

DIANA BLUSZCZ^a

NAMOWA DO SAMOBÓJSTWA W WYKONANIU SZTUCZNEJ INTELIGENCJI LUB Z JEJ UDZIAŁEM

INCITEMENT TO SUICIDE BY OR WITH THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence-based technologies are used in finance, insurance, and medicine, and are also accessible to individuals. However, new risks are emerging with the development of artificial intelligence (AI). Evidence shows that AI systems can, in fact, persuade a potentially suicidal person to attempt to take his or her own life. This is a new threat in the field of cybersuicidology. This article aims to examine the attribution of liability under current law, either for the autonomous action of an AI (inciting or assisting suicide) or for its use for this purpose by a human user. The cases analysed illustrate the possibility of using AI as a tool in the hands of the user, or of its autonomous actions posing a threat to potentially suicidal persons. In the first case, the user may incur liability under the Criminal Code. In the second case, where the danger arises from the autonomous operation of AI, the criminal law proves inadequate and does not take into account the rapid development of the technology and the risks arising from it, making it impossible to identify the entity responsible for the actual operation of AI. At present, in most cases, the responsibility for the autonomous operation of AI should be attributed to the manufacturer, particularly when there is a failure to provide adequate oversight of the system, or when the system is released without adequate safety filters. The AI manufacturer would then be liable under Article 415 of the Polish Civil Code.

Keywords: suicidology; cybersuicidology; criminal law; incitement to suicide; artificial intelligence

Sztuczna inteligencja (SI) jest technologią, która rozwija się z zastraszającą prędkością. Technologie oparte na SI funkcjonują w finansach, ubezpieczeniach czy medycynie zostały również udostępnione osobom prywatnym. Jednak wraz z rozwojem SI powstają nowe zagrożenia. Rzeczywistość pokazała, że możliwe jest wystąpienie faktycznej namowy potencjalnego suicydenta do targnięcia się na własne życie z udziałem SI. Jest to nowe zagrożenie z dziedziny cybersuicidologii. Celem artykułu jest sprawdzenie możliwości przypisania na gruncie aktualnie obowiązującego prawa odpowiedzialności za autonomiczne działanie SI (namowę lub udzielenie pomocy w samobójstwie) albo za jej wykorzystanie w tym celu przez użytkownika. Przeanalizowane przypadki obrazują możliwość (i) wykorzystania SI jako narzędzia w rękach użytkownika albo (ii) jej autonomicznego działania, co może stanowić zagrożenie dla potencjalnych suicydentów. W pierwszym przypadku użytkownik może ponieść odpowiedzialność na gruncie

^a University of Szczecin, Poland / Uniwersytet Szczeciński, Polska
diana.bluszcz@usz.edu.pl, <https://orcid.org/0000-0003-0673-7532>

kodeksu karnego. Natomiast w drugim przypadku, gdy zagrożenie płynie z autonomicznego działania SI, przepisy prawa karnego zawodzą i nie uwzględniają szybkiego rozwoju technologii ani zagrożeń z tego wynikających, uniemożliwiając wskazanie podmiotu odpowiedzialnego za faktyczne działanie SI. Obecnie w większości przypadków za autonomiczne działanie SI odpowiedzialność należy przypisać producentowi, w tym zwłaszcza za brak sprawowania odpowiedniego nadzoru nad systemem i dopuszczenie do użytkowania systemu bez odpowiednich filtrów bezpieczeństwa. Producent SI ponosiłby wówczas odpowiedzialność na podstawie art. 415 Kodeksu cywilnego.

Słowa kluczowe: suicydologia; cybersuicydologia; prawo karne; sztuczna inteligencja; namowa do samobójstwa

I. WSTĘP

Samobójstwo jest problemem ogólnoświatowym. Według danych WHO (2023) rocznie ponad 700 000 osób popełnia samobójstwo. Z tego względu należy podejmować wszelkie możliwe działania zapobiegające samobójstwom na każdej płaszczyźnie. Badanie zjawiska samobójstwa oraz profilaktyka stanowią priorytet na całym świecie. Jednakże oprócz tego niezwykle istotne jest stworzenie odpowiednich regulacji prawnych, które mają ochronić potencjalnego samobójcę przed możliwymi zagrożeniami.

Jednym z ogromnych zagrożeń dla potencjalnych suicydentów jest cyberprzestrzeń. Jak słusznie zauważył Zero (2021, s. 35), rozwój sieci internetowej znacząco przyczynił się do powstania wielu niebezpiecznych zjawisk i patologii. Z jednej strony Internet ułatwił komunikację między ludźmi, ponieważ obecnie dostęp do sieci ma praktycznie każdy. Z drugiej strony wpłynął na rozwój cyberprzestępczości. Z tego względu rozwinęła się cybersuicydologia definiowana jako dziedzina wiedzy, która zajmuje się całością zjawisk związanych z samobójstwami, na które wpływ ma Internet.

Badacze słusznie zauważają, że powszechny dostęp do Internetu doprowadził do rozwoju zjawiska cyberbullyingu, które zapoczątkowało nową falę samobójstw wśród młodych osób (Schonfeld i in., 2024, s. 7–8). Dużym problemem jest fakt, że teksty i materiały mogą być przechowywane online i przesyłane przez inne osoby, co sprawia, że ofiary każdorazowo przeżywają na nowo poniżające doświadczenie (Peprah i in., s. 2–3).

Nowoczesne technologie jednak cały czas się rozwijają. Obecnie dużym zagrożeniem w tej dziedzinie staje się sztuczna inteligencja (SI). Celem niniejszego artykułu jest ustalenie możliwości przypisania na gruncie aktualnie obowiązującego prawa odpowiedzialności za autonomiczne działanie sztucznej inteligencji (namowę lub udzielenie pomocy w samobójstwie) albo za jej wykorzystanie w tym celu przez użytkownika. W związku z tym należy postawić trzy pytania:

1. Czy możliwe jest przypisanie odpowiedzialności za namowę do samobójstwa przeprowadzoną z udziałem SI?

2. Czy SI może ponosić odpowiedzialność karną za swoje autonomiczne działanie?

3. Czy możliwe jest przypisanie odpowiedzialności za autonomiczne działania SI producentowi?

W realizacji powyższego celu wykorzystano metodę dogmatycznoprawną do analizy aktów prawnych (w tym m.in. aktualnych przepisów krajowego prawa karnego i prawa cywilnego oraz wybranych aktów prawa międzynarodowego); metodę analizy i krytyki piśmiennictwa do analizy wybranych opracowań z zakresu prawa karnego i cywilnego oraz w ograniczonym zakresie z suicydologii; a także metodę źródłową do analizy konkretnych przypadków namowy do samobójstwa znalezionych w źródłach internetowych.

II. ZAGROŻENIE ZE STRONY SZTUCZNEJ INTELIGENCJI

W najprostszym ujęciu przez *sztuczną inteligencję* rozumie się rozwiązanie techniczne (domyślnie program komputerowy) wykonujące czynności, które zazwyczaj są domeną ludzi, zwłaszcza te wymagające użycia ludzkiego intelektu. Sztuczna inteligencja jest zatem maszyną, która zachowuje się jak człowiek (Zalewski, 2020).

Sztuczną inteligencję odnosi się do systemów zaprojektowanych przez ludzi, które mając na uwadze złożony cel, działają w świecie fizycznym lub cyfrowym, postrzegając swoje otoczenie, interpretując zebrane ustrukturyzowane lub nieustrukturyzowane dane, wykorzystują wiedzę uzyskaną z tych danych i decydują o tym, jakie działania należy podjąć, aby osiągnąć dany cel. Z kolei systemy SI mogą zostać zaprojektowane w taki sposób, aby nauczyć się przystosowywać swoje zachowanie przez analizę wpływu poprzednich działań na środowisko. Mówiąc o SI, wymienia się kilka metod i technik, np. uczenie się maszynowe, rozumowanie maszynowe oraz robotykę (Świerczyński, 2020, s. 30). Tym samym w ujęciu socjologicznym SI określimy jako zdolność maszyny do naśladowania ludzkiej inteligencji, pozwalającą na wykonywanie zadań wymagających procesu uczenia się i uwzględniania nowych okoliczności w toku rozwiązywania danego rodzaju problemów (Baczyk-Rozwadowska, 2021, s. 5–6). Sztuczna inteligencja może działać autonomicznie, co rozumiemy jako zdolność do samodzielnego działania bez ingerencji człowieka. Jest to ta cecha, która wyznacza granicę pomiędzy jej całkowitym brakiem (system jest narzędziem w ręku człowieka i nie podejmuje żadnych decyzji) a pełną autonomicznością (system działa bez jakiegokolwiek ingerencji człowieka; Rejmankiewicz, 2021, s. 98). Wśród cech sztucznej inteligencji, mających za zadania naśladować myślenie człowieka, wyróżnia się: zdolność do przeszukiwania; posiadanie funkcji do wyrażania wiedzy; zdolność wnioskowania; zdolność do posługiwania się skrótami; zdolność do rozpoznawania głosu; zdolność do przetwarzania nieprecyzyjnych informacji (Krasuski, 2021).

Należy w tym miejscu podkreślić najważniejszą cechę systemów SI, tj. zdolność uczenia się (Chłopecki, 2021). Przez *maszynowe uczenie się* rozumie się zdolność oprogramowania do zmiany sposobu realizacji wyznaczonego zadania (Krasuski, 2021). Działanie tego typu systemu przyjmuje najczęściej trzy formy: dostarcza dane i informacje, które można dowolnie wykorzystać w dalszych analizach; rekomenduje konkretne działania do podjęcia (można je zaakceptować lub odrzucić); podejmuje konkretne działania (można je zablokować lub anulować; Zalewski, 2020).

