Edyta Sadowska

Uniwersytet Pedagogiczny w Krakowie

Studia Wschodnioeuropejskie

Nr ekspercki 19-t. 2/2023

Ewolucja Cyberzagrożeń: Deepfake i Media Syntetyczne w kontekście bezpieczeństwa energetycznego Europy Wschodniej

Wstęp

Współczesna era informacyjna, związana z dynamicznym rozwojem technologii cyfrowych, w szczególności systemów generatywnej sztucznej inteligencji (GAI), znacząco poszerza zakres potencjalnych cybezagrożeń[1]. Tym samym rosnące ryzyko wystąpienia działań mogących prowadzić do dezinformacji może mieć wpływ na kluczowe sektory gospodarki, takie jak infrastruktura energetyczna[2]. Europa Wschodnia, jako strategiczny obszar geopolityczny, nie pozostaje odseparowana od tej nowej rzeczywistości. W kontekście rosnącej niestabilności politycznej i geoekonomicznej regionu, podmioty zarówno państwowe, jak i niepaństwowe, zyskują możliwości wykorzystania zaawansowanych technik dezinformacyjnych. Przejście od tradycyjnych metod dezinformacyjnych do wykorzystania mediów syntetycznych stanowi kolejny etap ewolucyjny wojen informacyjnych, których skutki mogą dotknąć nie tylko przestrzeń wirtualną, ale także kluczowe sektory gospodarki realnej[3].

---

[1] Cyt: *Deepfakes are evolving from mere research and academic tools to a weapon to create high-stake warfare for creating social discord, increasing polarization, and, in some cases influencing the election outcome.* A. Jaiman, Deepfakes & Synthetic Media. Humanity at the Edge of an Uncanny Valley, s, 7.

[2] M. S. Rogers, D. Weinstein, *Protecting our critical infrastructure in the digital age,* on-line: https://thehill.com/opinion/cybersecurity/447596-protecting-our-critical-infrastructure-in-the-digital-age/, [10.07.2023]

[3] Cyt: *The nature of Information Warfare and its connections with changes in the information ecosystem, before moving to the main relevant types of synthetic media, discussing their potential Information Warfare uses and weighing their pros and cons, subsequently evaluating the likelihood of nefarious use by Russia and other aggressive states.* J. Kalpokiene, I. Kalpokas, *Synthetic Media and Information Warfare: Assessing Potential Threats,* The Russian Federation in Global Knowledge Warfare, ss. 33–50 por**.** H. Nasu**,** *Deepfake technology in the age of information warfare*, Lieber Institute, on-lain: https://lieber.westpoint.edu/deepfake-technology-age-information-warfare/, [14.07.2023] por. Ch. Perez, *Information Warfare in Russia's War in Ukraine. The Role of Social Media and Artificial Intelligence in Shaping Global Narratives*, on-line: https://foreignpolicy.com/2022/08/22/information-warfare-in-russias-war-in-ukraine/ [14.07.2023].

Media syntetyczne, w szczególności technika deepfake, umożliwiają manipulację treściami wideo i dźwiękowymi, mogącą prowadzić do rozprzestrzeniania autentycznie wydających się treści dezinformacyjnych, które są trudne do odróżnienia od rzeczywistych wydarzeń[4]. Szczególnego znaczenia w ostatnim czasie nabiera minimalny próg wejścia kompetencyjnego, pozwalający na tworzenie tego typu treści. Narzędzia nie wymagają specjalnych umiejętności programowania czy kompetencji z obszaru STEM. Oznacza to, że każdy użytkownik/użytkowniczka komputera z dostępem do Internetu może wygenerować łudząco przypominający autentyczny obraz, tekst czy video. Takie działania bazując na ludzkim okulocentrycznym pojmowaniu prawdy mogą w łatwy sposób wprowadzić w błąd jednostki, a za pomocą szybkiego rozpowszechniania informacji poprzez media społecznościowe dotrzeć do znaczącej liczby odbiorców[5].

Media syntetyczne wprowadzają tym samym nową broń do arsenału cyberzagrożeń. W sektorze energetycznym, złożonym i wrażliwym pod względem infrastruktury i ekonomicznej stabilności, tego rodzaju działania dezinformacyjne mogą prowadzić zarówno do zakłóceń operacyjnych, ale także utraty zaufania wśród interesariuszy oraz, w najbardziej ekstremalnych przypadkach, wywołania destabilizującego wpływu na całą gospodarkę i bezpieczeństwo państwa[6]. Niniejszy artykuł ma na celu analizę zagrożeń związanych z wykorzystaniem mediów syntetycznych, ze szczególnym uwzględnieniem wykorzystania mediów syntetycznych takich jak: techniki deepfake, w sektorze energetycznym Europy Wschodniej[7]. Istotnym założeniem bazowym jest podjęcie holistycznego spojrzenia na rosnące zagrożenia wynikające z nowoczesnych technik dezinformacyjnych w sektorze energetycznym Europy Wschodniej, które mają kluczowe znaczenie dla zachowania

