Mariusz Bogusz*
DOI: 10.34616/150987

Spotkanie tradycji z technologią, czyli garść refleksji na temat wykorzystania generatywnej sztucznej inteligencji w naukach prawnych

Dzieło Profesora Jerzego Supernata pt. *Administracja i prawo administracyjne. Słownik polsko-angielski. Krótki słownik terminów środowiska administratywistów*[1] jest książką niezwykłą. Pod skromnym tytułem kryje się imponujący (blisko ośmiusetstronicowy!), opracowany nakładem benedyktyńskiej pracy jednego człowieka (!), prawniczy słownik polsko-angielski (część pierwsza książki) oraz wciągający, nakreślony słowem oraz obejmujący kilkadziesiąt lat historii, portret środowiska polskich administratywistów. Portret złożony ze szkiców sylwetek ludzi i naukowych (choć nie tylko) spraw, którymi żyli, stopklatek z wydarzeń i błysków anegdot (część druga książki). Wszystko to nakreślone lekkim piórem uważnego, przenikliwego, ale niezwykle życzliwego ludziom Obserwatora i Niestrudzonego Kronikarza. Można wręcz powiedzieć, że czego nie ma w książce Profesora Jerzego Supernata, tego w polskiej administratywistyce nie było w ogóle. Po prostu się nie zdarzyło, bo nieopisane nie istnieje.

Słownik polsko-angielski Profesora Jerzego Supernata zawiera również hasło „sztuczna inteligencja – *artificial intelligence (AI)*"[2]. Termin ten, mimo że używany w naukowym dyskursie od kilkudziesięciu lat, prawdziwie zawrotną karierę robi od 30 listopada 2022 r., kiedy to amerykańskie laboratorium badawcze pod nazwą OpenAI, udostępniło w internecie usługę pod nazwą *Chat GPT* w wersji 3,5. Od tego momentu każdy użytkownik sieci może, na zasadzie otwartego dostępu (po utworzeniu konta i zalogowaniu się), prowadzić „dialog z maszyną", na podobieństwo dialogu prowadzonego z człowiekiem, w trybie „pytanie – odpowiedź". Chat GPT jest w stanie podać przepis na potrawę, przetłumaczyć tekst na dowolny język, rozwiązać problem życiowy lub naukowy z dowolnej dziedziny wiedzy, pocieszyć, napisać przemówienie, wymyślić dow-

* Dr hab., prof. Uniwersytetu Gdańskiego, ORCID: 0000-0002-3173-079X.

[1] J. Supernat, *Administracja i prawo administracyjne. Słownik polsko-angielski. Krótki słownik terminów środowiska administratywistów*, Wrocław 2022.

[2] *Ibidem*, s. 637.

cip itd. Chat GPT ma także swoje „rodzeństwo", tj. podobne aplikacje stworzone przez inne firmy[3], tym niemniej karierę Chata GPT, sądząc m.in. po masowej skali zainteresowania aplikacją, liczbie publikacji na jej temat oraz intensywności uniwersyteckich dyskusji jej poświęconych, należy uznać za najbardziej widowiskową. Chat GPT (i jego „rodzeństwo") są rodzajową postacią sztucznej inteligencji, a mianowicie jej postacią generatywną (dalej: generatywna AI), tj. taką, która umożliwia tworzenie nowych tekstów, obrazów, dźwięków i innych treści w oparciu o dane, na podstawie których jest szkolona. Z kolei sama sztuczna inteligencja (bez kwantyfikatora: generatywna) to – najogólniej rzecz ujmując (ponieważ ścisłej i powszechnie akceptowanej definicji brak) – system komputerowy (maszyna i oprogramowanie) zdolna do ciągłego udoskonalania się oraz samodzielnego/autonomicznego wykonywania zadań.

Sztuczna inteligencja doczekała się także swojej prawnej definicji i regulacji – pierwszym aktem regulującym ją w szerokim zakresie jest rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (dalej jako „rozporządzenie w sprawie sztucznej inteligencji"), a jego stosowanie, z pewnymi wyjątkami, ma się rozpocząć 2 lata po wejściu w życie[4].

