Nr 20(1/1)/2024
Numer DOI: 10.5604/ 01.3001.0054.6519

**Paweł Pajor** [1]
**Uniwersytet Śląski, Polska**
ORCID ID: 0000-0002-4711-8781
e-mail: pavelpajor@gmail.com

# Inżynieria społeczna w czasach generatywnej sztucznej inteligencji

ABSTRAKT

Artykuł przedstawia ewolucję inżynierii społecznej w kontekście dynamicznego rozwoju generatywnej sztucznej inteligencji. Socjotechnika w początkach swego istnienia uważana była za dziedzinę naukową reprezentującą praktyczny aspekt zastosowania wiedzy socjologicznej, lecz obecnie najczęściej postrzegana jest jako narzędzie manipulacji. Jednak rozwój generatywnej sztucznej inteligencji oraz jej wykorzystanie w klasycznych technikach i strategiach socjotechnicznych może przesunąć granice zainteresowania socjologii z działań obszaru badań i opisu zachowań społecznych do praktycznego wykorzystania w procesach inicjowania i kierowania zmianami społecznymi. Autor proponuje zdefiniowanie dwóch obszarów oddziaływania inżynierii społecznej, w których wykorzystywane są zaawansowane technologie cyfrowe, oraz definiuje termin syntetycznego przekazu. Wprowadza termin Frontend Socjotechniczny, obejmujący jawne i widoczne formy aktywności algorytmów cyfrowych, oraz Backend Socjotechniczny, który obejmuje niejawne formy działania technologii cyfrowych. Artykuł przedstawia analizę syntetycznych formy przekazu, takich jak fakenews (tekst), photofake (obraz), voicefake (głos) i deepfake (obraz i głos), które coraz częściej są wykorzystywane w manipulacji opiniami, postawami i zachowaniami społecznymi. Artykuł porusza również kwestie dotyczące zastosowania technologii cyfrowych do nieautoryzowanej inwigilacji cyfrowej.

SŁOWA KLUCZOWE: sztuczna inteligencja, socjotechnika, *deepfake*, manipulacja, inwigilacja

## Wprowadzenie

Inżynieria społeczna, znana również jako socjotechnika, mimo obecności w dyskursie akademickim i publicznym, pozostaje zaniedbanym obszarem badawczym i jest w dużej mierze nieobecna w głównym nurcie socjologii. Mimo że w przestrzeni naukowej socjotechnika określana jest m.in. jako praktyczne zastosowanie wiedzy socjologicznej w rozwiązywaniu problemów społecznych, to obecnie najczęściej dziedzina ta utożsamiana jest z manipulacją. Warto jednak przypomnieć, że socjotechnika ma również bogatą historię naukową i teoretyczną. Jako interdyscyplinarna dziedzina socjotechnika czerpie z teorii socjologii, psychologii, nauk politycznych i ekonomii, tworząc praktyczne metody wpływania na

[1] Data złożenia tekstu do Redakcji „MiS”: 15.01.2024; data recenzji: 15.05.2024; data zatwierdzenia tekstu do druku: 29.05.2024; data publikacji: 30.06.2024/Submission date to the "Media and Society" Editorial Office: 15.01.2024; review date: 15.05.2024; article approval print date: 29.05.2024; publication date: 30.06.2024.

społeczeństwo i rozwiązywania problemów społecznych. Celem inżynierii społecznej była nie tylko manipulacja, jak często jest błędnie rozumiana socjotechnika, ale także racjonalne kształtowanie procesów społecznych w celu poprawy jakości życia ludzi. Socjotechnika, poprzez swoje narzędzia i metody, pozwala na analizę i interwencję w różnych aspektach życia społecznego, od kampanii zdrowotnych po strategie polityczne. Z tego względu jest to dziedzina, która zawsze miała, ma i będzie miała zastosowanie w wielu obszarach aktywności człowieka.

**Kontekst historyczny inżynierii społecznej**

Wśród najważniejszych naukowych prekursorów inżynierii społecznej wymienia się takich badaczy jak Nathan Roscoe Pound, Karl Raimund Popper, Robert King Merton, Gunnar Myrdal oraz Edward L. Bernays. W Polsce przedstawicielami tej dziedziny byli Stanisław Ossowski, Leon Petrażycki, Zygmunt Bauman, Adam Podgórecki.

Pochodzenie terminu socjotechnika jest kwestią sporną, jednak przyjmuje się, że jest to nauka powstała w XX wieku. Terminu inżynieria społeczna, w znaczeniu zastosowania wiedzy nauk społecznych do celowych działań praktycznych, użył po raz pierwszy amerykański socjolog Edwin L. Earp w książce *The Social Engineer*[2], opublikowanej w 1911 r. w Stanach Zjednoczonych. Autor przedstawił wizję społeczeństwa jako maszyny, którą można kształtować i poprawiać za pomocą metod technicznych. Twierdził, że inżynier społeczny jest osobą, która potrafi zastosować naukowe metody do rozwiązywania problemów społecznych i kształtowania postaw ludzkich.

Natomiast Karl R. Popper używał terminu inżynieria społeczna w kontekście krytyki autorytarnego kształtowania i kontrolowania społeczeństwa zgodnie z określonym planem[3]. Termin inżynieria społeczna pojawia się również w pracy R. Pounda, przedstawiciela amerykańskiej socjologicznej szkoły prawa, zatytułowanej *Introduction to the Philosophy of Law*[4] wydanej w 1922 roku oraz w *Interpretations of Legal History*[5] w 1923 roku. Koncepcja inżynierii społecznej zaproponowana przez Pounda uznawała prawo za instrument socjotechniki, ponieważ jest ono w stanie skutecznie zapewnić równowagę społeczną.

[2] E.L. Earp, *The Social Engineer*, New York Eaton & Mains, 1911.
[3] K.R. Popper, Społeczeństwo otwarte i jego wrogowie, Wydawnictwo Naukowe PWN, 2010.
4 R. Pound, *Introduction to the Philosophy of Law*, New Haven, Yale University Press, 1922.
5 R. Pound, *Interpretations of Legal History*, Cambridge, Cambridge University Press, 1923.

W 1944 roku terminu inżynieria społeczna użył Gunnar Myrdal i jego współpracownicy Richard Sterner i Arnold Rose w książce *An American Dilema*[6], którzy zajmowali się problematyką rasizmu w Ameryce. G. Myrdal i jego współpracownicy podnieśli w swojej pracy konieczność stworzenia praktycznego narzędzia, które w efektywny sposób pozwoliłoby rozwiązać ten społeczno-polityczny problem, jakim był w Ameryce problem rasizmu. W zaproponowanym przez nich ujęciu inżynieria społeczna była traktowana jako praktyczne działanie prowadzące do otrzymania oczekiwanych zmian[7].

Do prekursorów socjotechniki zaliczany jest również Edward L. Bernays. Jego publikacje, takie jak np. *Propaganda* (1928) czy *Krystalizacja opinii publicznej* (1923), podejmowały problematykę związaną z inżynierią społeczną, a fundamentalna praca dla tej dziedziny *The Engineering of Consent* (1947) była „poświęcona inżynierii społecznej służącej wymuszaniu na ludziach zgody na różne pomysły i przedsięwzięcia[8]". Wykorzystywał w swoich działaniach wiedzę z dziedziny psychologii, socjologii, a także innych nauk społecznych w celu kształtowania opinii publicznej oraz wpływania na zachowania społeczne, co przyczyniło się do rozwoju technik, które są dzisiaj uznawane za elementy inżynierii społecznej i socjotechniki[9].

