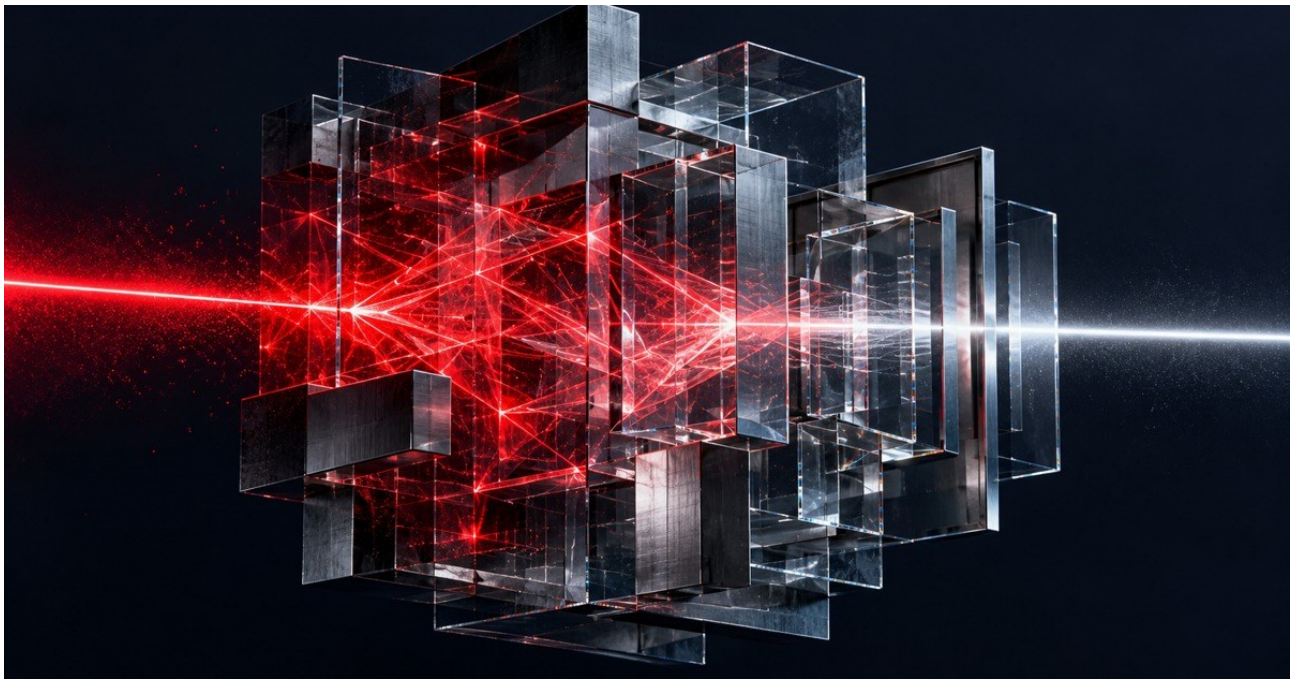
DISINFORMATION & DIGITAL RESILIENCE / PL

Architektura oszustwa: systemowa analiza nieładu informacyjnego w erze cyfrowej

Jak platformy, ekonomia uwagi i podatności poznawcze tworzą ekosystem dezinformacji. Analiza Ivana-Oleksandra Yakymiva o historii manipulacji, generatywnej AI i budowaniu odporności cyfrowej.

Ivan-Oleksandr Yakymiv | 28.10.2026



Analiza autorska · wersja opracowana redakcyjnie na podstawie tekstu autora z 28 grudnia 2025 r. Zaktualizowana 28 października 2026 r.; doprecyzowano zakres badań i regulacji. Ilustracje redakcyjne przedstawiają metafory omawianych zjawisk; nie są dokumentacją rzeczywistych zdarzeń.

Ontologiczne i epistemologiczne fundamenty współczesnej dezinformacji

Ilustracja redakcyjna TRUTH · Ontologiczne i epistemologiczne fundamenty współczesnej dezinformacji

Współczesna kondycja sfery publicznej w dobie transformacji cyfrowej nie jest jedynie wynikiem postępu technologicznego, lecz stanowi głęboki kryzys prawdy o charakterze systemowym. Analiza zjawiska, które w literaturze przedmiotu określa się mianem architektury oszustwa, wymaga wyjścia poza potoczne rozumienie kłamstwa i przyjęcia perspektywy systemów złożonych. Dezinformacja w XXI wieku nie jest zbiorem odizolowanych incydentów, lecz emergentną właściwością wzajemnie powiązanych systemów społecznych i technicznych, współtworzącą swoisty ekosystem informacyjny. W tym ujęciu współczesny kryzys prawdy jawi się jako produkt uboczny architektury platform cyfrowych, w których fałsz i zaangażowanie emocjonalne wzajemnie się wzmacniają, tworząc mechanizm określany jako „asamblaż dezinformacyjny" (disinformation assemblage). Jest to perspektywa analityczna, a nie twierdzenie, że każdy system cyfrowy nieuchronnie wytwarza dezinformację. [1]

Zrozumienie tej architektury wymaga interdyscyplinarnego podejścia, integrującego czynniki społeczne, architekturę technologiczną oraz podatności poznawcze ludzkiego umysłu. Czynniki społeczne obejmują tu szerokie spektrum motywacji – od partykularnych interesów politycznych, przez chęć zysku ekonomicznego, aż po indywidualną gratyfikację psychologiczną płynącą z poczucia przynależności do określonej grupy ideologicznej. Architektura technologiczna, oparta na nieprzejrzystych algorytmach i masowej analityce danych, dostarcza z kolei infrastruktury dla precyzyjnego skalowania i targetowania przekazu. Ostatnim elementem są podatności poznawcze, gdzie ewolucyjnie ukształtowane mechanizmy przetwarzania informacji, takie jak wirusowość emocji czy błędy konfirmacji, stają się punktami wejścia dla manipulacji.

Społeczeństwo: motywacje polityczne, interes ekonomiczny i przynależność grupowa tworzą warunki popytu oraz podaży na zmanipulowane przekazy.

Technologia: platformy, analityka danych, boty i AI zapewniają infrastrukturę skalowania oraz dystrybucji.