Przepis art. 3 pkt 1 rozporządzenia w sprawie sztucznej inteligencji definiuje „system AI" jako „system maszynowy, który został zaprojektowany do działania z różnym poziomem autonomii po jego wdrożeniu oraz który może

[3] Należy tutaj wymienić aplikacje: *Copilot* (twórca: *Microsoft*), *Gemini* (twórca: *Google*), *Grok* (twórca: *xAI Elona Muska*) – zob. Ł. Lamża, *Zajrzyjmy pod maskę sztucznej inteligencji*, „Tygodnik Powszechny" 2024, nr 14, 3–9 kwietnia 2024 r., s. 12.

[4] Dz. Urz. UE L 2024/1689 z 12.07.2024 r. W kontekście prawa Unii Europejskiej i organów UE należy odnotować oryginalne prace J. Supernata wskazujące na potrzebę obejmowania zakresem znaczeniowym pojęcia administracja publiczna również organów administracji europejskiej, co – zdaniem tego Autora – w dalszej perspektywie otworzy pole do badań tej administracji w „europejskiej przestrzeni naukowej przez międzynarodowe zespoły w globalnym języku" (zob. J. Supernat, *Założenia nauki administracji Franciszka Longchampsa de Bérier i administracja unijna*, „Państwo i Prawo" 2019, nr 5, s. 79–86; *idem*, *Komisja Europejska jako administracja publiczna*, [w:] A. Matan (red.), *Administracja w demokratycznym państwie prawa. Księga jubileuszowa Profesora Czesława Martysza*, Warszawa 2022, s. 712–725). Nie będzie to jednak zadanie proste – z uwagi na zróżnicowanie języków UE, zróżnicowanie terminologii języków prawnych, a także odmienność znaczeń nadawanych nawet takim samym terminom występującym w językach prawnych prawa UE i państw członkowskich (J. Supernat, *Sądy i kontrola sądowa w Unii Europejskiej*, [w:] W. Jakimowicz, M. Krawczyk, I. Niżnik-Dobosz (red.), *Fenomen prawa administracyjnego. Księga jubileuszowa Profesora Jana Zimmermanna*, Warszawa 2019, s. 871–872).

wykazywać zdolność adaptacji po jego wdrożeniu, a także który – na potrzeby wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie otrzymanych danych wejściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne". Trafnie wskazuje D. Flisak, że przyjęta w rozporządzeniu definicja „została oparta na kluczowych cechach odróżniających systemy AI od tradycyjnych systemów oprogramowania. Należą do nich zdolność do wnioskowania, autonomiczne działanie czy zdolność adaptacyjna. Na podstawie tych cech można stwierdzić, że zakres omawianego pojęcia nie obejmie w szczególności zautomatyzowanych systemów wykonujących określone operacje"[5].

Generatywna AI będąca – o czym była już mowa – rodzajową postacią sztucznej inteligencji istotnie robi wrażenie, zwłaszcza w tej jej odmianie, która – jak Chat GPT – imituje dialog z drugim człowiekiem: jedno zdanie wypowiadasz (piszesz) ty, rozmówca odpowiada (na temat!), potem znowu ty, potem znowu rozmówca... Iluzja rozmowy jest pełna.

W tym stanie rzeczy zasadne jest postawienie pytania o ewentualność wykorzystania generatywnej AI w naukach prawnych. Przechodząc do konkretów – rzecz sprowadza się do zagadnienia, czy jeżeli np. wyznaczymy generatywnej AI zadanie: „sformułuj aktualny problem prawny z zakresu gałęzi prawa [tu pada nazwa gałęzi]" albo „podaj wykaz pozycji literaturowych dotyczących problemu [tu zostaje wskazany problem]", albo „przedstaw relację przepisu art. … ustawy [tu pada nazwa ustawy] do przepisu art. … ustawy [tu pada nazwa ustawy]", albo „zreferuj główne tezy monografii [tu pada nazwisko autora i tytuł monografii]", to możemy się po generatywnej AI spodziewać prawidłowego wykonania tego zadania i prawidłowego wyniku, czy też nie.

