

Maciej Saskowski

Uniwersytet Komisji Edukacji Narodowej w Krakowie

ORCID: 0000-0003-1611-7062

Fact-checking jako praktyka medialna: między rzetelnością a cenzurą informacyjną

Wprowadzenie

W dobie informacji cyfrowej, w której każdy użytkownik internetu może stać się nadawcą treści docierających do tysięcy odbiorców, problem prawdziwości informacji staje się kwestią kluczową dla jakości debaty publicznej. W odpowiedzi na rosnące zagrożenie dezinformacją oraz fake newsami, rozwija się zjawisko fact-checkingu – działalności polegającej na weryfikacji prawdziwości wypowiedzi, danych statystycznych, informacji udostępnianych przez media, polityków oraz inne osoby publiczne.

Organizacje fact-checkingowe, początkowo działające głównie jako inicjatywy dziennikarskie, dziś często współpracują z globalnymi platformami społecznościowymi i instytucjami międzynarodowymi. Odgrywają one coraz istotniejszą rolę w kształtowaniu środowiska informacyjnego, co potwierdzają globalne programy, takie jak Third-Party Fact-Checking realizowany przez Meta¹.

W tym kontekście fact-checking można postrzegać nie tylko jako narzędzie weryfikacji, lecz także jako wyraźną formę praktyki medialnej – czyli systematycznego działania komunikacyjnego, które wpływa na dynamikę debaty publicznej, relacje nadawca-odbiorca oraz strukturę władzy

¹ *Third-Party Fact-Checking on Facebook*, Meta Platforms, 7.02.2025, [on-line:] <https://www.facebook.com/journalismproject/programs/third-party-fact-checking> – 26.05.2025.

informacyjnej. Jest to praktyka oparta na selekcji, interpretacji i klasyfikowaniu wypowiedzi, która – choć często deklaratywnie neutralna – posiada zdolność realnego kształtowania hierarchii tematów, opinii i aktorów życia publicznego.

Choć z założenia fact-checking ma służyć ochronie prawdy i przeciwdziałać manipulacjom, jego funkcjonowanie nie pozostaje wolne od kontrowersji. Pojawiają się pytania o metodologiczne podstawy weryfikacji, niezależność instytucji fact-checkingowych, ich transparentność oraz – co szczególnie istotne z punktu widzenia wolności słowa – konsekwencje, jakie niesie za sobą etykietowanie treści jako „fałszywych”, „manipulacyjnych” czy „dezinformacyjnych”. W praktyce bowiem takie działania często prowadzą do ograniczania zasięgu lub całkowitego usuwania treści z przestrzeni publicznej.

Celem niniejszego artykułu jest pogłębiona analiza zjawiska fact-checkingu, jego funkcji, mechanizmów i wpływu na konstytucyjnie chronioną wolność wypowiedzi. Przedmiotem szczególnej refleksji będą napięcia pomiędzy potrzebą ochrony przestrzeni informacyjnej przed szkodliwą dezinformacją a prawem jednostki do swobodnego wyrażania poglądów – nawet tych kontrowersyjnych, niepopularnych czy błędnych. W artykule podjęta zostanie próba odpowiedzi na pytanie, czy fact-checking może pełnić rolę strażnika prawdy, nie stając się zarazem instrumentem cenzury, a także jakie mechanizmy prawne i instytucjonalne mogłyby zagwarantować równowagę między tymi dwoma wartościami.

W celu lepszego zobrazowania omawianych problemów, w artykule zastosowano metodę studiów przypadku (*case studies*), która pozwala na szczegółowe przeanalizowanie konkretnych sytuacji, w których działania fact-checkerów istotnie oddziaływały na debatę publiczną, granice wolności słowa oraz dostęp do informacji. Wybór tej metody wynika z przekonania, że tylko poprzez analizę realnych i wielowymiarowych przykładów – osadzonych w konkretnych kontekstach społeczno-politycznych – możliwe jest uchwycenie złożoności mechanizmów weryfikacji informacji jako praktyki medialnej, oraz rzeczywistego wpływu tych działań na praktykę komunikacyjną i funkcjonowanie demokracji. Case studies umożliwiają również wskazanie typowych kontrowersji, napięć i konsekwencji towarzyszących stosowaniu narzędzi fact-checkingowych – zarówno po stronie nadawców, jak i odbiorców treści.

W dalszych częściach artykułu omówione zostały cztery przypadki: działania weryfikacyjne podejmowane w trakcie pandemii COVID-19, moderacja treści związanych z wyborami prezydenckimi w Stanach Zjednoczonych w 2020 r., kontrola informacji podczas wojny w Ukrainie oraz zjawisko fact-checkingu w debacie publicznej na temat osób LGBT+.

Każdy z tych przykładów ukazuje inne oblicze napięcia między walką z dezinformacją a ochroną wolności wypowiedzi – i stanowi punkt wyjścia do refleksji nad granicami odpowiedzialnej moderacji w społeczeństwie informacyjnym.

Geneza i rozwój fact-checkingu

Zjawisko fact-checkingu, choć zyskało na znaczeniu szczególnie w ostatniej dekadzie, ma swoje korzenie w klasycznych praktykach dziennikarskich XX w. Weryfikacja faktów była od zawsze istotnym elementem pracy redakcyjnej, jednak dopiero w epoce tzw. postprawdy – naznaczonej rozmyciem granic między informacją a opinią – zaczęła być traktowana jako odrębna dziedzina działalności medialnej².

Pierwsze profesjonalne organizacje fact-checkingowe powstały w Stanach Zjednoczonych. Już w 1994 r. założono Center for Media and Public Affairs, a w 2003 r. powstał FactCheck.org – inicjatywa Annenberg Public Policy Center przy Uniwersytecie Pensylwanii. W 2007 r. uruchomiono PolitiFact, prowadzony przez redakcję „Tampa Bay Times”, a w 2010 r. działalność rozpoczął The Washington Post Fact Checker³. Rozkwit tego typu inicjatyw wiąże się bezpośrednio z rozwojem internetu, zwłaszcza mediów społecznościowych, które z jednej strony przyspieszyły cyrkulację informacji, z drugiej – radykalnie zmniejszyły kontrolę nad jakością treści. W odpowiedzi na rosnącą dezinformację w sferze publicznej – zwłaszcza w kontekście kampanii wyborczych i wydarzeń kryzysowych – fact-checking zyskał instytucjonalne wsparcie i zaczął przyjmować bardziej zorganizowane formy.

W Polsce zjawisko to upowszechniło się później, ale w ostatnich latach zyskało znaczną dynamikę. Do najważniejszych inicjatyw należą m.in. Demagog.org.pl (powstały w 2014 r.), Konkret24 (TVN), FakeHunter (PAP) oraz komponenty fact-checkingowe serwisu OKO.press⁴. Działalność ta obejmuje nie tylko analizę wypowiedzi publicznych, ale także edukację medialną i raportowanie na temat strategii dezinformacyjnych.

Przełomowym momentem dla globalnej koordynacji działań fact-checkingowych było powołanie w 2015 r. International Fact-Checking Network (IFCN) przy Instytucie Poyntera. IFCN sformułowało zasady

² S. Lewandowsky, U.K.H. Ullrich, J. Cook, *Beyond Misinformation: Understanding and Coping with the 'Post-Truth' Era*, „Journal of Applied Research in Memory and Cognition” 2017, vol. 6, no. 4, s. 353–369.

³ L. Graves, *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*, Columbia University Press, New York 2016, s. 45–67.

⁴ O nas, Demagog.org.pl, [on-line:] <https://demagog.org.pl/o-nas/> – 26.05.2025.

etyczne, których przestrzeganie jest warunkiem uzyskania certyfikacji. Kluczowe z nich to: przejrzystość metodologii, jawność finansowania, deklaracja bezstronności i możliwość korekty błędów⁵.