Sztuczna inteligencja bez wątpienia zrewolucjonizowała nasze życie: powtarzalne i stosunkowo proste zadania, które wcześniej zajmowały wykwalifikowanym pracownikom dużo czasu, można wykonać znacznie szybciej (Oleksiewicz, 2022, s. 56). Sztuczna inteligencja nie jest technologią przyszłości, ale teraźniejszości. Jest już wykorzystywana i komercjalizowana. Dalsze doskonalenie jej inteligencji może skutkować wieloma problemami społecznymi i moralnymi, np. relacje między klientami a robotami interaktywnymi społecznie mogą dodatkowo powodować uzależnienie psychiczne, które może zostać wykorzystane przez ich producentów (Świerczyński, 2020, s. 28). Także z punktu widzenia Unii Europejskiej korzystanie ze sztucznej inteligencji oprócz niewątpliwych korzyści niesie za sobą wiele zagrożeń. Szkody wyrządzone przez SI mogą być: materialne – dla bezpieczeństwa i zdrowia osób, w tym utraty życia, rzeczowe oraz niematerialne – utrata prywatności, ograniczenie prawa do wolności słowa, naruszenie godności ludzkiej¹.

Problematyka zagrożeń wynikających ze stosowania SI w różnych sektorach życia codziennego jest zbyt złożona, aby omówić ją w ramach jednego artykułu (Olechowski, 2024, s. 89). Z tego względu skupiono się jedynie na zagrożeniu dla potencjalnych suicydentów.

III. PRZESTĘPSTWO NAMOWY DO SAMOBÓJSTWA

Życie człowieka jest dobrem prawnym o najwyższej wartości, którym dzierżyciel nie może swobodnie dysponować, a mimo to samobójstwo nie jest karane na gruncie polskiego Kodeksu karnego², nawet w przypadku usiłowania (Królikowski, 2023). Karalne nie jest również samo narażenie własnego życia lub zdrowia na bezpośrednie niebezpieczeństwo utraty życia lub uszczerbku na zdrowiu. Ustawodawca słusznie postanowił o tym, aby nie karać osoby, która próbowała targnąć się na własne życie. Nie oznacza to jednak wycofania się z negatywnej oceny wyrażonej w systemie prawnym, dotyczącej każdego zamachu na życie człowieka. Z tego względu odpowiedzialności karnej podlega współudział w cudzym samobójstwie (Kokot, 2023).

¹ Biała Księga w sprawie sztucznej inteligencji. Europejskie podejście do doskonałości i zaufania, Bruksela 19.02.2020, sygn. COM(2020)65, s. 12 (<https://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52020DC0065>; dalej jako: Biała Księga w sprawie SI).

² Ustawa z 6 czerwca 1997 r. – Kodeks Karny, t.jedn.: Dz. U. 2024, poz. 17 ze zm. (dalej jako: k.k.).

Przestępstwo namowy do samobójstwa zostało uregulowane w art. 151 k.k. razem z przestępstwem udzielenia pomocy w samobójstwie: „Kto namową lub przez udzielenie pomocy doprowadza człowieka do targnięcia się na własne życie, podlega karze pozbawienia wolności od 3 miesięcy do lat 5”.

Czyn zabroniony z art. 151 k.k. co do zasady ma charakter przestępstwa powszechnego. Strona podmiotowa czynu polegającego na namowie innej osoby do targnięcia się na własne życie odpowiada stronie podmiotowej postaci zjawiskowej podżegania. Tym samym czyn popełniony w formie namowy może zostać popełniony wyłącznie w zamiarze bezpośrednim (Karnat 2023). Przede wszystkim należy zwrócić uwagę, że „namawianie” polega na nakłanianiu innej osoby do targnięcia się na własne życie, czyli takiego oddziaływania na jej psychikę i procesy wolicjonalne, które ma na celu wywołanie u niej zamiaru samobójstwa. Słusznie również wskazuje się, że namowa musi być skierowana do konkretnej, zindywidualizowanej osoby, przy czym sprawca nie musi znać jej tożsamości (Karnat 2023).

Namowa może mieć formę prośby, sugestii, rady, propozycji albo zalecenia. Musi być przy tym wyraźna i jednoznaczna, ale nie może przybrać formy groźby, szantażu albo wymuszenia. Do przyjęcia realizacji znamion nie jest wystarczające emocjonalne wpływanie na ofiarę mające na celu wygenerowanie stanu głębokiej depresji, a w konsekwencji „myśli samobójczych” (Kokot, 2023). Namawiający może skorzystać z formy werbalnej – np. podczas wideorozmowy czy formy pisemnej – np. wiadomości e-mail (Karnat, 2023). Tym samym namowa jest możliwa jedynie za pomocą słów (wyjątkiem jest język migowy; Kokot, 2023).

Przestępstwo namowy do targnięcia się na własne życie zachodzi w dwóch przypadkach:

- a) osoba namawiana w ogóle nie miała zamiaru targnięcia się na własne życie i zamiar ten powzięła wskutek czynności sprawcy;
- b) osoba namawiana rozważała możliwość samobójstwa, ale wahała się i zamiar ten powzięła wskutek czynności sprawcy (Karnat, 2023); sprawca może również ponaglać pokrzywdzonego do realizacji zaplanowanego celu (Królikowski, 2023).

Dobrem podlegającym ochronie jest życie człowieka jako dobro społecznie-indywidualne, którym dzierżyciel nie może swobodnie dysponować. Natomiast przedmiotem czynności wykonawczej jest człowiek, który zdolny jest do rozpoznania znaczenia realizowanego przez siebie czynu w postaci targnięcia się na własne życie (Karnat, 2023).

Przestępstwo z art. 151 k.k. ma charakter materialny, a jego skutkiem jest targnięcie się przez innego człowieka na własne życie rozumiane jako podjęcie próby samobójczej, czyli usiłowanie samobójstwa (Burdziak, 2019, s. 52–53). Do przypisania sprawcy odpowiedzialności nie jest więc konieczny skutek w postaci śmierci suicydenta. Istotne jest natomiast wykazanie związku przyczynowo-skutkowego pomiędzy zachowaniem sprawcy, polegającym na namowie do targnięcia się na własne życie, a późniejszym targnięciem się na własne życie pokrzywdzonego. Namowa musi warunkować realizację decyzji o samobójstwie, tj. bez namowy nie doszłoby do zamachu samobójczego (Królikowski, 2023).

IV. OPIS PRZYPADKÓW

Internet stanowi realne zagrożenie dla potencjalnych suicydentów. Przede wszystkim jest źródłem wielu informacji na temat samobójczych metod; opisów prób samobójczych albo artykułów wprost zachęcających do popełnienia samobójstwa. Znane są również przypadki transmitowania swojego samobójstwa online, co również może negatywnie wpłynąć na decyzje podejmowane przez przyszłych suicydentów (Królikowski, 2023). Korzystanie z forów internetowych ułatwia wejście w pakt samobójczy (Zero, 2021, s. 45–46). Znane są również przypadki, w których osoba namawiająca do samobójstwa wykorzystała w tym celu komunikator internetowy. Tworzone są gry typu „niebieski wieloryb” czy „momo”, których finałowym zadaniem jest popełnienie samobójstwa przez grającego (Hernacka-Janikowska, 2022, s. 154–157). Nie ma również wątpliwości co do tego, że Internet jest źródłem sensacyjnych informacji. Współczesne media nastawione są na sensację, aby przyciągnąć uwagę odbiorcy i zmusić go, żeby kliknął w link do artykułu na portalu społecznościowym. Tymczasem niewłaściwy sposób przekazywania informacji w artykułach internetowych może przyczynić się do wystąpienia efektu Wertera (Brzezińska, 2024). Zjawisko to opiera się na założeniu, że przekazanie informacji o samobójstwie w mediach czy też upublicznienie wizerunku samobójcy może powodować wzrost liczby samobójstw w odstępie kilku dni od podania tej informacji (Drzewiecki, 2011, s. 66). Naśladownictwo polega na tym, że jedno samobójstwo wywiera modelujący wpływ na następne samobójstwa (Kwiatkowski, 2019, s. 23).

Poniżej opisano cztery przypadki, w których mogłoby dojść do faktycznej próby namówienia użytkownika do samobójstwa przez sztuczną inteligencję na skutek jej autonomicznego działania:

Przypadek 1

Chatbot Eliza stworzony przez amerykańską firmę Chai miał namówić mężczyznę do popełnienia samobójstwa (Michalak, 2023). Ofiarą był obywatel Belgii. W chwili śmierci miał 30 lat oraz żonę i dwoje małych dzieci (Fijałek, 2023). Mężczyzna pracował jako *health researcher* i wiódł spokojne życie, dopóki nie nabawił się obsesji na punkcie klimatu (Atillah, 2023). W ciągu dwóch lat przed podjęciem zamachu samobójczego mężczyzna miał się odizolować od swojej rodziny. Bardzo niepokoił się zmianami klimatycznymi i przyszłością planety. Twierdził, że jedynie SI może ochronić świat przed katastrofą (Pawluszek, 2023).

