mgr Dariusz Wachowicz
Wojskowa Akademia Techniczna w Warszawie
ORCID: https//orcid.org/ 0000-0003-0225-8219

ZARZĄDZANIE BEZPIECZEŃSTWEM INFORMACYJNYM W MEDIACH SPOŁECZNOŚCIOWYCH OKRESU TRANSFORMACJI CYFROWEJ

INFORMATION SECURITY MANAGEMENT IN SOCIAL MEDIA DURING THE DIGITAL TRANSFORMATION PERIOD

Streszczenie

W dobie dynamicznej transformacji cyfrowej media społecznościowe stały się kluczowym narzędziem komunikacji, informacji i wpływu społecznego. Poniższy artykuł przedstawia zagrożenia związane z bezpieczeństwem informacyjnym, które wynikają z otwartości, szybkości przepływu danych i powszechnego dostępu do treści generowanych przez użytkowników. Szczególny nacisk położono na problem dezinformacji, fake newsów oraz manipulacji opinią publiczną, które mogą destabilizować społeczeństwa i podważać zaufanie do instytucji. Autor podkreśla konieczność wdrażania kompleksowych strategii zarządzania ryzykiem informacyjnym, obejmujących zarówno rozwiązania technologiczne, jak i edukacyjne. Podkreśla rolę państw, instytucji międzynarodowych i samych użytkowników w budowaniu odporności informacyjnej. Artykuł wskazuje również na potrzebę współpracy międzysektorowej, aktualizacji regulacji prawnych oraz rozwijania kompetencji cyfrowych jako elementów kluczowych dla skutecznego zarządzania bezpieczeństwem informacji w erze cyfrowej.

Słowa kluczowe: dezinformacja, bezpieczeństwo informacyjne, media społecznościowe, zarządzanie bezpieczeństwem informacji, odporność społeczna, edukacja medialna

Summary

In the era of dynamic digital transformation, social media has become a key tool for communication, information exchange, and social influence. This article presents the threats to information security arising from the openness, speed of data flow, and widespread access to user-generated content. Particular emphasis is placed on the issues of disinformation, fake news, and manipulation of public opinion, which can destabilize societies and undermine trust in institutions. The author highlights the necessity of implementing comprehensive information risk management strategies that include both technological and educational solutions. The role of states, international institutions, and individual users in building informational resilience is emphasized. The article also points to the need for cross-sectoral cooperation, updates to legal regulations, and the development of digital competencies as key elements for effective information security management in the digital age

Keywords: disinformation, information security, social media, information security management, societal resilience, media literacy

Wstęp

Współczesne społeczeństwa funkcjonują w warunkach intensywnej cyfryzacji życia codziennego. Media społecznościowe, które z założenia postrzegane były głównie jako narzędzia komunikacji interpersonalnej – dziś odgrywają znacznie ważniejszą rolę w kształtowaniu opinii publicznej, mobilizacji społecznej, a także w transmisji zróżnicowanych gatunkowo informacji. Szczególną uwagę badaczy zajmujących się zarządzaniem bezpieczeństwem przyciąga rosnące znaczenie zjawiska informacji wprowadzających w błąd (złych i wadliwych), które w coraz większym stopniu oddziałuje na stabilność społeczną, zaufanie ludności wobec władz terytorialnych, samorządowych i centralnych oraz bezpieczeństwa narodowe. Analiza tego zjawiska wprowadzania w błąd i służącego manipulacji zaufaniem społecznym oraz wpływającego na sprawowanie władzy w kontekście funkcjonowania mediów społecznościowych oraz zarządzania bezpieczeństwem informacyjnym wymaga przedstawienie definicji i typologii informacji oraz wskazania mechanizmów jej rozpowszechniania w środowisku cyfrowym, a także identyfikacji narzędzi wykorzystywanych w kampaniach, które manipulują i wpływają na zaufanie społeczne podmiotów władzy i odporność instytucji etatystycznych. Dlatego autor podejmuje próbę ukazania roli systemów zarządzania bezpieczeństwem informacji, które mogą zwiększać odporność społeczeństwa na manipulację informacjami.

