

Krystian MULARZ 

ORCID: 0000-0002-5011-6576. *Mgr, Uniwersytet Rzeszowski, Kolegium Nauk Społecznych, Instytut Nauk Socjologicznych; Szkoła Doktorska Uniwersytetu Rzeszowskiego, al. T. Rejtana 16C; 35-959 Rzeszów; e-mail: krystianmu@dokt.ur.edu.pl;*

data złożenia tekstu do Redakcji DI: 3.04.2024; data wstępnej oceny artykułu: 10.04.2024

DEEPPFAKE JAKO GENERATOR FIKCYJNYCH TREŚCI I RZECZYWISTEJ NIEUFNOŚCI

DEEPPFAKE AS A GENERATOR OF FICTITIOUS CONTENT AND REAL DISTRUST

Słowa kluczowe: deepfake, dezinformacja, fakenews, manipulacja, Internet.

Keywords: deepfake, disinformation, fakenews, manipulation, Internet.

Streszczenie

Sztuczna inteligencja stała się tematem dyskursu naukowego ze względu na nieoczekiwane konsekwencje jej rozwoju i sposobów wykorzystania. Technologie AI mogą spełniać zarówno funkcje rozrywkowe i edukacyjne, jak i stwarzać zagrożenie w postaci szerzenia dezinformacji oraz dyskredytowania innych osób w przestrzeni publicznej. Jedną z niewłaściwych form wykorzystania sztucznej inteligencji jest tzw. *deepfake*. Niniejsza praca ma na celu omówienie pojęcia *deepfake* na podstawie literatury przedmiotu oraz analizy treści zawartej w sieci. Autor szczegółowo scharakteryzował pojęcie *deepfake*, omówił przykłady fikcyjnych treści oraz społeczne skutki ich publikowania. Ponadto dokonał refleksji na temat wpływu szerzącej się dezinformacji i zagrożeń z nią związanych na zanik zaufania, co skutkuje przekształceniem się kultury w taką, w której dominuje nieufność i podejrzliwość.

Abstract

Artificial intelligence has become a topic of scientific discourse due to the unexpected consequences of its development and uses. AI technologies can serve both entertainment and educational functions, as well as pose a threat in the form of spreading disinformation and discrediting others in public spaces. One of the inappropriate uses of artificial intelligence is the so-called deepfake. This paper aims to discuss the concept of deepfake based on the literature on the subject and the analysis of the content contained in the network. The author characterized the concept of deepfake in detail, discussed examples of fictitious content and the social consequences of its

publication. In addition, the author reflected on the impact of the spread of disinformation and the risks associated with it on the disappearance of trust, resulting in the transformation of culture in which distrust and suspicion prevail.

Wstęp

Nowe technologie nieustannie zmieniają świat społeczny, implikując wiele innowacyjnych rozwiązań, które jednocześnie stawiają przed ludźmi nowe wyzwania. Dzięki rozwojowi techniki, która umożliwiła nieograniczoną odległością lub porą dnia komunikację z innymi, ludzkość przełamała bariery czasowe i przestrzenne. Jan van Dijk zjawisko to nazywa „rewolucją komunikacyjną”, odnosząc się do stopniowego postępu technologicznego w zakresie możliwości interakcyjnych współczesnego społeczeństwa¹. Największy wpływ na postęp komunikacyjny wywarł Internet, którego najintensywniejszy rozwój datuje się na przełom XX i XXI w. To wówczas powstawały pierwsze blogi internetowe, YouTube, Twitch.tv oraz pierwsze portale społecznościowe, takie jak Facebook, Twitter lub Instagram². Pojawienie się tych platform oraz wzrost dystrybucji urządzeń mobilnych umożliwiło jednostkom komunikowanie się niezależnie od okoliczności, podnosząc tym samym znaczenie komunikacji zapośredniczonej i wspólnot wirtualnych³. Van Dijk upatruje w tym szansy na pogłębienie więzi w sieciach społecznych poprzez wzajemne wzmocnianie się kontaktów bezpośrednich i wirtualnych między ludźmi⁴. Inni z kolei dostrzegają w tym zagrożenia z powodu m.in. negatywnego wpływu nadmiernej częstotliwości korzystania z sieci na relacje w świecie rzeczywistym i wzrastające poczucie samotności⁵. Alvin Toffler relacje wirtualne określa wręcz mianem iluzji, która jedynie pogłębi zjawisko samotności⁶.

¹ J. van Dijk, *Społeczne aspekty nowych mediów*, przekł. J. Konieczny, Wydawnictwo Naukowe PWN, Warszawa 2020, s. 14.

² K. Mularz, *Media społecznościowe a budowanie wizerunku i relacji międzyludzkich* [w:] *Świat (po)cyfrowej rewolucji. W kręgu analiz socjologicznych*, red. A. Kotowska, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów 2019, s. 97–98.

³ J. van Dijk, *Społeczne aspekty...*, s. 234.

⁴ Tamże

⁵ A. Januszewska et al., *Uzależnienie od Internetu oraz samotność wśród studentów Warszawskiego Uniwersytetu Medycznego* [w:] *Wyzwania współczesnego pielęgniarstwa*, red. M. Wysokiński, Z. Sienkiewicz, W. Fidecki, Warszawski Uniwersytet Medyczny, Warszawa 2021, s. 105–115.

⁶ B. Piechota, *O atrakcyjności sieci* [w:] *Paradoksy Internetu. Konteksty społeczno-kulturowe*, red. M. Szpunar, Toruń 2011, s. 122–123.

Rozwój technologiczny wpłynął również na proces globalizacji społeczeństwa, znacząco ułatwiając przepływ informacji i wymianę międzykulturową⁷. Globalizacja charakteryzuje się wysokim stopniem urbanizacji i przyrostem populacji⁸. Sieć ludzka nie tyle uległa rozszerzeniu, co gęstnieniu poprzez wzrost natężenia i prędkości komunikacji. Manuel Castells używa terminu społeczeństwa informacyjnego, w którym „przetwarzanie i transmisja informacji staje się fundamentalnym źródłem produktywności i władzy”⁹.

Pojawienie się nowych środków komunikacji ułatwiło jednak nie tylko szeroko pojęty rozwój i przepływ wiedzy, ale także szerzenie dezinformacji. Badacze dla opisu współczesnej rzeczywistości wykorzystują termin „Świat VUCA”, który oznacza społeczeństwo cechujące się destabilizacją i lękiem społecznym spowodowanym tempem, chaotycznością i ciągłością zmian¹⁰. Do niedawna niepewna sytuacja związana z pandemią COVID-19 lub wciąż trwający konflikt zbrojny w Ukrainie to przykłady zmian społecznych, których długofalowe skutki są trudne do przewidzenia. To z kolei budzi społeczny niepokój ze względu na trudności w planowaniu i poszukiwaniu rozwiązań dla nowych wyzwań wynikających z niespodziewanej sytuacji społecznej¹¹.

Dodatkowe utrudnienie w tym zakresie stanowi fakt, że wiele informacji, które docierają do przeciętnego człowieka, jest nieprawdziwych, a nowe technologie wykorzystywane są celowo do szerzenia dezinformacji i społecznej paniki. Rozwój sztucznej inteligencji sprawił, że weryfikacja danych jest dla odbiorcy coraz trudniejsza ze względu na szerokie spectrum możliwości oferowane przez technologię. Współcześnie coraz trudniejsze wydaje się odróżnienie informacji oficjalnej, pochodzącej z rzetelnego źródła, od wymysłu mniej lub bardziej zorganizowanych grup mających na celu wzbudzenie niepokoju w społeczeństwie, osiągnięcia własnych korzyści bądź ośmieszenia innych. Trzeba bowiem zaznaczyć, że manipulacja dotyczy nie tylko ważnych społecznie spraw, jak chociażby wspomniany konflikt w Ukrainie, ale dotyka również wizerunku jednostek, stwarzając tym samym zagrożenie zarówno w kontekście bezpieczeństwa, jak i ochrony godności ludzkiej.

⁷ P. Sztompka, *Socjologia. Wykłady o społeczeństwie*, Znak Horyzont, Kraków 2021, s. 772–774.

⁸ J. van Dijk, *Społeczne aspekty...*, s. 40.

⁹ M. Castells, *Społeczeństwo sieci*, tłum. M. Marody et al., Wydawnictwo Naukowe PWN, Warszawa 2007, s. 36.

¹⁰ E. Dybowska, *Pedagogika ignacjańska w świecie VUCA*, „Horyzonty Wychowania” 2022, nr 21 (57), s. 13.

¹¹ K. Mularz, *Znaczenie przyjaźni we współczesnym świecie*, „Społeczeństwo. Edukacja. Język” 2023, nr 18, s. 98–99.

Celem niniejszego artykułu jest opis jednej z technik manipulowania informacją określanej mianem *deepfake*. Autor artykułu chce przybliżyć czytelnikowi zagrożenia z tym związane poprzez szczegółowe wyjaśnienie pojęcia *deepfake* oraz wskazanie przykładów zastosowania tej techniki. Argumenty przedstawione w artykule oparte są na literaturze przedmiotu i własnych obserwacjach.

Deepfake – charakterystyka pojęcia

Pojęcie *deepfake* jest stosunkowo nowym zjawiskiem opisywanym w literaturze naukowej. Oznacza wykorzystywanie sztucznej inteligencji do manipulacji obrazem poprzez nakładanie różnych materiałów¹². Nazwa wywodzi się z forum internetowego Reddit, na którym w roku 2017 użytkownik o nicku „deepfakes” opublikował film pornograficzny z udziałem m.in. aktorki Gal Gadot¹³. Materiał wideo został spreparowany przez wspomnianego użytkownika, który poprzez zastosowanie nowoczesnego oprogramowania umieścił twarz kobiety w dostępnych treściach pornograficznych. W ten sposób zapoczątkował niepokojący trend w Internecie – odtąd bowiem pojawiało się coraz więcej fałszywych treści z twarzami osób publicznych, które wykorzystywano zarówno w celach satyrycznych, jak i z zamiarem zdeprecjonowania danej jednostki lub zyskania poparcia dla konkretnej idei¹⁴.

