

MANIPULACJA I DEZINFORMACJA: FAKE NEWS I DEEPPFAKE

Wstęp

Nowoczesne środki komunikacji zapewniają światu szybką wymianę informacji. Ten proces jeszcze kilkadziesiąt lat temu był dostępny jedynie w sferze science fiction, a dzisiaj – w celu zobrazowania – można go porównać do prędkości światła. Newsy, na które odbiorcy czekali kilka godzin czy dni, dzisiaj stają się dostępne w ciągu kilku sekund – i to w najwyższej jakości obrazu i dźwięku. Szybkość i technologia rozprzestrzeniania się wiadomości doprowadziła również do wielu prób zniekształcania rzeczywistości bądź wręcz zakłamywania jej w sposób profesjonalny. Wiele newsów zawiera częściowo lub całkowicie zniekształcone elementy, a fakty nierzadko zajmują drugorzędne miejsce¹.

Fake news i deepfake, bo o nich mowa, są znaczącą cechą tak zwanej post-prawdy, czyli rzeczywistości częściowo wykreowanej przez media, która powoduje nieporozumienia w sferze publicznej. Szczególne trudności w rozróżnianiu fikcji od fałszu mają młodsze pokolenia. „Cyfrowi tubylcy” doskonale posługują się mediami społecznościowymi, ale jeśli chodzi o ocenę prawdziwości informacji przepływających przez te kanały, łatwo dają się oszukać². Również dojrzały odbiorcy przekazów medialnych mają kłopoty z odróżnieniem kłamstw i fałszu od prawdy, zważywszy na to, że zachowania manipulacyjne ewoluują i stają się coraz bardziej wyrafinowane (różnego rodzaju manipulacje przy użyciu sztucznej inteligencji).

Fake news

Rozwój Internetu spowodował, że w dzisiejszym świecie użytkownik może znaleźć online większość faktów, ale również wiele nieprawdziwych informacji, czyli fake newsów. Fake news to celowo przygotowane materiały, które z zasady mają wprowadzać odbiorcę w błąd; to wiadomość nieprawdziwa, niedokładna, nieodpowiadająca faktom. Fake news jako dezinformacja wpływa negatywnie na zaufanie odbiorców do danego medium³.

Raport Newseria podaje następującą definicję fake newsa: „(...) informacja opublikowana przez media, sprawiająca wrażenie zweryfikowanej oraz opisującej fakty, która w rzeczywistości wprowadza opinię publiczną w błąd, uwiarygadniając zawarte w niej niepotwierdzone informacje, dane oraz niezwerifikowane źródła”⁴.

Badacze Johnny Botha i Heloise Pieterse wyróżniają następujące rodzaje fake newsów⁵:

Clickbait: obejmuje wiadomości, które są celowo sfabrykowane, aby zdobyć więcej odwiedzających witrynę i zwiększyć przychody z reklam dla witryn internetowych.

¹ Anthony Le Duc, Catholic Church Communication in the Post-Truth Era: Intra-Religious and Inter-Religious Dimensions, 2019, <https://ssrn.com/abstract=3335641>, 1.

² Ibid, 3.

³ Marek Chyliński, Fałszywe wiadomości – antydobra w ekosystemie informacji, „Com.press. Pismo naukowe. Komunikacja i nauki o mediach. Journal of communication and media science”, 4 (2018), 9-10.

⁴ Raport Newseria, Fake news, czyli jak kłamstwo rządzi światem, Wrzesień 2017, 4.

⁵ Johnny Botha, Heloise Pieterse, Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security, 2020.

Satyra/parodia (satire/parody): wiadomości stworzone wyłącznie w celach rozrywkowych, bez zamiaru wyrządzenia jakiegokolwiek szkody, które mogą oszukać czytelnika.

Fałszywe powiązania/wprowadzające w błąd nagłówki (false connection/misleading headings): dotyczy wiadomości lub artykułów, które zawierają prawdziwą i dokładną treść, ale wykorzystują wprowadzające w błąd lub sensacyjne nagłówki.

Propaganda: odnosi się do wiadomości, które są celowo tworzone w celu wprowadzenia odbiorców w błąd lub promowania stronniczego punktu.

Nieobiektywne/przesadzone wiadomości (biased/slanted news): obejmują wiadomości, które opierają się na uprzedzeniach i przekonaniach adresatów.

Niechlujne dziennikarstwo/błąd (sloppy journalism/error): wiąże się z wiadomościami tworzonymi przy użyciu niewiarygodnych informacji lub niezaufanych źródeł, które mogą wprowadzać konsumentów w błąd.