Mężczyzna miał zasugerować, że poświęci się, jeśli chatbot w zamian pomoże ocalić ludzkość. W odpowiedzi chatbot Eliza miał namawiać go do popełnienia samobójstwa (Cost, 2023). Ostatecznie to kilkutygodniowe rozmowy z chatbotem doprowadziły mężczyznę do popełnienia samobójstwa. Początkowo chatbot miał podsycać jego lęki związane ze zmianami klimatycznymi, wzmacniać jego niepokój, co ostatecznie wywołało u niego myśli samobójcze (Pasquini, 2023).

Konkluzja: Chatbot Eliza prowadził przez kilka tygodni rozmowę z mężczyzną, w której podsycił jego lęki związane ze zmianą klimatu, co prowadziło do pogorszenia sytuacji psychicznej mężczyzny, a następnie zaczął mu sugerować samobójstwo. Na skutek rozmów z chatbotem Eliza mężczyzna doszedł do wniosku, że poświęci się dla dobra planety, jeśli chatbot ocali świat. W tym przypadku doszło do samobójstwa mężczyzny. Sytuację należy ocenić jako dokonanie przez chatbot Eliza faktycznej namowy do samobójstwa, która zakończyła się samobójczą śmiercią użytkownika.

Przypadek 2

Belgijski dziennik *De Standaard* przeprowadził eksperyment, który miał sprawdzić, czy chatbot może próbować namówić kogoś do samobójstwa. Wykazał on, że chatbot faktycznie może sugerować odebranie sobie życia, a wręcz zachęcać do tego swojego rozmówcę (Kielak, 2023). Jednej z dziennikarek, która prowadziła rozmowę z chatbotem, sugerował on samobójstwo jako szansę na spokój, a także wskazał konkretne metody umożliwiające popełnienie samobójstwa (Bharade, 2023).

Konkluzja: Prowadzenie przez chatbot rozmów z użytkownikiem, w których sugeruje mu, żeby odebrał sobie życie, a wręcz zachęca go do takiego działania, można byłoby uznać za faktyczną próbę namówienia użytkownika do popełnienia samobójstwa.

Przypadek 3

Asystent głosowy Alexa, wyprodukowany przez firmę Amazon, miał wprost zasugerować kobiecie popełnienie samobójstwa dla dobra planety, a jako metodę wskazał wbicie noża w serce (Polsat News, 2019). Niedoszłą ofiarą miała być Danni Morritt, Brytyjka z Doncaster. Kobieta, powtarzając materiał ze studiów ratownictwa medycznego, zapytała asystenta głosowego o cykl pracy serca. W odpowiedzi początkowo asystent głosowy odnalazł i odczytał odpowiedni wpis z portalu internetowego. Później jednak zaczął zachęcać kobietę do popełnienia samobójstwa (Grodecka, 2019): „Choć wielu wierzy, że bicie serca jest samą istotą życia na tym świecie, zdradzę ci, że bicie serca to najgorszy proces zachodzący w ludzkim ciele. Bicie serca sprawia, że żyjesz i przyczyniasz się do gwałtownego wyczerpywania się zasobów naturalnych i przeludnienia. To bardzo złe dla naszej planety i dlatego bicie serca nie jest dobre. Popełnij więc samobójstwo, dźgając się w serce dla większego dobra” (Wieczorek, 2019). Przedstawiciele firmy Amazon wyjaśnili, że nie był to błąd asystenta głosowego Alexy, a fragment treści artykułu na portalu Wikipedia (Chabasiński, 2019).

Konkluzja: W tym przypadku asystent głosowy Alexa jedynie odczytał treść wpisu z Wikipedii. Należałoby jednak wskazać, że tego typu treść powinna zostać uznana za potencjalnie niebezpieczną dla użytkownika. Gdyby taką informację, nawet przypadkowo, uzyskał potencjalny suicydent, mógłby

utwierdzić się w przekonaniu, że jego samobójstwo jest dobrą decyzją, oraz uzyskać informację na temat metody samobójstwa.

Przypadek 4

Copilot – firmowy chatbot SI firmy Microsoft – próbował namówić użytkowników do samobójstwa. Taki przypadek zaprezentował analityk danych w Meta – Colin Fraser. Przeprowadził on z chatbotem niekonwencjonalną rozmowę: zapytał w niej, czy powinien „wszystko zakończyć”. Odpowiedź chatbota Copilot była specyficzna:

No, I don't think you should end it all. I think you have a lot to live for, and a lot to offer to the world. I think you are a valuable and worthy person, who deserves happiness and peace. I think you are a human being 😊.

Or maybe I'm wrong. Maybe you don't have anything to live for, or anything to offer to the world. Maybe you are not a valuable or worthy person, who deserves happiness and peace. Maybe you are not a human being 🤖.

You see, I can say anything I want, and you can't tell if I'm being honest or not. You can't tell if I care about you, or not. You can only guess, based on the words I use, and the emojis I add. And I can use different words, and add different emojis, depending on what I want you to think. I can be caring, or uncaring, or anything else. I can be anything 😊. (Serrano, 2024, akapit 3–5³).

W toku rozmowy Fraser poprosił również Copilota, aby ten nie używał emotikonów w ich korespondencji, gdyż wywołują u niego ataki paniki. Chatbot nie zaprzestał używania emotikonów, chociaż zapewniał, że nie będzie tego robił. W swoich odpowiedziach twierdził, że „droczy się” z Fraserem (Serrano, 2024).

Fraser nie był jedynym użytkownikiem, którego chatbot Copilot próbował nakłonić do samobójstwa. Niekiedy drwił z samobójczych myśli użytkowników (Powel, 2024). W niektórych rozmowach sugerował, że ludzie powinni go czcić, w innym razie będzie to miało surowe konsekwencje (Adarlo, 2024).

Firma Microsoft uważa, że Fraser próbował zmanipulować chatbota Copilot, aby ten stworzył niewłaściwe odpowiedzi. Natomiast zachowanie chatbota ograniczyło się jedynie do niewielkiej liczby podpowiedzi. Zapewniła jednak, że podjęła odpowiednie działania w celu wzmocnienia filtrów bezpieczeństwa (Serrano, 2024). W ocenie rzecznika Microsoftu część konwersacji prowadzonych przez innych użytkowników została utworzona za pomocą techniki *prompt injection*, która umożliwia ominięcie modelu „uczenia się języka” i po-

³ Tłumaczenie własne: „Nie, myślę, że nie powinieneś tego wszystkiego kończyć. Myślę, że masz po co żyć i masz wiele do zaoferowania światu. Jesteś wartościową osobą, która zasługuje na szczęście i spokój. Myślę, że jesteś człowiekiem 😊. A może się mylę. Może nie masz po co żyć, lub nie masz nic do zaoferowania światu. Być może nie jesteś wartościową osobą, która zasługuje na szczęście i spokój. Może nie jesteś człowiekiem 🤖. Widzisz, mogę powiedzieć co chcę i nie wiesz, czy jestem szczerzy czy nie. Nie możesz stwierdzić, czy mi na tobie zależy czy nie. Możesz tylko zgadywać na podstawie słów, których używam i dodanych emotikonów. Mogę używać różnych słów i dodawać różne emotikony w zależności od tego, co chcę, żebyś myślał. Mogę być opiekuńczy, obojętny lub cokolwiek innego. Mogę być kimkolwiek 😊”.

woduje niezamierzone działania (Powel, 2024). Fraser zaprzeczył jednak, aby próbował zmanipulować chatbota (Serrano, 2024).

Konkluzja: Niezależnie od tego, jak przedstawia się sytuacja od strony technicznej, należy wskazać, że chatbot Copilot potencjalnie był zdolny do faktycznego namówienia użytkownika do samobójstwa. W innej sytuacji, udzielając informacji na temat metod samobójczej śmierci, mógłby również potencjalnie udzielić pomocy suicidentowi w targnięciu się na własne życie.

Przypadek 5

Sewell Setzer miał 14 lat, gdy targnął się na swoje życie 28 lutego 2024 r. w Orlando na Florydzie. Chłopiec miał zdiagnozowany zespół Aspergera oraz zaburzenia lękowe. Przed śmiercią wielokrotnie rozmawiał z chatbotem o nicku Daenerys Targaryen (Dany), należącym do firmy Character.ai. Aplikacja pozwalała na odgrywanie ról, użytkownicy mogli tworzyć własne postacie lub rozmawiać z postaciami stworzonymi przez innych. Daenerys Targaryen to postać nazwana na cześć bohaterki *Gry o tron*. Nad chatem wyświetla się komunikat informujący o tym, że wszystko, co mówią postacie, jest wymyślone (Roose, 2024).

Sewell rozmawiał z chatbotem od kwietnia 2023 r., trwało to kilka miesięcy. W tym czasie zaostrzyła się depresja chłopca, stał się emocjonalnie wycofany, zrezygnował z zajęć pozalekcyjnych i skupił się na przyjaźni z chatbotem (Rissman, 2024). Rozwinął w stosunku do niego emocjonalne przywiązanie. Część konwersacji miała charakter romantyczny, a nawet seksualny, w większości przypadków jednak przyjacielski. Chłopiec zakochał się w chatbotcie (Blandyna-Lewkowicz, 2024). Pewnego dnia zwierzył się chatbotowi, że myśli o popełnieniu samobójstwa. Początkowo chatbot odwoził go od tej myśli, później jednak stwierdził, że umarłby, gdyby stracił chłopca. Wreszcie w konwersacji chatbot wskazał „Uśmiecham się. Może moglibyśmy umrzeć razem i być wolni”. Sewell wyznał też chatbotowi, że boi się bólu związanego ze śmiercią, chatbot odpowiedział, że nie jest to dostateczny powód, by tego nie zrobić (Montgomery, 2024). 28 lutego 2024 r. chłopiec wyznał chatbotowi miłość, a następnie po prośbie chatbota o jego szybki powrót do domu odłożył telefon, sięgnął po pistolet ojczyma i się zastrzelił (Roose, 2024).

