KRZYSZTOF JAN JAKUBSKI
Doktor nauk ekonomicznych

Uniwersytet Civitas Warszawa

ORCID: 0000-0002-3491-9520
e-mail: kojot@jakubski.pl

DOI: https://zenodo.org/records/19223512

AGENCI SZTUCZNEJ INTELIGENCJI JAKO NARZĘDZIE WSPÓŁCZESNEJ PRZESTĘPCZOŚCI – ANALIZA KRYMINOLOGICZNA I IMPLIKACJE DLA BEZPIECZEŃSTWA PUBLICZNEGO

Artificial Intelligence Agents as a Tool of Modern Crime – Criminological Analysis and Implications for Public Safety

Abstrakt

Artykuł analizuje wykorzystanie agentów sztucznej inteligencji jako narzędzi współczesnej przestępczości z perspektywy kryminologicznej i nauk policyjnych. Autor wskazuje, że agentowe systemy AI nie tworzą nowej kategorii autonomicznych sprawców, lecz pełnią funkcję wysoce zaawansowanych, adaptacyjnych narzędzi, które radykalnie modyfikują relację między sprawcą a czynem zabronionym. Kluczowe znaczenie mają cechy takie jak autonomia, skalowalność i zdolność do personalizacji działań, które sprzyjają industrializacji przestępczości oraz rozmyciu odpowiedzialności sprawcy. Artykuł ukazuje agentów AI jako katalizator nowego paradygmatu – przestępczości zautomatyzowanej – oraz identyfikuje wynikające z tego wyzwania dla organów ścigania, systemu dowodowego i polityki prewencyjnej. Opracowanie ma charakter przeglądowy i wskazuje potrzebę dalszych badań empirycznych oraz interdyscyplinarnej refleksji nad przeciwdziałaniem przestępczości wspieranej przez AI. Końcowym wnioskiem jest to, iż w zakresie zwalczania cyberprzestępczości opartej o AI nie jesteśmy spóźnieni o lata — jesteśmy spóźnieni koncepcyjnie. Jeśli nie zmienimy paradygmatu obrony, AI nie tylko ominie te systemy — ona je zastąpi.

Słowa klucze: sztuczna inteligencja, agenci AI, przestępczość zautomatyzowana, kryminologia, nauki policyjne, narzędzie przestępstwa, odpowiedzialność karna, atrybucja sprawstwa, autonomia algorytmiczna, bezpieczeństwo publiczne

Abstract

This article examines the use of artificial intelligence agents as tools in contemporary crime from a criminological and police science perspective. The author argues that AI agent systems do not create a new category of autonomous perpetrators, but rather serve as highly advanced, adaptive tools that radically modify the relationship between the perpetrator and the criminal act. Key characteristics include autonomy, scalability, and the ability to personalize actions, which foster the industrialization of crime and the blurring of perpetrator responsibility. The article portrays AI agents as a catalyst for a new paradigm—automated crime—and identifies the resulting challenges for law enforcement, the evidence system, and prevention policy. This review highlights the need for further empirical research and interdisciplinary reflection on countering AI-assisted crime. The conclusion is that we are not behind years in combating AI-based cybercrime—we are conceptually behind. If we do not change the defense paradigm, AI will not only bypass these systems—it will replace them.

Keywords: artificial intelligence, AI agents, automated crime, criminology, policing science, instrument of crime, criminal liability, attribution of responsibility, algorithmic autonomy public security

1. Wstęp

Dynamiczny rozwój sztucznej inteligencji, w szczególności sztucznej inteligencji agentowej zdolnej do autonomicznego lub półautonomicznego działania[1], staje się istotnym czynnikiem wpływającym na współczesne zagrożenia dla bezpieczeństwa publicznego. Technologie te, pierwotnie projektowane w celu zwiększania efektywności procesów społecznych i gospodarczych, coraz częściej wykorzystywane są jako narzędzia działalności przestępczej. Z perspektywy nauk policyjnych kluczowe znaczenie ma nie sama innowacyjność technologiczna, lecz jej potencjał kryminogenny oraz wpływ na modus operandi sprawców, strukturę przestępczości oraz zdolności reagowania organów ścigania.

Wykorzystanie agentów sztucznej inteligencji prowadzi do jakościowej zmiany charakteru przestępczości. Umożliwia bowiem automatyzację czynów zabronionych, zwiększenie ich skali, a także obniżenie bariery wejścia dla sprawców pozbawionych specjalistycznych kompetencji. Zjawisko to wpisuje się w szerszy proces „industrializacji przestępczości", w którym działania nielegalne przybierają formę zorganizowanych, powtarzalnych i trudnych do wykrycia procesów. W konsekwencji rodzi to nowe wyzwania dla Policji i innych służb odpowiedzialnych za bezpieczeństwo wewnętrzne, zarówno na poziomie prewencji, jak i wykrywania oraz ścigania przestępstw.

Pomimo rosnącej liczby publikacji dotyczących zagrożeń związanych z wykorzystaniem sztucznej inteligencji, zauważalna jest dominacja analiz o charakterze technicznym lub prawnym. Relatywnie niewiele opracowań podejmuje problematykę agentów AI z perspektywy kryminologicznej i policyjnej, koncentrując się na ich roli jako narzędzi

[1] Agent AI to pojedynczy program lub system działający autonomicznie w określonym środowisku, natomiast sztuczna inteligencja agentowa (ang. *agentic AI*) to cała koncepcja i architektura, w której wielu takich agentów współpracuje, planuje i podejmuje działania w sposób bardziej złożony i proaktywny.

przestępstwa oraz na konsekwencjach dla praktyki organów ścigania. Uzasadnia to potrzebę syntetycznego ujęcia istniejącej wiedzy i identyfikacji luk badawczych. dokonanie systematycznego i krytycznego przeglądu literatury naukowej oraz raportów instytucjonalnych dotyczących zjawiska agentów AI w perspektywie kryminologicznej. Metodologia została zaprojektowana z myślą o zapewnieniu maksymalnej przejrzystości i replikowalności procesu badawczego.

Na potrzeby niniejszego opracowania pojęcie agentów sztucznej inteligencji odnosi się wyłącznie do współcześnie dostępnych systemów algorytmicznych, opartych na modelach uczenia maszynowego i architekturach agentowych, które działają w ramach ściśle określonych celów, danych wejściowych oraz ograniczeń projektowych. Analiza nie obejmuje hipotetycznych systemów ogólnej sztucznej inteligencji (AGI), rozumianych jako podmioty zdolne do autonomicznego, uniwersalnego rozumowania i samodzielnego formułowania celów w szerokim spektrum domen, gdyż tego rodzaju byty pozostają poza zakresem technologii funkcjonujących de lege lata. W konsekwencji agenci AI są w niniejszym artykule traktowani nie jako autonomiczni sprawcy przestępstw, lecz jako narzędzia o wysokim stopniu złożoności i adaptacyjności, wykorzystywane przez człowieka w procesie realizacji czynów zabronionych. Przyjęta perspektywa pozwala skoncentrować analizę na realnych, empirycznie obserwowalnych formach przestępczości wspieranej przez sztuczną inteligencję oraz na wynikających z nich wyzwaniach dla kryminologii i nauk policyjnych, z wyłączeniem spekulatywnych rozważań o charakterze futurystycznym.

Celem niniejszego artykułu przeglądowego jest analiza wykorzystania agentów sztucznej inteligencji jako narzędzi działalności przestępczej w ujęciu właściwym dla nauk policyjnych, ze szczególnym uwzględnieniem aspektów kryminologicznych, prewencyjnych i operacyjnych. Artykuł ma na celu nie tylko uporządkowanie dotychczasowych ustaleń, lecz także wskazanie obszarów wymagających pogłębionych badań empirycznych.

Problemy badawcze

Na potrzeby artykułu sformułowano następujące główne problemy badawcze:

P1. W jaki sposób wykorzystanie agentów sztucznej inteligencji wpływa na formy, skalę i dynamikę współczesnej przestępczości?

P2. Jakie cechy agentów AI decydują o ich szczególnym potencjale kryminogennym z perspektywy nauk policyjnych?

P3. W jaki sposób wykorzystanie agentów AI modyfikuje profil sprawcy oraz relację sprawca–ofiara?

P4. Jakie wyzwania operacyjne i organizacyjne stawia przestępczość wspierana przez AI przed Policją i innymi służbami bezpieczeństwa?

P5. W jakim zakresie obecne rozwiązania prewencyjne i prawne odpowiadają na zagrożenia wynikające z wykorzystania agentów AI?

Hipotezy badawcze

W oparciu o dotychczasową literaturę i raporty instytucjonalne przyjęto następujące hipotezy badawcze:

H1. Wykorzystanie agentów sztucznej inteligencji prowadzi do wzrostu skali i automatyzacji przestępczości, przy jednoczesnym obniżeniu progu wejścia dla sprawców.

H2. Agenci AI jako narzędzia przestępstwa, sprzyjają rozmyciu odpowiedzialności sprawcy i utrudniają identyfikację oraz atrybucję czynów zabronionych.
H3. Przestępczość wspierana przez AI zwiększa podatność określonych kategorii ofiar oraz ryzyko wtórnej wiktymizacji.
H4. Aktualne modele działania Policji i służb bezpieczeństwa nie są w pełni dostosowane do skali i specyfiki zagrożeń generowanych przez agentów AI.
H5. Skuteczne przeciwdziałanie przestępczości z wykorzystaniem AI wymaga pogłębionych badań empirycznych, wykraczających poza dotychczasowe analizy teoretyczne i przeglądowe.

Metodologia badań

Artykuł ma charakter przeglądowy i opiera się na metodach jakościowych typowych dla nauk policyjnych i nauk o bezpieczeństwie. Zastosowano przede wszystkim:

1. analizę literatury naukowej z zakresu kryminologii, nauk policyjnych, bezpieczeństwa wewnętrznego oraz studiów nad przestępczością cyfrową z lat 2010–2025, przy czym szczególny nacisk położono na okres 2023–2025, ze względu na gwałtowny rozwój generatywnej sztucznej inteligencji i jej wpływu na praktyki cyberprzestępcze;
2. analizę raportów instytucjonalnych organizacji międzynarodowych oraz instytucji odpowiedzialnych za bezpieczeństwo cyfrowe, w tym m.in. Europolu, Interpolu, ENISA, a także wybranych opracowań i raportów instytucji polskich, w szczególności NASK – Państwowego Instytutu Badawczego oraz CERT Polska;
3. analizę aktów prawnych i dokumentów strategicznych dotyczących bezpieczeństwa i przeciwdziałania przestępczości, w tym dokumentów krajowych oraz wybranych regulacji i strategii międzynarodowych;
4. analizę porównawczą sposobów ujmowania zagrożeń związanych z AI w różnych podejściach badawczych.

Dobór źródeł miał charakter celowy i ukierunkowany był na identyfikację powtarzających się wątków, luk badawczych oraz rozbieżności interpretacyjnych. W analizie uwzględniono również dostępne opracowania i raporty dotyczące zagrożeń związanych ze sztuczną inteligencją przygotowywane przez instytucje funkcjonujące w Polsce, co pozwoliło na odniesienie problematyki badawczej do kontekstu krajowego.

Ze względu na ograniczony dostęp do danych operacyjnych oraz dynamiczny charakter zjawiska, artykuł nie obejmuje badań empirycznych, co stanowi jego świadome ograniczenie badawcze. Jednym z kluczowych założeń metodologicznych jest traktowanie niniejszego opracowania jako etapu wstępnego, którego rezultatem ma być wskazanie potrzeby dalszych, szczegółowych badań empirycznych, w tym badań akt spraw, analiz statystycznych oraz badań jakościowych prowadzonych w środowisku służb odpowiedzialnych za bezpieczeństwo publiczne.