Tworzenie materiałów typu *deepfake* polega na wykorzystaniu sztucznej inteligencji w zakresie techniki głębokiego uczenia się. Sztuczna inteligencja to, według Witolda Marciszewskiego, „kolosalna moc obliczeniowa komputera, biorąca się z doskonałości sprzętu i programów”¹⁵. Pojęcie to zostało zastosowane w celu wzbudzenia w ludzkości refleksji na temat tempa rozwoju technologii i jej oddziaływania na życie człowieka. Jak pisze dalej autor: „Zwrot «sztuczna inteligencja» ukazuje coś, czego laik nie wydedukuje z obiegowego pojęcia komputera. Mianowicie, że inne niż rachowanie czynności inteligentne można również powierzyć maszynie – pod warunkiem, że uda się je sprowadzić do rachowania”¹⁶. Dylan W. Wheeler opisuje AI (z ang. *artificial intelligence*) jako inteligencję wykazywaną przez maszyny w postaci naśladowania funkcji

¹² O. Wasiuta, S. Wasiuta, *Deepfake jako skomplikowana i głęboko fałszywa rzeczywistość*, „Studia de Securitate” 2019, nr 9(3), s. 20.

¹³ I. Dąbrowska, *Deepfake – nowy wymiar internetowej manipulacji*, „Zarządzanie Mediami” 2020, t. 8(2), s. 91.

¹⁴ Tamże.

¹⁵ W. Marciszewski, *Sztuczna inteligencja*, Znak, Kraków 1998, s. 8.

¹⁶ Tamże.

poznawczych charakteryzujących człowieka, takich jak zdolność uczenia się i rozwiązywanie problemów¹⁷.

Przykładem takiej funkcji poznawczej, stanowiącej wspólny mianownik między człowiekiem a maszyną, jest właśnie technika głębokiego uczenia się. W literaturze w odniesieniu do człowieka technika ta definiowana jest jako proces, który umożliwia jednostce przyswajanie wiedzy i jej późniejsze wykorzystanie w praktyce¹⁸. Oznacza to, że poprzez pogłębioną edukację człowiek zyskuje wiedzę specjalistyczną w różnych dziedzinach. Technika głębokiego uczenia się pozwala jednostce nie tylko zapamiętać fakty, ale także zrozumieć, jak i w jakim celu wykorzystać zdobytą wiedzę oraz w jakich konkretnie sytuacjach zastosowanie tej wiedzy będzie skuteczne i pozwoli rozwiązać nowe problemy¹⁹.

W odniesieniu do sztucznej inteligencji technika głębokiego uczenia się jest metodą „nauczania” maszyny zainspirowaną strukturą ludzkiego mózgu²⁰. Mózg człowieka składa się z komórek nerwowych połączonych za pomocą aksonów, dendryt oraz synaps, które razem tworzą rozbudowaną sieć połączeń. Oprócz oczywistego zadania, jakim jest podtrzymywanie funkcji życiowych, komórki te poprzez liczne połączenia odbierają, przekazują i przetwarzają informacje, odpowiadając za ich zapamiętywanie²¹. Działanie mózgu ma charakter dynamiczny, ponieważ na jakość przetworzonych w sieci neuronowej informacji oddziałuje stan ludzkiej pamięci. Przykładem takiej dynamiki jest sytuacja, kiedy jednostka spotyka na swojej drodze nieznanego, wobec którego z powodu braku informacji na jego temat przejawia inne zachowania niż podczas kolejnych spotkań, kiedy posiada więcej danych wynikających z wcześniejszych interakcji²².

Osiągnięcia technologiczne umożliwiły utworzenie wysoko rozwiniętych oprogramowań „samouczących się” w postaci sztucznych sieci neuronowych, w których proces przetwarzania informacji odbywa się na podobnej zasadzie jak w mózgu człowieka²³. Oznacza to, że komputery posiadają elementy ludzkiej

¹⁷ D.W. Wheeler, *Artificial Intelligence and Its Disinformation Campaigns*, 2019, s. 3–4, <https://www.dylanwheeler.net/files/Disinformation.pdf> [dostęp: 27.03.2023].

¹⁸ E. Musiał, *Nowe trendy w edukacji – koncepcja głębokiego uczenia się*, „Zeszyty Naukowe Wyższej Szkoły Humanitas. Pedagogika” 2018, nr 16, s. 57.

¹⁹ Tamże.

²⁰ D.W. Wheeler, *Artificial Intelligence...*, s. 4.

²¹ Z. Stęgowski, *Sztuczne sieci neuronowe*, „Kernel.” 2004, nr 1(1), s. 17.

²² Tamże.

²³ Zdzisław Stęgowski zaznacza, że człowiek jest w zakresie zapamiętywania doskonalszy od maszyny. Wynika to z rozproszenia pamięci konkretnej informacji w różnych komórkach, co sprawia, że uszkodzenie jakiejś komórki nie jest równoznaczne z utratą danych. Z kolei w kompu-

inteligencji, za pomocą których można „uczyć” maszyny wykonywania zadań poprzez systematyczne gromadzenie danych²⁴. Tak jak człowiek, który dzięki nabytej wiedzy i własnemu doświadczeniu jest w stanie sprostać coraz trudniejszym wyzwaniom, tak również oprogramowania poprzez przetwarzanie danych są w stanie efektywniej spełniać swoje funkcje.

Tworzenie materiałów typu *deepfake* odbywa się właśnie na takiej zasadzie. Przytaczając wspomniany przykład nieetycznego wykorzystania wizerunku Gal Gadot, programowi komputerowemu dostarcza się sporą ilość materiałów w postaci zdjęć, nagrań dźwiękowych oraz wideo z udziałem aktorki. Na ich podstawie program rozpoczyna „nauczę” naśladowania kobiety, tak aby w jak najwierniejszy sposób oddać rzeczywiste jej zachowania, tonację głosu, mimikę i indywidualne cechy charakterystyczne wyróżniające ją spośród innych ludzi²⁵. W ten sposób program jest w stanie wygenerować fałszywe nagranie na podstawie autentycznych materiałów z udziałem konkretnej osoby, a także innych dostępnych w sieci treści. To stwarza przeciętnemu użytkownikowi możliwość wygenerowania nagrania o dowolnej treści i z dowolną osobą, o ile tylko dostarczy on oprogramowaniu wystarczająco dużo materiału na temat danej jednostki – im więcej źródeł, z których może skorzystać program, tym bardziej realny efekt, który jest trudny do zweryfikowania przez odbiorców. Do uwiarygodnienia *deepfake’ów* wykorzystuje się tzw. generatywne sieci przeciwstawne GAN (z ang. *Generative Adversarial Networks*). Sieci te składają się z dwóch podstawowych elementów: generatora oraz dyskryminatora. Zadaniem generatora jest tworzenie treści poprzez wielostopniowe filtrowanie i scalanie różnych elementów, na podstawie których ma powstać docelowy materiał. Dyskryminator z kolei na podstawie wzorców i dostępnych danych weryfikuje utworzone przez generatora treści. Proces tworzenia fikcyjnych materiałów kończy się sukcesem, gdy dyskryminator uzna, że utworzone nagranie jest rzeczywiste lub wysoce prawdopodobne²⁶.

Iłona Dąbrowska, dokonując analizy przykładów spreparowanych nagrań i obrazów, zaproponowała typologię materiałów typu *deepfake* na podstawie pierwotnego celu, w którym zostały wyprodukowane. W ten sposób autorka wyodrębniła cztery podstawowe kategorie *deepfake’ów*: rozrywkowe, edukacyjne, dezinformacyjne oraz dyskredytacyjne²⁷.

terach informacje zapamiętywane są w konkretnych nośnikach pamięci, których uszkodzenie często oznacza nieodwracalną utratę danych (zob. Z. Stęgowski, *Sztuczne sieci...*, s. 17–18).

²⁴ K. Różanowski, *Sztuczna inteligencja: rozwój, szanse i zagrożenia*, „Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki” 2007, nr 2(2), s. 112.

²⁵ O. Wasiuta, S. Wasiuta, *Deepfake...*, s. 21.

²⁶ I. Okulska, *Sieci neuronowe typu GAN i GPT-2, słowa zużyte i kreatywność, czyli literacki second-hand*, „Forum Poetyki” 2019, nr 18, s. 27–28.

²⁷ I. Dąbrowska, *Deepfake...*, s. 96.

Celem rozrywkowych treści, jak sama nazwa wskazuje, jest przede wszystkim dostarczenie przyjemności i humoru poprzez odniesienie do szeroko pojętej kultury, a bohaterami fałszywych nagrań są zarówno osoby fikcyjne, jak i aktorzy oraz inne osoby publiczne²⁸. Przykładem treści tego typu może być przerebiony przez AI utwór pt. „King Bruce Lee karate mistrz”, który po edycji jest „wykonywany” przez Tadeusza Sznuka²⁹. Innym przykładem jest fragment z filmowej adaptacji powieści J.R.R. Tolkiena pt. „Władca Pierścieni”, w którym wszystkim aktorom realnie występującym w danej scenie wstawiono twarz oraz podłożono głos Arnolda Schwarzenegera³⁰.

Deepfake edukacyjny ma za zadanie dostarczanie wiedzy i istotnych informacji oraz zachęcenie odbiorców do ich pogłębiania³¹. Taki cel spełnia np. film przedstawiający wygenerowany przez AI wizerunek Alberta Einsteina, który rzetelnie odpowiada na zadane pytania³².

Zdecydowanie bardziej niebezpiecznym rodzajem są treści dezinformacyjne. Powodem tworzenia tego typu nagrań jest chęć wzbudzenia niepokoju i lęku społecznego poprzez propagowanie fałszywych informacji. Przykładem są przekazy Federacji Rosyjskiej i oskarżenie Ukrainy m.in. o mistyfikację wydarzeń w Mariupolu, gdzie zbombardowany został szpital z ludnością cywilną. Według rosyjskich przekazów szpital nie funkcjonował, a rzekomym celem ataku było przeciwdziałanie przebywającemu tam radykalnemu oddziałowi „Azow”. Rosjanie w celu dezinformacji wykorzystali wizerunek influencerki Marianny Wyszomirskiej³³.

Ostatnim rodzajem materiałów typu *deepfake* wskazanym przez Ilonę Dąbrowską są treści dyskredytacyjne. Ich celem jest osłabienie pozycji i kreowanie negatywnej opinii publicznej wokół danej osoby. Najczęściej dotyczy to sceny politycznej³⁴. Przykładem jest spreparowane nagranie, w którym głos Joe Bidena zniechęca obywateli do udziału w prawyborach w stanie New Hampshire³⁵.

²⁸ Tamże.