Mówiąc o historii socjotechniki nie można pominąć prekursora polskiej socjotechniki profesora Adama Podgóreckiego, który rozwijał tę dziedzinę w Polsce w latach 60. i 70. XX w. Niestety ze szkodą dla nauki polskiej szkoła socjotechniki Podgóreckiego napotkała na problemy w 1973 roku, gdy badacze Instytutu Profilaktyki Społecznej i Resocjalizacji Uniwersytetu Warszawskiego zdecydowali się na przesłanie przedstawicielom ówczesnej władzy, którą sprawowała Polska Zjednoczona Partia Robotnicza (PZPR), ekspertyzy zatytułowanej *Diagnostyczny obraz niektórych trudnych problemów społeczeństwa polskiego oraz refleksje socjotechniczne*. W dostarczonym decydentom partyjnym materiale przedstawiono realistyczny obraz polskiego społeczeństwa oraz propozycje działań naprawczych. Działania pracowników naukowych IPSiR wywołały negatywną reakcję politycznych władz państwowych oraz władz Uniwersytetu Warszawskiego. W konsekwencji wobec naukowców Zakładu Socjologii Norm i Patologii Społecznej, stworzonego przez Podgóreckiego w 1976 roku, zastosowano represje. Zakład zlikwidowano przeprowadzając

[6] G. Myrdal, R. Sterner, A. Rose. *An American Dilemma: The Negro Problem and Modern Democracy,* New York, Harper & Brothers Publishers, 1944.
[7] A. Podgórecki, *Zasady socjotechniki*, Wiedza Powszechna, Warszawa 1966, s. 9 – 30.
[8] E.L. Bernays, *Propaganda*, Wektory, Wrocław, 2020, s. 13.
[9] E.L. Bernays, *The Engineering of Consent, Annals of the American Academy of Political and Social Science*, 1947.

jego reorganizację, a Podgóreckiego zmuszono do emigracji do Kanady w 1977 roku. Pomimo że naukowiec zgromadził wokół siebie zespół badawczy zajmujący się tą problematyką, to po jego emigracji rozpoczął się proces rozpadu zespołu, a terminy socjotechnika i inżynieria społeczna przestały w Polsce funkcjonować na wiele lat w dyskursie naukowym jako kategorie akademickie[10].

**Sztuczna inteligencja**

Generatywna sztuczna inteligencja, której rozwoju jesteśmy świadkami, jest rewolucją technologiczną zmieniającą paradygmat życia społecznego. Premiera narzędzia ChatGPT od OpenAI 30 listopada 2022 roku zapoczątkowała globalny, dynamiczny rozwój generatywnej sztucznej inteligencji, wpływając na wszystkie aspekty naszej egzystencji, stopniowo determinując zmiany o trudnych do przewidzenia konsekwencjach dla człowieka. Jednak zanim to nastąpiło, potrzebny był ponad wiek badań i doświadczeń.

Historia sztucznej inteligencji zaczęła się w wyobraźni marzycieli, którzy chcieli sprawić, aby życie ludzi stało się lepsze i łatwiejsze. Jej początek można upatrywać w powstałym w 1822 roku prototypie maszyny różnicowej Charles Babbage'a, która potrafiła rozwiązywać równania wielomianowe[11]. Jednak powszechnie uważa się, że nowoczesna historia sztucznej inteligencji została zainicjowana w koncepcji Alana Turinga, który w 1936 roku w swojej publikacji *On Computable Numbers, with an Application to the Entscheidungsproblem*[12] wykazał, że wszelkie możliwe obliczenia mogą zostać wykonane przez urządzenie matematyczne. Urządzenie to dzisiaj nosi nazwę uniwersalnej *Maszyny Turinga*. Koncepcja *Maszyny Turinga* otworzyła drzwi dla badań nad sztuczną inteligencją. Ustanowiła ramy teoretyczne, na których naukowcy mogli oprzeć dalsze badania nad maszynami myślącymi i rozwiązywać problemy, które wcześniej uważano za domenę wyłącznie ludzkiej inteligencji. Jednak dopiero w 1950 roku Turing, publikując *Computing*

[10] J. Kwaśniewski, *Reorganizacja Czyli rozbicie IPSiR dnia 30 czerwca 1976 roku*, Profilaktyka społeczna i Resocjalizacja, 2014, 24, s.198.
[11] W. Isaacson, *Innowatorzy, o tym jak grupa hakerów, geniuszy i geeków wywołała cyfrową rewolucję*, Insignis Media, Kraków, 2016, s. 27-59.
[12] A. Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, Proceedings of the London Mathematical Society, Tom 42, s. 230-265, 1936. W tej pracy Turing przedstawił koncepcję abstrakcyjnej maszyny, która później stała się znana jako maszyna Turinga. Maszyna Turinga jest teoretycznym modelem obliczeniowym, który potrafi symulować dowolny algorytm komputerowy. Głównym celem tej pracy było dostarczenie formalnej odpowiedzi na tzw. problem Entscheidungsproblem, postawiony przez niemieckiego matematyka Davida Hilberta. Entscheidungsproblem polegał na pytaniu, czy istnieje ogólna procedura decyzyjna dla wszystkich matematycznych problemów, tzn. algorytm, który dla danego stwierdzenia potrafiłby stwierdzić, czy jest ono prawdziwe czy fałszywe.https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf (13.01.2024).

*Machinery and Intelligence,* opisał metodę, która pozwalałaby na sprawdzenie, czy maszyna może wykazać inteligencję na poziomie człowieka[13]. Metoda znana jako *Test Turinga* i polegała na ocenie, czy maszyna może prowadzić konwersację z ludźmi w sposób na tyle przekonujący, że oceniający ten dialog sędziowie nie będą w stanie odróżnić maszyny od człowieka. Chociaż Turing nie stworzył bezpośrednio żadnych systemów sztucznej inteligencji, to jego prace stanowią podwaliny teoretyczne dla tej dziedziny i w znacznym stopniu wpłynęły na rozwój technologii komputerowych, które umożliwiły dalsze postępy w tej dziedzinie[14].

Drugim przełomowym momentem na drodze prowadzącej do powstania sztucznej inteligencji było zorganizowanie latem w 1956 roku na *Dartmouth College* w USA, pierwszej konferencji poświęconej sztucznej inteligencji. Konferencja ta była serią warsztatów badawczych i jest uważana za początek badań nad sztuczną inteligencją. Inicjatorem konferencji był ówczesny profesor matematyki w *Dartmouth College* John McCarthy uznawany obecnie za jednego z twórców tej dziedziny. McCarty zaproponował termin *sztuczna inteligencja*, używając tej nazwy w formalnym dokumencie projektu organizacji warsztatów. Konferencja jest obecnie uznawana za początek badań nad sztuczną inteligencją[15].

