

dr hab. Wojciech Filipkowski, prof. UwB
Katedra Prawa Karnego i Kryminologii
Pracownia Kryminalistyki
Wydział Prawa
Uniwersytet w Białymstoku

**Recenzja rozprawy doktorskiej mgr Gabrieli Garcia Vera
pt. „Prawnokarne aspekty sztucznej inteligencji”**

Przedstawioną do oceny rozprawę doktorską mgr Gabrieli Garcia Vera oceniam **negatywnie**.

I. Ocena wyboru tematu rozprawy doktorskiej

Obszarem badawczym rozprawy Doktorantki są prawnokarne aspekty funkcjonowania systemów sztucznej inteligencji (SI), ujmowane w perspektywie materialnego prawa karnego. W szczególności chodzi o badania na gruncie polskiego Kodeksu karnego (z odniesieniem do prawa Unii Europejskiej i wybranych państw poza unijnych), następujących zagadnień:

- ustalenie podmiotu czynu zabronionego w sytuacji wykorzystania systemów SI (człowiek, podmiot zbiorowy, ewentualnie koncepcja podmiotowości prawnej SI);
- przypisanie skutku i wina przy czynach związanych z działaniem systemów SI, w tym formy sprawcze i niesprawcze (sprawstwo, współsprawstwo, podleganie, pomocnictwo) oraz okoliczności wyłaczające winę;
- odpowiedzialność karna za czyny z wykorzystaniem SI w wybranych obszarach życia społecznego i gospodarczego, takich jak autonomiczne pojazdy,

medycyna, autonomiczne systemy uzbrojenia, *deepfake*, tworzenie dóbr niematerialnych (utwory artystyczne) oraz kradzież dóbr wirtualnych;

- identyfikacja luk i wyzwań dla przyszłych regulacji prawnokarnych w związku z dynamicznym rozwojem SI, na tle unijnych ram prawnych (w szczególności unijnego Aktu o sztucznej inteligencji) oraz regulacji i inicjatyw w Niemczech, Singapurze, USA i Chinach.

Z tego też względu należy skonkludować, iż projektowane badania lokują się w dogmatyce prawa karnego materialnego (z zauważalnym komponentem komparatystycznym i historyczno-prawnym). Wybór tego obszaru jako tematu rozprawy doktorskiej jest trafny i bardzo dobrze rokujący, zarówno z punktu widzenia wymogów ustawowych wobec rozprawy doktorskiej w naukach prawnych, jak i realnych potrzeb nauki oraz praktyki wymiaru sprawiedliwości.

W recenzowanej rozprawie Doktorantka trafnie identyfikuje lukę badawczą w postaci braku systemowych, kompleksowych opracowań naukowych w języku polskim dotyczących prawnokarnych aspektów posługiwania się systemami SI, przy dominacji ujęć cywilistycznych i licznych, ale rozproszonych opracowań anglojęzycznych. W związku z tym formułuje także wyraźną tezę badawczą, co pokazuje, że należy ją traktować jako lukę prawną. W swoich badaniach Doktoranta zmierza do przedstawienia nowego spojrzenia na zasad odpowiedzialności karnej z elementem wykorzystania systemów SI. To spełniałoby ustawowo wymagany wzorzec „oryginalnego rozwiązania”.

Jednocześnie wybrany obszar badawczy należy ocenić jako bardzo aktualny i społecznie istotny. Ma on znaczenie dla rozwoju dogmatyki i dla praktyki stosowania prawa ze względu na szybkie nasycanie się naszego życia tego typu systemami – od prostych aplikacji po zabiegi medyczne i systemy uzbrojenia – ale także w kontekście zachowań przestępczych. Z resztą recenzowana rozprawa prezentuje konkretne przykłady: pojazdy autonomiczne, medycynę, autonomiczne systemy uzbrojenia, *deepfake* oraz kradzież dóbr wirtualnych. Co powoduje, że przedstawione zostały zagadnienia odpowiedzialności karnej w realnych scenariuszach, a więc nie ma charakteru wyłącznie teoretycznego.

Biorąc pod uwagę wszystkie powyższe argumenty, wybór tytułowego, złożonego obszaru badawczego należy ocenić jako bardzo trafny, aczkolwiek bardzo rozległy. Mamy do czynienia z wyraźną luką badawczą, wysoką doniosłość teoretyczną i praktyczną, pozwala na oryginalne rozwiązanie problemu naukowego i koresponduje z klasycznym warsztatem dogmatyki prawa karnego.

II. Ocena założeń badawczych

Doktorantka zawarła we Wstępie szerokie spektrum celów badawczych od deskryptywnych, eksplanacyjnych i eksploracyjnych, po preskryptywne a nawet predykcyjne. Dzieli je na cele główne i dodatkowe. Gdy samodzielnie potraktujemy cele dodatkowe jako operacjonalizację i uszczegółowienie celów głównych, przedstawiona początkowo struktura staje się bardziej zrozumiała. Każdy cel główny znajduje swoje rozwinięcie w celach dodatkowych, co pozwala na pełniejsze zrozumienie zakresu planowanych badań. Niemniej jednak już ten pierwszy etap wskazuje na bardzo szeroki ich zakres, a nawet za szeroki jak na rozprawę doktorską.