W kolejnych latach fact-checking przekształcił się z oddolnych, niezależnych inicjatyw dziennikarskich w istotny komponent polityki informacyjnej platform cyfrowych. Firmy takie jak Meta, Google czy TikTok nawiązały współpracę z lokalnymi i międzynarodowymi weryfikatorami treści, delegując im odpowiedzialność za klasyfikację informacji jako rzetelnych bądź fałszywych⁶. To z kolei rodzi pytania o zakres niezależności fact-checkerów, ich odpowiedzialność oraz relację z mechanizmami kontroli społecznej.

Mechanizmy i modele działania weryfikatorów informacji

Organizacje zajmujące się fact-checkingiem deklarują działanie w oparciu o zasady przejrzystości, bezstronności i weryfikowalności. Niemniej analiza ich funkcjonowania pokazuje, że mechanizmy weryfikacyjne są zróżnicowane, a ich praktyczna implementacja nie zawsze spełnia wysokie standardy rzetelności i neutralności.

Proces weryfikacji zazwyczaj obejmuje kilka etapów: identyfikację wypowiedzi lub treści do sprawdzenia, pozyskanie danych źródłowych, konsultacje z ekspertami, dokonanie oceny oraz publikację wraz z uzasadnieniem. Choć brzmi to technicznie, wiele decyzji podejmowanych w tym procesie ma charakter interpretacyjny i zależy od przyjętej metodologii⁷.

Jednym z kluczowych elementów działania weryfikatorów jest skala ocen, która może przyjmować różne formy – od binarnych „prawda/fałsz” po złożone kategorie takie jak „częściowa prawda”, „manipulacja”, „brak kontekstu” czy „bez dowodów”. Każde z tych oznaczeń niesie ze sobą ładunek wartościujący, który wpływa na odbiór treści przez użytkowników⁸.

Problemem metodologicznym pozostaje selekcja wypowiedzi poddawanych ocenie. Większość organizacji deklaruje, że wybiera treści istotne społecznie, popularne lub viralowe. W praktyce jednak dobór może być

⁵ *Code of Principles*, International Fact-Checking Network (IFCN), [on-line:] <https://ifcncodeofprinciples.poynter.org/> – 26.05.2025.

⁶ T. Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press, New Haven 2018, s. 103–117.

⁷ L. Graves, dz. cyt., s. 122–135.

⁸ C. Wardle, H. Derakhshan, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Council of Europe, 2017.

nacechowany politycznie lub ideologicznie. Badania wykazują, że odbiorcy częściej zarzucają stronniczość organizacjom, które nie ujawniają jasnych kryteriów selekcji⁹.

Coraz częściej działalność fact-checkingowa nie ogranicza się do publikowania analiz, ale jest zintegrowana z mechanizmami moderacyjnymi platform społecznościowych. Przykładowo, Meta we współpracy z zewnętrznymi partnerami fact-checkingowymi ogranicza zasięg treści uznanych za fałszywe, a w niektórych przypadkach blokuje ich dalsze udostępnianie. Oznaczenia te mogą skutkować również demonetyzacją lub czasowym zawieszeniem konta¹⁰. W efekcie fact-checking zyskuje realną moc wpływania na dostępność treści, co oznacza, że decyzje weryfikatorów – mimo że formalnie nie są wiążące prawnie – wywierają skutki społeczne i informacyjne, analogiczne do sankcji.

Z perspektywy regulacyjnej istotna jest także kwestia relacji instytucjonalnych i finansowych pomiędzy weryfikatorami a platformami technologicznymi. Przykładowo, część organizacji otrzymuje wsparcie finansowe od tych samych firm, których treści pomagają moderować. Budzi to pytania o niezależność i potencjalne konflikty interesów. Niepokój budzi również brak jawnych i niezależnych mechanizmów odwoławczych. W wielu przypadkach twórcy treści nie są informowani o szczegółach procesu oceny ani nie mają realnej możliwości zakwestionowania decyzji. Odbiorcy treści widzą jedynie oznaczenie – bez pełnego kontekstu i bez dostępu do źródeł, które posłużyły weryfikatorom do sformułowania werdyktu¹¹.

Fact-checking a media cyfrowe i platformy społecznościowe

Rozwój mediów cyfrowych, zwłaszcza dominacja platform społecznościowych, takich jak Facebook, Twitter (obecnie X), YouTube czy TikTok, znacząco zmieniły sposób produkcji, dystrybucji i odbioru informacji. Użytkownicy stali się nie tylko odbiorcami, ale również producentami treści, a algorytmy platform decydują o tym, które informacje zyskują widoczność, a które są marginalizowane.

W reakcji na rosnącą skalę dezinformacji, platformy te wdrożyły mechanizmy moderacyjne, które obejmują współpracę z organizacjami

⁹ Zob. L. Dierickx, C.-G. Lindén, *Transparency and fact-checking in open societies*, „Journalism”, 15.10.2024, <https://doi.org/10.1177/14648849241292200>.

¹⁰ *How Facebook Works with Fact-Checkers*, Facebook Help Center, [on-line:] <https://www.facebook.com/help/publisher/182222309230722> – 26.05.2025.

¹¹ T. Gillespie, dz. cyt., s. 142–146.

fact-checkingowymi. Meta (Facebook) oraz Google podpisały partnerstwa z certyfikowanymi weryfikatorami zrzeszonymi w ramach International Fact-Checking Network (IFCN), umożliwiając im oznaczanie treści publikowanych przez użytkowników jako „fałszywe”, „wprowadzające w błąd” lub „częściowo prawdziwe”¹². W praktyce oznacza to, że po negatywnej ocenie przez fact-checkera, treść może zostać objęta sankcjami: jej zasięg zostaje ograniczony, widoczność zmniejszona, a autor może stracić dostęp do funkcji komercyjnych (np. reklam czy monetyzacji). W skrajnych przypadkach dochodzi do czasowego lub trwałego usunięcia treści z platformy¹³. Mechanizmy te wzmacniają wpływ fact-checkerów, czyniąc z nich nie tylko komentatorów debaty publicznej, ale współdecydentów o tym, co może być widoczne, a co – ukryte. Jak zauważa Tarleton Gillespie, „platformy społecznościowe nie są już neutralnymi pośrednikami, ale aktywnymi zarządcami treści i reputacji”¹⁴.

Działania moderacyjne oparte na ocenach weryfikatorów spotkały się z krytyką ze strony części użytkowników, ekspertów i organizacji zajmujących się ochroną praw obywatelskich. Podnoszone są wątpliwości dotyczące braku przejrzystości procedur, asymetrii w traktowaniu różnych grup oraz ryzyka arbitralności decyzji. Brakuje też jednolitych i skutecznych mechanizmów odwoławczych.

Z perspektywy medioznawczej szczególnie interesujące jest zjawisko tzw. algorytmicznego wpływu – treści oznaczone jako fałszywe nie są wprawdzie usuwane, ale ich zasięg drastycznie spada, co prowadzi do zjawiska „shadow banning”, czyli niewidzialnego wykluczania z obiegu informacyjnego¹⁵. W efekcie użytkownik nie zostaje bezpośrednio oceniany, ale jego głos traci skuteczność oddziaływania – bez wyroku, bez procesu i bez jasnego powodu. Tego rodzaju systemy – choć technicznie sprawne – mogą prowadzić do efektu mrożącego, w którym użytkownicy autocenzurują się w obawie przed sankcjami. Z kolei twórcy niezależnych mediów i komentatorzy polityczni mogą stracić zaufanie do platform jako przestrzeni demokratycznej wymiany opinii.

Ograniczenia obiektywizmu i transparentności w weryfikacji faktów

Chociaż organizacje fact-checkingowe deklarują przywiązanie do standardów obiektywności i przejrzystości, w praktyce ich działalność

¹² *Code of Principles...*

¹³ Tamże.