²⁹ Youtube.com, *Tadeusz Szuk – King Bruce Lee karate mistrz | AI Cover*, <https://www.youtube.com/watch?v=QW22vGPRvuo> [dostęp: 28.03.2024].

³⁰ Youtube.com, *The Schwarzenegers of the Ring Part 2*, <https://www.youtube.com/watch?v=Fko4z7lADw0> [dostęp: 27.03.2024].

³¹ I. Dąbrowska, *Deepfake...*, s. 96.

³² Youtube.com, *Einstein AI – Interactive Virtual Being*, <https://www.youtube.com/watch?v=c0o5s0DZeJl> [dostęp: 28.03.2024].

³³ Stowarzyszenie Demagog, *Ostrzał szpitala w Mariupolu był ustawką? Fake news!*, https://demagog.org.pl/fake_news/ostrzal-szpitala-w-mariupolu-by-l-ustawka-fake-news/ [dostęp: 28.03.2024].

³⁴ I. Dąbrowska, *Deepfake...*, s. 96.

³⁵ Bankier.pl, *Kolejny deepfake zebrał żniwa. Tym razem celem był Joe Biden*, <https://www.bankier.pl/wiadomosc/Kolejny-deepfake-zebral-zniwa-Tym-razem-celem--byl-Joe-Biden-8701675.html> [dostęp: 28.03.2024].

Zdaniem autora artykułu zaproponowana typologia nie wyczerpuje w pełni zagadnienia *deepfake*. Wiele z fikcyjnych produkcji można bowiem przyporządkować do więcej niż jednej kategorii. Przykładowo przeróbki z wizerunkiem Tadeusza Sznuka, który „wykonuje” popularne utwory muzyczne, może z jednej strony pełnić funkcję rozrywkową, z drugiej zaś dyskredytować bohatera tych nagrań. Oczywiście w tym przypadku zależy to od subiektywnych odczuć osoby, z której wizerunku skorzystano. Autor nie sądzi bowiem, że twórcy fikcyjnych nagrań mieli na celu osłabienie wizerunku Tadeusza Sznuka, który cieszy się dużą popularnością i szacunkiem w kraju. Niemniej, jeśli sam bohater czułby się urażony wykorzystaniem jego wizerunku do takich celów i nie odpowiadałoby to jego koncepcji tworzenia własnego PR-u, to należałoby rozważyć czy taki *deepfake* nie spełnia także funkcji dyskredytacyjnej.

Oczywiście Ilona Dąbrowska zaznacza, że zaproponowana typologia utworzona została na podstawie celu autorów tych nagrań, a nie odczucia osób z nimi związanych. Wydaje się jednak, że w niektórych sytuacjach trudno jest jednoznacznie zdefiniować taki cel. Na myśl przychodzą fikcyjne filmy pornograficzne z wizerunkiem osób, które nigdy nie wzięły w nich udziału, jak Scarlett Johansson, Emma Watson czy wspomniana wcześniej Gal Gadot. Należy bowiem zadać sobie pytanie, czy zostały one utworzone i opublikowane wyłącznie w celu dyskredytacji bohaterek nagrań, czy może cechują się dodatkową funkcją, jaką jest spełnienie własnych, egoistycznych fantazji seksualnych ich twórców? Oczywiście to nadal jest bardzo niepokojące, jednak należy brać pod uwagę taką ewentualność, choćby z uwagi na rosnącą popularność programów służących do generowania fikcyjnych postaci. Pod hashtagiem #aigirls lub #aierotica w aplikacji Twitter można znaleźć profile publikujące wygenerowane przez sztuczną inteligencję wizerunki kobiet z wyraźnym podtekstem seksualnym. Przykładami takich profili są AI Corner Hotties³⁶, kira_ai_gf³⁷ lub AI_Girls_Studio³⁸. Ten ostatni posiada nawet odnośnik do serwisu Patreon, na którym za odpowiednią opłatą można otrzymać treści niedostępne za darmo – wśród zablokowanych dla niesubskrybujących profilu kategorii są takie jak „Nude” lub „Breast”, których nazwy wyraźnie wskazują na treści, jakie można tam znaleźć³⁹. Oznacza to więc, że sztuczna inteligencja jest pewnego rodzaju rewolucją w kontekście kanonów piękna i seksualności – tworzy bowiem wize-

³⁶ Profil AI Corner Hotties na portalu Twitter, <https://twitter.com/AiCorner54628> [dostęp: 28.03.2024].

³⁷ Profil kira_ai_gf na portalu Twitter, https://twitter.com/kira_ai_gf [dostęp: 28.03.2024].

³⁸ Profil AI_Girls_Studio na portalu Twitter, https://twitter.com/AI_Girls_Studio [dostęp: 28.03.2024].

³⁹ Profil AI_Girls_Studio na portalu Patreon, <https://www.patreon.com/AIGirlsStudio> [dostęp: 28.03.2024].

runki kobiet wygenerowane na podstawie udostępnionych wcześniej zdjęć, które w rezultacie mają spełniać indywidualne preferencje użytkownika. To sprawia również, że fikcyjne materiały pornograficzne wykorzystywane są do generowania zysków (!), co widać na przykładzie wspomnianego profilu AI_Girls_Studio. Zdaniem Alexa Valatisa, który zajmuje się analizowaniem wpływu sztucznej inteligencji na życie społeczne, do 2025 r. większość kont na portalu OnlyFans będzie profilami przedstawiającymi sztucznie wygenerowane kobiety, za których zdjęcia nieświadomi tego (lub świadomi!) użytkownicy będą płacić pieniądze w ramach subskrypcji, a za tymi profilami w rzeczywistości stać będą mężczyźni⁴⁰.

Z tego względu zdaniem autora artykułu należy nieco poszerzyć lub doprecyzować zaproponowaną kategorię *deepfake*'ów. Biorąc pod uwagę przytoczone profile generujące zyski poprzez udostępnianie zdjęć roznegliżowanych fikcyjnych kobiet, należy uwzględnić również cel dochodowy takich produkcji. Ponadto ze względu na popularność materiałów pornograficznych generowanych przez sztuczną inteligencję, zarówno z wykorzystaniem wizerunku prawdziwych osób, jak i fikcyjnych postaci, autor sugeruje, że być może warto wyodrębnić osobną kategorię, którą byłby *deepfake* erotyczny.

Aleksander Piecuch wyróżnił kilka sposobów wykorzystania materiałów typu *deepfake*:

- rozrywka – celem jest edycja filmów lub obrazów z wykorzystaniem przerobionych lub podmienionych twarzy innych osób;
- kreacja wirtualnych postaci – tworzenie fikcyjnych wizerunków na bazie materiałów źródłowych, np. przedstawienie w chińskich mediach prezentera telewizyjnego wygenerowanego przez AI;
- *fake news* – publikowanie nieprawdziwych informacji, których rzekomymi autorami są politycy i inne ważne osoby publiczne;
- pornografia – edycja twarzy aktorów występujących w filmach dla dorosłych poprzez zastąpienie ich wizerunkiem innych osób;
- logowanie, autoryzacja – łamanie zabezpieczeń aplikacji, które wymagają od użytkownika wykonania konkretnych gestów lub wypowiedzenia krótkiej kwestii w celu uzyskania dostępu do konta;
- ataki finansowe – naśladowanie menedżerów poważnych firm i wydawanie poleceń podczas wideorozmów lub konwersacji telefonicznych, np. zlecenie przelewu⁴¹.

⁴⁰ B. Francuz, *AI zamiast modelek. OnlyFans mogą zalać profile kobiet, które nie istnieją – i ludzie za to zapłacą*, <https://android.com.pl/tech/562237-ai-onlyfans-generowane-modelki/> [dostęp: 28.03.2024].

⁴¹ A. Piecuch, *Życie smart – ma swoją cenę*, „Dydaktyka Informatyki” 2020, nr 15, s. 60–61.

Niezależnie od przyjętej perspektywy przedstawiony opis służyć ma czytelnikowi jako źródło wiedzy na temat nowych rozwiązań technologicznych, które poza innowacjami z zakresu edukacji i nowymi, na pozór niegroźnymi funkcjami implikują poważne zagrożenia, które znacząco wpływają na życie społeczne i których nie należy lekceważyć.

Zagrożenia dla wizerunku wynikające z rozwoju materiałów typu *deepfake*

Jednym z programów służących produkcji fikcyjnych treści jest FakeApp. Jego początek datuje się na styczeń 2018 r., kiedy to aplikacja została udostępniona w Internecie z możliwością darmowego jej pobrania. Program ten, wykorzystując opisane wcześniej techniki głębokiego uczenia się, umożliwił przeciętnemu, nieposiadającemu fachowej wiedzy informatycznej użytkownikowi tworzenie fałszywych nagrań jedynie na podstawie dostarczonej grafiki⁴². Do osiągnięcia najlepszych efektów w zakresie wiarygodności nagrania program wymaga ok. 2000–3000 zdjęć, choć samo wygenerowanie filmu nie wymaga tak dużej ilości materiału⁴³. Program ten był wykorzystany do tworzenia nie tylko treści pornograficznych z udziałem znanych osób, ale również do produkcji krótkich filmów humorystycznych, edukacyjnych lub innych różnorodnych materiałów. FakeApp zatem wykorzystywano w różnych celach, zarówno tych zdecydowanie uderzających w godność ludzką, jak i do publikowania treści niejednoznacznych pod tym względem, które mogły przynieść więcej pozytywnych niż negatywnych konsekwencji. Przykładem jest choćby film z wykorzystaniem wizerunku Baracka Obamy, który „przestrzegał” odbiorców przed dezinformacją, zachęcając do ostrożności i weryfikowania danych przetwarzanych w sieci⁴⁴. W rzeczywistości były prezydent Stanów Zjednoczonych nie wypowiedział tych słów, lecz wykorzystanie jego wizerunku miało na celu uwiarygodnić przestrożę, którą właściciele portalu BuzzFeed, gdzie opublikowany został materiał, chcieli przekazać użytkownikom Internetu⁴⁵.

Rozwój technologii wykorzystujących wizerunek innych osób doprowadził jednak do powstania programów służących wyłącznie nieetycznym celom. Przy-

⁴² O. Wasiuta, S. Wasiuta, *FakeApp jako nowe zagrożenie bezpieczeństwa politycznego i informacyjnego*, „Studia de Securitate” 2019, nr 9(3), s. 129.

⁴³ Tamże, s. 130.