Tradycyjne, dotychczas stosowane prawnicze programy komputerowe, oparte na stosunkowo prostych algorytmach, nie oferują tak zaawansowanych funkcji jak generatywna AI. P. Fik i P. Staszczyk wyjaśniają, że owe proste algorytmy „wykonują przydzielone im z góry zadania poprzez otrzymane instrukcje – linijki kodu", natomiast „sztuczna inteligencja to coś więcej, ponieważ oprócz wykonywania przypisanych jej zadań ma ona w swoim genie także dążność do ciągłego udoskonalania się i wyboru najefektywniejszych rezultatów na przyszłość (*machine learning*)"[6]. W konsekwencji prawnicze programy komputerowe, oparte na owych prostych algorytmach, realizując instrukcje wprowadzone

[5] D. Flisak, *3. „System AI" – podstawowe pojęcie Aktu*, [w:] *Akt w sprawie sztucznej inteligencji [projekt UE]*, Lex/el. 2024.

[6] P. Fik, P. Staszczyk, *Sztuczna inteligencja w unijnej koncepcji e-sprawiedliwości – teoria i możliwy wpływ na praktykę*, „Europejski Przegląd Sądowy" 2022 nr 7, s. 6.

przez użytkownika, zapewniają pewny, acz „niewyrafinowany" wynik. Jeżeli użytkownik nakaże programowi wyszukać w bazie danych określoną frazę (np. „rażące naruszenie prawa"), to program przeszuka bazę danych w celu odnalezienia tej frazy (ciągu znaków) i wyświetli wszystkie rekordy, w których ta fraza występuje. Program „powie", czy w bazie danych znajduje się ciąg znaków, którego użytkownik szuka i wskaże miejsca, w których ten ciąg znaków występuje. Tylko tyle, ale i aż tyle. Wynik automatycznego przeszukania określonej bazy danych jest przy tym pewny: algorytm niczego nie pominie, ale też i niczego nie „zmyśli".

Na podobnych, stosunkowo prostych algorytmach, umożliwiających weryfikację określonych czynności (pod kątem ich akceptacji, odrzucenia lub konieczności poprawienia) mogą być oparte niektóre, całkowicie lub częściowo zautomatyzowane, procedury sądowe lub administracyjne (np. niemieckie sądowe postępowanie upominawcze[7] czy – w Polsce – proces składania w formie elektronicznej sprawozdań finansowych przedsiębiorstw do Krajowego Rejestru Sądowego, postępowania rejestrowe, deklaracji podatkowych do urzędów skarbowych, deklaracji z zakresu ubezpieczeń społecznych i ubezpieczeń zdrowotnych do ZUS itp.).

Tymczasem generatywna AI oferuje funkcje bardziej zaawansowane, np. tworzy tekst, wkładając w wytworzony, a oczekiwany przez użytkownika wynik własny pierwiastek „twórczy". W tym miejscu wyłania się jednak pytanie o wiarygodność/prawdziwość owego wytworzonego przez generatywną AI wyniku. W celu udzielenia odpowiedzi na to pytanie konieczne jest wyjaśnienie zasady, na której opiera się współczesny mechanizm działania generatywnej AI. Jakkolwiek paradoksalnie by to nie brzmiało, generatywna sztuczna inteligencja jest osiągnięciem nauk… statystycznych, opartym na tzw. dużych modelach językowych (*large language model* – LLM). Mechanizm działania generatywnej sztucznej inteligencji przystępnie wyjaśnia Ł. Lamża[8]. LLM powstaje wtedy, gdy po zebraniu olbrzymiej próbki tekstów napisanych przez ludzi nakażemy komputerowi (z odpowiednim oprogramowaniem) poddanie tych tekstów analizie pod kątem ustalenia, które wyrazy występują najczęściej w jakim kontekście oraz wykrycia w tym zakresie powtarzalnych wzorców; w konsekwencji komputer jest w stanie – w oparciu o wcześniej obliczone w trakcie „uczenia się" prawdopodobieństwo – „odgadnąć" kolejność występowania wyrazów po sobie

[7] *Ibidem*, s. 7.