¹⁴ T. Gillespie, dz. cyt., s. 145.

¹⁵ C. Wardle, H. Derakhshan, dz. cyt.

napotyka na poważne ograniczenia metodologiczne, poznawcze i instytucjonalne. Te ograniczenia nie tylko wpływają na wiarygodność procesu weryfikacji, lecz także rodzą pytania o jego neutralność w kontekście debaty publicznej.

Weryfikacja informacji nigdy nie jest procesem całkowicie neutralnym – wymaga interpretacji, kontekstu i często również osądu wartościującego. Przykładowo, wypowiedź polityka, że „bezrobocie jest najniższe w historii”, może być oceniona jako prawdziwa lub manipulacyjna w zależności od tego, czy uwzględni się różne metodologie pomiaru, sezonowość lub dane historyczne. Granice między kategoriami „częściowa prawda” a „manipulacja” bywają płynne.

Dodatkowo, weryfikatorzy korzystają z określonych źródeł – oficjalnych danych, ekspertyz naukowych, publikacji medialnych – które same w sobie nie są wolne od błędów ani ideologii. W rezultacie fakt „sprawdzenia” informacji nie oznacza jej ostatecznej prawdziwości, lecz odzwierciedla stan wiedzy w danym momencie.

Funkcjonowanie fact-checkingu podlega również ograniczeniom zasobów: organizacje muszą wybierać, które treści analizują. Nie istnieją jednolite kryteria doboru materiału do weryfikacji. Jedne podmioty skupiają się na wypowiedziach polityków, inne na treściach viralowych, a jeszcze inne – na zdrowiu publicznym czy migracjach. Ta selektywność może prowadzić do zarzutów o stronniczość, zwłaszcza jeśli wypowiedzi jednej strony sporu politycznego są analizowane częściej lub bardziej krytycznie. W odbiorze społecznym może to budować obraz weryfikatorów jako uczestników, a nie neutralnych obserwatorów debaty.

Kolejnym czynnikiem podważającym zaufanie do weryfikatorów jest sposób ich finansowania. Część z nich otrzymuje środki od fundacji, korporacji technologicznych lub instytucji publicznych. Choć formalnie deklarują niezależność, w praktyce brak pełnej jawności i potencjalne konflikty interesów podważają ich legitymizację. Przykładowo, jeśli organizacja fact-checkingowa otrzymuje grant od platformy społecznościowej, której treści ma weryfikować, powstaje zależność strukturalna. Również powiązania personalne – np. członkowie rad programowych mający afiliacje polityczne – mogą wpływać na odbiór działalności organizacji.

Wiele organizacji fact-checkingowych nie publikuje informacji o swoich pomyłkach ani nie prowadzi jawnych rejestrów korekt. Brak transparentnych mechanizmów przyznawania się do błędów osłabia ich wiarygodność. Tymczasem właśnie gotowość do rewizji stanowiska powinna być jednym z fundamentów rzetelnej działalności weryfikacyjnej. Zdarza się, że treść uznana za fałszywą w danym momencie – np. podczas pandemii COVID-19 – później okazuje się prawdziwa lub przynajmniej

prawdopodobna. Brak aktualizacji ocen może prowadzić do trwałego stygmatyzowania wypowiedzi i osób.

Fact-checking a konstytucyjna wolność słowa – analiza prawna

Wolność słowa jest jednym z fundamentów państwa demokratycznego. W Polsce została zagwarantowana w art. 54 ust. 1 Konstytucji RP: „Każdemu zapewnia się wolność wyrażania swoich poglądów oraz pozyskiwania i rozpowszechniania informacji”¹⁶. Jest to prawo o charakterze pozytywnym – oznacza nie tylko ochronę przed cenzurą, ale również aktywne umożliwienie uczestnictwa w debacie publicznej. Prawo to znajduje potwierdzenie także w art. 10 Europejskiej Konwencji Praw Człowieka oraz w orzecznictwie Europejskiego Trybunału Praw Człowieka, który wielokrotnie podkreślał, że wolność słowa „dotyczy nie tylko informacji lub idei przychylnie odbieranych, lecz również tych, które obrażają, szokują lub niepokoją”¹⁷.

Wolność słowa nie jest absolutna – może podlegać ograniczeniom przewidzianym przez prawo, jeśli są one konieczne w demokratycznym państwie do ochrony interesów, takich jak bezpieczeństwo narodowe, porządek publiczny, zdrowie, moralność lub prawa innych osób¹⁸. Kluczowe jest jednak to, że ograniczenia te powinny być proporcjonalne i poddane kontroli sądowej.

W kontekście fact-checkingu pojawia się pytanie, czy działania organizacji weryfikujących fakty i platform cyfrowych można uznać za formę ingerencji w wolność słowa, nawet jeśli nie pochodzą one bezpośrednio od organów państwa. Z jednej strony, organizacje te działają formalnie jako podmioty prywatne. Z drugiej jednak – platformy społecznościowe, będące dominującym kanałem debaty publicznej, pełnią dziś funkcję infrastruktury komunikacyjnej¹⁹.

Niektórzy badacze wskazują na zjawisko transformacji platform cyfrowych w quasi-publiczne przestrzenie debaty, których regulacje wewnętrzne mają realny wpływ na realizację konstytucyjnych praw jednostki. Zdaniem Magdaleny Szpunar, platformy takie powinny być traktowane jako „funkcjonalni następcy państwowych instytucji informacji

¹⁶ Konstytucja Rzeczypospolitej Polskiej z 2 kwietnia 1997 r., Dz.U. 1997 Nr 78, poz. 483.

¹⁷ ETPCz, sprawa Handyside przeciwko Wielkiej Brytanii, wyrok z 7 grudnia 1976 r., skarga nr 5493/72.

¹⁸ Tamże; por. też art. 10 ust. 2 EKPC.

¹⁹ T. Gillespie, dz. cyt., 151–153.

publicznej” i podlegać pewnym wymogom jawności i proporcjonalności²⁰. Ograniczanie zasięgu lub usuwanie treści na podstawie ocen fact-checkerów może w takim przypadku zostać uznane za faktyczne ograniczenie wolności słowa, szczególnie jeśli nie towarzyszą mu jasne procedury odwoławcze, uzasadnienie decyzji oraz możliwość obrony przez nadawcę treści.

Decyzje o oznaczeniu treści jako „fałszywe” mają często charakter jednostronny i nie podlegają niezależnemu nadzorowi. Ich skutkiem może być tzw. efekt mrożący (*chilling effect*) – zjawisko, w którym użytkownicy rezygnują z wypowiadania kontrowersyjnych opinii w obawie przed represją informacyjną²¹. Trybunał w Strasburgu wielokrotnie uznawał takie efekty za niezgodne z zasadą proporcjonalności, jeśli nie są odpowiednio uzasadnione²².

Brak publicznie dostępnych procedur odwoławczych, niezależnych recenzji czy możliwości rewizji ocen fact-checkerów podważa standardy państwa prawa. W tej sytuacji niezbędne staje się wprowadzenie minimalnych gwarancji proceduralnych dla użytkowników, zwłaszcza w kontekście ich prawa do swobodnego wyrażania poglądów.

Case study: pandemia COVID-19 i cenzura pandemii

Pandemia COVID-19 była momentem przełomowym dla systemów weryfikacji informacji. Stan zagrożenia zdrowia publicznego i dynamicznie zmieniająca się wiedza naukowa sprawiły, że organizacje fact-checkingowe oraz platformy społecznościowe stanęły przed ogromnym wyzwaniem: jak przeciwdziałać dezinformacji, nie ograniczając jednocześnie wolności wypowiedzi?

