

**A U T O R E F E R A T**  
**PRZEDSTAWIAJĄCY DOROBEK**  
**I OSIĄGNIĘCIA HABILITANTA**

Niniejszy autoreferat został sporządzony zgodnie z art. 16 i 18a ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (tekst jedn. Dz. U z 2016 r., poz. 882, ze zm.) oraz § 12 ust. 2 rozporządzenia Ministra Nauki i Szkolnictwa Wyższego z dnia 26 września 2016 r. w sprawie szczegółowego trybu i warunków przeprowadzania czynności w przewodzie doktorskim, w postępowaniu habilitacyjnym oraz w postępowaniu o nadanie tytułu profesora (Dz. U. z 2016 r., poz. 1586).

Autoreferat zawiera opis dorobku i osiągnięć naukowych habilitanta, w szczególności określonych w art. 16 ust. 2 powołanej ustawy.

**1. Posiadane dyplomy i stopnie naukowe**

Studia na Wydziale Prawa Uniwersytetu Warszawskiego ukończyłem w 1972 z wynikiem bardzo dobrym i otrzymałem dyplom na podstawie pracy magisterskiej „Kara śmierci na świecie”. W tym samym roku rozpocząłem pracę jako asystent stażysta w Zakładzie Statystyki i Demografii, którym kierował prof. Z Rogoziński. W roku 1979 uzyskałem tytuł doktora nauk prawnych na podstawie pracy doktorskiej o tytule „Komponenty statystyczne w zautomatyzowanym wyszukiwaniu informacji prawnej”. Moim promotorem był prof. F. Studnicki z Uniwersytetu Jagiellońskiego, a recenzentami prof. J. Wróblewski (UŁ) oraz prof. Andrzej Mrózek (UT). Praca uzyskała III nagrodę w konkursie Państwa i Prawa na najlepszy doktorat w Polsce w roku 1979.

**2. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych**

Od początku pracy zawodowej jestem zatrudniony na Uniwersytecie Warszawskim kolejno na stanowiskach:

od 1.10.1972 do 30.09.1973 **asystenta stażysty** na Wydziale Prawa Uniwersytetu Warszawskiego w Zakładzie Statystyki i Demografii w Instytucie Nauk Administracyjnych

od 1.10.1973 do 30.09.1978 **asystenta** na Wydziale Prawa Uniwersytetu Warszawskiego w Zakładzie Statystyki i Demografii w Instytucie Nauk Administracyjnych

(w tym od 1.10.1974 do 30.09.1977 skierowany byłem na **studia doktoranckie**)

od 1.10.1978 do 30.09.1979 **starszego asystenta** na Wydziale Prawa Uniwersytetu Warszawskiego w Zakładzie Statystyki i Demografii w Instytucie Nauk Administracyjnych

od 1.10.1979 do 30.09.2008 **adiunkta** na Wydziale Prawa i Administracji Uniwersytetu Warszawskiego w Zakładzie Logiki i Informatyki Prawniczej w Instytucie Nauk o Państwie i Prawie (po przekształceniu Zakładu i zmianie jego podporządkowania)

od 1.10.2008 do chwili obecnej **docenta** na Wydziale Prawa i Administracji Uniwersytetu Warszawskiego w Zakładzie Logiki i Informatyki Prawniczej

w latach 1993 – 1999 pełniłem na Wydziale Prawa i Administracji funkcję **kierownika studiów wieczorowych**

w roku 1999 wybrany zostałem na **Prodziekana ds. Studenckich**, którą to funkcję pełniłem do roku 2005 (dwie kadencje)

### **3. Osiągnięcie naukowe w rozumieniu art. 16 ust. 2 ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki**

Jako osiągnięcie naukowe w dyscyplinie naukowej prawo, w rozumieniu art. 16 ust. 2 ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki pragnę wskazać monografię:

#### **„SYSTEMY WYSZUKIWANIA INFORMACJI PRAWNEJ”**

Warszawa 2017, Wydawnictwo Wolters Kluwer s. 488

ISBN 978-83-8124-255-4,

ISSN 1897-4392

## Uzasadnienie wyboru tematu pracy

Przedstawiona do oceny w postępowaniu habilitacyjnym, będąca monografią, książka „Systemy wyszukiwania informacji prawnej”, jest pracą z zakresu informatyki prawniczej. Podejmowana w książce problematyka dotycząca budowy komputerowych systemów „wyszukiwawczych” była już od początków powstania tej dyscypliny nauki, tj. od końca lat 50 – tych XX wieku, uznawana za podstawową. Jest tak również współcześnie. Wprawdzie w pewnym okresie czasu zwłaszcza w latach 80 – tych ubiegłego stulecia można było zauważyć pewien zastój w pracach badawczych wiążących się z problematyką budowy systemów wyszukiwawczych, co zostało zresztą określone przez J. Binga jako „unsatisfactory state of art.”, to już od początków XXI stulecia doszło do zmiany tej sytuacji. Zagadnienie wiążące się z budową tych systemów stały się znów centralnym punktem zainteresowania informatyki prawniczej. Stan ten E. Shweighofer określił jako „empire stikes back”.

Problematyka poruszana w monografii stanowi od wielu lat przedmiot moich zainteresowań naukowych, co znalazło wyraz w tematyce mojej rozprawy doktorskiej „*Komponenty statystyczne w zautomatyzowanym wyszukiwaniu informacji prawnej*”, publikacji szeregu artykułów, a także w wydanej w 1999 r. książce „*Informatyka prawnicza. Zagadnienia teorii i praktyki*”, której część poświęcona była opisowi stosowanych przy budowie systemów wyszukiwania informacji prawnych w tamtym czasie rozwiązań. Nie była jednak poświęcona wyłącznie systemom wyszukiwania informacji i dotyczyła także problematyki występującej w ramach innych działów informatyki prawniczej wiążących się z budową systemów automatycznego rozstrzygania, czy też systemów ekspertowych. Obecnie wydana monografia koncentruje się wyłącznie na problematyce budowy systemów wyszukiwania informacji i porusza obok problemów wiążących się z ich tworzeniem i organizacją również szereg innych istotnych dla funkcjonowania tych systemów zagadnień takich jak dla przykładu analiza sposobów ich realizowanie za pomocą modeli informacyjnych. Uzasadnieniem dla takiego doboru tematyki pracy było również to, że w polskiej literaturze informatyczno-prawnej szczegółowemu omówieniu problematyki budowy tego typu systemów jak dotychczas nie poświęcono zbyt wiele miejsca. Tymczasem problematyka ta jest dla informatyki prawniczej problematyką podstawową, co wiąże się z tym, że prawnik współcześnie nie jest w stanie wykonywać swojego zawodu bez posługiwania się komputerowymi systemami wyszukiwania informacji prawnej.

## **Cel pracy**

Celem książki było kompleksowe przedstawienie problematyki dotyczącej zarówno roli, miejsca i znaczenia komputerowych systemów wyszukiwania informacji prawnej dla funkcjonowania porządku prawnego jak też prezentacja zagadnień wiążących się z konstrukcją tych systemów i perspektywami ich rozwoju. Informatyka prawnicza w ostatnich dwudziestu latach przeżywała i przeżywa nadal okres intensywnego rozwoju, co przejawia się w powstawaniu szeregu nowych aplikacji, a także wielu nowych koncepcji wiążących się z tworzeniem systemów wyszukiwawczych, co znajduje wyraz w konstruowaniu systemów o tzw. nieklasycznym charakterze, w tym systemów wykorzystujących sztuczną inteligencję. Istotna część monografii poświęcona jest omówieniu tych zagadnień.

## **Struktura pracy**

Struktura monografii została podporządkowana wskazanym wyżej celom badawczym.