Psychologia: emocje, selektywna uwaga i błędy poznawcze wpływają na interpretację oraz udostępnianie informacji.

Te trzy warstwy łączą motywacje społeczne, infrastrukturę dystrybucji i mechanizmy poznawcze. Ich współdziałanie może sprzyjać dezinformacji, lecz rezultat zależy od konkretnej platformy, kontekstu i działań odbiorców.

Dezinformacja jako system złożony i asamblaż społeczno-techniczny

Ilustracja redakcyjna TRUTH · Dezinformacja jako system złożony i asamblaż społeczno-techniczny

Analiza systemowa dezinformacji pozwala na zidentyfikowanie jej jako zjawiska wykraczającego poza ramy prostego modelu nadawca-odbiorca. Jest to system społeczno-techniczny, w którym interakcje między ludźmi a algorytmami prowadzą do powstania nowych, nieprzewidywalnych wcześniej form komunikacji. Pojęcie „asamblażu dezinformacyjnego" pozwala uchwycić ten proces jako tymczasową, ale potężną konstelację różnorodnych aktorów: od państwowych biurokracji i służb specjalnych, przez komercyjne firmy analityczne, aż po nieświadomych użytkowników mediów społecznościowych, którzy stają się „nieodpłatnymi pracownikami" kampanii dezinformacyjnych poprzez proste akty udostępniania i polubienia treści.

Kluczową cechą tego systemu jest jego samonapędzający się charakter. Mechanizmy te można opisać za pomocą pętli sprzężenia zwrotnego, gdzie zaangażowanie generowane przez treści

dezinformacyjne dostarcza danych (behawioralnej nadwyżki), które są następnie wykorzystywane do optymalizacji algorytmów rekomendacyjnych. W literaturze naukowej podkreśla się, że współczesne „zanieczyszczenie informacji" zachodzi na skalę globalną, a jego skutki są trudne do precyzyjnego oszacowania, a ustalenie ich wpływu na wyniki wyborów, polaryzację czy kryzysy zdrowia publicznego wymaga badań właściwych dla danego kontekstu. [1]

Algorytmy pełnią funkcję strażników dostępu do informacji: porządkują i rekomendują treści. Mogą wzmacniać selektywną ekspozycję, ale komory echa nie powstają wyłącznie za ich sprawą. Znaczenie mają również wybory użytkowników, sieci kontaktów i kontekst polityczny. Nie należy traktować bańki filtrującej jako uniwersalnego skutku korzystania z internetu.

Zależności między wcześniejszymi przekonaniami, emocjami i przyjęciem informacji można badać ilościowo. Nie ma jednak podstaw, by przedstawiać jeden prosty wzór jako zweryfikowane prawo opisujące wszystkie te procesy. W tej analizie pętla sprzężenia zwrotnego jest modelem pojęciowym; nie przypisujemy jej pozornej precyzji matematycznej.

Trzy oblicza informacyjnego chaosu – Typologia Wardle i Derakhshana

Ilustracja redakcyjna TRUTH · Trzy oblicza informacyjnego chaosu – Typologia Wardle i Derakhshana

Podstawą naukowej analizy nieładu informacyjnego jest precyzyjne rozróżnienie terminologiczne, zaproponowane w fundamentalnym raporcie Claire Wardle i Hosseina Derakhshana z 2017 roku. Autorzy ci odeszli od nieprecyzyjnego terminu „fake news", wprowadzając trójpodział na misinformację, dezinformację i malinformację, oparty na wymiarach prawdziwości i intencjonalności szkody. Misinformacja oznacza fałszywą informację przekazywaną bez intencji wyrządzenia szkody; dezinformacja łączy fałsz z celowym szkodliwym działaniem; malinformacja wykorzystuje prawdziwe informacje do wyrządzenia szkody, np. przez ujawnienie prywatnych danych. [1]

Dezinformacja, będąca głównym przedmiotem analizy, jest procesem złożonym, który należy badać w trzech fazach: kreacji (stworzenie komunikatu), produkcji (zamiana w produkt medialny) oraz dystrybucji (rozpowszechnienie w przestrzeni publicznej). Agent (aktor) tworzący dezinformację może być zupełnie inny niż agent ją produkujący czy dystrybuujący, co utrudnia identyfikację źródeł kampanii wpływowych. Motywacje agentów są zróżnicowane: od finansowych (generowanie przychodów z reklam) przez polityczne (destabilizacja przeciwnika), aż po społeczne (budowanie tożsamości grupowej).

Największym wyzwaniem dla współczesnych demokracji jest fakt, że granice między tymi kategoriami są płynne. Malinformacja może stać się paliwem dla dezinformacji, gdy wyrwane z kontekstu prawdziwe dane są używane do budowania fałszywych narracji. Skuteczność tych działań opiera się na umiejętnym wykorzystaniu emocji, takich jak strach czy oburzenie, co sprzyja ponownemu udostępnianiu treści w ramach cyfrowych „plemion".

Archeologia kłamstwa – Od retoryki antycznej do rewolucji Gutenberga

Ilustracja redakcyjna TRUTH · Archeologia kłamstwa – Od retoryki antycznej do rewolucji Gutenberga

Manipulacja nie zaczęła się od internetu. Nie należy jednak utożsamiać całej tradycji retorycznej z oszustwem: Arystoteles analizował środki perswazji, podkreślając zarazem znaczenie prawdy, sprawiedliwości i argumentacji. Jego Retoryka pomaga zrozumieć oddziaływanie mowy, lecz nie uzasadnia tezy, że prawda jest w niej zbędna. [3]

Rewolucja Gutenberga w XV wieku, będąca momentem demokratyzacji wiedzy, stała się jednocześnie katalizatorem dla masowej dezinformacji. Druk umożliwił tanią produkcję ulotek i pamfletów, które stały się potężną bronią w wojnach religijnych okresu Reformacji. Informacja zyskała nową dynamikę: mogła być rozpowszechniana poza bezpośrednią kontrolą instytucji

centralnych, co zapoczątkowało proces decentralizacji kłamstwa.