[4] Y. Hwang , Ji Youn Ryu, Se-Hoon Jeong , Effects of Disinformation Using Deepfake: The Protective Effect of Media Literacy Education, Cyberpsychol Behav Social Network, por. I. Kalpokas, Problematising reality: the promises and perils of synthetic media, SN Social Sciance (2021), on-line: https://link.springer.com/content/pdf/10.1007/s43545-020-00010-8.pdf, [14.07.2023] por. D. A. Coccomini, R. Caldelli, F. Falchi, Cross-Forgery Analysis of Vision Transformers and CNNs for Deepfake Image Detection, Proceedings of the 1st International Workshop on Multimedia AI against Disinformation (MAD'22), June 27–30, 2022, Newark, NJ, USA. ACM, New York, NY, USA, on-lain: https://arxiv.org/pdf/2206.13829.pdf, [14.07.2023]

[5] Cyt. *On the other extreme, people can doubt the trouth of anythink they hear from the gouverment, corporation, friends. […] RAND call this phenomenon trouth decay. Foreign adversaries and nefarious domestic actors actively use disinformation to undemine governdego ment trust.* A.Jaiman, *Deepfakes*…dz.cyt. s. 5.

[6] R. Wang, Z. Huang, Z. Chen, *Anti-Forgery: Towards a Stealthy and Robust DeepFake Disruption Attack via Adversarial Perceptual-aware Perturbations*, on-lain: https://arxiv.org/pdf/2206.00477.pdf, [07.07.2023].

[7] Z Adams, M. Osman, Ch. Bechlivanidis, B. Meder, *(Why) Is Misinformation a Problem? Perspectives on Psychological Science* 1–28, on-line: https://eprints.whiterose.ac.uk/197813/3/17456916221141344.pdf, [07.07.2023].

stabilności regionalnej, bezpieczeństwa państw oraz integralności infrastruktury krytycznej. Poprzez analizę możliwości wykorzystania oraz perspektyw rozwoju zagrożeń ze strony mediów syntetycznych artykuł może przyczynić się do szerokiej i transdyscyplinarnej debaty nad skutecznymi środkami przeciwdziałania ewoluującym zagrożeniom cybernetycznym. W niniejszym artykule wykorzystano metodę analizy i syntezy literatury naukowej w szczególności anglojęzycznej celem zrozumienia obecnych zagrożeń i wykorzystania mediów syntetycznych.

Znaczen Problematyka cyberzagrożeń w sektorze energetycznym Europy Wschodniej

Współczesne wyzwania bezpieczeństwa, związane z rozwojem technologii oraz postępującą automatyzacją systemów – także tych wchodzących w skład infrastruktury krytycznej niosą za sobą nowe rodzaje zagrożeń. Podatny na nowego typu zagrożenia staje się także sektor energetyczny, będący kręgosłupem gospodarki i infrastruktury każdego kraju[8]. Jak słusznie zauważyli A. Zając, R. Balina D. Kowalski: Bezpieczeństwo energetyczne każdego kraju jest jednym z głównych czynników jego prawidłowego funkcjonowania (ang. The energy security of each country is one of the main factors of its proper functioning)[9]. Kluczowym elementem bezpieczeństwa energetycznego jest z jednej strony zaspokojenie potrzeb energetycznych i utrzymanie stabilności dostaw, ale także utrzymanie poczucia zaufania zarówno odbiorców jak i inwestorów. Zapewnienie bezpiecznych dostaw energii ma wpływ niemalże na każdy obszar funkcjonowania kraju: począwszy od zapewnienia ciągłości dostarczania energii dla gospodarki, poprzez rozwój naukowo technologiczny, ogólny rozwój ekonomiczny, czy utrzymanie potencjału militarnego. Przyjmując szeroką i multisektorową perspektywę bezpieczeństwo energetyczne zapewnia zdolność państwa do zapewnienia stabilnego dostępu do energii niezbędnej do podtrzymywania funkcji gospodarczych, społecznych i politycznych[10]. W ostatnich latach perspektywy w obszarze bezpieczeństwa

[8] Cyt. *For years, experts have expressed growing concerns that critical infrastructures are vulnerable to cyberattacks almost anywhere in the world. Remotely transmitted malware and viruses can compromise interconnected power systems, airports, hospitals, military systems, and anywhere internetconnected computers are used to manage critical functions. As the world heads toward more digitalization, integrating new communication technologies offers unlimited possibilities for the proper accumulation of information, creating a fundamental dependence on proper functioning in all areas of society. As such reliance could offer a great opportunity for a smart and efficient world, it also poses significant threats to the operation of critical infrastructures and vital facilities* T. M. Aljohani, *Cyberattacks on Energy Infrastructures: Modern War Weapons*, on-line: https://arxiv.org/ftp/arxiv/papers/2208/2208.14225.pdf, [07.07.2023].

[9] A. Zając, R. Balina D. Kowalski *Financial and Economic Stability of Energy Sector Enterprises as a Condition for Poland's Energy Security—Legal and Economic Aspects*, Energies 2023, 16(3), 1442.

[10] Cyt. *Ensuring energy stability highlights the responsibility for society, long-term development, lasting cooperation, protection of the natural environment, consumer protection, producer protection, etc* N. Dimitrov,

energetycznego wskazują naglącą konieczność dostosowania strategii bezpieczeństwa energetycznego do zmieniającego się kontekstu, zwłaszcza w odniesieniu do technologii. Wzrastająca liczba podłączonych do sieci urządzeń IoT oraz zastosowanie technologii inteligentnych sieci energetycznych (smart grids) stanowią atrakcyjne cele dla ataków cybernetycznych. Jednocześnie należy pamiętać, że rozwój technologii może prowadzić do powstawania coraz bardziej wyrafinowanych i trudnych do zatrzymania form ataków cybernetycznych. Przyszłość bezpieczeństwa energetycznego będzie wymagała stanowczego połączenia wysiłków na poziomie międzynarodowym, z wykorzystaniem najnowszych osiągnięć w dziedzinie cyberbezpieczeństwa.

