

Artificial intelligence i singularity w procesie stosowania prawa

dr Tomasz A. Zienowicz¹

Niewątpliwie pewne elementy rewolucji teleinformatycznej czy też informatycznej wkraczają w polski system prawa oraz w systemy prawne państw europejskich. Na polskim gruncie uwagę skupiają regulacje ustawy z 5.9.2016 r. o usługach zaufania oraz identyfikacji elektronicznej². Akt ten ma służyć implementacji rozporządzenia Parlamentu Europejskiego i Rady (UE) nr 910/2014 z 23.7.2014 r. w sprawie identyfikacji elektronicznej i usług zaufania w odniesieniu do transakcji elektronicznych na rynku wewnętrznym oraz uchylającego dyrektywę 1999/93/WE³. Rodzi się tu oczywiste pytanie: czy idziemy w nieznane, jakim jest zastosowanie sztucznej inteligencji w procesie stosowania, a może w przyszłości także stanowienia prawa oraz czy ww. nowelizacja jest lub może okazać się pierwszym krokiem na tej drodze?

Uwagi wstępne

Na wstępie, należy w szerszym, społecznym kontekście zwrócić uwagę na zagadnienia dotyczące szeroko pojętej problematyki elektronicznej. Jak już wspomniano, w ostatnim czasie mamy do czynienia ze zbiegającymi się zdarzeniami mogącymi mieć zasadniczy wpływ na życie w sferze informatycznej oraz ewolucję prawa w zakresie jego stosowania na omawianej płaszczyźnie. Google Inc. opracowało i przedstawiło system WaveNet. WaveNet jest obecnie największym generatorem mowy ludzkiej i syntezatorem dźwięków. System generuje mowę w celu jej naśladowania, a wyniki wskazują, że rezultaty są często bardziej naturalne niż proste naśladownictwo syntetyczne. WaveNet pobiera także dane z zasobów internetu w celu syntezy sygnałów audio i przedstawia automatycznie generowane utwory muzyczne. Od dawna zauważalne zdaje się być nasilenie działań nakierowanych na realizację marzeń o szeroko pojętej interakcji człowiek – komputer. Wytwarzanie zdolności komputerów do rozumienia mowy ludzkiej zostało zrewolucjonizowane w ostatnich latach poprzez zastosowanie tzw. głębokich sieci neuronowych, takich jak np. Google Voice Search. Jednak generowanie mowy przez komputery, tj. tworzenie procesów określanych jako syntezy mowy lub tekstu na mowę, nadal w dużej mierze opierają się na tworzeniu bardzo dużych baz danych krótkich fragmentów mowy, które są rejestrowane z pojedynczych źródeł, a następnie w procesie rekombinacji kompletowane w określonej wypowiedź. Doprowadziło to do powstania wielkiego zapotrzebowania na stan, gdzie wszystkie informacje wymagane do wygenerowania danych są przechowywane w parametrach modelu, a zawartość i właściwości mowy mogą być sterowane za pomocą wejść do modelu. Takie rozwiązanie proponuje WaveNet. Można by to ująć w ten sposób, że system funkcjonuje jako w pełni spłotowa sieć neuronowa, gdzie warstwy spłoty mają różne parametry, które pozwalają jej w sposób

wykładniczy na generowanie danych potrzebnych do budowy docelowego dźwięku⁴.

Drugim istotnym wydarzeniem wartym odnotowania jest pojedynek programu AlphaGo z mistrzem świata *Lee Sedolem* w Go. Pojedynek odbył się między 9 a 15.3.2016 r. Odbyło się pięć gier w Go. Człowiek został pokonany 4:1 przez program DeepMind opracowany przez Google⁵. Po pojedynku koreańska organizacja Go przyznała programowi AlphaGo najwyższą rangę Go arcymistrza i honorowy 9 dan, uznając, że było to przełomowe wydarzenie w dziejach tej gry. Wskazać należy, że aby pokonać gracza w Go, program komputerowy musi uwzględnić nie tylko sformalizowanie zasady strategii, lecz także zaprojektować program nauczania. Zasady tej gry są bowiem bardziej złożone niż w szachach⁶. Program AlphaGo różni się znacząco od poprzednich przede wszystkim poprzez zastosowanie sieci neuronowych. Działanie tego programu nie opiera się na heurystyce stosowanej przez człowieka, lecz w dużej mierze program uczy się sam przez analizę dziesiątków milionów ruchów wykonanych w poprzednich grach Go, wybierając dla siebie najbardziej korzystne do określonej sytuacji w meczu. Program AlphaGo ocenia pozycję gracza w celu przewidzenia jego ruchu na

¹ Autor jest doktorem nauk prawnych, adwokatem.

² T.j. Dz.U. z 2019 r., poz. 162 ze zm.

³ Dz.Urz. UE L Nr 257, s. 73; dalej jako: rozporządzenie 910/2014.

⁴ Zob. deepmind.com (dostęp z 4.4.2017 r.).

⁵ Pojedynek został porównany z historycznym meczem szachowym pomiędzy *Deep Blue* i *Garry Kasparowem* z 1997 r. Go to kompleksowa gra planszowa, która wymaga intuicji, kreatywnego i strategicznego myślenia. Przez wielu uważana jest za trudniejszą w dziedzinie sztucznej inteligencji niż szachy. Wielu uważa, że w dziedzinie sztucznej inteligencji Go wymaga analizy większej liczby elementów naśladowujących ludzką myśl niż szachy.

⁶ Przed 2015 r. najlepszy program komputerowy grający w Go przeciwko człowiekowi mógł wykazać przewagę tylko na amatorskim poziomie, na małej 9 x 9 planszy program radził sobie lepiej niż gracz rzeczywisty. Przed powstaniem AlphaGo niektórzy badacze twierdzili, że komputer nigdy nie pokona najlepszych graczy w Go. *Elon Musk*, poprzedni inwestor w Deepmind, stwierdził w 2016 r., że eksperci w dziedzinie myśli AI potrzebują ok. 10 lat od osiągnięcia zwycięstwa przeciwko mistrzowi świata w Go.

wiele posunięć w przyszłość⁷. Wskazać należy, że AlphaGo to program komputerowy opracowany przez Google DeepMind do celów gry w Go. Algorytm AlphaGo wykorzystuje kombinację uczenia maszynowego i wyszukiwania na zasadzie drzewa kombinacji w połączeniu ze szkoleniem ze zbiorów gier z komputerem i ludźmi. AlphaGo początkowo została stworzona, aby naśladować ludzką grę, próbując dopasować ruchy graczy poprzez analizę gier historycznych z bazy około 30 mln ruchów z 160 000 gier. Kolejnym etapem była analiza tych danych w celu ustanawiania własnych posunięć i tworzenia własnych strategii oraz scenariuszy gry⁸. Autorzy programu wskazują, że mimo iż mają zaprogramowane urządzenie do prowadzenia gier z ludźmi, to nie wiedzą, jak program tworzy własną strategię gry. Program stworzony jest ze zbiorów danych i algorytmów, natomiast gra prowadzona jest samodzielnie przez AlphaGo⁹.

Należałoby na bazie wskazanych zjawisk zastanowić się, czy tzw. *deep learning* dostępny dla poszczególnych programów komputerowych oraz postęp w syntezie mowy mogą i powinny mieć wpływ na wymiar sprawiedliwości, na stanowienie i stosowanie prawa w szerokim tego słowa znaczeniu. Na tle wskazanej ustawy oraz rozporządzenia 910/2014 wskazać należy na grupę przepisów wprowadzonych do polskiego systemu prawnego realizującego, zdaniem ustawodawcy, idee identyfikacji elektronicznej i usług zaufania w odniesieniu do transakcji elektronicznych na rynku wewnętrznym.

