieku 18-24 lat odsetek ten wzrasta do 72%[1]. Przełomem dla ich obecności w domenie politycznej były wybory prezydenckie z 2025 roku, w których aż 5 z 12 kandydatów można było zakwalifikować jako twórców internetowych. Goście kanału Sławomira Mentzena przed drugą turą, które zgromadziły ponad 3 mln widzów. To na YouTube obaj kandydaci omawiali i podpisywali Deklarację Toruńską, a o tych dyskusjach rozmawiano w mediach tradycyjnych. Sam Mentzen mówi, że zarobił na tych programach 350 tysięcy złotych i już od dawna więcej z działalności internetowej niż parlamentarnej[2].

Unikalnym natomiast przykładem stał się Krzysztof Stanowski i jego Kanał Zero, który w ciągu dwunastu miesięcy osiągnął 26,75 mln zł przychodu. W szczycie kampanii generował ponad 175 mln wyświetleń miesięcznie, łącząc rolę kandydata z rolą głównego prowadzącego debaty. Budzi to pytania: jak traktować bezpośrednie zarabianie na kampanii wyborczej? Jakie reguły powinny obowiązywać, gdy influencer gości u siebie polityka? Jak odbiorca ma wiedzieć, co jest autentyczną opinią, a co płatną współpracą?

Coraz więcej zależności

Na całym świecie obserwujemy już nowego typu polityków, którzy budują swoją pozycję poza tradycyjnymi strukturami partyjnymi i niezależnie od publicznych subwencji, opierając swój byt na produkcji treści cyfrowych. Taki “influ-polityk” nie musi prowadzić kampanii na konwencjach partyjnych i nie udziela wywiadów w tradycyjnych studiach telewizyjnych.Zamiast tego prowadzi wielogodzinne transmisje internetowe.

Prawdziwą skalę problemu widać jednak w jego wyciągu bankowym, gdzie pensja i dieta poselska okazują się zaledwie skromnymi dodatkami do przychodów generowanych z mechanizmów monetyzacyjnych platformy – opłat za wyróżnione wiadomości, prywatnych darowizn, udziału w zyskach z reklam, czy nawet współprac sponsorskich. Głównym źródłem utrzymania i politycznego bytu takiego posła nie jest komitet wyborczy, lecz algorytm globalnej korporacji technologicznej oraz stała gotowość do monetyzowania społecznych emocji. Jak będzie on głosować nad przepisami wpływającymi na strefę internetową, skoro żyje z dobrej ekspozycji przez platformy?

Mieszanie produkcji treści internetowych z polityką przybiera na skali na całym świecie, a jego skrajne konsekwencje widać na przykład w Korei Południowej. W kraju tym już przynajmniej 25 parlamentarzystów zdobyło prestiżowe „srebrne przyciski” YouTube za przekroczenie 100 tysięcy subskrypcji[3]. Podczas kryzysu wokół stanu wojennego pod koniec 2024 roku tamtejsze kanały polityczne zarabiały setki milionów wonów na transmisjach na żywo ze starć i protestów[4]. Nakręcał to niebezpieczny system zachęt: im bardziej polaryzujący, agresywny i nienawistny jest przekaz, tym większe zyski z wpłat widzów trafiają na konta twórców internetowych.

1 Polacy w social mediach, Social Media IAB Polska, styczeń 2026r. https://www.iab.org.pl/wp-content/uploads/2026/01/Raport-Polacy-w-social-mediach_IAB_Polska_2026.pdf

2 Sławomir Mentzen ujawnia. Tyle zarobił na rozmowach z Karolem Nawrockim i Rafałem Trzaskowskim, Business Insider, 12.01.2026 r. https://businessinsider.com.pl/wiadomosci/slawomir-mentzen-ujawnia-zarobki-z-wywiadow-z-kandydatami-na-prezydenta/v8e4qg1

3 Rise of YouTube politicians tests S. Korea’s democratic resilience, The Korea Herald, 14.02.2026 r. https://www.koreaherald.com/article/10675902

4 Korea’s political YouTubers earn massive profit out of impeachment crises, The Korea Times, 23.03, 2025 r. https://www.koreatimes.co.kr/southkorea/politics/20250323/koreas-political-youtubers-earn-massive-profit-out-of-impeachment-crises