Przykładowo:

- Cel główny 1, opisowy (identyfikacja problemów wynikających z luki prawnej) rozwija się przez: przybliżenie genezy i znaczenia pojęć SI, analizę szans i zagrożeń jej stosowania, co pozwala zidentyfikować konkretne obszary problemowe.
- Cel główny 2, opisowy i eksplanacyjny (wypełnienie luki badawczej) rozwija się przez: analizę wybranych zagadnień prawa karnego w kontekście SI (przypisanie skutku, strony podmiotowej, formy sprawstwa, okoliczności wyłączające winę), co stanowi rzeczywiste wypełnienie luki poznawczej.
- Cel główny 3, eksplanacyjny (określenie podmiotu czynu zabronionego) rozwija się przez: wyjaśnienie koncepcji nadzoru człowieka nad SI oraz szczegółowe omówienie zagadnienia podmiotu czynu zabronionego popełnionego z użyciem SI.
- Cel główny 4, eksploracyjny (analiza odpowiedzialności karnej w wybranych obszarach) rozwija się przez: szczegółowe omówienie problemów związanych

z odpowiedzialnością karną za czyny zabronione w konkretnych obszarach życia (pojazdy autonomiczne, medycyna, *deepfake*, dobra wirtualne, autonomiczne systemy uzbrojenia)

- Cel główny 5, predykcyjny i preskryptywny (identyfikowanie wyzwań dla przyszłych regulacji) rozwija się przez: przeprowadzenie analizy obowiązujących przepisów Kodeksu karnego pod kątem skuteczności jego stosowania wraz z wskazaniem konieczności wprowadzenia nowych rozwiązań prawnych (postulaty *de lege ferenda*).

Ten poziom złożoności i szczegółowość pokazują, że Doktorantka rzeczywiście zmierza do dogłębnej analizy zasad przypisywania człowiekowi odpowiedzialności karnej w kontekście użycia przez niego systemu SI. I jest to dominujący obszar badawczy, przeważający nad pozostałymi aspektami prawnokarnymi, tj. sankcjami oraz typami czynów zabronionych. W tym zakresie zachodzi istotna dysproporcja między tytułowymi aspektami, która prowadzi do pytania: czy nie ten jeden obszar badawczy powinien zostać uwzględniona w tytule rozprawy jako jedyny? Rozprawa zawiera tylko wybrane aspekty przypisywania odpowiedzialności karnej, a nawet i one nie zostało przeprowadzone w sposób kompleksowy. Ponadto, postawiona w dalszej części Wstępu teza główna dotyczy tylko tego aspektu. W związku z tym tytuł samej rozprawy powinien wskazywać wprost na wybrane aspekty odpowiedzialności karnej w kontekście używania systemów SI przez człowieka, gdyż prawnokarnej ocenie podlegają jedynie tylko jego czyny.

Z drugiej jednak strony w rozprawie znajdują się także rozważania kwestie wokół potencjalnego przypisywania odpowiedzialności karnej systemom SI (również wersji „ucieleśnionej” – czyli robotom) lub nadawaniu im jakiejś formy osobowości prawnej w ogólności (nie zawsze karnej, co też jest źródłem niekonsekwencji prowadzonych rozważań). Temat ponoszenia jakiejś formy odpowiedzialności prawnej jest obecny w literaturze polskiej i obcej, ale – zdaniem piszącego tę recenzję – nie powinien być to równoważny obszar badań zawartych w rozprawie, a jedynie uzupełniający lub realizujący cel badań predykcyjnych lub preskryptywnych. Obecnie obowiązujące prawo karne nie daje żadnych możliwości na prowadzenie tego typu rozważań

normatywnych, gdyż jego zasady dotyczą ponoszenia specyficznej odpowiedzialności prawnej przez osoby fizyczne za ich czyny zabronione (opisane w części szczególnej kodeksu karnego lub innych ustawach) oraz prawnokarnej reakcji wymierzonej względem takich osób. Bardzo kontrowersyjnym zagadnieniem jest ewentualne poszukiwanie form odpowiedzialności dla systemów SI akurat w tej gałęzi systemu prawa. Proponowane zmiany byłyby tak daleko idące, że odpowiedzialność karna straciłaby swoją specyfikę, coś co pozwala ją odróżnić od pozostałych gałęzi systemu prawa. O ile można odnaleźć skrócone przedstawienie argumentów za rozszerzeniem odpowiedzialności karnej na systemy SI (ewentualnie roboty), to brak jest szerszego ukazania argumentów przeciwnych.

W stopniu niewystarczającym w celach badawczych wyodrębniono pozostałe dwa aspekty prawa karnego: konstrukcja typów czynów zabronionych oraz prawnokarna reakcja na posługiwanie się systemami SI przez sprawców przestępstw. Tylko pierwsze z nich są zaledwie wspomniane w rozprawie i pełnią jedynie funkcję ilustracyjną. Natomiast brak jest systemowych rozważań nad zakresem i sposobem kryminalizacji zachowań z wykorzystaniem systemów SI przez człowieka, tak jak mamy do czynienia z np. systemami teleinformatycznymi będącymi środowiskiem (w rozumieniu narzędzia albo przedmiotu zamachu lub ochrony), w którym popełniane są przestępstwa. Brak jest też systemowych rozważań nad istniejącym lub proponowanym system sankcji: kar, środków karnych i innych instrumentów i instytucji prawnokarnej reakcji na popełnienie przestępstwa.