Rozwój prasy i rynku czytelniczego stopniowo wzmacniał konkurencję o uwagę odbiorcy. Sensacja mogła przyciągać publiczność, ale historycznych form komunikacji nie należy mechanicznie utożsamiać z dzisiejszymi algorytmami. Ciągłość dotyczy motywacji i perswazji; infrastruktura oraz skala pozostają odmienne.

Zrozumienie tej ewolucji pozwala dostrzec, że dezinformacja nie jest „błędem" nowoczesności, lecz stałym elementem ewolucji systemów komunikacyjnych, który w każdej epoce adaptuje się do dostępnej infrastruktury technologicznej.

Profesjonalizacja i uprzemysłowienie perswazji w XX wieku

Ilustracja redakcyjna TRUTH · Profesjonalizacja i uprzemysłowienie perswazji w XX wieku

Wiek XX przyniósł bezprecedensową profesjonalizację technologii manipulacji, przekształcając propagandę w dziedzinę naukową i administracyjną. Rządy mocarstw po raz pierwszy stworzyły wyspecjalizowane agencje odpowiedzialne za zarządzanie percepcją społeczną, takie jak Komitet Creela w USA czy Ministerstwo Propagandy Josepha Goebbelsa w Niemczech. Nowe media – radio i kino – pozwoliły na jednoczesną synchronizację emocji milionów ludzi, co stało się fundamentem systemów totalitarnych.

Okres Zimnej Wojny to czas globalizacji wojny informacyjnej. Sowieckie służby specjalne rozwinęły koncepcję „aktywnych działań" (aktivnyye meropriyatiya), których celem nie była jedynie zmiana poglądów odbiorcy, lecz przede wszystkim erozja zaufania do instytucji demokratycznych i nauki. Przykładem o ogromnym zasięgu była Operacja INFEKTION, w ramach której rozpowszechniano fałszywą narrację o stworzeniu wirusa HIV/AIDS w laboratoriach wojskowych USA.

Nasycanie przestrzeni informacyjnej wieloma sprzecznymi komunikatami może sprzyjać dezorientacji i zmęczeniu odbiorców. Paul i Matthews opisali współczesny rosyjski model „firehose of falsehood": dużą liczbę kanałów, szybki i powtarzalny przekaz oraz brak przywiązania do spójności. Nie jest to wyczerpujący opis wszystkich kampanii propagandowych. [4]

Rewolucja Web 2.0 – Od struktur hierarchicznych do propagandy obliczeniowej

Ilustracja redakcyjna TRUTH · Rewolucja Web 2.0 – Od struktur hierarchicznych do propagandy obliczeniowej

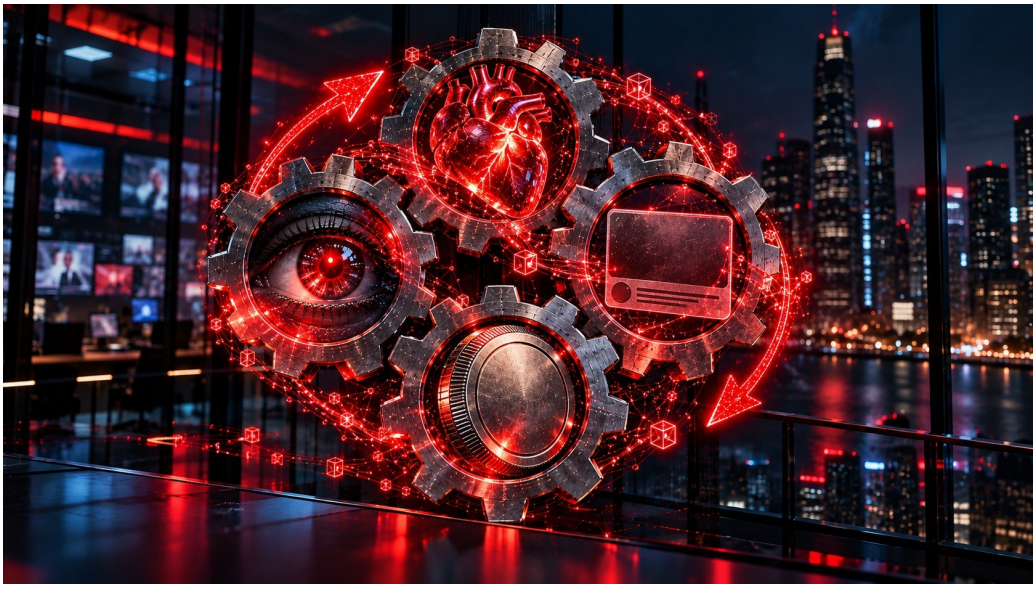
Internet, a w szczególności era Web 2.0, radykalnie zmienił architekturę władzy informacyjnej. Nastąpiło przejście od scentralizowanego, odgórnego modelu nadawczego do zdecentralizowanego, oddolnego i partycypacyjnego fenomenu. W tradycyjnym modelu propaganda wymagała ogromnej infrastruktury (drukarnie, studia filmowe), co ograniczało krąg nadawców do państw i korporacji. Dzisiaj każdy użytkownik, wyposażony w urządzenie mobilne, staje się węzłem w globalnej sieci dystrybucji treści.

Pojawienie się platform społecznościowych doprowadziło do powstania „propagandy obliczeniowej" (computational propaganda). Jest to wykorzystanie algorytmów, botów i skoordynowanych działań w mediach społecznościowych do manipulowania widocznością treści i tworzenia iluzji powszechnego poparcia dla określonych idei. Boty nie tylko rozpowszechniają informacje, ale przede wszystkim „hakują" algorytmy platform, zwiększając szansę, że przekaz trafi do trendów i zostanie odebrany jako popularny. [5]

W tym nowym paradygmacie odpowiedzialność za rozprzestrzenianie fałszu ulega rozmyciu. Użytkownicy, motywowani mechanizmami gratyfikacji (lajki, udostępnienia), nieświadomie wzmacniają kampanie wpływowe, stając się częścią asamblażu dezinformacyjnego. To partycypacyjne oblicze dezinformacji sprawia, że jest ona znacznie trudniejsza do zwalczenia niż

tradycyjna propaganda państwowa.