W Stanach Zjednoczonych profesjonalizacja tego sektora osiągnęła już skalę przemysłową. W cyklu wyborczym 2024 roku budżety największych komitetów politycznych przeznaczone na komunikację cyfrową i sieci twórców liczono w dziesiątkach milionów dolarów, a same konwencje Partii Demokratycznej i Republikańskiej oficjalnie akredytowały setki influencerów[5]. Zjawisko to zeszło już także na poziom lokalny i samorządowy. Administracja burmistrza Nowego Jorku Zohrana Mamdaniego zbudowała sieć powiązań z co najmniej 200 twórcami. Otrzymywali oni codziennie gotowe przekazy dnia, materiały wideo, zaproszenia na wydarzenia i bezpośredni dostęp do urzędników[6]. W 2025 roku w Nepalu, zaangażowanie tamtejszych influencerów oraz młodzieży w masowe protesty – których iskrą, obok sprzeciwu wobec korupcji, była próba zablokowania platform – doprowadziło we wrześniu 2025 roku do upadku rządu premiera Oliego[7]. Na Węgrzech przez lata prorządowe centrum szkoleniowe i producenckie Megafon pompowało miliony euro w rozwój prawicowych influencerów w zamian za przychylność wobec Viktora Orbána[8]. Gdy zaś Google i Meta wprowadziły ograniczenia dla oficjalnych reklam politycznych, środki finansowe nie zniknęły z systemu, lecz przepłynęły do niejawnych sieci wsparcia i profili twórców masowo obchodzących regulaminy platform. stanowi przestrogę przed prostym zakazywaniem reklamy politycznej bez jednoczesnego nadzoru nad ekonomią twórców.

Najpoważniejsze ostrzeżenie płynęło jednak z Rumunii, gdzie w pierwszej turze wyborów prezydenckich w listopadzie 2024 roku niespodziewanie zwyciężył Călin Georgescu. Odtajnione dokumenty rumuńskich służb oraz analizy śledcze ujawniły, że za jego sukcesem stała skoordynowana operacja cyfrowa wykorzystująca ponad 25 tysięcy kont na TikToku oraz giełdy mikro-influencerów, finansowane z nieujawnionych, w tym zagranicznych źródeł[9]. Choć kandydat oficjalnie zadeklarował zerowy budżet kampanijny, skala ukrytego finansowania i manipulacji algorytmicznej doprowadziła do bezprecedensowego unieważnienia wyborów przez rumuński Sąd Konstytucyjny i wywołała głęboki kryzys zaufania do instytucji państwa.

Jak zarabia się na politycznych zasięgach

Aby skutecznie uregulować ten rynek, należy rozróżnić poziomy, na których w ekosystemie twórców przepływają pieniądze. Granica między dzieleniem się opiniami a profesjonalną działalnością medialno-polityczną bywa płynna, a same platformy celowo bagatelizują ten problem. Przejrzystości wymaga każdy z trzech głównych sposobów monetyzacji:

Przychody z systemów reklamowych platform. Podstawowym mechanizmem jest bezpośrednie przeliczanie czasu uwagi widza na pieniądze. Na polskim rynku YouTube realny przychód twórcy wynosi **ok. 4,5 zł (ok. 1,15 USD) za tysiąc wyświetleń**[10]. Ponieważ wysokość zarobków zależy bezpośrednio od czasu spędzonego przed ekranem, algorytmy rekomendacji naturalnie premiują treści sensacyjne, konfrontacyjne i polaryzujące. Zjawisko uzależniającego projektowania interfejsów i monetyzacji uwagi ma zostać zaadresowane na poziomie unijnym w przygotowywanym Akcie o sprawiedliwości cyfrowej (Digital Fairness Act)[11].

Współprace i barter polityczny. Tak jak twórca ma obowiązek oznaczyć płatną promocję kosmetyku czy suplementu diety, tak samo powinien mieć bezwzględny obowiązek informowania o sponsorowanym charakterze przekazu politycznego lub społecznego. Współpraca polityczna często przybiera formę trudnego do wyceny barteru poza samą platformą: influencer otrzymuje od partii, fundacji lub organizatorów wydarzeń politycznych zlecenia okołomedialne, opłacone podróże, ekskluzywny dostęp do polityków na zapleczu czy sztuczne pompowanie widoczności przez partyjne konta. Tego rodzaju korzyści powinny być jawnie oznaczane przez twórców i wykazywane w rejestrach komitetów wyborczych oraz partii pod rygorem dotkliwych kar finansowych.