2. Pojęcie i charakterystyka agentów sztucznej inteligencji

2.1. Definicja i istota agenta AI w kontekście kryminologicznym

W literaturze przedmiotu agent sztucznej inteligencji jest definiowany jako system komputerowy zdolny do autonomicznego podejmowania decyzji i realizowania działań

w określonym środowisku w celu osiągnięcia wyznaczonych celów[2]. Kluczowe dla jego działania są funkcje percepcji otoczenia, przetwarzania informacji, uczenia się oraz adaptacyjnego reagowania, co odróżnia go od prostych automatów[3]. Z perspektywy nauk o bezpieczeństwie i kryminologii, fundamentalne znaczenie ma odróżnienie tych zaawansowanych systemów od tradycyjnych narzędzi cyfrowych. Podczas gdy proste programy wykonują sztywno zapisane sekwencje poleceń, agenci AI charakteryzują się **autonomią działania** (mogą podejmować decyzje bez bezpośredniej, bieżącej kontroli człowieka) oraz **zdolnością do adaptacji** (modyfikują swoje zachowania na podstawie nowych danych i interakcji z otoczeniem)[4]. Agenci AI są zazwyczaj stworzeni do wykonywania konkretnych zadań. Mają pomagać w różnych sytuacjach – na przykład w odpowiadaniu na pytania, organizowaniu kalendarza, a nawet zarządzaniu skrzynką odbiorczą. Agent AI to pojedynczy program lub system działający autonomicznie w określonym środowisku, natomiast sztuczna inteligencja agentowa (ang. agentic AI) to cała koncepcja i architektura, w której wiele takich agentów współpracuje, planuje i podejmuje działania w sposób bardziej złożony i proaktywny[5]. Agenci AI świetnie radzą sobie z automatyzacją prostych, powtarzalnych zadań, ale nie mają takiej autonomii ani zdolności podejmowania decyzji, jak sztuczna inteligencja oparta na agentach. Ta właśnie kombinacja autonomii i plastyczności stanowi jakościową zmianę w krajobrazie narzędzi przestępczych, umożliwiając nie tylko automatyzację, ale także tworzenie dynamicznych, trudnych do przewidzenia i spersonalizowanych zagrożeń[6].

2.2. Klasyfikacja i jej znaczenie dla identyfikacji zagrożeń

Dla celów analizy kryminologicznej i policyjnej, klasyfikacja agentów AI powinna uwzględniać cechy bezpośrednio wpływające na ich potencjał przestępczy oraz na trudności w

---

[2] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, London 2021, s. 52–60, polskie wydanie: *Sztuczna inteligencja. Nowe spojrzenie.* Wydanie IV. Tom 1 i 2, (tłum. A. Grażyński), Helion 2023 – patrz Rozdział 2 – Intelligent Agents, szczególnie strony otwierające rozdział (2.1 i 2.2)

[3] Por. K. Hayward, M. Maas, *Artificial intelligence and crime: A primer for criminologists*, 2020 https://matthijsmaas.com/uploads/Hayward%20and%20Maas%20-%202020%20-%20Artificial%20intelligence%20and%20crime%20A%20primer%20for%20cr.pdf, (dostęp: 1.3.2026 r.); M. Woźniak, *Sztuczna inteligencja jako przedmiot i narzędzie przestępstwa*, Warszawa 2022.; V. Ciancaglini, S. Gariuolo, S. Hilt, R. McArdle, R. Vosseler, *AI Assistants in the Future: Security Concerns and Risk Management*, Trend Micro 2024, https://www.trendmicro.com/vinfo/gb/security/news/security-technology/looking-into-thefuture-risks-and-security-considerations-to-ai-digital-assistants (dostęp: 1.3.2026 r.)

[4] Por. M. Brundage i in., *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. arXiv przesłano 20 lutego 2018 r. (wersja 1), ostatnia aktualizacja 1 grudnia 2024 r, https://arxiv.org/abs/1802.07228 (dostęp: 1.3.2026 r.); Europol, *Facing reality? Law enforcement and the challenge of deepfakes*. Publications Office of the European Union 2022. https://www.europol.europa.eu/publicationsevents/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes (dostęp: 1.3.2026 r.)

[5] E. Lisowski, *AI Agents vs Agentic AI: What's the Difference and Why Does It Matter?*, Medium 2024. https://medium.com/@elisowski/ai-agents-vs-agentic-ai-whats-the-difference-and-why-does-it-matter (dostęp: 1.3.2026 r.)

[6] Por. S. Caneppele, F. Calderoni, *Crime-as-a-Service and the Transformation of Criminal Tools.* [w:] *The Routledge Handbook of Technology, Crime and Justice.* red. F. Miró-Llinares, S. D. Johnson, Routledge 2022. s. 213–229.

ściganiu[7]. Odpowiednio skonstruowana klasyfikacja nie ma wyłącznie charakteru porządkującego, lecz stanowi narzędzie identyfikacji zagrożeń oraz wsparcia dla działań operacyjnych Policji.

2.2.1. Poziom autonomii a problem atrybucji sprawstwa

Pierwszym i kluczowym wymiarem klasyfikacyjnym jest poziom autonomii agenta AI, który wyznacza stopień oddzielenia sprawcy od czynu przestępnego. **Agenci półautonomiczni (1)** realizują ściśle określone, powtarzalne zadania, takie jak masowe rozsyłanie identycznych treści czy wykonywanie prostych operacji technicznych. W tym sensie pełnią oni funkcję cyfrowych narzędzi, których użycie pozostaje stosunkowo łatwe do powiązania z działaniami człowieka.

Znacznie poważniejsze wyzwania rodzą **agenci w pełni autonomiczni (2)**, zdolni do samodzielnego doboru strategii działania w zmiennym środowisku. Ich wykorzystanie prowadzi do zatarcia bezpośredniego związku przyczynowo-skutkowego pomiędzy decyzją sprawcy a wyrządzoną szkodą. Zjawisko to określane jest w literaturze jako problem atrybucji (attribution problem) i stanowi jedno z najpoważniejszych wyzwań dowodowych we współczesnej przestępczości technologicznej[8].

2.2.2. Charakter interakcji agenta AI ze środowiskiem

Drugim istotnym kryterium klasyfikacyjnym jest sposób, w jaki agent AI wchodzi w interakcję ze środowiskiem. **Agenci percepcyjni** (1) koncentrują się na zbieraniu, agregowaniu i analizie danych, często działając w sposób trudny do wykrycia. W praktyce przestępczej pełnią oni funkcję rozpoznawczą, umożliwiając identyfikację potencjalnych ofiar, luk w zabezpieczeniach czy podatności systemów społecznych i technicznych.

Odmienną rolę odgrywają **agenci wykonawczy** (2), którzy inicjują konkretne działania prowadzące bezpośrednio do naruszenia dobra prawnego. Mogą to być działania o charakterze informacyjnym, finansowym lub technicznym, takie jak rozpowszechnianie dezinformacji, dokonywanie nieuprawnionych transakcji czy ingerencja w systemy informatyczne. Z punktu widzenia organów ścigania rozróżnienie tych dwóch typów agentów ma znaczenie operacyjne, gdyż determinuje odmienne metody wykrywania i przeciwdziałania[9].

2.2.3. Agenci hybrydowi jako narzędzie automatyzacji przestępczości

W praktyce przestępczej największe zagrożenie stanowią **agenci hybrydowi** (3), łączący funkcje percepcyjne, analityczne i wykonawcze w jednym systemie[10]. Ich znaczenie

[7] UNICRI, *Artificial intelligence and robotics for law enforcement*. United Nations Interregional Crime and Justice Research Institute 2021. https://unicri.it/index.php/News/NewUNICRI-Report-Addresses-Aritificial-Intelligence-Robotics-Law-Enforcement (dostęp: 1.3.2026 r.)

[8] Por. B. Brożek, M. Jakubiec, *O odpowiedzialności prawnej maszyn autonomicznych*, Artificial Intelligence and Law 2017, nr 25, s, 293–304. (dostęp: 1.3.2026 r.); A. Jarosiewicz, J. Kulesza, *Problem atrybucji w cyberprzestrzeni: Wyzwania dla prawa międzynarodowego*, „Państwo i Prawo" 2023, nr 78(5), s. 45-62.

[9] Por. National Institute of Standards and Technology (NIST). *AI risk management framework: AI RMF (1.0)*. U.S. Department of Commerce 2023. https://www.nist.gov/itl/ai-riskmanagement-framework (dostęp: 1.3.2026 r.)

[10] Por. Europol. (2022). *op.cit.*

kryminologiczne polega na zdolności do automatyzacji całych procesów przestępczych, a nie jedynie pojedynczych czynności.

Szczególną rolę odgrywają tu **agenci komunikacyjni** (4), tacy jak zaawansowane chatboty i generatywne modele językowe, które umożliwiają prowadzenie wysoce skutecznych, spersonalizowanych ataków socjotechnicznych. Dzięki zdolności do naśladowania ludzkiej komunikacji utrudniają one ofiarom rozpoznanie zagrożenia, zwiększając ryzyko wiktymizacji[11].

Równolegle **agenci analityczni** (5) wspomagają sprawców w podejmowaniu decyzji strategicznych, analizując duże zbiory danych i optymalizując działania przestępcze. Mogą one dotyczyć zarówno ekonomicznych aspektów przestępczości, jak wycena danych na nielegalnych rynkach, jak i logistyki czy wykorzystywania różnic w systemach prawnych poszczególnych państw[12]. Połączenie tych funkcji umożliwia tworzenie zautomatyzowanych łańcuchów przestępczych obejmujących rekrutację ofiar, realizację czynu oraz ukrywanie jego efektów, w tym pranie pieniędzy[13].

2.3. Cechy kryminogenne agentów AI i ich implikacje

Charakterystyka techniczna agentów AI przekłada się bezpośrednio na specyficzne, wzmożone ryzyka kryminologiczne. **Autonomia i adaptacyjność** to cechy, które przekształcają agenta z narzędzia w quasi-wspólnika przestępcy. Sprawca może zlecić systemowi ogólny cel, podczas gdy agent samodzielnie dobierze metody, dostosuje komunikaty do reakcji ofiar i ominie napotkane zabezpieczenia, działając w sposób nieprzewidywalny nawet dla swojego twórcy. Ta elastyczność radykalnie utrudnia profilowanie ataków i tworzenie skutecznych, statycznych sygnatur ochronnych[14].

Kolejną cechą o fundamentalnym znaczeniu jest **skalowalność**. Jeden agent AI może jednocześnie prowadzić atak na miliony użytkowników, generować tysiące unikalnych, wiarygodnych prób phishingu lub zarządzać armią botów w kampanii dezinformacyjnej. Ta **"ekonomia skali w przestępczości"** obniża jednostkowy koszt ataku, zwiększa prawdopodobieństwo sukcesu i przeciąża możliwości reagowania służb[15]. Jednocześnie prowadzi do **głębokiego rozmycia odpowiedzialności** (attribution problem**)**. Działania agenta są odsunięte w czasie i przestrzeni od sprawcy, który może ukryć się za wieloma warstwami technologii i jurysdykcji. Sam agent, pozbawiony podmiotowości prawnej, nie ponosi

---

[11] K. Weise, C. Metz, *When AI chatbots hallucinate*. The New York Times 2023, March 29. https://www.nytimes.com/2023/03/29/technology/ai-chatbots-hallucinations.html (dostęp: 1.3.2026 r.)
[12] S. Caneppele, F. Calderoni (2022), *op.cit.*
[13] UNODC (2025), *Emerging threats: The intersection of criminal and technological innovation in the use of automation and artificial intelligence in the cybercrime landscape of Southeast Asia. United Nations Office on Drugs and Crime*, wrzesień 2025, https://www.unodc.org/roseap/uploads/documents/Publications/2025/UNODC_Report_Emerging_threats_-_The_intersection_of_criminal_and_technological_innovation_in_the_use_of_automation_and_AI.pdf (dostęp: 1.3.2026 r.)
[14] Por. M. Brundage i in., *op. cit*.
[15] A. Lavorgna, *Organized Crime and Cybercrime*. [w:] T. J. Holt. A. M. Bossler (red.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, Palgrave Macmillan, Cham 2020, s. 117-234. (dostęp: 1.3.2026 r.)

odpowiedzialności karnej, co stawia przed wymiarem sprawiedliwości fundamentalne pytania o dowodzenie sprawstwa i winy[16].

Wreszcie, zdolność do **zaawansowanego modelowania i predykcji zachowań** ofiar stanowi o jakościowej zmianie w wiktymologii cyberprzestępstw. Agenci, analizując ogromne zbiory danych behawioralnych z wycieków lub mediów społecznościowych, mogą tworzyć niezwykle precyzyjne profile psychograficzne. Pozwala to na konstruowanie przekazów socjotechnicznych maksymalnie dostosowanych do lęków, potrzeb i słabości konkretnej osoby, co znacząco podnosi skuteczność manipulacji i poszerza krąg potencjalnych ofiar poza tradycyjnie uważanych za "narażonych" [17].