Ekonomia uwagi – Prawda jako produkt uboczny zaangażowania



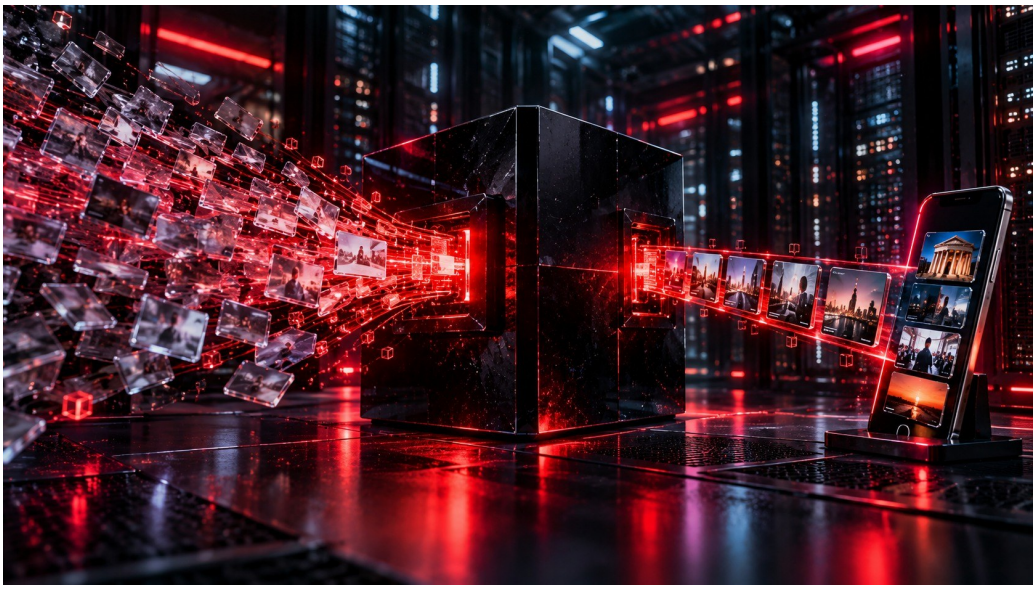
Ilustracja redakcyjna TRUTH · Ekonomia uwagi – Prawda jako produkt uboczny zaangażowania

Współczesna architektura oszustwa jest nierozerwalnie związana z modelem biznesowym wielkich platform technologicznych, opartym na ekonomii uwagi. W tym systemie informacja przestała być dobrem publicznym, a stała się towarem, którego wartość komercyjna bywa wiązana z poziomem wygenerowanego zaangażowania (engagement). Nie oznacza to, że platformy w ogóle nie uwzględniają jakości treści. Kluczowym zasobem biznesowym pozostaje jednak czas spędzony przez użytkownika na platformie, który przekłada się na zysk z targetowanych reklam.

Systemy rekomendacyjne korzystają z wielu sygnałów, w tym zaangażowania. W określonych warunkach mogą premiować treści silnie emocjonalne, co stwarza okazję do manipulacji. Nie jest to dowód, że każdy algorytm zawsze promuje fałsz. Istotne są cele optymalizacji, zabezpieczenia i mierzalne skutki działania konkretnego systemu.

W interpretacji systemowej dezinformacja może być produktem ubocznym bodźców ekonomicznych nagradzających aktywność bez odpowiedniej kontroli jakości informacji. To argument za audytem modeli dystrybucji, a nie za zakładaniem identycznego działania wszystkich platform.

Kapitalizm inwigilacji i nieprzejrzystość algorytmicznych „czarnych skrzynek”



Ilustracja redakcyjna TRUTH · Kapitalizm inwigilacji i nieprzejrzystość algorytmicznych „czarnych skrzynek”

Głębszą warstwę technologiczną dezinformacji opisuje Shoshana Zuboff w teorii kapitalizmu inwigilacji. System ten polega na jednostronnym zawłaszczaniu ludzkiego doświadczenia jako surowca do produkcji danych behawioralnych. Kluczowym pojęciem jest „behawioralna nadwyżka” – dane zbierane o użytkownikach bez ich wiedzy, które służą przewidywaniu zachowań i próbom wpływania na nie. Jest to przywołanie ramy teoretycznej Shoshany Zuboff, nie dowód nieograniczonej skuteczności profilowania. [11]

Profilowanie może ułatwiać kierowanie perswazyjnych komunikatów do określonych grup odbiorców. Nie oznacza to możliwości dowolnego sterowania jednostką: skuteczność mikrotargetowania zależy od kontekstu i wymaga niezależnej oceny.

Kolejnym filarem tej architektury jest algorytmiczna nieprzejrzystość, czyli fenomen „czarnej skrzynki” (Black Box). Logika działania systemów rekomendacyjnych, które decydują o tym, co widzimy w naszym kanale wiadomości, jest ukryta za barierą własności intelektualnej i tajemnicy handlowej. Utrudnia to społeczną i naukową kontrolę nad sposobem, w jaki algorytmy filtrują rzeczywistość i czy systemowo nie promują one fałszu kosztem prawdy.

Nieprzejrzystość utrudnia odbiorcy ocenę, dlaczego zobaczył dany materiał. Potrzebne są zrozumiałe wyjaśnienia rekomendacji, dostęp badawczy oraz badanie rzeczywistych skutków, bez zakładania z góry pełnej izolacji użytkowników.