Finansowanie społecznościowe i wpłaty bezpośrednie. Trzecim filarem są systemy bezpośredniego wsparcia od widzów – od wbudowanych narzędzi platform po zewnętrzne bramki płatnicze i serwisy subskrypcyjne. Jak pokazują badania nad TikTok LIVE, transmisje na żywo o charakterze politycznym są wykorzystywane do zarabiania na wirtualnych prezentach i windowania zasięgów. W ten sposób konta służą do transferowania trudnych do wyśledzenia środków finansowych bezpośrednio do wybranych twórców[12]. Brakuje dziś jasnych reguł określających, gdzie i w jaki sposób twórcy angażujący się w kampanię powinni raportować tego typu przepływy oraz jak weryfikować, czy za serią drobnych wpłat nie stoi skoordynowana operacja finansowa.

Polska w regulacyjnej szarej strefie

Na tym tle Polska pozostaje niebezpiecznie narażona, mimo wysokiej cyfryzacji polityki. Polscy parlamentarzyści publikują rocznie setki tysięcy wpisów w mediach społecznościowych,

5 Influencers are playing a big role in this year's election. There's no way to tell who's getting paid for their endorsements, CNN, 29.02.2024 r. https://edition.cnn.com/2024/10/29/tech/influencers-presidential-campaign-paid-disclosure/

6 Mayor Mamdani's Friendly Influencers and the Funding of Public Messaging in New York City, Columbia Journalism Review, 21.08.2026 r. https://www.cjr.org/tow_center/mayor-mamdanis-friendly-influencers-and-the-funding-of--public-messaging-in-new-york-city.php

7 Nepal Prime Minister Oli resigns amid violent protests over social media ban, corruption, TRT World, 9.09, 2025 r. https://www.trtworld.com/article/fa4525ce6186

8 How Hungary's richest man funneled millions into Orbán's propaganda machine, International Consortium of Independent Journalism, 7.10.2026 r. https://www.icij.org/news/2026/10/how-hungarys-richest-man-funneled-millions--into-orbans-propaganda-machine/

9 Romanian court cancels presidential election amid Russian influence fears, Politico, 6.12.2024 r. https://www.politico.eu/article/romania-court-cancels--presidential-election-runoff-tiktok-russian-influence-calin-georgescu/

10 YouTube AdSense RPM, Dynamoi, dostęp 18.10.2026 r. dynamoi.com/data/youtube-adsense-rpm

11 Towards a digital fairness act, European Parliamentary Research Service, 10.2026 r. https://www.europarl.europa.eu/RegData/etudes/BRIE/2026/791529/EPRS_BRI(2026)791529_EN.pdf

12 The Hidden Power of TikTok LIVE, Obserwatorium Demokracji Cyfrowej, marzec 2026 https://ceeddw.org/report-the-hidden-power-of-tiktok-live-now-available/

a rekordziści zamieszczają nawet do 3 tysięcy postów miesięcznie. Równolegle twórcy internetowi coraz chętniej zapraszają polityków do swoich programów, bo obecność wyrazistych postaci niemal gwarantuje wyświetlenia i wysokie zyski reklamowe.

Mimo tych zagrożeń polskie instytucje unikają wzięcia odpowiedzialności za ten obszar. Urząd Ochrony Konkurencji i Konsumentów, który wypracował skuteczne narzędzia dyscyplinowania twórców w przypadku ukrytych reklam komercyjnych, wykazuje niechęć wobec nadzoru nad polityczną działalnością influencerów, uznając, że jego ustawowe kompetencje kończą się na ochronie konsumenta. W efekcie kluczowa przestrzeń kształtowania opinii publicznej i decyzji wyborczych funkcjonuje w regulacyjnej próżni.