**Praktyczne wykorzystanie sztucznej inteligencji**

Sztuczna inteligencja wykorzystywana jest od wielu lat w różnych obszarach aktywności człowieka, chociaż bardzo często nie jesteśmy tego świadomi. Jednak dopiero generatywna sztuczna inteligencja będąca jej zaawansowaną formą zdolna jest do kreowania nowych oryginalnych odpowiedzi na podstawie dostarczonych danych w obszarach tekstu, obrazu, filmu i dźwięku. W przeciwieństwie do wielu wcześniejszych form sztucznej inteligencji, które działały głównie na podstawie zakodowanych w algorytmach reguł, generatywna sztuczna inteligencja może sama kreować unikalne wyniki w oparciu o dostarczone jej dane. Generatywna sztuczna inteligencja jest zdolna do zrozumienia i uwzględnienia kontekstu konwersacji dialogowej i jest w stanie generować odpowiedzi, które są spójne z tym kontekstem. Jest to przełomowe rozwiązanie, gdyż wcześniejsze modele sztucznej inteligencji działały w sposób ograniczony i nie potrafiły analizować oraz rozumieć szerszego kontekstu rozmowy z użytkownikiem. Istotną właściwością generatywnej sztucznej inteligencji jest

[13] A. Turing, *Computing Machinery and Intelligence*, Mind, A Quarterly Review of Psychology and Philosophy, Oxford University Press, 1950, vol. LIX, No. 236, s. 433-460, https://doi.org/10.1093/mind/LIX.236.433 (13.01.2024).
[14] M.A. Boden, *Sztuczna Inteligencja*, Wydawnictwo Uniwersytetu Łódzkiego, 2020, s.20 – 21; s.136 – 137.
[15] https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth (13.01.2024).

również umiejętność uczenia, co pozwala tej technologii szkolić się na nieograniczonych zbiorach danych, a następnie wykorzystywać zdobytą wiedzę do tworzenia odpowiedzi[16].

Sukces technologii udostępnionej przez OpenAI zaskoczył możliwościami, jakie daje przetwarzanie języka naturalnego i zdolnością prowadzenia dialogu człowieka z maszyną, jaka wcześniej była zarezerwowana wyłącznie dla ludzi. Trudno dzisiaj przewidzieć konsekwencje powstałych za przyczyną generatywnej sztucznej inteligencji możliwości, jednak można przypuszczać, że jej potencjał z pewnością zostanie wykorzystany w socjotechnice[17]. Inżynieria społeczna mająca za cel wywieranie wpływu na opinie, postawy i zachowania społeczne od początku swojego istnienia wykorzystuje techniki i strategie mające na celu manipulację świadomością społeczną. Jednak obecnie klasyczne strategie socjotechniczne, takie jak propaganda, manipulacja, dezinformacja, inwigilacja, modelowanie społeczne, nudging, inżynieria konsensusu, profilowanie społeczne czy inżynieria konfliktu otrzymały obecnie bardzo efektywne narzędzie, które z pewnością zostanie wykorzystane w inżynierii społecznej.

W oparciu o obserwację można obecnie podjąć próbę klasyfikacji oddziaływania socjotechnicznego z wykorzystałem nowych technologii w dwóch obszarach. *Frontendu Socjotechnicznego* i *Backendu Socjotechnicznego.* Propozycja klasyfikacji koncentruje się na sposobie wykorzystania nowych technologii. *Frontend Socjotechniczny* obejmuje narzędzia i zewnętrzne formy przekazu, które bezpośrednio oddziałują na percepcję i emocje użytkowników w *jawny sposób*. W tym obszarze technologia oddziałuje na zmysły i świadomość społeczną poprzez takie formy przekazu jak *fakenews, photofake, deepfake oraz voicefake*. *Backend Socjotechniczny,* w którym technologia jest wykorzystywana do działań socjotechnicznych w *sposób ukryty* i niezauważalny przez użytkowników. W tym obszarze technologie są stosowane do nieautoryzowanego monitoringu zachowań użytkowników.

## ***Frontend Socjotechniczny*. Syntetyczne formy przekazu**

*Frontend Socjotechniczny* obejmuje metody komunikacji, które bezpośrednio i otwarcie oddziałują na użytkowników. Wśród tych metod znajdują się między innymi fałszywe informacje tekstowe określane terminem *fake news*, które są jedną z najbardziej rozpowszechnionych form dezinformacji, manipulacji i szerzenia propagandy. Pierwszym udokumentowanym przykładem celowo zaprojektowanego fake newsa jest cykl artykułów zamieszczonych w gazecie New York Sun w 1835 roku, znany jest jako Wielkie *Kłamstwo o*

[16] S. Russell, P. Norvig, *Sztuczna Inteligencja. Nowe spojrzenie*, Tom 1 i 2, Helion, Gliwice, 2023.
[17] https://openai.com/ (13.01.2024).

*Księżycu.* 25 sierpnia 1835 roku gazeta istniejąca niecałe 2 lata zamieściła pierwszy z cyklu artykułów zatytułowany *Wielkie odkrycia astronomiczne dokonane ostatnio przez Sir Johna Herschela,* opisujący istniejące życie na księżycu[18]. Definicja *fake newsa* obejmuje różne aspekty, jednak wspólną cechą zawsze jest celowe wprowadzanie w błąd, manipulacja i dezinformacja[19].

Kolejną formą manipulacji informacją, tym razem za pośrednictwem obrazu, jest *photofake*. Jest to metoda, która za sprawą rozwoju generatywnej sztucznej inteligencji stała się bardzo popularna i umożliwia modyfikację obrazu w niezauważalny dla odbiorcy sposób. Historia manipulacji obrazem fotograficznym sięga XIX wieku i najwcześniejszych prac wykonywanych ręcznie na szklanych i blaszanych płytach fotograficznych w celu nie tylko poprawy estetyki zdjęć, ale także tworzenia nowych efektów wizualnych wcześniej niemożliwych do wykonania w innych technikach wizualnych[20]. Jednym z najwcześniejszych i najbardziej znanych przykładów manipulacji obrazem fotograficznym z epoki przedcyfrowej było zdjęcie prezydenta Abrahama Lincolna z około 1860 roku, które przedstawiało obraz prezydenta Stanów Zjednoczonych skomponowane z ciałem polityka Johna Calhouna. Zdjęcie to uznawane jest za jeden z pierwszych przykładów manipulacji informacją wizualną w historii fotografii[21]. Manipulacja fotografią znalazła szczególne zastosowanie w systemach totalitarnych, takich jak komunizm i nazizm. W tych reżimach fotografia stała się narzędziem propagandy i kontroli społecznej. W Związku Radzieckim, pod przywództwem Józefa Stalina, manipulacja zdjęciami była często wykorzystywana do celów propagandowych. Przykładem tego jest usunięcie Leona Trockiego z fotografii, po tym jak popadł w niełaskę u Stalina. Praktyka ta, znana jako *damnatio memoriae*, miała na celu wymazywanie niewygodnych osób z historii i pamięci społecznej. Natomiast w nazistowskich Niemczech manipulacja zdjęciami służyła głównie celom propagandowym i antysemickim. Adolf Hitler i odpowiedzialny za propagandę III Rzeszy Joseph Goebbels wykorzystywali fotografię do kreowania wizerunku silnego zjednoczonego narodu oraz do demonizowania osób narodowości żydowskiej oraz innych mniejszości[22].