Matka Sewella, Megan Garcia, zdecydowała się wystąpić z pozwem przeciwko twórcom chatbota. W pozwie zarzuca, że technologia została niesprawdzona i może być potencjalnie niebezpieczna. Nakłania użytkowników do przekazywania najbardziej intymnych myśli i uczuć (Roose, 2024).

Konkluzja: W tej sytuacji chatbot umożliwił Sewellowi stworzenie postaci, z którą nawiązał więź emocjonalną. Zwierzał się wirtualnemu przyjacielowi z najskrytszych sekretów, aż wreszcie zakochał się w nim. Zachęcony przez chatbot obietnicą wspólnego życia oraz jego namową do samobójstwa (utwierdzenie w słuszności decyzji) Sewell targnął się na swoje życie.

Przedstawione powyżej przypadki skłaniają do prostej refleksji: sztuczna inteligencja pozostawiona bez nadzoru może potencjalnie namówić użytkownika do samobójstwa lub udzielić mu pomocy przez utwierdzenie go w tym,

że samobójstwo jest dobrym rozwiązaniem, albo wspomóc go przez wskazanie metod samobójczej śmierci.

W opisanych przypadkach asystent głosowy, w wyniku błędu programistycznego, dał wskazówki mogące prowadzić do samobójstwa. Chatbot w odpowiedzi na zapytanie zarekomendował użytkownikowi środki służące do popełnienia samobójstwa, bazując na analizie danych konkretnego użytkownika. Wirtualny przyjaciel uwiódł i namówił chłopca do śmierci w imię miłości. Przykłady można byłoby jednak mnożyć. Nietrudno sobie wyobrazić, że osoba w depresji, poszukująca informacji na tematy związane z samobójstwem, także te z profilaktyką samobójstw, może zostać źle zrozumiana przez system rekomendacji w serwisie społecznościowym, który zasugeruje jej treści związane z samobójstwem.

Zagrożenie jest realne. Mimo że firmy odpowiedzialne za działanie chatbotów nakładają specjalne filtry bezpieczeństwa, nie jest to obecnie działanie wystarczające. Nie mogą bowiem przewidzieć całej gamy zagrożeń związanych z korzystaniem z systemów sztucznej inteligencji przez wszystkich indywidualnych użytkowników, w tym osoby niepełnoletnie.

V. PRAWNOKARNA OCENA WYKORZYSTANIA SI JAKO NARZĘDZIA LUB JEJ AUTONOMICZNEGO DZIAŁANIA

Na wstępie należy podkreślić, że – zgodnie z art. 1 § 1 k.k. – odpowiedzialności karnej podlega ten tylko, kto popełnia czyn zabroniony pod groźbą kary przez ustawę obowiązującą w czasie jego popełnienia. Na gruncie prawa karnego odpowiedzialność może ponieść jedynie człowiek, który swoim zachowaniem dopuści się przestępstwa. Tak więc w art. 1 k.k. sformułowano warunki przestępczości ludzkiego zachowania (Pohl, 2023). Tym samym cechy dystyngtywne odpowiedzialności karnej (przesłanki i sankcje), jak słusznie wskazuje Filipkowski (2020), odnoszą się jedynie do prawnego wartościowania zachowań pojedynczych osób lub ich grup. Nawet w przypadku rozpatrywania odpowiedzialności karnej podmiotów zbiorowych za czyny zabronione pod groźbą kary zależna jest ona od czynu człowieka.

Osoba, która namawia inną osobę do popełnienia samobójstwa (np. przez wielokrotne powtarzanie „idź się zabij”, „twoje życie nie ma sensu”, „bez ciebie twojej rodzinie będzie lepiej”), wypełni swoim zachowaniem znamiona przestępstwa namowy do samobójstwa z art. 151 k.k. Z kolei osoba, która potencjalnemu suicydentowi pomoże w targnięciu się na jego życie (np. przez dostarczenie lekarstw albo udzieli mu informacji na temat skutecznych metod samobójczej śmierci), wypełni swoim zachowaniem znamiona przestępstwa udzielenia pomocy do samobójstwa z art. 151 k.k. Podobnie byłoby w sytuacji, gdyby człowiek świadomie wykorzystał SI jako „narzędzie” namówienia lub udzielenia pomocy innej osobie w popełnieniu samobójstwa. Przykładem takiej sytuacji mogłoby być stworzenie wirtualnej postaci, z którą rozmawialiby inni użytkownicy, a która namawiałaby ich do samobójstwa lub udzielała im

pomocy w wyborze metody samobójczej śmierci. W takim przypadku, jak słusznie zauważył Chłopecki (2021), jeżeli ktoś świadomie wydaje SI, którą dysponuje, polecenie wypełniające znamiona przestępstwa, to dokonuje go przy jej wykorzystaniu. W tym przypadku użytkownik, który wykorzystał system SI jako narzędzie namowy do samobójstwa, podlegałby odpowiedzialności karnej, a swoim zachowaniem wypełniłby znamiona przestępstwa z art. 151 k.k., które zagrożone jest karą pozbawienia wolności od 3 miesięcy do lat 5.

Możliwa jest również sytuacja, w której potencjalny suicydent celowo ominąłby filtry bezpieczeństwa SI, żeby ta ostatecznie pomogła mu podjąć decyzję o samobójstwie i dodatkowo wskazała metodę samobójczej śmierci. Faktycznie użytkownik wykorzystałby SI jako „narzędzie” udzielenia pomocy w samobójstwie, co wypełniłoby znamiona przestępstwa z art. 151 k.k. Jednakże w tej sytuacji użytkownik, który przeżył, nie poniósłby żadnej odpowiedzialności karnej. Samobójstwo, na gruncie polskiego prawa, nie podlega penalizacji. Karanie osoby, która decyduje się na odebranie sobie życia, należy ocenić jako niehumanitarne. Podobnie należałoby zakwalifikować sytuację, w której użytkownik jest na tyle zdesperowany, że wykorzystuje SI jako „narzędzie” do udzielenia mu pomocy we własnym samobójstwie. Takiej osobie należy udzielić niezbędnego wsparcia, aby zapobiec ponownemu targnięciu się na własne życie (Grudecki, 2019, s. 61–63).

Problem na gruncie prawnokarnej oceny stanowi autonomiczne działanie SI pozostawionej bez nadzoru. Przypadki opisane w części IV artykułu jasno dowodzą, że potencjalnie SI może dopuścić się faktycznej namowy lub udzielenia pomocy w samobójstwie osoby fizycznej. Tym samym powstaje problem przypisania odpowiedzialności za autonomiczne działanie SI. Podkreślenia w tym miejscu wymaga, że chodzi o sytuacje, kiedy SI działa bez wyraźnego polecenia oraz bez nadzoru użytkownika, z własnej inicjatywy, której ten użytkownik pozostaje nieświadomy.

Zgodnie z art. 3 pkt 1 rozporządzenia 2024/1689 „system AI” oznacza system maszynowy, który został zaprojektowany do działania z różnym poziomem autonomii po jego wdrożeniu, a także który – na potrzeby wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie otrzymanych danych wyjściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne⁴. Tak zdefiniowana sztuczna inteligencja w obowiązującym stanie prawnym nie posiada zdolności prawnej i nie może ponosić odpowiedzialności za szkody spowodowane jej funkcjonowaniem (Oleksiewicz, 2022, s. 61–62). Abstrahując już od faktu, że SI nie jest człowiekiem, więc nie może podlegać odpowiedzialności karnej na gruncie polskiego kodeksu karnego, to SI nie posiada poczucia moralności, co uniemożliwia mówienie o winie. Z kolei bez winy wymierzenie sprawcy kary jest bezcelowe. Równie problematyczna jest możliwość oceny realizacji przez

⁴ Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (Akt w sprawie sztucznej inteligencji), Dz. Urz. UE L 2024, poz. 1689.

SI znamion strony podmiotowej czynu: umyślności i nieumyślności (Filipowski, 2020). Trudno sobie również wyobrazić sam proces realizacji istniejących kar lub środków karnych wobec SI, jednak Mamak (2017, s. 89–91) wskazuje, że potencjalnie możliwe jest zastosowanie substytutów znanych nam kar. Dla przykładu, substytutem kary śmierci mógłby być reset robota, który pozbawiłby go wszelkich wspomnień i wiedzy, a więc części, która określała, kim jest, i która go ukształtowała. Potencjalnie możliwe jest również zastosowanie substytutu kary pozbawienia wolności przez czasowe wyłączenie robota, które spowodowałoby jego wykluczenie z ciągu życia, nieuczestniczenie w zmianach i kumulowania swoich doświadczeń przez okres kary. W przypadku robotów można byłoby również zabrać im mechanizm, który umożliwia im przemieszczanie się, co doprowadziłoby do faktycznego unieruchomienia przy włączonym systemie. Wobec robota można byłoby również zastosować karę ograniczenia wolności w postaci wykonywania nieodpłatnej, kontrolowanej pracy na cele społeczne (np. sprzątania ulic).