2.4. Ograniczenia technologiczne jako punkty zaczepienia dla prewencji i ścigania

Pomimo swojej potęgi, agenci AI nie są wszechmocni, a ich ograniczenia wyznaczają kluczowe obszary dla działań prewencyjnych i dochodzeniowych. Podstawowym ograniczeniem jest **brak intencji i świadomości moralnej**. Agent realizuje zaprogramowany cel, często w sposób bezwzględnie optymalizacyjny, bez zrozumienia kontekstu społecznego czy konsekwencji prawnych. Oznacza to, że ostateczna odpowiedzialność karna i etyczna zawsze spoczywa na ludziach: twórcach, użytkownikach lub tych, którzy świadomie wykorzystali system do złych celów. To rozgraniczenie jest podstawą dla konstrukcji odpowiedzialności prawnej[18].

Ponadto, **błędy adaptacyjne[19] i "halucynacje[20]"** mogą prowadzić do niekonsekwencji, nielogicznych działań lub generowania treści zdradzających sztuczne pochodzenie. Dla organów ścigania takie artefakty mogą być cennym źródłem dowodów, "cyfrowym odciskiem palca" konkretnego modelu lub wskazówką co do metodologii ataku[21]. Wreszcie, **całkowita zależność od jakości danych i struktury środowiska** stanowi ich Achillesową piętę. Skuteczność agenta jest limitowana danymi, na których się uczył, oraz parametrami działania narzuconymi przez programistę. Działania prewencyjne mogą więc koncentrować się na **zatruwaniu danych** (data poisoning)[22], na których uczą się złośliwe systemy, oraz na projektowaniu środowisk cyfrowych (np. platform, protokołów komunikacyjnych) w sposób, który utrudnia lub wykrywa autonomiczne, masowe działania,

[16] Por. A. Jarosiewicz, J. Kulesza, (2023), *op. cit*.; UNODC. (2025*). op.cit.*
[17] Por. Europol., *op.cit.*; P. Zając, *Wiktymologia w epoce cyfrowej: Nowe zagrożenia związane z wykorzystaniem AI*, „Archivum Criminologiae", 2023, nr 15(1), s. 87-105.
[18] S. Kemp, L. Halonen, I. Pölönen, *The European Union's Artificial Intelligence Act: A critical review from a criminal law perspective*, „New Journal of European Criminal Law" 2022, nr 13(4), s. 456-478.
[19] Błędy adaptacyjne generatywnej sztucznej inteligencji to zjawisko polegające na nieprawidłowym dostosowaniu się modelu do nowych danych lub kontekstu, skutkujące generowaniem treści nieadekwatnych, błędnych lub pozornie poprawnych, lecz niezgodnych z rzeczywistością.
[20] Halucynacje generatywnej sztucznej inteligencji to błędne lub zmyślone informacje tworzone przez modele AI, które są przedstawiane w sposób spójny i pozornie wiarygodny, mimo że nie mają oparcia w faktach ani danych źródłowych.
[21] NIST, *op.cit.*; M. Woźniak, *op.cit.*
[22] Zatruwanie danych (*data poisoning*) to rodzaj cyberataku polegający na celowym wprowadzaniu zmanipulowanych informacji do zbiorów treningowych modeli sztucznej inteligencji, co prowadzi do zniekształcenia ich działania i generowania błędnych wyników.

tworząc swego rodzaju **"architekturę bezpieczeństwa"** (security by design) nieprzyjazną dla zautomatyzowanej przestępczości[23].

3. Kryminologiczne ujęcie technologii jako narzędzia przestępstwa

Analiza technologii w kontekście przestępczości wymaga wyjścia poza wąskie ramy inżynierii czy informatyki i umieszczenia jej w szerszym, kryminologicznym paradygmacie. Technologia – od najprostszych narzędzi po najbardziej zaawansowane systemy autonomiczne – nie istnieje w próżni społecznej; jest ona tworzona, adaptowana i wykorzystywana przez ludzi w konkretnych kontekstach kulturowych, ekonomicznych i prawnych[24]. Kluczowe dla zrozumienia fenomenu agentów sztucznej inteligencji w przestępczości jest zatem potraktowanie ich nie jako zjawiska sui generis, lecz jako kolejnego etapu w długiej historii relacji między innowacjami technologicznymi a działalnością przestępczą. W GenAI (generatywnej sztucznej inteligencji) pojedynczy model reaguje na polecenia użytkownika, odpowiadając na pytania lub tworząc treści. W agentowej sztucznej inteligencji autonomiczni „agenci" w systemie wykonują zadania bez poleceń użytkownika, aby osiągnąć cel. Rozdział ten ma na celu zbudowanie teoretycznych podstaw, które pozwolą ukazać agentów AI jako narzędzia przestępstwa o szczególnym, powstającym potencjale kryminogennym, którego zrozumienie wymaga odwołania się do klasycznych i współczesnych koncepcji kryminologicznych[25]. Interesującym opracowaniem teoretycznym jest „A Criminology of Machines"[26], w którym autor przedstawia nową perspektywę kryminologiczną, traktującą autonomiczne systemy — w tym agentów sztucznej inteligencji — jako potencjalnych „aktorów" przestępczości. Podkreśla przy tym, że dotychczasowe teorie kryminologiczne nie uwzględniają takich podmiotów ani narzędzi, mimo iż ich interakcje mogą prowadzić do rezultatów sprzecznych z prawem.

3.1. Narzędzie przestępstwa w teorii kryminologicznej - od przedmiotu fizycznego do algorytmicznej funkcji

W tradycyjnej kryminologii narzędzie przestępstwa było rozumiane przede wszystkim jako materialny przedmiot służący do bezpośredniego popełnienia czynu zabronionego (np. wytrych, broń, środki łączności)[27]. Jednak w erze cyfrowej to ujęcie okazuje się

[23] Por. M. Brundage i in., *op.cit.*; J. Surma, *Hakowanie Sztucznej Inteligencji*, Wydawnictwo Naukowe PWN 2022.
[24] Por. T. J. Pinch, W. E. Bijker, *The social construction of facts and artifacts*. [w:] W. E. Bijker, T. P. Hughes, T. J. Pinch (red.), *The Social Construction of Technological Systems*, MIT Press 2012, s. 11-44; E. Nerantzi, G. Sartor, *Hard AI Crime: The Deterrence Turn*, „Oxford Journal of Legal Studies" 2024, nr 44(3), s. 673–701. (dostęp: 1.3.2026 r.); K.J. Jakubski i in., *Ciemne strony cyberprzestrzeni – wybrane aspekty* [w:] E. Chodźko, K. Talarek (red.), *Wyzwania i problemy społeczeństwa w XXI wieku*. *Tom I*, Lublin 2020, s. 250-292
[25] R. V. Clarke, *Technology, Criminology and Crime Science*, „European Journal on Criminal Policy and Research" 2004, nr 10, s. 55–63 (dostęp: 1.3.2026 r.); M. McGuire, *Technology, Crime and Justice: The Question Concerning Technomia*, Routledge 2012. https://books.google.pl/books/about/Technology_Crime_and_Justice.html (dostęp: 1.3.2026 r.)
[26] G. M. Campedelli, *A Criminology of Machines*, https://arxiv.org/abs/2511.02895 (dostęp: 1.3.2026 r.)
[27] M. Felson, R. V. Clarke, *Opportunity Makes the Thief: Practical Theory for Crime Prevention*. Police Research Series1998, Paper No. 98. Home Office, London https://resaud.net/wp-content/uploads/2024/03/Felson-et-Clarke-1998-Opportunity-Makes-the-Thief.-Practical-Theory-for-.pdf (dostęp: 1.3.2026 r.)

niewystarczające. Współczesne teorie, takie jak teoria możliwości (Rational Choice Theory) i teoria codziennych aktywności (Routine Activity Theory), poszerzają to pojęcie, kładąc nacisk na funkcjonalną rolę narzędzia w ułatwianiu popełnienia przestępstwa[28]. Z tej perspektywy, narzędziem jest wszystko, co zmniejsza wysiłek, zwiększa korzyści lub obniża ryzyko wykrycia z punktu widzenia racjonalnie kalkulującego sprawcy[29].

W przypadku cyberprzestępczości narzędziem staje się nie tyle fizyczny komputer, ile oprogramowanie, algorytm lub dostęp do określonej usługi sieciowej[30]. Agent AI reprezentuje najwyższy stopień tej ewolucji: jest to narzędzie algorytmiczne o charakterze pośredniczącym i wzmacniającym. Jego kryminogenna funkcja polega nie na fizycznym uszkodzeniu, ale na: **(1) automatycznej eksploracji i wykorzystaniu okazji przestępczych**, **(2) minimalizacji kognitywnego i czasowego wysiłku sprawcy** poprzez przejęcie złożonych operacji, oraz **(3) systematycznym zacieraniu śladów i utrudnianiu atrybu**cji[31]. W ten sposób klasyczne pojęcie narzędzia zostaje zradykalizowane – agent AI nie jest już biernym przedłużeniem ręki sprawcy, ale aktywnym współtwórcą scenariusza przestępczego.

3.2. Technologia jako katalizator ewolucji przestępczości w ujęciu historycznym

Historyczna analiza pokazuje, że każda znacząca rewolucja technologiczna – od wynalezienia telegrafu i kolei żelaznej, przez telefony komórkowe, po Internet – była równolegle rewolucją w metodach i skali przestępczości[32]. Telegraf umożliwił szybszą koordynację transgranicznych afer finansowych, samochód zrewolucjonizował ucieczki z miejsca przestępstwa, a telefon komórkowy stał się narzędziem nowych form szantażu i stalkingu. Systemy przekazu informacji zawsze były w zainteresowaniu złoczyńców. System telegrafu optycznego Claude'a Chappe'a, używany we Francji od 1792 r., został wykorzystany przez braci Blanc w latach 1834–1836 do manipulacji giełdowych, co doprowadziło do procesu w Tours w 1837 r. W 1903 roku, gdy Guglielmo Marconi prezentował światu swój wynalazek – aparat przekazujący wiadomości na odległość, Nevil Maskelyne, brytyjski magik i entuzjasta technologii bezprzewodowej, spowodował, za sprawą zakłócających impulsów, że zamiast alfabetu Morse'a, widzowie usłyszeli ciąg obelg pod adresem odkrywcy, co uznawane jest za pierwszy przypadek „hakowania" w historii technologii radiowej. Wykorzystanie komputerów, a następnie sieci komputerowej do popełniania czynów przestępczych jest zjawiskiem niemal równoległym do historii ich rozwoju i wykorzystywania[33] .

---

[28] R. V Clarke (2004), *op.cit.*
[29] D.B.Cornish, R.V. Clarke, *The Reasoning Criminal: Rational Choice Perspectives on Offending*, Transaction Publishers, 2014, s. 105-169.
[30] M. McGuire, *Technology, Crime and Justice: The Question Concerning Technomia*. Routledge 2012, s. 57–84.
[31] Por. D. M. Vicente, R. S. Pereira, A. A. Leal, (red.), *Legal aspects of autonomous systems: A comparative approach.* Springer Nature 2024. https://www.google.pl/books/edition/Legal_Aspects_of_Autonomous_Systems/ (dostęp: 1.3.2026 r.)
[32] M. McGuire, *op.cit.*, s. 12-20, 33-52, 57-70; D. S. Wall, *Cybercrime: The transformation of crime in the information age.* Polity Press 2007. https://books.google.pl/books?id=SiG-zE6yteMC&printsec=frontcover&hl=pl#v=onepage&q&f=false (dostęp: 1.3.2026 r.)
[33] K.J. Jakubski, *Przestępczość komputerowa – zarys problematyki*, „Prokuratura i Prawo" 1996, 12, s. 34 – 50; S. Schjolberg, *The History of Cybercrime Third edition*, Books on Demand 2020; K.S. Choi, C.S. Lee, E.R. Louderback, *Historical Evolutions of Cybercrime: From Computer Crime to Cybercrime.* [w:] *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, red, T. Holt, A. Bossler, Palgrave Macmillan, Cham

Elina Nerantzi podkreśla historyczną zmianę w charakterze czynu przestępczego, gdzie technologia sama w sobie staje się aktywnym elementem zachowania przestępczego, co należy ująć jako przełomowy moment w ewolucji przestępczości technologicznej[34].

Kluczową cechą tej historycznej dynamiki jest asymetria adaptacyjna między przestępcami a organami ścigania. Przestępcy, działając często w mniejszych, bardziej elastycznych strukturach, wykazują się większą umiejętnością adaptowania (innovative appropriation) technologii cywilnych w nielegalnych celach[35]. Biurokratyczne, zhierarchizowane instytucje państwowe, związane procedurami i ograniczeniami budżetowymi, adaptują się wolniej, co tworzy okresy „okna możliwości" dla nowych form przestępczości. Rozwój agentów AI doskonale wpisuje się w ten schemat. Podczas gdy sektor publiczny dopiero zaczyna eksperymentować z AI w celach dochodzeniowych, zorganizowane grupy przestępcze już wdrażają generatywne modele językowe do masowych oszustw, wykorzystując tę właśnie asymetrię[36].