Ataki cybernetyczne stanowią dziś niezwykle zaawansowaną i dynamicznie ewoluującą formę zagrożenia, w szczególności kiedy zagrażają infrastrukturze krytycznej państwa stanowią coraz poważniejszy problem dla decydentów i konsumentów[11]. Postęp technologiczny, a zwłaszcza jego tempo umożliwia przestępcom wykorzystywanie coraz to nowszych technik i narzędzi w czasie, kiedy możliwości ich wykrycia są dopiero opracowywane. Wraz z dalszym rozwojem trudno określić jednoznaczną perspektywę ewolucji cyberzagrożeń, a rosnąca zależność społeczna od infrastruktury cyfrowej oraz rozwijający się internet rzeczy (IoT – Internet of things) czyni ten obszar niezwykle podatnym na użycie. Coraz częściej mówi się zatem o cyberbezpieczeństwie jako jednej z istotnych składowych bezpieczeństwa praktycznie na każdym poziomie podmiotowym począwszy od bezpieczeństwa jednostki, przez bezpieczeństwo społeczne, narodowe czy międzynarodowe. W konsekwencji, odpowiedź na rosnące spectrum zagrożeń w cyberprzestrzeni musi być holistyczna i dynamiczna. Konieczne jest nie tylko stałe adaptowanie technologii obronnych, wspierane przez wykorzystanie sztucznej inteligencji do wykrywania chociażby różnego rodzaju anomalii, ale przede wszystkim zwrócenie uwagi na palące potrzeby legislacyjne, przede wszystkim na płaszczyźnie międzynarodowej. W świetle tych wyzwań, przyszłość ataków cybernetycznych będzie stanowiła test dla zdolności społeczeństw i instytucji do

Methods for Enhancing and Evaluation of Energy Security and Economic Growth in Bulgaria, Economic Alternatives, 2019, Issue 4, s. 526.

[11] Cyt. *However, risks to these essential technology systems are increasing—in particular, malicious actors are becoming more willing and capable of carrying out cyberattacks. Such attacks could result in serious harm to human safety, national security, the environment, and the economy. Agencies and critical infrastructure owners and operators must protect the confidentiality, integrity, and availability of their systems and effectively respond to cyberattacks. Cybersecurity high-risk series: Challenges in Protecting Cyber Critical Infrastructure*, Gouverment Accessibility Office, on-lie: https://www.gao.gov/assets/gao-23-106441.pdf, [12.08.2023]

skutecznego przeciwdziałania zagrożeniom, zapewniając jednocześnie ciągłą innowację w obszarze cyberbezpieczeństwa.

Obszar Europy Wschodniej, staje obecnie przed wyjątkowymi wyzwaniami związanymi z cyberzagrożeniami, także w kontekście destabilizacji związanej z wojną w Ukrainie i Rosyjską agresją. które zdobywają nowe narzędzia na froncie cybernetycznej walki. W ostatnich latach rośnie też znaczenie zagrożeń cybernetycznych ukierunkowanych na sektor energetyczny Europy Wschodniej. Działania te obejmują szeroki zakres ataków, od operacji wywiadowczych i sabotażu po szpiegostwo przemysłowe i kradzież poufnych informacji[12], ale także zagrożenia związane z potencjalnymi atakami cybernetycznymi i wojną informacyjną. Infrastruktura energetyczna regionu, której znaczące elementy stanowią elektrownie, linie przesyłowe, magazyny paliwa i systemów dystrybucji, jest kluczowym celem dla agresywnych podmiotów dążących do osłabienia struktury państwowej czy też uzyskania przewagi strategicznej. Nowe, złożone techniki i narzędzia, wspomagane przez rozwijającą się technologię mogą prowadzić nie tylko do utraty kontroli nad systemami energetycznymi, ale mogą powodować skutki daleko wykraczających poza obszar sektora energetycznego. Sektor energetyczny Europy Wschodniej, pełniący kluczową rolę w utrzymaniu stabilności gospodarczej i strategicznej, staje obecnie w obliczu rosnących zagrożeń, które mają potencjał nie tylko osłabić integralność infrastruktury, ale znacząco zakłócić funkcjonowanie rynków energetycznych i wpłynąć na bezpieczeństwo wewnętrzne poszczególnych krajów. Ataki na infrastrukturę energetyczną stwarzają bowiem wyjątkowo poważne ryzyko dla dostaw energii, stabilności gospodarczej i bezpieczeństwa publicznego. Wśród zagrożeń warto wspomnieć chociażby o zagrożeniach natury operacyjnej, takich jak ataki DDoS (Distributed Denial of Service), które mogą prowadzić do zakłóceń w operacjach energetycznych poprzez przeładowanie systemów i sieci przesyłowych, wykorzystując techniki ransomware, które polegają na zaszyfrowaniu systemów i żądaniu okupu za ich odblokowanie[13]. Innym działaniem może być zainfekowanie systemów SCADA (Systemów

[12] D.C. Smith, *Enhancing cybersecurity in the energy sector: a critical priority*, Journal of Energy & Natural Resources Law Volume 36, 2018 - Issue 4, on-line: https://www.tandfonline.com/doi/epdf/10.1080/02646811.2018.1516362?needAccess=true&role=button, [12.08.2023] por. D.C. Smith, *Cybersecurity in the energy sector: are we really prepared?*, Journal of Energy & Natural Resources Law Volume 39, 2021.