[8] Ł. Lamża, *op. cit.*, s. 13. Na temat mechanizmów działania generatywnej AI zob. także: M. Rotkiewicz, *Dżin już wyleciał. Rozmowa z prof. P. Kazienko*, „Polityka" 2023, nr 15, 5–11 kwietnia 2023 r., s. 16–18.

(jeden po drugim); jak pisze Ł. Lamża, „gdy tego typu model języka zastosujemy do problemu zgadywania następnego słowa w urwanym tekście, powstanie «algorytm generatywny»”[9].

Generatywna AI potrafi więc tworzyć poprawne pod względem językowym teksty na podstawie obliczonego wcześniej (podczas uczenia się/treningu na gigantycznej próbce prawdziwych tekstów) prawdopodobieństwa występowania kolejnych sylab/wyrazów. Wygenerowany tekst powstaje na skutek naśladownictwa opartego na prawdopodobieństwie[10].

Taka „statystyczna” metoda generowania tekstu zapewnia jego językową poprawność, natomiast nie gwarantuje jego merytorycznej sensowności ani prawdziwości. Tekst wytworzony przez generatywną sztuczną inteligencję może zawierać zdania sensowne i prawdziwe, ale również i bzdury. Bzdurne wytwory generatywnej AI określa się mianem halucynacji. Stanowią one nieuniknioną konsekwencję fundamentalnego założenia, jakie leży u podstaw generatywnej AI, a mianowicie oparcia jej na zasadzie prawdopodobieństwa[11]. Jeżeli prawdopodobieństwo jest orłem, to halucynacja jest reszką, ale są to dwie, nierozerwalnie ze sobą związane strony tego samego medalu.

Generatywna AI nie ma świadomości, nie myśli, nie rozumuje, nie wie, co jest prawdziwe, a co nie, a nawet nie rozumie znaczenia wyrazów i tekstów, które tworzy. Ona jedynie generuje ciągi znaków, których następowanie po sobie jest oparte na prawdopodobieństwie wyliczone na podstawie setek milionów czy miliardów rzeczywistych tekstów[12]. Nie zmienia to faktu, że człowiek – użytkownik generatywnej AI ma wrażenie – w warunkach pełnej iluzji – dia-

---

[9] *Ibidem*, s. 13.

[10] Prof. A. Mądry wyjaśnia, że „*Chat GPT* – najprościej mówiąc – stara się podać odpowiedź, która w miliardach tekstów z Internetu, jest najbardziej prawdopodobnym ciągiem słów” (J. Żakowski, *Frankenstein otwiera oczy. Rozmowa z prof. A. Mądrym*, „Polityka” 2024, nr 1–2, 1–9 stycznia 2024 r., s. 27). Na temat naśladowniczych „umiejętności” generatywnej AI zob. B. Paszcza, *Algorytmy jak papugi*, „Rzeczpospolita”, dodatek „Plus Minus” z 21 stycznia 2023 r.

[11] Halucynacji generatywnej AI nie eliminuje nawet odpowiednie sprofilowanie tekstowej „bazy treningowej”. Opracowana przez inżynierów firmy Meta (d. Facebook) sztuczna inteligencja pod nazwą Galactica miała zrewolucjonizować naukę. „Przeszkolono” ją bowiem na podstawie 48 mln tekstów naukowych. Ruszyła 15 listopada 2022 r. Program zamknięto po trzech dniach – okazał się bowiem „generatorem losowych bzdur” (A. Kaźmierska, M. Gikowski, *Ludzkość puszcza stery*, „Tygodnik Powszechny” 2023, nr 16, 16 kwietnia 2023 r., s. 15).