Tworzenie ram dla zastosowania sztucznej inteligencji w procesie stosowania prawa

Zanim jednak uwaga czytelnika zostanie skierowana na kwestię materialnoprawną, omówienia wymaga, w wyselekcjonowanym zakresie, aktualna literatura z zakresu sztucznej inteligencji. W pierwszej połowie lat 90. XX w. podstawa badań nad sztuczną inteligencją i prawem w Stanach Zjednoczonych, mogąca mieć jakikolwiek związek z procesem stosowania prawa, koncentrowała się głównie na rozwiązywaniu poszczególnych spraw, nie obejmując swym zakresem całościowych badań nad przedmiotową problematyką¹⁰. Wszystkie badane sprawy traktowano jednorodnie, niezależnie od tego, w jakim okresie były rozpoznawane, oraz niezależnie, z jakiego rządu czy rodzaju sądu pochodziły. Jednym z podstawowych założeń tych badań było ustalenie polegające na tym, że traktowanie przypadków jako jednorodnych jest jednym z ważnych sposobów prowadzenia badań. Takie podejście umożliwiało zbieranie rozsądnej liczby interesujących spraw, wybieranych z poszczególnych sądów z rozróżnieniem na ich wagę oraz zakres terytorialny itp. Jednak, jak zauważono, ignorowanie czasu, miejsca i kontekstu proceduralnego może zniekształcać sposoby badania precedensów, a także

rolę tych badań, jaką odgrywają w decyzjach podejmowanych w przyszłości¹¹. Istotne dla tego typu badań jest wprowadzenie pojęcia, a ściślej rzecz ujmując sposobu segregacji danych w celu budowy decyzji (w tym procesowej w przyszłości) – ADF. ADF to *Abstract Dialectical Framework*, czyli dialektyczne abstrakcyjne ramy, w zakresie których przetwarzane są uogólnione argumenty, w których każdy tworzy węzeł z innym w warunkach ich wzajemnej akceptacji. Pozwala to na modelowanie różnych typów zależności, uwzględniając wzajemną argumentację na rzecz poparcia swego stanowiska, a także różne typy węzłów w zakresie jednej ramy. Według twórców tej koncepcji norma standardowa może być uogólniona do dialektycznych ram (ADF), w przypadku stabilnych i preferowanych (przez ich użytkowników) wypowiedzi mających semantyczny sens do nieco ograniczonej klasy, którą nazwać można ramą bipolarną¹². ADF-y mają zalety łączenia hierarchicznej struktury pojęć, przy czym tradycyjna ich reguła oparta jest na warunku wzajemnej akceptacji. Ponadto ADF stosowane z użyciem zasad użyteczności i logicznego podziału umożliwiają skuteczną modularyzację projektu. Korzyści związane z oprogramowaniem wykorzystującym ADF dają także podstawę do projektowania aktów prawnych na podstawie orzecznictwa¹³. Formalne modele argumentacji stały się inspiracją do badań dla różnych form sztucznej inteligencji, takich jak niekonsekwentne rozumowanie, systemy wielozadaniowe i projektowanie uzasadnień prawnych za pomocą sztucznej inteligencji. Zaznaczyć tu wypada, że ramy argumentacji stanowią szczególny sposób definiowania skutków niemonotonicznych przez konstruowanie i porównanie argumentów. Te ramy to zasadniczo wykresy, których węzły reprezentują argumenty. Według tej struktury argument zostaje zaakceptowany, chyba że został uwzględniony w innym zaakceptowanym argumentcie¹⁴. Dodać należy, że kolejność

⁷ Ekspertzi stwierdzili błędy w grze AlphaGo przeciwko Fan (październik 2015 r.), w szczególności odnoszące się do braku świadomości i analizy całej planszy, tak więc przed meczem otwarcia przeciwko Lee nie było wiadomo, o ile program poprawił swoją grę od swojego poprzedniego meczu. AlphaGo nie był dostosowany do gry przeciwko Lee Sedolowi, ponieważ szkolenie AlphaGo wymaga analizy dziesiątków milionów gier, a kilkaset lub tysiąc gier pobranych od danego gracza nie było na tyle miarodajnych, aby móc stwierdzić, że AlphaGo nauczył się gracza, przeciwko któremu występuje.

⁸ Zob. wired.com (dostęp z 12.3.2017 r.).

⁹ Ibidem.

¹⁰ M.J. Sergot, F. Sadri, R.A. Kowalski, F. Kriwaczek, P. Hammond, H.T. Cory, The British Nationality, Act as a logic program, Communication of the ACM 1986, Nr 29(5), s. 370–386.

¹¹ L. Al-Abdulkarim, K. Atkinso, T. Bench-Capon, Accommodating change, [16.11.2016 r.], s. 409 i n.

¹² G. Brewka, S. Woltran, Abstract Dialectical Frameworks, [w:] F. Lin, U. Sattler, M. Truszczynski (red.), Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, Toronto-Ontario-Canada 2010, s. 15.

¹³ L. Al-Abdulkarim, K. Atkinso, T. Bench-Capon, Accommodating change..., s. 417.

¹⁴ G. Brewka, S. Woltran, Abstract Dialectical Frameworks..., s. 15–16.

tekstów w warunkach akceptacji jest określona przez odpowiednie priorytety przypisane do węzłów poprzednich lub następnych. Zauważyć należy, że priorytety są zorganizowane w oparciu na węzeł-węzeł i nie ma potrzeby zachowania spójności preferencji w węzłach. Ograniczenie to mogłoby zostać nałożone, ale wydaje się, że ta elastyczność daje więcej korzyści niż strat¹⁵. Badania w zakresie sztucznej inteligencji zostały skierowane także na inne kwestie istotne dla przedmiotowej problematyki. Wskazać tu należy, że jednym z założeń przedmiotu badań jest ustalenie polegające na tym, że skuteczność argumentacji zależy od wartości wyznawanych przez interlokutora. Toteż jednym z tematów analizy argumentacji opartej na wartościach jest porównanie wartości. Modele formalne dotyczące porównywania wartości korzystają z szablonów jakościowych lub ilościowych. Niektórzy z badaczy zastosowali techniki łączące jakościowe i ilościowe szablony opracowane z myślą o dowodzeniu opartym na wartościach¹⁶.

Uzasadnienie czynności dokonywanych w toku postępowania nie jest łatwym zajęciem, dotyczy to zarówno czynności wpadkowych, podawanych ustnie w toku prowadzenia sprawy, jak i ostatecznych (kończących daną instancję) decyzji procesowych. Podejmowane decyzje – akty decyzyjne obarczone są częstokroć błędami mogącymi prowadzić do wypaczenia istoty sprawiedliwości i skutkować mogą powstaniem niejednokrotnie poważnych konsekwencji dla podmiotów będących stronami takich postępowań. Analityczne metody właściwego postępowania dowodowego zazwyczaj skupiają się na jednym z trzech narzędzi: argumentów, spostrzeżeń (scenariuszy) lub prawdopodobieństw (w szczególności na pewnych etapach postępowania cywilnego). Łączenie w ramach budowy struktury sztucznej inteligencji ww. narzędzi napotyka określone przeszkody. Główną z nich jest błędna interpretacja zakresów wskazanych wyżej narzędzi, a właściwie można byłoby stwierdzić, że owa błędna interpretacja zakresów dotyczy braku zgody co do znaczenia tych pojęć (braku powszechnej zgody). To pozwala stwierdzić, że formalna budowa struktury, jaką miałyby posługiwać się sztuczna inteligencja, nie może być oparta li tylko na ilościowych i jakościowych zbiorach argumentów łączonych z hipotezami stawianymi w określonych postępowaniach. Przykładowo wskazać tu należy, że ustalenie, jaki przebieg miało przestępcze działanie, jest częstokroć zadaniem o wysokim stopniu skomplikowania. Taki stan sprzyja powstawaniu błędów, które – można byłoby zaryzykować stwierdzenie – są wprost proporcjonalne do stopnia skomplikowania danego stanu rzeczy. Jak wskazują badania, ilość błędów popełnianych w postępowaniach karnych, np. w Holandii, jest na istotnym poziomie. Ostatnie badania wskazują, że szacuje się, iż liczba błędnych ustaleń w wyrokach skazujących w procesach karnych w Holandii jest rzędu 5–10%¹⁷. Taki odsetek błędnych wyroków jest nie tylko porażką wymiaru sprawiedliwości, ale także zagraża bezpieczeństwu społecznemu. W rezultacie

istnieje potrzeba zastosowania narzędzi analitycznych, które mogą zapobiegać błędom oraz, co istotniejsze, wprowadzenia systemu – może opartego właśnie na sztucznej inteligencji, który wyeliminowałby albo istotnie zredukował skalę tych uchybień. Jak już wspomniano, w literaturze dotyczącej prawidłowego rozumowania dowodowego w ramach propozycji budowy mechanizmów sztucznej inteligencji wyróżniono trzy ustrukturyzowane narzędzia analityczne: argumenty, scenariusze i prawdopodobieństwa¹⁸. Narzędzia te mają na celu pomóc w organizacji i uporządkowaniu materiału dowodowego, zapobiegając tym samym najczęściej występującym błędom, oraz dać podstawę do budowy algorytmów w ramach sztucznej inteligencji, mogących podjąć próbę skutecznego dania odporu narastającemu problemowi błędów wymiaru sprawiedliwości.