[13] W grudniu 2015 roku doszło do ataku cybernetycznego na elektrownię jądrową w Ukraińskim Czarnobylu. Atakujący wykorzystali zaawansowane malware, aby wyłączyć systemy sterowania elektrowni, co skutkowało przerwą w dostawach energii elektrycznej na znacznym obszarze. Ten incydent ilustruje potencjał ataków na infrastrukturę energetyczną, które mogą prowadzić do zakłóceń operacyjnych oraz destabilizacji.

Kontroli i Akwizycji Danych) które może doprowadzić do wyłączania elektrowni, nadmiernego obciążenia sieci czy też wstrzymywania dostaw energii. W efekcie, ataki te mogą skutkować poważnymi konsekwencjami zarówno dla funkcjonowania państwa i gospodarki, ale także życia codziennego jednostek. Działania zmierzające do prób infiltracji w systemy sterowania, poprzez wyłączanie poszczególnych elementów infrastruktury, aż po zniszczenie lub uszkodzenie urządzeń krytycznych mogą prowadzić do destabilizacji rynków energetycznych, manipulacji cenami energii, wprowadzeniem chaosu w mechanizmy dystrybucji oraz wywołania niepewności w relacjach międzynarodowych. Konwergencja tych działań może prowadzić do długoterminowych skutków ekonomicznych i politycznych, wpływając na zaufanie do instytucji rządowych i przedsiębiorstw energetycznych. Przykładem takiej destabilizacji rynków może być celowa manipulacja informacjami o dostawach paliwa czy awariach w elektrowniach, generująca szereg spekulacji na rynku, prowadząc do fluktuacji cen. W efekcie, region może stać się bardziej podatny na szantaż gospodarczy, co może przyczynić się do destabilizacji ekonomicznej, politycznej i społecznej. Patrząc szerzej na to działanie, może ono mieć wpływ na kształtowanie opinii publicznej, skłaniając ją do działań niezgodnych z interesem narodowym, co w perspektywie i przy zaostrzeniu sytuacji, a także szybkim rozpowszechnianiu się informacji może stanowić zagrożenie dla stabilności i suwerenności państw regionu.

Galopująca ewolucja technologiczna doprowadza do tego, że dezinformacja staje się zagadnieniem o coraz bardziej złożonym charakterze, a nowoczesne techniki, takie jak media syntetyczne i deepfake'i, odgrywają coraz ważniejszą rolę w kontekście ataków cybernetycznych i wojny informacyjnej[14]. Media syntetyczne mają kluczowe znaczenie dla propagowania dezinformacji, otwierając nowe i dostępne dla wszystkich możliwości manipulacji treściami[15]. Wojna informacyjna wydaje się być szczególnie istotna w kontekście Europy Wschodniej[16]. Dezinformacja oparta na deepfake'ach i mediach generowanych

[14] C. Vaccari, A. Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception*, Uncertainty, and Trust in News, Social Media + Society January-March 2020: 1–13, on-line: https://journals.sagepub.com/doi/pdf/10.1177/2056305120903408, [12.08.2023].

[15] R. Millière, *Deep learning and synthetic media*, on-line: https://arxiv.org/pdf/2205.05764.pdf, [12.08.2023].

[16] Cyt. *Russia, throughout history has experimented with its information along with cyber channels and have been able to generate fruitful results that are manifested in the contemporary era. Russian expertise towards merging cyber domain into its military capabilities is praiseworthy to the point that its Western adversaries, despite the fact being economically and technologically advanced compared to Russia, have not been able to effectively counter Russian aggression.* A.Rashid, A. Yar Khan, S.Wasif Azim, *Cyber hegemony and information warfare: A case of Russia*, Liberal Arts & Social Sciences International Journal (LASSIJ), Vol. 5, No. 1, (January-June 2021): 648-666, on-line:

syntetycznie może prowadzić do wywołania fałszywych kryzysów, szybkiego rozprzestrzeniania dezorientujących informacji (za pomocą mediów społecznych), czy też podważania wiarygodności instytucji rządowych[17]. W kontekście tych problemów, pytania badawcze, które towarzyszą analizie, stają się kluczowe. Jakie są konkretnie formy zagrożeń cybernetycznych dla sektora energetycznego Europy Wschodniej? W jaki sposób agresywne podmioty wykorzystują nowoczesne techniki dezinformacyjne, w tym deepfake do osiągnięcia swoich celów w regionie? Jakie są potencjalne skutki ekonomiczne, polityczne i społeczne wynikające z udanych ataków cybernetycznych i dezinformacyjnych w sektorze energetycznym?