[12] T. Stawiszyński trafnie konstatuje, że „model językowy, jak sama nazwa wskazuje, zajmuje się wytwarzaniem zdań”, zaś podstawowym zadaniem sztucznej inteligencji jest „generowanie spójnych wypowiedzi”. Autor ten pisze dalej, że sztuczna inteligencja robi to „[…] analizując wypowiedzi już wcześniej sformułowane i wytwarzając na tej podstawie coś, co będzie akurat najlepiej pasowało do kontekstu. Otóż zwróćmy uwagę, że kwestia prawdy i fałszu w ogóle w tym polu nie powstaje” (T. Stawiszyński, *Imitacja*, „Tygodnik Powszechny” 2024, nr 10, 6–12 marca 2024 r., s. 32).

logu z drugim człowiekiem (przede wszystkim z uwagi na rzeczowy charakter dialogu i poprawność języka używanego przez „interlokutora"). M. Rotkiewicz przytacza głosy sceptyków, zdaniem których wręcz „daliśmy się uwieść programowi [Chat GPT – M.B.], który układając tekst, kieruje się wyłącznie prawdopodobieństwem, z jakim jeden wyraz następuje po poprzednim"[13].

Dopiero na tak zarysowanym tle jest możliwe udzielenie odpowiedzi na postawione wcześniej pytanie o wiarygodność wyniku wytworzonego przez generatywną AI w obszarze nauk prawnych oraz o jej przydatność jako narzędzia badawczego. Nie powinna być zaskoczeniem konstatacja, że przydatność generatywnej AI w obszarze nauk prawnych wydaje się wątpliwa. Tekstowy (bo w naukach prawnych liczą się teksty) wytwór generatywnej AI może bowiem odpowiadać rygorom naukowej prawdy, ale równie dobrze może być poprawną pod względem językowym halucynacją. Ryzyko jest realne, co potwierdza osobiste doświadczenie: generatywna AI „potrafi" przypisać konkretnym jednostkom tekstu prawnego (artykułom, paragrafom) treści, których te przepisy nie zawierają; stworzyć fikcyjnych autorów opracowań naukowych; podać wykazy nieistniejących pozycji literaturowych; sformułować fałszywe wnioski czy popełniać jeszcze inne błędy. Jako narzędzie pracy prawnika-naukowca generatywna AI jest zatem narzędziem zawodnym. „Broni się" ona jedynie na poziomie poprawności językowej wytworzonego tekstu; jest to jednak niewystarczająca kwalifikacja dla potrzeb uznania jej za pewne i wiarygodne narządzie pracy prawnika-naukowca.

Po to, by dokonać oceny wytworu generatywnej AI w kategoriach: prawda albo fałsz, wywód logicznie poprawny albo nonsensowny, należy uprzednio zgromadzić wiedzę i posiąść umiejętności za pomocą tradycyjnych metod. Weryfikacja wytworu generatywnej AI wiąże się zatem z nakładem czasu i wysiłkiem, którego korzystanie z generatywnej AI miało akurat człowiekowi oszczędzić. Przypomina to trochę problem błędnego koła.

Jest przy tym wysoce wątpliwe, czy mechanizm generatywnej AI mógłby być – dla potrzeb pracy prawnika-naukowca – zasadniczo udoskonalony. Niezmieniona pozostanie bowiem sama zasada działania generatywnej AI, którą jest zasada prawdopodobieństwa, a nie zasada prawdy/fałszu oraz logicznego/nielogicznego rozumowania. Do dokonywania ocen w tym zakresie niezbędny jest bowiem ludzki mózg.

Wydaje się, że w obrębie nauk prawnych bardziej efektywne i pewniejsze jest wykorzystywanie opisanych wyżej stosunkowo prostych algorytmów, któ-

[13] M. Rotkiewicz, *AGI: trzy litery, które wstrząsną światem. Nadciąga ogólna sztuczna inteligencja*, „Polityka" 2024, nr 7, 7–13 lutego 2024 r., s. 62.

rych funkcje są co prawda mniej wyrafinowane niż funkcje generatywnej AI, ale za to umożliwiają osiągnięcie pewnych, wolnych od halucynacji wyników. Tytułowe spotkanie, a w istocie rzeczy pojedynek między tradycją a technologią w postaci generatywnej AI, należy – w obrębie nauk prawnych – rozstrzygnąć na korzyść tradycji.