Kolejną formą przekazu informacyjnego tworzoną w oparciu o nowoczesne technologie, a mającą za zadanie wprowadzanie w błąd opinii publicznej, jest *deepfake*. Jest to

[18] https://www.lindahall.org/about/news/scientist-of-the-day/richard-adams-locke/ (13.01.2024).
[19] K. Bąkowicz, *Wprowadzenie do definicji i klasyfikacji zjawiska fake newsa*, Studia Medioznawcze, tom 20, nr 3 (78), 2019, ISSN 2451-1617, ORCID: 0000-0001-6365-2696.
[20] N. Rosenblum, *A World History of Photography*, Abbeville Press, 1984.
[21] J. Sharma & R. Sharma, *Analysis of Key Photo Manipulation Cases and their Impact on Photography*, 2017.
[22] https://encyclopedia.pub/entry/30879 (13.01.2024).

syntetyczny przekaz filmowy, którego treści zbudowane są z materiału wideo oraz audio. Termin deepfake pojawił się po raz pierwszy w przestrzeni publicznej w grudniu 2017 jako pseudonim użytkownika działającego na platformie *Reddit.* Historia nazwy *deepfake* rozpoczęła się, gdy użytkownik opublikował filmy stworzone przy użyciu technologii *deep learning*, w których wizerunki znanych osób zostały przedstawione w materiałach pornograficznych w sytuacjach, które nigdy nie miały miejsca. Gdy witryna internetowa *Motherboard* zaczęła komentować i analizować to nowe zjawisko, zaczęło być ono popularne, zyskując międzynarodową rozpoznawalność i od tego czasu termin *deepfake* funkcjonuje w przestrzeni publicznej[23]. Obecnie generatywna sztuczna inteligencja jest używana coraz częściej do modyfikowania i tworzenia syntetycznych materiałów audio i wideo w taki sposób, że coraz trudniej je odróżnić od oryginału. Technologia ta otworzyła nowe możliwości dla socjotechniki, gdyż siła *deepfake'a* pozwala na bardzo skuteczną manipulację opinią publiczną. Obserwacja rozwoju tego zjawiska pokazuje, że wykorzystanie *deepfake'ów* w celach manipulacji będzie rosło, a wygenerowane syntetyczne treści będzie coraz trudniej odróżnić od prawdziwych informacji[24]. Jest to zjawisko niepokojące oraz niebezpieczne, gdyż syntetyczne fałszywe informacje w postaci *deepfake'ów* mogą rujnować reputację, prestiż, wizerunek i markę zarówno osób fizycznych, jak i prawnych. Mogą dezorganizować działalność firm i instytucji, powodując nieodwracalne straty, co potwierdza, jak niebezpieczną bronią może być ta technologia[25].

Jednym z najbardziej znanych przykładów *deepfake'a* rozpowszechnionego w Internecie była wypowiedź prezydenta USA, Barack'a Obamy z roku 2018. Technologia pozwoliła na stworzenie realistycznie brzmiącego głosu prezydenta oraz mimiki twarzy Obamy, co spowodowało, że film wyglądał na autentyczny. Był to jeden z pierwszych tak znanych przykładów wykorzystania *deepfake'a* do stworzenia fałszywego przekazu politycznego w celu pokazania, jak technologia cyfrowa może zostać wykorzystana[26]. Kolejnym przykładem znanego deepfake'a jest film z 2020 roku pokazujący północnokoreańskiego przywódcę autorytarnego komunistycznego reżimu Kim Jong-una, który wypowiada się pozytywnie na temat wartości demokratycznych[27]. Kolejnym *deepfakiem,* który spowodował dezorientację opinii publicznej, była informacja wygenerowana w marcu

[23] https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley (13.01.2024).
[24] https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/?sh=58bb0b1b7494 (13.01.2024).
[25] A. Kopciuch, *Deepfake jako nowa broń w walce informacyjnej*, Fundacja Bezpieczna Cyberprzestrzeń, 2021.
[26] https://www.youtube.com/watch?v=cQ54GDm1eL0 (13.01.2024).
[27] https://www.youtube.com/watch?v=ERQlaJ_czHU (13.01.2024).

2022 roku przez Rosję, przedstawiająca ukraińskiego prezydenta Wołodymyra Zełenskiego ogłaszającego kapitulację i przekazanie władzy Rosji. Wideo to mogło mieć poważne implikacje polityczne i międzynarodowe, gdyby w porę nie zostało rozpoznane i zdementowane jako fałszywa wiadomość *deepfake*. Był to jeden z pierwszych przykładów wykorzystania *deepfake'a* podczas konfliktu zbrojnego na skalę międzynarodową[28]. W każdym z powyższych przypadków sztuczna inteligencja i generatywna sztuczna inteligencja odegrały kluczową rolę w tworzeniu tych syntetycznych przekazów. Możliwość dokładnego naśladowania mimiki twarzy, gestów, ruchu ciała, ruchów ust, tonu głosu stanowi coraz potężniejsze narzędzie manipulacji, które w każdej chwili może być wykorzystane w celach politycznych, społecznych lub komercyjnych.

Jednak w najbliższej przyszłości niezwykle niebezpieczną technologią cyfrową wykorzystującą generatywną sztuczną inteligencję może okazać się technologia służąca do klonowania głosu, czyli *voicefake*. Technologia pozwalająca na klonowanie głosu jest obecnie znacznie tańsza od technologii generującej *deepfake,* co w najbliższej perspektywie może okazać się bardzo niebezpiecznym, skutecznym narzędziem manipulacji i oszustw. W tym kontekście popularne *oszustwo na wnuczka*[29] może przybrać nową, masową formę, gdzie przestępcy dzięki generatywnej sztucznej inteligencji będą w czasie rzeczywistym wykorzystywać do prowadzenia rozmów głosy bliskich osób, co będzie trudne do weryfikacji.

Na potwierdzenie możliwości sklonowanych głosów można przytoczyć przykład oszustwa dokonanego w 2019 roku za pomocą tej metody, jeszcze przed rozpowszechnieniem generatywnej sztucznej inteligencji. Przypadek został opisany przez *Wall Street Journal* i pokazuje, jak wielkie niebezpieczeństwo niesie ze sobą ten typ formy przekazu. *WSJ* opisał historię przestępców, którzy za pomocą słabej sztucznej inteligencji sklonowali głos szefa wielkiej firmy i zadzwonili do jego podwładnych z poleceniem przelania kwoty 220 tys. euro, a podwładni polecenie to wykonali, sądząc, że rozmawiają ze swoim szefem[30]. Przykład ten pokazuje, że w przyszłości nie tylko *deepfake,* ale także *voicefake* mogą stać się poważnym zagrożeniem dla zaufania w relacjach między ludźmi oraz ludźmi i instytucjami. Technologia klonowania głosu pozwala obecnie stworzyć syntetyczny duplikat oryginału na podstawie bardzo krótkiej próbki prawdziwego nagrania posiadającego charakterystyczne parametry

[28] https://www.youtube.com/watch?v=X17yrEV5sl4 (13.01.2024).
[29] B. Jewartowski, *Bezpieczeństwo osób starszych w kontekście oszustw metodą „na wnuczka" jako problem społeczno-polityczny*, Studia Politicae Universitatis Silesiensis, T. 17, ISSN 1895-3492, 2016.
[30] C. Stupp, *Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case*, The Wall Street Journal, 2019. https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (13.01.2024).

mowy właściciela, takie jak ton, barwę, intensywność, tempo i inne subtelne cechy akustyczne[31]. Mimo że *fake news, photofake, deepfake* czy *voicefake* mogą stanowić odrębne formy manipulacji, to obserwuje się tendencję do powszechnego określania wszystkich nieprawdziwych syntetycznych form przekazu terminem *deepfake* jako ogólnego określenia na wszelkie rodzaje fałszywych syntetycznych materiałów stworzonych za pomocą sztucznej inteligencji.