Już w pierwszych miesiącach pandemii wiodące platformy, takie jak Facebook, Twitter, YouTube i TikTok, nawiązały pogłębioną współpracę z organizacjami fact-checkingowymi oraz Światową Organizacją Zdrowia (WHO). Wprowadzono system oznaczania treści dotyczących COVID-19 jako „prawdziwych” lub „fałszywych”, usuwania wpisów zawierających „dezinformację zdrowotną” oraz blokowania kont powtarzających

²⁰ Zob. M. Szpunar, *Czym są nowe media – próba konceptualizacji*, „Studia Medioznawcze” 2008, nr 4 (35), s. 33.

²¹ Zob. W. Brzozowski, *Efekt mrożący*, „Studia Iuridica” 2023, nr 4, s. 38, [on-line:] https://ejournals.eu/pliki_artykułu_czasopisma/pelny_tekst/018f9af2-1603-7016-81f-f4475fe3e6d6c/pobierz – 26.05.2025.

²² ETPCz, sprawa Steel and Morris przeciwko Wielkiej Brytanii, wyrok z 15 lutego 2005 r., skarga nr 68416/01.

rzekomo nieprawdziwe informacje²³. Choć intencją było powstrzymanie fali fake newsów, efektem ubocznym była masowa eliminacja treści, które nie były jednoznacznie fałszywe, lecz prezentowały alternatywne opinie lub hipotezy, często dopiero weryfikowane przez naukę²⁴.

Jednym z najgłośniejszych przykładów była teoria, że wirus SARS-CoV-2 mógł pochodzić z laboratorium w chińskim Wuhan. Początkowo była ona uznawana przez większość fact-checkerów i instytucji zdrowotnych za teorię spiskową. W konsekwencji posty i artykuły poruszające ten temat były masowo oznaczane, usuwane lub blokowane przez algorytmy platform cyfrowych²⁵. Jednak w 2021 r. stanowisko to zaczęło się zmieniać. Pojawiły się publikacje w renomowanych mediach, takich jak Wall Street Journal, oraz oficjalne wypowiedzi naukowców i przedstawicieli wywiadu USA, sugerujące, że hipoteza laboratoryjna jest jedną z możliwych do rozważenia. WHO ogłosiła konieczność dalszych badań nad pochodzeniem wirusa, a Facebook przestał automatycznie usuwać wpisy o tej treści²⁶. To zdarzenie pokazało, że fact-checking może opierać się na chwilowym konsensusie, który z czasem ulega rewizji – co jest naturalne w nauce, ale problematyczne, gdy towarzyszy temu cenzura.

Kolejne pole konfliktu dotyczyło skuteczności masek, szczepionek oraz lockdownów. Przykładowo, początkowe rekomendacje WHO nie zalecały masek dla osób zdrowych, co później zostało zmienione. Osoby powołujące się na pierwsze stanowisko WHO były jednak oznaczane jako dezinformatorzy, mimo że ich wypowiedzi były zgodne z ówczesnym stanem wiedzy²⁷. Podobnie traktowano lekarzy i naukowców, którzy publicznie kwestionowali politykę lockdownów lub promowali inne metody leczenia COVID-19. Nawet jeśli ich wypowiedzi były merytoryczne i poparte danymi, często trafiały na czarne listy platform i były ograniczane zasięgowo lub usuwane²⁸.

²³ COVID-19 response and fact-checking, Meta Journalism Project [on-line:] <https://www.facebook.com/journalismproject/coronavirus-information> – 26.05.2025.

²⁴ S. Lewandowsky, U.K.H. Ullrich, J. Cook, dz. cyt., s. 353–369.

²⁵ Zob. A. Hern, *Facebook lifts ban on posts claiming Covid-19 was man-made*, „The Guardian”, 27.05.2021, [on-line:] <https://www.theguardian.com/technology/2021/may/27/facebook-lifts-ban-on-posts-claiming-covid-19-was-man-made> – 25.05.2025.

²⁶ M.R. Gordon, W.P. Strobel, *Intelligence on Sick Staff at Wuhan Lab Fuels Debate on Covid-19 Origin*, „The Wall Street Journal”, 23.05.2021, [on-line:] <https://www.wsj.com/world/china/intelligence-on-sick-staff-at-wuhan-lab-fuels-debate-on-covid-19-origin-11621796228> – 26.05.2025.

²⁷ Zob. *Coronavirus: WHO advises to wear masks in public areas*, BBC News, 6.06.2020, [on-line:] <https://www.bbc.com/news/health-52945210> – 25.05.2025.

²⁸ Zob. *Cenzura nie tylko covidowa. KULISY wielkiej operacji okłamywania społeczeństwa, narodów, Świata*, Strona Mirosława Dakowskiego, [on-line:] <https://www.dakowski>.

Wprowadzone przez platformy cyfrowe mechanizmy fact-checkingowe doprowadziły do istotnych konsekwencji społecznych: użytkownicy przestali ufać, że ich opinie mogą być uczciwie oceniane, a niektórzy zrezygnowali z aktywności online w obawie przed stygmatyzacją²⁹. Odbiorcy treści zaczęli postrzegać weryfikację jako narzędzie represji, nie ochrony. Zjawisko to przyczyniło się do pogłębienia polaryzacji społecznej i spadku zaufania do instytucji publicznych, mediów i ekspertów. W efekcie fakt, że jakaś informacja została oznaczona jako „fałszywa”, nie przekładał się już na jej dyskredytację – przeciwnie, bywał postrzegany jako dowód na „niewygodną prawdę”.

Case study: wybory prezydenckie w USA w 2020 r. i blokowanie treści politycznych

Wybory prezydenckie w Stanach Zjednoczonych w 2020 r. były jednym z najintensywniej monitorowanych wydarzeń politycznych pod kątem dezinformacji w historii internetu. Już na kilka miesięcy przed głosowaniem platformy społecznościowe – w tym Facebook, Twitter (obecnie X) oraz YouTube – zapowiedziały zaostrzenie polityki moderacyjnej wobec treści uznawanych za fałszywe lub „wprowadzające w błąd”. Na niespotykaną wcześniej skalę uruchomiono współpracę z zewnętrznymi weryfikatorami informacji (m.in. PolitiFact, USA Today Fact Check, Lead Stories). Oceniali oni wypowiedzi kandydatów, komentarze użytkowników oraz publikacje medialne. Wiele z tych analiz było automatycznie zintegrowanych z algorytmami platform – skutkując etykietami ostrzegawczymi, ograniczeniem zasięgów lub usunięciem treści³⁰.

Szczególną uwagę poświęcono wypowiedziom ówczesnego prezydenta Donalda Trumpa, który wielokrotnie kwestionował legalność głosowania korespondencyjnego. W odpowiedzi Twitter zaczął oznaczać jego wpisy jako „sporne” lub „zawierające dezinformację wyborczą”³¹. Działania te eskalowały po dniu wyborów, a kulminacją było trwałe zablokowanie konta Trumpa po wydarzeniach z 6 stycznia 2021 r. w Kapitolu.

pl/cenzura-covidowa-kulisy-wielkiej-operacji-oklamywaniaspoleczenstwa-narodow-swiata/ – 25.05.2025.

²⁹ A. Marwick, R. Lewis, *Media Manipulation and Disinformation Online*, Data & Society, 15.05.2017, [on-line:] <https://datasociety.net/library/media-manipulation-and-disinfo-online/> – 25.05.2025.

³⁰ *PolitiFact's coverage of Election 2020*, PolitiFact, [on-line:] <https://www.politifact.com/2020/> – 25.05.2025.

³¹ E. Chen, H. Chang, A. Rao, K. Lerman, G. Cowan, E. Ferrara, *COVID-19 misinformation and the 2020 U.S. presidential election*, Harvard Kennedy School (HKS) Misinformation Review, 3.03.2021, <https://doi.org/10.37016/mr-2020-57>.