Jednak najistotniejsze są błędy metodologiczne poczynione w początkowej fazie prowadzonych badań, rzutujących na całą treść rozprawy i jej naukowy charakter. W naukach społecznych przyjmuje się następujące etapy pracy naukowej. Badania powinny określać w pierwszej kolejności obszar lub przedmiot badań na podstawie gruntownego przeglądu literatury. Następnie dokonuje się wyboru celu badawczego (lub nielicznych celów) związanego z przedmiotem badań. Kolejnym etapem jest formułowanie pytania głównego i szczegółowych powiązanych z celem. Hipotezy badawcze stanowią hipotetyczną odpowiedź na stawiane pytania, które wymagają weryfikacji lub falsyfikacji za pomocą właściwie dobranych metod i technik

badawczych. To dopiero daje merytoryczną i weryfikowalną podstawę naukową do postawienia tezy. W recenzowanej rozprawie przedstawiona kolejność jest inna. Doktoranta wychodzi od tezy (s. 15) i następnie stawia pytania (s. 16), co należy uznać za błąd. W ogóle nie zostały przedstawione w założeniach hipotezy badawcze. Powstaje więc pytanie, po co badać problem, jeżeli od początku jesteśmy go pewni? Narusza to także zasadę obiektywizmu prowadzonych badań naukowych, gdyż dobiera się problemy do tezy badawczej.

Nawiązując także do zgodnego stanowiska prezentowanego przez komentatorów Ustawy Prawo o szkolnictwie wyższym i nauce K. Śtebza¹ oraz Ł. Kierznowskiego² w odniesieniu do kryterium „posiadanych umiejętności samodzielnego prowadzenia badań naukowych” należy przyjąć, że przesłanka ta byłaby spełniona, gdyby Doktorantka wykazał, że potrafi poprawnie formułować m.in. cele, problemy badawcze i stawiać hipotezy, dobierać metody badawcze i poprawnie je stosować, a w końcu – analizować wyniki i wyciągać z nich poprawne wnioski.

Należy przyznać, iż pytania są adekwatne względem celów badawczych, tzn. każde z kluczowych celów ma odpowiadające mu pytanie badawcze, a struktura pytań jest zasadniczo spójna z układem rozdziałów zapowiedzianym we Wstępie. Natomiast, częściowo dyskusyjnej jest ich powiązanie z postawioną tezą. Tylko problemy 4 - 7 bezpośrednio nawiązują do aspektów prawnokarnych. Ponadto zastrzeżenia mogą dotyczyć zakresu i stopnia skomplikowania niektórych pytań. Z perspektywy standardów rozpraw doktorskich (i szerzej metodologii badań społecznych) należałoby rozważyć rozdzielenie złożonych pytań na kilka prostszych oraz wyraźniejsze wskazanie jednego „pytania głównego”, a pozostałych jako szczegółowych. Natomiast samo brzmienie pytań wpisuje się jednak poprawnie w język nauk prawnych.

Teza badawcza sformułowana została w sposób zbyt ogólny, chociaż kategori czny: „regulacje prawne wynikające z obowiązującego Kodeksu karnego są niewystarczające dla ustalenia odpowiedzialności za czyny popełnione przy

¹ K. Śtebza [w:] K. W. Baran (red.), *Akademickie prawo zatrudnienia, Komentarz*, Wolters Kluwer, Warszawa 2020, LEX.

² Ł. Kierznowski, *Stopnie naukowe i stopnie w zakresie sztuki. Komentarz*, DfIn, Warszawa 2021, LEX.

wykorzystaniu sztucznej inteligencji" (s. 15). Ma ona charakter normatywno-aplikacyjny typowy dla rozpraw z zakresu nauk prawnych. Odpowiada ona obiegowym i intuicyjnym opiniom funkcjonującym w społeczeństwie, jak i fragmentarycznym badaniom prawnym prezentowanym w literaturze. Niemniej jednak powinna ona zostać wyartykułowana dopiero na końcu rozprawy, po przedstawieniu bieżącego stanu rozwoju systemów SI, przeprowadzonej weryfikacji aktualnego stanu prawnego oraz stać się podstawą do sformułowania postulatów zmian (*de lege ferenda*).

Do tego na pewno byłaby przydatna metoda dogmatyczną zastosowana względem norm krajowych, unijnych oraz obcych. Tak samo jak wzbogacenie jej o metodę egzegezy tekstu prawnego oraz hermeneutykę prawniczą, pozwalającą na rozszerzenie interpretacji unormowań prawnych o kontekst historyczny, społeczny i cel ustawodawcy. W rozprawie przedstawiono odniesienia do regulacji unijnych, polskich oraz rozwiązań w Niemczech, Singapurze, USA i Chinach, co jest metodologicznie uzasadnione, aczkolwiek nie mają one zawsze charakteru prawnokarnego. Natomiast brak jest wyraźnej informacji o racjonalnie uzasadnionych kryteriach doboru państw i organizacji międzynarodowych do porównania – czy wybór był podyktowany różnym poziomem nasycenia systemów SI w życiu społecznym, swoistym modelem regulacyjnym, czy innymi czynnikami. Mam jedynie lakoniczne stwierdzenie, że są to „najlepsze przykłady” działań legislacyjnych w obszarze odpowiedzialności karnej.

Można też sformułować kolejną uwagę, iż brak wyraźnej charakterystyki metody jakościowej studium przypadków we Wstępie, mimo że Rozdział IV obfituje w analizy konkretnych obszarów lub problemów (*deepfake*, autonomiczne pojazdy) uznanych przez Doktorantkę za najbardziej reprezentatywne obszary zastosowania systemów SI. Poza tym, brak jest wyjaśnienia kryteriów doboru i weryfikacji źródeł internetowych. Ze względu na dynamikę rozwoju systemów SI sięganie do nich jest zrozumiałe, ale nie powinno z metodologicznego punktu widzenia być dziełem przypadku, gdyż pojawiają się problemy oceny ich wiarygodności. To istotna luka wobec dynamiki prawa dotyczącego systemów SI. Dalej, rozprawa nie zawiera także charakterystyki ograniczeń metodologicznych przeprowadzonych badań (np. brak

dostępu do określonych źródeł, trudności w analizie właściwości technicznych systemów SI, dynamika zmian prawnych), co jest standardem w pracach naukowych. Wątpliwość budzi także konsekwencja i systematyczność postugiwanie się metodami badawczymi, które miały potwierdzać postawioną przedwcześnie tezę badawczą, a dokładniej pozwalałyby udzielić odpowiedzi na pytania badawcze i zaproponować postulaty, o czym w dalszych kolejnej części recenzji.