ZAKRES DEFINICYJNY I KONTEKST POJĘCIA DEZINFORMACJA

Pojęcie dezinformacji jest rodzajem informacji, która w semantycznej jej definicji lansowanej przez Luciano Floridi jest czymś, co składa się z danych i znaczeń prawidłowo ustrukturalizowanych,[1] zaś w definicji Mariana Turskiego jest to „wielkość abstrakcyjną, która może być przechowywana w pewnych obiektach, przesyłana między pewnymi obiektami i stosowana do sterowania pewnymi obiektami, przy czym przez obiekty rozumie się organizmy żywe, urządzenia techniczne oraz systemy takich obiektów".[2] Znaczenie pojęcia dezinformacji związać można z danymi, które są nieprawidłowo ustrukturalizowane i zakłócającymi sterowanie pewnymi obiektami. Ona dość powszechnie w literaturze przedmiotu odnosi się do celowego i intencjonalnego rozpowszechniania fałszywych, zmanipulowanych lub wprowadzających w błąd informacji, których celem jest osiągnięcie określonych korzyści politycznych, ekonomicznych bądź społecznych. Bohdan Pac definiuje dezinformację jako „rozpowszechnianie zmanipulowanych lub sfabrykowanych informacji w celu wywarcia wpływu na odbiorców i skłonienia ich do określonych zachowań na korzyść dezinformującego".[3] Z kolei w definicji przyjętej przez Komisję Europejską, za dezinformację uznaje się „możliwe do zweryfikowania nieprawdziwe lub wprowadzające w błąd informacje, tworzone, przedstawiane i rozpowszechniane w celu uzyskania korzyści gospodarczych lub wprowadzenia w błąd opinii publicznej, które mogą wyrządzić szkodę publiczną".[4] Należy mieć na uwadze jednak, że dezinformacja nie oznacza tego samego co malinformacja i misinformacja oraz fake news (rysunek 1). Malinformacja, czyli szkodliwa informacja, to treści oparte na faktach, które jednak zostają przedstawione w sposób zmanipulowany poprzez np. wyolbrzymienie lub wyrwanie z kontekstu, by celowo wprowadzić odbiorcę w błąd. Jest to informacja,

1 Zob.: L. Floridi, *The Ethcis of Information*, Oxford University Press, Oxford: 2013, s. 84-86.
2 M. Turski, *Propedeutyka informatyki*. PWN, Warszawa 1975, s. 8.
3 B. Pac, *Integracja wojny informacyjnej i hybrydowej w konfliktach międzynarodowych*, wydawnictwo Bellona, Warszawa, 1/2016, s. 64.
4 https://bip.brpo.gov.pl/pl/content/rpo-dezinformacja-walka-mc-odpowiedz (dostęp: 10. 06. 2025)

która choć oparta na rzeczywistości, jest „rozpowszechniana w celu wyrządzenia szkody osobie, grupie, organizacji lub narodowi. Ogólnie rzecz biorąc, odnosi się do strategicznego wykorzystania prawdziwych informacji, w tym wycieku informacji niejawnych, w celu spowodowania szkód".[5] Z kolei „misinformacja to treści nieprawdziwe, ale tworzone i rozpowszechniane bez intencji wyrządzenia szkody. Mogą powstawać w wyniku błędu lub braku rzetelności".[6] I wreszcie, fake newsy to wiadomości przeinaczone, określane także jako takie, które mają charakter określony przez Ralpha Keysa „post-prawdy". Według niego post-prawda (post-truth) opisuje sytuację, w której obiektywne fakty tracą znaczenie w kształtowaniu opinii publicznej, a emocje i osobiste przekonania odgrywają większą rolę. W kontekście politycznym i medialnym, post-prawda charakteryzuje się wykorzystywaniem przekazów, które są zgodne z emocjami i uprzedzeniami odbiorców, nawet jeśli są one sprzeczne z faktami.[7] Zatem można przyjąć, że zarówno fake newsy jak i post-prawda dotyczy subiektywnie przeinaczonych wiadomości i wiadomości. Wszystkie z wymienionych rodzajów dezinformacji rozpowszechniają treści świadomie i nieświadomie, celowo i przypadkowo, ale także z zamiarem wywołania skutków bardziej negatywnych niż pozytywnych. Przy czym misinformacja często nazywana nieintencjonalną dezinformacją wynika z szeregu różnych okoliczności. Do najczęściej spotykanych należą emocjonalna reakcja charakterystyczna dla post-prawdy oraz pobieżne zapoznanie się z informacją, które ograniczają zdolność do krytycznej oceny i sprzyjają impulsywnemu udostępnianiu. Zarówno misinformacja, dezinformacja i malinformacja oraz fake newsy swoim zakresem znaczeniowym się zazębiają, chociaż ujęcie tego zazębiania i różnic między nimi jest wciąż dyskutowane w literaturze przedmiotu. Niewątpliwie w dyskursie tym fake-newsy silnie wiązane są z post-prawdą. Dlatego one zdają się być tym, co obejmuje misinformacje, dezinformacje i malinformacje. Ujęcie takie, które jednak modyfikujące dość powszechne w literaturze, przedstawia poniższy rysunek 1.

Rys. 1 Różnice między dezinformacją a chybioną misinformacją, szkodliwą malinformacją i przeinaczoną informacją (post-prawdziwą)

Źródło: opracowanie własne na podstawie: C. Wardle, H. Derakhshan, Information Disorder: Toward on interdisciplinary framework for research and policy making, Council of Europe report, 2017, s. 20.

5 K. Mikulski, *Dezinformacja, misinformacja, malinformacja. Kilka słów o trollach i fake news!*, https://idzpodprad.pl/aktualnosci/dezinformacja-misinformacja-malinformacja-kilka-slow-o-trollach-i-fake-news/ (dostęp: 10. 06. 2025)

6 https://nask.pl/magazyn/dezinformacja-misinformacja-i-malinformacja-czym-sie-roznia (dostęp: 10. 06. 2025)

7 Zob. R. Keyes, *Czas postprawdy. Nieszczerość i oszustwa w codziennym życiu*, tłum. Paweł Tomanek, Wydawnictwo Naukowe PWN, Warszawa 2017, s. 390.

