

Activity 5

Analyzing the impact of disinformation on the society

D5.5. Raport

Dynamika i struktura przepływu (dez)informacji

Grzegorz Rzeczkowski, Collegium Civitas

Przemysław Witkowski, Collegium Civitas

Roland Zarzycki, Collegium Civitas

Wprowadzenie: koncepcja

Otoczająca nas rzeczywistość informacyjna zaroża się od fake newsów. Można odnieść wrażenie, że w XXI wieku dotykamy bezprecedensowych trudności komunikacyjnych wynikających z rozmycia



się granic prawdziwości, zaś gwałtowność rozwoju technologicznego i trudny do ogarnięcia zalew informacji czynią nas bezbronnymi w obliczu tych nowych wyzwań. Choć jednak termin fake news nie był używany przez lwią część historii ludzkości, nie oznacza to, że z przefałszowaniami rzeczywistości oraz rozprzestrzenianiem się plotek i fałszywych informacji nie borykaliśmy się w przeszłości. Co więcej, przez całe wieki tkwiliśmy w teoriach i modelach świata, które potem okazywały się fałszywe (Kuhn, 2012). Czy więc fake newsem jest to, że znaleźliśmy się w jakichś nowych okolicznościach? A jeśli okoliczności faktycznie są nowe, to na czym polega ta różnica?

Jak poniekąd ustalono, Donald Trump w okresie swojej czteroletniej kadencji prezydenckiej zamieścił na serwisie Twitter (obecnie X) 30573 false or misleading claims (Kessler, 2021). Interesujące jest to, że gorąca dyskusja polityczna wokół dezinformacji szerzonej przez Trumpa nie podejmuje w wiarygodny sposób zagadnienia sposobów weryfikacji prawdziwości lub nieprawdziwości tych 30753 false or misleading claims. Wydaje się, że polaryzacja pomiędzy stronami sporu jest tak skrajna, że sam spór nie ma już charakteru racjonalnego lub też, innymi słowy, toczy się w odniesieniu do dwóch różnych struktur racjonalności, które są wzajemnie nieprzystające (cf. MacIntyre, 1988). Dotykamy tu więc problemu, kto ma prawo i kompetencje, by decydować o prawdziwości.

W codziennym życiu stwierdzanie fałszywości bazuje na naszych przekonaniach co do istoty prawdziwości, osadzonych mocno w naszym światopoglądzie, statusie oraz rolach społecznych. O sukcesie nowożytnej metody naukowej w formule rozwijanej od około 17 wieku było to, że wypracowała ona schematy działania pozwalające w jakimś stopniu abstrahować od konkretnych światopoglądów, statusów oraz ról społecznych. Innym komponentem owego sukcesu była testowalność nauki i jej otwartość na falsyfikowalność. To z kolei oznaczało, że – wielokrotnie w przeciwieństwie do poznania religijnego – nauka godzi się na to, że żaden z przyjętych modeli naukowych nie jest ultimate i nie może rościć sobie prawa do prawdy absolutnej – każdy bowiem z modeli podlega falsyfikowalności.

Jedną z naszych podstawowych intuicji badawczych jest to, że w dzisiejszej debacie wokół fake newsów to nie zjawisko fake newsów jest bezprecedensowe, ale próba zastosowania nauki do orzeczenia o prawdzie i fałszu w sposób absolutny. Z perspektywy metodologii nauk istnieje

fundamentalna różnica pomiędzy błędnym rozumowaniem realizowanym w ramach danego paradygmatu, kiedy to w wyniku informed discussion, na bazie pewnych pryncypiów myślenia właściwych dla danego paradygmatu rozstrzygamy o błędzie w ramach danego modelu, a sytuacją, w której ścierają się ze sobą dwa niewspółmierne paradygmaty naukowe (Kuhn, 2012). W szerokim kontekście politycznym i ideologicznym owe dwa zagadnienia zlewają się w jedno, kiedy określone akторы życia społecznego próbują pewne kwestie rozstrzygnąć w sposób ostateczny nadużywając autorytetu nauki. Owo roszczenie do obiektywizmu napotyka jednocześnie opór, zarówno w łonie samej nauki, ale także w postaci niezgody szerokich formacji politycznych, które wskazują na szereg słabości świata nauki, w tym jego komodyfikację i instrumentalizację polityczną. Zjawisko to znajduje doskonałe przełożenie na debatę i praktykę funkcjonowania fake newsów w przestrzeni publicznej.

Nie jest zadaniem tego raportu rozstrzygać o kondycji współczesnej nauki. Nie sposób jednak zrozumieć dynamiki i struktury rozprzestrzeniania się dezinformacji, pomijając percepcję społeczną owych fundamentalnych zagadnień dotyczących autorytetu nauki. I tak według badania Digital Poland (2022) jedynie 11% badanych zgodziło się, że naukowcy w badaniach stosują rzetelne metody badawcze, a kolejne 42% raczej z tym stwierdzeniem się zgodziło, co w szczególności daje bardzo niewiele ponad połowę odpowiedzi oznaczających zaufanie wobec naukowych metod weryfikacji prawdziwości. Jednocześnie według tego samego badania 6% osób zdecydowanie uważa, że naukowcy prowadzą badania na zlecenie dużych korporacji przez co wyniki tych badań są zafałszowane, a kolejne 31% uważa, że raczej tak jest. Oznacza to, że w znacznej liczbie przypadków weryfikacja informacji metodami naukowymi nie jest kontestowana na gruncie naukowej polemiki, ale na gruncie niewiarygodności nauki jako takiej, co z kolei wydaje się być konsekwencją owego nadużywania autorytetu nauki do rozstrzygnięcia w sposób absolutny kwestii, które wykraczają poza akceptowany paradygmat.

W tym kontekście analizie poddaliśmy szereg istniejących modeli teoretycznych oraz praktycznych pozwalających analizować rozprzestrzenianie się informacji. Naszym celem było zbadanie dynamiki i struktury (dis)information flows, by lepiej zrozumieć zjawisko dezinformacji jako takiej. W wyniku tej refleksji nad literaturą tematu doszliśmy do wniosku, że wyjaśnienie kluczowych aspektów

zjawiska dezinformacji wymaga wyjścia poza te modele i zwrócenia uwagi na meta-czynniki w ramach analizy heurystycznej. Co standardowo rozpoznaje się jako czynniki pozwalające powstrzymać rozprzestrzenianie się dezinformacji? Edukacja, cenzura, fact-checking (Jackson et al., 2022). Wydaje się jednak, że niedostatecznie wiele uwagi poświęca się co najmniej następującym meta-czynnikom:

- (1) Ideologicznemu tłu oraz ideologicznej zbieżności paradygmatów, którymi posługują się odbiorcy (dez)informacji);
- (2) Politycznemu i ideologicznemu uwarunkowaniu działań ukierunkowanych na zwalczanie rozprzestrzeniania się (dez)informacji, w tym tzw. fact-checkingu;
- (3) Samej operacyjnej definicji fake newsa, pomijając bowiem dysputy akademickie, w praktyce fact-checkingowej zbyt lekko podchodzi się do samej możliwości orzekania o fałszywości lub prawdziwości danej informacji, a tym samym klasyfikowania jej jako fake news.

Głównym celem publikacji jest ustalenie wpływu tych czynników na obraz the dynamics and structure of the (dis)information flows, jaki wyłania się z szeregu badań empirycznych i analizy dostępnych modeli teoretycznych. Raport zorganizowany jest w następujący sposób. W sekcji [1] przedstawiamy metodologię naszego badania, które składa się z analizy teoretyczno-koncepcyjnej w oparciu o literaturę tematu oraz własne badanie empiryczne, stanowiące dopełnienie części teoretycznej. W sekcji przedstawiamy [2] omawiamy literaturę przedmiotu poświęconą modelowaniu the dynamics and structure of the (dis)information flows, jak również pewne istotne obserwacje dotyczące samych sposobów modelowania tych zjawisk. W sekcji trzeciej opisujemy stan wiedzy oraz dyskutujemy [3] znaczenie fact-checkingu dla rozprzestrzeniania się fake-newsów oraz ograniczenia i obciążenia samego fact-checkingu. Wyniki badania empirycznego przedstawiamy w sekcjach [4] funkcjonowanie polskich platform fact-checkingowych oraz w sekcji [5] elementy analizy statystycznej oraz social network analysis poświęconej dynamice rozprzestrzeniania się wybranych fake newsów w polskojęzycznym internecie. W sekcji [6] wnioski finalne przedstawianego raportu. Sekcja ostatnia zawiera spis cytowanej literatury.

Cele i metodologia badań

Przegląd istniejącej literatury tematu pozwolił nam ustalić szereg metod, które wykorzystywane są do modelowania dynamiki i struktury (dis)information flows. Kluczowymi czynnikami wpływającymi na tę dynamikę są, szeroko rozumiane, tło ideologiczne (w tym homofilia) oraz edukacja edukacja (Jackson et al., 2022) jako elementy statyczne oraz różnego rodzaju cenzura oraz fact-checking jako elementy dynamiczne. Cele badawcze naszego badania obejmują:

- (1) [cel koncepcyjny] krytyczna refleksja nad debatą wokół zysków i strat wynikających z aplikowania wzmiankowanych powyżej elementów dynamicznych (cenzura oraz fact-checking) wpływających na the dynamics and structure of the (dis)information flows;
- (2) [cel empiryczny] ustalenie charakterystyki polskiego fact-checkingu oraz jego wpływu na the dynamics and structure of the (dis)information flows w polskojęzycznym Internecie.

Z celem pierwszym wiążemy dwa szczegółowe pytania badawcze o charakterze spekulatywnym:

(Q1.1) W jakim stopniu funkcjonujący obecnie sposób tworzenia modeli opisujących the dynamics and structure of the (dis)information flows uwzględnia kwestie uwarunkowań politycznych, ideologicznych, w szczególności w ich napięciu względem zagadnień metodologii nauk oraz teorii prawdy?

(Q1.2) Czy efekty aplikowania cenzury oraz fact-checkingu dla ograniczania rozprzestrzenia się misinformation and disinformation są społecznie korzystne?

Z celem drugim wiążemy trzy szczegółowe pytania badawcze o charakterze empirycznym:

(Q2.1) Czy polskie portale fact-checkingowe dokładają należytej staranności, by funkcjonować w sposób możliwie jak najbardziej zobiektywizowany?

(Q2.2) Czy polskie portale fact-checkingowe różnicują metody weryfikacji misinformation and disinformation w zależności od typu weryfikowanej informacji?

(Q2.3) Czy fact-checking wpływa na rozprzestrzenianie się wiadomości opisywanych przez polskie portale fact-checkingowe w polskojęzycznej części funkcjonowania portalu X (dawniej: Twitter)?