Zakończenie stanowi rodzaj skrótowego podsumowania treści rozprawy, gdyż cząstkowe wnioski znajdują się na końcu każdego rozdziału w punktach o jednolitym tytule „Uwagi końcowe”. Doktorantka powtarza ogólne argumenty o fragmentaryczności norm prawa karnego oraz o ich niewystarczającym stopniu dla ustalenia odpowiedzialności karnej za czyny popełnione przy wykorzystaniu systemów SI. W ten sposób odnosi się do postawionej tezy badawczej, do której dobierała argumenty. Tradycyjne instytucje prawa karnego, wypracowane na gruncie antropocentrycznego modelu odpowiedzialności, okazują się niewystarczające wobec autonomicznych lub częściowo autonomicznych systemów decyzyjnych. Dynamiczny rozwój systemów SI stawia przed prawem karnym nowe wyzwania, których ustawodawca nie mógł przewidzieć. Natomiast brak przejrzystych zasad odpowiedzialności może prowadzić do nadużyć i utraty zaufania użytkowników do nowych technologii. Tym samym Zakończenie odnosi się do celów głównych i dodatkowych oraz do pytań badawczych.

Zakończenie prezentuje siedem głównych grup postulatów z ostatniego rozdziału rozprawy. Dotyczą one bezpośrednio aspektów prawnokarnych: rozszerzenia ustawy o sztucznej Inteligencji o regulacje karnoprawne, zasady odpowiedzialności nadzorcy systemów SI, a także na inne podmioty występujące w cyklu życia systemów SI, harmonizacji przepisów w płaszczyźnie międzynarodowej czy problem podmiotowości prawnej SI. Za pośrednio związane z aspektami prawnokarnymi należy uznać te charakterze procesowym: specyficzne procedury śledcze oraz administracyjnym: Polski Rejestr SI.

Całościowa ocena warsztatu badawczego prowadzi do wniosku, iż Doktorantka nie opanowała go na oczekiwanym poziomie rozprawy doktorskiej w naukach

prawnych. Wskazują na to wymienione powyżej błędy poczynione już na etapie projektowania badań.

III. Ocena realizacji założeń rozprawy doktorskiej

Przyjęte podejście do konstruowania rozprawy doktorskiej – poprzez kolejne instytucje prawa karnego – reprezentuje tradycyjne ujęcie prawnokarne przy prowadzeniu badań dogmatycznych, co oczywiście jest zasadne i przez to akceptowalne. Jednakże utrudnia realizację zakładanych celów w sposób konsekwentny i systemowy. Podejście analityczne do realizacji poczynionych założeń i w konsekwencji struktury rozprawy powinno opierać się na – omawiany z resztą w rozprawie – poziomie autonomiczności systemów SI, roli człowieka w pętli decyzyjnej albo roli poszczególnych osób w całym procesie powstawania, wdrażania i użytkowania systemów SI. W ten sposób ułatwiłoby to gruntowną analizę instytucji prawnokarnych i diagnozę „rzeczywistych problemów związanych z odpowiedzialnością karną za czyny zabronione popełnione z wykorzystaniem systemów sztucznej inteligencji” (s. 20). Brak takiego analitycznego podejścia w strukturze prowadzonych rozważań powoduje, że w wielu miejscach pracy brak jest kompleksowej analizy sytuacji i czytelnik ma wrażenie „przeskakiwania” z wątku na wątek w poszczególnych punktach i pomiędzy nimi. Prowadzi to do wniosków, które są często albo fragmentaryczne, gdyż dotyczą tylko wybranych aspektów: czyli grupy systemów SI, podmiotów w ich cyklu życia lub relacji człowiek-system, albo bardzo ogólne.

Zamiast niekonsekwentnego używania terminologii opisującej podmioty w cyklu życia systemów SI, można byłoby tutaj skorzystać z nomenklatury określonej przez unijny tzw. AI Act (który Doktoranta przecież wykorzystuje). Z punktu widzenia prawnokarnych badań najważniejszymi są: dostawca (*provider*) systemu lub modelu SI (rozвивa lub zleca rozwój i wprowadza system na rynek pod własną nazwą albo znakiem towarowym), podmiot stosujący/wdrażający (*deployer*) i użytkownik (*user*), każdy, kto wykorzystuje system SI w swojej działalności; obok importera, dystrybutora i upoważnionego przedstawiciela dostawcy. Zasady i zakres ponoszenia

odpowiedzialności (lub jej braku) każdej z nich powinien być analizowany w każdym z pozostałych aspektów opisujących np. poziom autonomiczności czy roli człowieka w pętli decyzyjnej.

Natomiast przyjęta struktura rozprawy jest zgodna z przedstawionymi problemami badawczymi. Ujawnia logiczny tok rozumowania: od fundamentów teoretycznych (geneza, terminologia) przez dogmatykę prawa karnego, studia przypadków, po postulatory *de lege ferenda*. Tytuły rozdziałów i paragrafów są relatywnie precyzyjne i funkcjonalne, wykorzystują nomenklaturę prawniczą adekwatną do rozprawy z zakresu prawa karnego materialnego. Numeracja wielopoziomowa zachowuje przejrzystość nawet przy swojej głębokiej strukturze.