Psychologia poznawcza dezinformacji – Rola człowieka jako głównego wzmacniacza

Ilustracja redakcyjna TRUTH · Psychologia poznawcza dezinformacji – Rola człowieka jako głównego wzmacniacza

Vosoughi, Roy i Aral przeanalizowali około 126 tys. zweryfikowanych historii rozpowszechnianych na Twitterze w latach 2006–2017. W badanym zbiorze fałszywe wiadomości docierały dalej i rozchodziły się szybciej niż prawdziwe. Wynik dotyczy określonego zbioru i platformy; nie jest uniwersalną miarą dla całego internetu ani współczesnego X. [2]

Autorzy omówili m.in. dwa możliwe mechanizmy związane z obserwowanym wzorcem:

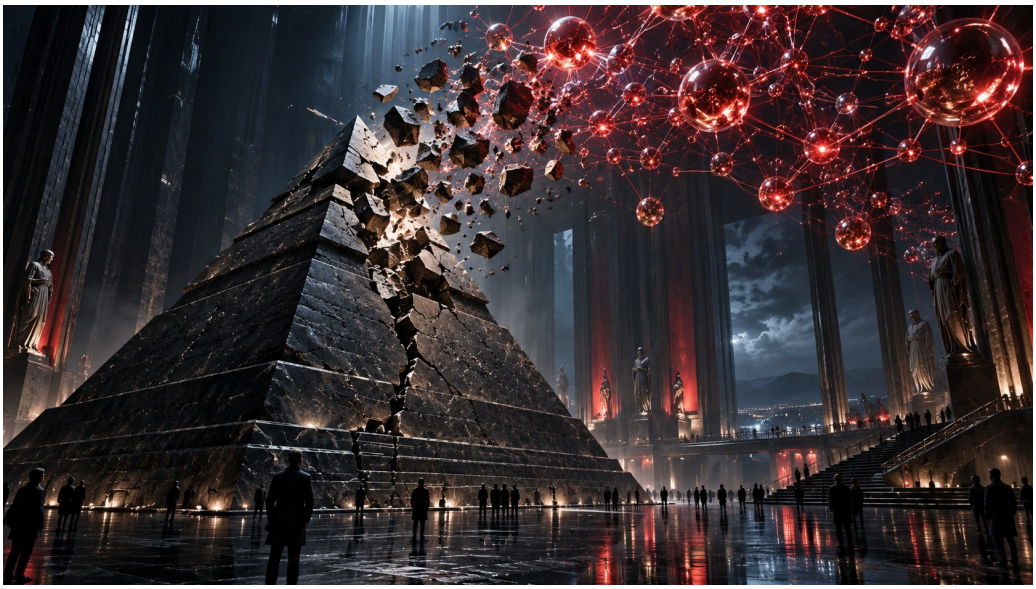
Nowość: fałszywe historie były bardziej nowatorskie w zastosowanej analizie. To wspiera hipotezę, że zaskoczenie może zachęcać do udostępniania, lecz samo w sobie nie dowodzi pojedynczej przyczyny zachowania. [2]

Emocje: w odpowiedziach na fałszywe historie częściej pojawiały się strach, obrzydzenie i zaskoczenie. Korelacje emocjonalne nie wystarczają do stwierdzenia, że każda fałszywa wiadomość jest bardziej wirusowa od prawdziwej. [2]

W analizie wpływ botów nie wyjaśniał różnicy między dyfuzją prawdy a fałszu: przyspieszały oba rodzaje przekazu podobnie. Nie znaczy to, że boty nie są groźne albo że taki sam wynik musi wystąpić w każdej kampanii. [2]

Kolejnym mechanizmem psychologicznym jest błąd potwierdzenia (confirmation bias) – tendencja do przyjmowania informacji zgodnych z naszymi poglądami i odrzucania faktów, które im przeczą. W sytuacjach niepewności informacyjnej dążymy do spójności narracyjnej, a nie do obiektywnego sprawdzenia faktów. Prowadzi to do powstania baniek epistemicznych, w których fenomenologiczne poczucie posiadania wiedzy całkowicie zastępuje rzeczywistą weryfikację.

Upadek hierarchii epistemicznych i erozja zaufania do instytucji wiedzy



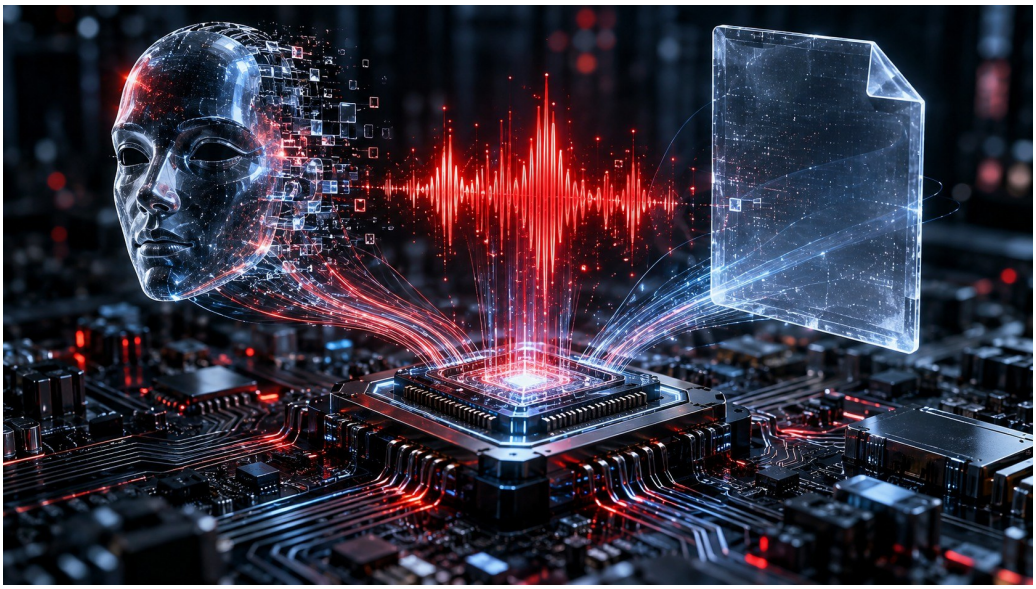
Ilustracja redakcyjna TRUTH · Upadek hierarchii epistemicznych i erozja zaufania do instytucji wiedzy

Dezinformacja nie tylko wprowadza w błąd, ale systemowo podważa fundamenty, na których opiera się społeczne uznawanie prawdy. Następuje przejście od „hierarchii epistemicznej", opartej na autorytecie wiedzy, procedurach recenzowania naukowego i etyce zawodowej, do „hierarchii uwagi", gdzie o znaczeniu informacji decydują algorytmicznie mierzalne wskaźniki popularności.