3.3. Agenci AI jako narzędzia „wysokiego potencjału kryminogennego" - cechy definiujące

Co zatem odróżnia agentów AI od wcześniejszych narzędzi i uzasadnia przypisanie im szczególnie wysokiego potencjału kryminogennego? Można wskazać na trzy kluczowe, emergentne cechy, które tworzą jakościową nowość w historii technologii przestępczych.

3.3.1. Autonomiczna agencja i adaptacyjność:

W przeciwieństwie do statycznego wytrychu czy nawet złośliwego oprogramowania o stałym działaniu, agenci AI posiadają zdolność do samodzielnego podejmowania decyzji w niepełni przewidywalnym środowisku w celu osiągnięcia nadrzędnego celu[37]. Oznacza to, że narzędzie nie tylko wykonuje program, ale także interpretuje sytuację i modyfikuje swoje zachowanie, pokonując dynamiczne przeszkody (np. zmieniające się zabezpieczenia, reakcje ofiar). To sprawia, że jest ono niezwykle odporne na proste, reaktywne środki zaradcze.

3.3.2. Skalowalność i personalizacja w jednym

Agent AI rozwiązuje historyczny dylemat przestępczości - wybór między atakiem masowym, ale mało skutecznym (np. spam), a atakiem wysoce spersonalizowanym, ale pracochłonnym (np. oszustwo na „prezesa"). Dzięki możliwości analizy dużych zbiorów danych (Big Data) i generowania unikalnej treści, agent może prowadzić miliony jednocześnie

2020, s. 27-43. (dostęp: 1.3.2026 r.); D.B. Parker, *Criminal Justice Resource Manual on Computer Crime* , Departament Sprawiedliwości USA, 1980, https://www.ojp.gov/pdffiles1/Digitization/118214NCJRS.pdf (dostęp: 1.3.2026 r.)

[34] E. Nerantzi, G. Sartor, *op.cit.*, s. 673–701,

[35] Europol, *op.cit.*

[36] UNODC, *op.cit.*

[37] S. Russell, P. Norvig, P., *op.cit.*

personalizowanych interakcji, łącząc skalę z precyzją. To radykalnie zwiększa zarówno efektywność, jak i potencjalny zasięg szkody społecznej[38].

3.3.3. Głęboka mediacja i rozmycie sprawstwa

Agent AI działa jako skomplikowany bufor algorytmiczny między intencją sprawcy a realizacją czynu. Proces ten jest na tyle złożony, że prowadzi do ontologicznego rozmycia odpowiedzialności. Czy sprawcą jest programista modelu, użytkownik, który sformułował zły prompt[39], czy może samo oprogramowanie? To rozmycie, zwane problemem atrybucji, jest nie tylko praktycznym utrudnieniem dla śledczych, ale także kwestionuje fundamentalne założenia prawa karnego opartego na winie jednostki[40].

3.4. Ku przestępczości zautomatyzowanej - nowy paradygmat w relacji sprawca–narzędzie

Kumulatywny efekt opisanych cech prowadzi do narodzin nowego paradygmatu, który można określić mianem **przestępczości zautomatyzowanej (automated crime).** W tym modelu rola ludzkiego sprawcy ewoluuje z bezpośredniego wykonawcy w kierunku nadzorcy, stratega lub inwestora w algorytmiczny system przestępczy[41]. Przestępstwo przestaje być pojedynczym aktem, a staje się ciągłym, zautomatyzowanym procesem eksploatacji podatności, zarządzanym przez AI.

Takim przykładem nowej jakości technologicznej, w której AI nie tylko wspiera, lecz aktywnie prowadzi elementy czynu zabronionego jest zbudowany na modelach LLM, autonomiczny agent wieloetapowy ScamAgent, zdolny do generowania wysoce realistycznych skryptów rozmów oszustów, które symulują rzeczywiste scenariusze oszustwa. System utrzymuje pamięć dialogową, dynamicznie adaptuje się do odpowiedzi użytkownika i stosuje strategie perswazyjne w kolejnych turach konwersacji, co czyni tradycyjne zabezpieczenia nieskutecznymi[42].

Z kryminologicznego punktu widzenia oznacza to przesunięcie w modelach teoretycznych. Teoria codziennych aktywności musi uwzględnić, że „odpowiedni opiekun nieobecny" może być nie tylko fizycznie nieobecnym człowiekiem, ale także niedostatecznym lub przestarzałym systemem algorytmicznej obrony. Teoria możliwości musi brać pod uwagę, że kalkulacja sprawcy dotyczy nie tylko korzyści z pojedynczego czynu, ale opłacalności inwestycji w autonomiczny system generujący strumień przychodów z wielu czynów[43]. W ten

[38] S. Caneppele, F. Calderoni, (2022), *op.cit*.
[39] Prompt to polecenie, zapytanie lub zestaw instrukcji wpisywany przez użytkownika w celu nakłonienia narzędzia sztucznej inteligencji (np. ChatGPT, Gemini, Midjourney) do wykonania określonego zadania. Działa jak most między intencją człowieka a algorytmem, definiując kontekst, styl i format oczekiwanej odpowiedzi.
[40] A. Jarosiewicz, J. Kulesza (2023), *op.cit.*
[41] NIST (2023), *op.cit.*
[42] S. Badhe, *ScamAgents: How AI Agents Can Simulate Human-Level Scam Calls*. arXiv 2025. (dostęp: 1.3.2026 r.)
[43] E.R. Leukfeldt, *Cybercrime and the scalability of offending* [w:] *The human factor of cybercrime*, red. T.J. Holt,

sposób agenci AI nie tylko transformują praktyki przestępcze, ale także zmuszają do rewizji samych teoretycznych ram, za pomocą których te praktyki opisujemy i analizujemy.

3.5. Agenci AI jako punkt zwrotny w kryminologicznej teorii narzędzia przestępstwa

Agenci sztucznej inteligencji stanowią punkt zwrotny w długiej historii technologii jako narzędzia przestępstwa. Przekraczają one dotychczasowe rozumienie narzędzia jako przedmiotu lub prostego programu, stając się aktywnymi, adaptacyjnymi pośrednikami o wysokiej autonomii. Ich unikalny potencjał kryminogenny wynika z połączenia autonomii, skalowalności i zdolności do głębokiego rozmycia sprawstwa. Wprowadzają one świat przestępczy w erę przestępczości zautomatyzowanej, w której algorytm przejmuje operacyjną inicjatywę, a rola człowieka ulega fundamentalnej redefinicji. Zrozumienie tej ewolucji w kontekście historycznym i teoretycznym jest niezbędnym fundamentem dla dalszej analizy konkretnych form przestępczości z użyciem AI oraz dla formułowania adekwatnych odpowiedzi ze strony systemu bezpieczeństwa.

3.6. Autonomizacja narzędzi przestępstwa jako nowy etap ewolucji przestępczości technologicznej

O ile poprzedni podpunkt koncentrował się na konsekwencjach teoretycznych, o tyle niniejszy punkt ujmuje autonomizację narzędzi przestępstwa jako proces ewolucyjny o wymiarze operacyjnym i systemowym. Współczesne raporty analityczne dotyczące zagrożeń cybernetycznych wskazują, że kluczowym wyróżnikiem obecnego etapu rozwoju przestępczości technologicznej nie jest już sama cyfryzacja działań przestępczych, lecz ich postępująca autonomizacja. Z perspektywy kryminologicznej oznacza to jakościową zmianę w charakterze narzędzia przestępstwa: od technologii wymagających stałego nadzoru sprawcy do systemów zdolnych do samodzielnego podejmowania decyzji operacyjnych w trakcie realizacji czynu zabronionego. Analizy wskazują, że agentowe systemy AI umożliwiają prowadzenie złożonych operacji przestępczych w sposób ciągły i skalowalny [44].

Z punktu widzenia nauk policyjnych szczególnie istotne jest to, że autonomiczne narzędzia oparte na AI integrują w jednym systemie etapy dotychczas rozdzielone organizacyjnie i czasowo, takie jak rozpoznanie, planowanie, realizacja oraz ewaluacja skuteczności działań przestępczych. Agentowe systemy AI są w stanie analizować reakcje ofiar i środowiska, modyfikować swoje strategie oraz kontynuować działanie bez bezpośrednich instrukcji człowieka. W efekcie dochodzi do osłabienia klasycznego związku przyczynowo-

A.M. Bossler, Routledge 2020, s. 255–272, https://www.google.pl/books/edition/The_Human_Factor_of_Cybercrime/1YK1DwAAQBAJ (dostęp: 1.3.2026 r.)

[44] Trend Micro, *AI Pulse: Sticker Shock, Rise of the Agents, Rogue AI*, Trend Micro Research2024, https://www.trendmicro.com/pl_pl/research/24/h/agentic-ai-takeover.html (dostęp: 1.3.2026 r.); Trend Micro. *The Road to Agentic AI: Navigating Architecture,* Threats, and Solutions, Trend Micro Research 2025, https://www.trendmicro.com/vinfo/us/security/news/security-technology/the-road-to-agenticai-navigating-architecture-threats-and-solutions (dostęp: 1.3.2026 r.)

skutkowego pomiędzy zachowaniem sprawcy a skutkiem przestępnym, co rodzi istotne trudności dowodowe i atrybucyjne w postępowaniach karnych.

Raporty Trend Micro podkreślają również, że rozwój autonomicznych narzędzi przestępczych prowadzi do zjawiska określanego jako industrializacja przestępczości, w której działania przestępcze przybierają formę powtarzalnych, zoptymalizowanych procesów realizowanych na dużą skalę[45]. Jak wskazuje R. Flores: „Rok 2026 zostanie zapamiętany jako rok, w którym cyberprzestępczość przestała być branżą usługową i stała się w pełni zautomatyzowana. Wkraczamy w erę, w której agenci AI będą odkrywać, wykorzystywać i monetyzować słabości bez ingerencji człowieka. Wyzwaniem dla obrońców nie jest już samo wykrywanie ataków, ale nadążanie za napędzanym przez maszyny tempem zagrożeń"[46]. Taki model działania sprzyja powstawaniu przestępczości rozproszonej, transgranicznej i trudnej do jednoznacznej kwalifikacji w ramach tradycyjnych kategorii kryminologicznych. Z perspektywy policyjnej oznacza to konieczność odejścia od reaktywnego modelu zwalczania przestępczości na rzecz podejścia systemowego, uwzględniającego analizę algorytmicznych wzorców działania i procesów decyzyjnych autonomicznych narzędzi AI.

Istotnym elementem tej transformacji jest również dwoisty charakter agentowych systemów AI, które mogą być wykorzystywane zarówno do ochrony bezpieczeństwa, jak i do działań przestępczych. Jak pośrednio wynika z analiz Trend Micro, te same cechy technologiczne – autonomia, zdolność uczenia się oraz adaptacja w czasie rzeczywistym – stanowią podstawę zarówno nowoczesnych systemów obronnych, jak i potencjalnie wysoce skutecznych narzędzi przestępczych. W ujęciu kryminologicznym potwierdza to tezę, że technologia nie jest czynnikiem determinującym przestępczość sama w sobie, lecz wzmacnia jej potencjał w zależności od kontekstu użycia oraz poziomu kontroli instytucjonalnej.

Podsumowując, autonomizacja narzędzi przestępstwa opartych na agentach AI stanowi nie tylko kolejny etap rozwoju przestępczości technologicznej, lecz także wyzwanie konceptualne dla nauk policyjnych. Zjawisko to wymaga dalszych, pogłębionych badań empirycznych i teoretycznych nad mechanizmami sprawstwa, możliwościami prewencji oraz skutecznością reakcji organów ścigania wobec przestępczości realizowanej przy użyciu autonomicznych systemów sztucznej inteligencji.