Znaczenie mediów syntetycznych i deepfake'ów w kontekście dezinformacji i ataków cybernetycznych

W dzisiejszym świecie, w którym informacje i komunikacja odgrywają kluczową rolę, rozwój technologii stawia przed nami nowe wyzwania i możliwości. Jak zauważa O. Wasiuta i S. Wasiuta: *Zazwyczaj skuteczną metodą wojny informacyjnej jest uwolnienie dezinformacji lub podanie informacji w korzystny sposób dla agresora. Metody tepozwalają na zniekształcenie oceny tego, co się dzieje, demoralizację obywateli i potencjalnie zapewniają przejście na stronę agresora informacyjnego*[18]. Wojna informacyjna[19], stanowiąca element wojny hybrydowej[20], nabiera szczególnego znaczenia w kontekście sektora energetycznego Europy Wschodniej[21], której region, obdarzony strategicznym znaczeniem geopolitycznym, stał się polem rywalizacji państw i aktorów pozapaństwowych. Wojna informacyjna, jako forma agresji niemilitarnej, stwarza nową jakość zagrożeń poprzez wpływanie na opinię

https://pdfs.semanticscholar.org/c53e/04fcae8573bec52c358aad566be8a6449a26.pdf?_gl=1*j6k4hl*_ga*NzM5MDM3MTMyLjE2ODQ3NjU3NDk.*_ga_H7P4ZT52H5*MTY5MTU2NjIwNy45LjEuMTY5MTU2NzQ4NS41NS4wLjA, [12.08.2023].

[17] H. Etienne, *The future of online trust (and why Deepfake is advancing it),* AI and Ethics (2021) 1:553–562on-line: https://link.springer.com/content/pdf/10.1007/s43681-021-00072-1.pdf, [12.08.2023].

[18] O. Wasiuta, S. Wasiuta, *Wojna informacyjna zagrożeniem dla bezpieczeństwa ludzkości*, Repozytorium Uniwersytet Pedagogiczny https://rep.up.krakow.pl/xmlui/bitstream/handle/11716/2026/06--Wojna-informacyjna-zagrozeniemWasiuta--Wasiuta.pdf?sequence=1&isAllowed=y, [12.08.2023].

[19] C. B. Lewis *Information Warfare*, on-line: https://irp.fas.org/eprint/snyder/infowarfare.htm, [12.08.2023].

[20] A. Bilal, *Hybrid Warfare – New Threats, Complexity, and 'Trust' as the Antidote,* NATO Review, on-line: https://www.nato.int/docu/review/articles/2021/11/30/hybrid-warfare-new-threats-complexity-and-trust-as-the-antidote/index.html, [12.08.2023] por. S. D. Bachmann *Hybrid Wars: The 21st-century new threats to global peace and security*, Scientia Militaria, South African Journal of Military Studies, Vol 43, No. 1, 2015, ss. 77 – 98.

[21] Szerzej: M. Snegovaya, *Putin's information warfare inUkraine Soviet origins of Rusiia's hybrid warfare*, Institute for the Study of War, 2015.

publiczną, kreowanie napięć społecznych i podsycanie konfliktów[22]. Atakujący wykorzystują narzędzia wojny informacyjnej, w tym dezinformację i propagandę, aby wprowadzić chaos, dezorientację i nieufność. W sektorze energetycznym, dezinformacja może dotyczyć fałszywych informacji na temat dostaw energii, planów strategicznych czy też stanu technicznego infrastruktury. Manipulując percepcją społeczeństwa i decydentów, atakujący mogą wpłynąć na kształtowanie polityki energetycznej, destabilizować relacje m iędzynarodowe i wprowadzić dezorganizację w zarządzaniu kryzysowym[23].

W tym kontekście, media syntetyczne i technika deepfake wyróżniają się jako narzędzia o znaczącym potencjale wpływania na społeczeństwo, politykę i gospodarkę. Dziś, pojęcia takie jak deepfake'i i media syntetyczne stały się nieodłącznym elementem dyskusji na temat manipulacji treściami wizualnymi i dźwiękowymi. Nowoczesne technologie oparte na sztucznej inteligencji, które pozwalają na tworzenie autentycznie wyglądających treści multimedialnych, mogą być trudne a coraz częściej niemożliwe do odróżnienia od rzeczywistych źródeł. Deepfake to neologizm powstały z połączenia słów "deep learning" (głęboka nauka) i "fake" (fałszywy). N. Schick definiuje deepfaki jako media zmanipulowane lub w całości wygenerowane przez sztuczną inteligencję[24]. Technika deepfake polega na wykorzystaniu algorytmów uczenia maszynowego, zwłaszcza głębokiego uczenia, do tworzenia fałszywych treści wideo i dźwiękowych. Przy użyciu dostarczonych danych, takich jak wideo źródłowe z osobą, która ma być sfabrykowana, algorytmy te są w stanie generować dowolny przekaz na przykład wideo, w którym twarz i głos osoby źródłowej są przypisywane innej postaci lub sytuacji. Deepfake'i w prosty sposób mogą służyć celom dezinformacyjnym a w sektorze energetycznym, mogą być wykorzystane do stworzenia fałszywych wideo prezentujących chociażby awarie infrastruktury, ataki na elektrownie czy fałszywe wypowiedzi liderów politycznych lub ekspertów. Te wideo mogą być trudne do wykrycia

[22] Szerzej: B. D. Johnson, A. Draudt, i.inn. Information warfare and the future of conflict, Arizona State University, on-line: https://threatcasting.asu.edu/sites/default/files/2020-07/threatcasting-2020-The%20Future%20of%20Information%20Warfare-WEB_0.pdf, [12.08.2023].