Rozdział I książki dotyczy przedmiotu i zadań informatyki prawniczej, a także ukazania miejsca jakie zajmuje w ramach tej dyscypliny problematyka budowy systemów wyszukiwania informacji prawnej. Rezultatem przeprowadzonych w ramach tego rozdziału rozważań jest sformułowanie definicji informatyki prawniczej jak i określenie jej współczesnej struktury, co pozwoliło na ukazanie miejsca jakie zajmuje w jej ramach problematyka budowy systemów wyszukiwawczych. Przedmiotem szczegółowych rozważań jest w tym rozdziale określenie wzajemnego stosunku jaki zachodzi między informatyką prawniczą a prawem informatycznym. Przedstawione zostały poglądy, głównie nauki polskiej, lecz także anglosaskiej i niemieckiej na przedmiot informatyki prawniczej. W rezultacie uznano, że informatyka prawnicza i prawo informatyczne powinny być uznane za odrębne dyscypliny nauki. Podjęta została również polemika z poglądami wiążącymi się z koncepcją tworzenia nowej dyscypliny wiedzy, jaką ma być informatyka prawa, której autorem jest J. Janowski. W rezultacie przedstawiony w monografii pogląd uznaje, iż informatyka prawnicza jest dziedziną w stosunku do tzw. informatyki prawa odrębną. Podejście to nie neguje konieczności prowadzenia badań o charakterze interdyscyplinarnym nie oznacza jednak akceptacji poglądu, w myśl którego informatyka prawnicza powinna być uznawana jedynie za jedną z płaszczyzn informatyki prawa. Tego typu rozumienie informatyki prawniczej pozwoliło na jej zdefiniowanie i uznanie, że informatyka prawnicza jest dziedziną badającą możliwości stosowania, stosowanie i pozaprawne skutki stosowania komputerów na gruncie prawa. Tego

typu zdefiniowanie informatyki prawniczej nawiązuje do tradycyjnego jej ujęcia, prezentowanego przez F. Studnickiego.

Jak chodzi o strukturę informatyki prawniczej wyodrębnione zostały cztery jej działy. Pierwszy z nich zajmować się ma zagadnieniami związanymi z tworzeniem komputerowych systemów wyszukiwania informacji, drugi problematyką wiążącą się z tworzeniem systemów automatycznego rozstrzygnięcia, trzeci zagadnieniami tworzenia systemów ekspertowych. Dział czwarty obejmuje problematykę wiążącą się z tworzeniem na gruncie prawa systemów o innym niż wskazane wyżej charakterze (systemy legislacyjne, systemy skomputeryzowanych rejestrów, systemy wiążące się z elektronizacją sądów i administracji, systemy elektronicznego obrotu prawnego). Struktura ta uwzględnia aktualny stan wiedzy informatyczno-prawnej i ma charakter typologii biorącej jako kryterium wyodrębnienia działów przedmiotowy charakter tworzonych systemów.

Rozdział II książki poświęcony jest zagadnieniom informacji prawnej. Przedstawione zostały koncepcje rozumienia pojęcia „informacja” na gruncie lingwistyki, cybernetyki, teorii informacji i informatyki. W konsekwencji za właściwą definicję informacji przydatną do stosowania na gruncie informatyki prawniczej uznano definicję podkreślającą jej semantyczny charakter, uznając że informacja to pewna treść, wyrażana przy pomocy znaków językowych przekazywana przez jakiegoś nadawcę, która po przekształceniu na dane jest przedmiotem przetwarzania komputerowego. W rozdziale tym omówione zostały również poglądy wiążące się z rozumieniem pojęcia „informacja o prawie” wskazując, iż w świetle współczesnego rozwoju informatyki prawniczej pojęcie to nie może być utożsamiane z pojęciem „informacja prawna” i konieczne jest wyodrębnienie drugiej kategorii informacji prawnej, jaką jest „informacja w prawie”. Koncepcję wyodrębniania tych dwóch rodzajów informacji stworzoną przez J. Janowskiego uznano za w pełni uzasadnioną. Wskazano jednocześnie, że przedmiotem monografii są zagadnienia wiążące się z „informacją o prawie”, a za informację tę uznano informację zawartą w różnorodnych tekstach będących aktami normatywnymi, wyrokami sądowymi i decyzjami administracyjnymi oraz publikacjami naukowymi z dziedziny prawa.

Rozdział III książki omawia problematykę szeroko rozumianego języka prawnego. Konieczność poświęcenia temu zagadnieniu szczególnej uwagi wynika z faktu, że podlegające przetwarzaniu w komputerowych systemach wyszukiwawczych teksty pisane są w takim właśnie języku. W rozdziale przedstawiono występujące w polskiej teorii prawa koncepcje teoretyczno-prawne i logiczne wiążące się z wyodrębnianiem na gruncie języka prawnego sensu *largo* jego odmian jakimi są język prawny i prawniczy. Te dwa języki po raz pierwszy

wyodrębnione zostały przez B. Wróblewskiego w 1948 roku i od tego czasu toczy się szeroka dyskusja dotycząca nie tylko kryteriów ich wyodrębniania lecz także tego czy w ogóle istnieją.