W przypadku pytań Q1.1 oraz Q1.2 badanie realizujemy na drodze krytycznej refleksji oraz dyskusji na bazie dostępnej literatury bez formułowania szczegółowych hipotez.

Z pytaniami Q2.1, Q2.2 oraz Q2.3 wiążemy trzy hipotezy badawcze:

(H2.1) Główne polskie portale fact-checkingowe działają w sposób tendencyjny, realizując określoną agendę polityczną.

(H2.2) Polskie portale fact-checkingowe nie różnicują w należyty sposób metod weryfikacji fake-newsów.

(H2.3) Fact-checking nie ma wpływu na rozprzestrzenianie się wiadomości opisywanych przez polskie portale fact-checkingowe w polskojęzycznej części funkcjonowania portalu X (dawniej: Twitter).

Metodologia części empirycznej badania obejmuje dwa komponenty badawcze:

(A) Analizę zawartości wszystkich opublikowanych na oficjalnych stronach internetowych wyników procesów fact-checkingowych realizowanych przez wiodące polskie platformy fact-checkingowe w okresie jednego miesiąca (08.09.2023-07.10.2023). Dla wyznaczenia wiodących polskich platform fact-checkingowych wykorzystano raport Digital Poland (2022) wskazujący listę serwisów, które określają się mianem fact-checkingowych (w kolejności zależnej od deklarowanej przez respondentów popularności): oko.press (37%), konkret24 (35%), sprawdzimy.info (33%), fakenews.pl (32%), demagog.org.pl (23%), sprawdzam.afp.com (20%), pravda.org.pl (16%), fakehunter.pap.pl (12%) oraz odfekuj.info (8%). Oprócz serwisów wzmiankowanych przez raport Digital Poland funkcjonują w Polsce także inne serwisy, jak na przykład Demaskator24, jednak z powodu ich znikomego udziału w rynku, nie zostały poddane analizie.

(B) Analizę trzech case studies wybranych spośród wiadomości analizowanych jako fake news: (1) dezinformacja związana z cenami paliw w państwowej sieci stacji benzynowych Orlen w okresie przedwyborczym; (2) dezinformacja związana z kondycją ekonomiczną polski w komunikatach Narodowego Banku Polskiego; (3) dezinformacja związana z wadliwością szczepionek podawanych Polakom w okresie pandemii COVID-19 w raportach polskiej Najwyższej Izby Kontroli. Badanie

zrealizowano na bazie wyszukanej po słowach kluczach (odpowiednio: szczepienia + NIK; ceny + Orlen; ceny + NBP) wśród wiadomości zawartych na serwisie X (dawniej: Twitter) z wykorzystaniem narzędzia Hoaxy oraz wśród korpusu wiadomości skonstruowanego z wykorzystaniem narzędzia Sketchengine.

Hipoteza H2.1 weryfikowana była poprzez analizę zawartości i profilu treści weryfikowanych przez każdą z wiodących polskich platform fact-checkingowych względem ich deklaracji z jednej strony oraz standardów etycznych dla platform fact-checkingowych z drugiej. Hipoteza H2.2 weryfikowana była poprzez analizę typów treści weryfikowanych przez każdą z wiodących polskich platform fact-checkingowych względem klasyfikacji UNHCR w zestawieniu ze stosowaną metodą fact-checkingu. Hipoteza H2.3 weryfikowana była poprzez analizę tła dyskursywnego danego komunikatu z wykorzystaniem SketchEngine, a następnie dynamiki rozprzestrzeniania się tweetów dotyczących danego komunikatu z wykorzystaniem Hoaxy.

Modelowanie przepływu fake newsów

Historia modelowania riot behavior oraz rumor diffusion jest długa, a u jej koncepcyjnego podłoża leży pojęcie "threshold", który musi zostać przekroczony, by dany przekaz stał się wiarygodny (Granovetter, 1978). Tym samym, z samej już tej konstrukcji wynika, że o percepcji danego przekazu nie decyduje istota rzeczy lub zobiektywizowane kryteria prawdziwości, ale liczba osób. Jednocześnie wiemy, że dobór wiadomości, które jednostki upowszechniają i same biorą uważają za znaczące, wynika ze z góry ustalonego światopoglądu ((Hahn & Iyengar, 2009). Livan et al. (2020) proponują learning model, który wyjaśnia w jaki sposób individuals are empowered to self-select narratives they want to be exposed to and a minority of them reject information that not convergent with their worldview. Problemem jest oczywiście bezpodstawne założenie, że jest to mniejszość. Jest to spójne z tak zwaną koherencyjną teorią prawdy, a także ugruntowane przez sprzyjający polaryzacji i tworzeniu się komunikacyjnych bubbles rozwoju nowych mediów.

Od dłuższego czasu tworzonych jest wiele kwantytatywnych modeli rozprzestrzeniania i detekcji fake newsów (Africano Franco et al., 2023; Aldini, 2021; Apostolidis et al., 2022; Arini et al. 2023;

Batistela et al. 2020; Franceschi & Pareschi, 2022; Franceschi et al., 2022; Liming et al, 2021; Qian et al. 2021; Ren & Zhang, 2021). Sama integracja know how wynikające z każdego z tych setek badań wydaje się ważnym, potrzebnym, ale i trudnym zadaniem badawczym. Z perspektywy niniejszego raportu, na szczególną uwagę zasługuje kilka modeli. Interesujące, podejście zorientowane na analizę motywacji proponują Alwan et al. (2021), analizując wielowymiarowo motywację do rozpowszechniania informacji przez użytkowników Internetu ze względu na siedem wyszczególnionych czynników. Z kolei model wypracowany przez Aramaki et al. (2021) uwzględnia dynamikę rozprzestrzeniania się fake newsów ze względu na analizę dynamiki w podziale na czas do momentu rozpoznania danej wiadomości jako fake i od tego momentu. Z drugiej strony Havlin et al. (2021) pokazują, że wychwytywanie fake newsów jest możliwe jeszcze przed tym momentem, bowiem badanie dynamiki ich rozprzestrzeniania wskazuje na strukturalne różnice wobec dynamiki rozprzestrzeniania zwykłych wiadomości. Interesująca klasyfikacja faz rozprzestrzenienia dezinformacji zaproponowana została także przez Lynnette & Taeihagh (2021). Brady et al. (2020) diagnozują zjawisko *moral contagion*, które ma znaczący wpływ na szerzenie się fake newsów, oraz proponują objaśniający je model MAD. Kluczową konkluzją tych obserwacji jest zwrócenie uwagi, że group-identity-based motivations są kluczowe przy komunikowaniu treści o zabarwieniu moral-emotional, co z kolei współgra z samą architekturą mediów społecznościowych, które właśnie na bazie tego fenomenu budują swoją potęgę. W tym kontekście warto wspomnieć także badanie Centola et al. (2005), w którym pokazano, że w sytuacjach odseparowania jednostek i wyeksponowania ich na komunikację z immediate neighbors o ekstremistycznych poglądach, highly questionable norms can emerge locally and be disseminated. Ważne podejście do modelowania fake newsów zaprezentowane zostało w pracy Kopp et al. (2018), gdzie tradycyjne podejście do modelowania zachowań komunikacyjnych wzbogacone zostało o ujęcie deception z perspektywy teorii gier.

Fundamentalnie intrygujące jest jednak to, że każdy z tych modeli zakłada, że w przypadku pojedynczej wiadomości możliwe jest ustalenie, czy jest ona fałszywa, czy nie. Co więcej, sieci neuronowe służące do konstrukcji tych modeli bazują na takich elementarnych rozróżnieniach, jak również na doborze materiału empirycznego służącego ich uczeniu. To z kolei oznaczałoby

możliwość jasnego ustalenia, co w skład takiego materiału powinno zostać włączone. Tymczasem wydaje się, że na poziomie czysto koncepcyjnym nie mamy uniwersalnego praktycznego narzędzia odróżniania prawdziwego newsa od fake newsa. Tu kluczowe jest rozróżnienie, o którym wspominaliśmy w poprzedniej sekcji, a mianowicie, czy fałszywość danej wiadomości weryfikujemy w ramach pewnego przyjętego paradygmatu, bowiem taki jest jej charakter, czy też fałszywość danej wiadomości powinna być rozpatrywana w terminach konkurujących paradygmatów. Niestety, oba te typy przypadków są w praktyce komunikacyjnej mieszane. Jednocześnie założenie, że w przypadku pojedynczej, atomowej wiadomości możliwe jest rozstrzygnięcie o jej prawdziwości, które jest kluczowe przy konstrukcji wielu modeli, jest bardzo dyskusyjne w praktyce.

Podejście do kwestii definicji misinformation and disinformation z perspektywy pragmatycznej teorii prawdy prowadzi nas do ważnych obserwacji. Wśród 10 wyróżnionych przez UNHCR (2021) typów misinformation and disinformation mamy trzy, które są weryfikowalne na bazie technicznej analizy treści (manipulated content, imposter content, sponsored content), cztery, które są weryfikowalne w sposób spekulatywno-dyskusyjny na bazie krytyczno-retorycznej analizy dyskursu (misleading content, false context, false connections, error) oraz trzy (fabricated content, satire and parody, propaganda), w przypadku których weryfikacja jest praktycznie niemożliwa bez bardzo mocnych dodatkowych założeń dotyczących koncepcji prawdy, natury komunikacji, a niejednokrotnie wręcz rozstrzygnięć metafizycznych (przy czym warto zaznaczyć, że nie oznacza to, że wszystkie przypadki z kategorii fabricated content są trudne do rozstrzygnięcia). Niestety, w praktyce debaty publicznej oraz popularnego obecnie fact-checkingu nie czyni się tutaj żadnego rozróżnienia i często każdy typ misinformation and disinformation traktowany jest jako weryfikowalny w oparciu o dominujące źródła.

Jak wskazują ważne źródła (Ireton & Posetti, 2018), dobrze rozpoznane zjawiska takie jak confirmation bias, motivated reasoning czy availability heuristic mogą znacząco wpłynąć na percepcję faktów. Jak przy tym zauważają autorki, "Fact-checking itself, it should be noted, is an imperfect instrument. Something can be 100% accurate, and yet still leave out important context. Facts are invariably constructed, ordered and reordered meaningfully within broader narrative structures which can provide different significance to the same basic facts. The truth, moreover, is

more than a collection of facts. Fact-checking is not a tool to be deployed to shut down alternative interpretations as much as underwriting a series of facts that can impact on narrative and individual predispositions, so as to ground rational debate."