⁴⁴ I. Dąbrowska, *Deepfake...*, s. 93.

⁴⁵ Tamże.

kładem takiego programu jest *DeepNude*, który umożliwił usunięcie ubrań z ciał osób widocznych na obrazach⁴⁶. Procedura jest bardzo prosta – wystarczy w aplikacji dodać interesującą użytkownika grafikę, aby następnie, po kliknięciu jednego przycisku, wygenerować jego fikcyjny odpowiednik z nagimi bohaterami⁴⁷. Aplikacja szybko została uznana za niemoralną i szkodliwą, uderzającą zwłaszcza w godność kobiet – 96% materiałów wygenerowanych przez *DeepNude* przedstawiało właśnie kobiety⁴⁸.

Pornografia jest zdecydowanie jedną z najgroźniejszych konsekwencji niewłaściwego wykorzystania sztucznej inteligencji. Choć fikcyjne materiały erotyczne z udziałem osób publicznych zaczęły pojawiać się już kilka lat temu, a aplikacje umożliwiające dokonywanie nieetycznych przeróbek zostały uznane za niemoralne, to nadal istnieje możliwość ich pobrania i wielokrotnego wykorzystywania, a portale społecznościowe, pomimo zakazu publikowania takich treści, nie rozwiązały problemu w pełni. Wystarczy cofnąć się do stycznia 2024 r., kiedy to na Twitterze pojawiły się zdjęcia roznegliżowanej Taylor Swift wygenerowane przez sztuczną inteligencję. Według portalu „The Verge” zdjęcia te zostały usunięte dopiero po przekroczeniu granicy ponad 45 mln ponownych udostępnień przez użytkowników Twittera⁴⁹. Popularność kontrowersyjnych wpisów skutkowało nie tylko ostatecznym usunięciem zidentyfikowanych treści. Przez jakiś czas administracja zablokowała frazę „Taylor Swift” w swojej wyszukiwarce, po której wpisaniu portal generował komunikat o błędzie. Po paru dniach serwis odblokował możliwość wyszukiwania informacji pod daną frazę⁵⁰. Nadal jednak Twitter uniemożliwia wygenerowanie jakichkolwiek wyników poprzez wyrażenie „Taylor Swift AI”, po którego wpisaniu wyświetla się informacja o usterce.

Choć treści w jawny sposób uderzające w osoby publiczne są niepokojące i nie powinny mieć miejsca, to zaznaczyć należy, że takie jednostki, pomimo iż są bardziej narażone na ataki ze strony innych, to jednak dysponują również

⁴⁶ K. Karski, *Korzyści i zagrożenia wynikające z implementacji sztucznej inteligencji*, „Zeszyty Naukowe Akademii Górnośląskiej” 2023, nr 6, s. 63.

⁴⁷ CY. Yeh et al., *Disrupting Image-Translation-Based DeepFake Algorithms with Adversarial Attacks* [w:] *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*, 2020, s. 53.

⁴⁸ K. Karski, *Korzyści i zagrożenia...*, s. 63.

⁴⁹ Stowarzyszenie Demagog, *Pornografia tworzona przez AI, czyli deepfake z Taylor Swift*, https://demagog.org.pl/analizy_i_raporty/pornografia-tworzona-przez-ai-czyli-deepfake-z-taylor-swift/ [dostęp: 28.03.2024].

⁵⁰ billboard.com, *Taylor Swift Searchable on X Again After Brief Blockage Over DeepFake Images*, https://www.billboard.com/business/tech/taylor-swift-searchable-x-twitter-ai-generated-deepfake-explicit-photos-1235592389/?utm_medium=social&utm_source=twitter [dostęp: 28.03.2024].

większymi środkami i możliwościami na zwalczanie oraz dementowanie fikcyjnych treści. Inaczej sprawa wygląda, gdy fałszywe nagrania i obrazy dotyczą osób prywatnych. Aleksander Piecuch przy opisie sposobów wykorzystywania materiałów typu *deepfake* wspomina o zjawisku tzw. *revenge porn*⁵¹. Jest to forma tzw. pornografii w odwecie, która polega na publikowaniu treści erotycznej z udziałem konkretnej osoby. Najczęstszymi sprawcami takiego zabiegu są byli partnerzy, którzy zachowali materiały jeszcze z czasów trwania wspólnego związku. Celem takich czynów jest poniżenie i osłabienie reputacji danej osoby w akcie zemsty⁵². Niektórzy autorzy w tym kontekście używają terminu *cyber revenge*, który stanowi element *seksbullyingu*, czyli publicznego prześladowania innych za pośrednictwem Internetu z wykorzystaniem materiałów pornograficznych⁵³.

Pornografia z zemsty do niedawna była stosowana głównie przez byłych partnerów lub osoby, które w jakiś sposób uzyskały realne treści erotyczne z udziałem jednostek, jakie obrały sobie za cel. W dobie rozwiniętej sztucznej inteligencji, która jest w stanie generować dowolny rodzaj multimediów z dowolnymi osobami, zjawisko *revenge porn* stanowi duże i realne zagrożenie. Taka sytuacja miała miejsce w jednej ze szkół w Londynie. Grupa chłopców prześladowała dziewczęta poprzez publiczne naśmiewanie się z ich zdjęć w portalach społecznościowych. Uczniowie ci posunęli się nawet do wykorzystania wizerunku dziewcząt do spreparowania realnych obrazów pornograficznych, które następnie, już z twarzami swoich koleżanek, publikowali na czacie grupowym. Doprowadziło to do sytuacji, w której jedna z uczennic, 14-letnia Mia Janin, popełniła samobójstwo⁵⁴. Jak wyznała jedna z pozostałych dziewcząt, Mia nie cieszyła się w ostatnich dniach swojego życia popularnością, o czym świadczyły m.in. negatywne komentarze pod jej ostatnim autorskim filmem na TikToku. Fałszywe i poniżające dziewczynę obrazy oraz brak sympatii rówieśników doprowadził w tym przypadku do tragedii⁵⁵. Opisana historia może znamionować problemy, z którymi społeczeństwo będzie zmagać się coraz częściej. Nieustannie rozwijająca się technologia umożliwia użytkownikom preparowanie grafiki z coraz większą precyzją, a to sprawia, że osoby, które padną ofiarą ta-

⁵¹ A. Piecuch, *Życie smart...*, s. 61.

⁵² K. Groszkowska, *Cyberprzemoc wobec kobiet*, Instytut Prawa Ustrojowego, s. 6, https://ipu.org.pl/wp-content/uploads/2021/02/Cyberprzemoc-wobec-kobiet_Kamila-Groszkowska.pdf [dostęp: 28.03.2024].

⁵³ S. Kania, *Seksting – diagnoza i profilaktyka pozytywna* [w:] *Psycholog w szkole: koncepcje*, red. Z.B. Gaś, Innovatio Press, Lublin 2024, s. 157.

⁵⁴ S. Hooper, *Girl killed herself when bullies shared fake nudes of Her*, <https://metro.co.uk/2024/01/24/teen-took-life-online-bullying-shared-fake-nudes-20162284/> [dostęp: 28.03.2024].

⁵⁵ Tamże.

kiej manipulacji, będą miały duże trudności, aby sytuację zdementować i przekonać innych do swojej racji. Stwarza to więc pole do dyskryminacji i piętnowania osób na podstawie nieprawdziwych informacji.

Jak zostało opisane wcześniej, dyskredytowanie innych osób może odbywać się także poprzez przypisywanie im słów, których nigdy nie powiedziały. 28 marca 2024 r. na Twitch.tv odbyła się transmisja na żywo z udziałem fikcyjnych postaci prezentujących Donalda Trumpa oraz Joe Bidena. Transmisja miała charakter debaty przed wyborami prezydenckimi w Stanach Zjednoczonych, podczas której wygenerowane przez sztuczną inteligencję postacie konkurowały ze sobą. W rzeczywistości wypowiediane kwestie były komentarzami widzów obecnych podczas transmisji i komunikujących się za pośrednictwem czatu⁵⁶. Sytuacja ta nosi znamiona rozrywkowego *deepfake* 'a, a może nawet i edukacyjnego, biorąc pod uwagę to, że poza żartobliwymi wypowiedziami użytkowników pojawiały się również merytoryczne dyskusje na temat przyszłości Stanów Zjednoczonych, co mogło skłonić widzów do większego zainteresowania się sprawami swojego kraju. Niemniej należy zaznaczyć, że choć transmisja ta raczej nie niesie za sobą negatywnych skutków, to sama jakość wygenerowanych postaci i ton ich wypowiedzi przywodzi na myśl rzeczywistych polityków. Biorąc pod uwagę fakt, że technologia może zostać wykorzystana do różnych celów i z coraz większą precyzją, niepokojące jest to, że w zasadzie każdy człowiek może paść ofiarą manipulacji i ujrzeć siebie mówiącego rzeczy, o których być może nigdy by nie pomyślał, lub zachowującego się w niewłaściwy sposób. Poza opisaną transmisją z wykorzystaniem wizerunku znanych polityków wymienić można choćby przypadek Nancy Pelosi, byłej spikerki Izby Reprezentantów, której wystąpienie zostało spowolnione i zmontowane w taki sposób, że na pierwszy rzut oka kobieta wydaje się być pod wpływem środków psychoaktywnych⁵⁷.

Dagmara Mateja wyróżnia kilka kategorii jednostek zagrożonych materiałami typu *deepfake*:

- celebryci – głównym zagrożeniem dla tej grupy jest rozpowszechnianie fałszywych treści erotycznych z ich udziałem i stale rozwijająca się pornografia;
- politycy i reprezentanci instytucji państwowych – ta kategoria nierzadko pada ofiarą manipulacji słowem poprzez fałszowanie wypowiedzi, jak również manipulacji obrazem za pośrednictwem fałszywych treści, np. ze spotkań ze szpiegami i przestępcami lub poświadczającej obecność w innych kontrowersyjnych miejscach;

⁵⁶ Zapis transmisji na kanale TrumpOrBiden2024 na Twitch.tv, <https://www.twitch.tv/videos/2104948045> [dostęp: 28.03.2024].

⁵⁷ I. Dąbrowska, *Deepfake...*, s. 94.