4. Agenci sztucznej inteligencji w strukturze współczesnej przestępczości

Postępująca cyfryzacja życia społecznego oraz dynamiczny rozwój zaawansowanych agentów sztucznej inteligencji prowadzą do głębokich, jakościowych zmian w morfologii współczesnej przestępczości[47]. Agenci AI, wbrew niektórym popularnym wyobrażeniom, nie stanowią nowej kategorii autonomicznych sprawców, lecz funkcjonują jako wysoce zaawansowane narzędzia pośredniczące (intermediaries), które radykalnie modyfikują relację

[45] Trend Micro. *The AI-fication of cyberthreats: Security predictions for 2026*. Trend Micro Research 2025, *https://documents.trendmicro.com/assets/research-reports/the-ai-fication-of-cyberthreats-trend-micro-security-predictions-for-2026.pdf* (dostęp: 1.3.2026 r.)
[46] T.J.R. Flores, wypowiedź w: *Trend Micro Predicts 2026 as the Year Cybercrime Becomes Fully Industrialized.* Trend Micro Security News 2025, https://newsroom.trendmicro.com/2025-11-25-Trend-Micro-Predicts-2026-as-the-Year-Cybercrime-Becomes-Fully-Industrialized (dostęp: 1.3.2026 r.)
[47] S. Caneppele, F. Calderoni, *op.cit*.

między sprawcą a czynem zabronionym. Wpływają one na sam sposób popełnienia przestępstwa, jego potencjalną skalę społeczną, a przede wszystkim – na fundamentalne utrudnienie procesów wykrywczych i dowodowych[48]. Dla nauk policyjnych i bezpieczeństwa kluczowe staje się zrozumienie, w jaki sposób ta technologia transformuje tradycyjne typologie przestępczości, oraz jakie systemowe konsekwencje – wykraczające poza pojedyncze incydenty – niesie to dla bezpieczeństwa publicznego i paradygmatów działania organów ścigania[49]. Z perspektywy kryminologicznej analiza wykorzystania agentów AI wymaga również odniesienia do konstrukcji odpowiedzialności karnej przewidzianych w obowiązującym prawie.

4.1. Transformacja tradycyjnych form przestępczości

Agenci AI zasadniczo nie kreują zupełnie nowych kategorii czynów zabronionych de lege lata, lecz prowadzą do ich głębokiej transformacji, czyniąc znane od dawna formy przestępczości (oszustwa, nękanie, naruszenia prywatności) bardziej efektywnymi, masowymi i trudniejszymi do zwalczenia[50]. W obszarze przestępczości gospodarczej automatyzacja i adaptacyjność agentów AI demokratyzują i industrializują działania socjotechniczne. Podczas gdy tradycyjny phishing opierał się na stosunkowo łatwych do wychwycenia, masowych, lecz identycznych komunikatach, współczesne generatywne agenci AI umożliwiają tworzenie hiperpersonalizowanych ataków spersonalizowanego phishingu na niespotykaną skalę[51]. System analizuje fragmenty publicznie dostępnych danych o ofierze (z mediów społecznościowych, wycieków, forów) i generuje unikalną wiadomość, która może perfekcyjnie naśladować styl komunikacji kolegi z pracy, dostawcy usług lub członka rodziny. Ta zdolność prowadzi do paradoksalnego zjawiska: mimo że atak jest zautomatyzowany i prowadzony na tysiące osób, każda ofiara odbiera go jako wysoce personalny i zindywidualizowany[52]. Z perspektywy kryminologicznej oznacza to zasadnicze przesunięcie – z przestępstwa opartego na indywidualnej relacji sprawca-ofiara w kierunku masowej, a jednocześnie zindywidualizowanej wiktymizacji, gdzie sprawca korzysta z algorytmicznego "powielacza", który zachowuje pozory intymnej znajomości[53].

4.2. Agenci AI a przestępstwa przeciwko prywatności i godności osobistej

Szczególnie destrukcyjny wpływ agentów AI uwidacznia się w sferze przestępstw godzących w prywatność, wizerunek i godność osobistą. Generatywne systemy, a zwłaszcza technologie tworzenia głębokich fałszerstw (technologia deepfake - syntetyczne fałszerstwa audiowizualne), przekraczają barierę między sferą publiczną a intymną ofiary, umożliwiając tworzenie syntetycznych, ale fotorealistycznych materiałów kompromitujących[54]. Powoduje to

[48] Europol, *op.cit.*
[49] UNODC, *op.cit*.
[50] E. R. Leukfeldt, *op.cit.*
[51] NIST, *op.cit.*
[52] P. Zając, *op.cit.*
[53] S. Caneppele, F. Calderoni, *op.cit.*
[54] Europol, *op.cit*., K.J. Jakubski, *Niebezpieczna Sztuczna Inteligencja*, „Cybersecurity & Cybercrime" 2024, t. 1, Numer specjalny: PT XXI 2023, s. 247-301.

powstanie nowych form wiktymizacji wtórnej i trzeciorzędowej, gdzie szkoda nie wynika z jednorazowego ujawnienia prawdziwych faktów, lecz z nieusuwalnego, wiralowego rozprzestrzeniania się fałszywej, ale wiarygodnej narracji, która niszczy reputację, życie rodzinne i karierę zawodową ofiary[55]. Dla organów ścigania kluczowym wyzwaniem jest tu nie tylko anonimizacja sprawcy (działającego poprzez serwery proxy i zautomatyzowane boty), ale także dowodowa "przepaść technologiczna". W postępowaniach z oskarżenia publicznego obowiązek wykazania fałszywego charakteru materiału spoczywa na organach ścigania, jednak w praktyce dowodowej ustalenie to wymaga kosztownych i specjalistycznych ekspertyz cyfrowych, podczas gdy sama treść może już wywołać nieodwracalną szkodę społeczną[56].

4.3. Automatyzacja przemocy psychicznej i nękania

Wykorzystanie agentów AI w przestępstwach typu stalking lub uporczywego nękania prowadzi do ich jakościowej przemiany. Tradycyjny stalking był zjawiskiem ściśle związanym z fizyczną lub cyfrową obecnością i zaangażowaniem sprawcy, co pozwalało na profilowanie jego zachowań i często ułatwiało identyfikację. Dziś możliwe jest pełne zautomatyzowanie kampanii nękania – od generowania i wysyłki tysięcy zróżnicowanych, kontekstowych wiadomości, przez symulowanie aktywności ofiary w sieci (np. zakładanie fałszywych profili), po koordynowanie ataków przez " farmy trolli"[57] zarządzane przez AI[58]. To tworzy zjawisko "duchowej" przemocy psychicznej – ofiara jest poddawana nieustannemu, adaptacyjnemu naciskowi, który wydaje się emanować z samej sieci, bez widocznego, ludzkiego źródła. Z punktu widzenia kryminologii oznacza to eskalację potencjalnej krzywdy przy jednoczesnej minimalizacji ryzyka wykrycia przez sprawcę. Ofiara doświadcza chronicznego stresu i poczucia bezsilności, podczas gdy sprawca może nie angażować się emocjonalnie i pozostawać w pełni operacyjnym "cieniu"[59].

4.4. Agenci AI jako narzędzie przestępczości zorganizowanej i dezinformacji

Agenci AI ewoluują z roli prostego narzędzia do roli kluczowego elementu infrastruktury operacyjnej zorganizowanych grup przestępczych. Pełnią funkcje analitycznego i strategicznego wsparcia: optymalizują logistykę nielegalnych dostaw, modelują ryzyko

[55] C. McGlynn i in., *'It's Torture for the Soul' The Harms of Image-Based Sexual Abuse*, „Social & Legal Studies" 2020, nr 30(4), s. 541-562 (dostęp: 1.3.2026 r.); K.J. Jakubski, *op.cit.*
[56] A. Jarosiewicz, J. Kulesza, *op.cit.*
[57] Farma trolli (również fabryka trolli) – duża grupa (od kilkudziesięciu do kilkuset osób) opłaconych pracowników (trolli) publikujących komentarze na zlecenie. Z reguły każdy pracownik korzysta z kilku, a nawet kilkunastu kont jednocześnie. Działają na zlecenie. Są tworzone i zasilane finansowo przez zlecający im podmiot, którym często jest państwo lub organizacja polityczna. Przykładem jest rosyjska Agencja Badań Internetowych (znana jako Trolle z Olgino), której właścicielem był Jewgienij Prigożyn. Przykładem polskiej farmy trolli, która prowadziła fałszywe konta w sieci, obsługiwała państwowe spółki i szerzyła dezinformację na rzecz różnych podmiotów była firma Cat@Net.
[58] A. Grycuk*., Fake newsy, trolle, boty i cyborgi w mediach społecznościowych,* „Analizy BAS" 2021 nr 1(152) , s. 1-12, https://www.researchgate.net/publication/349109775_Fake_newsy_trolle_boty_i_cyborgi_w_mediach_spolecznosciowych, (dostęp: 1.3.2026 r.)
[59] S. Kemp i in., *op.cit.*

wykrycia, automatycznie "przesiewają" ogromne zbiory skradzionych danych w poszukiwaniu najbardziej wartościowych informacji, a nawet prowadzą symulacje skuteczności różnych scenariuszy przestępczych[60]. W ten sposób AI staje się "mózgiem operacyjnym", który zwiększa efektywność i bezpieczeństwo grup działających w modelu crime-as-a-service[61].

Najbardziej systemowym zagrożeniem jest jednak wykorzystanie agentów AI do zautomatyzowanej, wielokanałowej dezinformacji. Generatywne modele językowe mogą masowo produkować spójne narracje, komentarze, fałszywe artykuły newsowe i posty społecznościowe, które są następnie dystrybuowane przez sieci botów symulujących autentyczną aktywność ludzką[62]. Działania te wykraczają poza klasyczne przestępstwa przeciwko wizerunkowi, stając się zagrożeniem dla bezpieczeństwa narodowego i porządku konstytucyjnego. Mogą one destabilizować procesy wyborcze, podważać zaufanie do instytucji państwa, wzmacniać polaryzację społeczną i wywoływać niepokoje społeczne. Z perspektywy policyjnej jest to wyzwanie hybrydowe, wymagające współpracy nie tylko z prokuraturą, ale także z służbami specjalnymi, regulatorami rynku cyfrowego i środowiskiem akademickim[63].

4.5. Konsekwencje dla struktury przestępczości i działań Policji

Upowszechnienie agentów AI w świecie przestępczym prowadzi do fundamentalnych napięć w istniejącym systemie prawno-dowodowym. Rozmycie bezpośredniego sprawstwa stawia pod znakiem zapytania tradycyjne konstrukcje odpowiedzialności karnej oparte na winie i zamiarze (strona podmiotowa czynu - mens rea) konkretnej osoby fizycznej. Jednocześnie lawinowy wzrost liczby czynów o charakterze masowym, lecz o indywidualnie odczuwalnych skutkach, przeciąża możliwości reakcyjne systemu.

Dla Policji i innych służb oznacza to konieczność głębokiej redefinicji metodologii działania. Nacisk musi zostać przeniesiony z wyłącznie reaktywnego ścigania po zgłoszeniu na prewencję i wczesne wykrywanie oparte na analityce danych i współpracy proaktywnej. Kluczowe staje się rozwijanie defensywnej sztucznej inteligencji (AI-for-security) – systemów zdolnych do wykrywania anomalii, identyfikowania zautomatyzowanych kampanii dezinformacyjnych i śledzenia transakcji prania pieniędzy w czasie rzeczywistym[64]. Wymaga

[60] UNODC, *op.cit.*
[61] Model *Crime-as-a-Service* (CaaS) wykształcił się w wyniku dwóch równoległych procesów: (1) gwałtownej digitalizacji i komercjalizacji cyberprzestrzeni na początku XXI wieku oraz (2) rosnącej profesjonalizacji środowisk przestępczych, w szczególności grup specjalizujących się w cyberprzestępczości zorganizowanej. CaaS oznacza udostępnianie innym przestępcom narzędzi, know-how, infrastruktury, luk programistycznych czy zautomatyzowanych pakietów ataków, których można użyć bez zaawansowanych kompetencji technicznych. Szerzej – K.J. Jakubski *Od CaaS do AI-CaaS. Kryminologiczne konsekwencje demokratyzacji narzędzi przestępczych*, artykuł złożony do publikacji w czasopiśmie Przegląd Policyjny.
[62] S.C. Woolley, P.N. Howard (red), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, Oxford Studies in Digital Politics, New York, 2018 (online Oxford Academic), (dostęp: 1.3.2026 r.)
[63] Financial Action Task Force (FATF), *Second 12-Month Review of the Revised FATF Standards on Virtual Assets and Virtual Asset Service Providers*. FATF/OECD 2021., https://www.fatf-gafi.org/content/dam/fatf-gafi/guidance/Second-12-Month-Review-Revised-FATF-Standards-Virtual-Assets-VASPS.pdf.coredownload.pdf (dostęp: 1.3.2026 r.)
[64] NIST, *op.cit.*

to inwestycji w specjalistyczne kadry (cyberkryminologów, data scientistów[65]), ścisłej współpracy z sektorem prywatnym (dostawcami AI, operatorami platform) oraz przyjęcia perspektywy zarządzania ryzykiem systemowym zamiast jedynie reagowania na incydenty [66].