Nawiązując do wcześniej wymienionego porządku narzędzi, należy wskazać, że w analizie argumentacyjnej rozważeniu podlega struktura powiązanych ze sobą dowodów, przyczyn i hipotez. Zazwyczaj materiał dowodowy nie jest jednolity, tj. dowody zebrane w sprawie często są sprzeczne ze sobą. Wskazać należy, że poddanie analizie niejednolitego materiału dowodowego ma istotny sens badawczy, ponieważ tylko wtedy, kiedy właśnie ten materiał dowodowy nie jest jednolity, wyłaniają się zagadnienia istotne dla omawianej problematyki. Niektóre dowody w takich zbiorach są sprzeczne z możliwymi wnioskami. Analiza argumentacyjna pomaga w rozwiązywaniu takich konfliktów¹⁹. W analizie scenariuszy dochodzi do weryfikacji rozmaitych hipotetycznych scenariuszy, które następnie są rozważane w świetle dowodów. Analiza dokonywana poprzez hipotetyczne scenariusze pomaga spójnej interpretacji wszystkich dowodów²⁰. W analizie probabilistycznej pewne okoliczności uprawdopodobniają rzeczywisty stan rzeczy. Analiza probabilistyczna określa różne stopnie prawdopodobieństwa w dowodach. Samo usuwanie prawdopodobieństwa na rzecz dowodzenia odbywa się

¹⁵ L. Al-Abdulkarim, K. Atkinso, T. Bench-Capon, Accommodating change..., s. 419.

¹⁶ B. Verheij, Formalizing value-guided argumentation for ethical systems design, *Artif Intell Law* 2016, Nr 24 (4), s. 387–407.

¹⁷ T. Derksen, *Onschuldig vast*. ISVW, Leusden 2016, s. 38.

¹⁸ T. Anderson, D. Schum, W. Twining, *Analysis of evidence*, 2nd edn. Cambridge University Press, Cambridge 2005, s. 33; A.P. Dawid, W. Twining, M. Vasiliki (red.), *Evidence, inference and enquiry*, Oxford 2011, s. 12; H. Kaptein, H. Prakken, B. Verheij (red.), *Legal Evidence and Proof: Statistics, Stories, Logic (Applied Legal Philosophy)*, Farnham 2009, s. 68.

¹⁹ J.H. Wigmore, *The principles of judicial proof as given by logic, psychology, and general experience, and illustrated in judicial trials*, 2nd edn 1931, 3rd edn *The science of judicial proof* 1937, Boston 1913, s. 37.

²⁰ W.L. Bennett, M.S. Feldman, *Reconstructing reality in the courtroom*, London 1981, s. 71; N. Pennington, R. Hastie, *Reasoning in explanation-based decision making*, *Cognition* 1993, Nr 49(1–2), s. 123–163; W.A. Wagenaar, P.J. van Koppen, H.F.M. Crombag, *Anchored narratives. The psychology of criminal evidence*, London 1993, s. 13.

metodami naukowymi pozwalającymi przenieść określone stany do zbioru dowodów (chodzi tu np. o badania DNA)²¹.

Rola sztucznej inteligencji w tworzeniu narzędzi do jej wykorzystania w stosowaniu prawa

Badania nad sztuczną inteligencją dostarczają wiele tematów natury poznawczej. Sztuczna inteligencja w zakresie stosowania prawa, połączona z badaniem stanu faktycznego, wykazuje wiele niedociągnięć natury formalnej, tj. występujących w funkcjonowaniu algorytmów sztucznej inteligencji. Głównym źródłem różnic, na którym opierają się nieprawidłowości w działaniu sztucznej inteligencji, są jakościowe i ilościowe niedobory w analizie materiału dowodowego²². Niektórzy naukowcy w związku z takim stanem rzeczy proponują podejście probabilistyczne do badań nad systemem²³. Zwrócić tu wypada uwagę na wiele zagadnień ściśle związanych z problemami, przed którymi stoi proces tworzenia sztucznej inteligencji i jej zastosowania w omawianym zakresie. Wskazuje się m.in. na to, że w zakresie ilościowym występowanie określonych zjawisk można uznać za wsparcie dla probabilistycznej metody funkcjonowania sztucznej inteligencji w ocenie stanu faktycznego, akcentując probabilistyczne adaptacje argumentacji abstrakcyjnej opartej na określonych założeniach²⁴. Innym istotnym zagadnieniem jest zbadanie rozkładu prawdopodobieństwa w supozycjach argumentacji abstrakcyjnej i logicznej, co prowadzi do wyeksponowania różnych rodzajów niespójności, które mogą się pojawić podczas analizy określonych stanów faktycznych²⁵. Problematyka badawcza obejmuje także badanie rozumowania niemonotonicznego pod względem domyślnego rozumowania, opierającego się np. na semantyce Adamsa pod kątem ekstremalnych prawdopodobieństw²⁶.

W literaturze dotyczącej analiz scenariuszy wykorzystano także kilka pojęć w celu analizy „jakości” rozważanych scenariuszy. Trzy pojęcia, jak się wydaje, odgrywają tu zasadnicze znaczenie, a mianowicie: spójność scenariusza, prawdopodobieństwo scenariusza i kompletność scenariusza. Scenariusz jest spójny, jeśli nie zawiera sprzeczności. Przykładowo podejrzany nie może być zarówno w domu, jak i na miejscu zbrodni. Scenariusz jest wiarygodny, gdy pasuje do wiedzy o świecie (ogólnej, powszechnej), np. w scenariuszu morderstwa ofiara umiera z powodu strzału z broni palnej. Scenariusz jest kompletny, gdy wszystkie jego elementy znajdują się w odpowiednich relacjach logicznych, tj. pasują do siebie²⁷.

Jednakowoż wskazać także należy, że według niektórych badaczy omawianej problematyki, w opozycji do prezentowanego podejścia, a w związku z wnioskowaniem bayesowskim, działanie sztucznej inteligencji nie wymaga modelowania dowodów pod kątem ich prawdopodobieństwa²⁸. W ten

sposób rozwiązano np. problem polegający na konieczności podania większej liczby danych niż są aktualnie dostępne w czasie rzeczywistym, gdy dany problem jest rozwiązywany. Zastrzec tu jednak należy, że istota problemu wyłania się dopiero przy bardziej złożonych przykładach. Bardziej realistyczne (częściej występujące) przykłady mogą szybciej wykazać, czy i jak tak zaprojektowana sztuczna inteligencja radzi sobie z zadaną materią. Kolejną kwestią jest pochodzenie stanów faktycznych poddanych analizie. Z formalnego punktu widzenia ocena argumentów i scenariuszy zachodzi w kontekście modelu danego przypadku. Taka ocena wymaga, aby ten model miał charakter weryfikowalny. W obecnych badaniach widzimy próby wyjaśnienia, w jaki sposób argumenty, scenariusze i rachunek prawdopodobieństwa zostają zaprzęgnięte w analizę materiału dowodowego. Poprzez wyjaśnienie formalnych powiązań między jakościowymi i ilościowymi modelami analitycznymi można zrozumieć, dlaczego