Definicja misinformation and disinformation staje się problematyczna w praktyce operacyjnej, w związku z czym przy tworzeniu modeli jej rozprzestrzeniania się stosuje się matematyczne uproszczenia, które wprawdzie pozwala generować atrakcyjne modele, ale niestety pozostawia wątpliwości co do adekwatności tych modeli wobec modelowanej rzeczywistości. Przykładowo, Brody et al. (2022) definiują model przyjmując, że fake newsem jest każdy szum zaburzający informację o niezerowej wartości oczekiwanej zaburzenia. Takie ujęcie zagadnienia bazuje na ważnej presupozycji, że dysponujemy kryterium pozwalającym na oddzielać informację od szumu.

Część zespołów badawczych konstruuje swoje modele, próbując uchwycić istotę tej trudności. Jak zauważa Cowen et al. (2022) abstrakcyjne modele, które sprawdzają się w przypadku epidemiologii, mają ograniczone zastosowanie, kiedy przetwarzanie sygnałów przez sieć zależy od wewnętrznej struktury psychologicznej czy systemu wierzeń każdego jej wierzchołka. Jak z kolei argumentują Baldassarri & Bearman (2007) modele muszą uwzględniać wielowymiarowość światopoglądu.

Jednocześnie dużym i dynamicznie rozwijającym się obszarem badawczym jest weryfikacja skuteczności działań ukierunkowanych na powstrzymanie rozprzestrzenienia się misinformation and disinformation. Jak zauważa Patwa et al. "compared to traditional news media, the lack of effective regulation and fact-checking mechanisms over internet posts enables the creation and publication of false news to influence opinions and obtain low-cost revenues" (Patwa, et al., 2021). Thus, różnego rodzaju platformy internetowe seem to provide a conducive environment for the propagation of false news (Ghorbani & Zhang, 2020). Co istotne, o ile flagging posts is not an effective way to stop spread of news, już ich cenzurowanie jest efektywne (Clark et al., 2022). Problem polega na tym, że jako nietransparentne budzi wiele ważnych wątpliwości, w tym w terminach wolności słowa oraz otwartości nauki na krytyczne pozycje. Rozwiązaniem w pewnym sensie pośrednim pomiędzy flagging a censoring wydaje się być fact-checking, o którym pisać będziemy szczegółowo w kolejnych sekcjach.

Na uwagę szczególną zasługują badania Acemoglu et al. (2021), którzy pokazują, że homophily powstrzymuje szerszą cyrkulację artykułów, ale jednocześnie tworzy echo chambers, w ramach których obowiązują obniżone standardy rzetelności. Co ciekawe, jak wskazują autorzy platformy cyfrowe same wpuszczają news o obniżonej wiarygodności do przestrzeni, w których mamy greater polarization, zaś bardziej neutralne i nie budzące wiadomości news trafiają do szerokiego odbiorcy. W tym sensie o rozprzestrzenianiu się fake newsów nie decyduje wyłącznie homofilia, ale algorytm, który na homofilii bazuje, by zwiększać zasięgi. Oznacza to, że z perspektywy ekonomicznej rozprzestrzenianie się misinformation and disinformation jest zbieżne z interesem wielkich platform cyfrowych, co z kolei otwiera fundamentalnie ważny polityczny w dyskusji.

W kontekście dynamicznego rozwoju sztucznej inteligencji, warto wspomnieć także, że jest ona jednocześnie sprzymierzeńcem i wrogiem walki z rozprzestrzenianiem misinformation and disinformation: z jednej strony może ona wyłapywać podejrzane treści, ale z drugiej strony może też je generować. Jak zauważają Menczer & Yang (2023), state-of-the-art LLM-content detectors cannot effectively distinguish between human and LLM-powered bots in the wild, co wskazuje na ważne niebezpieczeństwa. Oczywiście istotną kwestią wciąż pozostaje kwestia tego, jak w materiale badawczym, na którym sieci neuronowe są uczone, rozróżniono prawdę i fałsz.

Jednocześnie w obszarze prawodawstwa platformom cyfrowym stawia się coraz większe wymagania. Mają one dokładać należytej staranności i uzasadnić środki stosowane do combat disinformation on their services, jak również włączać w swoje działania szereg praktyk służących fact-checkingowi (Miguel, 2023). Tworzy to politycznie kontrowersyjną sytuację, w której pewnego rodzaju cenzura jest wymuszana przez legislatora, ale odpowiedzialność za ograniczanie wolności słowa spada na operatorów przestrzeni komunikacyjnej.

Warto wspomnieć, że obecnie badacze mają do dyspozycji szeroką paletę oprogramowania, które pozwala monitorować the dynamics and structure of the (dis)information flows, w tym między innymi programy: FindExif, Hoaxy, Intelligence X, Maltego, SketchEngine, TraceRoute, Twitter Trails, których funkcjonowanie analizowaliśmy w kontekście naszych badań oraz korzystaliśmy w ograniczonym zakresie dla modelowania jego części empirycznej.

Branża fact-checkingowa: wpływ i bias

Fact-checking odgrywa w dzisiejszych czasach zauważalną i rosnącą rolę. Jak już w 2016 roku zdiagnozowali Cherubini & Graves (2016) na całym świecie powstawać zaczęło mnóstwo agencji fakt-checkingowych. I choć fakty fałszowane były od zawsze i podobnie od zawsze były weryfikowane, XXI wiek z pewnością przyniósł co najmniej zmianę ilościową w tym zakresie. Zalew informacji doprowadził do konieczności szybkiej profesjonalizacji w tej branży (zjawisko pozytywne), ale także uczynił to dość lukratywnym i politycznie wpływowym biznesem (zjawisko potencjalnie niebezpieczne). Jakość i rodzaj komunikacji i obecności fake newsów jest istotnie uwarunkowany geopolitycznie (Graves & Cherubini, 2016). Sądząc po naszym badaniu i zestawiając je z wynikami badań realizowanych w USA, można byłoby zaryzykować stwierdzenie, że także jakość funkcjonowania podmiotów fact-checkingowych jest geopolitycznie uwarunkowana, choć nie jest to przedmiotem naszych analiz.

Badania Cha et al. (2021) pokazują, że fact-checking wpływa na young adults, jednocześnie jednak szereg czynników, jak na przykład negatywne pierwsze wrażenie, ma znaczący wpływ na siłę tego efektu. Jak wskazują badacze, platformy fact-checkingowe analizują wiele newsów, które są niejednoznaczne, a ich weryfikacja jest dyskusyjna. An meta-analytic review realizowany przez Cohen et al. (2020) pokazał, że fact-checking ma duży wpływ na political beliefs, gdy o prawdziwości rozstrzyga się zero-jedynkowo ($d = 0.29$), ale staje się to znacznie mniej jednoznaczne, gdy zastosowane zostają "truth scales". To samo badanie potwierdziło, że kluczowe znaczenie dla percepcji fake newsów i rezultatów ich weryfikacji w istotnym stopniu zależy od preexisting beliefs, ideology, and knowledge. Wiąże się to bezpośrednio z problemem, na który wskazywaliśmy w pierwszej sekcji niniejszego raportu badawczego, a mianowicie brakiem wyraźnych granic definicyjnych w przypadku niektórych typów fake newsów, a także pojawianiem się wielu przypadku pośrednich. Percepcję fake newsów z jednej strony, a raportów fact-checkingowych z drugiej opisywać można z wykorzystaniem mechanizmów psychologicznych. I tak, jak pokazuje Cha et al. (2021), eksperyment z udziałem 11 145 young adults pokazał, że osób tych nie przekonują komunikaty niejednoznaczne, co z kolei wiąże się z szeroko rozpoznanym zjawiskiem, jakim jest

uncertainty-aversion. Jak jednak twierdzą autorzy, fact-checking dobrze sprawdza się przy relaxing extreme views.

W badaniu przeprowadzonym przez Portera i Wooda (2021) badano skuteczność platform fact-checkingowych. Ustalono, że fact-checking znacząco reduced belief in misinformation i utrzymywał się także przez parę tygodni. Autorzy oszacowali ten wpływ na 0.59 points on a 5-point scale. Jak jednak zauważa Tompkins (2020), fact-checking sprawdza się, ale jedynie w dość specyficznych okolicznościach. Powołując się na Dan Kahan, wskazuje, że ludzie korzystają z fałszywej informacji, jeśli potwierdza to ich stanowisko, ale jeśli okazuje się fałszywa, to przestają z niej korzystać, co wcale nie oznacza, że zmieniają swoje stanowisko. Inną kwestią jest właśnie to, czy zmiana stanu wiedzy pociąga za sobą zmianę zachowania, czy preferencji wyborczych. Jak pokazują badania Barrera et al. (2020) wcale tak być nie musi. Według badań Adamica et al. (2015) realizowanych na grupie ponad 10 milionów amerykańskich użytkowników Facebooka, indywidualne wybory użytkowników miały większe znaczenie dla limiting exposure to cross-cutting content niż algorithmic ranking.

Aral et al. (2018) pokazali, że tweety zawierające fałsz were 70% more likely to be retweeted than truthful tweets. Co ciekawe, nie wynikało to, jak wydawało się na początku, z tego, że osoby rozprzestrzeniające fałszywe tweety mają więcej followersów (bo tak nie jest), ale tweety zawierające fałszywą informację były po prostu ciekawsze i more novel, a więc, z socjologicznego punktu widzenia, były bardziej użyteczne w konwersacjach. Podobne badanie Anderson et al. (2021) pokazało, że szerzenie fake newsów na Facebooku jest habitualne, a nie wynika z lenistwa, czy obciążenia, co z kolei stoi po części w sprzeczności z wynikami badań Pennycook & Rand (2019). Warto zwrócić przy tym uwagę, że czymś innym jest szerzenie wiadomości, a czymś innym jest wiara w ich treść.

W wielkim projekcie badawczym "Globalism Project 2021" (Rogers de Waal, 2021) ustalono, że nie wszystkie tak zwane spiskowe teorie dziejów są równie popularne oraz że ich popularność rozkłada się bardzo różnie w zależności od kraju (Kirk, 2022). Co istotne, przykładowo w przypadku kwestii pandemii COVID-19 obserwuje się wielkie rozbieżności w postawie wobec stwierdzenia "coronavirus is a myth created by some powerful forces, and the virus does not really exist" oraz stwierdzenia

"the truth about the harmful effects of vaccines is being deliberately hidden from the public". To także pokazuje, jak istotną rolę dla kształtowania dyskursu wokół misinformation and disinformation odgrywają platformy fact checkingowe oraz sama dyskusja o fake newsach promowana przez prominentne portale. W tym sensie fact-checking ma także wpływ na dynamikę i strukturę trajektorii rozprzestrzeniania się fake newsów. Ważnym, a niedostatecznie zbadanym problemem staje się więc obciążenie samych fact-checking platforms.