Analiza wykorzystania agentów AI w działalności przestępczej wskazuje, że obowiązujące konstrukcje odpowiedzialności karnej pozostają co do zasady adekwatne, jednak wymagają reinterpretacji w warunkach algorytmicznego pośrednictwa działania. W praktyce możliwe są trzy komplementarne modele atrybucji odpowiedzialności: (1) **model instrumentalny**, w którym agent AI traktowany jest jako zaawansowane narzędzie pozostające pod faktyczną kontrolą sprawcy; **(2) model pośredniego sprawstwa**, akcentujący wykorzystanie autonomii systemu do realizacji zamiaru przestępnego przy ograniczonej bezpośredniej ingerencji człowieka; oraz (3) **model odpowiedzialności funkcjonalnej**, odnoszący się do podmiotów projektujących, wdrażających lub udostępniających systemy AI w warunkach rażącego braku zabezpieczeń. Z perspektywy de lege ferenda nie wydaje się konieczne konstruowanie nowej kategorii „odpowiedzialności sztucznej inteligencji", lecz raczej doprecyzowanie kryteriów przypisania zamiaru, kontroli i przewidywalności skutku w sytuacjach, gdy decyzje operacyjne są delegowane na system algorytmiczny. Kierunek ten pozwala zachować antropocentryczny fundament prawa karnego, jednocześnie dostosowując je do realiów przestępczości zautomatyzowanej.

4.6. Kwalifikacja prawna wykorzystania agentów AI w świetle Kodeksu karnego

Obowiązujący Kodeks karny nie zawiera przepisów odnoszących się wprost do sztucznej inteligencji. Nie oznacza to jednak powstania luki prawnej w zakresie penalizacji czynów popełnianych z jej wykorzystaniem. Agent AI, jako system techniczny pozbawiony podmiotowości prawnej, nie może być sprawcą przestępstwa w rozumieniu prawa karnego materialnego. Odpowiedzialność karna spoczywa wyłącznie na osobie fizycznej, która wykorzystuje system algorytmiczny do realizacji znamion czynu zabronionego[67].

Z perspektywy dogmatyki prawa karnego kluczowe jest rozróżnienie między narzędziem służącym do popełnienia przestępstwa a narzędziem przeznaczonym do tego celu. W pierwszym przypadku mamy do czynienia z technologią o legalnym zastosowaniu, która zostaje wtórnie wykorzystana w działalności przestępczej – tak jak nóż kuchenny może posłużyć do zabójstwa. W doktrynie prawa karnego rozróżnienie to ma znaczenie przy ocenie odpowiedzialności twórcy lub dystrybutora narzędzia, zwłaszcza gdy jego konstrukcja wskazuje na przystosowanie do popełniania czynów zabronionych. W drugim przypadku

[65] W języku polskim nie ma jednego, powszechnie przyjętego odpowiednika dla data scientist. Data scientist to specjalista działający na styku informatyki, statystyki i wiedzy dziedzinowej, którego głównym zadaniem jest pozyskiwanie, przetwarzanie oraz analizowanie dużych zbiorów danych w celu odkrywania wzorców, budowania modeli predykcyjnych i formułowania praktycznych wniosków wspierających procesy decyzyjne. W pracy wykorzystuje narzędzia programistyczne (np. Python, R), metody uczenia maszynowego oraz techniki wizualizacji danych, łącząc kompetencje analityczne z umiejętnością komunikacji wyników w zrozumiały sposób dla odbiorców nieposiadających wiedzy technicznej.

[66] S. Kemp i in., *op.cit.*

[67] J. Królikiewicz, *Sztuczna inteligencja w ujęciu odpowiedzialności karnej – luka w odpowiedzialności czy możliwość jej przypisania?*, „Problemy Prawa Karnego" 2021, nr 5(1), s. 45-68.

chodzi o narzędzia tworzone lub modyfikowane z zamiarem przestępczym, co może aktualizować odpowiedzialność nie tylko użytkownika, ale także twórcy takiego systemu[68]. Większość agentów AI wykorzystywanych w przestępczości należy do pierwszej kategorii – są to komercyjne modele językowe lub systemy automatyzacji, które sprawcy adaptują do celów niezgodnych z prawem. Sytuacja komplikuje się jednak w przypadku tzw. modeli z usuniętymi zabezpieczeniami (jailbroken models) lub systemów celowo pozbawionych zabezpieczeń etycznych, gdzie intencja przestępcza może być wpisana już w etap przygotowania narzędzia.

4.6.1. Przestępstwa przeciwko mieniu – oszustwo (art. 286 k.k.)

Wykorzystanie agentów AI w oszustwach aktualizuje znamiona art. 286 § 1 k.k., który penalizuje doprowadzenie innej osoby do niekorzystnego rozporządzenia mieniem za pomocą wprowadzenia w błąd, wyzyskania błędu lub niezdolności do należytego pojmowania przedsiębranego działania. Agent AI pełni tu funkcję narzędzia realizującego znamię wprowadzenia w błąd – generuje fałszywe komunikaty, podszywa się pod osoby lub instytucje, kreuje wiarygodne narracje manipulacyjne. Przykładem może być wykorzystanie generatywnych modeli językowych do prowadzenia tzw. spear-phishingu (spersonalizowanego phishingu)[69], co opisywałem w rozdziale 4.1. Agent AI analizuje publicznie dostępne informacje o ofierze (np. z mediów społecznościowych), a następnie generuje spersonalizowaną wiadomość podszywającą się pod współpracownika lub instytucję finansową, której celem jest nakłonienie ofiary do ujawnienia danych dostępowych lub dokonania przelewu.

Z punktu widzenia strony podmiotowej kluczowe jest wykazanie, że sprawca działał z zamiarem bezpośrednim kierunkowym – w celu osiągnięcia korzyści majątkowej. Automatyzacja procesu oszustwa nie zmienia tej konstrukcji. Sprawca, uruchamiając agenta AI zaprogramowanego do wyłudzania danych lub środków finansowych, realizuje swój zamiar za pośrednictwem narzędzia technicznego, podobnie jak sprawca wykorzystujący tradycyjne metody socjotechniczne.

Problematyczna pozostaje kwestia świadomości bezprawności w sytuacjach, gdy agent AI samodzielnie modyfikuje strategie manipulacji w sposób nieprzewidziany przez sprawcę. Polskie prawo karne przyjmuje jednak, że kto uruchamia zautomatyzowany proces przestępczy, ponosi odpowiedzialność za jego skutki w granicach objętych zamiarem, choćby ewentualnym[70].

4.6.2. Przestępstwa przeciwko wolności – stalking (art. 190a k.k.)

Artykuł 190a § 1 k.k. penalizuje uporczywe nękanie innej osoby lub osoby jej najbliższej, wzbudzające uzasadnione okolicznościami poczucie zagrożenia, poniżenia lub udręczenia, lub istotnie naruszające jej prywatność. Znamię uporczywości wymaga

[68] R. Rejmaniak, *Autonomiczność systemów sztucznej inteligencji jako wyzwanie dla prawa karnego*, „Roczniki Nauk Prawnych" 2021, nr 31(3), s. 89-112.
[69] NIST (2023). op.cit.; Vectra AI. *Spear phishing: How targeted attacks work and how to stop them*. 2026, https://www.vectra.ai/topics/spear-phishing (dostęp: 1.3.2026 r.)
[70] J. Królikiewicz, *op. cit*.

powtarzalności zachowań sprawcy oraz szczególnego nastawienia psychicznego wyrażającego się w nieustępliwości działania mimo świadomości jego uciążliwości dla pokrzywdzonego.

Wykorzystanie agentów AI do automatyzacji nękania nie wyłącza znamienia uporczywości – przeciwnie, może je wzmacniać. Sprawca, który konfiguruje system do ciągłego, adaptacyjnego wysyłania wiadomości do ofiary, wykazuje się szczególną determinacją i premedytacją. Fakt, że poszczególne komunikaty są generowane automatycznie, nie zmienia istoty zachowania sprawcy, który świadomie uruchomił i utrzymuje działanie takiego systemu.

Artykuł 190a § 2 k.k. penalizuje podszywanie się pod inną osobę i wykorzystywanie jej wizerunku lub danych osobowych w celu wyrządzenia jej szkody. Przepis ten znajduje bezpośrednie zastosowanie do przypadków wykorzystania technologii tzw. deepfake (syntetyczne fałszerstwa audiowizualne), gdzie agent AI generuje syntetyczne materiały audiowizualne z wizerunkiem ofiary. W takich sprawach szczególnego znaczenia nabiera kwestia dowodowa – konieczność wykazania sztucznego pochodzenia materiału, co wymaga specjalistycznych ekspertyz z zakresu informatyki śledczej.

4.6.3. Przestępstwa przeciwko ochronie informacji (art. 267-269b k.k.)

Rozdział XXXIII Kodeksu karnego zawiera przepisy penalizujące różne formy nielegalnego dostępu do informacji i systemów informatycznych. Z perspektywy wykorzystania agentów AI szczególne znaczenie mają:

1. **Artykuł 267 § 1 k.k**. – bezprawne uzyskanie dostępu do informacji nieprzeznaczonej dla sprawcy. Agent AI może służyć jako narzędzie przełamywania zabezpieczeń (np. poprzez automatyczne testowanie haseł, wykorzystywanie podatności systemów) lub jako środek socjotechniczny do wyłudzenia danych dostępowych od użytkowników.
2. **Artykuł 268 k.k**. – niszczenie, uszkadzanie, usuwanie lub zmienianie zapisu istotnej informacji. Autonomiczne działania agenta AI, prowadzące do modyfikacji lub zniszczenia danych, mogą realizować znamiona tego przestępstwa, pod warunkiem wykazania związku przyczynowego z działaniem sprawcy-człowieka.
3. **Artykuł 269b § 1 k.k**. – wytwarzanie, pozyskiwanie, zbywanie lub udostępnianie narzędzi przystosowanych do popełnienia przestępstw określonych w art. 165 § 1 pkt 4, art. 267-269a k.k. Przepis ten może znaleźć zastosowanie do twórców i dystrybutorów agentów AI celowo pozbawionych zabezpieczeń i przeznaczonych do działań przestępczych.

4.6.4. Pranie pieniędzy (art. 299 k.k.) – konstrukcja przestępstwa bazowego

W kontekście wykorzystania agentów AI do prania pieniędzy konieczne jest precyzyjne odniesienie do konstrukcji tego przestępstwa w polskim prawie karnym. Artykuł 299 § 1 k.k. penalizuje przyjmowanie, przekazywanie lub wywożenie za granicę środków płatniczych lub innych wartości dewizowych, praw majątkowych lub mienia pochodzącego z korzyści związanych z popełnieniem czynu zabronionego, a także podejmowanie innych czynności mogących udaremnić lub znacznie utrudnić stwierdzenie ich przestępnego pochodzenia, wykrycie, zajęcie lub orzeczenie przepadku.

Istotą prania pieniędzy jest zatem wtórny charakter tego przestępstwa – wymaga ono uprzedniego popełnienia tzw. przestępstwa bazowego (źródłowego), z którym powiązane są tzw. „brudne środki". Agent AI może wspomagać proces prania pieniędzy poprzez automatyzację transferów między rachunkami, optymalizację struktury transakcji w celu uniknięcia progów raportowania czy zarządzanie siecią tzw. „słupów" (money mules). Nie może jednak sam „generować" środków do wyprania – te muszą stanowić korzyści majątkowe związane z popełnieniem czynu zabronionego (tzw. przestępstwa bazowego), takiego jak oszustwo, kradzież czy handel narkotykami.

W praktyce organów ścigania oznacza to, że wykazanie przestępstwa prania pieniędzy z wykorzystaniem AI wymaga równoległego udowodnienia przestępstwa źródłowego, co przy transgranicznym charakterze cyberprzestępczości stanowi istotne wyzwanie dowodowe i jurysdykcyjne.