Zmanipulowane wiadomości (manipulated news): odnoszą się do wiadomości, które manipulują treścią prawdziwych i opartych na faktach historii, aby oszukać czytelników.

Sfabrykowane wiadomości (fabricated news): obejmują wiadomości zawierające w 100% fałszywe treści stworzone w celu oszukania i wyrządzenia szkody.

Treści sponsorowane (sponsored content): to wiadomości lub reklamy, które są zamaskowane jako treści redakcyjne, które mogą wprowadzać użytkowników w błąd.

Fake newsy są rozpowszechniane za pośrednictwem tradycyjnych mediów oraz mediów związanych z Internetem. Tego rodzaju treści czasami stają się bardzo niebezpieczne, ponieważ wpływają na całe społeczeństwa i są podstawą błędnych decyzji. Stąd tak ważna jest edukacja związana z właściwym korzystaniem z mediów i umiejętnością oceny informacji. Krytyczne myślenie, weryfikacja źródeł czy promowanie umiejętności cyfrowych, to podstawy w demaskowaniu fake newsów, które stają się coraz bardziej wyrafinowane z powodu rozwoju technologicznego.

Deepfake

Fałszywe treści wideo, czyli deepfake, generowane przez sztuczną inteligencję stają się coraz bardziej powszechne i przekonujące. Deepfake to „ewolucja” fake newsa, która wykorzystuje sztuczną inteligencję. Nowe techniki pozwalają również niewykwalifikowanym osobom na tworzenie deepfake, prezentując wideo i głos znanych osób. Deepfake narodził się w 2017 roku w obrębie filmów pornograficznych. Jego istotą są zaawansowane algorytmy, które uczą się podobieństw między prawdziwą i fałszywą twarzą czy głosem, redukują całość do wspólnych cech i tworzą fake⁶. Termin deepfake pochodzi od akronimu Deep Learning i Fake. Deepfake to technika stosowana w treściach wideo, która umożliwia edycję filmów, których końcowy wynik należy uznać za fałszywy. W wielu sytuacjach wykorzystywany jest wizerunek osób znanych (celebrytów, polityków), na przykład Barack Obama, Władimir Putin, Nancy

⁶ Ian Sample, What are deepfakes – and how can you spot them?, The Guardian, <https://tiny.pl/cwbrc>.

Pelosi czy Mark Zuckerberg⁷. Efektem jest tworzenie treści wideo, opartych na istniejącym już contencie dostępnym w Internecie. Z pomocą sztucznej inteligencji oraz dzięki materiałowi wyjściowemu z nagraniami danej osoby, można wygenerować nowe materiały wideo tej osoby, mówiącej dowolne rzeczy. Na YouTube można znaleźć coraz więcej tego rodzaju filmów⁸.

Dostępne filmy deepfake można umiejscowić w jednej z następujących grup⁹:

Demonstracja technologii (technology demonstration): obejmuje filmy deepfake, które są tworzone jako przykłady demonstrujące działanie technologii.

Satyryczne/memy (satirical/meme): odnoszą się do filmów deepfake, które są humorystyczne lub prześmiewcze, stworzone jako forma komentarza politycznego lub społecznego.

Pornograficzne (pornographic): wiążą się z filmami deepfake, które często zawierają twarze znanych celebrytów na ciele aktorów pornograficznych.

Zwodniczy deepfake (deceptive deepfake): fałszywe filmy prezentujące polityków lub inne autorytety z zamiarem wywołania skandalu.

Dane wejściowe mogą być stosowane przez algorytmy i odpowiednie oprogramowanie na różne sposoby przy tworzeniu deepfake. Oto przykłady: Zamiana twarzy (face swap) – przeniesienie twarzy jednej osoby na twarz innej osoby na filmie; edycja atrybutów (attribute editing) – zmiana cech charakterystycznych osoby; odwzorowanie twarzy (face reenactment) – przeniesienie ekspresji mimicznej z twarzy jednej osoby na osobę w docelowym filmie wideo; w pełni syntetyczny materiał (fully synthetic material) – prawdziwy materiał filmowy jest używany do treningu oprogramowania, ale materiał wynikowy jest całkowicie nowy¹⁰.

