

Toruń, dnia 3 lutego 2026r.

Dr hab. Jerzy Lachowski, prof. ucz.
Katedra Prawa Karnego UMK
Ul. Władysława Bojarskiego 3
87-100 Toruń

**Recenzja rozprawy doktorskiej Pani mgr Gabrieli Garcii Very pt.
Prawnokarne aspekty sztucznej inteligencji, napisanej pod kierunkiem
naukowym Pani dr hab. Anny Golonki, prof. UR**

Pani mgr Gabriela Garcia Vera przedstawiła do oceny rozprawę doktorską dotyczącą prawnokarnych aspektów sztucznej inteligencji. Nie ulega wątpliwości, że prawne aspekty AI to duże wyzwanie dla współczesnych badaczy prawa. Wynika to z dynamicznego rozwoju AI, jej różnych odmian, poziomów samodzielności, a także ze skąpego materiału normatywnego zwłaszcza w obszarze prawa karnego. Co prawda na szczeblu europejskim można odnotować pewne regulacje, których celem jest określenie standardu bezpieczeństwa w zakresie wykorzystywania AI (mam tutaj na myśli Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 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), a w prawie krajowym – choćby w ustawie- Prawo o ruchu drogowym znajdujemy dość lakoniczną definicję pojazdu autonomicznego, to polskiemu porządkowi prawnemu daleko jeszcze do kompleksowego uregulowania odpowiedzialności karnej za skutki działania sztucznej inteligencji. Zagadnienie jest wyjątkowo aktualne, wszak można odnotować w mediach informacje na temat tego, że robot zabił człowieka, czy też czat GPT doprowadził nastolatka do samobójstwa. W takich przypadkach fundamentalną kwestią jest wskazanie podmiotu odpowiedzialnego za takie skutki zwłaszcza w sytuacji, gdy były one całkowicie nieprzewidywalne dla producenta czy sprzedawcy takiego systemu AI. Funkcja

ochronna prawa karnego domaga się odpowiedzi na takie pytanie. O wadze tego zagadnienia świadczy również to, że cały XXI Międzynarodowy Kongres Prawa Karnego zorganizowany przez Międzynarodowe Stowarzyszenie Prawa Karnego (AIDP) w 2024r. został poświęcony problematyce sztucznej inteligencji. Pokłosiem tego wydarzenia są niezwykle interesujące materiały naukowe, które poruszają te kwestie, choć przyznać trzeba- że w dość ogólny sposób. Z tych powodów wybór tematu i jego sformułowanie przez Doktorantkę zasługuje na najwyższą ocenę i świadczy o dużych ambicjach, odwadze oraz pasji badawczej.

Rozprawa doktorska składa się z pięciu rozdziałów poprzedzonych wykazem skrótów i wstępem, które zamykają: zakończenie, wykaz literatury przedmiotu, aktów prawnych, orzecznictwa, netografia, wykaz rysunków i tabel.

Rozdział I został poświęcony powstaniu i znaczeniu sztucznej inteligencji. Autorka wyciągnęła niejako przed nawias zagadnienia podstawowe związane z etapami kształtowania się AI, obszarami jej zastosowania, kwestiami terminologicznymi.

Rozdział II dotyczy zagadnienia podmiotu czynu zabronionego, w tym roli człowieka w procesie decyzyjnym związanym z działaniem sztucznej inteligencji. Tutaj omówiono również zagadnienie podmiotu czynu zabronionego popełnionego z wykorzystaniem systemów AI i wreszcie problem o charakterze fundamentalnym, a mianowicie podmiotowość prawnokarną samej sztucznej inteligencji.

Rozdział III poświęcony został wybranym zagadnieniom prawnym z obszaru prawa karnego w stosunku do sztucznej inteligencji. W tym miejscu szczegółowym rozważaniom poddano kwestię przypisania skutku w związku z działaniem AI, zagadnienie realizacji znamion podmiotowych czynu zabronionego przez AI, form zjawiskowych popełnienia czynu zabronionego, a także problem przypisywalności winy w relacji do sztucznej inteligencji.

W rozdziale IV omówiono problematykę związaną z odpowiedzialnością karną za czyny zabronione popełnione z wykorzystaniem sztucznej inteligencji w różnych obszarach życia społecznego i gospodarczego. Autorka skupiła swoją uwagę tutaj na odpowiedzialności karnej związanej z użyciem pojazdu

autonomicznego, wykorzystywaniu AI w medycynie, wojsku, omówiła problematykę *deepfake'ów*.