²¹ F. Taroni, C. Champod, P. Margot, Forerunners of Bayesianism in early forensic science, *Jurimetrics* 1998, Nr 38, s. 183–200. Teoria argumentacji skupia się na rozważeniu argumentów, będących argumentami presumptycznymi (zwanymi także wzmacniającymi), w zakresie logicznego wykraczania poza ich przesłanki. Na tle klasycznej logiki argument ze zbioru P do wniosków Q wykracza poza jego rozbiór, gdy Q nie wynika logicznie z P. Wskazać należy, że wiele argumentów stosowanych w praktyce jest domniemywanych i nie mają one oparcia w zgromadzonym w sprawie materiale dowodowym. Przykładowo świadek może twierdzić, że podejrzany był na miejscu zbrodni. Fakt, że świadek tak zeznawał, nie uzasadnia logicznie faktu, że podejrzany był na miejscu zbrodni. W szczególności gdy zeznania świadków są celowo fałszywe albo oparte na niedokładnych sposrzeniach lub niemożności przypomnienia sobie, co u niektórych wymusza konfabulację, by uniknąć przyznania, że się nie pamięta, gdy tymczasem faktycznie podejrzanego nie było w ogóle na miejscu zbrodni. Niekiedy organ rozpoznający sprawę w takiej sytuacji zawęża czynności poznawcze do rozstrzygnięcia, która z wersji jest prawdziwa, a którą należy wykluczyć, z góry zakładając, że jedna z nich musi być prawdziwa, a druga nie. W takiej sytuacji odbywa się niejako ważenie argumentów, tj. dowodów zgromadzonych w sprawie. I te, które są „mocniejsze”, zdają się potwierdzać rzeczywisty stan rzeczy.

²² B. Verheij, *Formalizing...*, s. 387–407.

²³ J. Kohlas, B. Anrig, R. Haenni, P.-A. Monney, Model-based diagnostics and probabilistic assumption-based reasoning, *Artif Intell* 1998, Nr 104 (1–2), s. 71–106.

²⁴ P.M. Dung, P. Thang, Towards probabilistic argumentation for jury-based dispute resolution, [w:] *Computational Models of Argument: Proceedings of COMMA 2010*, Desenzano del Garda, Italy, 8–10 Sept 2010, Amsterdam 2010, s. 171–182.

²⁵ A. Hunter, A probabilistic approach to modelling uncertain logical arguments, *Int J Approx Reason* 2013, Nr 54, s. 47–81; S. Benferhat, A. Safiotti, P. Smets, Belief functions and default reasoning, *Artif Intell* 2000, Nr 122(1–2), s. 1–69.

²⁶ *Ibidem*.

²⁷ N. Pennington, R. Hastie, Reasoning..., s. 123–163; W.A. Wagenaar, P.J. van Koppen, H.F.M. Crombag, Anchored narratives..., s. 39; F.J. Bex, P.J. van Koppen, H. Prakken, B. Verheij, A hybrid formal theory of arguments, stories and criminal evidence, *Artif Intell Law* 2010, Nr 18, s. 1–30; B. Verheij, F.J. Bex, Accepting the truth of a story about the facts of a criminal case, [w:] H. Kaptein, H. Prakken, B. Verheij (red.), *Legal Evidence and Proof...*, s. 161–193; C.S. Vlek, H. Prakken, S. Renooij, B. Verheij, Representing the quality of crime scenarios in a Bayesian network, [w:] A. Rotolo (red.), *Legal Knowledge and Information Systems: JURIX 2015: the twenty-eighth annual conference*, Amsterdam 2015, s. 131–140.

²⁸ B. Verheij, Proof with and without probabilities, *Artif Intell Law* 2017, Nr 25, s. 127–154.

niektórzy wolą racjonalnie analizować dowody jakościowo, a inni ilościowo (liczbowo)²⁹.

Rosnąca potrzeba tworzenia norm, które mogłyby być interpretowane przez sztuczną inteligencję w procesie stosowania prawa, zaczyna być coraz bardziej zauważalna. Obecne podejście do budowy norm w systemach prawa wielu krajów jest tradycyjne, tj. norma ma po prostu zakazywać działań niedozwolonych lub tworzyć zespół reguł (zasad), które podmiot ma obowiązek przestrzegać, zarówno w ramach praw i obowiązków, jak i w celu uniknięcia sankcji i kar. Gdyby jednak budowa norm dawała możliwość wykorzystywania zdolności do znajomości systemu prawa, poprzez zrozumienie wartości społecznych i moralnych, którym normy mają służyć, tzn. gdyby przepisy umożliwiały taką interpretację prawa, która uwzględniałaby moralność w powszechnym rozumieniu, w szczególności chodzi tu o takie rozumowanie, które pozwala przy konkurencyjnych normach zastosować rozsądek przeciętnego adresata tych norm, to dawałoby to szanse na wykorzystanie sztucznej inteligencji do celów stosowania prawa³⁰. Jednym ze sposobów na wdrożenie takiego rozwiązania jest skorzystanie z systemu argumentacji opartego na wartościach i zaprojektowanie go do praktycznego „rozumowania” – użytku na zasadzie wygenerowania argumentów za i przeciw działającym przy użyciu schematu, w którym adresaci norm wybierają między działaniami opartymi na ich preferencjach w ramach stosowania prawa opartego na wartościach³¹. Rozumowanie moralne wymaga, aby podmioty prawa miały akceptowalny zestaw wartości i akceptowalny sposób ich egzekwowania. Tu pojawia się zagadnienie wykraczające poza ramy tego opracowania, polegające na próbie udzielenia odpowiedzi, czy norma prawna musi być oznajmiana poprzez język, a nie zapis cyfrowy. Zmiana sposobu zapisu normy prawnej przeniosłaby system w całkowicie inny wymiar, otwierając możliwości, o których obecnie nie sposób myśleć i których przewidzieć się nie da.

Niektórzy autorzy podają w wątpliwość zagadnienie, aby kiedykolwiek udało się doprowadzić do tego, że sztuczna inteligencja może stosować zasady moralne³². W literaturze przedmiotu od lat wskazywano różnicę pomiędzy domyślnymi systemami etycznymi (w zakresie stosowania norm) a systemami opartymi na zasadach etycznych³³. To rozróżnienie stosuje się obecnie do większości podejść do analizy zastosowania sztucznej inteligencji na podstawie stosowania norm etycznych. Takiej argumentacji towarzyszyły obawy polegające na tym, że systemy etyczne muszą rozwijać także normy chroniące ludzkość przed ekspansją sztucznej inteligencji i skierowaniem jej przeciwko ludziom. W rzeczywistości jednak okazało się³⁴, że autonomiczne systemy oparte na sztucznej inteligencji, działające w realnym świecie, stały się coraz bardziej częścią naszej (ludzkiej) rzeczywistości i jak na razie nie stanowią zagrożenia ani też nie są systemami samorozwijającymi się³⁵.

Inne prezentowane stanowisko polega na tym, że np. w projekcie *Moora* określone schematy są ustanowione do działania w ograniczonym, przewidywalnym środowisku, podczas gdy rzeczywisty świat nie jest przewidywalny i pojawiają się nieprzewidziane sytuacje w nieograniczonym zakresie. Trudno jest więc zaaplikować do systemu mającego za zadanie stosowanie określonych norm opartych na określonych wartościach taki zbiór dopuszczalnych działań, które mogłyby odnieść się do nieskończonej ilości nieprzewidywanych okoliczności³⁶. W związku z tym autorzy wskazują, że podmioty prawa w otwartych systemach, w których dopuszcza się kierowanie normami moralnymi, powinny mieć możliwość wyraźnego uzasadnienia działań, jakie podejmują, a w szczególności mieć prawo do uzasadnienia swojego stanowiska dotyczącego określonego zachowania w ramach prawa jako zgodne z normami etycznymi³⁷.

Budowa systemów wartości w zastosowaniu sztucznej inteligencji

Niewątpliwie sztuczna inteligencja, mająca być mechanizmem rozwiązującym problematykę związaną z bardziej złożonymi problemami merytorycznymi, powinna opierać się na systemach etycznych. Powstaje jednak pytanie, jakie są kryteria etyczne, za pomocą których należy wybrać zarówno normy etyczne będące podstawą późniejszych decyzji, jak i ich stopień złożoności dopasowany do rozwiązywanych problemów. Chociaż ilościowe podejście może wystarczyć w prostych scenariuszach, to w złożonych wymagających kalkulacji nie da się zastosować prostych rozwiązań. Pamiętać należy, że w rzeczywistym stanie rzeczy charakterystycznymi scenariuszami są te o bardziej złożonych konstrukcjach, stąd pochodną są zbiory dylematów moralnych rozbudowane

²⁹ *Ibidem*.