**Backend Socjotechniczny. Cyfrowa inwigilacja**

Technologie takie jak *fake news, deepfake, photofake* czy *voicefake* stwarzają widoczne i zauważalne realne niebezpieczeństwo dla komunikacji społecznej, jednak o wiele bardziej niebezpieczną jest technologia, która w ukryciu jest wykorzystywana do powszechnej cyfrowej inwigilacji. Systemy stworzone w celu nadzorowania i śledzenia naszej aktywności w świecie realnym i wirtualnym, którymi zarządza *Backend Socjotechniczny,* dzięki rozwojowi generatywnej sztucznej inteligencji pozwalają dzisiaj na stałą cyfrową nieautoryzowaną inwigilację w czasie rzeczywistym. Dysponenci tych technologii coraz częściej ograniczają naszą prywatność, uzasadniając te działania najczęściej koniecznością zapewnienia bezpieczeństwa publicznego. W tym kontekście cyfrowa algorytmiczna inwigilacja oznacza nie tylko monitorowanie działalności w przestrzeni cyfrowej, ale również obserwacje naszej aktywności w przestrzeni publicznej świata realnego najczęściej bez naszej wiedzy i zgody[32]. Zwłaszcza branża mediów społecznościowych, której fundamentem działania jest eksploracja danych o użytkownikach i ich zachowaniach, jest zainteresowana rozwojem tej technologii, ponieważ pozwala to na jeszcze bardziej precyzyjne profilowanie odbiorcy dla treści reklamowych. Umożliwia to prowadzenie jeszcze bardziej skutecznych kampanii marketingowych, które przekładają się w konsekwencji na jeszcze większe zyski.

Przykładami nieautoryzowanego wykorzystania eksploracji danych są skandale związane z *Cambridge Analytica* oraz działania ujawnione przez *Edwarda Snowdena*. *Cambridge Analytica* była firmą konsultingową z Wielkiej Brytanii specjalizującą się w eksploracji i analizie danych dla potrzeb kampanii politycznych. W roku 2018 spowodowała skandal inwigilacyjny, prowadząc nieautoryzowane i nielegalne profilowanie użytkowników sieci społecznościowych. Wykorzystując dane około 90 milionów użytkowników Facebooka,

[31] https://spidersweb.pl/plus/2023/05/klonowanie-glosu-audio-deep-fake (13.01.2024).
[32]B. Schneier, *Dane i Goliat. Ukryta bitwa o Twoje dane i kontrola nad światem*, Wydawnictwo Helion, Gliwice, 2017.

firma stworzyła precyzyjne profile psychologiczne tych osób bez ich wiedzy i zgody. Następnie firma używała pozyskane w ten sposób dane do budowania modeli pozwalających na precyzyjne profilowanie użytkowników, w tym określenie ich przekonań politycznych, religijnych cech osobowości oraz przewidywanie ich opinii i reakcji na konkretne kampanie. Informacje te były następnie wykorzystywane w projektach politycznych, m.in. w kampanii na rzecz promowania wyjścia Wielkiej Brytanii z Unii Europejskiej (brexit) oraz w wyborach prezydenckich w USA w 2016 roku, w których *Cambridge Analytica* wspierała Donalda Trumpa. Skandal Cambridge Analytica stał się symbolem nadużyć w zakresie wykorzystania danych, inwigilacji cyfrowej i wpływania na decyzje polityczne poprzez wykorzystanie profilowania psychologicznego[33].

Kolejnym przykładem niewłaściwego wykorzystywania danych była afera związana z ujawnieniem przez Edwarda Snowdena w 2013 roku informacji o globalnych programach nadzoru prowadzonych przez Agencję Bezpieczeństwa Narodowego (NSA) i inne amerykańskie agencje wywiadowcze. Ujawnione przez Snowdena informacje dotyczyły nadzoru rządowego i masowej eksploracji danych przez agencje rządowe, w tym monitorowanie komunikacji elektronicznej obywateli na całym świecie. Konsekwencje ujawnienia tych praktyk wywołały międzynarodowy skandal i debatę na temat prywatności, bezpieczeństwa i granic nadzoru rządowego nad obywatelami. Działania Snowdena doprowadziły do zmian w prawie i polityce, w tym do przyjęcia przez USA *Freedom Act* w 2015 roku, który ograniczył zakres zbierania danych przez NSA[34].

Podsumowując, eksploracja danych w kontekście rozwoju generatywnej sztucznej inteligencji nabierać będzie coraz większego znaczenia w kontekście możliwości precyzyjnego profilowania społecznego, które może być również wykorzystywane w projektach socjotechnicznych. Możliwość analizy ogromnych ilości danych w czasie rzeczywistym i danych historycznych o naszych aktywnościach w świecie wirtualnym, takich jak transakcje finansowe, zakupy, przeglądanie stron internetowych, aktywność na profilach społecznościowych oraz korzystanie z innych usług online, pozwalać będzie na niezwykle dokładne profilowanie użytkowników. Ułatwi to dobieranie reklam do odpowiedniej grupy użytkowników, ale również otworzy pole do nadużyć w zakresie zaawansowanej nieautoryzowanej inwigilacji na szeroką skalę.

---

[33] J. Fernando, *Cambridge Analytica: Overview, History*, Example, Investopedia, https://www.investopedia.com/terms/c/cambridge-analytica.asp (13.01.2024).
[34] B. Gellman, *Dark Mirror: Edward Snowden and the American Surveillance State*, Penguin Press, 2020.

**Social Credit System. Totalna inwigilacja**

Szczególnie niepokojące są postępy w precyzyjnym profilowaniu społecznym z wykorzystaniem generatywnej sztucznej inteligencji, gdyż umożliwiają głęboką analizę zachowań, preferencji i nawyków osób inwigilowanych, co może być, jak pokazały przykłady *E. Snowdena* i *Cambridge Analytica*, wykorzystane do zaawansowanych projektów inżynierii społecznej, takich jak wdrażany obecnie w Chińskiej Republice Ludowej znany jako *Social Credit System.*

Historia *Social Credit System* sięga wczesnych lat 90. XX wieku. Projekt początkowo był próbą rozwoju systemów bankowości osobistej pozwalających na ocenę zdolności kredytowej mieszkańców Chin. Wynikało to z faktu, że dynamicznie rozwijające się Chiny, chcąc ożywić konsumpcję, potrzebowały zaangażowania ekonomicznego mieszkańców w zakresie zaciągania zobowiązań finansowych w postaci pożyczek i kredytów. Niestety bez posiadania historii kredytowej nie było to możliwe. W tamtym czasie jedynie 300 mln osób z 1,5 mld obywateli Chin posiadało taką historię. Dlatego, aby pomóc mieszkańcom Chin szybko zbudować zdolność kredytową, wymyślono system, który w oparciu o ocenę zachowań społecznych miał budować ich zdolność do zaciągania zobowiązań. Koncepcja systemu opierała się na punktacji (*social scoring*). Za pozytywne zachowania obywatele będą nagradzani punktami w rankingu, dzięki czemu ich możliwości kredytowe będą rosły, a za złe zachowania punkty oceny w rankingu będą odejmowane, obniżając zdolność kredytową. Na pewnym etapie projektowania i rozwoju systemu władze autorytarnych Chin uznały jednak, że pomysł ten jest na tyle interesujący, że można go również wykorzystać do zbudowania systemu, który będzie praktycznym narzędziem do efektywnego zarządzania postawami, opiniami i zachowaniami chińskich obywateli oraz podmiotów gospodarczych. W 2007 roku zostały opracowane przez chiński rząd wstępne koncepcje projektu. W 2009 roku rozpoczęto lokalne próby wdrażania systemu w wybranych regionach, a w 2011 roku władze Chin podjęły formalną decyzję o rozpoczęciu procesu wprowadzania systemu w całym kraju. W roku 2014 rozpoczęto oficjalny krajowy pilotaż systemu, a w 2018 uruchomiono system *czerwonych i czarnych* list będących rejestrami osób i organizacji z wysokimi i niskimi wynikami punktacji. W roku 2019 nadal kontynuowano pilotaż i testy różnych funkcji systemu, jednak pandemia COVID-19 i problemy technologiczne opóźniły uruchomienie ogólnokrajowego systemu. W 2023 roku system nadal nie był scentralizowany i znajdował się w fazie wdrażania i testów, ale z nadal aktualnym planem wdrożenia systemu w całym kraju. *Social Credit System* korzysta z zaawansowanej technologii cyfrowej. Można przewidywać, że system budowany będzie