[23] Cyt: India, April 2018: A video goes viral on WhatsApp, the world's most popular mobile instant messaging platform. The footage, seemingly from a CCTV camera, shows a group of children playing cricket in the street. Suddenly, two men on a motorbike ride up and grab one of the smallest kids then speed away. This "kidnapping" video creates widespread confusion and panic, spurring an 8-week period of mob violence that kills at least nine innocent people). C.Vaccari, A.Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News,* Social Media + Society January-March 2020: 1–13.

[24] Cyt. *A deepfake is a type of synthetic media, meaning media (including images, audio and video) that is either manipulated or wholly generated by AI*. N. Schock, *Deepfakes and the Infoapokalypse. What you urgently need to know*. London 2020, s. 8.

jako fałszywe, co sprawia, że za pomocą mediów społecznościowych w erze infoapokalipsy osiągają natychmiastowy zamierzony efekt dezinformacyjny wprowadzając destabilizację w społeczeństwie, prowadzącą do dezorientacji i reakcji paniki. Tego rodzaju manipulacja wizualna ma potencjał wywołania niekontrolowanego chaosu i zakłóceń w funkcjonowaniu sektora energetycznego. W sektorze energetycznym Europy Wschodniej, te nowoczesne technologie mogą stanowić element coraz bardziej złożonej układanki zagrożeń otwierając nowe horyzonty dla wykorzystania tych narzędzi w celach dezinformacyjnych.

Wraz z rozwojem narzędzi takich jak modele GPT 3, 3.5, 4 (do generowania treści), modele DALL-E/ Midjourney/ Lensa/ Stable Diffusion do generowania obrazu czy szeregu podobnych aplikacji do generowania video (Wonder Studio, Synthesia) i dźwięku (Prime Voice AI) kluczowe staje się pytanie o przyszłość środowiska bezpieczeństwa związanego z płaszczyzną komunikacji informacyjnej. Na pierwszy plan wysuwa się niezwykle istotny dla bezpieczeństwa aspekt łatwości generowania treści syntetycznych, a w szczególności audio, foto i video. Szeroki i niemalże darmowy dostęp użytkowników do narzędzi pozwalających na generowanie treści syntetycznych (deep fake'ów i fake newsów) wymusza konieczność reakcji zarówno legislacyjnej jak i edukacyjnej zmierzającej do zmiany mocno okulocentrycznego dotąd paradygmatu prawdy[25]. Unia Europejska w przygotowywanym Akcie o sztucznej inteligencji (AI act) wpisała systemy takie jak ChatGPT na listę systemów wysokiego ryzyka[26], których dostawcy będą musieli brać odpowiedzialność za skutki ich użycia, określając że kraje w ciągu 18 miesięcy od uchwalenia regulacji, będą musiały wdrożyć ramy dotyczące etycznej AI do swoich systemów prawnych. Co istotne w kontekście deepfaków generowanych za pomocą AI: łatwy próg wejścia – nie wymagający specjalistycznych kompetencji, a także generowanie treści o spersonalizowanych charakterze pogłębia zagrożenia. Możliwe staje się sfałszowanie informacji używając do tego głosu bliskich lub znanych nam osób – co uautentycznia materiał. Tego typu działania mogą doprowadzić do chaosu informacyjnego, ale także stanowią doskonałe narzędzie do manipulacji społecznej i zwiększenia zarówno skali, jak i efektywności ataków cybernetycznych opartych na personalizowanych treściach syntetycznych.

[25] S. Venkataramakrishnan (2019). *Can you believe your eyes? How deepfakes are coming for politic,* Financial Times, 24.X. 2019 on-lain: https://www.ft.com/content/4bf4277c-f527-11e9-a79c-bc9acae3b654, [12.08.2023] por. *When seeing is no longer believing*, CNN, on-lain: https://edition.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/, [12.08.2023].

[26] *Akt ws. sztucznej inteligencji: pierwsze przepisy regulujące sztuczną inteligencję Aktualności Parlament Europejski*, on-lain: https://www.europarl.europa.eu/news/pl/headlines/society/20230601STO93804/akt-ws-sztucznej-inteligencji-pierwsze-przepisy-regulujace-ai, [12.08.2023].

W kontekście sektora energetycznego, media syntetyczne mogą zostać wykorzystane do stworzenia realistycznych wizualizacji infrastruktury energetycznej, które sugerują potencjalne zagrożenia lub awarie. Przykładowo, atakujący mogą wytworzyć wideo przedstawiające fałszywy wyciek substancji toksycznych z elektrowni, prowadząc do dezinformacji i paniki społecznej[27]. Rola mediów syntetycznych polega bowiem na stwarzaniu iluzji rzeczywistości, co może prowadzić do zniekształcenia percepcji rzeczywistych zagrożeń, opierając się na fałszywych przemówieniach, wywiadach czy wypowiedziach ekspertów. Takie działania mogą prowadzić do kreowania fałszywych narracji na temat planów energetycznych, dostaw energii czy awarii w sektorze energetycznym, wprowadzając dezorientację whhmśród decydentów i opinii publicznej, co może wpływać na procesy decyzyjne, a nawet destabilizować obszar energetyczny. Efektem tego rodzaju dezinformacji może być panika, spekulacje na rynkach energetycznych oraz dezorganizacja działań zarządzania kryzysowego.