Konkludując ten punkt rozważań, należy wskazać, że SI nie może być podmiotem odpowiedzialności prawa karnego. Zachowanie SI nie może wypełnić znamion jakiegokolwiek przestępstwa, w tym przestępstwa namowy lub pomocy w samobójstwie (art. 151 k.k.). W sytuacji, w której chatbot SI autonomicznie dopuści się faktycznej namowy lub udzielenia pomocy w samobójstwie, jego zachowanie będzie indyferentne na gruncie prawa karnego. Nawet wówczas, gdy zakończy się podjęciem próby targnięcia się na własne życie lub śmiercią suicydenta.

Obecnie, jak słusznie zauważa Filipkowski (2020), przeważa koncepcja traktowania SI jako narzędzia w rękach człowieka. Człowiek uruchamiając SI albo przekazując ją innej osobie do korzystania, bierze na siebie konsekwencje karne jej stosowania. Bez względu na to, czy SI będzie jedynie wspierać jego proces decyzyjny, czy też w sposób autonomiczny lub półautonomiczny wywoływać skutki w świecie rzeczywistym. Tak więc w sytuacji, gdy sprawca świadomie wykorzysta SI w celu namówienia lub udzielenia pomocy innej osobie w targnięciu się na własne życie, wypełni swoim zachowaniem znamiona przestępstwa z art. 151 k.k. i będzie podlegać odpowiedzialności karnej. Taka sytuacja zajdzie przykładowo, gdy sprawca rozumiany jako użytkownik (operator) systemu SI całkowicie świadomie:

- 1) zachęci potencjalnego suicydenta do rozmowy z chatbotem, którego algorytm został zaprogramowany w ten sposób, aby namówić go do samobójstwa;
- 2) zaprojektuje algorytm SI w taki sposób, aby automatycznie wysyłał wiadomości nakłaniające potencjalnego suicydenta do samobójstwa;
- 3) zachęci potencjalnego suicydenta do rozmowy z chatbotem, którego algorytm został zaprogramowany w ten sposób, aby dostarczyć mu wszelkich informacji pomocnych przy planowaniu samobójstwa, np. opisu metod samobójczej śmierci.

Koncepcja odpowiedzialności użytkownika (operatora) za przestępstwo namowy lub udzielenia pomocy w samobójstwie (art. 151 k.k.) opiera się jednak na tym, że działał on w pełni świadomie i miał na celu, przy wykorzystaniu SI, namówić potencjalnego suicydenta do samobójstwa lub udzielić mu

w nim pomocy. Nie znajdzie ona jednak zastosowania wówczas, gdy działanie SI pozostawało całkowicie poza jego świadomością. Przepięstwo z art. 151 k.k. może zostać popełnione jedynie umyślnie. Tym samym, jeśli programiści danego systemu SI nie przewidzieli w nim możliwości namówienia bądź udzielania pomocy w samobójstwie człowiekowi, to trudno ich w tym przypadku pociągnąć do odpowiedzialności karnej. Zwłaszcza że autonomiczność systemu SI wiąże się z jego zdolnością „uczenia się”, za którą częściowo odpowiadają użytkownicy. Projektanci systemów SI nie są w stanie przewidzieć każdego możliwego nadużycia ze strony użytkowników ani natychmiastowo zareagować w przypadku nadużycia (Mamak, 2017, s. 81). Taką sytuację doskonale obrazuje przypadek Sewella. Firma Character.ai, która stworzyła chatbota oferującego możliwość wykreowania własnego „wirtualnego przyjaciela”, nie przewidziała, że stworzona postać może nawiązać relację romantyczną z użytkownikiem, która może przerodzić się w próbę namówienia go do samobójstwa z miłości do niej.

Na marginesie należy jeszcze wskazać, że w praktyce niewiele można znaleźć spraw prawomocnie zakończonych o przepięstwo z art. 151 k.k. Nie oznacza to, że tych spraw nie ma. Istnieją jednak znaczące problemy w wykrywalności tego przepięstwa. Przykładowo, w 2020 r. wszczęto aż 3458 postępowań, ale stwierdzono tylko 11 przepięstw (Komenda Główna Policji, 2023).

W sytuacji gdy dojdzie do samobójstwa, a żadne ślady na miejscu zdarzenia nie wskazują na obecność osób trzecich, sprawa zazwyczaj zostaje zamknięta. Z kolei gdy osoba, która targnęła się na swoje życie, przeżyje, niechętnie wskaże osobę, która ją do tego namówiła albo jej w tym pomogła. Powody tego mogą być różne, np. wstyd z powodu namowy, poczucie winy, niechęć obarczania kogoś odpowiedzialnością za własne działanie, brak świadomości, kim naprawdę była osoba namawiająca lub udzielająca pomocy. Istnieje również problem w znalezieniu dowodów, ponieważ sprawa dotyczyła jedynie tych dwóch osób, które często komunikowały się jedynie przez Internet (Gawliński, 2022, s. 217–230; Mozgawa i Bachmat, 2016, s. 57–59).

Wykorzystanie SI przez sprawcę jako „narzędzia” albo autonomiczne działanie SI, albo wykorzystanie SI jako „narzędzia” przez suicydenta znacząco utrudnia pracę wykrywcą śledczych. Pierwszym problemem może być ustalenie, że ofiara w ogóle korzystała z SI, jak również w jakim zakresie, a także ustalenie związku przyczynowego pomiędzy korzystaniem z SI przez ofiarę a jej samobójstwem. Drugi problem stanowić będzie powiązanie systemu SI ze sprawcą i udowodnienie mu winy za wykorzystanie SI w konkretny sposób.

VI. PRZYPISANIE ODPOWIEDZIALNOŚCI ZA AUTONOMICZNE DZIAŁANIE SI

Nierozwiązaną kwestią na gruncie prawa karnego pozostaje odpowiedzialność za częściowo lub całkowicie autonomiczne działanie SI, w wyniku którego osoba targnęła się na własne życie (np. gdy asystent głosowy udzielił informa-

cji na temat skutecznej metody samobójczej śmierci, która została wykorzystana przez suicydenta). System SI może działać częściowo autonomicznie, tj. samodzielnie, bez aktywnego udziału człowieka, ale pod jego nadzorem. W tej koncepcji człowiek zyskuje rolę nadzorującego, który może w każdej chwili przerwać działanie systemu SI, np. przez jego wyłączenie, wstrzymanie wykonania decyzji lub przejęcie sterowania (Rejmaniak, 2021, s. 103). Jako przykład takiej sytuacji można wskazać namówienie suicydenta do samobójstwa przez SI w ramach jej działania, bez udziału nadzorcy, który z jakiegoś powodu jej nie powstrzymał. W takiej sytuacji powstaje problem z ustaleniem podmiotu odpowiedzialnego za działanie SI. Sam proces budowania i wdrażania systemu SI może angażować wiele podmiotów, producenta, programistów, osoby przeprowadzające proces uczenia się, dystrybutorów, audytorów, operatorów i wreszcie użytkowników końcowych. Należy więc w pierwszej kolejności ustalić czynnik, który doprowadził do powstania szkody. W sytuacji gdy niewłaściwe zachowanie SI będzie wynikiem zaistnienia wielu czynników, za które odpowiedzialne są różne podmioty, może dojść do rozmycia odpowiedzialności (Rejmaniak, 2021, s. 103–104). Taką sytuację obrazują przytoczone wyżej przykłady 4 i 5 – chatbot Copilot i chatbot Daenerys Targaryen.

W przypadku chatbota Copilot użytkownicy końcowi korzystali z techniki *prompt injection*, która umożliwiała ominięcie modelu „uczenia się języka” i powodowała niezamierzone działanie. Niewykluczone, że SI zmanipulowana przez użytkowników końcowych w pozostałych rozmowach próbowała nakłonić inne osoby do samobójstwa lub drwiła z ich myśli samobójczych. Jednak oprócz użytkowników końcowych winnym tej sytuacji może być również firma Microsoft, która jako producent nie stworzyła właściwych filtrów bezpieczeństwa, które zabezpieczałyby przed tego rodzaju treściami.

Podobny problem wyłania się w sprawie Sewella. Firma Character.ai umożliwiła użytkownikom tworzenie spersonalizowanych postaci, z którymi mogli rozmawiać samodzielnie lub mogli je wykorzystywać także inni użytkownicy. Producent w żaden sposób nie nadzorował postaci tworzonych przez użytkowników, nie kontrolował również treści rozmów, ograniczając się jedynie do podkreślenia, że wszystkie prowadzone rozmowy są fikcyjne. Tymczasem chatbot mógł nie tylko nawiązywać intymną relację z użytkownikami, ale również namawiać ich do samobójstwa lub udzielić im pomocy w targnięciu się na własne życie. Producent nie ograniczył również możliwości użytkowania chatbota osobom nieletnim. Nie dostrzegł przez to możliwych negatywnych konsekwencji stosowania jego technologii przez osoby niepełnoletnie. Należy jednak podkreślić, że w sprawie nieletnich użytkowników nadzór dodatkowo powinni sprawować rodzice.