³⁰ T. Bench-Capon, S. Modgil, Norms and value based reasoning: justifying compliance and violation, *Artif Intell Law* 2017, Nr 25, s. 29–64.

³¹ *Ibidem*.

³² J.H. Moor, The nature, importance, and difficulty of machine ethics, *IEEE Intell Syst* 2006, Nr 21(4), s. 18–21.

³³ *Ibidem*.

³⁴ Moore wskazywał na opisane obawy ponad dekadę temu.

³⁵ S. Russell, D. Dewey, M. Tegmark, Research priorities for robust and beneficial artificial intelligence: an open letter, *AI Mag* 2016, Nr 36(4), s. 3–4. Moor uznał, że domyślne systemy etyczne można uznać za moralne. Ale nie wszyscy zgodziliby się na takie postawienie sprawy, na przykład Immanuel Kant (1785 r.) twierdzi, że działania podejmowane w zgodzie z prawem mają sens tylko wówczas, gdy są one również wykonane dla dobra prawa.

³⁶ M. Esteva, D. De La Cruz, C. Sierra, *Islander: an electronic institutions editor*, [w:] *Proceedings of 1st international joint conference on autonomous agents and multiagent systems 2002*, s. 1045–1052; T. Agotnes, W. van der Hoek, M. Tennenholtz, M. Wooldridge, Power in normative systems, [w:] *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems 2009*, vol. 1, s. 145–152; W. van der Hoek, M. Roberts, M. Wooldridge, Social laws in alternating time: Effectiveness, feasibility, and synthesis, *Synthese* 2007, Nr 156(1), s. 1–19.

³⁷ *Ibidem*.

proporcjonalnie do tej problematyki³⁸. Zatem niewątpliwie wymagane jest tu podejście obliczeniowe określające uzasadnienie takiego rozumowania. Próba takiego obliczeniowego rozumowania została zbadana zarówno w grupie podmiotów prawa, jak i w kontekście przekonań, pragnień oraz intencji (BDI)³⁹, a przede wszystkim poprzez połączenie deontologii, epistemologii i logiki⁴⁰. Zaznaczyć tu należy, że podejścia te nie odnoszą się bezpośrednio do kwestii racjonalnego wyboru opcji działania poprzez odwołanie się do kryteriów etycznych. Jednak od 2006 r. rodzi się coraz większe zainteresowanie korzystaniem z teorii argumentacji w kontekście praktycznego rozumowania w celu rozwiązania takiej problematyki⁴¹. Dużą rolę odegrały w tym aspekcie badania nad rozumowaniem i argumentacją na bazie wartości⁴². W badaniach tych kryteria etyczne odwoływały się do przyjętych wartości promowanych przez działania i względne znaczenie oparte na rankingu tych wartości tworzony w zakresie częstotliwości posługiwania się nimi⁴³. Na tle tych rozważań trudno byłoby przyjąć za podstawę tezę uzasadniającą budowę sztucznej inteligencji na algorytmie opartym tylko na argumentacji i uzasadnianiu stawianych tez, z pominięciem wartości, jako wystarczającym dla formalnej realizacji celu, jakim jest zastosowanie konkretnej normy. Kompleksowe spojrzenie na omawianą problematykę może prowadzić do wniosku polegającego na tym, że modele takie mogą wymagać włączenia BDI i logiki deontycznej, a być może także integracji ilościowej i jakościowej w zakresie wartościowania⁴⁴. Dlatego, jak się wydaje, celem badań nad samym stworzeniem i zastosowaniem sztucznej inteligencji powinno być rozumowanie i argumentacja oparta na wartościach. Takie działania mogłyby odegrać ważną rolę w zdolności do zrozumienia norm i związanych z nimi względów etycznych, a tym samym także do radzenia sobie z nieprzewidywanymi sytuacjami⁴⁵. W badaniach nad sztuczną inteligencją analizie poddana została także teoria gier w zakresie dylematu więźnia. Badanie rozmaitych scenariuszy na podstawie tej teorii wskazuje na większą efektywność uzyskiwaną w warunkach współpracy, nie zaś w warunkach konkurencji⁴⁶. Jednak należy tu zwrócić uwagę na fakt, że badania prowadzono na wysoce jednorodnej społecznie grupie i stosowano tę samą metodologię dla każdego eksperymentu. W konsekwencji tak przyjętej metodologii nie jest prawdopodobne, aby teoretyczne modele gier i ich kryteria były uzasadnieniem dla norm postępowania napotykanych w codziennym życiu społeczeństw⁴⁷. Innym podejściem do omawianej problematyki jest modelowanie scenariuszy jako stanowych⁴⁸ schematów przejściowych (STD) oraz zbadanie, w jaki sposób normy można projektować w takich sytuacjach tak, aby uniknąć niepożądanych wyników⁴⁹. W tych podejściach czynniki są typowe i występują przy użyciu modelu *Belief-Desire-Intention* (BDI)⁵⁰. Model BDI zakłada, że podmioty prawa mają zestaw przekonań i zestaw celów (pragnień). Działania są wybierane przez identy-

fikowanie pragnień, które można zrealizować w konkretnej sytuacji (zamiary kandydata), a następnie podporządkować się jednej lub kilku intencjom i wybrać taki sposób działania, aby zrealizować związane z tym cele. To jednak rodzi pytanie, skąd pochodzą pragnienia, a w systemach BDI nie ma żadnego wyjaśnienia tej kwestii. Często są one całkowicie niedefiniowalne, a nawet w sytuacjach, w których obecnie znajduje się podmiot, istnieje stały zestaw potencjalnych pragnień, z których niektóre są aktywne w danej sytuacji, a inne nie. Ta niezdolność do uzasadnienia pragnień znacznie ogranicza zdolność rozumowania moralnego w przejrzysty sposób i na precyzyjnie określonych zasadach⁵¹.

W literaturze przedmiotu występują także alternatywne podejścia do kwestii rozumowania praktycznego i sformalizowanego jako rozumowania opartego na wartościach, traktowanego jako środek umożliwiający podmiotom prawa uzasadnienie norm poprzez rozumienie wartości społecznych i moralnych⁵². Założenie tu jest takie, że podmioty prawa powiązane są szeregiem wartości społecznych, aspiracjami lub celami, jak np. wolność, równość, braterstwo, bogactwo, zdrowie i szczęście. Wartości te stanowią powody, dla których pewne zachowania są uważane za cele podmiotu, a zatem pozwalają na uzasadnienie jego pragnień⁵³. Podstawową ideą jest tu założenie, że podmioty prawa mają zestaw określonych wartości, a ich aspiracje i preferencje można usystematyzo-

³⁸ D. Dennet, *Darwin's Dangerous Idea*, Simon & Schuster, New York 1995, s. 11

³⁹ A.S. Rao, M. Georgeff, Modeling rational agents within a BDI-architecture, [w:] *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR'91)* 1991, s. 473–484.

⁴⁰ J. van Den Hoven, G.J. Lokhorst, Deontic logic and computer-supported computer ethics, [w:] J. Moor, T.W. Bynum (red.), *Cyberphilosophy: the intersection of philosophy and computing*, Blackwell 2002, s. 376–386.

⁴¹ I. Rahwan, L. Amgoud, An argumentation based approach for practical reasoning, *Proc AAMAS 2006*, Nr 6, s. 347–354.

⁴² K. Atkinson, T. Bench-Capon, Practical reasoning as presumptive argumentation using action based alternating transition systems, *Artif Intell* 2007, Nr 171(10–15), s. 855–874; B. Verheij, *Formalizing...*, s. 387–407.

⁴³ *Ibidem*.

⁴⁴ T. Bench-Capon, S. Modgil, Norms and value..., s. 29–64.

⁴⁵ *Ibidem*.