- żołnierze – zagrożenie dla nich stanowią nieprawdziwe treści dotyczące rzekomych egzekucji na niewinnych ludziach;
- funkcjonariusze – niebezpieczeństwem dla tej grupy są fałszywe oskarżenia natury obyczajowej lub o łapówkarstwo;
- użytkownicy Internetu – zagrożeniem są wszystkie dezinformujące treści dostępne w Internecie, w tym również szerzenie za ich pośrednictwem przemy i pornografii, której sami mogą paść ofiarą⁵⁸.

Jako że obecne społeczeństwo jest, według terminologii Manuela Castellsa, społeczeństwem informacyjnym charakteryzującym się wszechobecnością nowej technologii, która ułatwia i przyspiesza przepływ informacji, zdecydowana większość ludzi może określić siebie mianem „użytkowników Internetu”. Przypadek Mii Janin ukazuje, jakie mogą być konsekwencje publikowania materiałów na swój temat np. w portalach społecznościowych typu Facebook czy Instagram. Gdyby takich treści zabrakło w Internecie, prześladowcom 14-latkibyłoby o wiele trudniej stworzyć fałszywe materiały z jej udziałem. Czy to oznacza, że powinniśmy wstrzymać się od korzystania z portali społecznościowych? Zdaniem autora artykułu absolutnie nie, ponieważ negatywne konsekwencje nie powinny przysłańiać tych pozytywnych, jakimi zdecydowanie są ułatwienie procesów komunikacyjnych i możliwość pozyskiwania informacji. Należy jednak zachować szczególną ostrożność zarówno w kontekście własnego bezpieczeństwa, jak i wiarygodności odbieranych przez nas danych, które mogły zostać sfalszowane.

Pozostając przy temacie wizerunku i publikacji materiałów w mediach społecznościowych, trzeba zaznaczyć, że tak naprawdę wielu użytkowników fałszuje informacje na swój temat, być może nawet nie zdając sobie z tego sprawy. Ilona Dąbrowska zwraca uwagę, że od kilku lub nawet kilkunastu lat powszechnie dostępne są technologie umożliwiające korektę grafiki, w tym własnych zdjęć. Za ich pośrednictwem użytkownik jest w stanie poprawić obraz według własnych preferencji⁵⁹. W ten sposób osoby publikujące zdjęcia w mediach dokonują poprawek poprzez sztuczne podkreślenie własnych atutów fizycznych bądź maskowanie negatywnych cech własnego wyglądu. Wszystko to ma na celu zaimponowanie innym i podniesienie własnej popularności. Piotr Sztompka używa terminu „społeczeństwa autoprezentacji”, w którym ludzie mają świadomość tego, że ich życie odbywa się w otoczeniu innych i w relacjach z innymi, a pragnąc jak najlepszej opinii na własny temat, dążyć będą do jej uzyskania wszelkimi możliwymi sposobami⁶⁰. Być może słowo „fałsz” w przypadku takich

⁵⁸ D. Mateja, *Fake newsy w społeczeństwie informacyjnym a organizacje fact-checkingowe*, „Naukowy Przegląd Dziennikarski” 2019, nr 3, s. 25.

⁵⁹ I. Dąbrowska, *Deepfake...*, s. 91–92.

⁶⁰ P. Sztompka, *Socjologia...*, s. 815.

działań jest trochę przesadzone, jeśli weźmiemy pod uwagę omówione w niniejszym artykule historie związane z wykorzystaniem cudzego wizerunku do nieetycznych celów. Trudno jednak nie zgodzić się z tym, że jest to pewien rodzaj zakłamywania rzeczywistości, który jednocześnie wpływać może destrukcyjnie na społeczeństwo. Przez takie praktyki bowiem to, co jest nieidealne, zostaje zamaskowane sztucznym obrazem, a promowane jest wyłącznie to, co wpisuje się w kanon piękna. Magdalena Szpunar, pisząc o chęci ludzi zaistnienia w sieci, używa terminu „kultury cyfrowego narcyzmu”, która oznacza dążenie do uwielbienia i sławy⁶¹. Jak pisze autorka: „Niepoddana pod publiczny osąd jednostka ginie w odmętach innych, sławnych profili [...]. Nigdy też nie można spocząć w walce o zainteresowanie, gdyż internetowe gwiazdy szybko zastępowane są przez nowsze”⁶². Tym samym współczesny człowiek podlega presji zewnętrznej związanej z koniecznością dostosowania się do współczesnych norm i oczekiwań społecznych. Porażka na tym polu zaś nie rzadko oznacza dyskryminację i obniżenie wartości jednostki. Z tego też powodu wiele osób wciąż korzysta z różnych filtrów upiększających swój własny naturalny wygląd. Magdalena Szpunar nie bez przyczyny zastosowała słowo „narcyzm”, ponieważ tak funkcjonująca sieć wykreowała jednostki narcystyczne, które dążą do osiągnięcia własnych korzyści bez względu na okoliczności i nie rzadko kosztem innych⁶³. Ludzie zakłámują informacje na swój temat, chcąc przypodobać się innym, lecz wszystko tak naprawdę oparte jest na fałszu, który nie jest korzystny dla innych, a wynika jedynie z własnych, egoistycznych potrzeb.

Kultura nieufności w dobie popularyzacji fikcyjnych treści

Poza zagrożeniami i zmianami związanymi z manipulacją własnym wizerunkiem niepokój budzi również możliwość skutecznego szerzenia dezinformacji na ważne społecznie tematy, co w dzisiejszym niepewnym świecie stanowi kolejne poważne zagrożenie. Obecny świat jest światem VUCA, cechującym się wzmożoną obawą o przyszłość, lękiem związanym z tempem przemian i niedostateczną wiedzą w zakresie rozwiązywania nowych problemów⁶⁴. Na akronim VUCA składają się cztery elementy: *volatility* – ciągłość i szybkie tempo zmian; *uncertainty* – niepewność związana z nieprzewidywalnością nowych wyzwań;

⁶¹ M. Szpunar, *Kultura cyfrowego narcyzmu*, Wydawnictwo AGH, Kraków 2016, s. 67–68.

⁶² Tamże.

⁶³ K. Mularz, *Media społecznościowe...*, s. 99.

⁶⁴ E. Dybowska, *Pedagogika ignacjańska...*, s. 13.

complexity – złożoność świata charakteryzująca się brakiem jasnych związków przyczynowo-skutkowych; *ambiguity* – niejednoznaczność cechująca się wysokim prawdopodobieństwem błędu i dezorientacji⁶⁵. Wszystkie te elementy, z naciskiem zwłaszcza na ten ostatni, mają swoje odzwierciedlenie również w zjawisku popularyzacji technologii *deepfake*.

22 marca 2024 r. miał miejsce atak terrorystyczny na salę koncertową pod Moskwą. Według danych z 29 marca 2024 r. w zamachu życie straciły 143 osoby, a 382 zostały ranne⁶⁶. Do ataku przyznało się Państwo Islamskie, które zagroziło Federacji Rosyjskiej przeprowadzeniem kolejnych ataków w ramach odwetu za zniewolenie członków ich organizacji⁶⁷. Zanim jednak nastąpiło potwierdzenie tej informacji, w tym przez agencje wywiadowcze Stanów Zjednoczonych, na prorosyjskich kanałach wybrzmiało jakoby odpowiedzialna za atak terrorystyczny była Ukraina. Informacje były powielane przez różne konta w mediach społecznościowych i szybko zyskały na popularności. Do manipulacji wykorzystana została m.in. fałszywa informacja, według której odnaleziono pojazd z ukraińskimi tablicami rejestracyjnymi, który zdaniem autorów najprawdopodobniej należał do zamachowców⁶⁸. W sieci pojawił się również *deepfake* przedstawiający sekretarza Rady Bezpieczeństwa Ukrainy Ołeksija Daniłowa, który „potwierdził” udział Ukrainy w zamachu terrorystycznym⁶⁹. To nie jedyny przykład manipulowania informacją przez władze rosyjskie. W pierwszym roku konfliktu zbrojnego pomiędzy Rosją a Ukrainą zidentyfikowano wiele fałszywych przekazów w mediach prorosyjskich, które miały na celu zdyskredytowanie narodu ukraińskiego i polaryzowanie społeczeństwa poprzez wzbudzanie niechęci do Ukrainy w innych krajach Unii Europejskiej⁷⁰. W Internecie opublikowano np. informację o wycieku kart do głosowania dotyczących odłączenia Lwowa od Ukrainy i przyłączenia go do Polski⁷¹. Również częstym

⁶⁵ P. Pacek, *Znaczenie informacji dla przetrwania w świecie VUCA* [w:] *Edukacja w świecie VUCA. Szanse, wyzwania, zagrożenia*, red. D. Kaźmierczak et al., LIBRON, Kraków 2021, s. 304.

⁶⁶ wp.pl, *Zamach pod Moskwą. Sąd kolejne zatrzymania*, <https://wiadomosci.wp.pl/zamach-pod-moskwa-sa-kolejne-zatrzymania-7011164756249536a> [dostęp: 29.03.2024].

⁶⁷ rmf.pl, *Państwo Islamskie grozi Rosji zamachami. „Strzeżcie się”*, https://www.rmf24.pl/fakty/swiat/swiat-news-panstwo-islamskie-grozi-rosji-zamachami-strzezcie-sie,nId,7411533#crp_state=1 [dostęp: 29.03.2024].

⁶⁸ Profil Piotr Panasiuk na portalu Twitter, <https://twitter.com/panasukpetr/status/1771264599028846817> [dostęp: 29.03.2024].

⁶⁹ Profil WarNewsPL na portalu Twitter, <https://twitter.com/WarNewsPL1/status/1771464202311352611> [dostęp: 29.03.2024].

⁷⁰ A. Majchrzak, *Rosyjska dezinformacja i wykorzystanie obrazów generowanych przez sztuczną inteligencję (deepfake) w pierwszym roku inwazji na Ukrainę*, „Media. Biznes. Kultura” 2023, nr 1(14), s. 76.

⁷¹ Tamże, s. 76–77.

argumentem, pojawiającym się w sieci do dnia dzisiejszego, jest inicjacja konfliktu z Rosją przez tzw. Państwa Zachodu i Ukrainy⁷². Wykorzystano także *deepfake* przedstawiający wygenerowany przez AI wizerunek prezydenta Ukrainy Wołodimira Zeleńskiego, który mówi m.in. o konieczności kapitulacji Ukrainy i zwróceniu Rosji wschodniego regionu Donbasu⁷³.