Ostatni rozdział merytoryczny poświęcony został wyzwaniom, które stoją przed ustawodawcą krajowym, międzynarodowym, w tym unijnym w zakresie wykorzystywania AI.

Taki układ rozprawy doktorskiej nie budzi najmniejszych zastrzeżeń. Jest on logiczny, konsekwentny, dobrze przemyślany. Świadczy o tym, że Autorka świetnie zna się na problematyce, którą zbadała w rozprawie doktorskiej, co czyni z Niej jedną z nielicznych w kraju specjalistek z tego zakresu.

W rozprawie posłużono się kilkoma metodami badawczymi tj. dogmatyczno-prawną, prawno-porównawczą, historycznoprawną. Doktoranta czyni to w sposób swobodny, prawidłowy. Korzystanie z tych metod zawsze służy określonemu, z góry założonemu celowi. Trudno w tym zakresie formułować pod adresem Doktorantki jakiegokolwiek zastrzeżenia.

Autorka rozprawy używa w niej klarownego, poprawnego języka. Prace bardzo dobrze się czyta. O zagadnieniach trudnych potrafi pisać w sposób bardzo przystępny dla czytelnika, co świadczy o dobrym warsztacie naukowym. Przypisy, gdy idzie o ich wartość merytoryczną, nie budzą najmniejszych zastrzeżeń. Doktorantka robi je tam, gdzie wymaga tego logika wywodu, dzięki czemu nie mają one charakteru jedynie ornamentu, którego celem jest zrobienie wrażenia na czytelniku lecz stanowią gruntowną podbudowę dla prowadzonych rozważań. Na szczególne wyróżnienie zasługuje ilość wykorzystanej faktycznie literatury przedmiotu. W wykazie na s. 378-425 znalazła się imponująca ilość źródeł w liczbie prawie 700 (sic!). Jest to nie tylko literatura polska, ale również obcojęzyczna, którą Doktorantka twórczo wykorzystuje w trakcie swoich rozważań. Błędy literowe i stylistyczne pojawiają się bardzo rzadko.

Przechodząc do merytorycznej oceny analizowanej rozprawy należy wskazać, że ambicją Autorki nie było poruszenie wszystkich zagadnień związanych z kwestią odpowiedzialności karnej za działania sztucznej inteligencji. Wybrała Ona tylko te problemy, które Jej zdaniem mają charakter fundamentalny. Uczyniła to w sposób bardzo uporządkowany: najpierw

skupiła się na podstawach odpowiedzialności karnej, aby następnie przejść do zagadnień nieco bardziej szczegółowych. Takie podejście do tematu niewątpliwie ustrzegło Doktorantkę przed ryzykiem powtórzeń. Wręcz zaskakująca jest teza, że rozwój sztucznej inteligencji sięga aż starożytnej Grecji (s. 24-25 choć prawdą jest, że ówczesne wyobrażenia na ten temat mają niewiele wspólnego z tym, jak AI jest pojmowana obecnie). Doktorantka trafnie akcentuje normatywny wymóg sprawowania nadzoru nad systemami AI przez człowieka, co może prowadzić do tezy, że niedopuszczalne jest wykorzystywanie w życiu społecznym i w otoczeniu gospodarczym takich systemów sztucznej inteligencji, które byłyby zupełnie autonomiczne i funkcjonowałyby w oderwaniu od nadzoru ze strony człowieka. Konieczność nadzoru ze strony człowieka ułatwia niewątpliwie pociągnięcie do odpowiedzialności karnej konkretnej osoby fizycznej w przypadku spowodowania prawnie relewantnego skutku przez działanie systemu AI. Słusznie zauważono w pracy, że jakkolwiek fakt nadzoru człowieka nad działaniem AI może zapobiegać negatywnym skutkom jej działania, to jednak nie jest to bezwzględna gwarancja uniknięcia takich skutków (s. 98). Kontrola taka może mieć niekiedy charakter iluzoryczny. Niezwykle interesujący jest fragment pracy dotyczący podmiotowości prawnokarnej AI (s. 104 i nast.). Autorka zadaje sobie pytanie, czy słusznie w prawie UE wyklucza się możliwość nadania podmiotowości prawnej sztucznej inteligencji. Akcentuje, że znane są w historii prawa przypadki nadawania podmiotowości prawnej zwierzętom (s. 105). Jak podnosi słusznie, co prawda obecnie kładzie się nacisk na nadzór człowieka nad systemami AI, ale z drugiej strony, co wskazano w rozdziale I, daje się zaobserwować tendencję do tworzenia systemów myślących, działających samodzielnie. Trafnie zauważa, że nie można wykluczyć w przyszłości tworzenia inteligentnych modeli, które będą kreowały jeszcze bardziej inteligentne systemy, których konstrukcja będzie kompletnie niezrozumiała dla człowieka. Autorka twierdzi, że skoro zwierzęta mogą mieć podmiotowość prawną, to tym bardziej podmiotowość taką będzie można nadawać AI. Można w tym miejscu dodać, wspierając argumentację Doktorantki, że skoro powstają projekty ustaw o podmiotowości prawnej rzeki Odry, która przecież nie jest konstruktem wyposażonym w jakąkolwiek