⁴⁶ E. Ullmann-Margalit, *The emergence of norms*, Oxford 1977, s. 36–38.

⁴⁷ T. Bench-Capon, S. Modgil, Norms and value..., s. 29–64.

⁴⁸ W zasadzie nie można tego sformułowania inaczej przetłumaczyć.

⁴⁹ M. Wooldridge, W. van der Hoek, On obligations and normative ability: towards a logical analysis of the social contract, *J Appl Log* 2005 Nr 3, s. 396–420; W. van der Hoek, M. Roberts, M. Wooldridge, Social laws in alt (2010) Optimal social laws, [w:] *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems 2010*, vol. 1, s. 667–674.

⁵⁰ A.S. Rao, M. Georgeff, Modeling rational agents..., s. 473–484; M. Wooldridge, *An introduction to multiagent systems*, Wiley, Hoboken 2009, s. 171.

⁵¹ I. Rahwan, L. Amgoud, An argumentation..., s. 347–354.

⁵² J. Raz, *Practical Reasoning*, Oxford University Press, Oxford 1979, s. 76.

⁵³ K. Atkinson, T. Bench-Capon, States, goals and values: revisiting practical reasoning, *Argum Comput* 2016 Nr 7(2–3), s. 135–154; K. Atkinson, T. Bench-Capon, Value based reasoning and the actions of others, [w:] 22nd European Conference on Artificial Intelligence 2016, s. 680–688.

wać, tj. m.in. ułożyć w kolejności wartości społecznych (sklasyfikować). Próba promowania lub unikania preferowanych wartości stanowi powód dla pożądania niektórych stanów i wybierania konkretnych działań. W kontekście opartego na takich argumentach modelach rozumowania, preferowane wartości mogą być postrzegane jako publiczne, tj. są akceptowane przez ogół⁵⁴. Zatem argumentacja (jej przyjęcie) na rzecz zachowania honorującego określoną wartość społeczną zależy nie tylko od samego argumentu, ale musi być także akceptowana społecznie i tak powinna być postrzegana przez systemy sztucznej inteligencji w procesie stosowania prawa. Argumentacja jest zbijana poprzez inne argumenty tylko wówczas, gdy kontrargumenty są tej samej wartości w zakresie akceptacji społecznej⁵⁵. Korzystanie z tej metody umożliwia wyjaśnienie (i obliczanie) różnych argumentów zaakceptowanych przez różne grona odbiorców. Założenie wartościowe zostało wykorzystane jako podstawa praktycznego rozumowania w wielu opracowaniach⁵⁶. Podstawą rozważań staje się tu podejście polegające na tym, że podmioty prawa muszą stawić czoła sytuacjom wymagającym dokonania wyboru co do sposobu działania. Zwykle wybierają działanie, które ma najbardziej korzystne konsekwencje dla nich samych. Ponieważ różne podmioty będą miały różne preferencje, toteż dokonywane przez nich wybory będą zgoła odmienne. Jednak podmioty te nie istnieją w próżni, muszą mieć świadomość tego, że ich wybory wpływają na inne czynniki, a inne czynniki wpływają (m.in. wtórnie) na konsekwencje własnych działań. W systemach norm, jakimi kieruje się grupa społeczna, poszczególne działania zostaną zaakceptowane lub nie, co dla określonego działania ma zasadnicze znaczenie. Poszczególne działania mogą być zatem sprzeczne z interesem grupy lub mogą pobudzać chęć zgromadzenia dodatkowych informacji o tym, czego inni mogą się spodziewać poprzez podjęte działania. Szeroko pojętymi normami mogą być normy społeczne lub prawne. W obydwu przypadkach są one zazwyczaj wzmacniane sankcjami. Pamiętać jednak należy, że normy prawne są wspierane sankcjami formalnymi, w szczególności sankcją karną, podczas gdy normy społeczne są wspierane przez nieformalne sankcje, takie jak ostracyzm, wyśmiewanie lub dezaprobata. Natomiast normy moralne powinny być zabezpieczone normami prawnymi oraz powinny być z nimi zgodne, jednak normy moralne nie są zazwyczaj obwarowane sankcjami. Prawo, wprowadzając instytucję sankcji i kar, często zmierza do realizacji celu, jakim jest promowanie trwałości grupy lub społeczeństwa oraz wspólnych wartości⁵⁷.

Co do wartości, to wskazać należy, że już samo projektowanie algorytmiczne sztucznej inteligencji mogącej podejmować zadania decyzyjne w oparciu na określonych wartościach, napotyka zasadniczy problem właśnie na etapie definiowania i wartościowania (katalogowania) wartości, na których miałby być zbudowany ów system. Jak się wskazuje

w literaturze przedmiotu, w społeczeństwie nie ma tzw. czystych przyjemności (tzw. przyjemności netto), zawsze są one okupione pracą lub wyrzeczeniami do nich prowadzącymi. Niezadowolenie z pracy, która pozwala zarobić, jest niwelowane przyjemnościami podczas odpoczynku od niej, co prowadzi także do wniosku, że nie ma samych obowiązków netto. Już na tym przykładzie wskazać można problematykę jakościową, tj. czy norma ma preferować zwiększoną ochronę pracy czy wypoczynku, czy może próbować znaleźć równowagę pomiędzy tymi postawami społecznymi. Ponadto wskazać należy, że istnieją przyjemności wyższego rzędu, których konsumentami jest odpowiednio wąska grupa osób w danym społeczeństwie, więc ochrona ich stanów zawsze budzić będzie określone kontrowersje⁵⁸.

Tu pojawia się inne zagadnienie dotyczące nieomawianej dotąd problematyki. Otóż problemy natury etycznej, aksjologicznej występują także na etapie tworzenia prawa, więc, jak się wydaje, samo ich występowanie nie może być podstawą dopatrywania się złożoności problematyki z zakresu sztucznej inteligencji jedynie w zakresie stosowania prawa. Uwagę należy skierować na jurystyczny aspekt tworzenia prawa. Traktowanie orzecznictwa, w tym precedensu, w krajach, gdzie system prawa nie jest, co do zasady, systemem prawa stanowionego, jako źródło stanowiące podstawę do ustaleń subsumcyjnych, jest zasadniczo pożądanym i powszechnym zachowaniem. Cyto-

⁵⁴ C. Perelman, *The new rhetoric*, Berlin 1971, s. 56–58.

⁵⁵ P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artif Intell* 1995, Nr 77(2), s. 321–357; T. Bench-Capon, G. Sartor, A model of legal reasoning with cases incorporating theories and values, *Artif Intell* 2003, Nr 150(1), s. 97–143.

⁵⁶ *Ibidem*; A. Garcez, D. Gabbay, L. Lamb, Value-based argumentation frameworks as neural-symbolic learning systems, *J Log Comput* 2005, Nr 15(6), s. 1041–1058; J. Tremblay, I. Abi-Zeid, Value-based argumentation for policy decision analysis: methodology and an exploratory case study of a hydroelectric project in Quebec, *Ann Oper Res* 2016, Nr 236(1), s. 233–253; F. Nawwab, T. Bench-Capon, P. Dunne, A methodology for action-selection using value-based argumentation, [w:] P. Besnard, S. Doutre, A. Hunter (red.), *Computational Models of Argument – Proceedings of COMMA 2008*, vol. 172 *Frontiers in Artificial Intelligence and Applications*, Amsterdam 2008, s. 264–275; B. Burgemeestre, J. Hulstijn, Y.-H. Tan, Value-based argumentation for justifying compliance, *Artif Intell Law* 2011, Nr 19(2–3), s. 149–186; T. Payne, V. Tamma, Using preferences in negotiations over ontological correspondences [w:] *PRIMA 2015: Principles and Practice of Multi-Agent Systems* 2015, s. 319–334; C. Trojahn, P. Quaresma, R. Vieira, An extended value-based argumentation framework for ontology mapping with confidence degrees, [w:] I. Rahwan, S. Parsons, C. Reed, *Argumentation in Multi-Agent Systems* 2007, vol. 4946 of *lecture notes in computer science*, 2008, s. 132–144. Autorzy uciekają się do opowiadania o mrówce i koniku polnym. Latem konik polny wypoczywa, śpiewa i je to, co jest dla niego osiągalne. Mrówka zaś ciężko pracuje, gromadząc zapasy na zimę. Zimą mrówka żyje w ukryciu w dostatku, natomiast konik polny umiera z głodu, ponieważ nie zadbał o zgromadzenie należytego pożywienia. Dając inny przykład, a mianowicie biblijną przypowieść o synu marnotrawnym, wskazują, że syn ten został przyjęty na łono rodziny, a jego występki zostały mu przebaczone. Tu następuje konstatacja, że podobne zachowania mają inne zakończenia i towarzyszą im odmienne skutki (sankcje na płaszczyźnie prawa).