Wiele sprzecznych informacji pojawiało się również w okresie pandemii COVID-19. Na stronie internetowej stowarzyszenia Demagog można znaleźć listę ponad 100 wyróżnionych fałszywych informacji na temat koronawirusa⁷⁴. Niepokój społeczny, jaki wywołała niespodziewana sytuacja związana z pojawieniem się nieznannej dotąd jednostki chorobowej, sprawił, że do dyskursu naukowego przeniknęło pojęcie infodemii. Światowa Organizacja Zdrowia (WHO) zdefiniowała ją w następujący sposób: „Infodemia to nadmierna ilość informacji, która utrudnia znalezienie rozwiązania problemu. Infodemiccy mogą szerzyć dezinformację i plotki podczas stanu zagrożenia zdrowia. Może to utrudniać skuteczną reakcję w zakresie zdrowia publicznego oraz powodować zamieszanie i brak zaufania wśród ludzi”⁷⁵. Nieufność społeczeństwa była potęgowana przez sprzeczne komunikaty publiczne, które tylko pogłębiały stan destabilizacji. Władze krajów demokratycznych ograniczyły aktywność i wolność obywateli. W celu zapewnienia bezpieczeństwa ludności wprowadzono m.in. obostrzenia dotyczące publicznych debat na temat trafności zastosowania ograniczenia wolności obywatelskiej⁷⁶. To wszystko przyczyniło się do powstania licznych teorii spiskowych na temat choroby, szczepionek, instytucji farmakologicznych i rządowych oraz konsekwencji izolacji społecznej. Zdaniem Joanny Wróblewskiej-Jachny przyczyniło się to do powstania tzw. mniejszości kognitywnych. Autorka, powołując się na definicję Andrzeja Kasperka, mniejszości kognitywne opisuje jako takie, które „definiują w szczególny sposób rzeczywistość, w opozycji do [...] mainstreamowych definicji”⁷⁷. Podczas trwania pandemii koronawirusa wykreowało się wiele tego typu grup, które były przeciwstawne do oficjalnej narracji specjalistów w zakresie medycyny i władz państwowych. Podważano wiarygodność istnienia wirusa oraz skuteczność szczepionek, które w opinii członków mniejszości kognitywnych miałyby powodować więcej szkód

⁷² Tamże, s. 77.

⁷³ Youtube.com, *Deepfake video of Volodymyr Zelensky surrendering surfaces on social media*, <https://www.youtube.com/watch?app=desktop&v=X17yrEV5sl4> [dostęp: 28.03.2024].

⁷⁴ Stowarzyszenie Demagog, *Koronawirus – zestawienie fałszywych informacji*, https://demagog.org.pl/analizy_i_raporty/koronawirus-zestawienie-falszywych-informacji/ [dostęp: 29.03.2024].

⁷⁵ who.int, <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf> [dostęp: 29.03.2024].

⁷⁶ J. Wróblewska-Jachna, *Pandemia i infodemia. Refleksje nad jakością ekosystemu medialnego podczas pandemii COVID-19*, „Media i Społeczeństwo” 2022, nr 16, s. 43.

⁷⁷ Tamże.

dla zdrowia niż pożytku. Jedną z takich organizacji jest np. „Stolik Wolności”, który organizował liczne protesty, konferencje oraz publikacje mające na celu zdyskredytowanie oficjalnych informacji i działań podejmowanych w celu zwalczania pandemii⁷⁸.

Powyższe przykłady to jedynie fragment materiałów dezinformacyjnych dostępnych w sieci. Jak wskazano w poprzednich podrozdziałach, zakres i cel treści typu *deepfake* jest bardzo szeroki. Niewątpliwie fikcyjne informacje dotyczące ważnych dla społeczeństwa tematów, które implikują zmiany i budzą niepewność, jeszcze bardziej wzmacniają poczucie lęku i nieufności, wpływając destabilizująco na życie społeczne. Jednym z terminów służących opisaniu takiej rzeczywistości jest kultura nieufności, w której brakuje podstawowego elementu spajającego społeczność – zaufania.

Russell Hardin wskazuje, że w literaturze wyróżniono trzy różne koncepcje zaufania. Pierwszą z nich jest model oparty na korzyściach. Model ten zakłada, że dla jednostki, która została obdarzona zaufaniem, motywacją do bycia wiarygodnym i uczciwym jest możliwość pielęgnowania i utrzymywania korzystnej dla niej relacji z osobą ufającą. Drugi model zaufania odnosi się do subiektywnego poczucia wypełnienia moralnych zobowiązań względem ufającego. Trzeci koncentruje się na indywidualnej osobowości i jej psychologicznych oraz charakterologicznych skłonnościach do tego, aby w oczach innych okazać się osobą, której warto zaufa⁷⁹. Hardin pisze dalej, że „zaufanie jest [...] kategorią kognitywną, ponieważ pobudza je oczekiwana motywacja ze strony innych. Wszystkie trzy typowe koncepcje należy rozpatrywać w kategoriach kognitywnych, ponieważ wszystkie opierają się na ocenie wiarygodności osoby potencjalnie zaufanej”⁸⁰.

Piotr Sztompka definiuje zaufanie jako pewnego rodzaju zakład o niepewną przyszłość, która jest związana z postępowaniem innych ludzi. Jak pisze: „wszelkie działania ludzkie są zorientowane ku przyszłości i kształtowane w swoim przebiegu przez antycypacje przyszłości. Skutki naszych obecnych działań zależą w istotnym stopniu od przyszłych i często od nas niezależnych okoliczności”⁸¹. Te okoliczności, o których wspomina autor, wynikają z szeroko pojętej aktywności człowieka, organizacji i instytucji społecznych. Oznacza to, że jednostka funkcjonuje w świecie, który jest stworzony oraz modyfikowany przez innych i ich działania. Inni, o których mowa, również posiadają swoją autonomię i poprzez wolność oraz zdolność do podejmowania inicjatyw zmie-

⁷⁸ Tamże, s. 44.

⁷⁹ R. Hardin, *Zaufanie*, przeł. A. Gruba, Wydawnictwo Sic!, Warszawa 2009, s. 25.

⁸⁰ Tamże.

⁸¹ P. Sztompka, *Socjologia...*, s. 278–279.

niąją rzeczywistość społeczną. Trudno jest, a czasem wręcz niemożliwe, przewidzieć, jakie działania podejmą inni i jakie będą tego konsekwencje. W związku z tym zaufanie jest ściśle powiązane z ryzykiem, czy działania podejmowane przez pozostałych członków społeczeństwa będą dla jednostki szkodliwe, czy też wpłyną na nią korzystnie⁸². W tym kontekście jest to również zakład o to, czy inni ludzie okażą się wiarygodni w swoim działaniu, o czym pisał wspomniany wcześniej Hardin.

Sztompka zaufanie nazywa fundamentalną relacją, która stanowi ważną część kapitału moralnego będącego szczególną odmianą więzi społecznych. Moralność składa się z trzech kluczowych elementów znaczeniowych:

- stosunku jednych do drugich;
- relacji w istotny sposób oddziałującej na dobro innej jednostki;
- norm i zobowiązań charakteryzujących daną relację, których naruszenie oznacza szkodę dla partnera⁸³.

Relacje międzyludzkie są więc ściśle związane z moralnością, tworząc swą własną przestrzeń moralną, a jednostki działające w owej przestrzeni można określić jako posiadające kapitał moralny, kapitał, którego istotnym elementem jest właśnie zaufanie⁸⁴.

Rola zaufania jest niezwykle istotna dla budowy stabilnego, prawidłowo rozwijającego się społeczeństwa. Zdaniem Sztompki istnieje co najmniej siedem powodów takiego stanu rzeczy. Po pierwsze świat jest stale kształtowany i modyfikowany przez ludzi, co oznacza, że na rzeczywistość, w której żyjemy, nieustannie wpływ mają zarówno pojedyncze jednostki, jak i mniejsze bądź większe grupy społeczne. To oznacza, że aby rozwiązywać problemy i efektywnie radzić sobie z wyzwaniem przyszłości, trzeba ściśle współpracować z innymi, a poprzez współpracę należy rozumieć również okazywanie innym zaufania⁸⁵. Drugim powodem jest rozwój sieci społecznej, w tym sieci współzależności między różnymi systemami (np. politycznymi lub finansowymi), co spowodowane jest ciągłym procesem globalizacji społeczeństw, a prawidłowe funkcjonowanie tych systemów wymaga dużych pokładów zaufania. Trzecim powodem są osiągnięcia technologiczne, których wdrożenie przyniosło, poza korzyściami, również nieoczekiwane skutki i nowe problemy, a tym należy przeciwdziałać poprzez zwiększenie zasobów zaufania. Większe możliwości oferowane przez świat są czwartą przyczyną wzrastającego znaczenia zaufania. Mianowicie wzrost możliwości sprawił, że coraz mniej przewidywalne stają się decyzje podejmowane przez

⁸² Tamże.

⁸³ Tamże, s. 272

⁸⁴ Tamże.

⁸⁵ P. Sztompka, *Zaufanie. Fundament społeczeństwa*, Znak, Kraków 2007, s. 45.

ludzi⁸⁶. Zygmunt Bauman określił współczesność mianem płynnej ponowoczesności, której immanentną cechą staje się niepewność. Ciągłe przemiany i brak stabilizacji osłabiły poczucie bezpieczeństwa i pewność wobec zamiarów innych⁸⁷. Niepewność w tym kontekście związana jest z niewiedzą na temat tego, komu należy zaufać np. podczas wyborów parlamentarnych, a w drugą stronę i bazując na tym samym przykładzie – jaką politykę ostatecznie realizować będzie wybrane ugrupowanie. Zaufanie staje się więc obecnie niezbędne dla podjęcia jakichkolwiek działań, ponieważ nieufność oznaczałaby rezygnację z aktywności społecznej, np. brak udziału w wyborach⁸⁸. Piątym powodem jest złożoność współczesnych systemów, w których coraz trudniej odnaleźć się przeciętnemu człowiekowi. Nieprzejrzystość złożonego życia społecznego skutkuje wzrostem potrzeby zaufania, bez którego ludzie byłiby niezdolni do podejmowania jakichkolwiek działań⁸⁹. Szóstym powodem jest anonimowość jednostek stanowiących sieć społeczną człowieka. Obecnie często jesteśmy zależni od osób, których osobiście nie znamy i których nie możemy kontrolować. Niezbędne zatem dla codziennego funkcjonowania staje się zaufanie, którym obdarzamy innych, np. że nikt bezprawnie nie wykorzysta naszego wizerunku. Wreszcie siódmym powodem, niejako powiązaniem z poprzednimi, jest coraz większa liczba obcych osób w najbliższym otoczeniu i wymiana międzykulturowa. Wzrost migracji jest źródłem niepewności związanej z powstaniem kulturowej przestrzeni, w której to, co znajome i ukształtowane przez tradycyjne wartości, konkuruje z nieznanymi dotąd normami i obyczajami reprezentowanymi przez obce jednostki. Tak różnorodne środowisko wymaga zaufania, ponieważ, jak pisze Sztompka: „[...] staje się niezbędnym zasobem pozwalającym poradzić sobie z obecnością obcych”⁹⁰.