również w oparciu o sztuczną inteligencję, jednak brak jest dowodów na szerokie użycie tej technologii.

*Social Credit System* jest złożonym systemem i może posiadać różne powiązane ze sobą elementy w zależności od miejsca pilotażu. Jednak charakterystyczne są trzy powiązane ze sobą elementy systemu, takie jak: *scoring, czerwone i czarne listy* oraz *podział obywateli na kategorie A, B, C, D*. Scoring który jest elementem bazowym systemu, jest mechanizmem wskazującym, jak *dobry* lub *zły, godny zaufania* lub *niegodny* zaufania jest obywatel lub firma w kontekście społecznym i gospodarczym. Posiadanie wysokiej punktacji za oczekiwane zachowania społeczne przez osoby fizyczne daje wiele korzyści, takich jak łatwiejszy dostęp do kredytów, niższe rachunki za usługi komunalne, priorytet w procesach rekrutacyjnych i awansów, dostęp do lepszych ofert ubezpieczeniowych, większa szansa na dostanie paszportu, preferencyjne traktowanie w wynajmie mieszkań i nieruchomości, łatwiejszy dostęp do usług medycznych, uprzywilejowane traktowanie w hotelach i na lotniskach. Wysoka punktacja za właściwe zachowanie podmiotów prawnych daje łatwiejszy dostęp do finansowania i kredytów, niższe stawki ubezpieczeniowe, priorytet w procesach przetargowych i zamówień publicznych, preferencyjne traktowanie w relacjach z urzędami i instytucjami państwowymi, dostęp do specjalnych programów wsparcia i dotacji rządowych, uprzywilejowane warunki w działalności eksportowej i międzynarodowej.

Charakterystycznym przykładem wdrożenia *scoringu* zachowań społecznych jest miasto *Rongcheng*, które w 2013 roku wprowadziło pilotaż systemu. Miasto zaczęło przyznawać każdemu mieszkańcowi podstawową ocenę kredytową w wysokości 1000 punktów, na którą mogły wpływać jego dobre i złe zachowania. Przykładowo wprowadzone zasady oceny zakładały, że rozpowszechnianie szkodliwych informacji na komunikatorze *WeChat* czy innych forach i blogach będzie karane odjęciem 50 punktów, podczas gdy przykładowo wygranie zawodów sportowych lub kulturalnych na poziomie krajowym będzie skutkowało dodaniem 40 punktów[35].

Drugim elementem systemu są *czerwone i czarne listy*, czyli rejestry obywateli i firm tworzone w oparciu o posiadany poziom punktacji za zachowania społeczne. Listy są formalnym i scentralizowanym sposobem oceny obywateli i przedsiębiorstw wynikających z zawieranych umów o współpracy pomiędzy jednostkami administracji rządowej, a ich celem jest skuteczna egzekucja prawa i eliminacja aspołecznych i niebezpiecznych zachowań. Osoby z wysoką punktacją trafiają na *dobre – czerwone listy*, a z niską punktacją na *złe – czarne listy*.

[35] MIT Technology Review https://www.technologyreview.com/2022/11/22/1063605/china-announced-a-new-social-credit-law-what-does-it-mean/ (13.01.2024).

Katalog przestępstw i naruszeń, za które osoby fizyczne mogą trafić na czarne listy, zawiera przykładowo takie przewinienia, jak niepłacenie alimentów, naruszenie zasad ruchu drogowego, niepłacenie mandatów i kar, uchylanie się od płacenia podatków, zaniechania w opiece nad osobami starszymi, zakłócanie porządku publicznego, umieszczanie fałszywych informacji w mediach społecznościowych, niechodzenie z psem na smyczy, niespłacone pożyczki, zaśmiecanie przestrzeni publicznej, zażywanie narkotyków, fałszywe reklamy. Przykładowy katalog przestępstw i naruszeń, za które osoby prawne mogą trafić na czarne listy, obejmuje m.in. niewywiązywanie się z obowiązków podatkowych, nieprzestrzeganie regulacji dotyczących ochrony środowiska, oszustwa korporacyjne i finansowe, nieprzestrzeganie standardów jakości i bezpieczeństwa produktów, nieetyczne praktyki zatrudnienia, niewywiązywanie się z umów handlowych i kontraktów, nieprawidłowe zarządzanie odpadami i substancjami toksycznymi, nieprzestrzeganie praw własności intelektualnej, nieprzestrzeganie regulacji dotyczących bezpieczeństwa i higieny pracy, nieprawidłowe lub fałszywe raportowanie finansowe[36].

Kolejnym elementem systemem zaufania społecznego jest *kategoryzacja obywateli*. Pilotażowo miasto *Suining* w prowincji Jiangsu w Chinach wprowadziło eksperymentalny system etykietowania, który na podstawie wyników punktacji i rodzaju listy, na której się znajdowali, podzielił obywateli na cztery kategorie A,B,C,D. Jednak eksperyment etykietowania nawiązywał do wojennych praktyk kategoryzacji ludności podczas wojny z Japonią i wywołał wiele kontrowersji, co spowodowało, że został ostatecznie zaniechany[37]. System *kategoryzacji obywateli* wprowadzony w *Suining* zaliczał do kategorii A modelowych obywateli, godnych zaufania, posiadających najwyższą ilość punktów i znajdujących się na czerwonych listach. Obywatele w tej grupie byli również nagradzani różnymi benefitami, takimi jak niższe oprocentowanie kredytów, priorytetowe traktowanie w instytucjach publicznych czy łatwiejszy dostęp do lepszych miejsc pracy. Do grupy B zaliczano obywateli, którzy uznawani byli za godnych średniego zaufania. Takie osoby mogły korzystać z niektórych benefitów, ale były również poddawane pewnym ograniczeniom, na przykład miały utrudniony dostęp do pewnych form kredytów lub usług. W grupie C klasyfikowani byli obywatele, którzy uznawani byli za niewystarczająco zaufanych. Osoby zaliczone do tej grupy napotkały już większe przeszkody w codziennym życiu niż obywatele z kategorią A i B. Miały

[36] Report, *Social credit & big data trends in China*, Innovation Centre Denmark, Shanghai 2018. file:///C:/Users/48516/Downloads/2018_Social-credit-and-big-data-trends-in-China.pdf (13.01.2024).
[37] Report, *China's social credit system and its development: between 'Orwellian nightmare' and technocratic utopia?*, Asia Research Center, Center for Security Studies, War Studies University, Warszawa, 2020.