Levine et al. (2023) przeprowadzili badanie zbieżności platform fact-checkingowych w USA, koncentrując się na zbieżności i rozbieżności doboru wiadomości do sprawdzenia oraz rezultatów tego sprawdzania pomiędzy The Washington Post and PolitiFact. Jak wskazali, "the selection disagreement was indicated at the level of 22.6% [...] moderate agreement was observed for deceptiveness scaling and nearly complete agreement was observed for bottom-line attributed veracity" (Levine et al., 2023). Badacze wyciągnęli szereg istotnych wniosków, a w szczególności, że efekt fact-checkingu zależy w istotnym stopniu od different fact-checkers oraz algorithmic bias. W dość podobnym badaniu Nieminen & Sankari (2021) analizowali proces fact-checkingowy stosowany przez PolitiFact weryfikując próbkę 858 materiałów. Dwoma problemami zdiagnozowanymi przez badaczy jest jednoznaczna weryfikacja complex propositions (33%) oraz weryfikacja stwierdzeń, których truthfulness cannot be defined in practice, e.g., claims about the future and vague claims (11%) (Nieminen & Sankari, 2021). Podobne badania wcześniej przeprowadzone zostały pomiędzy Fact Checker and Politifact, a w ich wyniku, podobnie, stwierdzono wysoką zbieżność w przypadku tych wiadomości, które były jednoznaczne, ale niemałe rozbieżności w przypadkach niejednoznacznych (Lim, 2018).

Przykładowo, jako side-result naszej analizy wspomnianych 10 typów misinformation and disinformation wyróżnionych przez UNHCR (2021) polskie platformy fact-checkingowe niemal w ogóle nie interesują się sponsored content (advertising or PR), mimo że wydają się one stanowić pokaźny share w ogóle misinformation and disinformation. Można spekulować, że wynika to z obawy o pozwy o zniesławienie i konfrontację z prawnikami wielkich korporacji.

Jak zauważa Fraser, sam fact-checking jest elementem gry politycznej oraz gry interesów (Fraser, 2023), co jest również kwestią pozostającą w centrum naszej uwagi.

Polskie platformy fact-checkingowe: analiza

Badaniu poddane zostały wszystkie procesy fact-checkingowe zrealizowane i opublikowane przez wiodące polskie platformy fact-checkingowe w okresie 08.09.2023-07.10.2023. Poniżej krótka charakterystyka badanych podmiotów:

- oko.press – to serwis informacyjny o tematyce polityczno-społecznej, specjalizujący się w dziennikarstwie śledczym i sprawdzaniu faktów funkcjonujący od roku 2016 w ramach działalności prowadzonych przez Fundację Ośrodek Kontroli Obywatelskiej „OKO”¹; na stronie głównej znajduje się link do podstrony "prawda czy fałsz", która zawiera materiały fact-checkingowe;
- konkret24 – to również serwis informacyjny o tematyce polityczno-społecznej będący wyspecjalizowaną komórką polskiej komercyjnej stacji telewizyjnej TVN należącej od roku 2022 do Warner Bros. Discovery; cały serwis poświęcony jest fact-checkingowi;
- sprawdzimy.info oraz odfejkuj.info (oba linki prowadzą do tej samej strony internetowej) został założony przez Ośrodek Analiz Cegielskiego powstał, aby weryfikować i korygować kłamstwa i manipulacje; portal powstał przy wsparciu Narodowego Instytutu Wolności – Centrum Rozwoju Społeczeństwa Obywatelskiego ze środków Programu Rozwoju Organizacji Obywatelskich na lata 2018–2030; całość strony poświęcona jest weryfikacji fake news;
- fakenews.pl to portal internetowy dotyczący szeroko pojętej tematyki fake news prowadzony przez fundację Przeciwdziałamy Dezinformacji; strona podzielona jest na podstrony odpowiadające kategoriom tematycznym wiadomości, których prawdziwość jest weryfikowana przez zespół redakcyjny; charakterystyczne dla funkcjonowania strony są bardzo nachalne komunikaty z prośbą o wsparcie finansowe;
- demagog.org.pl – strona prowadzona przez Stowarzyszenie Demagog, pierwszą polską organizacją factcheckingową, która od 2014 roku weryfikuje wypowiedzi polityków,

¹ oko means eye in Polish;

sprawdzamy obietnice wyborcze, walczymy z fake newsami i dezinformacją; strona dzieli się na dwie główne sekcje: "wypowiedzi" (polityków) oraz "fakenewsy", które są weryfikowane przez zespół redakcyjny;

- sprawdzam.afp.com – to wyspecjalizowana komórka francuskiej Agence France-Press (AFP); serwis koncentruje się na weryfikacji prawdziwości wiadomości z obszaru społeczno-politycznego;
- pravda.org.pl – serwis prowadzony przez Stowarzyszenie, które działa na rzecz zwiększenia świadomości o dezinformacji poprawiła się jakość debaty publicznej i jest wspierane przez jest wspierany przez Europejski Fundusz Mediów i Informacji; strona zawiera materiały w mocno zróżnicowanym formacie, co sprawiło, że w trakcie badania nie zawsze udawało się ustalić datę pochodzenia danego procesu fact-checkingowego;
- fakehunter – to społeczny projekt weryfikacji treści publikowanych w internecie, uruchomiony przez Polską Agencję Prasową (spółka Skarbu Państwa, której działalność opiera się na ustawie i ma charakter publicznej agencji prasowej) wspólnie z GovTech Polska, którego celem pierwotnym celem było demaskowanie nieprawdziwych wiadomości dotyczących wirusa SARS-CoV-2; obecnie strona zajmuje się weryfikacją zgłoszeń nadsyłanych przez użytkowników.

Z perspektywy badania istotną kwestią jest weryfikacja, czy wzmiankowane serwisy deklarują się jako fact-checkingowe oraz czy deklarują stosowanie fact-checkingowych standardów. Analiza opisów dostępnych na stronach wskazuje, że deklarują się jako serwisy fact-checkingowe i przynajmniej na poziomie deklaracji mają ambicje, by respektować środowiskowe standardy. Ważnym punktem odniesienia w tym zakresie jest kodeks opracowany przez europejską sieć standardów weryfikacji faktów (EFCSN, 2022) jak również, kodeks opracowany w 2016 przez International Fact-Checking Network (IFCN) (Babraj, 2019). Jednocześnie warto zwrócić uwagę, że sformułowania takie jak "Ze zbioru wszystkich wypowiedzi polityków wybieramy te, które są merytoryczne i dotyczą rzeczywistości obiektywnie sprawdzalnej. Dobór próby badawczej jest ściśle jakościowy i oparty o kryterium dostępności materiału.", które możemy znaleźć np. na stronie portalu Demagog mają charakter wysoce spekulatywny i niejednoznaczny.

W ramach przeprowadzonego badania (analizy zamieszczanych raportów z procesów fact-checkingowych) zauważono, że serwisy te przyjmują bardzo różne podejścia (głębsze, płytsze, bardziej rozległe, mniej rozległe, bardziej jednoznaczne, bardziej spekulatywne) oraz różnoraki dobór tematyki wiadomości, które poddawane są weryfikacji. Odnotowano następujące znaczące cechy funkcjonowania wymienionych powyżej platform:

- oko.press: specyfiką serwisu, w odróżnieniu od większości polskich platform factcheckingowych, zdecydowanie jest to, że zamieszczane w nim materiały są raczej (poza wyjątkami) opiniami i polemiką z pewnymi stanowiskami, nie zaś weryfikacją prawdziwości jakichś konkretnych stwierdzeń; w tym sensie same procedury fact-checkingowe stosowane przez oko.press mają charakter spekulacyjny; zwraca uwagę także fakt, że dobór treści, wobec który przedstawiana jest polemika, ma skrzywienie prawicowe i w sensie samego doboru treści portal jest nieobiektywny; jednocześnie waga dyskutowanych tematów jest znaczna i o szerokim zasięgu; częstotliwość (w sekcji poświęconej weryfikacji prawdziwości): 13 / miesiąc (niska)
- konkret24: specyfiką serwisu jest znaczny udział materiałów o charakterze spekulacyjnym, a częstym typem weryfikowanych wiadomości jest manipulacja ("misleading content"), a więc kwestie same w sobie spekulatywne; dobór tematyki jest wyraźnie polityczny (innego typu dezinformacja jest wyraźnie marginalizowana), przy czym rażąco jest proliberalny / antyprawicowy profil serwisu: wśród wszystkich zweryfikowanych postępowań fact-checkingowych przeszło 36% miało wydźwięk antyliberalny i w 84% tych przypadków werdyktem była ich fałszywość; jednocześnie, podobnie, 36% miało wydźwięk antyprawicowy, a fałszywość orzeczona została w 5% przypadków; oczywiście redakcja serwisu może argumentować, że źródłem takiego stanu rzeczy jest częstsza dezinformacja w przekazie prawicowym, jednak nie wydaje się to korespondować z rzeczywistością, biorąc pod uwagę ustalenia wszystkich innych serwisów; częstotliwość: 53 / miesiąc (znaczna)
- sprawdzimy.info oraz odfejkuj.info: w analizowanym okresie serwis nie zamieścił ani jednego wyniku weryfikacji fact-checkingowej; częstotliwość 0 / miesiąc (portal prawdopodobnie zawiesił działalność)