Na zakończenie rozważań, w ramach *licentia poetica* właściwej formie eseju, a przez to wolnej od rygorów rozprawy naukowej, ośmielę się – jako laik – sformułować prognozę. Prognozę bezpieczną, bo – w mojej ocenie – nieweryfikowalną w dostępnym horyzoncie czasowym. Otóż sztucznej inteligencji będącej odpowiednikiem ludzkiego umysłu człowiekowi nie uda się stworzyć nigdy. Przyczyna tego stanu rzeczy jest dosyć prozaiczna: stworzenie sztucznego tworu na wzór ludzkiego mózgu wymagałoby posiadania przez człowieka pełnej wiedzy o ludzkim mózgu. Tutaj sięgamy jednak granic poznania, których jako ludzie nie jesteśmy i nie będziemy w stanie przekroczyć. Nie chodzi przy tym o wiedzę o anatomii/strukturze mózgu (bo pod tym względem wiedza jest już – jak się wydaje – zupełna lub prawie zupełna), ale wiedzę o jego funkcjonowaniu[14]. W tym zakresie postęp wiedzy naukowej na drodze, która liczy miliardy lat świetlnych, liczyć należy w mikronach. Funkcjonowanie mózgu jest i pozostanie tajemnicą. Zawartości mózgu (rozumianej jako suma świadomości, osobowości, pamięci, doświadczeń, postaw i wielu innych elementów wraz z mechanizmami ich działania) nie da się zapisać w postaci ciągów zer i jedynek ani skopiować, ani przenieść do drugiego „komputera", żeby działała tak samo jak w pierwszym „komputerze". Ta CAŁOŚĆ wymyka się granicom poznania. Nasuwają się tutaj oczywiste analogie z pytaniami o początek Wszechświata: jak i skąd wzięło się to pierwsze COŚ, z którego powstała cała RESZTA. Fizycy, mimo nieustannie podejmowanych prób i budowanych teorii, mówią, że i tak na końcu stają wobec niewytłumaczalnej Tajemnicy, wobec Czegoś, co się w ramach ludzkiej świadomości, wiedzy, a nawet wyobraźni, nie mieści i ich granice przekracza. Jest niepoznawalne. Inteligencji sztucznej nie da się zbudować, bo jej oryginał jest niepoznawalny.

Oczywiście, czysto subiektywna prognoza laika nie jest równoznaczna z tezą o daremności wysiłków zmierzających do stałego poszerzania granic poznania i docierania do prawdy ani też z tezą o niecelowości tworzenia przez człowieka coraz lepszych narzędzi służących dobrostanowi jego i planety, w tym także tzw. sztucznej inteligencji. Jej rozwijanie jest celowe i uzasadnione

[14] Na temat wysiłków badawczych neuroinformatyków zob. W. Brzeziński, *Nie czytamy w myślach. Rozmowa z prof. Piotrem Durką, fizykiem biomedycznym*, „Katalog Festiwalowy Copernicus Festiwal 2024", dodatek do „Tygodnika Powszechnego" 2024, nr 20, s. 36–39.

(a przy tym nie do zatrzymania), choć termin użyty do jej nazwania ma w sobie więcej marketingowej przesady aniżeli nomenklatury odzwierciedlającej realia; tzw. sztuczna inteligencja – „odczarujmy" chociaż tę „tajemnicę" – nie jest bowiem żadną syntetyczną, krzemową, sztuczną świadomością, lecz jedynie skomplikowanym i zaawansowanym programem komputerowym.

Człowiek jest jednak tak skonstruowany, że zawsze będzie się starał sięgnąć gwiazd i będzie się do tego celu stale zbliżał. Będzie także doskonalił tzw. sztuczną inteligencję, by dorównała ona tej prawdziwej (co nigdy nie nastąpi). Niech człowiek to czyni – byle ze świadomością złożoności rzeczy i zjawisk oraz z pokorą, która nie stoi w sprzeczności z ambicją.