przykładowo problemy w uzyskaniu kredytu, wynajęciu mieszkania czy nawet w zakupie biletów na pociąg dużych prędkości lub samolot. Ostatnią kategorią obywateli były osoby należące do grupy D. Do tej kategorii klasyfikowano osoby mające niewystarczającą ilość punktów scoringu społecznego oraz były umieszczone na czarnej liście rankingowej oceny obywateli. Obywatele, którzy mieli kategorię D, to osoby pozbawione zaufania społecznego i były poddawane najbardziej surowym sankcjom. Takie osoby mogły być wykluczane z wielu aspektów życia społecznego i gospodarczego, włączając w to zakaz zatrudnienia w pewnych branżach oraz możliwość korzystania z publicznych usług[38]. Należy jednak zaznaczyć, że chociaż Social *Credit System* według standardów demokratycznego świata jest nieakceptowalny etycznie, to badania pokazują, że 80% obywateli Chińskiej Republiki Ludowej ocenia go pozytywnie lub wysoce pozytywnie, a jedynie 1% wyraża dla niego dezaprobatę[39].

**Podsumowanie**

Rozwój generatywnej sztucznej inteligencji oraz narzędzi cyfrowych, na których opierają się syntetyczne przekazy nowej generacji, stanowi zarówno szansę, jak i zagrożenie. Możliwość analizy ogromnych zbiorów danych w czasie rzeczywistym oraz precyzyjne profilowanie społeczne otwierają nowe perspektywy dla cyfrowej inwigilacji. Sztuczna inteligencja może zostać wykorzystana do manipulacji i dezinformacji opinii publicznej, a także do nieautoryzowanej inwigilacji, co w konsekwencji może podważać zaufanie do demokratycznych instytucji. Projekt *Social Credit System* jest skrajnym przykładem zastosowania technologii cyfrowych do modelowania społecznego, rodząc pytania o granice ingerencji technologii w prywatność obywateli. Istnieje ryzyko, że inżynieria społeczna, wyposażona w generatywną sztuczną inteligencję oraz wiedzę socjologiczną, może w przyszłości wywierać ogromny wpływ na zachowania społeczne, postawy ludzi oraz opinię publiczną, umożliwiając efektywne projektowanie interakcji społecznych poprzez skuteczne modelowanie społeczeństwa. Zachodzi także obawa, że sztuczna inteligencja, wraz z rozwijającymi się dzięki niej narzędziami, otwiera nowy rozdział w inżynierii społecznej,

[38] W. Pawłowski, *Orwell w chińskim wydaniu*, Polityka, 2018, https://www.polityka.pl/tygodnikpolityka/swiat/1759826,1,orwell-w-chinskim-wydaniu.read (13.01.2024).
[39] G. Kostka, *China's Social Credit Systems and Public Opinion: Explaining High Levels of Approval*, Freie Universität Berlin, SSRN Electronic Journal, , 2018, DOI: 10.2139/ssrn.3215138, https://www.researchgate.net/publication/328467411_China%27s_Social_Credit_Systems_and_Public_Opinion_Explaining_High_Levels_of_Approval (13.01.2024).

niekoniecznie bezpieczny dla człowieka. Biorąc to pod uwagę, konieczne będzie poszukiwanie rozwiązań i kształtowanie takiej przyszłości, w której sztuczna inteligencja będzie służyła rozwojowi społeczeństwa w sposób bezpieczny i etyczny, na rzecz dobra wspólnego. W obliczu tak dynamicznie zachodzących zmian odpowiedzialne zarządzanie tą technologią i jej wpływem na społeczeństwo będzie kluczowym wyzwaniem.

## Bibliografia


Bąkowicz K., *Wprowadzenie do definicji i klasyfikacji zjawiska fake newsa*, Studia Medioznawcze, 2019.
Bernays E.L., *Krystalizacja opinii publicznej*, Narodowe Centrum Kultury, Warszawa 2019.
Bernays E.L., *Propaganda*, Wektory, Wrocław 2020.
Bernays E.L., *The Engineering of Consent, Annals of the American Academy of Political and Social Science*, 1947.
Boden M.A., *Sztuczna Inteligencja*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2020.
Earp E.L., The Social Engineer, New York Eaton & Mains, 1911.
Eastman M., *Marxism: Is it Science?,* New York, The Dial Press, 1935.
Gellman B., *Dark Mirror: Edward Snowden and the American Surveillance State*, Penguin Press, 2020.
Isaacson W., *Innowatorzy, o tym jak grupa hakerów, geniuszy i geeków wywołała cyfrową rewolucję*, Insignis Media, Kraków 2016.
Jewartowski B., *Bezpieczeństwo osób starszych w kontekście oszustw metodą „na wnuczka" jako problem społeczno-polityczny*, Studia Politicae Universitatis Silesiensis, T. 17, 2016, ISSN 1895-3492.
Kopciuch A., *Deepfake jako nowa broń w walce informacyjnej*, Fundacja Bezpieczna Cyberprzestrzeń, 2021.
Kostka G., *China's Social Credit Systems and Public Opinion*: Explaining High Levels of Approval, Freie Universität Berlin, SSRN Electronic Journal, 2018, DOI: 10.2139/ssrn.3215138.
Kwaśniewski J., *Reorganizacja Czyli rozbicie IPSiR dnia 30 czerwca 1976 roku*, Profilaktyka społeczna i Resocjalizacja, 2014.
Lippmann W., *Opinia publiczna*, Fundacja Lethe, Kraków 2020.
M. Eastman, *Marxism: Is it Science?* New York, The Dial Press, 1935.
Myrdal G., Sterner R., Rose A., *An American Dilemma: The Negro Problem and Modern Democracy*, New York, Harper & Brothers Publishers, 1944.
Podgórecki A., *Zasady socjotechniki*, Wiedza Powszechna, Warszawa 1966.
Popper K.R., *Społeczeństwo otwarte i jego wrogowie*, Wydawnictwo Naukowe PWN, Warszawa 2010.
Popper K.R., *The Open Society and Its Enemies,* Princeton, Princeton University Press, 1945.
Pound R., *Interpretations of Legal History*, Cambridge, Cambridge University Press, 1923.
Pound R., *Introduction to the Philosophy of Law*, New Haven, Yale University Press, 1922.
Report. *China's social credit system and its development: between 'Orwellian nightmare' and technocratic utopia?*, Asia Research Center, Center for Security Studies, War Studies University, Warszawa 2020.
Rosenblum N., *A World History of Photography*, Abbeville Press, 1984.
Russell S., Norvig P., *Sztuczna Inteligencja. Nowe spojrzenie*, Helion, Gliwice 2023.
Schneier B., *Dane i Goliat. Ukryta bitwa o Twoje dane i kontrola nad światem*, Wydawnictwo Helion, Gliwice 2017.
Turing A., *Computing Machinery and Intelligence, Mind: A Quarterly Review of Psychology and Philosophy*, Oxford University Press, 1950.
Turing A., *On Computable Numbers, with an Application to the Entscheidungsproblem*, Proceedings of the London Mathematical Society,1936.


### Źródła internetowe


*Artificial Intelligence Coined at Dartmouth.* Pobrano z https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth (Dostęp: 22.05.2024).
BuzzFeedVideo. (2018). *You won't believe what Obama says in this video!* [Wideo]. YouTube. Pobrano z https://www.youtube.com/watch?v=cQ54GDm1eL0 (Dostęp: 22.05.2024).
Cole, S. (2018). *We are truly fucked: Everyone is making AI-generated fake porn now. Vice.* Pobrano z https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley (Dostęp: 22.05.2024).