- fakenews.pl: dobór materiałów jest zróżnicowany, pochodzi z różnych obszarów tematycznych i nie ma jednoznacznej orientacji politycznej, co jednak jest trudne do stwierdzenia ze względu na małą próbę, bowiem portal działa z bardzo niską intensywnością; zwraca uwagę stosunkowo jednoznaczne formułowanie weryfikowanych treści, co kolei stwarza możliwość weryfikacji zero-jedynkowej; jednocześnie dobór materiałów na tym portalu wydaje się kontrowersyjny na inny sposób: większość z nich wydaje się mało znacząca, co z kolei można byłoby przypisać stronięcia od tematów dużych i trudnych; częstotliwość: 8 / miesiąc (bardzo niska)
- demagog.org.pl: serwis przedstawia procesy fact-checkingowe w dwóch kategoriach: (fake)newsy oraz wypowiedzi; w przypadku sekcji wypowiedzi przedstawiane są wypowiedzi polityków, które zwykle mają charakter dość jednoznaczny, co z kolei pozwala na jednoznaczną ich weryfikację i takiej też, bardzo rzetelnej w porównaniu z innymi serwisami, weryfikacji dostarcza serwis Demagog; jednocześnie w tej kategorii portal nie unika tematów ważnych i trudnych; zwraca uwagę zbalansowanie polityczne doboru wiadomości, które są weryfikowane (wśród 35 wypowiedzi znalazło się: 26% wypowiedzi polityków centro-liberalnych, 31% prawicowych, 31% lewicowych; 11% niezależnych), natomiast kontrowersje budzić może ocena prawdziwości dobranych wypowiedzi (11% fałszu w wypowiedziach polityków centro-liberalnych, 36% fałszu w wypowiedziach polityków prawicowych; 18% fałszu w wypowiedziach polityków lewicowych i 25% fałszu w wypowiedziach osób niezależnych) (tu warto zwrócić uwagę, że dopełnieniem zbioru odpowiedzi jednoznacznie fałszywych niekoniecznie są odpowiedzi prawdziwe, bo portal stosuje także inne kategorie, jak "brak rozstrzygnięcia" czy "częściowy fałsz"); gorsze wrażenie tworzy druga kategoria prezentowana przez serwis dotycząca (fake)newsów, gdzie dominuje kilka specyficznych tematów w tym przede wszystkim kwestia (nie)szkodliwości technologii 5G, a także kwestia szczepień; dodatkowo istotna część weryfikowanych wiadomości pochodzi z jednego konkretnego portalu, co nosi znamiona walki ideowej; wiele weryfikowanych wiadomości wydaje się przy tym zupełnie nieznacząca; najważniejszym jednak zastrzeżeniem jest wzmiankowany już wcześniej problem z doбором metody weryfikacji, a mianowicie

wiadomości, w których kwestionuje się ustalenia naukowego mainstreamu weryfikuje się w wielu przypadkach poprzez odwoływanie się do ustaleń tegoż mainstreamu; częstotliwość: 97 / miesiąc (bardzo wysoka)

- sprawdzam.afp.com – serwis charakteryzuje się rzetelnością w zakresie weryfikowania przedstawianych wiadomości, bazując na dobrze uargumentowanych źródłach, a także wykorzystując analizy technologiczne (na niespotykaną na innych platformach skalę: 47%); weryfikowane wiadomości mają dość jednoznaczny charakter, co pozwala formułować jednoznaczne wyniki prowadzonych śledztw; portal raczej nie podejmuje tematów politycznych w sensie partyjnym; jednocześnie warto zwrócić uwagę, że dobór weryfikowanych treści jest problematyczny, bowiem są one w wielu wypadkach nieznaczące (pojawia się pytanie: po co je weryfikować, skoro wiele ważniejszych nie jest poddawana weryfikacji), ale też dlatego, że rysuje się wyraźna dominacja wiadomości o tematyce antyukraińskiej, które serwis weryfikuje; dyskusyjna jest wreszcie decyzja serwisu, by zamieszczać wyłącznie te treści, które zostały zweryfikowane jako fałsz; częstotliwość: 15 / miesiąc (niska)
- pravda.org.pl – w badanym okresie serwis weryfikował wyłącznie treści antyukraińskie i robił to w sposób niedokładny i dyskusyjny; trudno wyciągać jednak szersze wnioski, biorąc pod uwagę bardzo niską aktywność redakcji serwisu; częstotliwość: 4 / miesiąc (bardzo niska)
- fakehunter – w odróżnieniu od innych serwisów fakehunter weryfikuje treści nadsyłane do redakcji z prośbą o weryfikację przez użytkowników; ponieważ trudno określić, jak (i czy w ogóle) redakcja dodatkowo filtruje te zapytania, bazując na zasadzie domniemania niewinności, należy uznać, że fakehunter dobiera treści w sposób rzetelny; same zapytania formułowane przez użytkowników mają bardzo zróżnicowany charakter, przy czym warto zwrócić uwagę, że znaczna ich część jest mało znacząca, a niekiedy wręcz banalna; w zdecydowanej większości przypadków redakcja portalu przedstawia rzetelną argumentację przy weryfikowaniu zapytań użytkowników; częstotliwość: 78 / miesiąc (wysoka).

Znaczące obserwacje ogólne:



- w całej próbie 268 przeanalizowanych wiadomości jedynie niecałe 11% zakwalifikowano jako jednoznacznie prawdziwe; w tym kontekście pojawia się kluczowe pytanie, czy: (a) jest to zjawisko strukturalne i jedynie 11% wiadomości w publicznym obiegu jest jednoznacznie prawdziwa; (b) z jakichś powodów serwisy fact-checkingowe dobierają raczej wiadomości nieprawdziwe; (c) z jakichś powodów serwisy fact-checkingowe proces weryfikacji jest asymetrycznie przechylony w stronę podważania prawdziwości analizowanych treści;
- wśród analizowanych przez polskie portale fact-checkingowe treści odnotowano śladową (<1%) ilość misinformation oraz disinformation z kategorii "sponsored content (advertising or PR)" (według UNHCR); podobnie pojawiają się pytania, czy: (a) jest to zjawisko strukturalne i wiadomości tego typu nie pojawiają się w publicznym obiegu; (b) z jakichś powodów serwisy fact-checkingowe nie analizują tego typu treści;
- niektóre serwisy, jak Demagog, analizują także żarty ("satire and parody" w klasyfikacji UNHCR), poniekąd słusznie wskazując, że również przyczyniają się one do dezinformacji; co interesujące, ich udział w całości próby stanowi około 2%, co wydaje się wartością niezwykle niską, biorąc pod uwagę, jak wielka ilość różnego rodzaju memów zalewa Internet; innymi słowy, nie wydaje się, by materiały humorystyczne i sarkastyczne miały jedynie 2% udziału w informacji, co z kolei na inny sposób ujawnia tendencyjność w doborze kategorii;
- już sama analiza najpopularniejszych tematów podejmowanych przez serwisy fact-checkingowe w 2022 roku zrealizowana przez Centrum Edukacji Medialnej (CEM, 2023) pokazała, że dobór tych wiadomości nosi znamiona tendencyjnego, w szczególności wśród 20 najpopularniejszych tweetów wyraźnie uwidoczniły się trzy tematy: bieżącej polskiej polityki, COVID i spraw związanych z uchodźcami z Ukrainy; co ciekawe większość z tych publikacji pochodziła z trzech serwisów: Demagog, Konkret24 i FakeNews.pl (CEM, 2023); ponownie pojawia się pytanie, czy wynika to z czynników strukturalnych, czy właśnie z pewnego rodzaju świadomej asymetrii doboru;
- co do zasady analizowane serwisy nie czynią różnicy pomiędzy różnymi rodzajami misinformation and disinformation (zgodnie z przyjętą klasyfikacją UNHCR), a więc w szczególności: (a) weryfikują wiadomości z szeroko pojętej kategorii spiskowej teorii dziejów

odwołując się do oficjalnych komunikatów instytucji publicznych; (b) weryfikują podważanie powszechnie przyjmowanych ustaleń współczesnej nauki poprzez odwoływanie się do tychże powszechnie przyjmowanych ustaleń współczesnej nauki;

- co interesujące, jedynie około 7% stanowiły wiadomości weryfikowane poprzez analizę techniczną (fotomontaż, deep fake, etc.)
- w powszechnej debacie na temat dezinformacji często mówi się o wpływie rosyjskiej dezinformacji cyfrowej; przebadany materiał nie daje podstaw do potwierdzenia takiego stanu rzeczy, bowiem w marginalnej liczbie przypadków dziennikarze śledczy wskazywali na rosyjskie pochodzenie wiadomości zweryfikowanych jako fałszywe
- choć trudno ocenić polityczną lub światopoglądową skalę wagi danej wiadomości, wydaje się, że platformy fact-checkingowe przeplatają weryfikację wiadomości zupełnie banalnych z wiadomościami o wielkim ładunku politycznym; niejednokrotnie weryfikowane są wiadomości nieznaczące, a z kolei stwierdzenia bardzo znaczące w ogóle weryfikacji nie podlegają; oczywiście, można tłumaczyć to trudnością z weryfikacją, jednak werdykt "nierozstrzygnięte" lub "częściowa prawda" pojawia się w analizowanych raportach, jak widać rozważanie trudnych w weryfikacji wiadomości jest praktykowane; pojawia się pytanie, o intencjonalność takiego działania.

Wskazywane powyżej cechy praktyk fact-checkingowych mają istotny wpływ na dynamics and structure of the (dis)information flows zarówno w sensie technicznym, jak i ontologicznym.

Krótkoterminowe efekty zwalczania fake-newsów

Do badania wybrano trzy wiadomości, które zostały zakwalifikowane jako fakenewsy przez wybrane portale fact-checkingowe:

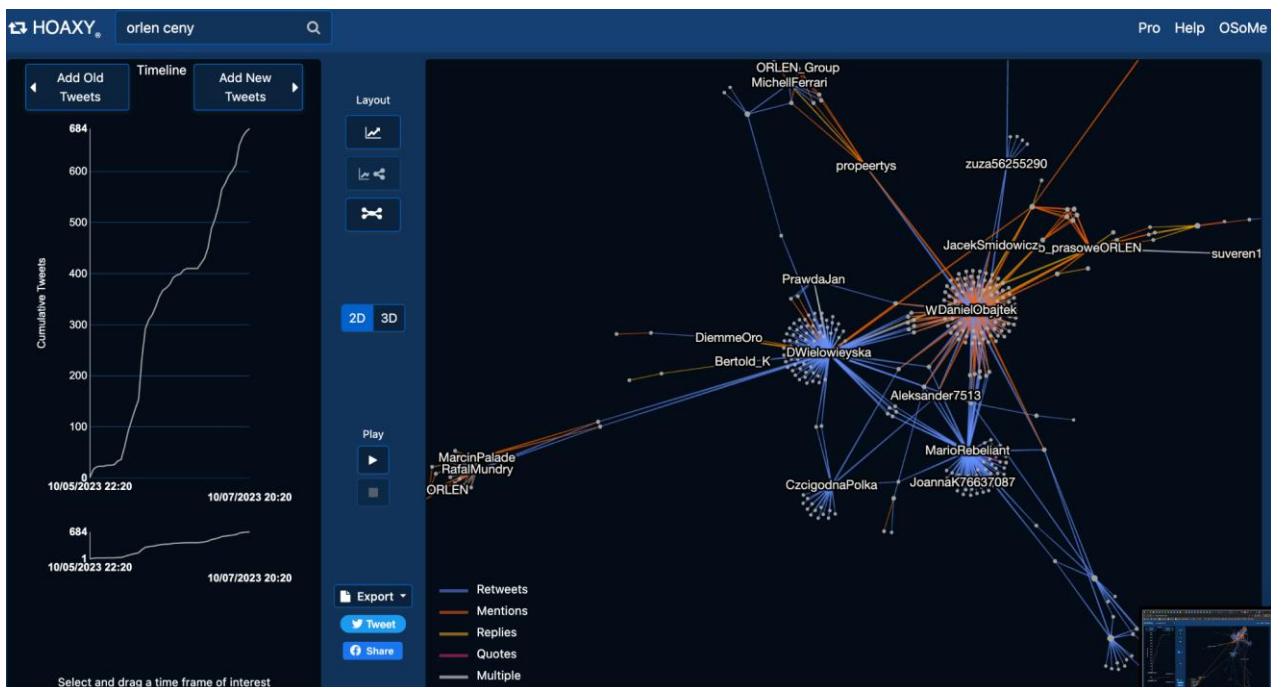
- [A] stwierdzenie, że tania benzyna na stacjach w Polsce nie jest zjawiskiem nowym (i specyficznym dla okresu przedwyborczego) [w powiązaniu z polityką grupy Orlen]
- [B] stwierdzenie, że ceny w Polsce od dawna nie rosną [w powiązaniu z polityką Narodowego Banku Polskiego]