Należy przyznać, że Doktoranta wykazuje się stosunkowo wysokim stopniem opanowania metody badawczej, jaką jest dogmatyka prawa, zwłaszcza karnego. W sposób obszerny prezentuje swoje umiejętności interpretacyjne z wykorzystaniem analizy norm, stanowisk doktryny i orzecznictwa. Jednak chociażby w Rozdziale III można odnieść wrażenie, że prowadzone analizy mają charakter opracowania typowego dla podręcznika do prawa karnego materialnego lub ewentualnie komentarza (powielanie treści znanych w nauce prawa karnego), niż analizy problemowej, służącej przede wszystkim poszukiwaniu odpowiedzi na pytania badawcze. Na pewno jednak istnieje powtarzalna dysproporcja w poszczególnych jednostkach technicznych tego rozdziału (punktach) między opisem instytucji prawa karnego a odpowiedzią na postawione pytania.

Ponadto, Doktoranta nie dochowuje obiektywizmu w prowadzeniu badań. Widoczne jest to w zakresie przedstawienia argumentacji za nadaniem osobowości prawnej systemom SI, a w zasadzie robotom (tzw. ucieleśnionej SI). Nie jest jasne, czy rozciąga się ona także na odpowiedzialność karną, czy tylko dotyczy kwestii bycia podmiotem praw. Natomiast trafne są rozważania na przykładzie ewolucji statusu zwierząt czy niewolników. Jednocześnie brak jest przedstawienia argumentów przeciwnych, obecnych także w polskiej literaturze np. w opracowaniach autorstwa prof. K. Mamaka. Wydaje się, że w tym zakresie brakuje Doktorantce konsekwencji w rozróżnianiu sytuacji systemów SI (w tym wersji *agentic AI* lub *AI Agent*, ale także

słabej i mocnej albo szczególnej i ogólnej SI, czasami nazywanej przez Doktorantkę „mocno zaawansowaną” – s. 200) i robotów w nią wyposażonych. Wynika to z stosunkowo obszernego Rozdziału I, który nie zawiera typowej w pracach operacjonalizacji pojęć, które następnie byłyby konsekwentnie stosowane w prowadzonych analizach.

Na marginesie za duży skrót myślowy uznaję pisanie, że robot popełnia przestępstwo lub ponosić będzie odpowiedzialność karną. Systemy SI mogą wywoływać skutki w świecie rzeczywistym (wspominana płaszczyzna ontologiczna) lub wirtualnym, które – gdyby wywołał je człowiek swoim czynem – byłyby objęte prawnokarnym wartościowaniem (płaszczyzna normatywna).

Dodatkowo warto byłoby wyraźnie i strukturalnie rozgraniczać w rozprawie rozważania o charakterze istotnym dla obszaru badawczego, czyli prawnokarnym od innych, tj. cywilistycznego czy administracyjnego, aby nie gmatwały odbioru prowadzonych analiz i zgłaszanych postulatów – np. kwestia nadawania osobowości prawnej systemom SI (lub robotom), propozycja rejestru systemów SI. Ewentualnie wskazać, jak tego typu rozważania wpływają lub mogą wpływać na zmiany w prawie karnym w kontekście poszukiwania rozwiązań dla zdarzeń z udziałem systemów SI albo tworzenie zasad ich odpowiedzialności typu *sui generis* (jak było w przypadku odpowiedzialności podmiotów zbiorowych). Ponadto, w kontekście wszystkich prowadzonych rozważań (nie tylko analizy podmiotu przestępstwa lub strony podmiotowej) należałoby konsekwentnie uwzględnić także sytuację, gdy system SI jest tworzony od początku do celów przestępczych (opisywane w literaturze kryminologicznej *AI-Crimes* albo *AI-Crime-as-a-Service*) albo został on z jakiś powodów niewłaściwie użyty. To rozróżnienie ma istotne znaczenie dla oceny odpowiedzialności karnej osób na poszczególnych etapach cyklu życia systemów SI.

Mimo obszernego Rozdziału I, nie zawiera on operacjonalizacji pojęć technologicznych użytych w rozprawie, nie wskazuje np. na przyjętą typologię systemów SI, która zostanie wykorzystana do prowadzenia badań dogmatycznych nad prawnokarną oceną czynów osób posługujących się takimi systemami, czy nawet nad *ponoszeniem odpowiedzialności przez takie systemy. W konsekwencji przyczynia się*

to do fragmentaryczności rozważań prawnokarnych w kolejnych rozdziałach dotyczących poszczególnych instytucji prawnokarnych lub wybranych aspektów odpowiedzialności karnej, gdyż nie są one prowadzone w sposób konsekwentny dla każdego rodzaju systemu SI wyróżnionego wg kryterium ich autonomiczności czy wytłumaczalności.

Podobnie jest w kontekście roli człowieka w pętli decyzyjnej. W zasadzie najczęściej Doktorantka opisuje przypadki *human-on-the-loop* czy *human-out-of-the-loop* (choć nie używa akurat tych pojęć). Brakuje przypadku, *human-in-the-loop*, kiedy człowiek aktywnie uczestniczy w procesie działania systemu – dostarcza dane, podejmuje lub zatwierdza decyzje, może korygować system i jest częścią samej pętli decyzyjnej (jako swoistej sieci relacji), a nie jedynie nadzoruje go lub blokuje (*human-on-the-loop*). To kryterium jest istotne dla kompleksowej oceny przypadków w całym kontinuum zdarzeń stanowiących podstawę ich prawnokarnej oceny. Natomiast nie pokrywa się z typologią autonomiczności systemów SI.