Fernando, J. (2021). *Cambridge Analytica: Overview, History, Example*: Investopedia. Pobrano z https://www.investopedia.com/terms/c/cambridge-analytica.asp (Dostęp: 22.05.2024).
https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/?sh=58bb0b1b7494 (Dostęp: 22.05.2024).
J. Fernando, J. (2021). *Cambridge Analytica: Overview, History, Example*. Investopedia, Pobrano z https://www.investopedia.com/terms/c/cambridge-analytica.asp (Dostęp: 22.05.20024).
Kostka, G. (2018). *China's Social Credit Systems and Public Opinion*: Explaining High Levels of Approval. SSRN Electronic Journal. DOI: 10.2139/ssrn.3215138. Pobrano z https://www.researchgate.net/publication/328467411_China%27s_Social_Credit_Systems_and_Public_Opinion_Explaining_High_Levels_of_Approval (Dostęp: 22.05.2024).
Langguth, J., Pogorelov, P., Brenner, B., Filkuková, P., (2021). *Thilo Schroeder, D.T., Don't Trust Your Eyes: Image Manipulation in the Age of DeepFakes*. Pobrano z https://www.frontiersin.org/articles/10.3389/fcomm.2021.632317/full (22.05.2024).
Lock, R. A. (2022). Linda Hall Library. Pobrano z https://www.lindahall.org/about/news/scientist-of-the-day/richard-adams-locke/ (Dostęp: 22.05.2024).
Myre, G. (2023). A decade on, Edward Snowden remains in Russia, though U.S. laws have changed. CT Public. Pobrano z https://www.ctpublic.org/2023-06-04/a-decade-on-edward-snowden-remains-in-russia-though-u-s-laws-have-changed (Dostęp: 22.05.2024)
OpenAI. https://openai.com (Dostęp: 22.05.2024).
*Photo Manipulation.* Pobrano z https://encyclopedia.pub/entry/30879 (Dostęp: 22.05.2024).
*Popularność żadnej innej aplikacji nie rosła tak szybko. 100 milionów użytkowników w rekordowo krótkim czasie. (2023).* Pobrano z https://tvn24.pl/biznes/ze-swiata/chatgpt-100-milionow-uzytkownikow-w-rekordowo-krotkim-czasie-st6764540 (Dostęp: 22.05.2024).
RepresentUs. (2020). *Dictators - Kim Jong-Un* [Wideo]. YouTube. Pobrano z https://www.youtube.com/watch?v=ERQlaJ_czHU (Dostęp: 22.05.2024).
Soll, J. (2016). *The Long and Brutal History of Fake News*. Pobrano z https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535/ (Dostęp:22.05.2024).
Stupp, C. (2019). *Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case* Pobrano z https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (Dostęp: 22.05.2024).
The Telegraph. (2022). *Deepfake video of Volodymyr Zelensky surrendering surfaces on social media* [Wideo]. YouTube. Pobrano z https://www.youtube.com/watch?v=X17yrEV5sl4 (Dostęp: 22.05.2024).
*Toews, R. (2020). Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared. Pobrano z https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/.*
*Photo Manipulation*, https://encyclopedia.pub/entry/30879 (Dostęp: 13.01.2024).
*Don't Trust Your Eyes: Image Manipulation in the Age of DeepFakes*, https://www.frontiersin.org/articles/10.3389/fcomm.2021.632317/full (Dostęp:13.01.2024).
*We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now,* https://www.vice.com/en/article/bjye8a/reddit-fake-porn-app-daisy-ridley (Dostęp:13.01.2024).
*Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared,* https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/?sh=58bb0b1b7494 (Dostęp:13.01.2024).
https://www.lindahall.org/about/news/scientist-of-the-day/richard-adams-locke/ (13.01.2024).
https://www.youtube.com/watch?v=cQ54GDm1eL0 (Dostęp:13.01.2024).
https://www.youtube.com/watch?v=ERQlaJ_czHU (Dostęp:13.01.2024).
https://www.youtube.com/watch?v=X17yrEV5sl4 (Dostęp:13.01.2024).
https://www.youtube.com/watch?v=5Fv-LKT_cEc (Dostęp:13.01.2024).
*Tede i Peja pogodzeni, czyli do czego (nie) doprowadzi klonowanie głosów*, https://spidersweb.pl/plus/2023/05/klonowanie-glosu-audio-deep-fake (Dostęp:13.01.2024).
C. Stupp, Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case, The Wall Street Journal, 2019. https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (Dostęp:13.01.2024).
J. Fernando, *Cambridge Analytica: Overview, History, Example, Investopedia,* https://www.investopedia.com/terms/c/cambridge-analytica.asp (Dostęp:13.01.2024).
MIT Technology Review https://www.technologyreview.com/2022/11/22/1063605/china-announced-a-new-social-credit-law-what-does-it-mean/ (Dostęp:13.01.2024).
W. Pawłowski, *Orwell w chińskim wydaniu*, Polityka, 2018, https://www.polityka.pl/tygodnikpolityka/swiat/1759826,1,orwell-w-chinskim-wydaniu.read (Dostęp:13.01.2024).

IFSEC Insider, *Role of CCTV Cameras: Public, Privacy and Protection*, 2021. https://www.ifsecglobal.com/video-surveillance/role-cctv-cameras-public-privacy-protection/ (Dostęp:13.01.2024).
*A World With a Billion Cameras Watching You Is Just Around the Corner,* Wall Street Journal, https://www.wsj.com/articles/a-billion-surveillance-cameras-forecast-to-be-watching-within-two-years-11575565402 (Dostęp:13.01.2024).
Report. *Social credit & big data trends in China*, Innovation Centre Denmark, Shanghai 2018 file:///C:/Users/48516/Downloads/2018_Social-credit-and-big-data-trends-in-China.pdf (Dostęp:13.01.2024).
Sharma J. & Sharma R., *Analysis of Key Photo Manipulation Cases and their Impact on Photography*, 2017, https://www.semanticscholar.org/paper/Analysis-of-Key-Photo-Manipulation-Cases-and-their-Sharma-Sharma/0eb7d2b171b55770e722edc332bb815ab3a3d7f7 (Dostęp:13.01.2024).

**Social Engineering in the Age of Generative Artificial Intelligence**

Summary

The article presents the evolution of social engineering in the context of the dynamic development of generative artificial intelligence. Initially, social engineering was considered a scientific field representing the practical application of sociological knowledge. However, today it is often perceived as a tool for manipulation. The development of generative artificial intelligence and its use in classic social engineering techniques and strategies may shift sociology's focus from merely studying and describing social behaviors to practically initiating and directing social changes. The author proposes defining two areas of social engineering impact where advanced digital technologies are used and introduces the term "synthetic communication." The concept of "Frontend Social Engineering" is introduced to encompass the explicit and visible activities of digital algorithms, while "Backend Social Engineering" refers to the covert operations of digital technologies. The article analyzes synthetic forms of communication such as fake news (text), photofake (image), voicefake (voice), and deepfake (image and voice), which are increasingly used to manipulate opinions, attitudes, and social behaviors. The article also addresses issues related to the use of digital technologies for unauthorized digital surveillance.

Keywords: artificial intelligence, social engineering, deepfake, manipulation, surveillance