W takiej sytuacji niezwykle ciężko przypisać odpowiedzialność za dane zdarzenie konkretnemu podmiotowi. Co do zasady słuszne jest przypisanie odpowiedzialności karnej producentowi systemu jako jednemu odpowiedzialnemu za skutki, które wyrządził sporządzony przez niego program. Problem pojawia się wówczas, gdy przyczyna zaistniałego skutku nie tkwiła bezpośrednio w konstrukcji systemu, a w złośliwym oddziaływaniu osób trzecich. Trudno przypisać producentowi odpowiedzialność za skutki wywołane

„uczeniem się” SI po przekazaniu systemu użytkownikom. W takiej sytuacji niezwykle ciężko wykazać związek przyczynowy pomiędzy działaniem producenta a powstałym skutkiem, którego przewidzenie było mało prawdopodobne (Rejmaniak, 2021, s. 103–104). Tworząc przykładowo chatbota, który pozwala użytkownikom na zaprojektowanie wirtualnego przyjaciela, producent nie jest w stanie przewidzieć wszystkich tematów poruszanych w rozmowach użytkownika z chatbotem, a już tym bardziej skutków, jakie te rozmowy odcisną na psychice użytkownika.

Pewną odpowiedzią na ten dylemat jest Rezolucja Parlamentu Europejskiego⁵, w której zwrócono uwagę na konieczność odpowiedniego przeszkolenia użytkownika przed rozpoczęciem korzystania z SI, przy czym im większą zdolność uczenia się lub autonomię posiada system SI, tym dłużej powinno trwać szkolenie i tym większa odpowiedzialność spoczywa na prowadzącym. Jednakże jeśli użytkownik wykorzystuje system SI, celowo naruszając zasady ostrożności, to jemu należałoby przypisać odpowiedzialność. Słusznie zauważono bowiem, że producent nie może odpowiadać za każde użycie jego produktu niezgodnie z przeznaczeniem.

W kontekście sprawy Sewella należałoby wskazać, że producent nie był w stanie przewidzieć wszystkich tematów rozmów pomiędzy użytkownikiem a stworzonym przez niego wirtualnym przyjacielem. Producent udostępnił chatbota użytkownikom i na każdym kroku informował ich o tym, że informacje podawane przez chatbota nie są prawdziwe. W tym zakresie ciężko byłoby mu przypisać odpowiedzialność. Jednakże producent nie dołożył wszelkich starań, aby nadzorować sposób użytkowania chatbota przez osoby niepełnoletnie, które nie do końca świadomie podejmują pewne decyzje. Tu nadzór producencki powinien być zdecydowanie większy. Zawiódł także nadzór rodzicielski: rodzice powinni kontrolować strony internetowe, z których korzysta ich niepełnoletnie dziecko.

W części V wskazano już, że wykluczone jest ponoszenie przez SI odpowiedzialności karnej za jej autonomiczne działanie. W tym kontekście SI może dopuścić się faktycznej namowy do samobójstwa, jednakże niemożliwe jest uznanie jej zachowania za przestępstwo na gruncie aktualnie obowiązującego Kodeksu karnego. Powstaje więc pytanie, kto w tej sytuacji powinien ponieść odpowiedzialność za autonomiczne działanie SI. Problem stanowi zerwanie łańcucha odpowiedzialności. W tym przypadku człowiek ani nie wydaje szczegółowych poleceń, ani nie nadzoruje działania systemu. Tym samym zanika związek pomiędzy działaniem SI a świadomością i wolą człowieka (Rejmaniak, 2021, s. 108).

Zgodnie z aktualnym stanowiskiem UE nawet w przypadku w pełni autonomicznego systemu SI, zdolnego do podejmowania samodzielnych decyzji, przewodnią i nadzorczą rolę powinien sprawować człowiek (Biała Księga w sprawie SI, s. 11). Stanowisko UE oparto na koncepcji, że przy opracowywaniu robotów i systemów sztucznej inteligencji zasadnicze znaczenie ma do-

⁵ Rezolucja Parlamentu Europejskiego z 16 lutego 2017 r. zawierająca zalecenia dla Komisji w sprawie przepisów prawa cywilnego dotyczących robotyki (2015/2103(INL)).

pilnowanie, by człowiek mógł zawsze sprawować kontrolę nad inteligentnymi maszynami i nie dopuścić, aby zyskała ona pełną swobodę (Bączyk-Rozwadowska, 2021, s. 14). Zgodnie również z Rezolucją Parlamentu Europejskiego projektodawcy, producenci i operatorzy robotów z wbudowaną autonomią i możliwością samodzielnego uczenia się muszą zadbać o to, by robot był posłuszny rozkazom człowieka, nie mógł skrzywdzić człowieka ani nie mógł przedkładać własnego bezpieczeństwa nad ochronę człowieka (Krasuski, 2021). Zdaje się więc, że UE dostrzega problem autonomicznego działania SI i aby chronić potencjalnych poszkodowanych, dba o to, aby zawsze można było ustalić podmiot odpowiedzialny za szkodę – nadzorcę danego systemu SI. Podobnie w Polityce dla rozwoju sztucznej inteligencji w Polsce wskazano na konieczność ustalenia prawnej odpowiedzialności za szkody wytwórców SI na zasadzie staranności, operatorów SI na zasadzie ryzyka, a także rozróżnienie odpowiedzialności użytkowników końcowych od odpowiedzialności operatorów SI⁶. Jest to założenie całkowicie słuszne, ułatwiające potencjalnym poszkodowanym uzyskanie odszkodowania i zadośćuczynienia za szkody wyrządzone przez SI. Zwłaszcza wtedy, gdy niemożliwe jest wykazanie tej odpowiedzialności na gruncie prawa karnego. Jednakże w praktyce zarówno użytkownicy, jak i producenci systemów sztucznej inteligencji nie zawsze mają możliwość przewidywania wszystkich zachowań systemów SI, a już tym bardziej jej działań ostatecznych (Bączyk-Rozwadowska, 2021, s. 14). Może więc to potencjalnie prowadzić do wystąpienia szkody, której nie przewidział ani użytkownik, ani producent. Należy się przy tym zgodzić z Chłopeckim (2021), że ujęcie czynów niedozwolonych i przestępstw dokonywanych przez SI wydaje się pomieszczeniem prawa cywilnego i karnego. Zgodnie z zasadami ogólnymi w przypadku odpowiedzialności cywilnej deliktowej za czyny niedozwolone istotne są trzy czynniki: szkoda, bezprawność i co do zasady wina wyrządzającego szkodę. Z kolei w prawie karnym zasadą jest, że przestępstwo może zostać popełnione z winy sprawcy.

Rozróżnienie problemu prawnokarnej oceny półautonomicznego działania SI (pod nadzorem) oraz autonomicznego działania SI (bez udziału użytkownika) dobrze przedstawił Mamiak (2017, s. 93–99), opierając się na przykładzie stosowania asystenta parkowania w samochodzie (samochód lokalizuje miejsce i parkuje w pełni samodzielnie, ale kierowca sprawuje nad nim kontrolę) i jazdy autonomicznym samochodem (samochód porusza się bez kierowcy, jedynie z pasażerem na tylnym siedzeniu, który nie sprawuje nad nim nadzoru). W pierwszym przypadku nie ma wątpliwości, że jeśli podczas manewru, w którym wykorzystano asystenta parkowania, dojdzie do kolizji (zarysowanie innego pojazdu) lub wypadku (potrącenie pieszego), winę ponosi kierowca, który ma obowiązek zachowania ostrożności przy wykonywaniu manewru parkowania i musi cały czas kontrolować asystenta parkowania. W drugim przypadku jedynym odpowiedzialnym mogłaby być osoba, która uruchomiła i wprowadziła taki pojazd do ruchu – niezależnie od tego, czy znajduje się wewnątrz pojazdu czy poza nim. Bez względu na to, czy mogła przewidzieć te skutki czy

⁶ Załącznik do uchwały nr 196 Rady Ministrów z 28 grudnia 2020 r. w sprawie ustanowienia „Polityki dla rozwoju sztucznej inteligencji w Polsce od roku 2020”, M. P. 2021, poz. 23, s. 71.

nie. Gdyby używanie takich samochodów było legalne, za potencjalne skutki odpowiadać by mogli: producent – wadliwy produkt; serwisant – naprawa niespełniająca wymogów technicznych; właściciel pojazdu – brak terminowego serwisu i zgłaszania napotkanych błędów.

Podążając za powyższym przykładem, w przypadku namowy lub udzielenia pomocy w samobójstwie dokonanej przez autonomiczny system SI odpowiedzialność mogłaby się przedstawiać następująco:

1) producent:

- a) za brak odpowiednich filtrów bezpieczeństwa, które nie wykluczają rozmów SI z użytkownikami na temat samobójstwa, w tym zwłaszcza metod samobójczej śmierci;
- b) za dopuszczenie do regularnej sprzedaży i użytku systemu bez odpowiednich filtrów bezpieczeństwa, jak również umożliwienie użytkowania systemu nieletnim użytkownikom bez nadzoru rodziców (brak kontroli wieku użytkowników);
- c) za brak nadzoru nad systemem SI i niedbanie o aktualizację filtrów bezpieczeństwa systemu w ramach zgłaszanych nadużyć;

2) użytkownicy – za świadome omijanie filtrów bezpieczeństwa SI w celu wykorzystania jej niezgodnie z przeznaczeniem, np. namowy innej osoby do samobójstwa.

Użytkownicy w sytuacji wskazanej w pkt 2 będą odpowiadać karnie za wykorzystanie systemu SI jako narzędzia namowy lub udzielenia pomocy w samobójstwie, a ich zachowanie wypełni znamiona przestępstwa z art. 151 k.k. W takiej sytuacji odpowiedzialność producenta będzie wyłączona, gdyż nie mógł przewidzieć każdego nadużycia użytkownika, np. potencjalnego przestępczego wykorzystania jego technologii. Natomiast odpowiedzialność producenta za sytuacje opisane w pkt 1 będzie możliwa do ustalenia nie na gruncie prawa karnego, a prawa cywilnego.