- [C] stwierdzenie, że w Polsce podano wadliwe szczepionki przeciwko COVID-19 [w powiązaniu z raportem Najwyższej Izby Kontroli]

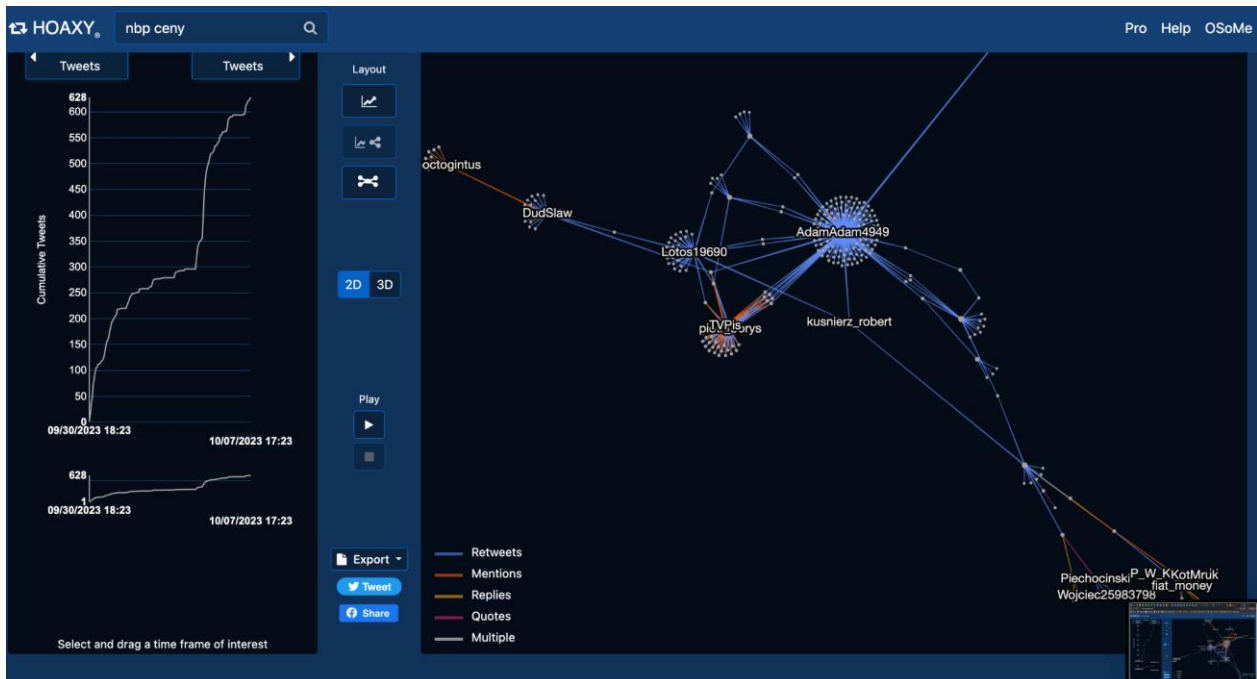
Wiedząc (na bazie przeprowadzonej analizy literatury oraz refleksji teoretycznej), że ideologiczne tło ma kluczowe znaczenie dla recepcji dyskusyjnych treści, z wykorzystaniem oprogramowania SketchEngine przeprowadzono badanie korpusu językowego treści internetowych powiązanych z każdym z trzech zagadnień. W przypadkach [A] i [B] stwierdzono istnienie zróżnicowanej palety komentarzy prasowych w polskiej części Internetu, przy czym komentarze pojawiające się na platformach wiodących portali informacyjnych (businessinsider, money.pl, tvn.24, polsat.news) miały pejoratywny wydźwięk. W przypadku tematu [B] zaobserwowano pewne wyciszenie aktywności informacyjnej po prawej stronie sceny dziennikarskiej. W przypadku [C] zauważono pewną powściągliwość w propagowaniu informacji ogłoszonych przez Najwyższą Izbę Kontroli, choć wiele ważnych portali podało do publicznej wiadomości te ustalenia w stosunkowo neutralnym tonie, w szczególności bez sugestii o potencjalnej dezinformacji zawartej w komunikacie organu kontrolnego. Wiadomość [C] szeroko nagłaśniana była przez portale informacyjne powiązane z formacjami politycznymi zorientowanymi antyszczepionkowo.

Analiza przy pomocy oprogramowania Hoaxy pokazała następującą dynamikę flow informacyjnego wygenerowanego wokół każdego z trzech wzmiankowanych przypadków wiadomości:

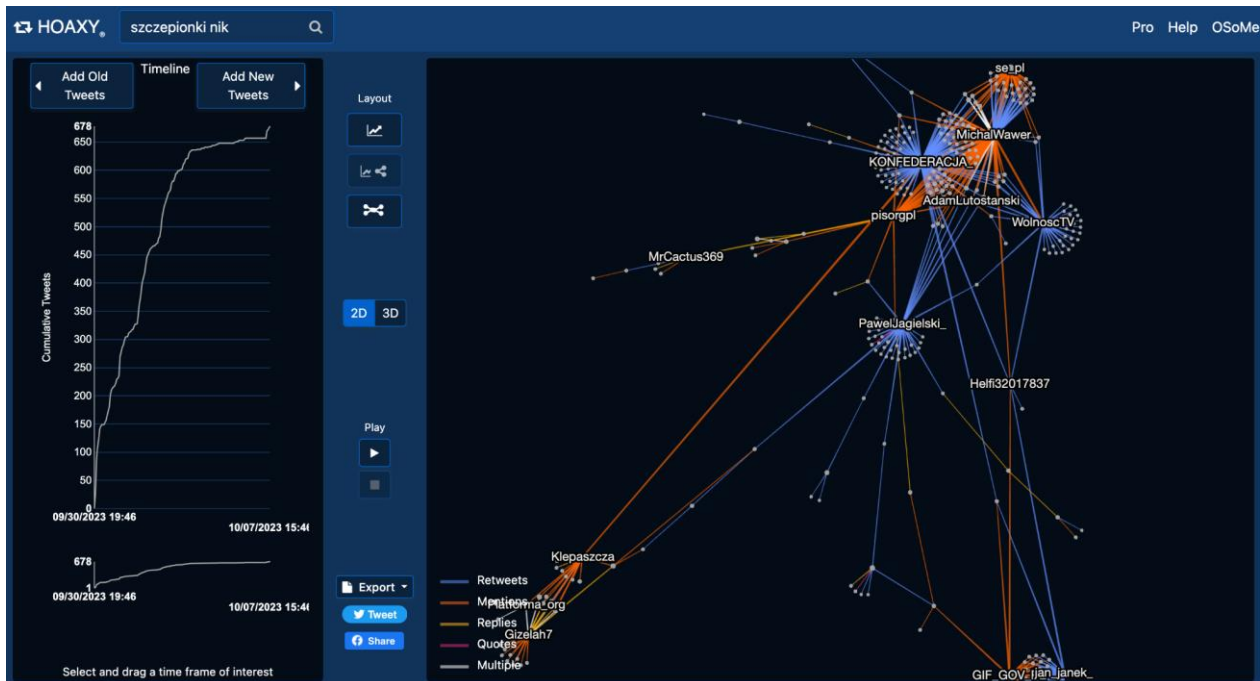
[A] Dynamika rozprzestrzeniania się wiadomości miała dwa wzmożenia, przy czym jej wygasanie i wzmaganie przebiegało równoległe wśród kont wspierających dyskutowane stwierdzenie, jak i je kwestionujących. Nie odnaleziono śladów zastosowania cenzury. Nie zauważono wpływu weryfikacji wiadomości przez platformy fact-checkingowe. Wydaje się, że o dynamice rozprzestrzeniania się wiadomości zadecydowała wyłącznie homofilia oraz tło ideologiczne. Dodatkowe parametry charakteryzujące strukturę: 684 tweetów wygenerowanych przez 401 kont w okresie 2023-10-07T17:31:34.000Z – 2023-10-05T20:20:23.000Z, w tym: 398 retweets, 204 mentions, 66 replies, 16 quotes.



[B] Bardzo podobnie wyglądała dynamika rozprzestrzeniania się wiadomości [B]: miała ona dwa wzmożenia, przy czym jej wygasanie i wzmaganie przebiegało równoległe wśród kont wspierających dyskutowane stwierdzenie, jak i je kwestionujących. Nie odnaleziono śladów zastosowania cenzury. Nie zauważono wpływu weryfikacji wiadomości przez platformy fact-checkingowe. Wydaje się, że o dynamice rozprzestrzeniania się wiadomości zdecydowała wyłącznie homofilia oraz tło ideologiczne. Dodatkowe parametry charakteryzujące strukturę: 629 tweetów wygenerowanych przez 410 kont w okresie 2023-10-07T14:59:51.000Z – 2023-09-30T16:23:52.000Z, w tym: 306 retweets, 167 mentions, 145 replies, 10 quotes.



[C] W przypadku dynamiki rozwoju tweetów dotyczących raportu Najwyższej Izby Kontroli różnica w dynamice sprowadza się wyłącznie do tego, że przyrost liczby tweetów nie ma spowolnienia w połowie okresu (jak ma to miejsce w przypadku wiadomości [A] i [B]); wzmaganie przebiegało równoległe wśród kont wspierających dyskutowane stwierdzenie, jak i je kwestionujących. Nie odnaleziono śladów zastosowania cenzury. Nie zauważono wpływu weryfikacji wiadomości przez platformy fact-checkingowe. Ponownie, wydaje się, że o dynamice rozprzestrzeniania się wiadomości zadecydowała wyłącznie homofilia oraz tło ideologiczne. Dodatkowe parametry charakteryzujące strukturę: 650 tweetów wygenerowanych przez 341 kont w okresie 2023-10-07T13:14:41.000Z – 2023-09-30T17:46:23.000Z, w tym: 368 retweets, 264 mentions, 43 replies, 3 quotes.