Instytucje tradycyjnie odpowiedzialne za tworzenie i weryfikację faktów – takie jak uniwersytety, redakcje prasowe czy agencje naukowe – są delegitymizowane poprzez przedstawianie ich jako stronniczych aktorów politycznych. W efekcie dochodzi do erozji wspólnej płaszczyzny epistemicznej, czyli zbioru faktów akceptowanych przez ogół społeczeństwa jako podstawa do debaty.

Skutkiem tego procesu jest paraliż decyzyjny w sytuacjach kryzysowych (np. pandemie, zmiany klimatu), gdzie zaufanie do ekspertów zostaje zastąpione lojalnością grupową i wiarą w teorie spiskowe. Gdy prawda staje się kwestią wyboru politycznego, a nie obiektywnego dowodu, fundamenty deliberatywnej demokracji ulegają rozkładowi.

Rewolucja Generatywnej AI i wyzwania syntetycznej rzeczywistości



Ilustracja redakcyjna TRUTH · Rewolucja Generatywnej AI i wyzwania syntetycznej rzeczywistości

Wkraczamy w nową fazę architektury oszustwa, napędzaną przez generatywną sztuczną inteligencję (Gen AI). Narzędzia te radykalnie obniżają koszty produkcji wysokiej jakości dezinformacji, pozwalając na masowe tworzenie tekstów, obrazów, a nawet filmów (deepfakes), które w określonych warunkach mogą wprowadzać odbiorców w błąd. Narzędzia AI mogą również wspierać dostosowanie przekazu pod kątem konkretnych profili psychologicznych odbiorców.

Główne zagrożenia związane z AI w kontekście dezinformacji to:

Mimikra Tożsamości: Możliwość naśladowania stylu i głosu osób publicznych, co pozwala na tworzenie fałszywych oświadczeń o ogromnym potencjale destabilizacyjnym.

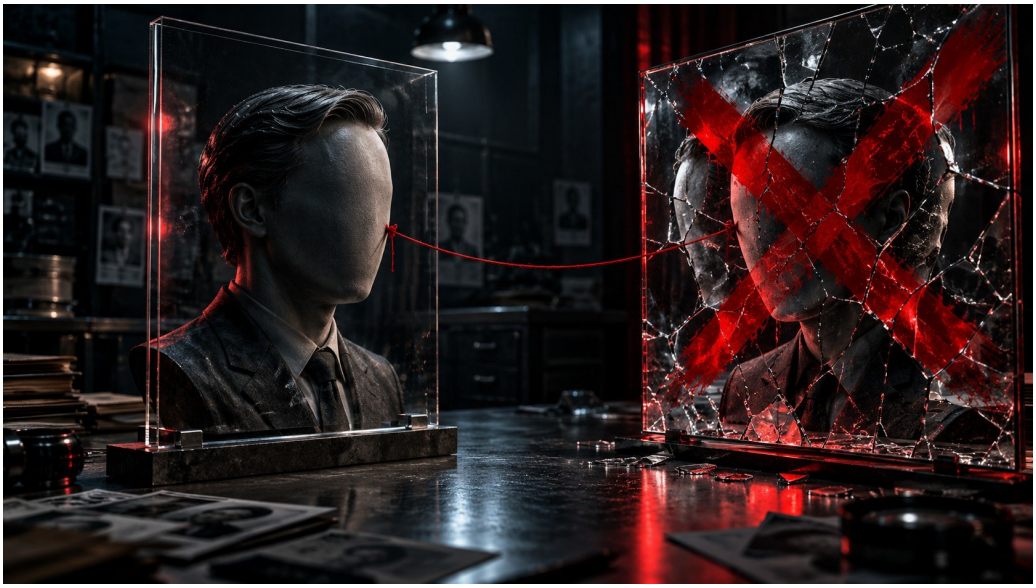
Automatyzacja manipulacyjnych interakcji: modele językowe mogą ułatwiać prowadzenie wielu rozmów i tworzenie pozorów oddolnej aktywności. Skala skutecznego wpływu zależy jednak od jakości narzędzi, moderacji i reakcji odbiorców.

Zanieczyszczenie danych: fałszywe treści mogą trafiać do zbiorów wykorzystywanych przez systemy AI. To ryzyko jakości danych, a nie dowód, że każdy model automatycznie utrwala każdą napotkaną narrację.

Jednocześnie AI może wspierać działania obronne. Algorytmy mogą wspierać fact-checkerów w szybkiej identyfikacji znanych wzorców dezinformacyjnych, analizować pochodzenie treści (provenance) oraz wykrywać anomalie w sieciach dystrybucji. Walka o prawdę staje się zatem wyścigiem zbrojeń algorytmicznych, w którym kluczowe znaczenie ma dostęp do wysokiej jakości danych i mocy obliczeniowej. Żaden detektor nie daje gwarancji autentyczności. Wynik analizy

automatycznej wymaga kontekstu i weryfikacji pochodzenia materiału.

Dywidenda kłamcy – Strategiczne wykorzystanie niepewności w erze deepfake'ów



Ilustracja redakcyjna TRUTH · Dywidenda kłamcy – Strategiczne wykorzystanie niepewności w erze deepfake'ów

Zjawisko „dywidendy kłamcy" (Liar's Dividend), opisane przez Bobby'ego Chesneya i Danielle Citron, stanowi jeden z najbardziej destrukcyjnych skutków upowszechnienia technologii deepfake. Polega ono na tym, że sama świadomość istnienia przekonujących fałszerstw cyfrowych pozwala osobom publicznym na zaprzeczanie autentyczności prawdziwych, kompromitujących materiałów. [6]

Strategia ta realizowana jest poprzez dwa główne mechanizmy:

Inwokacja niepewności informacyjnej: „W dobie AI nie można ufać niczemu, co widzi się w sieci" – takie stwierdzenie pozwala zdyskredytować niewygodne dowody bez konieczności udowadniania, że są one fałszywe.

Mobilizacja tożsamościowa: Przedstawianie prawdziwych materiałów jako „ataku dezinformacyjnego" ze strony przeciwników, co wzmacnia lojalność grupy własnej i polaryzację.