Dezinformacja może przyjmować różnorodne formy: od sensacyjnych nagłówków (clickbaitów), przez zmanipulowane zdjęcia i filmy (np. deepfake, AI)[8], po skoordynowane kampanie prowadzone przez tzw. farmy trolli. Kluczowym elementem dezinformacji jest jej zdolność do szybkiego rozprzestrzeniania się oraz wpływania na postawy i zachowania jednostek oraz grup społecznych. W literaturze przedmiotu zjawisko dezinformacji jest analizowane zarówno w ujęciu psychologicznym (np. wpływ poznawczy), socjologicznym (element polaryzacji społecznej), ale także z perspektywy zarządzania ryzykiem i bezpieczeństwem informacyjnym (np. identyfikacja zagrożeń, przeciwdziałanie, edukacja informacyjna).

MEDIA SPOŁECZNOŚCIOWE I ZARZĄDZANIE BEZPIECZEŃSTWEM W SIECI A DEZINFORMACJA

Media społecznościowe, takie jak Facebook, X, TikTok czy YouTube, oferują użytkownikom niespotykaną wcześniej łatwość tworzenia i udostępniania treści. Takie ułatwienie dostępu do narzędzi komunikacji sprzyja korzystnym procesom, takim jak aktywizacja społeczna czy większe zaangażowanie obywateli w życie publiczne. Jednocześnie jednak stwarza środowisko, w którym łatwiej rozprzestrzeniają się niezweryfikowane, a często również zmanipulowane informacje, czyli wadliwie ustrukturalizowane dane. Bardzo mocno sprzyja temu budowa platform społecznościowych, która oparta jest na algorytmach personalizacji, rekomendacjach treści i mechanizmach tzw. „viralów" – wszystkie te wskazane wzmacniają efekt bańki informacyjnej i sprzyjają szybkiemu rozprzestrzenianiu się treści emocjonalnych, kontrowersyjnych lub nieprawdziwych i powodujących zakłócenia sterownicze w zarządzaniu bezpieczeństwem. Zjawisko to określane jest często jako infodemia – czyli nadmiar informacji, w którym trudno odróżnić informacje sprzyjające bezpieczeństwu i jego sterowaniu, co prowadzi do dezinformacyjnego chaosu informacyjnego. W związku z tym należy podkreślić, że media społecznościowe nie są jedynie neutralnym kanałem przekazu, lecz aktywnym kreatorem całego systemu komunikacyjnego, mającym wpływ na przebieg debat publicznych, nastroje społeczne, a nawet procesy demokratyczne, jak wybory czy referenda.

System zarządzania bezpieczeństwem informacji (SZBI) zgodnie z ISO/IEC 27000 to „zbiór polityk, procedur, wytycznych oraz przydzielonych zasobów oraz aktywności, zarządzanych wspólnie przez organizację w celu ochronnych swoich zasobów informacyjnych"[9]. Ma on na celu ochronę informacji, systemów informacyjnych oraz użytkowników przed różnego rodzaju zagrożeniami, w tym przed dezinformacją. W związku z tym szczególnego znaczenia nabierają - gdy mowa o mediach społecznościowych - konkretne obszary zarządzania takie jak:

- identyfikacja i analiza ryzyk informacyjnych, która polega głównie na rozpoznawaniu źródeł i mechanizmów dezinformacji,

- ochrona użytkowników i instytucji, czyli implementacja rozwiązań technicznych (np. oznaczanie wprowadzających w błąd treści) i edukacyjnych (np. edukacja medialna),

- tzw. zarządzanie incydentami informacyjnymi obejmujące procedury reagowania na przypadki masowego rozpowszechniania błędnych treści,

8 Zob. https://www.pap.pl/aktualnosci/politycy-ofiarami-technologii-deepfake-ekspert-o-zagrozeniach (dostęp: 12. 06. 2025).
9 ISO/IEC 27001:2007 System Zarządzania Bezpieczeństwem Informacji

• współpraca międzysektorowa – współdziałanie administracji publicznej, sektora prywatnego i społeczeństwa obywatelskiego w celu zabezpieczenia użytkowników przed zjawiskiem dezinformacji – wprowadzaniem w błąd.

Dezinformacja w mediach społecznościowych stanowi jedno z najpoważniejszych wyzwań współczesnego zarządzania bezpieczeństwem informacyjnym. Jej złożony charakter, dynamiczne formy i globalny zasięg sprawiają, że skuteczne przeciwdziałanie wymaga interdyscyplinarnego podejścia, integrującego wiedzę z zakresu nauk społecznych, technologii informacyjnej oraz zarządzania. Stąd zasadność analizy praktycznych mechanizmów i strategii reagowania na dezinformację w środowisku cyfrowym.