O ile podrozdział III.1 zawiera rozważania na temat przypisania skutku osobom fizycznym występującym na różnych etapach cyklu życia systemu SI, to kolejny podrozdział ma zupełnie inny charakter. Doktorantka skupia się na możliwości realizacji znamion strony podmiotowej przez system SI w aspekcie intelektualnym i woluntatywnym. Są to na pewno „ciekawe” (s. 129) rozważania, ale utrzymując spójność i ciągłość analizy wybranych aspektów odpowiedzialności karnej, powinien dotyczyć jednak oceny zachowań osób fizycznych. Jest to bardzo ważny i praktyczny aspekt zwłaszcza w kontekście funkcjonowania systemów SI na różnym poziomie autonomiczności, działających na zasadzie czarnych skrzynek (*black-box*), systemów SI wytłumaczalnych lub rozliczanych, kwestii naruszenia reguł ostrożności (wynikających z analiz ryzyka i obowiązków *compliance* – opisywanych częściowo w rozprawie w kontekście unijnego AI Act – nałożonych na osoby będące na różnych etapach cyklu życia systemu SI) w przypadku nieumyślności. Brak rozważań nad zakresem interpretacji znamion strony podmiotowej – umyślnością, nieumyślnością lub jej postaci mieszanej osób posługujących się systemami SI należy uznać za poważny brak rozprawy. To jest i powinien być kluczowy element oceny, czy „regulacje

prawne wynikające z obowiązującego Kodeksu karnego są niewystarczające dla ustalenia odpowiedzialności za czyny popełnione przy wykorzystaniu sztucznej inteligencji” (s. 15).

W kolejnych częściach jest podobnie, Doktoranta częściowo skupia się nad futurystycznymi wizjami odpowiedzialności systemów SI. Analizuje problemy adaptacji rozważań doktrynalnych nad formami zjawiskowymi względem systemów SI, winą i okolicznościami ją wyłączającymi. Ponownie mamy rozbudowane i niebudzące większych zastrzeżeń badania dogmatyczne poparte literaturą i często orzecznictwem. Natomiast brak jest uzasadnienia dla częściowego pominięcia analizy tych instytucji prawa karnego w odniesieniu do osób posługujących się systemami SI. Prowadzi to do konstatacji, iż wadą rozprawy przedstawionej do recenzji jest brak konsekwencji i systemowości prowadzonych rozważań. Można byłoby ten stan poprawić – jak to już zostało wcześniej przedstawione – stosując podejście analityczne według kryteriów nawiązujących do charakterystyki systemów SI. W efekcie nie mielibyśmy do czynienia z fragmentarycznymi rozważaniami ilustrowanymi ciekawymi przykładami (ale nadal brak jest określenia metody ich doboru), ale logiczny i całościowy przegląd kontinuum zdarzeń stanowiących podstawę ich prawnokarnej oceny. Natomiast, należy także podkreślić, iż pojawiają się trafne rozważania dotyczące odpowiedzialności osób na różnych etapach cyklu życia systemów SI.

Jeden aspekt prawnokarny został pominięty w rozprawie. Jest nim okoliczność wyłączająca bezprawność czynu osoby fizycznej, gdy działa w ramach ryzyka nowatorstwa. Akurat ta instytucja prawa karnego – jak chyba żadna inna – nawiązuje do nowatorstwa systemów SI w różnych aspektach życia społecznego. W szczególności chodzi wymieniane w art. 3 pkt 55 AI Act tzw. piaskownice regulacyjne (*regulatory sandbox*). Nie wyłączają one odpowiedzialności karnej, ale stanowią kontekst regulacyjny dla ryzyka nowatorstwa. Dla oceny odpowiedzialności karnej kluczowe będzie, czy: ryzyko mieściło się w granicach przewidzianych i nadzorowanych w piaskownicy, zostało zredukowane przy użyciu dostępnych środków oraz nie doszło do przekroczenia jej warunków. Ma to bezpośrednie

przełożenie na określenie zakresu obowiązków osoby fizycznej (np. zapobiegawczych jako gwarant niewystąpienia skutku) lub przewidywalności skutku (ocena ryzyka ich wystąpienia). Nawiązuje to także do kwestii projektowania, rozwoju i wdrażania systemów SI, które od początku uwzględniają ograniczenia prawne, kwestie etyczne lub bezpieczeństwa (np. *Privacy* – lub *Security-by-design*).

Pod względem rozważań prawno Karnych wyżej należy ocenić Rozdział IV rozprawy. Różnorodne przykłady pokazują problemy zakresu odpowiedzialności osób fizycznych (choć nadal nie zostało przedstawione kryterium lub kryteria ich wyboru). Doktorantka wykazuje się umiejętnością ich rozwiązywania z zakresu prawa karnego z wykorzystaniem metody dogmatycznej, w wybranych wariantach. Ilustrują one różne konsekwencje związane z odpowiedzialnością karną interakcji osoby z systemami SI, które są na różnym poziomie autonomiczności, ale także o różnych rodzajach umiejscowienia człowieka w pętli decyzyjnej (na etapie projektowania lub wdrożenia modelu/systemu SI lub podejmowania już decyzji z jego udziałem/wsparciem) – *human-in-the-loop*, *human-on-the-loop* czy *human-out-of-the-loop*. W tym zakresie należy poczynić jedną uwagę, otóż Doktorantka nie przedstawia przekonujących argumentów, uzasadniających zawarcie rozważań na temat działalności SI jako *quasi* artysty. W tej części rozprawy brak jest jakichkolwiek rozważań prawno Karnych. Ponadto, nie nawiązuje ona w żaden sposób do kolejnej części o kradzieży dóbr wirtualnych z wykorzystaniem systemów SI.