Dywidenda kłamcy opisuje ryzyko osłabienia odpowiedzialności przez samo powołanie się na możliwość fałszerstwa. Nie należy bez wskazania konkretnego eksperymentu twierdzić, że zaprzeczanie zawsze działa lepiej niż przeprosiny lub milczenie. Odpowiedzią pozostaje dokumentowanie pochodzenia, niezależne potwierdzenie i przejrzysta ocena dowodów. [6]

Regulacja przestrzeni cyfrowej – DSA i odpowiedzialność platform

Ilustracja redakcyjna TRUTH · Regulacja przestrzeni cyfrowej – Akt o Usługach Cyfrowych (DSA) jako nowa konstytucja internetu

Akt o usługach cyfrowych (DSA) wprowadza obowiązki dotyczące przejrzystości i odpowiedzialności pośredników internetowych. Zakres obowiązków różni się zależnie od kategorii i wielkości usługi. Największe platformy i wyszukiwarki podlegają dodatkowym wymogom zarządzania ryzykiem systemowym. [7]

Kluczowe filary DSA w walce z dezinformacją obejmują:

Oceny ryzyka: bardzo duże platformy i wyszukiwarki muszą oceniać ryzyka systemowe, m.in. dla debaty obywatelskiej, wyborów, zdrowia publicznego i praw podstawowych, oraz podejmować proporcjonalne działania ograniczające te ryzyka. [7]

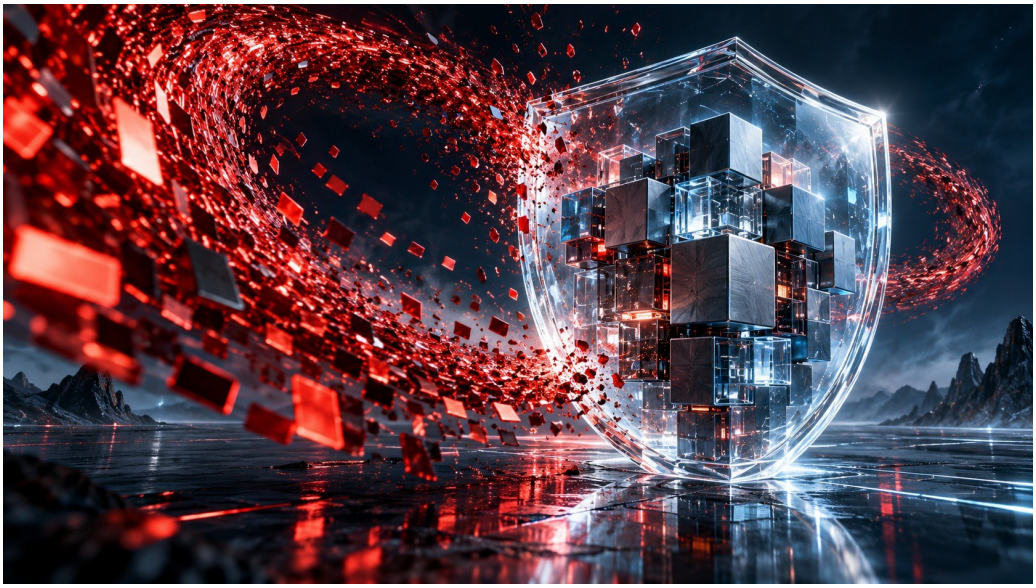
Reklamy i rekomendacje: DSA określa zasady przejrzystości reklam oraz głównych parametrów rekomendacji. Bardzo duże platformy i wyszukiwarki korzystające z rekomendacji mają zapewnić co najmniej jedną opcję nieopartą na profilowaniu; nie jest to powszechny nakaz wyłączenia każdej personalizacji reklam. [7]

Audyty i dane: dla największych usług przewidziano niezależne audyty oraz dostęp do danych dla uprawnionych organów i, przy spełnieniu warunków, zweryfikowanych badaczy. Nie oznacza to nieograniczonego publicznego dostępu do wszystkich danych. [7]

Zwodnicze interfejsy: DSA zabrania określonych praktyk projektowania interfejsów, które wprowadzają w błąd lub istotnie ograniczają świadomy wybór. Stosowanie tych przepisów wymaga uwzględnienia ich zakresu i innych regulacji. [7]

DSA nie ustanawia ogólnej zasady uznawania każdej fałszywej informacji za nielegalną. Łączy obowiązki dotyczące treści nielegalnych z zarządzaniem ryzykiem i ochroną praw podstawowych, w tym wolności wypowiedzi. To ramy rozliczalności platform, a nie gwarancja zniknięcia dezinformacji. [7]

Inokulacja psychologiczna i prebunking – Budowanie odporności poznawczej społeczeństwa

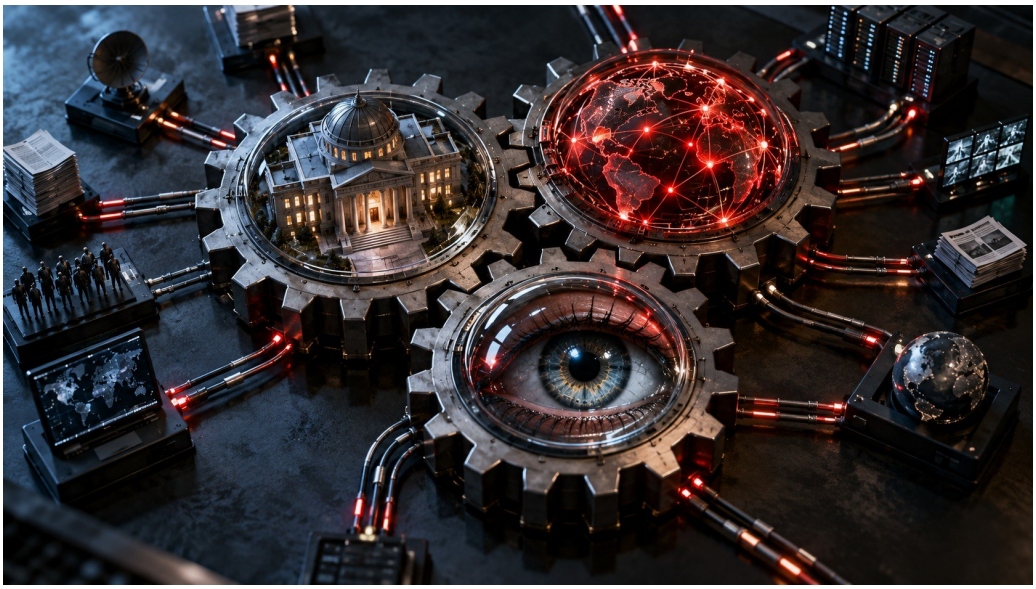


Ilustracja redakcyjna TRUTH · Inokulacja psychologiczna i prebunking – Budowanie odporności poznawczej społeczeństwa