⁵⁷ T. Bench-Capon, S. Modgil, *Norms and value...*, s. 29–64.

⁵⁸ *Ibidem*.

wanie zapadłego orzecznictwa czy to w pismach procesowych, czy też w orzecznictwie innych sądów, jeśli nie jest ono *contra legem*, nie budzi kontrowersji i stanowi obraz sposobu myślenia i stosowania technik doboru potrzebnej do przedstawienia danego stanowiska argumentacji prawniczej. Proceder ten zapewnia także stabilność rozumowania okołonormatywnego, mającą być wartością samą w sobie, wpływającą zasadniczo np. na stabilność norm społecznych, które prawa ma chronić i zabezpieczać. Nie sposób wskazać, jaki wpływ na kształtowanie się lini orzeczniczej ma analiza zapadłych orzeczeń, zwłaszcza w sytuacjach nieoczywistych, tj. takich, w których zapadłe orzeczenie nie jest bezpośrednio cytowane, po prostu nie koresponduje z omawianym stanem faktycznym. Jednak w sytuacjach, w których sąd wprost powołuje się na dane orzeczenie albo na utrwaloną linię orzecznictwa, oczywisty wydaje się ich wpływ na daną sprawę, w której owe orzeczenia są cytowane i do których sąd się odwołuje. Bazy orzeczeń istnieją w każdym systemie prawa – zwłaszcza takim, które nie było dotknięte zmianami ustrojowymi. Powstające wyszukiwarki i ich kwalifikowane postaci zdają się wprowadzać intuicyjne odniesienia do gromadzonego materiału zarówno w zakresie prawa, jak i orzecznictwa, a także w zakresie doktryny komentującej normy prawne i zapadłe orzecznictwo. Już sama budowa wyszukiwarek tego typu daje podstawy do tworzenia algorytmów mogących stanowić podstawę do tworzenia pozaludzkich procesów decyzyjnych przy ściśle określonym stanie faktycznym. W zasadzie zgodnie z regułami matematycznymi wskazać należy, że im więcej orzecznictwa na określony temat posiada system, tym jest ono bardziej reprezentatywne w przedmiotowym zakresie. Ta reprezentatywność mogłaby mieć wpływ na legitymizację określonych decyzji wydawanych przez „system”. Niewątpliwie takie rozwiązania mogłyby mieć wpływ na jedną z naczelných wartości prawa, a mianowicie na przewidywalność w procesie jego stosowania. W sprawach bardziej złożonych, których złożoność w tym kontekście polega przede wszystkim na tym, że norma prawna i zgromadzone orzecznictwo nie odpowiadają w pełni na złożoność stanu faktycznego przedstawionego do rozstrzygnięcia, tj. historycznie subsumpcja w tym zakresie nie miała miejsca lub miała miejsce, ale nie w odniesieniu do identycznego stanu faktycznego, należałoby tak zaprojektować algorytmy stosujące dane normy do określonego stanu faktycznego, aby mogły one w jak najpełniejszym zakresie znaleźć zastosowanie w danej sprawie. Poprzez gromadzenie wskazanych wyżej danych tworzy się taka baza danych, której parametry umożliwiają ich przetworzenie w tym stopniu i w takich zakresach, które dają podstawę do użycia kwantyfikatora wielomianowego Bayesa, przy użyciu funkcji językowych skierowanych na automatyczną identyfikację opisu stanów faktycznych umożliwiających stosowanie określonych i adekwatnych norm prawnych. Pozytywne wyniki realizacji tych zadań byłyby dowodem na to, że system jest w stanie zaadaptować informację napływającą od człowie-

ka i przetworzyć ją w celu wydania orzeczenia „sądowego” uwzględniającego stan faktyczny i podstawę prawną wynikającą z procesu zastosowania adekwatnej normy prawnej.

Należy w tym miejscu odnieść się do budowy systemu prawa *sensu largo*. Na gruncie prawa polskiego będą to źródła prawa powszechnie obowiązującego, w tym prawa lokalne, orzeczenia mające moc zasad prawa, pozostałe orzecznictwo, literatura przedmiotu oraz ewentualnie uzasadnienia do projektów aktów prawnych, które weszły w życie – to w szerszej formie interpretacji, w której poszukuje się znaczenia *ratio legis* interpretowanego przepisu. Natomiast na gruncie prawa *common law* wskazać należy, że jako pierwszorzędne źródła postrzega się ustawę (przepis prawa i orzecznictwo), drugorzędne to literatura i komentarz do ustawy (szerzej prawa), natomiast trzeciorzędne to digesta i paremie odzwierciedlające ogólny sens całych bloków regulacyjnych – na tle polskiej nomenklatury, powiedzieć by można, instytucji prawnych⁵⁹. Jakkolwiek by nie postrzegać owych zbiorów, wszystkie one dla danego systemu wchodziłyby w bazę, która dokonywałaby subsumcji stanów faktycznych na potrzeby wymiaru sprawiedliwości⁶⁰. W tym miejscu powstaje pytanie, co ze sprawami, które nie są przedmiotem szczegółowego orzecznictwa ani nie stanowią interesującego dla doktryny przedmiotu analizy. Zgodzić się wypada, że dla pełnego zastosowania omawianego systemu musiałby on poradzić sobie także i z tymi aspektami spraw⁶¹. Ta sama konstatacja dotyczy zasadniczo sytuacji,

⁵⁹ O. Shulayeva, A. Siddharthan, A. Wyner, Recognizing cited facts and principles in legal judgements, *Artif Intell Law* 2017, Nr 25, s. 107–126.

⁶⁰ Obecnie największymi dostawcami takich treści są LexisNexis i West-Law – za O. Shulayeva, A. Siddharthan, A. Wyner, Recognizing..., s. 110.

⁶¹ Zauważyć wypada, że pierwsze próby usystematyzowania informacji o cytowaniu innych dzieł bądź dokumentów zostały przeprowadzone w dziedzinie prawa przez Franka Sheparda w 1873 r., który polegając na zebranej analogowo wiedzy, w tym ze słowników, encyklopedii, czasopism etc., starał się usystematyzować cytaty pochodzące z orzecznictwa. Reminiscencją tej pracy jest serwis Citations Service LexisNexis Shepard. Pomimo że prawnicy byli pionierami analizy cytowań (Ogden 1993 r.), badania nad analizą cytowań w prawie nie rozwijały się tak szybko jak analiza cytowań dzieł w innych dziedzinach naukowych. Garfield (i jego praca z 1955 r.) jest często wymieniany jako jeden z pionierów i głównych uczestników analizy cytatów w nauce. Garfield zainspirowany pracami nad cytatami Sheparda twierdził, że podobne metody mogą być przydatne do podsumowania cytowania dzieł naukowych (Garfield 1955 r.). Garfield zastosował podejście bibliograficzne do tworzenia indeksów cytowań ICI, a dane z indeksów cytowań były później wykorzystywane do wielu badań bibliometrycznych (Moed 2005 r.). Uważał, że analiza cytowania może być wykorzystana do oceny osiągnięć naukowych, np. przy klasyfikacji czasopism na podstawie częstotliwości cytowania. Jak zauważył Moed (2005 r.), dane ilościowe z badań bibliometrycznych są szeroko stosowane do oceny wyników poszczególnych naukowców, czasopism naukowych, instytucji badawczych oraz ogólnych, strukturalnych aspektów systemu naukowego (np. pomiaru trendów w krajowym nakładzie publikacji). Niedawno wzrosło zainteresowanie studiami nad procesem cytowania dzieł, gdzie wydają się tworzyć dwa zasadnicze kierunki: zastosowanie analizy sieci do cytowań (Zhang i Koppaka 2007 r., Leicht i in. 2007 r., Winkels i in., Lupu i Voeten 2012 r., van Opijnen 2012 r., Neale 2013 r.) oraz systemy klasyfikacji umożliwiających oszacowanie statusu cytowanego przypadku (Jackson i in. 2003 r., Galgani i in. 2015 r.). „Modną” metodą jest także podejście oparte na RFC (powód powołania – *reason for citation*), opatentowanej technologii, która pozwala ustalić powody, dla których przytaczano dany fragment (Humphrey i in. 2005 r.).

w których dochodzi do wykładni systemowej albo funkcjonalnej – tu problematyka zaprojektowania logarytmów, które będą w stanie poradzić sobie z takimi sposobami interpretacji prawa, wydaje się tworzyć największe wyzwania dla budowy sztucznej inteligencji w procesie stosowania prawa.