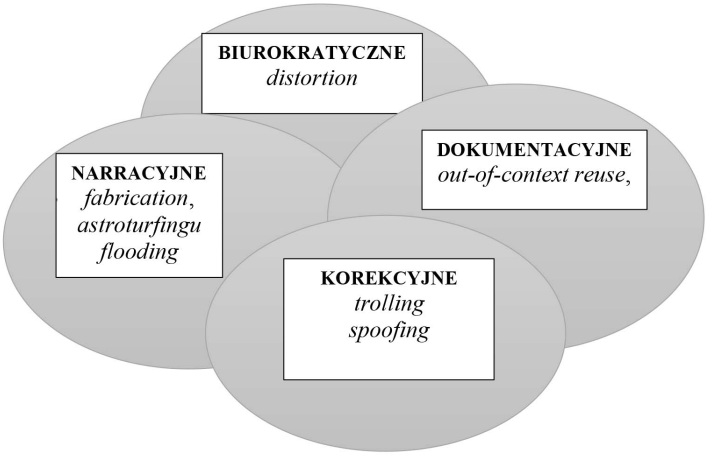
TECHNOLOGIE I NARZĘDZIA DEZINFORMACJI W ŚRODOWISKU MEDIÓW SPOŁECZNOŚCIOWYCH

Narzędzia dezinformacji jako rodzaju informacji są analogiczne z technologiami informacyjnymi, których Yuval Noah Harrari wyróżnił cztery[10]. Są to technologie narracyjne (tworzące informacje), dokumentacyjne (gromadzące informacje) i biurokratyczne (przechowujące informacje) oraz korekcyjne (wykorzystujące informacje i eliminujące błędy). Stąd zasadność rozróżnienia takich narzędzi dezinformacyjnych jak: narracyjne, dokumentacyjne, biurokratyczne i korekcyjne. Niewątpliwie dezinformacja w mediach społecznościowych nie jest efektem przypadku ani spontanicznego chaosu – coraz częściej przyjmuje formę zorganizowaną, celową i wspieraną zaawansowanymi technologiami. W literaturze oraz w analizach ośrodków badawczych zajmujących się bezpieczeństwem informacyjnym (np. EUvsDisinfo[11], NATO StratCom COE[12], Stanford Internet Observatory[13]) identyfikuje się pięć technik dezinformacyjnych wykorzystywanych w mediach społecznościowych. Do najbardziej popularnych według literatury przedmiotu należą: zniekształcanie faktów i zjawisk (distortion), polegające na częściowym przedstawieniu informacji w sposób selektywny, by zmienić jej wydźwięk – korelujące z technologiami biurokratycznymi, fabrykowanie treści (fabrication), tzn. tworzenie całkowicie błędnych informacji, często z wykorzystaniem fikcyjnych źródeł[14] - korelującymi z technologiami narracyjnymi, out-of-context reuse, czyli wykorzystywanie autentycznych materiałów (zdjęć, cytatów) w zmienionym kontekście, prowadzącym do błędnych wniosków – korelujące z technologiami dokumentacyjnymi, flooding – masowe rozprzestrzenianie dużej liczby sprzecznych komunikatów w celu osłabienia zdolności do rozróżniania informacji sprzyjających i niesprzyjających bezpieczeństwu, błędnych i niebłędnych – korelujące z technologiami narracyjnymi, trolling informacyjny, czyli celowe publikowanie prowokacyjnych treści w celu wywołania chaosu informacyjnego lub polaryzacji – korelujące z technologiami korekcyjnymi. Dodatkowo często stosowaną i wyróżnianą jest również szósta i siódma technika, a mianowicie: astroturfing, czyli symulowania spontanicznego poparcia społecznego dla określonych idei lub poglądów np. za pomocą botów inicjowanych przez zorganizowane grupy interesu – korelujące z technologiami narracyjnymi, a także tzw. Spoofing (ang. spoof – naciąganie, podszywanie) tożsamościowy, czyli

10 Zob. Y. N. Harari, *Nexus. Krótka historia informacji od epoki kamienia do sztucznej inteligencji*, tłum. Justyna Hunia, Wydawnictwo Literackie, Kraków 2024.
11 https://euvsdisinfo.eu/pl/ (dostęp: 14. 06. 2025)
12 https://stratcomcoe.org/ (dostęp: 14. 06. 2025)
13 https://cyber.fsi.stanford.edu/ (dostęp: 14. 06. 2025)
14 Por. E. Sadowska, *Ewolucja Cyberzagrożeń: Deepfake i Media Syntetyczne w kontekście bezpieczeństwa energetycznego Europy Wschodniej*, „Studia Wschodnioeuropejskie, Nr Ekspercki 19-t", 2023 nr 2, s. 8-18.

podszywanie się pod osoby lub instytucje, co bywa szczególnie skuteczne w kampaniach politycznych i wyborczych – korelujące z technologiami korekcyjnymi. Przeto syntetycznie technologie i skorelowane z nimi narzędzia dezinformacyjne można na bazie wiodącej i przywołanej wyżej literatury przedstawić tak, jaka na poniższym rysunku 2.