Rozdział V przedstawia wiele interesujących informacji o polityce legislacyjnej wybranych państw i organizacji międzynarodowej wokół systemów SI, cyberprzestępczości, cyberbezpieczeństwie na podstawie pozyskanych źródeł bibliograficznych o różnorodnym charakterze. Miał on ukazać „działania legislacyjne do wprowadzenia rozwiązań prawnych zabezpieczających zasady odpowiedzialności karnej w obszarze funkcjonowania sztucznej inteligencji”. Są to przede wszystkim rozważania o charakterze administracyjnoprawnym, częściowo cywilistyczne, a Doktorantka nie zawsze wskazuje, dlaczego lub w jaki sposób mają one przełożenie lub związek z instytucjami prawa karnego materialnego. Do pewnego stopnia wyjątkiem

jest przypadek Chin, gdzie wprost wskazano przepisy o odpowiedzialności dostawców i osób wdrażających i utrzymujących systemy SI.

Natomiast ostatnie fragmenty tego rozdziału stanowią bezpośrednie nawiązanie do głównego przedmiotu badań i próbę zrealizowania preskryptywnego celu rozprawy. Nadat widoczna jest fragmentaryczność rozważań i propozycji, które nie zawsze mają charakter prawnokarny. Brak jest wykazania „konieczności” wprowadzenia przepisów prawnokarnych do polskiej ustawy uszczegóławiającej unijny AI Act. Ponadto, propozycje polegające na wprowadzeniu przepisów karnych umożliwiających pociągnięcie do odpowiedzialności karnej osób za czyny zabronione popełnione z użyciem systemów AI jest co najmniej zastanawiające (V.3.1.3). Wady metodologiczne i fragmentaryczność prowadzonych badań dogmatycznych nie dają podstaw do proponowania nowych zasad odpowiedzialności karnej pod względem podmiotów ją ponoszącym (Rozdział V.3.2). Dlaczego nie rozważyć wprowadzenia np. przestępstw indywidualnych ze względu na etap życia cyklu systemu SI lub typów kwalifikowanych istniejących czynów zabronionych ze względu na cechę podmiotu? Natomiast propozycje dotyczące Procedur śledczych i dowodowych (V.3.1.5) i Współpracy międzynarodowej i harmonizacji standardów (V.3.1.6) są zdawkowe i bardzo ogólne. Wykraczają poza zakładany materialnokarny przedmiot badań, wchodząc w kwestie procedury karnej i kryminalistyki. Ponadto, nie nawiązują one w sposób bezpośredni do wyników prowadzonych badań. Propozycja związana z oceną przewidywania i zarządzania ryzykiem (V.3.1.4) jest w zasadzie powtórzeniem zapisów AI Act i istniejących już licznych zaleceń, będących wynikiem prac zarówno organizacji międzynarodowych, jak i sektorowych. Brak jest tutaj elementu oryginalności.

Zakończenie rozprawy posiada przejrzystą strukturę tematyczną, obejmującą podsumowanie badań naukowych, skrótową prezentację siedmiu kluczowych spostrzeżeń, kompleksową analizę *de lege lata* oraz postulaty *de lege ferenda*. Natomiast brak jest wyraźnego, metodycznego podsumowania realizacji celu głównego i celów szczegółowych. W naukach społecznych przyjmuje się za dobrą praktykę, aby w Zakończeniu wprost i w sposób skondensowany odnieść się do tego, czy i jak zrealizowano cel główny i cele szczegółowe wraz z uzasadnieniem, najlepiej z

użyciem ich pierwotnego brzmienia. To samo dotyczyłoby kwestii weryfikacji lub falsyfikacji hipotez szczegółowych i w konsekwencji hipotezy głównej – których zostały całkowicie pominięte. W obecnym kształcie Zakończenie zawiera treści pozwalające jedynie rekonstruować realizację celów badawczych przez samego czytelnika. Ponadto, z perspektywy standardów metodologicznych w naukach prawnych można oczekiwać wyraźniejszego, redakcyjnie odseparowanego *passusu*, który wprost ukazuje, które pytania badawcze zostały w pełni rozstrzygnięte, a które tylko zarysowane jako agenda dalszych badań.

IV. Ocena formalnej strony rozprawy

Na rozprawę składa się pięć rozdziałów o szczegółowej strukturze, na którą składa się punkty i podpunkty. Stopień jej rozbudowania jest różny w poszczególnych rozdziałach. Inne wyszczególnione części techniczne rozprawy to Wstęp, Zakończenie, Literatura przedmiotu, Wykaz aktów prawnych, Orzeczenia sądowe, Netografia oraz Wykaz rysunków i tabel. W przypadku Neografii często mają one postać hiperlinków bez opis, co utrudnia ocenę ich wiarygodności lub zawartości. Tylko Rysunek 1. Dyscypliny AI został obszernie opisany w rozprawie. Nie można tego powiedzieć o zawartości Tabeli 1. i Tabeli 2 (która powinna być zatytułowana jako Rysunek 2) odnoszących się do systemów ekspertowych.