Skoło zapadłoby orzeczenie w ramach stosowania prawa wydane przez sztuczną inteligencję, to i pojawiłaby się problematyka uzasadnienia tego orzeczenia. Formalne modele argumentacji stały się przedmiotem zainteresowania różnorodnych zespołów pracujących nad sztuczną inteligencją. Głównymi zagadnieniami, wokół których skupia się generalna uwaga badaczy, są: rozumowanie monotoniczne, systemy wielopodmiotowe i tworzenie uzasadnień prawnych. Ramy argumentacji stosowane na potrzeby tworzenia uzasadnień o charakterze prawnym stanowią także szczególny sposób definiowania skutków niemonotonicznych przez konstruowanie i porównanie argumentów oraz określenie ich wzajemnej pozycji w zbiorze, który tworzą na potrzeby danego uzasadnienia, jak też w relacji do innych uzasadnień w celu zapewnienia spójności w wartościowaniu argumentacyjnym. Jak się wydaje, szczególnie skuteczne i powszechnie stosowane są abstrakcyjne ramy aranżacji wprowadzone przez P.M. Dunga (1995 r.). Te ramy są zasadniczo wykresami, których węzły reprezentują argumenty. Treść tych argumentów nie jest dalej analizowana. Linki na wykresach do tych ram przedstawiają argumenty. Intuicyjnie, w takim szablonie argument zostaje zaakceptowany, chyba że zostanie on uwzględniony w innym zaakceptowanym argumentcie. Dunga zdefiniował kilka zasad semantyki formalizujących różne nauki oraz wskazał, które argumenty można zaakceptować w oparciu na konkretnych ramach, w szczególności, jak to nazywa, ugruntowaną, preferowaną i stabilną semantykę. Ramy Dunga są bardzo przydatne jako narzędzia analityczne do porównywania różnych form argumentacji. Są one powszechnie używane jako systemy docelowe do dokonywania tłumaczenia z mniej abstrakcyjnych systemów formalnych. Te formalizmy następnie transponują, poprzez tłumaczenie na ramy Dunga, semantykę. Wielu naukowców, takich jak A. Wyner, C. Bench-Capon, i T. Dunne (2009 r.) czy H. Prakken (2009 r.), podążyło za tym podejściem i zbudowało systemy argumentacji oparte na pracach Dunga. Jednak oczywiste jest, że ramy Dunga nie mają pewnych cech, które są powszechnie stosowane w niemal każdej formie argumentacji. Jedną z takich funkcji są standardy dowodowe. Literatura rozumowania prawnego jest pełna dyskusji na temat standardów, takich jak siła dowodów, przewaga dowodów, ich jakość i moc wykluczeniowa w stosunku do innych itp. Również w codziennym stosowaniu standardów dowodowych odgrywają zasadniczą rolę, np. w sytuacjach związanych z ryzykiem stosujemy wyższe standardy niż w przypadkach, w których go nie ma, tj. niejako bardziej zwracamy uwagę np. na prawdopodobność kontrahenta. Warunki akceptacji w ramach Dunga obejmują

bowiem dowolną funkcję określającą stan węzła w oparciu na status węzłów nadrzędnych. Obejmuje to warunki uzależnione od twierdzeń, a tym samym wykracza poza kilka dobrze znanych standardów prawnych. Ten stan zostanie osiągnięty na podstawie niektórych właściwości/typów łączących się na poszczególnych wykresach. Warunki akceptacji pozwalają na wprowadzenie różnych typów węzłów i łączy. Wpływ węzła na inny węzeł jest określony całkowicie przez warunki akceptacji. W przeciwieństwie do ram Dunga, w których łączy reprezentują jeden szczególny rodzaj relacji, a mianowicie argument, gdzie węzły są akceptowane, chyba że zostały one niejako wyargumentowane, inne ramy pozwalają na elastyczne odwzorowanie różnych odmian argumentacji: mogą one mieć węzły, które nie są akceptowane bez dodatkowego uzasadnienia oraz odpowiadające im łączy. Możemy też mieć linki (łączniki) o różnych wartościach argumentacyjnych, a nawet linki (łączniki), które wspierają lub niweczą węzeł w zależności od kontekstu. Poniższy wzór charakteryzuje to podejście:

abstrakcyjne dialektyczne ramy = wykresy zależności
+ warunki akceptacji.

Podsumowanie

Ważnym aspektem jest uogólnienie standardowej semantyki Dunga. Okazuje się, że ugruntowana semantyka może rzeczywiście być uogólnieniem do arbitralnych metod dialektycznych. Do ustanowienia stabilnej i preferowanej semantyki potrzebna jest koncepcja wsparcia i ataku, która jest obecna w nieco ograniczonej, ale wciąż silnej i elastycznej klasie ram, którą nazywamy dwubiegunową. W celu odzwierciedlenia tej ogólności występuje pojęcie „struktur dwukierunkowych”. W filozofii klasycznej dialektyka jest formą rozumowania opartą na wymianie argumentów i argumentów przeciwstawnych, opowiadając się za zdaniami (wnioskami – tezami) i przeciwstawnymi (antytezy). W strukturach Dunga natomiast argument zostaje zaakceptowany, chyba że zostanie zaatakowany przez zaakceptowany węzeł. Jak wspomniano powyżej, w tych strukturach Dunga opiera się całkowicie na warunku akceptacji, tj. rozstrzyga, czy węzeł jest akceptowany, czy nie. Nadal możemy mieć węzły, które zachowują się jak argumenty – w gruncie rzeczy są to szczególne przypadki ram, ale też mogą istnieć węzły, które są odrzucane, jeśli nie są obsługiwane przez niektóre akceptowane węzły⁶².

⁶² G. Brewka, S. Woltran, Abstract Dialectical Frameworks..., s. 15–16.

Słowa kluczowe: sztuczna inteligencja, stosowanie prawa, algorytmiczna analiza dowodów, definiowanie wartości, abstrakcyjne ramy dialektyczne.

Artificial intelligence and singularity in the process of applying law

There is no question that certain elements of information and communications technology entered the Polish system of law and other legal systems of European countries. In Poland, it is worth mentioning regulations of the act of 5.9.2016, on trust services and electronic identification. This act is to serve the implementation of the regulation of the European Parliament and the Council no 910/2014 of 23.7.2014, regarding electronic identification on external market and amending directive 1999/93/EC. A question arises: are we going into uncharted territory, which is application of artificial intelligence in the process of law, and maybe in the future also making law, or is the abovementioned amendment the first step towards that direction, or may be one in the future?

Keywords: artificial intelligence, applying law, algorithmic analysis of evidence, defining values, abstract dialectical frames.



Monografie Prawnicze

MONOGRAFIE PRAWNICZE

**OCHRONA KLIENTA NA RYNKU
USŁUG FINANSOWYCH
W ŚWIELE AKTUALNYCH
PROBLEMÓW
I REGULACJI PRAWNYCH**

Pod redakcją
EDYTY RUTKOWSKIEJ-TOMASZEWSKIEJ

C.H. Beck
Wydawnictwo C.H. Beck

www.ksiegarnia.beck.pl

Zadzwoń: 81 46 13 300 • E-mail: kontakt@beck.pl