inteligencję, to tym bardziej uzasadnione może być nadawanie takiej podmiotowości AI. Autorka jednocześnie dostrzega konieczność istnienia różnych reżimów prawnych w zależności od rodzaju systemu AI (s. 109). Z drugiej strony w rozprawie podkreśla się trafnie, że nawet zaawansowanym systemom AI brakuje „tego czegoś” (choć nie nazywa tego- s. 111), co nie pozwala nadać im pełni praw i obowiązków na równi z człowiekiem. Słusznie Doktorantka diagnozuje, że przy obecnym tempie rozwoju technologii, systemy AI będą coraz bardziej autonomiczne, co będzie uzasadniało jeszcze większą otwartość dla koncepcji podmiotowości prawnej AI. Trafnie Autorka dostrzega, że przypisanie skutku musi uwzględniać dwie płaszczyzny: ontologiczną oraz normatywną (s. 118). Podkreślenie tego będzie miało istotne znaczenie z punktu widzenia odpowiedzi na pytanie, czy skutek może być przypisany AI. Uzasadnione wątpliwości zgłasza Doktorantka w stosunku do koncepcji, zgodnie z którą skutek spowodowany przez działanie AI powinien być z reguły przypisywany jej twórcy (s. 122). Taki pogląd stanowi jednak niedopuszczalne uproszczenie pomijające to, że granicą przypisania skutku jest jego obiektywna przewidywalność. Im bardziej autonomiczny będzie system AI, tym mniej przewidywalne dla twórcy mogą być następstwa jego działania. Słusznie dostrzega, że w przypadku wyznaczenia podmiotu zobowiązanego do nadzoru nad AI, to ten podmiot- jako gwarant w rozumieniu przepisu art. 2kk- może ponosić odpowiedzialność za skutek spowodowany przez zaniechanie nadzoru (s. 125). Zgodnie z zaproponowanym przez Doktorantkę rozumieniem pojęcia świadomości stanowiącej element konstytutywny zamiaru, faktycznie wypada się zgodzić z tezą, że nie można wykluczyć zrealizowania tej przesłanki przez zwłaszcza zaawansowane systemy sztucznej inteligencji (s. 137). Trafnie Autorka podkreśla na s. 144, że system AI może wyrządzać szkody w świecie rzeczywistym i wirtualnym, a w przypadku, gdy zostanie zaprogramowany przez jego twórców na postępowanie wyrządzające takie skutki, wówczas winien być traktowany jako narzędzie w rękach człowieka (s. 144). Zasadniczo pozytywnie oceniam fragmenty rozprawy, które poświęcone zostały problematyce pomocnictwa i podżegania oraz niewykończonym formom sprawstwa. Tutaj Doktoranta odpowiednio szczegółowo przedstawia te konstrukcje prawne, aby następnie odnieść je do sytuacji wykorzystywania