[6] Conclusions

Podstawowym wnioskiem z badania jest dyskusyjność i polityczno-światopoglądowe uwarunkowanie dyskursu na temat misinformation and disinformation, w tym tak zwanych fake-newsów. Choć bardzo powszechnie mówi się w dzisiejszych czasach o olbrzymiej skali tego zjawiska, określenie w praktyce, co do niej należy okazuje się być dalece problematyczne. Wydaje się, że istotnym głosem w tej dyskusji powinien być głos, powołanych w celu formułowania eksperckich wglądów, platform fact-checkingowych. Zarówno jednak literatura omawiająca to zagadnienie w skali globalnej, jak i zrealizowane badanie empiryczne polskiej części Internetu, wskazują na ważne ograniczenia takiej perspektywy. Po pierwsze, jak wynika z badań Digital Poland (2022) mało (5%) użytkowników korzysta ze wskazań platform fact-checkingowych, po drugie rzetelność funkcjonowania polskich platform tego typu pozostawia bardzo wiele do życzenia, co jest bezpośrednim wnioskiem z niniejszego badania, po trzecie zaś (co być może należy postrzegać jako konsekwencję dwóch pierwszych punktów) analiza wybranych case studies pokazała, że wpływ tychże dyskusyjnych platform na rozprzestrzenianie się wiadomości jest niezauważalny.

Jeśli obszar misinformation and disinformation w polskim Internecie określać zgodnie ze wskazaniem platform fact-checkingowych, to należałoby uznać, że jest to zjawisko marginalne.



Jednak w praktyce większość użytkowników działa inaczej i używa innych metod do rozstrzygnięcia o prawdziwości i nieprawdziwości rozmaitych treści. Wśród tych na pewno warto wspomnieć techniki takie jak samodzielna weryfikacja, opinie internetowych influencerów, czy innego rodzaju treści publicystyczne, jak na przykład programy telewizyjne typu "W tyle wizji" realizowane przez formacje prawicowe. Jest to o tyle znaczące, że nie odnotowujemy na polskiej scenie fact-checkingowej platform zorientowanych prawicowo i deklarujących się nominalnie jako fact-checkingowe. Jest to poniekąd zrozumiałe w świetle ideałów Oświecenia.

Zgodnie z przedstawionymi analizami, stosowane obecnie nowoczesne modele ilustrujące the dynamics and structure of the (dis)information flows wskazują na kluczowe znaczenie tła ideologicznego oraz homofilii (wśród czynników statycznych), wysoką skuteczność cenzurowania treści oraz ograniczoną skuteczność fact-checkingu. Wydaje się prowadzić to do napięć ideologiczno-światopoglądowych przy weryfikacji prawdziwości treści typu fabricated content, satire and parody, propaganda, a w konsekwencji do nadużywania autorytetu nauki jako takiej, szkodząc jej wizerunkowi.

Cenzura, także jako ambiwalentna moralnie, okazuje się skutecznym narzędziem tłumienia misinformation and disinformation w danej przestrzeni cyfrowej, co nie znaczy, że dane przekonania zanikają. Z drugiej strony warto zauważyć, że collective intelligence works better when mechanisms encourage confrontation between distinct viewpoints (Greenstein et al., 2020).

Jak pokazują najnowsze badania może być to obcinanie zasięgów (Jackson et al., 2022), co od jakiegoś czasu stosuje Whatsapp. Ma to o tyle istotne znaczenie, że jest to miejsce, gdzie zaczyna manipulować się czymś, co z istoty chcielibyśmy by było wolne. To także miejsce, gdzie kwestie finansowe (bo zasięgi to zyski) zaczynają przenikać się z kontrolą dyskursu. Zarazem wydaje się, że powstrzymywanie dyskusji może mieć wręcz skutek katalizujący rozprzestrzenianie się fake newsów w bubble osób, które zostały z danego dyskursu wykluczone. Co ciekawe, cenzura niejednokrotnie jest wynikiem nacisków politycznych (cf. Money, 2023), kiedy politycy określonych formacji próbują scedować na biznes wymuszanie pewnych postaw w sferze komunikacyjnej. Uwagę zwraca niejednokrotnie zapalczywość, z jaką alternatywny dyskurs jest tłumiony, czego wiele przykładów można było zaobserwować podczas pandemii COVID-19.

Mimo significant dangers związanych z traktowaniem fact-checkingu jako źródła niepodważalnych faktów sygnalizowanych przez Ireton & Posetti (2018), obecna debata polityczna wokół fact-checkingu skręciła w kierunku cenzurowania. Wydaje się, że w dużej mierze jest to konsekwencją okresu pandemii COVID-19, podczas której próbowano w irracjonalnie emocjonalny sposób całkowicie usunąć z przestrzeni publicznego dyskursu wszelkie treści mogące poddawać w wątpliwość kształt przyjętych polityk anty-epidemicznych.

W świetle powyższych rezultatów niepokojem napawać może postawa ankietowanych wyrażona w badaniu Digital Poland (2022). Za dobrą metody walki z dezinformacją 41% zdecydowanie uznało "usuwanie lub blokowanie kont w mediach społecznościowych osób / organizacji najczęściej publikujących fałszywe informacje (fake news)", a kolejne 30% raczej się z tym zgodziło, zaś 69% za dobre wyjście uznało "automatyczne kasowanie fałszywej informacji przez gigantów internetu typu facebook, instagram, google po 1 godzinie od publikacji w mediach społecznościowych". Wyniki te wskazują na wysoki poziom nieświadomości przemocy strukturalnej i jej konsekwencji.

Badanie funkcjonowania polskich platform fact-checkingowych wskazało na szereg rozbieżności pomiędzy stosowanymi przez te platformy praktykami a standardami obowiązującymi w branży. Obraz polskiego fact-checkingu wyłaniający się z badania wskazuje na polityczne i światopoglądowe zaangażowanie redakcji tych serwisów dalece odbiegające od oczekiwań co do ich obiektywności. Ujawnia się ono na kilku płaszczyznach, w tym: w samym doborze weryfikowanych treści pod kątem politycznym, w wadze weryfikowanych treści, w pomijaniu szeregu znaczących tematów, w metodach stosowanych celem weryfikacji. Można spekulować co do intencjonalności tych działań, które mogą być nieświadomione, ale mogą też mieć na celu bądź komercyjne zwiększenie atrakcyjności serwisów, bądź realizowanie określonej agendy ideologicznej. W tym sensie słuszność przyjętej hipotezy H2.1 zależy od rozstrzygnięcia kwestii intencji. Bez popadania w psychologizm uznać jednak można, że aktywność polskich platform fact-checkingowych wskazuje na ich światopoglądową nieobiektywność.

Hipoteza H2.2 została potwierdzona w badaniu, bowiem wszystkie typy wiadomości weryfikowane były przez wiodące serwisy tymi samymi metodami, a więc w szczególności poprzez odwołanie się do mainstreamowego przekazu oraz oficjalnych stanowisk instytucji, nawet w okolicznościach, w

których dane wiadomości wiarygodność tegoż mainstreamowego przekazu kwestionowały. Zarazem znane są przypadki sytuacji, w których mainstreamowy przekaz okazywał się całkowicie mylny (Healio, 2009). Weryfikowanie wiarygodności danego źródła poprzez odwołanie do tego źródła można do „kup[na] kilku egzemplarzy dzisiejszej gazety po to, by się upewnić, że pisze ona prawdę” (Wittgenstein 1953, § 265).

Wreszcie analiza dynamiki trzech wybranych wiadomości rozpoznanych jako fake news przez polskie platformy fact-checkingowe wskazuje na brak związku fact-checkingu z dynamiką rozprzestrzeniania się tychże wiadomości (H2.3).

Wydaje się, że owa tendencyjność fact-checkingu jest rozpoznawana przez użytkowników sieci i traktowana jako narzędzie ideologicznego wpływu w ramach wojen dyskursywnych. Weryfikowanie wiadomości banalnych i banalnie błędnych z wiadomościami o wielkiej złożoności i znacznej wadze, przy użyciu analogicznych metod, dodatkowo zaciemnia obraz fact-checkingu. W związku z owym niewiarygodnym wizerunkiem i podejrzeniami o instrumentalne wykorzystanie, zwłaszcza osoby reprezentujące odmienne nastawienie światopoglądowe i/lub racjonalności nie są skłonne do przyjmowania wyników śledztw dotyczących dezinformacji za istotny punkt odniesienia. Ogólny poziom zaufania do platform fact-checkingowych w Polsce należy uznać za niski. Wydaje się także, że pośrednio ów brak zaufania udziela się także nauce jako takiej, na autorytet której platformy fact-checkingowe bardzo często się powołują. Ważnym pytaniem jest wpływ szerokiego upowszechnienia modeli bazujących na sieciach samouczących się (AI) na ten sektor oraz kondycję debaty o prawdzie. Oczywiście, ze względu na ograniczony zasięg wielu kwestii niniejszy raport wyjaśnić nie był w stanie, stanowi on natomiast krytyczny punkt wyjścia do dalszych badań.

Literatura

Acemoglu, D., Ozdaglar, A., Siderius, J. (2021) *A Model of Online Misinformation*, NBER Working Paper No. 28884; nber.org/papers/w28884 (access: 12.10.2023)

- Adamica, L.A., Bakshy, E., Messing, S. (2015), *Exposure to ideologically diverse news and opinion on Facebook*, Science, Vol 348, Issue 6239, pp. 1130-1132, doi: 10.1126/science.aaa116
- Africano Franco, D.R., Barragán Moreno, S., Carrillo Barbosa, R.L., Guzmán Rincón, A., Rodríguez-Canovas, B. (2023) *Social networks, disinformation and diplomacy: a dynamic model for a current problem*, Humanit Soc Sci Commun 10, 505, doi.org/10.1057/s41599-023-01998-z
- Aldini, A. (2022), *On the modeling and verification of the spread of fake news, algebraically*, Journal of Logic and Computation, Volume 32, Issue 6, p. 1272–1291, doi.org/10.1093/logcom/exac015
- Alwan, A., Garcia, E.P., Kirakosian, A.T., Weiss, A.P. (2021) *Toward a Comprehensive Model of Fake News: A New Approach to Examine the Creation and Sharing of False Information*, Societies 11, no. 3: 82., doi.org/10.3390/soc11030082
- Anderson, I.A., Ceylana, G., Wood, W. (2021), PNAS 2023, *Sharing of misinformation is habitual, not just lazy or biased*, Vol. 120 No. 4
- Apostolidis, L., Kompatsiaris, I., Makedas, T., Papadopoulos, S., Papadopoulou, O., Poldi, F. (2022) *MeVer NetworkX: Network Analysis and Visualization for Tracing Disinformation*, Future Internet 14, no. 5: 147. doi.org/10.3390/fi14050147
- Aral, S., Roy, D., Vosoughi, S. (2018) *The spread of true and false news online*, Vol 359, Issue 6380, pp. 1146-1151, DOI: 10.1126/science.aap95
- Aramaki, E., Kobayashi, R., Murayama, T., Wakamiya, S. (2021) *Modeling the spread of fake news on Twitter*, PLoS ONE 16(4): e0250419. doi.org/10.1371/journal.pone.0250419
- Arini, H.M., Liu, X., Mohawesh, R., Wu, Y., Yin, H. (2023), *Semantic graph based topic modelling framework for multilingual fake news detection*, AI Open 4 (2023) 33–41
- Babraj, R. (2019), *Czym jest fact-checking? – zarys inicjatyw na świecie i w Polsce*, Nask Cyber Policy, bit.ly/3tvRwbw (access: 12.10.2023)
- Baldassarri, D., Bearman, P. (2007), *Dynamics of Political Polarization*, American Sociological Review, Volume 72, Issue 5, https://doi.org/10.1177/000312240707200507