Opis ten miał na celu przybliżenie roli, jaką pełni zaufanie w życiu społecznym. Trzeba jednak zastanowić się, czy w obecnym społeczeństwie, pełnym dezinformacji, niepewności i nadużywania prywatności, możliwe jest zbudowanie wzajemnego zaufania. Jego brak oznacza osłabienie kapitału moralnego, co sprawia, że rozpadowi ulegają również więzi społeczne. Skrajnym przypadkiem zaniku elementów składających się na kapitał moralny jest całkowita atrofia więzi, która stanowi antytezę moralności⁹¹. Społeczności przejawiające zanik moralności cechuje nieufność i podejrzliwość wobec innych,

⁸⁶ Tamże, s. 45–47.

⁸⁷ Z. Bauman, *Ponowoczesność, czyli dekonstruowanie nieśmiertelności* [w:] *Postmodernizm a filozofia*, red. S. Czerniak, IFiS PAN, Warszawa 1996, s. 55

⁸⁸ P. Sztompka, *Zaufanie...*, s. 47.

⁸⁹ Tamże, s. 48.

⁹⁰ Tamże, s. 49.

⁹¹ P. Sztompka, *Socjologia...*, s. 276.

manipulowanie innymi, obojętność wobec doświadczanych przez innych krzywd i egoizm jednostek⁹².

Analizując przykłady manipulowania informacją za pośrednictwem technologii *deepfake* wydaje się, że obecnie coraz trudniej o zaufanie wobec innych. Szerzenie dezinformacji w zakresie pandemii COVID-19 czy konfliktu zbrojnego w Ukrainie znacznie ograniczyło zasób zaufania wśród ludzi. Działa ono pod tym względem destabilizująco, a społeczność coraz częściej przejawia cechy rozpadu więzi moralnych, zwłaszcza w obliczu przejawów wykorzystania wizerunku innych do nieetycznych celów. Kultura nieufności, inaczej nazywana kulturą cynizmu, cechuje się powszechną i uogólnioną podejrzliwością jednostek wobec instytucji publicznych i siebie nawzajem. Ludzie nieustannie kontrolują działania innych, a to wszystko w obawie przed oszustwem, nadużyciami i szeroko pojętymi krzywdami, które mogą im wyrządzić⁹³. Jak zostało opisane powyżej, do stabilizacji i prawidłowego funkcjonowania systemów społecznych niezbędne jest zaufanie jako fundament więzi społecznych. Kultura nieufności grozi ich całkowitym rozpadem, a bez nich niemożliwe jest zachowanie ładu społecznego i wzajemnej współpracy⁹⁴. Warto w tym miejscu zaznaczyć, że w aspekcie psychologicznym długotrwała ekspozycja na sytuacje stresowe, w której jednostka czuje się bezradna, stanowi poważne zagrożenie i czynnik ryzyka dla zdrowia psychicznego⁹⁵. Rozwój technologiczny sprawił, że współczesne społeczeństwo pełne jest różnych zagrożeń, które zmuszają nas do zachowania szczególnej ostrożności i wzbudzają niepokój. Sytuacji dodatkowo nie poprawia fakt, że obecnie wciąż zmagamy się ze skutkami pandemii COVID-19, a do tego wciąż trwa konflikt zbrojny w Ukrainie, który znacząco oddziałuje na społeczne nastroje ze względu na możliwość jego eskalacji na szerszą skalę. Według badań realizowanych przez Free Press for Eastern Europe 79% Polaków odczuwa lęk związany z wojną⁹⁶. A to wszystko w obliczu niepokojących prognoz ze strony Światowej Organizacji Zdrowia (WHO), której eksperci przewidują, że depresja stanie się do 2030 r. jedną z najpoważniejszych jednostek chorobowych, z którymi przyjdzie zmagać się ludzkości⁹⁷.

⁹² Tamże.

⁹³ M. Kosiorek, *Zaangażowanie społeczne nauczycieli w obliczu kultury nieufności*, „Przełom Pedagogiczny” 2020, nr 2, s. 51–52.

⁹⁴ Tamże.

⁹⁵ P. Pustułka, J. Radzińska, A. Kwiatkowska, *Strach w obliczu wojny w Ukrainie: osobiste i geopolityczne obawy młodych Polek i Polaków*, „Kultura i Społeczeństwo” 2023, nr 4, s. 209.

⁹⁶ Tamże.

⁹⁷ D.J. Szadkowski, J. Podemska, *Depresja i zachowania samobójcze wśród dzieci i młodzieży w dobie współczesnych zawirowań społecznych*, „Społeczeństwo. Edukacja. Język” 2022, nr 16, s. 219.

Zakończenie

Ulrich Beck określa współczesną cywilizację mianem „społeczeństwa ryzyka”, które cechuje niepewność związana ze skutkami ubocznymi postępu, jakimi są nieprzewidywalność zmian i zagrożeń, na które coraz trudniej się w należyty sposób przygotować⁹⁸. Zaufanie, które stanowi fundament społeczeństwa i które jest niezbędne do jego rozwoju oraz trwania w ładu i porządku, jest z kolei podjęciem ryzyka w kontekście oczekiwań wobec zachowań innych. Analizując zagrożenia związane z siecią i nowymi technologiami, należy zadać sobie pytanie: czy w dobie dezinformacji i narzędzi służących bezprawnemu wykorzystywaniu wizerunku innych społeczność jest zdolna do trwania we wzajemnym zaufaniu i budowaniu lepszej przyszłości?

Artykuł miał na celu nie tylko pogłębienie świadomości na temat zagrożeń w postaci fałszywych treści godzących w prywatność i implikujących niepokój. Autor żywi nadzieję, że czytelnik podda ten temat własnej refleksji i zastanowi się nad tym, w jaki sposób przeciwdziałać podobnym zjawiskom. Należy bowiem poszukiwać skutecznych rozwiązań, które stanowiłyby remedium na pojawiające się w związku z dynamicznym rozwojem technologii niebezpieczeństwa. Najbardziej zagrożoną grupą jest młodzież, która większość swojego wolnego czasu spędza w Internecie, przez co jest szczególnie narażona na skutki dezinformacji. Nie mówiąc już o tym, że sama może paść ofiarą wykorzystania własnego wizerunku. Nauczyciele, pedagodzy i rodzice powinni edukować uczniów w zakresie prawidłowego i odpowiedzialnego korzystania z sieci, a także poszanowania prywatności innych osób, tak aby nie powtórzyła się tragiczna historia z 14-letnią Mią Janin. Należy również weryfikować razem z młodzieżą informacje pozyskiwane z sieci, tak aby lęk i niepewność, które autorzy fikcyjnych treści starają się zasiać w umysłach odbiorców, nie dotknęły młodych osób, wciąż znajdujących się na etapie wczesnego rozwoju psychospołecznego.

Ponadto technologie i programy służące do obróbki filmów i obrazów stworzyły nowe możliwości do produkcji materiałów pornograficznych, przyczyniając się tym samym do powstania fałszywej pornografii. Jak wskazano w artykule, narzędzia służące do tworzenia fikcyjnych treści pornograficznych umożliwiły ich autorom czerpanie korzyści materialnych, co samo w sobie jest postępowaniem nieetycznym nie tylko ze względu na publikowanie szkodliwych treści erotycznych, ale również z powodu wykorzystywania w tym celu wizerunku innych osób. W tym kontekście należy więc wskazać młodzieży właściwe metody postępowania na drodze rozwoju osobistego, które nie godzą w godność drugie-

⁹⁸ U. Beck, *Społeczeństwo światowego ryzyka. W poszukiwaniu utraconego bezpieczeństwa*, Scholar, Warszawa 2012, s. 21.

go człowieka i stanowią rzeczywistą wartość rozwojową zarówno dla jednostki, jak i całego społeczeństwa. Promowanie fałszywych treści pornograficznych jest szkodliwe ze względu na postawę, w której jednostka wykorzystuje bez poszanowania wizerunek innych jako narzędzie do generowania własnych zysków, co nie sprzyja zaufaniu i podtrzymywaniu więzi społecznych. Ponadto treści pornograficzne same w sobie są niezwykle szkodliwe, bowiem, jak zaznacza Anna Krawula-Ptaszyńska, wczesny kontakt młodego człowieka z pornografią zaburza prawidłowe rozumienie roli mężczyzny i kobiety w życiu społecznym, zwłaszcza gdy informacje w tym zakresie przekazane przez wychowawców są znikome i zastępowane przez treści pornograficzne, które dostarczają fałszywego obrazu rzeczywistości, co również nie sprzyja budowaniu relacji społecznych opartych na zaufaniu i szacunku⁹⁹.

Wreszcie należy również uczyć młode osoby wzajemnej współpracy i zaufania, ponieważ jest to kluczowe w kontekście prawidłowego funkcjonowania w społeczeństwie i zapewnienia jego stabilizacji. Zaufanie jest również czymś, czego nieustannie powinni uczyć się wszyscy, włącznie z nauczycielami i rodzicami. Jak bowiem sugeruje przedstawiona analiza, cierpimy na deficyt zaufania i prospołecznych cech, których rozwój jest niezwykle istotny dla odwrócenia negatywnych tendencji związanych ze zdrowiem psychicznym i budowania lepszej przyszłości dla następnych pokoleń.