Deepfake to hiperrealistyczne wideo, które zostało cyfrowo zmanipulowane w celu przedstawienia ludzi mówiących lub robiących rzeczy, które w rzeczywistości nie miały miejsca. Tego rodzaju контент generowany przy użyciu technik komputerowych opartych na sztucznej inteligencji (AI), stanowi poważne zagrożenie dla prywatności, i w którym ryzyko wynikające z kradzieży tożsamości jest coraz większe¹¹.

Edukacja społeczeństwa w zakresie rozpoznawania deepfake ma kluczowe znaczenie przy rozwijającej się technologii, w celu uniknięcia dezinformacji. Warto wyjaśniać odbiorcom proces powstawania tego rodzaju treści, podawać przykłady i promować odpowiedzialne i krytyczne korzystanie z mediów, ponieważ deepfake stanowi realne cyberzagrożenie dla użytkowników. Należy jednocześnie pamiętać, że ta technologia jest również wykorzystywana w formie pozytywnej, rozrywkowej czy edukacyjnej¹².

⁷ Inesdi, Deep Fake y la manipulación de procesos, <https://tiny.pl/cwbrl>.

⁸ Deepfake Videos Are Getting Terrifyingly Real, <https://www.youtube.com/watch?v=T76bK2t2r8g>.

⁹ Johnny Botha, Heloise Pieterse, Fake News and Deepfakes...

¹⁰ Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, 9.

¹¹ Francisco José García-Ull, Deepfakes: el próximo reto en la detección de noticias falsas, *Anàlisi: Quaderns de Comunicació i Cultura*, 64 (2021), 103.

¹² Museum creates deepfake Salvador Dalí to greet visitors, <https://tiny.pl/cwbrq>.

Franciszek – fake news i deepfake

Fake news i deepfake to treści, które są całkowicie lub częściowo fałszywe, ale mimo to są publikowane w Internecie i innych mediach w celach często nacechowanych przemyślaną manipulacją. Nierzadko są rozpowszechniane z powodów politycznych, finansowych lub rozrywkowych i bywa, że mają tak szeroki zasięg, że w dużej mierze fałszują rzeczywistość lub wizerunek firmy, instytucji lub osoby. Również informacje dotyczące Kościoła nie są wolne od tego rodzaju treści.

Papież Franciszek wielokrotnie odnosił się do tego tematu. W 2018 r. poświęcił zjawisku fake news orędzie na 52. Światowy Dzień Środków Społecznego Przekazu. „W grę wchodzi bowiem nasza chciwość. Fake news stają się często wirusowe, to znaczy rozprzestrzeniają się w szybki i trudny do powstrzymania sposób, nie ze względu na logikę dzielenia się, która charakteryzuje media społecznościowe, ile raczej ze względu na ich oparcie w nienasyconej chciwości, która łatwo rozpala się w człowieku”¹³.

Franciszek zachęcił do inicjatyw edukacyjnych i prawnych, które będą przeciwdziałać dezinformacji w nowym kontekście komunikacyjnym. „Dlatego też bardzo godne pochwały są inicjatywy edukacyjne, które pozwalają nauczyć się, jak czytać i oceniać kontekst komunikacyjny, ucząc by nie być nieświadomymi propagatorami dezinformacji, ale przyczyniać się do jej odkrycia. Równie godne pochwały są inicjatywy instytucjonalne i prawne, starające się określić regulacje mające na celu powstrzymanie tego zjawiska (...)”¹⁴. Papież inspirował również do propagowania dziennikarstwa pokoju, które będzie wolne od udawania, fałszu i przemocy słownej, a jednocześnie będzie poszukiwało zrozumienia zjawisk oraz popularyzowało tłumaczenie dziejących się procesów¹⁵.

Z kolei w posynodalnej adhortacji apostolskiej *Christus vivit* Franciszek zwrócił uwagę na zagrożenia związane z fake newsami i niewłaściwym korzystaniu z nowoczesnych środków przekazu, które mogą skutkować samotnością, manipulacją, przemocą, izolacją, postępującym brakiem kontaktu z rzeczywistością, trudnościami w relacjach czy cyberprzemocą. „Rozpowszechnianie fake news jest wyrazem kultury, która zatraciła poczucie prawdy i nagina fakty do interesów partykularnych. Reputacja ludzi jest zagrożona poprzez procesy doraźne, prowadzone online. Zjawisko to dotyczy także Kościoła i jego pasterzy”¹⁶.

Natomiast w Orędziu na Światowy Dzień Środków Społecznego przekazu w roku 2020 Papież wspominał o sieci komunikacyjnej, w której mogą powstawać destrukcyjne historie prowadzące do niszczenia relacji. Użył jednocześnie pojęcia deepfake, które staje się coraz bardziej znane, osiąga zastraszający poziom i niesie za sobą wyrafinowany poziom manipulacji obrazem i dźwiękiem¹⁷.