Barrera, O., Guriev, S., Henry, E., Zhuravskaya, E. (2020) *Facts, alternative facts, and fact checking in times of post-truth politics*, Journal of Public Economics, Volume 182, February 2020, 104123

Batistela, C.M., Piqueira, J.R.C., Zilbovicius, M. (2020), *Daley-Kendal models in fake-news scenario*, Physica A: Statistical Mechanics and its Applications, Volume 548, article id. 123406

Brady, W.J., Crockett, M.J., Van Bavel, J.J. (2020) *The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online*, Perspectives on Psychological Science, 15(4), 978-1010. doi.org/10.1177/1745691620917336

Brody, D., Hughston, L., Macrina, A., (2022), *Modelling Information Flows in Financial Markets*, [in:] Brody, D., Hughston, L., Macrina, A., (2022) *Financial Informatics: An Information-Based Approach to Asset Pricing*, World Scientific Publishing

CEM (2023) *Polityka i Ukraina: o czym pisały polskie serwisy fact-checkingowe w 2022 roku?*, Centrum Edukacji Medialnej, bit.ly/48WqOJ2 (access: 12.10.2023)

Centola, D., Macy, M., Willer, R. (2005), *The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms*, American Journal of Sociology, Volume 110, Number 4

Cha, M., Kang, J.-H., Park, J.Y., Park, S. (2021), *The presence of unexpected biases in online fact-checking*, Harvard Kennedy School Misinformation, January 2021, Volume 2, Issue 1, DOI: <https://doi.org/10.37016/mr-2020-53>

Cherubini, F., Graves, L. (2016) *The Rise of Fact-Checking Sites in Europe*, Reuters Institute for the Study of Journalism

Clark, N., Ko, S., Sai Z. (2022) *Modeling the Spread of Misinformation on Social Media*, YT: [youtube.com/watch?v=8b258RJR1TM](https://www.youtube.com/watch?v=8b258RJR1TM) (access, 12.10.2023)

Cohen, J., Holbert, L., Morag, Y., Walter, N., (2020) *Fact-Checking: A Meta-Analysis of What Works and for Whom*, Political Communication, 37:3, 350-375, DOI: 10.1080/10584609.2019.1668894

Cowen, L., de Ruiter, J.-P., Rabb, N., Scheutz, M. (2022) *Cognitive cascades: How to model (and potentially counter) the spread of fake news*, PLoS ONE 17(1): e0261811, doi.org/10.1371/journal.pone.0261811

- Digital Poland (2022), *Dezinformacja oczami Polaków*, Fundacja Digital Poland pod redakcją P. Mieczkowskiego
- EFCSN (2022) *The European Code of Standards for Independent Fact-Checking Organisations*, European Fact-Checking Standards Network, efcsn.com/code-of-standards (access: 12.10.2023)
- Franceschi, J., Pareschi, L. (2022), *Spreading of fake news, competence and learning: kinetic modelling and numerical approximation*, Phil. Trans. R. Soc. A.3802021015920210159, doi.org/10.1098/rsta.2021.0159
- Franceschi, J., Pareschi, L., Zanella, M.. (2022), *From agent-based models to the macroscopic description of fake-news spread: the role of competence in data-driven applications*, SN Partial Differ Equ Appl. 2022;3(6):68. doi: 10.1007/s42985-022-00194-z
- Fraser, M. (2023) *Czy Facebook wykorzystuje fact-checking do własnej gry politycznej?*, CyberDefense24.pl, bit.ly/3tqgpVO (access: 12.10.2023)
- Ghorbani, A.A., Zhang, X. (2020) *An overview of online fake news: characterization, detection, and discussion*, Inf. Process. Manag. 57, 102025
- Granovetter, M. (1978) *Threshold Models of Collective Behavior*, American Journal of Sociology, 83(6), 1420–1443
- Greenstein, Sh., Gu, G., Zhu, F. (2020) *Ideology and Composition Among an Online Crowd: Evidence from Wikipedians*, Management Science 67(5):3067-3086, doi.org/10.1287/mnsc.2020.3661
- Hahn, K.S., Iyengar, S. (2009), *Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use* Get access Arrow, Journal of Communication, Volume 59, Issue 1, March 2009, Pages 19–39, doi.org/10.1111/j.1460-2466.2008.01402.x
- Havlin, Sh., Levy, O., Li, D., Sano, Y., Takayasu, H., Takayasu, M., Wu, J., Zhao, J., Zhao, Z. (2020) *Fake news propagates differently from real news even at early stages of spreading*, EPJ Data Sci. 9, 7, doi.org/10.1140/epjds/s13688-020-00224-z
- Healio (2009), *Cigarettes were once ‘physician’ tested, approved*, Hemonc Today, Issue: March 10, 2009, bit.ly/3S6ucLw (access: 12.10.2023)

- Ireton, Ch., Posetti, J. (eds) (2018), *Journalism, 'fake News' & Disinformation. Handbook for Journalism Education and Training*, United Nations Educational, Scientific and Cultural Organization
- Jackson, M.O., Malladi, S., McAdams, D. (2022) *Learning through the grapevine and the impact of the breadth and depth of social networks*, August 15, 2022, 119 (34), doi.org/10.1073/pnas.2205549119
- Kessler G. (2021), *Trump made 30,573 false or misleading claims as president. Nearly half came in his final year*, The Washington Post., bit.ly/46PQcyz (access: 12.10.2023)
- Kirk, I. (2022) *What conspiracy theories did people around the world believe in 2021?*, YouGov, bit.ly/46vrovW (access: 12.10.2023)
- Kopp, C., Korb, K.B., Mills, B.I. (2018) *Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to "fake news"*, PLoS ONE 13(11): e0207383. <https://doi.org/10.1371/journal.pone.0207383>
- Kuhn, Th.S. (2012) *The Structure of Scientific Revolutions*, 50th ed. Chicago, IL: University of Chicago Press
- Levine, T.R., Markowitz, D.M., Moore, A.D., Serota, K.B. (2023) *Cross-checking journalistic fact-checkers: The role of sampling and scaling in interpreting false and misleading statements*, PLoS One. 2023; 18(7): e0289004, doi: 10.1371/journal.pone.0289004
- Lim, Ch. (2018), *Checking how fact-checkers check*, doi.org/10.1177/20531680187868
- Liming W., Pengwei Z., Yang W., Yunjian Z., Zhen X. (2021) *Multimodal Fusion with Co-Attention Networks for Fake News Detection* [in:] *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, p. 2560–2569, online, Association for Computational Linguistics
- Livan, G., Sikder, O., Smith, R.E., Vivo, P. (2020) *A minimalistic model of bias, polarization and misinformation in social networks*, Scientific Reports 10, 5493, doi.org/10.1038/s41598-020-62085-w
- Lynnette, H.X.Ng., Taeihagh, A. (2021) *How does fake news spread? Understanding pathways of*

disinformation spread through APIs, Policy Internet. 2021, 13:560–585

McIntyre, A. (1988) *Whose Justice? Which Rationality?*, University of Notre Dame Press

Menczer, F., Yang, K.-Ch. (2023) *Anatomy of an AI-powered malicious social botnet*, Observatory on Social Media Indiana University, Bloomington | arXiv: 2307.16336v1

Miguel, R. (2023), *Platforms' Policies On Ai-Manipulated And Generated Misinformation*, EU DisinfoLab

Money (2023), *Wojna w Izraelu to problem dla Elona Muska. Platformie X grozi całkowite wyłączenie*, money.pl, 12.10.2023, bit.ly/46qLnvH (access: 12.10.2023)

Nieminen, S., Sankari, V. (2021) *Checking PolitiFact's Fact-Checks*, Journalism Studies, 22:3, 358-378, DOI: 10.1080/1461670X.2021.1873818

Pennycook, G., Rand, D.G. (2019), *Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning*, Cognition, 2019 Jul, 188:39-50, doi: 10.1016/j.cognition.2018.06.011

Porter, E., Wood, T.J. (2021) *The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom*, Proceedings of the National Academy of Sciences of the United States of America, doi.org/10.1073/pnas.2104235118

Qian, Y., Ye, Q., Yuan, H., Zhang, Y., Zheng, J. (2021) *Improving fake news detection with domain-adversarial and graph-attention neural network*, Decis. Support Syst. 151, 113633

Ren, Y., Zhang, J. (2021) *Fake news detection on news-oriented heterogeneous information networks through hierarchical graph attention* [in:] International Joint Conference on Neural Networks (IJCNN). IEEE, 2021, pp. 1–8

Rogers de Waal, J. (2021) *Globalism Project 2021: Highlights & Methodology*, YouGov, bit.ly/46wNT3p (access: 12.10.2023)

Tompkins, A. (2020), *Is fact-checking effective? A critical review of what works – and what doesn't*, DW Akademie (permalink: p.dw.com/p/3joav)

UNHCR (2021) *Using Social Media in Community Based Protection: a Guide*, Office of the United Nations High Commissioner for Refugees, January 2021

Wittgenstein, L. (1953) *Philosophical investigations*, New York, NY, USA: Wiley-Blackwell

Ten raport jest częścią międzynarodowego projektu współfinansowanego przez Unię Europejską (działanie nr 2020-EU-IA-0267) oraz przez polskie Ministerstwo Edukacji i Nauki w ramach programu Ministra Nauki i Szkolnictwa Wyższego pt. "PMW" w latach 2021 - 2024 (umowa nr 5213/CEF/2021/2).

Poglądy i wyrażone opinie są jednak wyłącznie poglądami autorów i niekoniecznie odzwierciedlają poglądy Unii Europejskiej. Ani Unia Europejska, ani Ministerstwo Edukacji i Nauki nie ponoszą za nie odpowiedzialności.