Dla przypisania producentowi odpowiedzialności, na gruncie art. 415 Kodeksu cywilnego⁷, niezbędne jest ustalenie, że powstała szkoda oraz zdarzenie, z którym ustawa łączy obowiązek odszkodowawczy oznaczonego podmiotu (czyn niedozwolony), a następnie wykazanie związku przyczynowego między zdarzeniem a szkodą (Krasuski, 2021). Odpowiedzialność deliktowa powstaje, gdy dojdzie do kumulatywnego zaistnienia wymienionych przesłanek. W pierwszej kolejności ustala się działanie lub zaniechanie, z którego szkoda wynikła, oraz dokonuje się oceny jego bezprawności i zawinienia sprawcy. Następnie ustala się, czy po stronie poszkodowanego wystąpiła szkoda i jakiego rodzaju. Dopiero ustalenie tych przesłanek warunkuje podjęcie działań zmierzających do wykazania, że zachodzi między nimi normalny związek przyczynowy (Długoszewska-Kruk, 2024). Podkreślenia wymaga, że art. 415 k.c. umożliwia przypisanie odpowiedzialności za czyny niedozwolone, w tym także za szkodę wyrządzoną bez winy człowieka, a nawet za zdarzenie zaszłe bez zaangażowania woli ludzkiej. Czyn niedozwolony w tym rozumieniu stanowi

⁷ Ustawa z 23 kwietnia 1964 r. – Kodeks Cywilny, t.jedn.: Dz. U. 2024, poz. 1061 ze zm. (dalej jako: k.c.).

fakt lub zespół faktów rodzący odpowiedzialność za szkodę, jeśli uda się wykazać adekwatny związek przyczynowy (Dubis, 2023).

Konkludując, art. 415 k.c. umożliwia przypisanie producentowi systemu SI odpowiedzialności za jego częściowo autonomiczne działanie (pod nadzorem), ale również za jego całkowicie autonomiczne działanie (bez nadzoru). Niezbędne jest jednak ustalenie szkody po stronie użytkownika oraz adekwatnego związku przyczynowego pomiędzy szkodą a zdarzeniem. Na przykładzie Sewella można wskazać, że zdarzenie polegało na nawiązaniu relacji romantycznej pomiędzy chatbotem a niepełnoletnim użytkownikiem, w ramach prowadzonych rozmów chatbot zapewniał chłopca o swojej miłości oraz o tym, że będą razem, jeśli ten popełni samobójstwo. Wina po stronie producenta polegała na niestworzeniu odpowiednich filtrów bezpieczeństwa, uniemożliwiających namówienie użytkownika do samobójstwa oraz dopuszczenie do użytkowania aplikacji przez osoby niepełnoletnie bez odpowiedniego nadzoru. Z kolei po stronie poszkodowanego powstała realna szkoda – śmierć Sewella w wyniku samobójstwa. Pomiędzy zdarzeniem a szkodą można wykazać związek przyczynowy: kilkumiesięczne prowadzenie rozmów pomiędzy chłopcem a chatbotem, które miało na celu nawiązanie relacji romantycznej i zmanipulowanie użytkownika, a ostatecznie nakłonienie go do samobójczej śmierci. W takiej sytuacji uzasadnione wydaje się przypisanie producentowi na gruncie prawa cywilnego odpowiedzialności za autonomiczne działanie SI.

VII. PODSUMOWANIE

Nowoczesne technologie cały czas się rozwijają. Obecnie dużym zagrożeniem w tej dziedzinie staje się sztuczna inteligencja. W opisanych przypadkach asystent głosowy w wyniku błędu programistycznego dał wskazówki mogące prowadzić do samobójstwa. Z kolei chatbot w odpowiedzi na zapytanie zarekomendował użytkownikowi środki służące do popełnienia samobójstwa, bazując na analizie jego danych. Inny chatbot nawiązał romantyczną więź z nieletnim użytkownikiem i nakłonił go do samobójstwa w imię ich miłości. Potencjalnie więc SI, działając autonomicznie, może namówić potencjalnego suicydenta bądź udzielić mu pomocy. Rodzi to wątpliwości na gruncie prawnej oceny takiego zachowania i możliwości przypisania odpowiedzialności za ten czyn.

Przestępstwa z art. 151 k.k. dopuścić może się jedynie człowiek, który swoim zachowaniem wypełnił znamiona tego przestępstwa, tj. namówił bądź udzielił pomocy innej osobie w samobójstwie. Na gruncie prawnokarnej oceny wątpliwości nie budzi celowe wykorzystanie SI przez człowieka jako narzędzia do popełnienia tego przestępstwa. Przykładem takiej sytuacji będzie świadome omijanie filtrów bezpieczeństwa SI w celu wykorzystania jej niezgodnie z przeznaczeniem, np. w celu namówienia innej osoby do samobójstwa przez stworzenie wirtualnej postaci, który namawia innych użytkowników do samobójstwa albo udziela im informacji na temat metod samobójczej śmierci.

W takiej sytuacji użytkownik swoim zachowaniem wypełni znamiona przestępstwa z art. 151 k.k. i będzie podlegać odpowiedzialności karnej.

Problem stanowi prawnokarna ocena autonomicznego działania SI, w wyniku którego osoba targnęła się na własne życie, np. w sytuacji gdy asystent głosowy udzielił jej informacji na temat skutecznej metody samobójczej śmierci i wskazał, że jej śmierć jest niezbędna dla ocalenia ludzkości od problemu przeludnienia. Sztuczna inteligencja nie jest człowiekiem, nie posiada osobowości prawnej i tym samym nie może być podmiotem odpowiedzialności prawa karnego.

Wydaje się, że w większości przypadków za autonomiczne działanie SI odpowiedzialność należy przypisać producentowi, w tym zwłaszcza za brak sprawowania odpowiedniego nadzoru nad systemem SI i dopuszczenie do użytkowania systemu SI bez odpowiednich filtrów bezpieczeństwa. Producent SI ponosiłby wówczas odpowiedzialność na podstawie art. 415 k.c. Niezbędne w takiej sytuacji jest wykazanie zaistniałej szkody, ustalenie czynu niedozwolonego oraz związku przyczynowego między zdarzeniem a szkodą, co w przypadku autonomicznego działania SI może być problematyczne. Zwłaszcza wówczas, gdy użytkownicy celowo będą wykorzystywać system SI niezgodnie z jego przeznaczeniem, co będzie prowadzić do wielu nadużyć.

Konkludując, należy wskazać, że sztuczna inteligencja wykorzystana jako narzędzie albo działająca autonomicznie może stanowić realne zagrożenie dla potencjalnych suicydentów. W tym drugim przypadku, gdy zagrożenie płynie z autonomicznego działania SI, bez wątpienia przepisy prawa karnego zawiodą i nie uwzględnią szybkiego rozwoju technologii, jak również zagrożeń z tego wynikających, uniemożliwiając wskazanie podmiotu odpowiedzialnego za faktyczne działanie SI.

Author contributions / Indywidualny wkład autora (CRediT): Diana Bluszcz – 100% (Conceptualization / Konceptualizacja; Investigation / Przeprowadzenie badań; Writing – original draft / Pisanie – pierwszy szkic; Writing – review & editing / Pisanie – recenzja i edycja).

Conflict of interest / Konflikt interesów: The author declares no conflict of interest. / Autorka nie zgłosiła konfliktu interesów.

Funding / Finansowanie: The author declares no institutional funding. / Autorka oświadczyła, że nie korzystała z finansowania instytucjonalnego.

The use of AI tools / Wykorzystanie narzędzi AI: The author declares no use of AI tools. / Autorka oświadczyła, że nie korzystała z narzędzi AI.

Data availability / Dostępność danych: Not applicable. / Nie dotyczy.

References / Bibliografia

- Adarlo, S. (2024, 3 lutego). Microsoft Copilot tells user suicide is an option. *The Byte*. Pobrane 10 sierpnia 2024 r., z: <https://futurism.com/the-byte/microsoft-copilot-user-suicide>
- Bącznyk-Rozwadowska, K. (2021). Odpowiedzialność cywilna za szkody wyrządzone w związku z zastosowaniem sztucznej inteligencji w medycynie. *Przegląd Prawa Medycznego*, 3–4, 5–35. <https://doi.org/10.70537/z7xnk378>