¹³ Messaggio del Santo Padre Francesco per la 52ª Giornata Mondiale delle Comunicazioni Sociali, 2.

¹⁴ Ibid.

¹⁵ Ibid, 4.

¹⁶ Francesco, *Christus vivit*, 89.

¹⁷ Messaggio del Santo Padre Francesco per la 54ª Giornata Mondiale delle Comunicazioni Sociali.

W 2022 roku, podczas audycji dla międzynarodowego konsorcjum mediów katolickich „Catholic fact-checking”, Franciszek mówił o fake newsach w kontekście Covid-19. Przedstawiając temat wspominał, że obok pandemii rozprzestrzenia się „infodemia”, czyli zniekształcanie rzeczywistości oparte na lęku. Prawem człowieka jest otrzymywanie sprawdzonych informacji bazujących na danych naukowych, a nie rozprzestrzanianych w postaci fake newsów. Zachęcił również, aby pracownicy mediów nigdy nie zmęczeni się weryfikowaniem danych i poszukiwaniem prawdy¹⁸.

W tym samym roku Papież skierował przesłanie do uczestników obrad międzynarodowej organizacji SIGNIS, która działa w 140 krajach na całym świecie. Temat spotkania brzmiał: Pokój w świecie cyfrowym. Franciszek przypomniał, że niekiedy media cyfrowe stają się miejscami toksycznymi, obszarem mowy nienawiści i fake newsów. Zachęcił do odpowiedniej edukacji medialnej i rozwijania zmysłu krytycznego, dzięki którym społeczeństwo będzie mogło przeciwdziałać kłamstwu i dezinformacji¹⁹.

Zakończenie

Fake news i deepfake to stosunkowo nowe elementy globalnej układanki komunikacyjnej, a profesjonaliści w dziedzinie manipulacji chętnie sięgają po te narzędzia. Efektem końcowym tych działań jest dezinformacja, która czasami przybiera gigantyczne rozmiary, wprowadzając w błąd nie tylko jednostki, ale całe społeczeństwa. Według raportu ITAL Communications – CENSIS z 2023 roku, 76,5% Włochów uważa, że fałszywe wiadomości są coraz bardziej wyrafinowane i trudne do wykrycia, 20,2% uważa, że nie posiada umiejętności niezbędnych do ich zdemaskowania, a 61,1% uważa, że posiada tylko niektóre z nich. Tylko mniejszość (18,7%) jest przekonana, że jest w stanie natychmiast rozpoznać fałszywkę²⁰.

Z pewnością sztuczna inteligencja przynosi wiele korzyści w świecie marketingu oraz mediów, natomiast używanie jej do manipulacji obrazem wideo jest nowym wyzwaniem dla świata komunikacji. Niewłaściwe i nieetyczne używanie tego narzędzia może prowadzić do ogromnego zamieszania wśród odbiorców przekazów medialnych. Dodatkowo sztuczna inteligencja przynosi zmianę paradygmatu i modyfikuje procesy w wielu sektorach (edukacja, zdrowie, finanse, technologia, zakupy). Stąd tak ważne jest pogłębianie znajomości zjawiska fake news i deepfake oraz dążenie do prawdy w przekazach informacyjnych. Również Kościół podejmuje konkretne działania w celu zapobieżenia negatywnym skutkom tych procesów; w związku z tym ogromnym wyzwaniem staje się edukacja medialna.

¹⁸ Address of His Holiness Pope Francis to participants in the meeting promoted by the international consortium of Catholic Media "Catholic fact-checking", Vatican 2022, <https://tiny.pl/cwbr7>.

¹⁹ Message of His Holiness Pope Francis to the Participants in The Signis World Congress, Vatican 2022, <https://tiny.pl/cwbrw>.

²⁰ Disinformazione e fake news in Italia, Terzo rapporto ITAL Communications – CENSIS, 2023, 5-6.

Bibliografia

Address of His Holiness Pope Francis to participants in the meeting promoted by the international consortium of Catholic Media "Catholic fact-checking", Vatican 2022, <https://tiny.pl/cwbr7>.

Almanza Andres (2021). El poder del algoritmo y la vida social. *Sistemas*, (161), 24-47, <https://sistemas.acis.org.co/index.php/sistemas/article/view/166>.