Rys. 2 Technologie i narzędzia dezinformacji



Źródło: opracowanie własne

Efektywność kampanii dezinformacyjnych opiera się w głównej mierze na zdolności do szybkiego i szerokiego dotarcia z przekazem. Aby to osiągnąć twórcy wykorzystują infrastrukturę techniczną, która obejmuje zarówno konta użytkowników rzeczywistych, jak i sztucznie generowane podmioty:

- wspomniane boty społecznościowe – automatyczne konta generujące lub udostępniające treści, często w sposób zsynchronizowany, by zwiększyć zasięg i wiarygodność komunikatu,

- farmy trolli – zorganizowane grupy osób odpowiedzialnych za zarządzanie kontami i aktywne uczestnictwo w dyskusjach online w sposób sugerujący autentyczność społecznego poparcia,

- konta deepfake'owe – wykorzystujące syntetycznie generowane obrazy, wideo lub głos, by tworzyć fikcyjne osoby, które mogą stać się wpływowymi uczestnikami debaty (działające bardzo często na tej samej zasadzie co boty społecznościowe),

- fora dyskusyjne, komunikatory, grupy zamknięte oraz rozproszone sieci udostępniania danych (np. Telegram, Reddit, Discord, Messenger, WhatsApp), które nie są domeną wyłącznie jednostek działających, aby zaszkodzić innej jednostce lub wprowadzić ją w błąd.

Obecnie można spotkać kampanie dezinformacyjne, które realizowane są przez różnorodne podmioty np. państwa narodowe, grupy interesu czy też niezależnych twórców

internetowych. Wspomniane już państwa mogą siać dezinformację w konkretnych celach politycznych zarówno zewnętrznych (np. destabilizacja przeciwnika geopolitycznego) jak i wewnętrznych (wpływ na wyniki wyborów, polaryzacja społeczeństwa czy podważanie autorytetu instytucji demokratycznych)[15]. „Newsy w najpopularniejszych polskich programach informacyjnych różnią się stylistyką i sposobem prezentowania treści. Widać to zwłaszcza w okresach ważnych, w których rozgrywają się wydarzenia wysokiej rangi, np. parlamentarna kampania wyborcza[16]. Znanym trendem w XXI wieku staje się próba uatrakcyjnienia przekazów przez ograniczenie ekspozycji problematyki złożonej i skomplikowanej, która mogłaby znudzić odbiorców, a wyolbrzymienie elementów, które mają wzmocnić pozycję komunikujących informacje i dezinformacje. Z kolei grupy interesu mogą prowadzić kampanie dezinformacyjne motywowane ekonomicznie, ponieważ kontrowersyjne treści są najbardziej nastawione na generowanie zysków z kliknięć, reklam i subskrypcji, bez względu na ich wartość informacyjną. Jeszcze innym przykładem są motywacje ideologiczne i społeczne. Działania te inspirowane są przez radykalne grupy społeczne lub religijne, których celem jest szerzenie określonych narracji światopoglądowych. Społeczności te są określane często jako wyznawcy teorii spiskowych czy dezinformujące ruchy społeczne. „Społeczności te zaprzeczają oficjalnym narracjom, przeciwstawiając im alternatywne i fałszywe wyjaśnienia zjawisk społecznych. Badania w tym nurcie wskazują, że tego typu społeczności wirtualne stanowią kluczowy element w rozpowszechnianiu dezinformacji oraz przyczyniają się do spadku zaufania do sfery informacyjnej i ważnych instytucji społecznych, których rolą jest m.in. ochrona obywateli"[17]. Radzenie sobie z tymi narracjami wymaga stosowania technologii korekcyjnych i nie należy do łatwych, chociaż obejmuje wiele procedur rekomendowanych w intrenecie,[18] wśród których jest i wskazanie 7 sposobów praktycznych na to, jak ochronić się przed dezinformacją.

ZARZĄDZANIE BEZPIECZEŃSTWEM INFORMACYJNYM W KONTEKŚCIE DEZINFORMACJI W INTERNECIE

Zarządzanie bezpieczeństwem informacyjnym w dobie wzrostu znaczenia zjawiska dezinformacji wymaga zintegrowanego podejścia, obejmującego zarówno działania reaktywne, jak i te proaktywne. Media społecznościowe, ze względu na swoją strukturę, skalę zagrożenia oraz podatność na manipulacje, stanowią dziś jedno z kluczowych pól operacyjnych dla tego rodzaju działań. Podstawą skutecznego reagowania na zagrożenia informacyjne jest ich wczesna identyfikacja. Istnieje wiele narzędzi, które mają spełniać rolę pomocnika użytkowników podczas surfowania w Internecie wykrywającego potencjalne zagrożenia. Jednym z nich jest system wczesnego ostrzegania (early warning systems EWS), który „wykorzystuje sztuczną inteligencję i zaawansowane algorytmy do monitorowania, wykrywania i przeciwdziałania rozprzestrzenianiu się fałszywych treści w intrenecie. Jego celem jest szybka identyfikacja zagrożeń, takich jak manipulacje medialne, które mogą pro-

15 Por. Ch.Wells, D. Freelon, *Disinformation as Political Communication*, Political Communication, 2020, vol. 37, no. 2, s. 145-156.

16 H. Batorowska, R. Klepka, O. Wasiuta, *Media jako instrument wpływu informacyjnego i manipulacji społeczeństwem*, Wydawnictwo LIBRON, Kraków, 2019, s. 328-329.

17 J. Czerwiński, *„Pożyteczni idioci" kampanii dezinformacyjnych. Mechanizmy tworzenia się społeczności wirtualnych rozpowszechniających fałsz i manipulację w sieci*, Konteksty Społeczne/Social Contexts, 2022, Tom 10, nr 2(20), s.7.