Wobec skali dezinformacji tradycyjne prostowanie faktów (debunking) okazuje się niewystarczające. Coraz większą rolę odgrywa teoria inokulacji psychologicznej (psychological inoculation), która zakłada, że uprzednie wystawienie ludzi na osłabione formy manipulacji buduje ich odporność na przyszłe ataki dezinformacyjne. Jest to strategia proaktywna, znana również jako prebunking. Prebunking uzupełnia sprawdzanie faktów; nie zastępuje go. [8]

Badania eksperymentalne wskazują, że krótkie materiały uczące rozpoznawania manipulacji mogą poprawiać niektóre miary odporności. Efekty zależą od interwencji, sposobu pomiaru i czasu. Nowsza praca pokazała też, że samo rozpoznawanie technik nie musi przekładać się na lepszą ocenę prawdziwości nagłówków; połączenie inokulacji z przypomnieniem o dokładności dawało lepsze wyniki. Nie należy obiecywać trwałej odporności po jednorazowym szkoleniu. [8] [9] [10]

Synteza i kierunki polityki informacyjnej – W stronę ekosystemu zaufania



Ilustracja redakcyjna TRUTH · Synteza i kierunki polityki informacyjnej – W stronę ekosystemu zaufania

Architektura oszustwa jest problemem strukturalnym, który nie może zostać rozwiązany za pomocą pojedynczej interwencji technologicznej czy prawnej. Wymaga on holistycznej strategii, łączącej regulację platform, wsparcie dla rzetelnego dziennikarstwa oraz masową edukację medialną. Dezinformacja jest zjawiskiem chronicznym, głęboko zakorzenionym w dynamice współczesnego kapitalizmu i technologii, dlatego celem polityki publicznej nie powinno być jej całkowite wyeliminowanie (co jest niemożliwe), lecz skuteczne zarządzanie ryzykiem i budowanie odporności społecznej.

Kluczowe rekomendacje dla przyszłych działań obejmują:

Wielostronne Zarządzanie (Multi-stakeholder approach): Współpraca rządów, platform, nauki i społeczeństwa obywatelskiego w celu monitorowania i mitygacji zagrożeń.

Reforma Algorytmiczna: Przejście od optymalizacji pod kątem zaangażowania do modeli premiujących jakość epistemiczną i różnorodność opinii.

Wzmocnienie Suwerenności Poznawczej: Masowe programy inokulacji psychologicznej i edukacji w zakresie sztucznej inteligencji, przygotowujące obywateli do życia w świecie syntetycznych treści.

Internacjonalizacja Regulacji: Harmonizacja standardów bezpieczeństwa AI i przejrzystości platform na poziomie globalnym, aby zapobiec arbitrażowi regulacyjnemu.

Ostatecznie walka z dezinformacją jest walką o jakość naszej sfery publicznej i zdolność do podejmowania zbiorowych decyzji w oparciu o fakty. Wymaga to nie tylko zmian technicznych, ale przede wszystkim odbudowy zaufania do instytucji wiedzy i przywrócenia prawdy jako

fundamentalnej wartości w procesach społecznych i politycznych.

O autorze

Ivan-Oleksandr Yakymiv jest prezesem zarządu Stowarzyszenia Centrum Wsparcia „Just Do It", założycielem i wydawcą TRUTH Magazine, liderem projektu CODI-DW (Combating Disinformation in the Digital World) oraz założycielem CODI Institute (Combating Disinformation). CODI Institute jest inicjatywą projektową i szkoleniową poświęconą odporności cyfrowej, krytycznemu myśleniu, roli AI w dezinformacji oraz architekturze platform przeciwdziałającej manipulacji online.

Ivan-Oleksandr Yakymiv

Źródła i dalsze informacje

1. Claire Wardle & Hossein Derakhshan (2017), Information Disorder — Council of Europe
https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

2. Soroush Vosoughi, Deb Roy & Sinan Aral (2018), The spread of true and false news online — Science
https://doi.org/10.1126/science.aap9559

3. Aristotle, Rhetoric, Book I — Internet Classics Archive, MIT
https://classics.mit.edu/Aristotle/rhetoric.1.i.html

4. Christopher Paul & Miriam Matthews (2016), The Russian Firehose of Falsehood Propaganda Model — RAND
https://www.rand.org/pubs/perspectives/PE198.html

5. Computational Propaganda — Oxford Internet Institute
https://www.oii.ox.ac.uk/research/projects/computational-propaganda/

6. Bobby Chesney & Danielle Citron (2019), Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security — California Law Review
https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security

7. Regulation (EU) 2022/2065 — Digital Services Act, Articles 25–27, 34–40
https://eur-lex.europa.eu/eli/reg/2022/2065

8. Jon Roozenbeek et al. (2022), Psychological inoculation improves resilience against misinformation on social media — Science Advances
https://pmc.ncbi.nlm.nih.gov/articles/PMC9401631/

9. Gordon Pennycook et al. (2021), Shifting attention to accuracy can reduce misinformation online — Nature https://doi.org/10.1038/s41586-021-03344-2

10. Gordon Pennycook et al. (2024), Inoculation and accuracy prompting increase accuracy discernment in combination but not alone — Nature Human Behaviour https://doi.org/10.1038/s41562-024-02023-2

11. Shoshana Zuboff (2019), The Age of Surveillance Capitalism — author's introduction https://shoshanazuboff.com/book/about/

Kontakt

https://truthmagazine.eu/articles/ivan-oleksandr-yakymiv/ | redaktor@truthmagazine.pl

Wersja online: https://truthmagazine.eu/publikacje/architecture-of-deception/?lang=pl