Decyzje platform wzbudziły gorącą debatę. Zwolennicy działań moderacyjnych argumentowali, że miały one na celu ochronę procesu demokratycznego przed próbami jego delegitymizacji. Krytycy natomiast podkreślali, że usuwanie konta urzędującego prezydenta USA bez wyroku sądowego stanowi niebezpieczny precedens dla wolności słowa w przestrzeni cyfrowej. Dodatkowo, treści zawierające relacje o nieprawidłowościach wyborczych – nawet jeśli opierały się na faktach lub zawierały opinie dziennikarzy śledczych – były często traktowane jako „podejrzane” i oznaczane, co zniechęcało do ich dalszego rozpowszechniania.

Wielu komentatorów zwracało uwagę na asymetrię moderacyjną – treści konserwatywne były częściej oznaczane jako dezinformacja niż liberalne, co doprowadziło do oskarżeń o stronniczość polityczną. Niektóre platformy, takie jak Parler czy Gab, odnotowały gwałtowny wzrost użytkowników, którzy opuścili mainstreamowe media społecznościowe w poszukiwaniu „wolnej” przestrzeni informacyjnej³².

Case study: wojna w Ukrainie i moderacja informacji o działaniach wojennych

Inwazja Rosji na Ukrainę w lutym 2022 r. wywołała nie tylko konflikt zbrojny, ale również bezprecedensową wojnę informacyjną. Media społecznościowe, platformy wideo oraz komunikatory stały się głównym polem walki o interpretację wydarzeń, mobilizację opinii publicznej i dezinformację. W odpowiedzi na zagrożenie, platformy cyfrowe zintensyfikowały współpracę z fact-checkerami oraz agencjami rządowymi, starając się kontrolować obieg treści związanych z wojną.

YouTube, Facebook, TikTok i Twitter wdrożyły automatyczne mechanizmy wykrywania treści powiązanych z Rosją, zablokowały konta rosyjskich mediów państwowych (RT, Sputnik) i rozpoczęły oznaczanie informacji uznanych za prorosyjską propagandę. Dużą rolę odegrały tu międzynarodowe organizacje fact-checkingowe, takie jak StopFake.org, EUvsDisinfo, Reuters Fact Check i AFP Fact Check. Zjawisko to wpisuje się w szerszy trend: platformy nie tylko reagują na dezinformację, ale również aktywnie kształtują przekaz wojenny, promując określone narracje (np. heroizm Ukraińców, zbrodnie wojenne Rosjan), przy jednoczesnym filtrowaniu materiałów „nieweryfikowalnych” lub „szkodliwych dla morale”³³.

³² Zob. J. Shieber, *Parler jumps to No. 1 on App Store after Facebook and Twitter ban Trump*, TechCrunch, 9.01.2022, [on-line:] <https://techcrunch.com/2021/01/09/parler-jumps-to-no-1-on-app-store-after-facebook-and-twitter-bans/> – 26.05.2025.

³³ Zob. C. de Guzman, *Meta Is Globally Banning Russian State Media on Its Apps, Citing 'Foreign Interference'*, „Time”, 17.09.2024, [on-line:] <https://time.com/7021897/meta-facebook-instagram-whatsapp-rt-russian-state-media-global-ban/> – 25.05.2025.

Część użytkowników i dziennikarzy zwracała uwagę, że platformy zaczęły usuwać także autentyczne relacje z miejsc walk, jeśli zawierały przemoc, treści drastyczne lub niejednoznaczne przekazy (np. jeńcy wojenni, obrazy zniszczeń po stronie rosyjskiej). Mimo że niektóre z tych treści były publikowane przez media lub świadków, zostały zaklasyfikowane jako „wzmacniające dezinformację” lub „naruszające standardy społeczności”. Taka praktyka stawia pytania o granice między ochroną użytkowników a dezaktywizacją przekazu informacyjnego. Czy filtrowanie rzeczywistości wojennej przez algorytmy i fact-checkerów nie prowadzi do wymazywania ważnych aspektów konfliktu?

W kontekście wojny w Ukrainie szczególnie jaskrawo widoczne stały się asymetrie narracyjne – treści wspierające działania Ukrainy były wzmacniane, podczas gdy te krytyczne wobec NATO, UE lub polityki zbrojeniowej Zachodu były często marginalizowane lub etykietowane jako „prorosyjskie”. Dotyczyło to także publicystów, analityków i ekspertów, których wypowiedzi były merytoryczne, lecz niezgodne z dominującym przekazem. W rezultacie fact-checking, choć spełniał funkcję ochrony przed manipulacją, mógł też przyczynić się do formowania jednolitego frontu informacyjnego, który ograniczał dostęp do alternatywnych interpretacji wydarzeń.

Ostatecznie wojna w Ukrainie unaoczniała, że platformy cyfrowe są nie tylko firmami technologicznymi, ale także aktorami geopolitycznymi. Ich decyzje dotyczące oznaczania treści, blokowania mediów i wspierania konkretnych narracji mają charakter strategiczny. W tym sensie fact-checking staje się nie tyle mechanizmem neutralnej analizy, co instrumentem globalnego zarządzania informacją.

Case study: debata o osobach LGBT+ i granice fact-checkingu

Tematy związane z osobami LGBT+ są przedmiotem intensywnych sporów kulturowych i politycznych, zarówno w mediach tradycyjnych, jak i w przestrzeni cyfrowej. Z jednej strony to obszar walki o prawa człowieka, równość i akceptację. Z drugiej – temat podatny na dezinformację, moralną panikę oraz silne emocje społeczne. Fakt, że wiele treści ma charakter opiniotwórczy i światopoglądowy, sprawia, że fact-checking w tej dziedzinie jest wyjątkowo trudny i delikatny.

W kontekście LGBT+ często dochodzi do mieszania faktów z opiniami. Przykładowo, wypowiedzi dotyczące edukacji seksualnej w szkołach, praw adopcyjnych par jedнопłciowych czy tożsamości płciowej bywają przedstawiane jako dezinformacja lub „teorie spiskowe” – mimo że

odnoszą się do zagadnień, w których istnieje szerokie spektrum przekonań i interpretacji.

Problem ten był szczególnie widoczny w Polsce w 2019 i 2020 r., kiedy niektóre wypowiedzi polityków i publicystów o „ideologii LGBT” zostały oznaczone przez weryfikatorów jako „manipulacyjne” lub „fałszywe” – mimo że opierały się na ocenie ideowej, nie zaś na faktach poddających się prostemu sprawdzeniu.

W 2020 r. wiele polskich samorządów przyjęło uchwały sprzeciwiające się „ideologii LGBT”. Kampania ta została nagłośniona międzynarodowo i spotkała się z reakcją ze strony Komisji Europejskiej, a także globalnych platform cyfrowych. Google i Facebook ograniczyły widoczność niektórych treści promujących te uchwały, a organizacje fact-checkingowe (np. Demagog.org.pl) opublikowały analizy wskazujące na błędy logiczne lub nieprawdziwe tezy zawarte w wypowiedziach samorządowców³⁴.

Z drugiej strony, przeciwnicy tych kampanii – w tym aktywiści LGBT+ – zgłaszali, że ich relacje z protestów były również usuwane lub oznaczane jako treści kontrowersyjne, zwłaszcza gdy zawierały nagrania z interwencji policji, obraźliwe transparenty lub obrazy przemocy. Platformy społecznościowe powoływały się na „naruszenia standardów społeczności”, co – w oczach aktywistów – stanowiło przykład niezamierzonej cenzury.