Pierwszoplanową kwestią jest posługiwanie się terminologią związaną z obszarem badawczym. Należy podkreślić, iż Doktoranta prawidłowo i sprawnie porusza się w nomenklaturze prawnokarnej. Jednakże nie można tego samego powiedzieć o nomenklaturze związanej z częścią technologiczną rozprawy. Już we Wstępie widoczne jest zamienne posługiwanie się przez Doktorantkę różnymi pojęciami „sztuczna inteligencja”, AI, „technologia AI”, „systemy sztucznej inteligencji”, które nie zawsze są tożsame. W dalszych częściach rozprawy do tego dochodzi także pojęcie „system AI”, czy też „robot” jako swoiste „ucieleśnienie” SI – pojawiające się w nowszej literaturze przedmiotu tzw. *embodied AI*, często też *„embodied artificial intelligence”* lub *„embodied agents”*. Z rozprawy nie dowiadujemy się o wzajemnych relacjach pojęciowych między nimi. Powstaje pytanie, czy

rozważania odnoszące się do systemów SI w wybranych aspektach prawnokarnej odpowiedzialności osób odnoszą się także do robotów? I odwrotnie, czy rozważania odnoszące się do jednej odmiany autonomiczności systemów SI odnoszą się także do innych? Albo dlaczego inne odmiany zostały pominięte w analizie?

Można sformułować uwagę techniczno-redakcyjną. Rozdział IV.5 łączy dwa odrębne zagadnienia („Dzieła tworzone przez AI” i „kradzież dóbr wirtualnych”), które mają odmienną naturę prawną (prawa autorskie a mienie). Merytorycznie właściwym byłoby rozdzielenie tych obszarów w kolejnym podpunkcie. Natomiast wracając do wcześniejszych rozważań merytorycznych, ta część tylko pośrednio, jeżeli w ogóle, odnosi się do tematu rozprawy. Można byłoby ją usunąć bez szkody dla treści.

Bibliografia jest ogromna i różnorodna. Nie sposób się nie zgodzić z Doktorantką, iż opracowań prawnokarnych dotyczących posługiwania się przez ludzi systemami sztucznej inteligencji jest niewiele. Jednakże nie udało się jej dotrzeć i uwzględnić ważnych publikacji na temat polskiego prawa karnego materialnego w tym zakresie. I problemem nie jest dynamicznie zmieniająca się rzeczywistość, ale niewystarczająca kwerenda biblioteczna. Przede wszystkim brak jest kompleksowego opracowania autorstwa A. Błachnio-Parzych, A. Kosyło, A. Sakowicz, M. Woiński, P. Zakrzewski, *Polish national report on criminalisation of AI-related offences*, [w:] F. Miro-Llinares, C. Duvac, T. Toader, M. Santisteban Galarza (red.), *Criminalisation of AI-related offences, Reveue Internationale de Droit Penal*, Maklu, Apeldoorn 2024, s. 331-358 oraz polskiej literatury tam wymienionej, jak również pozostałych raportów krajowych zawartych w tej pracy zbiorowej; F.J. Szkitadź, *Czy obowiązujący kodeks karny nadąga za algorytmami sztucznej inteligencji?*, *Kwartalnik Krajowej Szkoły Sądownictwa i Prokuratury* 2024, nr 4(56), s. 115–128 (obie pozycje wydane rok wcześniej niż przygotowano rozprawę). Te publikacje wprost odnoszą się do tytułu rozprawy i problemów, które powinny być poddane badaniu.

Brakuje także pionierskich prac prof. K. Mamaka, *Odpowiedzialność karna za wypadek drogowy z udziałem samochodu bez kierowcy*, *Paragraf na Drodze* 2015, nr 4; tegoż, *Prawo karne przyszłości*, Wolters Kluwer, Warszawa 2017; tegoż, *Robotics, AI, and Criminal Law*, Routledge 2024. A także bardziej szczegółowych np. J.

Lachowski, *Prawnokerne oceny eksploatacji prostytucji uprawianej przez roboty humanoidalne*, *Ius Novum* 2024, nr 1 (18), s. 1–15; J. Kaczmarek, A. Sampolski, *Wybrane zagadnienia odpowiedzialności karnej pojazdów autonomicznych*, *Monitor Prawniczy* 2018, nr 9.

Przy tej objętości rozprawy występowanie literówek zdarza się, ale są one nieliczne. Poza powyższymi uwagami terminologicznymi, język rozprawy jest komunikatywny. Natomiast za błąd należy uznać odsyłanie czytelnika do dysku na komputerze, np. pozycje w Literaturze przedmiotu: 76, 110, 221, 224, 254, 329, 392, 594, 638, 671 zamiast linków zasobów internetowych (w tym numerów DOI). Zdarza się brak konsekwencji w konsekwentnym opisie pozycji bibliograficznych (np. pisownia tytułów czasopism, opisu stron internetowych zawierających posty lub pliki z tekstem artykułów naukowych).

V. Wnioski końcowe

Konkludując przytoczone w niniejszej recenzji uwagi, pozwalają one dojść do przekonania, że przedstawiona do recenzji rozprawa doktorska mgr Gabrieli Garci Vera nie spełnia wymagań określonych w ustawie z dnia 20 lipca 2018 r. roku Prawo o szkolnictwie wyższym i nauce (tekst jednolity Dz.U. z 2024 r., poz. 1571 t.j. ze zm.) i zdaniem recenzenta nie może być podstawą dalszego postępowania w przewodzie doktorskim.

W szczególności przedstawiona do recenzji rozprawa nie spełnia wymogu z ustępu 1 art. 187 tej ustawy, czyli nie prezentuje umiejętności samodzielnego prowadzenia pracy naukowej ze względu na wskazane w recenzji błędy metodologiczne.

Białystok, dnia 31 stycznia 2026 roku.