sztucznej inteligencji. Ze szczególnym zainteresowaniem czytałem fragment dotyczący problematyki winy. Tutaj Autorka przekrojowo pokazuje, w jaki sposób Kodeks karny posługuje się tym pojęciem i skupia się na normatywnej koncepcji winy. Niezwykle interesujące są rozważania na temat przedstawionego na s. 186 kazusu z zakresu bezpieczeństwa w komunikacji, w którym to system sztucznej inteligencji, w który wyposażony jest pojazd autonomiczny w istocie ma dokonać wyboru „mniejszego zła”. Zgodzić się wypada z diagnozą zawartą na s. 199-200, że obecne uregulowania prawnokarne nie są w ogóle przygotowane do radzenia sobie z wyzwaniami zaawansowanych systemów AI. Zgadzam się z Doktorantką, że w przypadku pojazdów autonomicznych z tzw. poziomu 4 pociąganie do odpowiedzialności karnej osoby, która może ingerować w jazdę tylko wówczas, gdy pojazd zbliża się do granicy swoich możliwości operacyjnych, byłoby niezasadne (s. 222). Słusznie zauważono, że trudno byłoby pociągać do odpowiedzialności osobę, która nie miała wpływu na poruszający się pojazd i na której nie ciążył obowiązek nadzoru nad ruchem pojazdu (tamże). Podobnie trafne jest inne spostrzeżenie Autorki wyrażone na s. 244, że granicą odpowiedzialności lekarza posługującego się AI w procesie diagnostycznym jest przewidywalność negatywnych skutków działania sztucznej inteligencji. Pociąganie lekarza do odpowiedzialności karnej za nieprzewidywalne następstwa działania takich systemów mogłoby świadczyć o przedmiotowym traktowaniu i naruszać przepis art. 30 Konstytucji RP, który chroni godność osoby ludzkiej. W takim przypadku powstaje jednak pytanie, kto winien ponieść odpowiedzialność za negatywne skutki pozostające w związku przyczynowym z działaniem AI jeśli były one nieprzewidywalne dla człowieka? Zasadniczo prawidłowo zidentyfikowano w pracy te typy czynów zabronionych, których popełnienie może polegać na wykorzystywaniu tzw. *deepfake’ów*. Wskazano tutaj typ przestępstwa z art. 132kk, art. 224akk, 286kk, 190a§2kk, art. 212kk, art. 216kk. Do tej egzemplifikacji należy dodać koniecznie typ z art. 202§4b kk, w którym wyraźnie mówi się o wytworzonym lub przetworzonym wizerunku małoletniego uczestniczącego w czynności seksualnej. Poważnego rozważenia wymaga sformułowany na s. 293 postulat *de lege ferenda* wprowadzenia do kk odrębnego przestępstwa polegającego na tworzeniu i rozpowszechnianiu

treści typu *deepfake*, które mogą wywołać niepokój publiczny. Istotne problemy poruszono w części dotyczącej dzieł tworzonych przez AI i przywłaszczania sobie ich autorstwa. Tutaj rodzi się podstawowe pytanie, czy człowiek, który przypisuje sobie autorstwo dzieła stworzonego przez AI popełnia czyn zabroniony z art. 115 ust. 1 ustawy o Prawie autorskim i prawach pokrewnych. Czy w takim przypadku przywłaszcza sobie autorstwo cudzego utworu, skoro AI nie posiada podmiotowości w sferze prawa, a więc nie mogą jej przysługiwać prawa autorskie? Oczywiście słuszny jest postulat Doktorantki wprowadzenia do projektu ustawy o systemach AI rozdziału dotyczącego odpowiedzialności karnej za działalność przestępczą z tytułu jej funkcjonowania. Ze szczególną uwagą analizowałem proponowane przez Autorkę konkretne rozwiązania prawnokarne dotyczące odpowiedzialności za działania AI, które znajdują się na s. 354-355. Interesująca jest propozycja wyłączenia odpowiedzialności człowieka w przypadku , gdy dojdzie do przerwania związku przyczynowego między działaniami podmiotów odpowiedzialnych za AI a powstaniem szkody w postaci *novus actus interveniens*. Co prawda, Autorka dopuszcza w takim przypadku odpowiedzialność projektanta, osób wdrażających lub dystrybutorów systemów AI ale od razu trzeba zaznaczyć, że skoro taka odpowiedzialność miałaby być oparta na zasadach określonych w kk, to jej granicą będzie znowu przewidywalność skutków przez takie podmioty. Powstaje zatem od razu pytanie, kto miałby w takim przypadku za skutek odpowiedzieć gdyby się okazało, że dla tych podmiotów był on całkowicie nieprzewidywalny?

Lektura pracy rodzi również pewne refleksje o charakterze krytycznym lub polemicznym. I tak, trudno zrozumieć dlaczego na s. 117 Autorka ogranicza swoje rozważania tylko do przestępstw materialnych tj. znamienych skutkiem. Również w przypadku przestępstw formalnych skutek może mieć znaczenie, tyle że nie decyduje on o możliwości przypisania dokonania, lecz stanowi okoliczność związaną z czynem, która w świetle przepisu art. 115§2kk wpływa na stopień społecznej szkodliwości. W części dotyczącej przypisania skutku, przy analizie koncepcji tzw. wielu rąk zabrakło mi głębszej refleksji nad zagadnieniem istotności przyczynienia się do spowodowania skutku. W przypadku, gdy kilka podmiotów zaangażowanych