18 Zob. S. Adamczyk, *Jak przeciwdziałać dezinformacji? Poradnik dla administracji publicznej*, NASK , https://ochronaludnosci.edu.pl (dostęp: 15. 06. 2025)

wadzić do destabilizacji społecznej, przemocy lub zakłócenia procesów demokratycznych"[19]. W tej aplikacji analizie poddaje się takie wskaźniki jak nagły wzrost częstotliwości pojawiania się określonych fraz lub haseł, nienaturalne wzorce publikacji treści (np. wysoka aktywność botów) czy rozchodzenie się treści w zamkniętych sieciach lub grupach tematycznych. Coraz większe znaczenie zyskują również narzędzia do detekcji deepfake (fałszywych wideo, imitujących autentyczne nagrania), a także tzw. mapowanie sieci narracyjnych, pozwalające na zlokalizowanie źródeł dezinformacji i ich kanałów dystrybucji. „Nowoczesne metody łączą analizę sieciową, przetwarzanie języka naturalnego i uczenie maszynowe, aby skutecznie wykrywać, analizować i wizualizować rozprzestrzenianie się fałszywych narracji oraz identyfikować kluczowych aktorów i mechanizmy dystrybucji dezinformacji"[20].

Kolejnym bardzo ważnym elementem systemowej strategii zarządzania bezpieczeństwem informacyjnym jest budowanie odporności społecznej przez edukację medialną i informacyjną. Rozpoznawanie manipulacji, analiza źródeł oraz rozumienie procesów algorytmicznych stanowią fundament odpornego społeczeństwa cyfrowego. Żadne narzędzie czy aplikacja nie jest w stanie zastąpić krytycznego myślenia człowieka. Poniżej wskazanych zostanie kilka programów edukacyjnych w tym zakresie realizowanych w Polsce:

- zajęcia dydaktyczne w szkołach i uczelniach. Realizowane lekcje uczą krytycznego myślenia i podstaw fact-checkingu. W szkołach najczęściej są one integrowane z przedmiotami humanistycznymi, a na uczelniach występują jako odrębne przedmioty w ramach niektórych kierunków studiów;

- otwarte kursy online - dla użytkowników Internetu dostępne są darmowe kursy, które omawiają mechanizmy dezinformacji, techniki manipulacji, rolę algorytmów oraz sposoby identyfikowania fałszywych treści. Przykładowe kursy dostępne są na platformach takich jak Coursera, FutureLearn czy edukacyjne portale europejskie (np. EUvsDisinfo);

- kampanie społeczne mające na celu podniesienie świadomości społecznej na temat zagrożeń związanych z dezinformacją. Wykorzystują one różne kanały komunikacji – spoty w mediach, grafiki w Internecie, plakaty, lub działania w przestrzeni publicznej. Jednym z przykładów takiej akcji jest ostatnia kampania Fundacji PZU „Czysty przekaz". Ambasadorami kampanii są Szymon Marciniak, Magda Gessler oraz Mikołaj Roznerski, którzy „pokazują w spocie telewizyjnym, że choć w swojej pracy stawiają na czystość i autentyczność – w przepisach kulinarnych, decyzjach na boisku, grze aktorskiej – to w Internecie rozpoznanie czystej prawdy bywa trudniejsze. Pomaga w tym filtrowanie informacji, sprawdzanie źródeł oraz uważność w udostępnianiu treści"[21].

19 M. Yankoski, T. Weninger, W. Scheirer, *An AI early warning system to monitor online disinformation, stop violence, and protect elections*, Bulletin of the atomic scientists, 2020, vol. 76, no. 2, s. 85-86, tłum. własne.
20 Ibidem, s. 88 (tłum. własne)
21 https://czystyprzekaz.pl/ (dostęp: 14. 06. 2025)

Rys. 3 Plakat promujący kampanię społeczną #czystyprzekaz

Źródło: https://czystyprzekaz.pl

- aktywności organizacji fact-checkingowych. Organizacje fact-checkingowe prowadzą bieżącą weryfikację treści krążących w mediach i Internecie. Udostępniają publicznie raporty, analizy i sprostowania, często też prowadzą działania edukacyjne – warsztaty, webinary, kampanie informacyjne. W Polsce znaną organizacją fact-checkingową jest Demagog, który publikuje codzienne analizy wypowiedzi polityków.