Tematy transpłciowości, zwłaszcza w kontekście dzieci i młodzieży, są kolejnym polem napięć. Treści dotyczące terapii hormonalnej, sportu kobiet transpłciowych czy identyfikacji płciowej w szkołach są często przedmiotem sporów światopoglądowych. W niektórych krajach (np. w Wielkiej Brytanii i USA) wpisy krytyczne wobec niektórych aspektów polityk „affirmative care” były oznaczane jako dezinformacja, mimo że odwoływały się do publikacji naukowych lub raportów rządowych. W tym kontekście fact-checking zaciera granice między recenzją merytoryczną a kontrolą ideologiczną, zwłaszcza jeśli platformy działają w oparciu o nieprzejrzyste zasady.

Weryfikowanie treści związanych z LGBT+ to wyzwanie wymagające ekstremalnej wrażliwości na kontekst kulturowy, językowy i prawny. Oznaczenie wypowiedzi jako „fałszywej” może prowadzić do wyciszenia ważnych głosów – zarówno tych walczących o prawa, jak i tych reprezentujących obawy społeczne. Prawdziwa debata publiczna powinna dopuszczać kontrowersje, pod warunkiem, że nie są one formą nienawiści.

³⁴ *Dezinformacja na temat osób LGBT+. Pięć fałszywych narracji*, Demagog.org.pl, 13.07.2023, [on-line:] https://demagog.org.pl/analizy_i_raporty/dezinformacja-na-temat-osob-lgbt-piec-falszywych-narracji/ – 26.05.2025.

Podsumowanie case studies: praktyka fact-checkingu w konfliktach społeczno-politycznych

Analiza czterech przypadków działania mechanizmów fact-checkingowych w kontekście kluczowych wydarzeń politycznych i społecznych – pandemii COVID-19, wyborów prezydenckich w USA (2020), wojny w Ukrainie oraz debat wokół osób LGBT+ – ujawnia szereg wspólnych cech oraz systemowych wyzwań związanych z weryfikacją informacji w przestrzeni cyfrowej.

W każdym z omawianych przypadków weryfikacja faktów okazywała się czymś więcej niż tylko sprawdzeniem zgodności informacji z rzeczywistością. Była zarazem formą symbolicznego działania: potwierdzania lub delegitymizowania określonych przekazów, wpływania na reputację autorów, a także aktywnego udziału w sporach światopoglądowych i politycznych. W efekcie fact-checking często działał nie tylko jako strażnik prawdy, lecz także jako narzędzie kontroli narracji – zarówno w rękach instytucji publicznych, jak i platform cyfrowych.

We wszystkich analizowanych przypadkach szczególnie trudna okazała się weryfikacja treści opiniotwórczych, emocjonalnych lub odnoszących się do złożonych zjawisk (np. skuteczność szczepień, uczciwość wyborów, zbrodnie wojenne, tożsamość płciowa). W takich sytuacjach próby jednoznacznego ocenienia wypowiedzi jako „fałszywej” często prowadziły do uproszczeń lub wykluczeń głosów mniejszościowych. Zarówno zwolennicy, jak i przeciwnicy dominujących narracji, skarżyli się na brak pluralizmu i asymetrię w traktowaniu różnych punktów widzenia.

Zarówno podczas pandemii, jak i w kontekście wojny informacyjnej w Ukrainie czy wyborów w USA, platformy cyfrowe wprowadzały sankcje informacyjne bez jasnych procedur odwoławczych. Treści były usuwane, ograniczane lub etykietowane w sposób nieprzejrzysty dla odbiorców, a decyzje te często zapadały na podstawie niepublicznych ocen fact-checkerów. To prowadziło do efektu mrożącego i utraty zaufania użytkowników – nawet tych, którzy popierali cel walki z dezinformacją.

Choć organizacje weryfikacyjne często deklarują bezstronność, ich działalność jest osadzona w kontekście instytucjonalnym, finansowym i ideologicznym. Współpraca z platformami, granty od rządów lub korporacji, wybór tematów i język ocen – wszystko to wpływa na odbiór ich pracy jako potencjalnie stronnictwej lub represyjnej. Dotyczy to w szczególności treści politycznych, religijnych, moralnych i kulturowych, w których ocena faktu jest nierozzerwalnie związana z interpretacją.

Case studies pokazują, że fact-checking, choć potencjalnie korzystny dla debaty publicznej, może – w zależności od kontekstu – osłabiać

pluralizm, wzmacniać dominujące ideologie i marginalizować głosy kontrowersyjne, lecz legalne. Dlatego niezbędne jest wprowadzenie: przejrzystych zasad działania weryfikatorów, mechanizmów odwoławczych i korekty, edukacji użytkowników, która wzmacnia krytyczne myślenie, zamiast zastępować je autorytarnym etykietowaniem. Tylko wtedy fact-checking będzie w stanie pełnić funkcję rzetelnego strażnika prawdy, a nie instrumentu dyscyplinowania debaty.

Fact-checking jako narzędzie wpływu i kontroli informacyjnej – perspektywa medioznawcza

Fact-checking, mimo swojej funkcji ochrony debaty publicznej przed dezinformacją, może być analizowany również jako narzędzie kontroli narracji i wpływu symbolicznego. Z perspektywy medioznawczej działalność weryfikatorów nie odbywa się w próżni – wchodzi w relację z mechanizmami władzy, hegemonii informacyjnej i formowania opinii publicznej.

Organizacje fact-checkingowe nie tylko sprawdzają prawdziwość wypowiedzi – dokonują również selekcji, które informacje zasługują na weryfikację. Sam akt wyboru jest aktem redakcyjnym, który wpływa na strukturę debaty. To, co zostaje oznaczone, zyskuje status „problemu społecznego”, a to, co pominięte – znika z radarów opinii publicznej.

Z perspektywy teorii agendy (*agenda-setting*), media nie tyle mówią ludziom, co mają myśleć, co raczej – o czym mają myśleć. Podobną rolę odgrywa dziś fact-checking: wpływa na percepcję ważności i wiarygodności danych tematów. Oznaczenie wypowiedzi jako „fałszywej” przez uznanego fact-checkera niesie skutki nie tylko dla treści, ale i dla nadawcy. Powstaje efekt reputacyjny, w którym osoba lub organizacja zostaje społecznie napiętnowana. W przypadkach granicznych może to prowadzić do zjawiska ostracyzmu cyfrowego³⁵. Użytkownicy objęci takimi sankcjami często nie mają realnych możliwości obrony – ich wpisy są ukrywane, a konta ograniczane. Nawet jeśli ocena fact-checkera zostaje później zakwestionowana, wizerunkowe szkody bywają trudne do odwrócenia.

Fact-checking może – często nieintencjonalnie – wzmacniać narracje dominujące w debacie publicznej i eliminować te, które wykraczają poza główny nurt. Dotyczy to zwłaszcza kwestii spornych: migracji, polityki klimatycznej, zdrowia publicznego, wojny informacyjnej. W takich

³⁵ Zob. M.E. McCombs, D.L. Shaw, *The Agenda-Setting Function of Mass Media*, „Public Opinion Quarterly”, 1972, vol. 36, no. 2, s. 176-187; H. Habib, S. Elsharawy, R. Rahman, „I don't trust them”: Exploring Perceptions of Fact-checking Entities for Flagging Online Misinformation, arXiv, 2024, [on-line:] <https://arxiv.org/abs/2410.00866> – 26.05.2025.

przypadkach rozróżnienie między „prawdą” a „opinią” bywa szczególnie trudne. Jak zauważa Peter Pomerantsev, w świecie nasyconym informacją „prawda nie musi być zakazana – wystarczy, że zostanie zatopiona w natłoku konkurencyjnych narracji”³⁶. Fact-checking, o ile nie jest stosowany z rozważą, może przyczyniać się do tej właśnie sytuacji: zamiast rozjaśniać spór, wzmacnia jedno stanowisko kosztem pluralizmu.