Nowe wyzwania dla sektora energetycznego

Rola mediów syntetycznych i deepfake'ów w kontekście sektora energetycznego Europy Wschodniej jest zatem w szczególności związana z zagrożeniem dezinformacji, manipululacji informacji i destabilizacji. W obszarze energetycznym, gdzie bezpieczeństwo dostaw energii jest kluczowe dla stabilności społecznej i gospodarczej, media syntetyczne i deepfake'i stanowią istotne wyzwanie. Otwierają one drzwi do fałszerstw wizualnych i dźwiękowych, które są trudne do wykrycia, a co za tym idzie, trudne do obrony przed nimi. W związku z tym, koniecznym elementem strategii w sektorze energetycznym musi być opracowanie skutecznej obrony przed tymi zagrożeniami.

Jednym z kluczowych aspektów skutecznej walki z dezinformacją jest budowania i ciągłe zwiększenie świadomości na temat potencjalnych zagrożeń związanych z mediami syntetycznymi i deepfake'ami. Edukacja na temat technik manipulacji wideo i dźwięku oraz metod wykrywania fałszywych treści może pomóc w zwiększeniu odporności na dezinformację. Ponadto, konieczne jest wzmocnienie systemów monitorowania i wykrywania potencjalnych ataków opartych na mediach syntetycznych i deepfake'ach. Zaawansowane narzędzia analizy wizualnej i dźwiękowej mogą pomóc w identyfikacji fałszywych treści oraz szybkiej reakcji na potencjalne zagrożenia. Współpraca międzynarodowa odgrywa kluczową

[27] *Ukraine's Lessons for the Future of Hybrid Warfare,* RAND Corporation, on-line: https://www.rand.org/blog/2022/11/ukraines-lessons-for-the-future-of-hybrid-warfare.html, [12.08.2023].

rolę ponieważ państwa powinny wymieniać się informacjami, najlepszymi praktykami oraz wspólnie opracowywać strategie obrony przed tymi zagrożeniami, a instytucje międzynarodowe mogą odegrać ważną rolę w regulacji i monitorowaniu działań cyberprzestępczych oraz dezinformacyjnych. Strategia wymaga kompleksowego podejścia, obejmującego edukację, monitorowanie, wykrywanie i współpracę międzynarodową. Sektor energetyczny Europy Wschodniej musi być gotów na stawienie czoła nowym wyzwaniom, jakie niosą za sobą media syntetyczne i deepfake'i, aby utrzymać integralność infrastruktury, stabilność rynków i bezpieczeństwo narodowe.

Streszczenie:

Dynamiczny rozwój technologii cyfrowych, zwłaszcza generatywnych systemów sztucznej inteligencji (GAI), wprowadza szereg nowych zagrożeń dla bezpieczeństwa państwa. Niniejszy artykuł naukowy ma na celu analizę rosnącego ryzyka dezinformacji związanego z mediami syntetycznymi, szczególnie z techniką deepfake, w sektorze energetycznym Europy Wschodniej. Omówiony zostanie nowy aspekt wojen informacyjnych jakim stanowią media syntetyczne oraz wyzwania, jakie niesie manipulacja treściami wizualnymi i dźwiękowymi.

Słowa kluczowe:

Media syntetyczne, deepfake, dezinformacja, sektor energetyczny, Europa Wschodnia, wojna informacyjna, bezpieczeństwo energetyczne, zagrożenia cybernetyczne, technologia, sztuczna inteligencja.

Keywords:

Synthetic media, deepfake, misinformation, energy sector, Eastern Europe, information warfare, energy security, cyber threats, technology, artificial intelligence

Bibliografia:

1. Adams Z., Osman M., Bechlivanidis Ch., Meder B., (Why) Is Misinformation a Problem? Perspectives on Psychological Science 1–28, on-line: https://eprints.whiterose.ac.uk/197813/3/17456916221141344.pdf, [07.07.2023].
2. Akt ws. sztucznej inteligencji: pierwsze przepisy regulujące sztuczną inteligencję Aktualności Parlament Europejski, on-lain: https://www.europarl.europa.eu/news/pl/headlines/society/20230601STO93804/akt-ws-sztucznej-inteligencji-pierwsze-przepisy-regulujace-ai, [12.08.2023].
3. Aljohani T. M., Cyberattacks on Energy Infrastructures: Modern War Weapons, on-line: https://arxiv.org/ftp/arxiv/papers/2208/2208.14225.pdf, [07.07.2023].