Rys. 4 Analizy wypowiedzi kandydatów na Prezydenta RP, podczas debaty prezydenckiej

Źródło: https://demagog.org.pl/analizy_i_raporty/debata-tvp-trzaskowski-konta-nawrocki-fact-checking-na-zywo/ (dostęp: 14. 06. 2025)

Jednym z ważniejszych elementów zarządzania bezpieczeństwem informacyjnym są działania platform społecznościowych, które są kanałem dystrybucji dezinformacji, a jednocześnie aktorem, który może i powinien próbować przeciwstawiać się temu zagrożeniu. W odpowiedzi na presję społeczną, medialną i legislacyjną platformy wdrożyły szereg me-

chanizmów mających na celu ograniczenie rozprzestrzeniania się fałszywych informacji np. Facebook i Meta przez kilka lat współpracowały z IFCN (International Fact-Checking Network). Niestety w 2025 współpraca ta została zakończona przez głównie rzekome ograniczanie wolności słowa na wskazanych portalach przez IFCN.[22] Dodatkowo media społecznościowe używają algorytmów Ograniczających zasięgi lub widoczność treści ocenionych jako fałszywe, czy też oznaczają treści dezinformujące lub zawierające niezweryfikowane informacje (np. podczas pandemii COVID-19 lub wyborów). Skuteczność tych działań jest różna i często oceniana krytycznie. Główne problemy to brak przejrzystości algorytmicznej, niski poziom moderacji w językach innych niż angielski oraz niejednolite standardy między platformami.

Kluczowym aspektem zarządzania bezpieczeństwem informacyjnym jest przygotowanie instytucji do reagowania na incydenty dezinformacyjne, czyli chybione i złośliwe oraz wprowadzające w błąd ustrukturalizowanie danych. Tego rodzaju sytuacje, często o charakterze kryzysowym, wymagają zastosowania procedur analogicznych do zarządzania kryzysowego lub cyberbezpieczeństwa. Jednym z zasadniczych elementów tego procesu jest systematyczne zbieranie i analiza danych. Obejmuje to zarówno dane pochodzące z otwartych źródeł oraz danych telemetrycznych generowanych wewnętrznie przez systemy informacyjne organizacji. Analiza tych danych pozwala na szybkie wykrycie anomalii, identyfikację źródeł dezinformacji oraz ocenę potencjalnego zagrożenia dla reputacji i funkcjonowania instytucji.[23] Równie ważnym zdaje się być opracowanie kontr-narracji, czyli komunikatów informacyjnych, które w sposób wiarygodny i transparentny prostują (korygują) dezinformacje oraz przeciwdziałają ich dalszemu rozpowszechnianiu. Kontr-narracje muszą być precyzyjnie dostosowane do odbiorcy i kontekstu sytuacyjnego, aby skutecznie neutralizować wpływ dezinformacji i przywracać zaufanie do oficjalnych źródeł informacji. W kontekście środowiska cyfrowego niezbędna jest również współpraca z platformami społecznościowymi, jak i operatorami treści w celu szybkiego usuwania lub oznaczania dezinformujących materiałów. Ostatnim, chyba najważniejszym jest systematyczna ocena skutków incydentu oraz jego dokumentowanie. Działania te stanowią istotny komponent zarządzania wiedzą instytucjonalną i pamięcią organizacji, umożliwiając analizę skuteczności podjętych działań, wyciągnięcie wniosków oraz aktualizację procedur reagowania na podobne zdarzenia w przyszłości.

Podsumowanie

Zjawisko dezinformacji w mediach społecznościowych stanowi dziś jedno z najbardziej złożonych i dynamicznych wyzwań dla bezpieczeństwa informacyjnego. Przeprowadzana analiza wskazuje jednoznacznie, że skuteczne zarządzanie w tym obszarze wymaga zastosowania wielopoziomowych strategii i technologii, które łączyć będą narzędzia informatyczne, działania edukacyjne, odpowiedzialność instytucjonalną i regulacyjną oraz mechanizmy reagowania kryzysowego (korekcyjne). Dezinformacja nie tylko zakłóca racjonalny proces podejmowania decyzji, ale także podważa zaufanie społeczne (ów nexus Harariego), niszczy autorytet władczy instytucji publicznych i destabilizuje sferę publiczną. W konsekwencji tego bezpieczeństwo informacyjne przestaje być jedynie domeną informatyków i staje się coraz bardziej interdyscyplinarnym obszarem, wymagającym ścisłej współpracy między

22 Zob.: https://indid.pl/facebook-pozbywa-sie-fact-checkerow/ (dostęp: 15. 06. 2025).
23 Zob. https://nask.pl/magazyn/rola-instytucji-w-zwalczaniu-dezinformacji-wizja-dzialania-osrodka-analizy-dezinformacji-nask (dostęp: 15. 06. 2025).

sektorem publicznym, prywatnym i społecznym. Niezbędnym wydaje się instytucjonalizacja zarządzania dezinformacją, zarówno na poziomie krajowym, jak i międzynarodowym. Takowe powinno objąć tworzenie dedykowanych jednostek analitycznych, rozwój kompetencji w strukturach zarządzania kryzysowego oraz budowę formalnych procedur reagowania korekcyjnego. Należy również skupić się nad obecnie dostępnymi narzędziami informacyjnymi, które wciąż są niewystarczające. Algorytmy wykrywania dezinformacji, systemy klasyfikacji treści czy narzędzia do analizy narracji powinny być uzupełniane o czynniki ludzkie, takie jak wiedza ekspercka, analiza kontekstowa, czy decyzje etyczne, czego sztuczna inteligencja („nie-ludzka inteligencja") jest pozbawiona. Należy znaleźć równowagę pomiędzy automatyzacją a czynnikiem ludzkim i stosownymi regulacjami prawnymi. Ważnym jest również stałe budowanie fundamentów bezpieczeństwa przez masową edukację, budowania nawyków weryfikacji informacji i promowania kultury informacyjnej. Nawet najbardziej zaawansowane algorytmiczne systemy techniczne Sztucznej Inteligencji nie dają 100% pewności uniknięcia dezinformacji. Należy także skupić się na większym penalizowaniu platform cyfrowych, które muszą ponosić realną odpowiedzialność za bezpieczeństwo informacyjne swoich użytkowników. Dotychczasowe działania są dość powierzchowne i niedostosowane do skali zagrożeń. Niezbędne są skuteczne i egzekwowalne ramy prawne, takie jak Digital Services Act, które definiować będą granice odpowiedzialności i standardy ochrony informacyjnej.