W sposób bardziej szczegółowy umówione zostały koncepcje lingwistyczne i socjolingwistyczne wiążące się pojmowaniem języka prawnego. Problematykę tę przedstawiono na tle badań prowadzonych przez T. Gizberta – Studnickiego, który analizował początkowo problematykę możliwości wyodrębniania języka prawnego jako języka w rozumieniu koncepcji strukturalistycznej, której głównym przedstawicielem był N. Chomsky. Uznał on słusznie, że język prawny nie może być uznany za język na gruncie tej teorii lingwistycznej. Przedstawione zostały również prowadzone przez niego analizy charakteru języka prawnego sensu *largo* i jego odmian przy zastosowaniu teorii aktów performatywnych J. L. Austina. W bardziej szczegółowy sposób omówiona została również możliwość wyodrębnienia tych języków na gruncie teorii rejestrów językowych M. A. K. Hallidaya. Przedstawiona również została koncepcja języka prawnego jako socjolektu stworzona przez B. Hałas. Zaprezentowano także propozycję klasyfikacji języków występujących na gruncie prawa stworzoną przez M. Zielińskiego. Wskazano przy tym na brak jednolitego kryterium przeprowadzenia tej kwalifikacji. Uznano jednak, iż sama idea wyodrębnienia wielu różnych języków na gruncie prawa znajduje uzasadnienie. Opisane wyżej koncepcje wiążące się z pojmowaniem języka prawnego, mające charakter głównie teoretyczny stanowiły tło dla przedstawienia w dalszej części rozdziału koncepcji informatyczno-prawnych języków występujących na gruncie prawa. Wskazano przy tym, że informatyka prawnicza akceptować musi nie rekonstrukcjonistyczne, lecz deskrypcjonistyczne koncepcje rozumienia języka prawnego. Informatyka prawnicza utożsamiać bowiem musi język prawny z tekstem. Oznacza to konieczność oderwania się od lingwistycznego rozumienia języka w myśl którego tekst jest jedynie wytworem języka i egzemplifikacją jego użycia, tj. jest produktem języka, a nie językiem.

Przedstawione zostały także istotne z punktu widzenia informatyki prawniczej rozważania wiążące się z wzajemnym stosunkiem języka prawnego i języka powszechnego, uznając za słuszny pogląd, iż język prawny charakteryzuje duży stopień odrębności semantycznej. Jest to z punktu widzenia informatyki prawniczej istotne ustalenie, oznacza bowiem konieczność budowy przy tworzeniu systemów wyszukiwania informacji prawnej tezaursów o specjalistycznym charakterze.

W dalszej części rozdziału przedstawione zostały przeprowadzone w Polsce badania, wiążące się z ustaleniem empirycznych właściwości języka prawnego, z których najistotniejsze

znaczenie miała przeprowadzona przez A. Malinowskiego komputerowa analiza lingwistyczna polegająca na porównaniu korpusu Słownika Frekwencyjnego Języka Prawnego oraz korpusu Słownika Frekwencyjnego Polszczyzny Współczesnej. Z badań tych wynika, że język prawny posiada odrębną leksykę w stosunku do języka współczesnej polszczyzny, co uznano w monografii za ustalenie istotne również dla teorii lingwistyki. Pozwala bowiem uznać język prawny za rejestr języka w jego rygorystycznym rozumieniu, mimo iż nie był on uprzednio w taki sposób klasyfikowany i uznawano, że może być on uznany za rejestr jedynie przy liberalnym pojmowaniu rejestru. Omówiono również, istotne dla informatyki prawniczej wyniki tych badań dotyczących właściwości – statystycznych języka prawnego (tj. języka aktów normatywnych). Przedstawione zostały również w tym rozdziale metody analizy języka prawnego wiążące się z NLP (Natural Language Processing) oraz text mining. Dokonana została również analiza różnic występujących w ramach tekstów składających się na wyodrębniane zazwyczaj podbazy baz informacji o prawie tj. tekstów aktów normatywnych, tekstów stanowiących orzeczenia sądowe oraz tekstów naukowych, oparta na badaniu ich „chwiejności semantycznej”. W konkluzji uznano, że wobec występujących między tymi tekstami różnic w tym zakresie, powinno się przy przeszukaniu baz danych w ramach każdej z tych podbaz stosować w części odmienne rozwiązania, co jak się wydaje nie jest do końca brane pod uwagę przez twórców wielu baz komercyjnych, którzy w większości wypadków stosują dla przeszukania każdej z nich standardowe rozwiązania pełnotekstowe.