jest w jego sprowadzenie, konieczne jest ustalenie, który z nich w stopniu istotniejszym przyczynił się do jego sprowadzenia. Co najmniej nieściśle jest stwierdzenie zawarte na s. 129 dotyczące potrzeby „rozważenia strony podmiotowej AI”. Nie chodzi wszak tutaj o stronę podmiotową sztucznej inteligencji lecz o możliwość wyczerpania przez AI znamion strony podmiotowej czynu zabronionego, co przecież jednoznacznie wynika z tytułu podrozdziału na s. 129. Mam poważne wątpliwości, czy możliwe jest przypisanie zamiaru bezpośredniego systemom AI (zob. s. 139). Nie wydaje mi się, aby można było wprost przyłożyć do takich systemów wolę, która charakteryzuje działania człowieka. Wynika to z faktu, że AI działa w oparciu o algorytmy, jej decyzje są co do zasady zautomatyzowane, co trudno uznać za działanie na zasadzie wolności woli, swobody podejmowania decyzji, możliwości powstrzymania się od działania ale też możliwości podjęcia działania. Żałować należy, że Doktorantka nie dokonuje bardziej pogłębionej analizy nieumyślności i możliwości realizacji tego znamienia podmiotowego przez AI. Zabrakło mi tutaj odniesienia się do wszystkich przesłanek z art. 9§2kk w relacji do sztucznej inteligencji i odpowiedzi na pytanie, które z nich mogą być zrealizowane przez AI, a które nie i dlaczego. Szkoda, że Autorka nie udziela pogłębionej odpowiedzi na pytanie, czy AI może być jednym z „podmiotów” w ramach współsprawstwa, zwłaszcza, gdy jej działanie w sposób istotny przyczyniło się do sprowadzenia prawnokarnie relewantnego skutku. Co prawda Doktorantka zdaje się odrzucać teorię istotności roli wyjaśniającą istotę współsprawstwa, ale w ramach tej koncepcji szczególnie interesujące mogłyby być rozważania zmierzające do rozwiązania rzeczonego problemu. W części dotyczącej problematyki winy zabrakło mi nieco analizy ujęcia winy w świetle teorii relacyjnej, funkcjonalnej i dyskursywnej w stosunku do sztucznej inteligencji. Warto było postawić sobie pytanie, czy wina w rozumieniu tych koncepcji da się odnieść do systemu AI, czy też wciąż jedynym podmiotem zawinienia może być człowiek. Na s. 220 Autorka dość arbitralnie stwierdza, że przestępstwo spowodowania wypadku komunikacyjnego należy do kategorii powszechnych przywołując w tym zakresie tylko jeden (sic!) pogląd P. Zakrzewskiego abstrahując tym samym od poglądu przeciwnego, który postrzega to przestępstwo jako indywidualne niewłaściwe, które może być popełnione tylko

przez uczestnika ruchu. Można zadać Doktorantce pytanie, czy osoba nie będąca uczestnikiem ruchu może naruszyć zasady w nim obowiązujące, zgodnie z treścią przepisu art. 177§1kk. Pogląd ten wyrażony został w dość specyficzny sposób albowiem z jednej strony Autorka wskazuje na powszechny charakter tego czynu, z drugiej zaś (w ostatnim zdaniu na s. 220) wymienia tylko uczestników ruchu jako potencjalnych sprawców. Na s. 222 Doktorantka rozważa pojęcie „prowadzenia pojazdu” w przypadku pojazdów autonomicznych. Żałować należy, że nie odnosi swojego wyводу do przestępstwa z art. 178a§1kk. Nasuwa się wszak pytanie, czy osoba będąca w stanie nietrzeźwości znajdująca się w pojeździe autonomicznym z poziomu 3 czy 4 ponosi odpowiedzialność karną za ten czyn, jako prowadzący pojazd w ruchu.

Wskazane wyżej uwagi polemiczne, czy też niekiedy krytyczne w żaden sposób nie podważają wysokiej oceny recenzowanego opracowania. Doktoranta wykazała w nim umiejętność dokonania trafnego wyboru zagadnienia badawczego, prawidłowego formułowania tez badawczych, dostrzegania problemów naukowych, a co najważniejsze- umiejętność ich rozwiązywania. Udowodniła, że posiada ogólną wiedzę z zakresu prawa karnego, zdolność samodzielnego prowadzenia badań naukowych, a przedstawiona przez Nią rozprawa rozwiązuje oryginalnie istotny problem badawczy, przyczyniając się do rozwoju nauki w obszarze problematyki AI w relacji do nauk penalnych. Tym samym nie mam najmniejszych wątpliwości, że praca doktorska spełnia wymagania wynikające z art. 187 ust. 1 i 2 Prawa o szkolnictwie wyższym i wnoszę o dopuszczenie Pani mgr Gabrieli Garcii Very do dalszych etapów postępowania w przedmiocie nadania stopnia naukowego doktora.