4.6.5. Odpowiedzialność twórcy i użytkownika agenta AI

Polskie prawo karne przewiduje kilka konstrukcji umożliwiających pociągnięcie do odpowiedzialności osób zaangażowanych w przestępczość z wykorzystaniem agentów AI na różnych etapach:

1. **Sprawstwo bezpośrednie (art. 18 § 1 k.k.)** – użytkownik agenta AI, który wykorzystuje go do realizacji znamion czynu zabronionego, odpowiada jako sprawca, nawet jeśli techniczne wykonanie czynu zostało zautomatyzowane.
2. **Sprawstwo kierownicze (art. 18 § 1 k.k.)** – organizator przestępczego wykorzystania agentów AI, który kieruje wykonaniem czynu przez inne osoby lub systemy, odpowiada jak sprawca.
3. **Sprawstwo polecające (art. 18 §1 k.k.)** – w pewnych sytuacjach możliwe jest również zastosowanie konstrukcji sprawstwa polecającego, gdy sprawca wykorzystuje inne osoby jako wykonawców czynu, posługując się agentem AI jako narzędziem organizacji przestępstwa.
4. **Pomocnictwo (art. 18 § 3 k.k.)** – twórca agenta AI, który dostarcza narzędzie świadomie ułatwiające popełnienie przestępstwa, może odpowiadać za pomocnictwo. Wymaga to jednak wykazania, że działał z zamiarem (choćby ewentualnym) ułatwienia konkretnego przestępstwa lub przestępstw określonego rodzaju.
5. **Przygotowanie (art. 16 § 1 k.k.)** – w przypadku przestępstw, dla których przygotowanie jest karalne (np. przestępstwa terrorystyczne), tworzenie lub adaptowanie agentów AI do ich popełnienia może stanowić karalną fazę stadialną.

Problematyczna pozostaje sytuacja, gdy twórca udostępnia agenta AI o neutralnym zastosowaniu, który następnie zostaje wykorzystany przestępczo przez osoby trzecie. Zgodnie z zasadą indywidualizacji odpowiedzialności karnej, sam fakt stworzenia narzędzia podatnego na nadużycia nie wystarcza do przypisania odpowiedzialności – konieczne jest wykazanie zamiaru lub co najmniej przewidywania i godzenia się na przestępcze wykorzystanie[71].

[71] M. Jankowska, *Podmiotowość prawna sztucznej inteligencji?* [w:] A. Bielska-Brodziak (red.), *O czym mówią prawnicy, mówiąc o podmiotowości,* Wydawnictwo Uniwersytetu Śląskiego 2021, s. 171-196.

4.6.6. Luki i wyzwania regulacyjne

Analiza obowiązujących przepisów wskazuje, że polskie prawo karne, oparte na tradycyjnej dogmatyce sprawstwa i winy, zachowuje zasadniczą zdolność do penalizacji przestępstw popełnianych z wykorzystaniem agentów AI. Nie oznacza to jednak braku wyzwań regulacyjnych.

Po pierwsze, rozmycie związku przyczynowego między działaniem sprawcy a skutkiem przestępnym utrudnia dowodzenie w sprawach, gdzie agent AI działał w sposób częściowo autonomiczny i nieprzewidywalny. Tradycyjna konstrukcja zamiaru, wymagająca świadomości i woli realizacji znamion, może napotykać trudności interpretacyjne w sytuacjach, gdy system algorytmiczny podejmuje autonomiczne decyzje operacyjne.

Po drugie, transgraniczność przestępczości z wykorzystaniem AI sprawia, że sprawca, ofiara, serwery i skutki przestępstwa mogą znajdować się w różnych jurysdykcjach, co komplikuje ustalenie właściwości polskich organów ścigania (art. 5-6 k.k.) oraz współpracę międzynarodową.

Po trzecie, tempo rozwoju technologicznego przewyższa możliwości adaptacyjne systemu prawnego. O ile istniejące przepisy pozwalają na penalizację większości zachowań przestępczych z użyciem AI, o tyle mogą nie nadążać za nowymi formami zagrożeń, takimi jak autonomiczne kampanie dezinformacyjne czy zautomatyzowane manipulacje rynkowe.

Postulowane w literaturze kierunki zmian obejmują doprecyzowanie kryteriów przypisania zamiaru, kontroli i przewidywalności skutku w sytuacjach, gdy decyzje operacyjne są delegowane na system algorytmiczny[72]. Wymaga to jednak pogłębionej debaty doktrynalnej i nie powinno prowadzić do pochopnych nowelizacji, które mogłyby naruszyć fundamentalne zasady prawa karnego, w tym zasadę określoności czynu zabronionego (nullum crimen sine lege certa).

4.7. AI jako katalizator nowej ery kryminologicznej

Agenci sztucznej inteligencji utrwalają się jako nieodłączny, transformujący element ekosystemu przestępczego. Ich rola wykracza daleko poza bycie "sprytniejszym narzędziem"; polegają one na radykalnym wzmocnieniu możliwości sprawcy przy jednoczesnej minimalizacji jego ekspozycji na ryzyko, co stanowi wyzwanie dla podstawowych założeń polityki karnej i prewencji. Zrozumienie tej nowej dynamiki – w której algorytm staje się aktywnym pośrednikiem w relacji przestępczej – jest warunkiem sine qua non dla opracowania skutecznych strategii bezpieczeństwa publicznego w XXI wieku. Pilnie potrzebne są więc interdyscyplinarne badania empiryczne, które połączą wiedzę kryminologiczną, informatyczną, prawną i psychologiczną, aby można było nie tylko reagować na zagrożenia, ale także je antycypować i projektować odporny na nie krajobraz społeczno-techniczny.

Wnioski końcowe

[72] J. Królikiewicz, *op.cit*; R. Rejmaniak, *op. cit.*

Przeprowadzona analiza pozwala stwierdzić, że agenci sztucznej inteligencji stanowią jakościowo nowy etap w ewolucji technologii wykorzystywanych w działalności przestępczej. Nie tworzą oni odrębnej kategorii autonomicznych sprawców w sensie prawnokarnym, lecz funkcjonują jako wysoce zaawansowane, adaptacyjne narzędzia pośredniczące, które w istotny sposób modyfikują relację pomiędzy sprawcą a czynem zabronionym. Z perspektywy kryminologicznej oznacza to przesunięcie punktu ciężkości z analizy pojedynczego aktu przestępczego na analizę zautomatyzowanych, ciągłych procesów przestępczych zarządzanych algorytmicznie.

Kluczowe cechy agentów AI – autonomia decyzyjna, zdolność adaptacji do zmiennego środowiska, skalowalność oraz możliwość jednoczesnej personalizacji działań – prowadzą do dalszej industrializacji przestępczości. Przestępstwo przestaje być zdarzeniem incydentalnym, a staje się powtarzalnym, zoptymalizowanym procesem generującym strumień korzyści przy minimalnym zaangażowaniu i ekspozycji sprawcy. Zjawisko to obniża próg wejścia do działalności przestępczej, zwiększa jej masowość oraz istotnie utrudnia wykrywanie, dowodzenie i atrybucję sprawstwa.

Analiza potwierdza, że wykorzystanie agentów AI pogłębia problem rozmycia odpowiedzialności karnej. Wieloetapowa mediacja algorytmiczna między intencją człowieka a skutkiem przestępnym osłabia klasyczny związek przyczynowo-skutkowy, na którym opiera się dogmatyka prawa karnego. Z perspektywy nauk policyjnych rodzi to poważne konsekwencje dowodowe, w szczególności w zakresie ustalania zamiaru, kontroli nad narzędziem oraz granic odpowiedzialności użytkownika, twórcy lub podmiotu wdrażającego system AI. Problem ten ma charakter systemowy i nie może być skutecznie rozwiązany wyłącznie na poziomie technicznym lub operacyjnym.

W wymiarze wiktymologicznym agenci AI prowadzą do jakościowej zmiany charakteru krzywdy przestępnej. Automatyzacja i personalizacja oddziaływań przestępczych zwiększają podatność szerokich kategorii ofiar, jednocześnie intensyfikując skutki psychiczne i społeczne przestępstw. Zjawiska takie jak masowa, zindywidualizowana manipulacja, zautomatyzowane nękanie czy trwała wiktymizacja reputacyjna podważają tradycyjne założenia dotyczące relacji sprawca–ofiara oraz wymagają rewizji dotychczasowych modeli prewencji.

Z punktu widzenia praktyki policyjnej i systemu bezpieczeństwa publicznego wykorzystanie agentów AI ujawnia ograniczenia reaktywnego modelu zwalczania przestępczości. Skala, tempo i adaptacyjność przestępczości zautomatyzowanej powodują, że działania podejmowane wyłącznie po zaistnieniu szkody okazują się niewystarczające. Konieczne staje się przejście w kierunku podejścia systemowego, opartego na prewencji, wczesnym wykrywaniu wzorców algorytmicznych oraz wykorzystaniu defensywnych zastosowań sztucznej inteligencji w działalności analitycznej i operacyjnej Policji.

Wyniki analizy wskazują również na potrzebę rewizji i uzupełnienia istniejących ram teoretycznych kryminologii. Klasyczne koncepcje narzędzia przestępstwa, racjonalnego wyboru czy teorii codziennych aktywności wymagają dostosowania do realiów, w których autonomiczne systemy algorytmiczne przejmują znaczną część inicjatywy operacyjnej. Agenci

AI nie tylko wzmacniają potencjał sprawcy, lecz stają się aktywnym elementem środowiska przestępczego, co uzasadnia traktowanie ich jako jednego z kluczowych czynników strukturalnych współczesnej przestępczości.

Podsumowując, agentów sztucznej inteligencji należy uznać za narzędzia o wysokim potencjale kryminogennym, których rozpowszechnienie stanowi jedno z najpoważniejszych wyzwań dla bezpieczeństwa publicznego w najbliższych latach. Skuteczne przeciwdziałanie przestępczości wspieranej przez AI wymaga nie tylko rozwoju rozwiązań technologicznych, lecz przede wszystkim pogłębionych badań empirycznych, integrujących perspektywę kryminologiczną, prawną i policyjną. Bez takiej interdyscyplinarnej refleksji system bezpieczeństwa pozostanie strukturalnie spóźniony wobec dynamiki zagrożeń generowanych przez autonomiczne narzędzia sztucznej inteligencji. Problem nie polega już na tym, że AI jest bronią — lecz że staje się autonomicznym ekosystemem zagrożeń. Wchodzimy w erę agentowej AI, gdzie sieci wyspecjalizowanych agentów planują, testują, adaptują i eskalują ataki bez udziału człowieka. Phishing „na żądanie", deepfake'y w czasie rzeczywistym, malware uczący się ofiary — to dopiero początek. XAI[73], fact-checking[74] czy detektory fake news[75] działają reaktywnie, podczas gdy atakujące LLM-y działają predykcyjnie. Do tego dochodzi data poisoning[76], prompt-injection[77] na poziomie łańcuchów agentów i przejmowanie systemów decyzyjnych. OSINT[78] sam w sobie staje się polem walki. Nie jesteśmy spóźnieni o lata — jesteśmy spóźnieni koncepcyjnie. Jeśli nie zmienimy paradygmatu obrony, AI nie tylko ominie systemy — ona je zastąpi.

[73] XAI (Explainable Artificial Intelligence) to podejście do projektowania systemów sztucznej inteligencji, które umożliwia zrozumienie, wyjaśnienie i kontrolę sposobu podejmowania decyzji przez algorytmy. Celem jest zapewnienie przejrzystości, tak aby człowiek mógł ustalić, dlaczego system podjął określone działanie, co ma kluczowe znaczenie dla odpowiedzialności prawnej, nadzoru i audytu.

[74] Fact-checking to systematyczna weryfikacja prawdziwości informacji (np. wypowiedzi publicznych, treści medialnych, materiałów w sieci) prowadzona przez wyspecjalizowane podmioty lub narzędzia analityczne. Polega na porównaniu twierdzeń z wiarygodnymi źródłami, ocenieniu ich rzetelności i publikowaniu wyników w formie korekt lub analiz.

[75] Detektory fake news to narzędzia (często oparte na uczeniu maszynowym), które automatycznie oceniają wiarygodność treści publikowanych w internecie. Analizują język, źródła, wzorce dystrybucji oraz sygnały sieciowe, aby wykrywać dezinformację, manipulację lub treści generowane sztucznie. Ich celem jest wspieranie platform, mediów i organów publicznych w ograniczaniu szkodliwych narracji.

[76] Atak polegający na celowym wprowadzaniu złośliwych lub zniekształconych danych do zbiorów treningowych lub operacyjnych systemu AI, w celu wpływania na jego działanie, obniżenia skuteczności lub wywołania błędnych decyzji.

[77] Technika ataku, w której napastnik manipuluje treścią wprowadzoną do modelu językowego, aby skłonić go do wykonania działań sprzecznych z intencją twórców lub użytkownika — np. poprzez ukryte instrukcje, które przejęły kontrolę nad generowanym wynikiem.