Arencibia Mario González, Cardero Dagmaris Martínez (2021). Soluciones educativas frente a los dilemas éticos del uso de la tecnología deep fake. *Revista Internacional De Filosofía Teórica Y Práctica*, 1(1), 99-126. <http://riftp.editic.net/index.php/riftp/article/view/22>.

Botha Johnny, Pieterse Heloise, Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security, 2020.

Campbell Colin., Plangger Kirk, Sands Sean, Kietzmann Jan (2021) *Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising*. Journal of Advertising. <https://doi.org/10.1080/00913367.2021.1909515>.

Chyliński Marek, Falszywe wiadomości – antydobra w ekosystemie informacji, „Com.press. Pismo naukowe. Komunikacja i nauki o mediach. Journal of communication and media science”, 4 (2018), 9-10.

David Hunt, Derek Robertson, Allison Pow, The Counselor's role in the age of social media and fake news, *Journal of Creativity in Mental Health*, 13:4 (2018), 405-417.

Disinformazione e fake news in Italia, Terzo rapporto ITAL Communications – CENSIS, 2023.

El Mundo, Tomasso Debenedetti, el 'Rey de las fake news' que ha 'matado' a J. K. Rowling, Vargas Llosa, Isabel Allende o el Papa Benedicto, <https://www.elmundo.es/f5/descubre/2022/11/30/63872f01fc6c8318108b45a6.html>.

Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union.

Fadia Hussein, Hussin Hejase, Artificial Intelligence and Fake News, *Žurnalistikos tyrimai*, 16 (2023), 39-71.

Francesco, Christus vivit, Vaticano 2019, https://www.vatican.va/content/francesco/it/apost_exhortations/documents/papa-francesco_esortazione-ap_20190325_christus-vivit.html.

Francisco José García-Ull, Deepfakes: el próximo reto en la detección de noticias falsas, *Anàlisi: Quaderns de Comunicació i Cultura*, 64 (2021), 103-120.

Giansiracusa Noah (2021). *How Algorithms Create and Prevent Fake News: Exploring the Impacts of Social Media, Deepfakes, GPT-3, and More*, Apress, USA, <https://link--springer--com.us.debiblio.com/book/10.1007%2F978-1-4842-7155-1>.

Gomes-Gonçalves Sónia, Deepfakes: a new form of corporate disinformation, *IROCAMM - International Review Of Communication And Marketing Mix*, 5 (2) 2022, 22-38.

Inesdi, Deep Fake y la manipulación de procesos, <https://tiny.pl/cwbrl>.

Le Duc Anthony, Catholic Church Communication in the Post-Truth Era: Intra-Religious and Inter-Religious Dimensions, 2019, <https://ssrn.com/abstract=3335641>.

Magallón Rosa, R. (2019). 'Unfaking News'. *Cómo combatir la desinformación*. Madrid. Pirámide.

Maldita.es. (2021). Por qué los 'deepfakes' no son el problema (aún) y sí los 'cheapfakes'. 21/01/2021. <https://maldita.es/malditobulo/20210121/por-que-los-deepfakes-no-son-el-problema-aun/>.

Message of His Holiness Pope Francis to the Participants in The Signis World Congress, Vatican 2022, <https://tiny.pl/cwbrw>.

Messaggio del Santo Padre Francesco per la 52ª Giornata Mondiale delle Comunicazioni Sociali, https://www.vatican.va/content/francesco/it/messages/communications/documents/papa-francesco_20180124_messaggio-comunicazioni-sociali.html.

Messaggio del Santo Padre Francesco per la 54ª Giornata Mondiale delle Comunicazioni Sociali, https://www.vatican.va/content/francesco/it/messages/communications/documents/papa-francesco_20200124_messaggio-comunicazioni-sociali.html.

Paris Britt, Donovan Joan, (2021). *Deepfakes and cheap fakes*. Thousand Oaks: Sage (=Data & Society's Media Manipulation research initiative). <https://datasociety.net/library/deepfakes-and-cheap-fakes/>.

Raport Newseria, Fake news, czyli jak kłamstwo rządzi światem, Wrzesień 2017.

Sample Ian, What are deepfakes – and how can you spot them?, The Guardian, <https://tiny.pl/cwbrc>.

Vizoso Ángel, Vaz-Álvarez Martín, López-García Xosé, (2021). Fighting deepfakes: Media and internet giants' converging and diverging strategies against hi-tech misinformation. *Media and Communication*, 9(1), 291–300. <https://doi.org/10.17645/MAC.V9I1.3494>.