4. Bachmann S. D. Hybrid Wars: The 21st-century new threats to global peace and security, Scientia Militaria, South African Journal of Military Studies, Vol 43, No. 1, 2015.
5. Bilal A., Hybrid Warfare – New Threats, Complexity, and 'Trust' as the Antidote, NATO Review, on-line: https://www.nato.int/docu/review/articles/2021/11/30/hybrid-warfare-new-threats-complexity-and-trust-as-the-antidote/index.html, [12.08.2023].
6. Coccomini D. A., Caldelli R., Falchi F., Cross-Forgery Analysis of Vision Transformers and CNNs for Deepfake Image Detection, Proceedings of the 1st International Workshop on Multimedia AI against Disinformation (MAD'22), June 27–30, 2022, Newark, NJ, USA. ACM, New York, NY, USA, on-lain: https://arxiv.org/pdf/2206.13829.pdf, [14.07.2023].
7. Cybersecurity high-risk series: Challenges in Protecting Cyber Critical Infrastructure, Gouverment Accessibility Office, on-lie: https://www.gao.gov/assets/gao-23-106441.pdf, [12.08.2023].
8. Dimitrov N., Methods for Enhancing and Evaluation of Energy Security and Economic Growth in Bulgaria, Economic Alternatives, 2019, Issue 4.
9. Etienne H., The future of online trust (and why Deepfake is advancing it), AI and Ethics (2021) 1:553–562on-line: https://link.springer.com/content/pdf/10.1007/s43681-021-00072-1.pdf, [12.08.2023].
10. Hwang Y., Ji Youn Ryu, Se-Hoon Jeong , Effects of Disinformation Using Deepfake: The Protective Effect of Media Literacy Education, Cyberpsychol Behav Social Network,
11. Jaiman A., Deepfakes & Synthetic Media. Humanity at the Edge of an Uncanny Valley.
12. Johnson B. D., Draudt A., i.inn. Information warfare and the future of conflict, Arizona State University, on-line: https://threatcasting.asu.edu/sites/default/files/2020-07/threatcasting-2020-The%20Future%20of%20Information%20Warfare-WEB_0.pdf, [12.08.2023].
13. Kalpokiene J., Kalpokas I., Synthetic Media and Information Warfare: Assessing Potential Threats, The Russian Federation in Global Knowledge Warfare
14. Kalpokas I. , Problematising reality: the promises and perils of synthetic media, SN Social Sciance (2021), on-line: https://link.springer.com/content/pdf/10.1007/s43545-020-00010-8.pdf, [14.07.2023].
15. Millière R., Deep learning and synthetic media, on-line: https://arxiv.org/pdf/2205.05764.pdf, [12.08.2023].
16. Nasu H., Deepfake technology in the age of information warfare, Lieber Institute, on-lain: https://lieber.westpoint.edu/deepfake-technology-age-information-warfare/, [14.07.2023].
17. Perez Ch., Information Warfare in Russia's War in Ukraine. The Role of Social Media and Artificial Intelligence in Shaping Global Narratives, on-line: https://foreignpolicy.com/2022/08/22/information-warfare-in-russias-war-in-ukraine/ [14.07.2023].
18. Rashid A., Yar Khan A., Wasif Azim S., Cyber hegemony and information warfare: A case of Russia, Liberal Arts & Social Sciences International Journal (LASSIJ), Vol. 5,

No. 1, (January-June 2021): 648-666, on-line: https://pdfs.semanticscholar.org/c53e/04fcae8573bec52c358aad566be8a6449a26.pdf?_gl=1*j6k4hl*_ga*NzM5MDM3MTMyLjE2ODQ3NjU3NDk.*_ga_H7P4ZT52H5*MTY5MTU2NjIwNy45LjEuMTY5MTU2NzQ4NS41NS4wLjA, [12.08.2023].

19. Rogers M. S., Weinstein D., Protecting our critical infrastructure in the digital age, on-line: https://thehill.com/opinion/cybersecurity/447596-protecting-our-critical-infrastructure-in-the-digital-age/, [10.07.2023].
20. Schick N., Deepfakes and the Infoapokalypse. What you urgently need to know. London 2020, s. 8.
21. Smith D.C., Enhancing cybersecurity in the energy sector: a critical priority, Journal of Energy & Natural Resources Law Volume 36, 2018 - Issue 4, on-line: https://www.tandfonline.com/doi/epdf/10.1080/02646811.2018.1516362?needAccess=true&role=button, [12.08.2023].
22. Smith D.C., Cybersecurity in the energy sector: are we really prepared?, Journal of Energy & Natural Resources Law Volume 39, 2021.
23. Snegovaya M., Putin's information warfare inUkraine Soviet origins of Rusiia's hybrid warfare, Institute for the Study of War, 2015.
24. Ukraine's Lessons for the Future of Hybrid Warfare, RAND Corporation, on-line: https://www.rand.org/blog/2022/11/ukraines-lessons-for-the-future-of-hybrid-warfare.html, [12.08.2023].
25. Vaccari C., Chadwick A., Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News, Social Media + Society January-March 2020: 1–13, on-line: https://journals.sagepub.com/doi/pdf/10.1177/2056305120903408, [12.08.2023].
26. Venkataramakrishnan S. Can you believe your eyes? How deepfakes are coming for politic, Financial Times, 24.X. 2019 on-lain: https://www.ft.com/content/4bf4277c-f527-11e9-a79c-bc9acae3b654, [12.08.2023].
27. Wang R., Huang Z., Chen Z., Anti-Forgery: Towards a Stealthy and Robust DeepFake Disruption Attack via Adversarial Perceptual-aware Perturbations, on-lain: https://arxiv.org/pdf/2206.00477.pdf, [07.07.2023].
28. Wasiuta O., Wasiuta S., Wojna informacyjna zagrożeniem dla bezpieczeństwa ludzkości, Repozytorium Uniwersytet Pedagogiczny https://rep.up.krakow.pl/xmlui/bitstream/handle/11716/2026/06--Wojna-informacyjna-zagrozeniemWasiuta--Wasiuta.pdf?sequence=1&isAllowed=y, [12.08.2023].
29. When seeing is no longer believing, CNN, on-lain: https://edition.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/, [12.08.2023].
30. Zając A., Balina R. Kowalski D. Financial and Economic Stability of Energy Sector Enterprises as a Condition for Poland's Energy Security—Legal and Economic Aspects, Energies 2023, 16(3), 1442.