Działania weryfikatorów informacji są często splecione z interesami platform cyfrowych, które funkcjonują w logice kapitalizmu nadzoru Shoshany Zuboff – modelu, w którym dane użytkowników są towarem, a moderacja treści narzędziem zarządzania nastrojami społecznymi³⁷. Oznacza to, że fact-checking bywa używany nie tylko do ochrony prawdy, lecz również jako narzędzie optymalizacji przekazu: eliminowania treści „ryzykownych” dla stabilności systemu, reklamodawców lub interesów geopolitycznych. W tym sensie można mówić o ekonomii oznaczania prawdy.

Perspektywy, rekomendacje i modele regulacyjne

Dalszy rozwój fact-checkingu jako elementu ochrony debaty publicznej wymaga nie tylko doskonalenia metodologii i transparentności, ale także stworzenia ram regulacyjnych i etycznych. W tym kontekście kluczowe jest pogodzenie dwóch fundamentalnych wartości demokratycznych: waliki z dezinformacją i ochrony wolności słowa.

Podstawą zaufania społecznego do organizacji fact-checkingowych powinna być pełna transparentność ich działalności. Obejmuje to: publikację metodologii oceny treści, jawność źródeł finansowania, wskazanie kryteriów wyboru wypowiedzi do analizy, informowanie o ewentualnych korektach ocen i błędach. Wprowadzenie standardów, takich jak „Code of Principles” International Fact-Checking Network (IFCN) to krok w dobrym kierunku, ale ich egzekwowanie pozostaje wyzwaniem.

Organizacje weryfikujące informacje powinny być finansowane w sposób, który gwarantuje ich niezależność od podmiotów zainteresowanych wynikami ich pracy – zwłaszcza platform cyfrowych, rządów czy instytucji ideologicznie zaangażowanych. Warto rozważyć stworzenie funduszy pluralistycznych, finansowanych z różnych źródeł (np. publicznych, prywatnych, unijnych), ale zarządzanych przez niezależne ciała eksperckie. Taki model pozwalałby zminimalizować ryzyko konfliktu interesów i zależności strukturalnych.

³⁶ P. Pomerantsev, *This Is Not Propaganda*, Faber & Faber, London 2019, s. 54–58.

³⁷ S. Zuboff, *The Age of Surveillance Capitalism*, PublicAffairs, New York 2019.

Użytkownicy, których treści zostały oznaczone lub usunięte na podstawie oceny fact-checkerów, powinni mieć możliwość odwołania się od tej decyzji. Systemy odwoławcze mogłyby działać na poziomie: samej platformy (panel ekspercki), organizacji weryfikacyjnej (tryb rewizji oceny), niezależnego ciała arbitrażowego (np. publiczny ombudsman cyfrowy). Brak takich mechanizmów pogłębia poczucie arbitralności i niesprawiedliwości wśród użytkowników.

W dłuższej perspektywie najskuteczniejszym narzędziem walki z dezinformacją jest kształcenie odbiorców – nie tyle w kierunku biernego przyjmowania ocen fact-checkerów, co samodzielnego weryfikowania informacji. Edukacja medialna powinna obejmować: umiejętność rozpoznawania źródeł, krytyczne myślenie, rozumienie mechanizmów algorytmicznych, świadomość zjawisk manipulacyjnych. Organizacje fact-checkingowe mogłyby w tym zakresie pełnić rolę edukatorów społecznych – pod warunkiem zachowania neutralności i pluralizmu.

Na poziomie Unii Europejskiej i państw członkowskich pojawiają się pierwsze próby uregulowania działań weryfikacyjnych:

- Digital Services Act (DSA) wprowadza obowiązek raportowania działań moderacyjnych, w tym współpracy z fact-checkerami³⁸.
- Kodeks postępowania przeciwko dezinformacji (UE) zawiera zobowiązania dobrowolne, których skuteczność jest obecnie monitorowana³⁹.
- W Niemczech funkcjonuje NetzDG, nakładający obowiązki szybkiego reagowania na nielegalne treści, choć jego nadmierna restrykcyjność budzi kontrowersje⁴⁰.

Polska nie posiada obecnie kompleksowych regulacji w tym zakresie, co powoduje brak jasnych ram działania zarówno dla organizacji fact-checkingowych, jak i użytkowników platform.

Zakończenie

Fact-checking jest bez wątpienia jednym z najważniejszych narzędzi w walce z dezinformacją w epoce cyfrowej. Może służyć jako mechanizm ochrony debaty publicznej, budowania kultury informacyjnej oraz

³⁸ *Digital Services Act*, European Commission, [on-line:] <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package> – 26.05.2025.

³⁹ *The 2022 Code of Practice on Disinformation*, European Commission, [on-line:] <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation> – 26.05.2025.

⁴⁰ *Network Enforcement Act Regulatory Fining Guidelines*, 22.03.2018, [on-line:] https://www.bundesjustizamt.de/SharedDocs/Downloads/DE/NetzDG/Leitlinien_Geldbussen_en.pdf?__blob=publicationFile&v=3 – 26.05.2025.

wzmacniania standardów dziennikarskich. Jednak jego skuteczność i legitymizacja zależą od sposobu, w jaki jest realizowany, oraz od ram instytucjonalnych i komunikacyjnych, które mu towarzyszą.

Analiza przeprowadzona w niniejszym artykule pokazuje, że fact-checking – rozumiany jako praktyka medialna – nie jest wolny od ryzyka i ograniczeń. Problemy dotyczą zarówno kwestii metodologicznych (np. selekcji treści do weryfikacji czy kryteriów ocen), jak i strukturalnych, związanych z finansowaniem, współpracą z platformami cyfrowymi oraz brakiem przejrzystości i mechanizmów społecznej kontroli.

Szczegółne napięcie ujawnia się w relacji między działalnością fact-checkingową a konstytucyjnie chronioną wolnością słowa. Choć samo istnienie weryfikatorów nie stanowi naruszenia tego prawa – przeciwnie, może być jego przejawem – zagrożenie pojawia się wtedy, gdy ich werdykty stają się podstawą do ograniczania dostępu do legalnych treści, bez jasnych procedur odwoławczych i bez odpowiedzialności za skutki decyzyjne.

W świetle analizowanych przypadków – zwłaszcza z okresu pandemii COVID-19 – widoczny staje się efekt mrozący, który nie polega na formalnej cenzurze, lecz na algorytmicznej marginalizacji poglądów odbiegających od dominujących narracji. Tego typu praktyki medialne wpływają na redystrybucję uwagi w przestrzeni cyfrowej, a w konsekwencji – na to, co uznawane jest za „wiarygodne”, „naukowe” czy „społecznie akceptowalne”.

Z tego względu niniejszy artykuł formułuje szereg rekomendacji:

- wprowadzenie przejrzystych standardów i procedur odwoławczych;
- zapewnienie niezależności i jawności organizacji weryfikacyjnych;
- wspieranie edukacji medialnej i informacyjnej jako alternatywy wobec kontroli treści;
- wypracowanie ram prawnych równoważących walkę z dezinformacją i prawa jednostki.

W ostatecznym rozrachunku fact-checking może pełnić rolę strażnika prawdy – ale tylko pod warunkiem, że pozostanie praktyką przejrzystą, odpowiedzialną i demokratycznie nadzorowaną. Jego legitymizacja nie wynika z deklaracji misji, lecz z konkretnej praktyki medialnej – osadzonej w strukturze władzy komunikacyjnej, lecz podlegającej krytycznej refleksji.