Ramy regulacyjne

Na tym tle Polska pozostaje niebezpiecznie narażona, mimo Polska nie musi budować tego systemu od zera. Dysponujemy już wytycznymi Prezesa UOKiK dotyczącymi oznaczania treści komercyjnych, unijnym Aktem o usługach cyfrowych (DSA) oraz Rozporządzeniem o przejrzystości i targetowaniu reklamy politycznej (TTPA), które, choć wciąż czeka na polską ustawę wykonawczą określającą krajowe sankcje, wyznacza standardy jawności. Problem dostrzega również Krajowa Rada Radiofonii i Telewizji (KRRiT), która wydała specjalny informator dla twórców internetowych, zorganizowała poświęconą temu konferencję i zachęca influencerów – jako dostawców audiowizualnych usług medialnych na żądanie (VOD) – do rejestrowania swojej działalności w oficjalnym wykazie[13].

Skutecznych wzorców dostarczają także inne państwa Unii Europejskiej, m.in. Dania, która z powodzeniem egzekwuje jawność płatnych wpisów politycznych na gruncie przepisów o ochronie konsumentów. Kluczowe jest przy tym takie skalibrowanie przepisów, aby nie krępować spontanicznej, organicznej wypowiedzi obywatelskiej, lecz nałożyć największy ciężar obowiązków sprawozdawczych na profesjonalnych graczy: partie polityczne, komitety wyborcze, agencje marketingowe i wielkie platformy cyfrowe.

Szansą na ujednolicenie tych zasad w skali całej Unii jest nadchodząca rewizja dyrektywy o audiowizualnych usługach medialnych. Europejska Rada ds. Usług Medialnych w swoich założeniach do nowelizacji dyrektywy z lipca 2026 roku wprost wskazuje, że brak zharmonizowanych ram dla twórców internetowych rodzi niepewność prawną i utrudnia transgraniczne egzekwowanie przepisów. Rada rekomenduje bezpośrednie włączenie twórców w zakres dyrektywy oraz wypracowanie wspólnych unijnych standardów opartych na odpowiedzialności redakcyjnej, przejrzystości i ochronie odbiorców[14].

13 Odpowiedzialność, transparentność, wiarygodność - nowy informator KRRiT dla influencerów, KRRiT, 26.02.2026 r. https://www.gov.pl/web/krrit/odpowiedzialnosc-transparentnosc-wiarygodnosc--nowy-informator-krrit-dla-influencerow

14 The Media Board presents its suggestions for a revision of the AVMSD, Europejska Rada ds. Usług Medialnych, 30.07.2026 r. https://media-board.europa.eu/news-0/media-board-presents-its-suggestions-revision-avmsd-2026-07-30_en

Rekomendacje

1. Pełna przejrzystość monetyzacji w okresie wyborczym. Należy nałożyć na platformy cyfrowe, sieci partnerskie i agencje marketingowe **obowiązek bieżącego raportowania** w trakcie kampanii wszelkich przepływów finansowych trafiających do twórców zaangażowanych w agitację lub debaty z kandydatami, od przychodów z systemów typu AdSense po darowizny cyfrowe, oraz zintegrować te dane z europejskim repozytorium reklamy politycznej.

2. Jednoznaczne oznaczanie materiałów politycznych pod rygorem utraty monetyzacji. Każdy materiał o charakterze politycznym, który powstał w wyniku płatnej współpracy, barteru lub wsparcia organizacyjnego komitetu czy partii, musi posiadać **czytelne oznaczenie wizualne** już na samym początku nagrania. Naruszenie tego obowiązku powinno skutkować natychmiastowym wyłączeniem monetyzacji danego materiału oraz karami finansowymi dla zlecającego komitetu.

3. Oddanie obywatelom realnej kontroli nad algorytmami Platformy powinny zostać zobowiązane do wprowadzenia do**myślnego trybu obniżonej amplifikacji** skrajnie polaryzujących i toksycznych treści społeczno-politycznych, a także do zapewnienia użytkownikom możliwości całkowitego wyłączenia spersonalizowanych rekomendacji.

4. Bezwzględna **demonetyzacja dezinformacji** i uszczelnienie finansowania. Konieczne jest odcięcie od zysków reklamowych podmiotów systematycznie i celowo wprowadzających odbiorców w błąd. Jednocześnie należy zlikwidować luki w polskim prawie wyborczym, które pozwalają dziś na ukryte finansowanie przychylnych twórcom ekosystemów medialnych przez Spółki Skarbu Państwa oraz fundacje.