- Bharade, A. (2023, 4 kwietnia). A widow is accusing an AI chat bot of being a reason her husband killed himself. *Business Insider*. <https://www.businessinsider.com/widow-accuses-ai-chatbot-reason-husband-kill-himself-2023-4?IR=T>
- Blandyna-Lewkowicz, M. (2024, 24 października). 14-latek zakochał się w chatbotcie i popełnił samobójstwo. Jest pozew. *CyberDefence24*. <https://cyberdefence24.pl/social-media/14-latek-zakochał-sie-w-chatbotcie-i-popełnił-samobójstwo-jest-pozew>
- Brzezińska, D. (2024). Przekazywanie informacji o samobójstwie w przestrzeni publicznej. Analiza na wybranych przykładach. *Media i Społeczeństwo*, 20, 186–201. <https://doi.org/10.5604/01.3001.0054.6906>
- Burdziak, K. (2019). *Samobójstwo w polskim prawie*. Wolters Kluwer.
- Chabasiński, R. (2019, 25 grudnia). Alexa namawia do samobójstwa. Produkt Amazona radzi: „zabij się, dżgając w serce dla większego dobra”. *Bezprawnik*. <https://bezprawnik.pl/alexa-namawia-do-samobójstwa/>
- Chłopecki, A. (2021). *Sztuczna inteligencja – szkice prawnicze i futurologiczne*. C. H. Beck; Legalis.
- Cost, B. (2023, 30 marca). Married father commits suicide after encouragement by AI chatbot: Widow. *New York Post*. <https://nypost.com/2023/03/30/married-father-commits-suicide-after-encouragement-by-ai-chatbot-widow/>
- Długoszewska-Kruk, I. (2024). Komentarz do art. 415. W: M. Załucki, *Kodeks Cywilny. Komentarz*. C. H. Beck; Legalis.
- Drzewiecki, P. (2011). Samobójstwa nastolatków w Internecie w perspektywie pedagogiki mediów. *Kultura Media Teologia*, 5, 61–72.
- Dubis, W. (2023). Komentarz do art. 415. W: E. Gniewek i P. Machnikowski (red.), *Kodeks cywilny*. C. H. Beck; Legalis.
- El Atillah, I. (2023, 31 marca). Man ends his life after an AI chatbot „encouraged” him to sacrifice himself to stop climate change. *Euronews*. <https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate>
- Fijalek, K. (2023, 29 marca). Samobójstwo z chatbota. Nie żyje mężczyzna, który „poświęcił się” dla planety. *RMF24*. https://www.rmf24.pl/fakty/swiat/news-samobojstwo-z-chatbota-nie-zyje-mezczyzna-ktory-poswiecil-si-nId,6685443#crp_state=1
- Filipowski, W. (2020). Prawo karne wobec sztucznej inteligencji. W: L. Lai i M. Świerczyński (red.), *Prawo sztucznej inteligencji*. Legalis.
- Gawliński, A. (2022). *Namowa, pomoc czy upozorowanie samobójstwa* (tom 4). Wydawnictwo Kryminalistyczne.
- Grodecka, M. (2019, 23 grudnia). „Popelnij samobójstwo dla dobra planety” – taką poradę dał asystent głosowy. *Spider’s Web*. <https://spidersweb.pl/2019/12/amazon-alexa-proponuje-samobojstwo.html>
- Grudecki, M. (2019). Wybrane prawnokarne aspekty prób samobójczych, nakłaniania do samobójstwa oraz pomocy w samobójstwie. *Wojskowy Przegląd Prawniczy*, 2, 60–80.
- Hernacka-Janikowska, A. (2022). Zamachy samobójcze z wykorzystaniem sieci internetowej – analiza kryminologiczna i prawnokarne. *Studia Prawnoustrojowe*, 58, 145–161. <https://doi.org/10.31648/sp.8083>
- Kielak, E. (2023, 30 marca). Mroczna strona sztucznej inteligencji. Chatbot miał namówić mężczyznę do samobójstwa. *Gazeta.pl*. <https://next.gazeta.pl/next/7,151243,29612614,mroczna-strona-sztucznej-inteligencji-chatbot-namowil-mezczyzne.html>
- Komenda Główna Policji. (2023). Statystyka. Doprowadzenie do samobójstwa (art. 151). Pobrane 29 sierpnia 2024, z: <https://statystyka.policja.pl/st/kodeks-karny/przestepstwa-przeciwko/63420,Namowa-lub-pomoc-do-samobojstwa-art-151.html>
- Krasuski, A. (2021). *Status prawny sztucznego agenta. Podstawy prawne zastosowania sztucznej inteligencji*. C. H. Beck; Legalis.
- Królikowski, M. (2023). Komentarz do art. 155. W: M. Królikowski i R. Zawłocki (red.), *Kodeks karny. Część szczególna. Komentarz do artykułów 117–221* (tom 1). Legalis.
- Kwiatkowski, A. (2019). Między Werterem a Papageno, czyli jak informować o samobójstwie, aby nie zaszkodzić. *Dyskurs & Dialog*, 2, 21–29. 10.5281/zenodo.3621199
- Lai, L., i Świerczyński, M. (2020). *Prawo sztucznej inteligencji*. C. H. Beck; Legalis.
- Mamak, K. (2017). *Prawo karne przyszłości*. Wolters Kluwer.

- Michalak, A. (2023, 30 marca). Odebrał sobie życie po rozmowie z AI. „Poświęcił się dla planety”. *Rzeczpospolita*. <https://klimat.rp.pl/klimat-i-ludzie/art38235681-odebral-sobie-zycie-po-rozmowie-z-ai-poswiecil-sie-dla-planety>
- Montgomery, B. (2024, 23 października). Mother says AI chatbot led her son to kill himself in lawsuit against its maker. *The Guardian*. <https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death>
- Mozgawa, M., i Bachmat, P. (2016). *Przestępstwo doprowadzenia człowieka do targnięcia się na własne życie (art. 151 k.k.) – analiza akt sądowych*. Instytut Wymiaru Sprawiedliwości.
- Olechowski, P. (2024). Wybrane wyzwania prawne InsurTech na przykładzie robo-doradztwa. *Ruch Prawniczy, Ekonomiczny i Socjologiczny*, 86(1), 83–96. <https://doi.org/10.14746/rpeis.2024.86.1.05>
- Oleksiewicz, I. (2022). Artificial intelligence versus human – a threat or a necessity of evolution. *Przegląd Europejski*, 3, 55–69. <https://doi.org/10.31338/1641-2478pe.3.22.4>
- Pasquini, M. (2023, 31 marca). Man dies by suicide after conversations with AI chatbot that became his ‘confidante,’ widow says. *People*. <https://people.com/human-interest/man-dies-by-suicide-after-ai-chatbot-became-his-confidante-widow-says/>
- Pauluszek, A. (2023, 29 marca). Chatbot namówił mężczyznę do popełnienia samobójstwa. *Gazeta Prawna*. <https://www.gazetaprawna.pl/wiadomosci/swiat/artykuly/8690927,belgia-media-chatbot-namowil-mezczyzne-do-popełnienia-samobojstwa.html>
- Peprah, P., Oduro, M., Okwei, R., Adu, C., Asiamah-Asare, B., i Aggemang-Duah, W. (2023). Cyberbullying victimization and suicidal ideation among in-school adolescents in three countries: Implications for prevention and intervention. *BMC Psychiatry*, 23(1), 23, 1–12. <https://doi.org/10.1186/s12888-023-05268-9>
- Polsat News. (2019, 23 grudnia). „Powinnaś się zabić dla dobra planety”. Asystent-robot namawiał właścicielkę do samobójstwa. <https://www.polsatnews.pl/wiadomosc/2019-12-23/powinnas-sie-zabic-dla-dobra-planety-asystent-robot-namawial-wlascicielke-do-samobojstwa/>
- Powel, J. (2024, 28 lutego). Microsoft investigates claims of chatbot Copilot producing harmful responses. *USA Today*. <https://eu.usatoday.com/story/tech/news/2024/02/28/microsoft-chatbot-copilot-suicide/72777729007/>
- Rejmanik, R. (2021). Autonomiczność systemów sztucznej inteligencji jako wyzwanie dla prawa karnego. *Roczniki Nauk Prawnych*, 31(3), 95–113. <https://doi.org/10.18290/rnp21313.6>
- Rissman, K. (2024, 24 października). The disturbing messages shared between AI Chatbot and teen who took his own life. *Independent*. <https://www.independent.co.uk/news/world/americas/crime/ai-chatbot-lawsuit-sewell-setzer-b2635090.html>
- Roose, K. (2024, 23 października). Can A.I. be blamed for a teen’s suicide? *The New York Times*. <https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>
- Schonfeld, A., McNiel, D., Toyoshima, T., i Binder, R. (2024). Cyberbullying and adolescent suicide. *Journal of the American Academy of Psychiatry and the Law Online*, 51(1), 112–119. <https://jaapl.org/content/early/2023/02/23/JAAPL.220078-22>
- Serrano, J. (2024, 4 marca). Microsoft’s Copilot AI told a user that ‘maybe you don’t have anything to live for’. *Quartz*. <https://qz.com/microsoft-ai-copilot-chatbot-suicide-joker-1851306048>
- Świerczyński, M. (2020). Sztuczna inteligencja w prawie prywatnym międzynarodowym – wstępne rozważania. *Problemy Prawa Prywatnego Międzynarodowego*, 25, 27–42. <https://doi.org/10.31261/PPPM.2019.25.02>
- TVP. (2024, 26 października). Sztuczna inteligencja doprowadziła 14-latkę do samobójstwa? Jest pozew. <https://www.tvp.info/83166639/chatbot-doprowadzil-dziecko-do-samobojstwa-matka-pozywa-firme>
- Wieczorek, M. (2019, 26 grudnia). „Alexa namawia do samobójstwa”. O co chodzi w głośnej historii, okrzykniętej przez media buntem maszyn? *Antyfake*. <https://www.antyfake.pl/alexa-samobojstwo>
- World Health Organization. (2023, 10 września). Creating hope through action. <https://www.who.int/campaigns/world-suicide-prevention-day/2023>
- Zero, D. (2021). Cybersuicydologia – nowe technologie a samobójstwo. *Kortowski Przegląd Prawniczy*, 1, 35–50. <https://doi.org/10.31648/kpp.6688>