Bibliografia

- Brzozowski W., *Efekt mrożący*, „Studia Iuridica”, nr 4 (2023), s. 33–47, [on-line:] https://ejournals.eu/pliki_artykulu_czasopisma/pelny_tekst/018f9af2-1603-7016-81ff-4475fe3e6d6c/pobierz – 26.05.2025.
- Cenzura nie tylko covidowa. KULISY wielkiej operacji okłamywania społeczeństwa, narodów, Świata*, Strona Mirosława Dakowskiego, [on-line:] [https://www.dakowski.pl/cenzura-covidowa-kulisy-wielkiej-operacji-oklamywan-s-po-lecz-en-s-twa-narodow-swiata/](https://www.dakowski.pl/cenzura-covidowa-kulisy-wielkiej-operacji-oklamywan-spo-lecz-en-s-twa-narodow-swiata/) – 25.05.2025.
- Chen E., Chang H., Rao A., Lerman K., Cowan G., Ferrara E., *COVID-19 misinformation and the 2020 U.S. presidential election*, Harvard Kennedy School (HKS) Misinformation Review, 3.03.2021, <https://doi.org/10.37016/mr-2020-57>.
- Code of Principles*, International Fact-Checking Network (IFCN), [on-line:] <https://ifcncodeofprinciples.poynter.org/> – 26.05.2025.
- Coronavirus: WHO advises to wear masks in public areas*, BBC News, 6.06.2020, [on-line:] <https://www.bbc.com/news/health-52945210> – 25.05.2025.
- Dezinformacja na temat osób LGBT+. Pięć fałszywych narracji*, Demagog.org.pl, 13.07.2023, [on-line:] https://demagog.org.pl/analizy_i_raporty/dezinformacja-na-temat-osob-lgbt-piec-falszywych-narracji/ – 26.05.2025.
- Dierickx L., Lindén C.G., *Transparency and Fact-Checking in Open Societies*, „Journalism”, 2024, <https://doi.org/10.1177/14648849241292200>.
- Digital Services Act*, European Commission, [on-line:] <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package> – 26.05.2025.
- Gillespie T., *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press, New Haven 2018.
- Gordon M.R., Strobel W.P., *Intelligence on Sick Staff at Wuhan Lab Fuels Debate on Covid-19 Origin*, „The Wall Street Journal”, 23.05.2021, [on-line:] <https://www.wsj.com/world/china/intelligence-on-sick-staff-at-wuhan-lab-fuels-debate-on-covid-19-origin-11621796228> – 26.05.2025.
- Graves L., *Deciding What’s True: The Rise of Political Fact-Checking in American Journalism*, Columbia University Press, New York 2016.
- Guzman C. de *Meta Is Globally Banning Russian State Media on Its Apps, Citing ‘Foreign Interference’*, „Time”, 17.09.2024, [on-line:] <https://time.com/7021897/meta-facebook-instagram-whatsapp-rt-russian-state-media-global-ban/> – 25.05.2025.
- Habib H., Elsharawy S., Rahman R., *„I don’t trust tchem”: Exploring Perceptions of Fact-checking Entities for Flagging Online Misinformation*, arXiv, 2024, [on-line:] <https://arxiv.org/abs/2410.00866> – 26.05.2025.
- Hern A., *Facebook lifts ban on posts claiming Covid-19 was man-made*, „The Guardian”, 27.05.2021, [on-line:] <https://www.theguardian.com/technology/2021/may/27/facebook-lifts-ban-on-posts-claiming-covid-19-was-man-made> – 25.05.2025.
- How Facebook Works with Fact-Checkers*, Facebook Help Center, [on-line:] <https://www.facebook.com/help/publisher/182222309230722> – 26.05.2025.

- Lewandowsky S., Ecker U.K.H., Cook J., *Beyond Misinformation: Understanding and Coping with the 'Post-Truth' Era*, „Journal of Applied Research in Memory and Cognition” 2017, vol. 6, no. 4, s. 353–369.
- Marwick A., Lewis R., *Media Manipulation and Disinformation Online*, Data & Society, 15.05.2017, [on-line:] <https://datasociety.net/library/media-manipulation-and-disinfo-online/> – 25.05.2025.
- McCombs M.E., Shaw D.L., *The Agenda-Setting Function of Mass Media*, „Public Opinion Quarterly” 1972, vol. 36, no. 2, s. 176–187.
- Network Enforcement Act Regulatory Fining Guidelines*, 22.03.2018, [on-line:] https://www.bundesjustizamt.de/SharedDocs/Downloads/DE/NetzDG/Leitlinien_Geldbussen_en.pdf?__blob=publicationFile&v=3 – 26.05.2025.
- PolitiFact's coverage of Election 2020*, PolitiFact, [on-line:] <https://www.politifact.com/2020/> – 25.05.2025.
- Pomerantsev P., *This Is Not Propaganda*, Faber & Faber, London 2019.
- Shieber J., *Parler jumps to No. 1 on App Store after Facebook and Twitter ban Trump*, „TechCrunch”, 9.01.2021, [on-line:] <https://techcrunch.com/2021/01/09/parler-jumps-to-no-1-on-app-store-after-facebook-and-twitter-bans/> – 26.05.2025.
- Szpunar M., *Czym są nowe media – próba konceptualizacji*, „Studia Medioznawcze” 2008, nr 4 (35)s. 25–38.
- The 2022 Code of Practice on Disinformation*, European Commission, [on-line:] <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation> – 26.05.2025.
- Third-Party Fact-Checking on Facebook*, Meta Platforms, 7.02.2025, [on-line:] <https://www.facebook.com/journalismproject/programs/third-party-fact-checking> – 26.05.2025.
- Wardle C., Derakhshan H., *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*, Council of Europe, 2017.
- Zuboff S., *The Age of Surveillance Capitalism*, PublicAffairs, New York 2019.

Streszczenie

Artykuł podejmuje problematykę fact-checkingu jako współczesnej praktyki medialnej, analizując jego funkcje, ograniczenia oraz wpływ na wolność słowa w erze cyfrowej. Choć fact-checking deklaratywnie pełni rolę strażnika prawdy i narzędzia walki z dezinformacją, w praktyce wiąże się z licznymi kontrowersjami dotyczącymi przejrzystości, selekcji treści, relacji z platformami cyfrowymi oraz braku społecznej kontroli. Autor wskazuje, że działania weryfikacyjne – mimo potencjalnej wartości informacyjnej – mogą prowadzić do marginalizacji legalnych, choć niepopularnych poglądów, a nawet do efektu mrożącego w debacie publicznej. Na podstawie czterech studiów przypadku (COVID-19, wybory w USA, wojna w Ukrainie, debata o osobach LGBT+) pokazano, że fact-checking nie jest wyłącznie procesem technicznym, lecz stanowi istotny element walki o kontrolę nad narracją publiczną. Artykuł zawiera również rekomendacje dotyczące konieczności

ustanowienia ram instytucjonalnych, które zagwarantują przejrzystość, proporcjonalność i pluralizm w stosowaniu narzędzi weryfikacyjnych.

Słowa kluczowe: fact-checking, wolność słowa, dezinformacja, media cyfrowe, moderacja treści

Fact-checking as a media practice: between information reliability and censorship

Abstract

This article explores fact-checking as a contemporary media practice, analyzing its functions, limitations, and impact on freedom of expression in the digital era. While fact-checking is often presented as a tool to protect truth and combat disinformation, in practice it raises concerns related to transparency, content selection, relationships with digital platforms, and the lack of public oversight. The study argues that fact-checking, despite its potential informational value, can lead to the marginalization of legal yet unpopular views and contribute to a chilling effect in public debate. Based on four case studies (COVID-19, the 2020 U.S. presidential election, the war in Ukraine, and the public discourse on LGBT+ individuals), the article demonstrates that fact-checking is not merely a technical process but a form of narrative power in the information space. It concludes with recommendations for institutional safeguards that ensure transparency, proportionality, and pluralism in the use of verification mechanisms.

Keywords: fact-checking, freedom of expression, disinformation, digital media, content moderation