Niniejszy artykuł stanowi punkt wyjścia do dalszych badań i analiz, które powinny obejmować studia przypadków kampanii dezinformacyjnych, ewaluację skuteczności narzędzi przeciwdziałania dezinformacji oraz badania porównawcze strategii wobec dezinformacji w różnych państwach i Unii Europejskiej. Ochrona przed dezinformacją wymaga interdyscyplinarnego podejścia, które pomoże wypracować trwałe i efektywne mechanizmy ochrony przed jednym z najpoważniejszych zagrożeń dla społeczeństwa transformacji cyfrowego.

BIBLIOGRAFIA

Literatura

Batorowska H., Klepka R., Wasiuta O., *Media jako instrument wpływu informacyjnego i manipulacji społeczeństwem*, Wydawnictwo LIBRON, Kraków, 2019.

Czerwiński J., *„Pożyteczni idioci" kampanii dezinformacyjnych. Mechanizmy tworzenia się społeczności wirtualnych rozpowszechniających fałsz i manipulację w sieci*, Konteksty Społeczne/Social Contexts, 2022, Tom 10, Nr 2(20).

Floridi L., *The Ethcis of Information*, Oxford University Press, Oxford: 2013.

Harari Y. N., *Nexus. Krótka historia informacji od epoki kamienia do sztucznej inteligencji*, tłum. J. Hunia, Wydawnictwo Literackie, Kraków 2024..

Keyes R., *Czas postprawdy. Nieszczerość i oszustwa w codziennym życiu*, tłum. Paweł Tomanek, Wydawnictwo Naukowe PWN, Warszawa 2017.

Pac B., *Integracja wojny informacyjnej i hybrydowej w konfliktach międzynarodowych*, wydawnictwo Bellona, Warszawa, 1/2016.

Sadowska E., *Ewolucja Cyberzagrożeń: Deepfake i Media Syntetyczne w kontekście bezpieczeństwa energetycznego Europy Wschodniej*, Studia Wschodnioeuropejskie, Nr Ekspercki 19-t. 2/2023, Warszawa, 2023.

Turski M., *Propedeutyka informatyki*, PWN, Warszawa 1975.

Wardle C., Derakhshan H., *Information Disorder: Towards an Interdisciplinary Framework for Research and Policy making*, Council of Europe report, 2017.

Wells Ch., Freelon D., *Disinformation as Political Communication*, Political Communication, 2020, vol. 37, no. 2.

Yankoski M., Weninger T., Scheirer W., *An AI early warning system to monitor online disinformation, stop violence, and protect elections*, Bulletin of the Atomic Scientists, 2020, vol. 76, no. 2.

Źródła internetowe

Mikulski K., *Dezinformacja, misinformacja, malinformacja. Kilka słów o trollach i fake news!*, https://idzpodprad.pl/aktualnosci/dezinformacja-misinformacja-malinformacja-kilka-slow-o-trollach-i-fake-news/ (dostęp: 10. 06. 2025).

Adamczyk S., *Jak przeciwdziałać dezinformacji? Poradnik dla administracji publicznej*, NASK, https://ochronaludnosci.edu.pl (dostęp: 15. 06. 2025)

https://bip.brpo.gov.pl/pl/content/rpo-dezinformacja-walka-mc- (dostęp: 15. 06. 2025).

https://nask.pl/magazyn/dezinformacja-misinformacja-i-malinformacja-czym-sie-roznia (dostęp: 15. 06. 2025).

https://www.pap.pl/aktualnosci/politycy-ofiarami-technologii-deepfake-ekspert-o-zagrozeniach (dostęp: 12. 06. 2025).

https://euvsdisinfo.eu/pl/ (dostęp: 14. 06. 2025).

https://stratcomcoe.org/ (dostęp: 14. 06. 2025).

https://cyber.fsi.stanford.edu/ (dostęp: 14. 06. 2025).

https://czystyprzekaz.pl/ (dostęp: 14. 06. 2025).

https://demagog.org.pl/analizy_i_raporty/debata-tvp-trzaskowski-konta-nawrocki-fact-checking-na-zywo/ (dostęp: 15. 06. 2025).

https://indid.pl/facebook-pozbywa-sie-fact-checkerow/ [dostęp: 15 czerwca 2025 r.].

https://nask.pl/magazyn/rola-instytucji-w-zwalczaniu-dezinformacji-wizja-dzialania-osrodka-analizy-dezinformacji-nask (dostęp: 15. 06. 2025).