Bibliografia

- Bauman Z., *Ponowoczesność, czyli dekonstruowanie nieśmiertelności* [w:] *Postmodernizm a filozofia*, red. S. Czerniak, IFiS PAN, Warszawa 1996.
- Beck U., *Spoleczeństwo światowego ryzyka. W poszukiwaniu utraconego bezpieczeństwa*, Scholar, Warszawa 2012.
- Castells M., *Spoleczeństwo sieci*, tłum. M. Marody et al., Wydawnictwo Naukowe PWN, Warszawa 2007.
- Dąbrowska I., *Deepfake – nowy wymiar internetowej manipulacji*, „Zarządzanie Mediami” 2020, t. 8(2).
- Dijk J. van, *Spoleczne aspekty nowych mediów*, przeł. J. Konieczny, Wydawnictwo Naukowe PWN, Warszawa 2020.
- Dybowska E., *Pedagogika ignacjańska w świecie VUCA*, „Horyzonty Wychowania” 2022, nr 21 (57).
- Hardin R., *Zaufanie*, przeł. A. Gruba, Wydawnictwo Sic!, Warszawa 2009.
- Januszewska A. et al., *Uzależnienie od Internetu oraz samotność wśród studentów Warszawskiego Uniwersytetu Medycznego* [w:] *Wyzwania współczesnego pielęgniarstwa*, red. M. Wysokiński, Z. Sienkiewicz, W. Fidecki, Warszawski Uniwersytet Medyczny, Warszawa 2021.

⁹⁹ A. Krawula-Ptaszyńska, *Pornografia w procesie socjalizacji dzieci i młodzieży*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny” 2004, nr 2, s. 284.

- Kania S., *Seksting – diagnoza i profilaktyka pozytywna* [w:] *Psycholog w szkole: koncepcje*, red. Z.B. Gaś, Innovatio Press, Lublin 2024.
- Karski K., *Korzyści i zagrożenia wynikające z implementacji sztucznej inteligencji*, „Zeszyty Naukowe Akademii Górnośląskiej” 2023, nr 6.
- Krawula-Ptaszyńska A., *Pornografia w procesie socjalizacji dzieci i młodzieży*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny” 2004, nr 2.
- Kosiorek M., *Zaangażowanie społeczne nauczycieli w obliczu kultury nieufności*, „Przegląd Pedagogiczny” 2020, nr 2.
- Majchrzak A., *Rosyjska dezinformacja i wykorzystanie obrazów generowanych przez sztuczną inteligencję (deepfake) w pierwszym roku inwazji na Ukrainę*, „Media. Biznes. Kultura” 2023, nr 1(14).
- Marciszewski W., *Sztuczna inteligencja*, Znak, Kraków 1998.
- Mateja D., *Fake newsy w społeczeństwie informacyjnym a organizacje fact-checkingowe*, „Naukowy Przegląd Dziennikarski” 2019, nr 3.
- Mularz K., *Media społecznościowe a budowanie wizerunku i relacji międzyludzkich* [w:] *Świat (po)cyfrowej rewolucji. W kręgu analiz socjologicznych*, red. A. Kotowska, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów 2019.
- Mularz K., *Znaczenie przyjaźni we współczesnym świecie*, „Społeczeństwo. Edukacja. Język” 2023, nr 18.
- Musiał E., *Nowe trendy w edukacji – koncepcja głębokiego uczenia się*, „Zeszyty Naukowe Wyższej Szkoły Humanitas. Pedagogika” 2018, nr 16.
- Okulska I., *Sieci neuronowe typu GAN i GPT-2, słowa zużyte i kreatywność, czyli literacki second-hand*, „Forum Poetyki” 2019, nr 18.
- Pacek P., *Znaczenie informacji dla przetrwania w świecie VUCA* [w:] *Edukacja w świecie VUCA. Szanse, wyzwania, zagrożenia*, red. D. Kaźmierczak et al., LIBRON, Kraków 2021.
- Piecuch A., *Życie smart – ma swoją cenę*, „Dydaktyka Informatyki” 2020, nr 15.
- Piechota B., *O atrakcyjności sieci* [w:] *Paradoksy Internetu. Konteksty społeczno-kulturowe*, red. M. Szpunar, Wydawnictwo Adam Marszałek, Toruń 2011.
- Pustułka P., Radzińska J., Kwiatkowska A., *Strach w obliczu wojny w Ukrainie: osobiste i geopolityczne obawy młodych Polek i Polaków*, „Kultura i Społeczeństwo” 2023, nr 4, s. 205–230.
- Różanowski K., *Sztuczna inteligencja: rozwój, szanse i zagrożenia*, „Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki” 2007, nr 2(2).
- Stęgowski Z., *Sztuczne sieci neuronowe*, „Kernel.” 2004, nr 1(1), s. 16–19.
- Szadkowski D.J., Podemska J., *Depresja i zachowania samobójcze wśród dzieci i młodzieży w dobie współczesnych zawirowań społecznych*, „Społeczeństwo. Edukacja. Język” 2022, nr 16, s. 215–240.
- Szpunar M., *Kultura cyfrowego narcyzmu*, Wydawnictwo AGH, Kraków 2016.
- Sztompka P., *Zaufanie. Fundament społeczeństwa*, Znak, Kraków 2007.
- Sztompka P., *Socjologia. Wykłady o społeczeństwie*, Znak Horyzont, Kraków 2021.
- Wasiuta O., Wasiuta S., *Deepfake jako skomplikowana i głęboko fałszywa rzeczywistość*, „Studia de Securitate” 2019, nr 9(3).
- Wasiuta O., Wasiuta S., *FakeApp jako nowe zagrożenie bezpieczeństwa politycznego i informacyjnego*, „Studia de Securitate” 2019, nr 9(3).
- Wróblewska-Jachna J., *Pandemia i infodemia. Refleksje nad jakością ekosystemu medialnego podczas pandemii COVID-19*, „Media i Społeczeństwo” 2022, nr 16.
- Yeh CY. et al., *Disrupting Image-Translation-Based DeepFake Algorithms with Adversarial Attacks* [w:] *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*, Computer Vision Foundation, 2020.

Netografia

- Bankier.pl, *Kolejny deepfake zebrał zniwa. Tym razem celem był Joe Biden*, <https://www.bankier.pl/wiadomosc/Kolejny-deepfake-zebral-zniwa-Tym-razem-celem-byl-Joe-Biden-8701675.html> [dostęp: 28.03.2024].
- billboard.com, *Taylor Swift Searchable on X Again After Brief Blockage Over DeepFake Images*, https://www.billboard.com/business/tech/taylor-swift-searchable-x-twitter-ai-generated-deepfake-explicit-photos-1235592389/?utm_medium=social&utm_source=twitter [dostęp: 28.03.2024].
- Francuz B., *AI zamiast modelek. OnlyFans mogą zalać profile kobiet, które nie istnieją – i ludzie za to zapłacą*, <https://android.com.pl/tech/562237-ai-onlyfans-generowane-modelki/> [dostęp: 28.03.2024].
- Groszkowska K., *Cyberprzemoc wobec kobiet*, Instytut Prawa Ustrojowego, https://ipu.org.pl/wp-content/uploads/2021/02/Cyberprzemoc-wobec-kobiet_Kamila-Groszkowska.pdf [dostęp: 28.03.2024].
- Hooper S., *Girl killed herself when bullies shared fake nudes of Her*, <https://metro.co.uk/2024/01/24/teen-took-life-online-bullying-shared-fake-nudes-20162284/> [dostęp: 28.03.2024].
- Profil AI Corner Hotties na portalu Twitter, <https://twitter.com/AiCorner54628> [dostęp: 28.03.2024].
- Profil AI_Girls_Studio na portalu Twitter, https://twitter.com/AI_Girls_Studio [dostęp: 28.03.2024].
- Profil AI_Girls_Studio na portalu Patreon, <https://www.patreon.com/AIGirlsStudio> [dostęp: 28.03.2024].
- Profil kira_ai_gf na portalu Twitter, https://twitter.com/kira_ai_gf [dostęp: 28.03.2024].
- Profil Piotr Panasiuk na portalu Twitter, <https://twitter.com/panasukpetr/status/1771264599028846817> [dostęp: 29.03.2024].
- Profil WarNewsPL na portalu Twitter, <https://twitter.com/WarNewsPL1/status/1771464202311352611> [dostęp: 29.03.2024].
- rmf.pl, *Państwo Islamskie grozi Rosji zamachami. „Strzeżcie się”*, https://www.rmf24.pl/fakty/swiat/news-panstwo-islamskie-grozi-rosji-zamachami-strzezcie-sie,nId,7411533#crp_state=1 [dostęp: 29.03.2024].
- Stowarzyszenie Demagog, *Ostrzał szpitala w Mariupolu był ustawką? Fake news!*, https://demagog.org.pl/fake_news/ostrzal-szpitala-w-mariupolu-by-l-ustawka-fake-news/ [dostęp: 28.03.2024].
- Stowarzyszenie Demagog, *Pornografia tworzona przez AI, czyli deepfake z Taylor Swift*, https://demagog.org.pl/analizy_i_raporty/pornografia-tworzona-przez-ai-czyli-deepfake-z-taylor-swift/ [dostęp: 28.03.2024].
- Stowarzyszenie Demagog, *Koronawirus – zestawienie fałszywych informacji*, https://demagog.org.pl/analizy_i_raporty/koronawirus-zestawienie-falszywych-informacji/ [dostęp: 29.03.2024].
- Youtube.com, *Deepfake video of Volodymyr Zelensky surrendering surfaces on social media*, <https://www.youtube.com/watch?app=desktop&v=X17yrEV5sl4> [dostęp: 28.03.2024].
- Youtube.com, *Tadeusz Szuk – King Bruce Lee karate mistrz | AI Cover*, <https://www.youtube.com/watch?v=QW22vGPRvuo> [dostęp: 28.03.2024].
- Youtube.com, *The Schwarzenegers of the Ring Part 2*, <https://www.youtube.com/watch?v=Fko4z7lADw0> [dostęp: 28.03.2024].
- Youtube.com, *Einstein AI – Interactive Virtual Being*, <https://www.youtube.com/watch?v=c0o5s0DZeJl> [dostęp: 28.03.2024].

Wheeler D.W., *Artificial Intelligence and Its Disinformation Campaigns*, 2019, <https://www.dylanwheeler.net/files/Disinformation.pdf> [dostęp: 27.03.2023].

who.int, <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf> [dostęp: 29.03.2024].

wp.pl, *Zamach pod Moskwą. Są kolejne zatrzymania*, <https://wiadomosci.wp.pl/zamach-pod-moskwa-sa-kolejne-zatrzymania-7011164756249536a> [dostęp: 29.03.2024].

Zapis transmisji na kanale TrumpOrBiden2024 na Twitch.tv, <https://www.twitch.tv/videos/2104948045> [dostęp: 28.03.2024].