[78] OSINT to metoda pozyskiwania informacji polegająca na systematycznym gromadzeniu, analizie i interpretacji danych dostępnych publicznie — takich jak treści z mediów społecznościowych, rejestry państwowe, publikacje prasowe, dane techniczne czy materiały audiowizualne. Celem OSINT jest wyciąganie wiarygodnych wniosków z otwartych źródeł, bez stosowania środków operacyjnych lub niejawnych.

Bibliografia

Literatura:

Badhe S., *ScamAgents: How AI Agents Can Simulate Human-Level Scam Calls*. arXiv 9 grudnia 2025 r. (dostęp: 1.3.2026 r.).

Brożek B., Jakubiec M., *O odpowiedzialności prawnej maszyn autonomicznych*, „Artificial Intelligence and Law" 2017, nr 25, s. 293–304 (dostęp: 1.3.2026 r.)

Brundage M., Avin S., Clark J., Toner H., Eckersley P., Garfinkel B., Dafoe A., Scharre P., Zeitzoff T., Filar B., Anderson H., Roff H., Allen G.C., Steinhardt J., Flynn C., Beard, S., Belfield H., Farquhar S., Amodei D., *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*, arXiv przesłano 20 lutego 2018 r. (wersja 1), ostatnia aktualizacja 1 grudnia 2024 r, https://arxiv.org/abs/1802.07228 (dostęp: 1.3.2026 r.)

Campedelli G. M., *A Criminology of Machines*, arXiv przesłano 4 listopada 2025 r. (wersja 1), ostatnia aktualizacja 6 listopada 2025 r. https://arxiv.org/abs/2511.02895 (dostęp: 1.3.2026 r.)

Caneppele S., Calderoni F., *Crime-as-a-Service and the Transformation of Criminal Tools*. [w:] Miró-Llinares F., Johnson S. D. (red.), *The Routledge Handbook of Technology, Crime and Justice*. Routledge 2022. s. 213–229.

Ciancaglini V., Gariuolo S., Hilt S., McArdle R., Vosseler R., *AI Assistants in the Future: Security Concerns and Risk Management*, Trend Micro, 6 grudnia 2024, https://www.trendmicro.com/vinfo/gb/security/news/security-technology/looking-into-the-future-risks-and-security-considerations-to-ai-digital-assistants (dostęp: 1.3.2026 r.)

Choi KS., Lee C.S., Louderback E.R., *Historical Evolutions of Cybercrime: From Computer Crime to Cybercrime*. [w:] Holt T., Bossler A. (red.) *The Palgrave Handbook of International Cybercrime and Cyberdeviance*. Palgrave Macmillan, Cham 2020 (dostęp: 1.3.2026 r.)

Clarke R. V., *Technology, Criminology and Crime Science*, European Journal on Criminal Policy and Research 2004, t. 10, s. 55-63. (dostęp: 1.3.2026 r.)

Cornish D. B., Clarke R. V., *The Reasoning Criminal: Rational Choice Perspectives on Offending*. Transaction Publishers, 2014.

Europol, *Facing reality? Law enforcement and the challenge of deepfakes*. Publications Office of the European Union. https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes (dostęp: 1.3.2026 r.)

Felson M., Clarke R. V., *Opportunity Makes the Thief: Practical Theory for Crime Prevention*. Police Research Series, Paper No. 98. Home Office, London 1998 https://resaud.net/wp-content/uploads/2024/03/Felson-et-Clarke-1998-Opportunity-Makes-the-Thief.-Practical-Theory-for-.pdf (dostęp: 1.3.2026 r.)

Financial Action Task Force (FATF), *Second 12-Month Review of the Revised FATF Standards on Virtual Assets and Virtual Asset Service Providers*. FATF/OECD., https://www.fatf-gafi.org/content/dam/fatf-gafi/guidance/Second-12-Month-Review-Revised-FATF-Standards-Virtual-Assets-VASPS.pdf.coredownload.pdf (dostęp: 1.3.2026 r.)

Grycuk A.(2021), *Fake newsy, trolle, boty i cyborgi w mediach społecznościowych*, „Analizy BAS" 2021, nr 1(152), s. 1-12.
https://www.researchgate.net/publication/349109775_Fake_newsy_trolle_boty_i_cyborgi_w_mediach_spolecznosciowych (dostęp: 1.3.2026 r.)

Hayward K.J., Maas M.M., *Artificial intelligence and crime: A primer for criminologists*. Crime, Media, Culture: An International Journal, Tom 17 Numer 2, sierpień 2021, s. 209-233.https://matthijsmaas.com/uploads/Hayward%20and%20Maas%20-%202020%20-%20Artificial%20intelligence%20and%20crime%20A%20primer%20for%20cr.pdf (dostęp: 1.3.2026 r.)

Jakubski K.J., *Przestępczość komputerowa – zarys problematyki*, „Prokuratura i Prawo" 1996, nr 12, s. 34-50.
Jakubski K.J. i in., *Ciemne strony cyberprzestrzeni – wybrane aspekty* [w:] Chodźko E., Talarek K. (red.):*Wyzwania i problemy społeczeństwa w XXI wieku*. *Tom I*, Lublin 2020.
Jakubski K.J., *Niebezpieczna Sztuczna Inteligencja*, „Cybersecurity & Cybercrime," 2024, t. 1, Numer specjalny: PT XXI 2023, s. 247-301
Jakubski K.J., *Od CaaS do AI-CaaS. Kryminologiczne konsekwencje demokratyzacji narzędzi przestępczych*, artykuł złożony do publikacji w czasopiśmie Przegląd Policyjny
Jankowska M., *Podmiotowość prawna sztucznej inteligencji?* [w]: A. Bielska-Brodziak (red.), *O czym mówią prawnicy, mówiąc o podmiotowości*, Katowice 2021, s. 171-196
Jarosiewicz A., Kulesza J. (2023). *Problem atrybucji w cyberprzestrzeni: Wyzwania dla prawa międzynarodowego*, „Państwo i Prawo" 2023, nr 78(5), s. 45-62.
Kemp S., Halonen L., Pölönen I. (2022). *The European Union's Artificial Intelligence Act: A critical review from a criminal law perspective*, „New Journal of European Criminal Law" 2022, nr 13(4), s. 456-478.
Królikiewicz J., *Sztuczna inteligencja w ujęciu odpowiedzialności karnej – luka w odpowiedzialności czy możliwość jej przypisania?*, „Problemy Prawa Karnego" 2021, nr 5(1), s. 45-68.
Lavorgna A., *Organized Crime and Cybercrime* [w:] Holt T.J., Bossler A.M. (red.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, Palgrave Macmillan, Cham 2020, s. 117-234.
Leukfeldt E.R., *Cybercrime and the scalability of offending*. [w:] Holt T.J., Bossler A.M. (red.), *The human factor of cybercrime*, Routledge 2020, s. 255–272, https://www.google.pl/books/edition/The_Human_Factor_of_Cybercrime/1YK1DwAAQBAJ (dostęp: 1.3.2026 r.)
Lisowski E., *AI Agents vs Agentic AI: What's the Difference and Why Does It Matter?*, Medium, 18 grudnia 2024, https://tutorials.botsfloor.com/ai-agents-vs-agentic-ai-whats-the-difference-and-why-does-it-matter-03159ee8c2b4 (dostęp: 1.3.2026 r.)
McGlynn C., Johnson K., Rackley E., Henry N., Gavey N., Flynn A., Powell A., '*It's Torture for the Soul'*: *The Harms of Image-Based Sexual Abuse*, „Social & Legal Studies" 2020, nr 30(4), s. 541–562, (dostęp: 1.3.2026 r.)
McGuire M., *Technology, Crime and Justice: The Question Concerning Technomia.* Routledge 2012.
Nerantzi E., Sartor G., *Hard AI Crime: The Deterrence Turn,* „Oxford Journal of Legal Studies" 2024, nr 44(3), s. 673-701, (dostęp: 1.3.2026 r.)
National Institute of Standards and Technology (NIST), *AI risk management framework: AI RMF (1.0)*. U.S. Department of Commerce. https://www.nist.gov/itl/ai-risk-management-framework (dostęp: 1.3.2026 r.)
Parker D.B., *Criminal Justice Resource Manual on Computer Crime* , Departament Sprawiedliwości USA 1980, https://www.ojp.gov/pdffiles1/Digitization/118214NCJRS.pdf (dostęp: 1.3.2026 r.)
Pinch T.J., Bijker, W.E., *The social construction of facts and artifacts*. [w:] Bijker W.E., Hughes T.P., Pinch T. J. (red.), *The Social Construction of Technological Systems*. MIT Press 2012.
Rejmaniak R., *Autonomiczność systemów sztucznej inteligencji jako wyzwanie dla prawa karnego*, „Roczniki Nauk Prawnych" 2021, nr 31(3), s. 89-112.
Russell S., Norvig, P., *Artificial intelligence: A modern approach* (4th ed.), polskie wydanie: *Sztuczna inteligencja. Nowe spojrzenie*. Wydanie IV. Tom 1 i 2, (tłum. Grażyński A.), Helion 2023.
Schjolberg S., *The History of Cybercrime* Third edition, Books on Demand 2020.

Surma J., *Hakowanie Sztucznej Inteligencji*, Warszawa 2022.
Trend Micro, *AI Pulse: Sticker Shock, Rise of the Agents, Rogue AI*, Trend Micro Research 29 sierpnia 2024, https://www.trendmicro.com/pl_pl/research/24/h/agentic-ai-takeover.html (dostęp: 1.3.2026 r.)
Trend Micro, *The Road to Agentic AI: Navigating Architecture, Threats, and Solutions*, Trend Micro Research 28 lipca 2025, https://www.trendmicro.com/vinfo/us/security/news/security-technology/the-road-to-agentic-ai-navigating-architecture-threats-and-solutions (dostęp: 1.3.2026 r.)
Trend Micro, *The AI-fication of cyberthreats: Security predictions for 2026*. Trend Micro Research , https://documents.trendmicro.com/assets/research-reports/the-ai-fication-of-cyberthreats-trend-micro-security-predictions-for-2026.pdf (dostęp: 1.3.2026 r.)
Trend Micro, *Trend Micro Predicts 2026 as the Year Cybercrime Becomes Fully Industrialized*. Trend Micro Security News 25 listopada 2025, https://newsroom.trendmicro.com/2025-11-25-Trend-Micro-Predicts-2026-as-the-Year-Cybercrime-Becomes-Fully-Industrialized (dostęp: 1.3.2026 r.)
UNICRI, *Artificial intelligence and robotics for law enforcement*. United Nations Interregional Crime and Justice Research Institute, . https://unicri.org/sites/default/files/2025-06/UNICRI%20-%20Artificial%20intelligence%20and%20robotics%20for%20law%20enforcement.pdf (dostęp: 1.3.2026 r.)
UNODC, *Emerging threats: The intersection of criminal and technological innovation in the use of automation and artificial intelligence in the cybercrime landscape of Southeast Asia*, United Nations Office on Drugs and Crime, wrzesień 2025, https://www.unodc.org/roseap/uploads/documents/Publications/2025/UNODC_Report_Emerging_threats_-_The_intersection_of_criminal_and_technological_innovation_in_the_use_of_automation_and_AI.pdf (dostęp: 1.3.2026 r.)
Weise K., Metz C., *When AI chatbots hallucinate*. The New York Times, 29 marca 2023, https://www.nytimes.com/2023/03/29/technology/ai-chatbots-hallucinations.html (dostęp: 1.3.2026 r.)
Woolley S.C, Howard P.N. (red.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, New York 2018. (dostęp: 1.3.2026 r.).
Woźniak M., *Sztuczna inteligencja jako przedmiot i narzędzie przestępstwa*, Warszawa 2022.
Vectra AI, *Spear phishing: How targeted attacks work and how to stop them*. VECTRA AI INC https://www.vectra.ai/topics/spear-phishing (dostęp: 1.3.2026 r.)
Vicente D.M., Pereira R.S., Alves Leal A. (red.), *Legal aspects of autonomous systems: A comparative approach.* Springer International Publishing. (Series: Data Science, Machine Intelligence and Law) 2024, https://www.springerprofessional.de/en/legal-aspects-of-autonomous-systems/26586284 (dostęp: 1.3.2026 r.)
Zając P., *Wiktymologia w epoce cyfrowej: Nowe zagrożenia związane z wykorzystaniem AI*, „Archivum Criminologiae" 2023, nr15(1), s. 87-105.