Rozdział IV poświęcony jest roli systemów informowania o prawie. Na wstępie omówione w nim zostały zagadnienia wiążące się z wyodrębnianiem różnorodnych kategorii informacji prawnych. W szczególności zaprezentowane zostały dotyczące tej problematyki poglądy F. Studnickiego, a także T. Gizberta – Studnickiego. Przedstawiono również wyniki badań empirycznych przeprowadzonych w Polsce przez G. Wierczyńskiego, które dotyczyły dostępu do informacji o prawie w kontekście potrzeb prawnych obywateli. Sformułowana została teza o konieczności położenia nacisku na tworzenie powszechnie dostępnych niekomercyjnych, komputerowych baz wyszukiwania informacji prawnej. Tylko bowiem w przypadku oparcia modelu informowania o prawie obywateli na możliwości korzystania z takich baz możliwa stała by się realizacja idei równości w dostępie do informacji prawnej a w konsekwencji realizacja idei równości wobec prawa. Na przeszkodzie w realizacji tych celów stoi występująca na rynku światowym dominacja baz o charakterze komercyjnym i postępująca monopolizacja. W rozdziale omówione są koncepcje rozwiązań alternatywnych polegające z jednej strony na budowie wolno dostępnych baz informacji prawnych o charakterze publicznym, tworzonych

przez organizacje państwowe, w szczególności rządy i parlamenty, z drugiej na realizacji idei powszechnego dostępu do informacji prawnych w oparciu o bazy tworzone przez występujące w wielu krajach instytucje informacji prawnej, które już w chwili obecnej tworzą oparte o Internet globalne sieci informacyjne.

W rozdziale tym omówione zostały również dwa konkurencyjne modele udostępniania informacji stworzone przez J. Binga (zwolennika oparcia systemów informowania na tworzeniu narodowych baz informacyjnych) i G. Greenleafa (zwolennika tworzenia globalnych sieci informacyjnych wykorzystujących zasoby informacyjne instytucji informacji prawnej). Wskazano też, iż innym możliwym rozwiązaniem jest proponowane przez T. Bruce'a zastąpienie koncepcji tworzenia narodowych systemów informacyjnych jak i globalnej sieci informacyjnej przez system całkowicie zdecentralizowany oparty o wolny dostęp do źródłowych dokumentów prawnych, który wykorzystywać będzie Google, *web spiders* i techniki rangowania dokumentów. Jak się wydaje brak jest jednak ciągle odpowiednio zaawansowanych narzędzi technicznych do realizacji tego zadania.

W rozdziale tym omówione są również modele informowania o prawie w Polsce, które jak chodzi o realizację idei powszechnego dostępu do informacji prawnej realizowane są głównie w oparciu o bazy tworzone przez instytucje państwowe. W tym kontekście została omówiona problematyka elektronicznego ogłaszania aktów normatywnych, a także główne założenia funkcjonowania baz RCL (Rządowe Centrum Legislacyjne) oraz bazy ISAP (Internetowy System Aktów Prawnych). Omawianie problemów wiążących się z koniecznością zapewnienia obywatelom dostępu do aktualnej informacji o prawie za pośrednictwem baz o charakterze publicznym wskazuje, że działające w Polsce systemy nie pozwalają na właściwą realizację tej idei. Po pierwsze ze względu na brak właściwych metod przeszukiwania baz danych, w szczególności brak możliwości przeszukiwania pełnotekstowego, a także ze względu na problemy dostępu do posiadających walor autentyczności tekstów, jakie powstają po dokonaniu nowelizacji aktu prawnego. Proponowanym rozwiązaniem jest realizacja postulatu, aby w momencie uchwalenia nowelizacji aktu normatywnego sporządzany byłby jednocześnie tekst ujednolicony, również uchwalany przez parlament, który miałby walor obowiązującego tekstu autentycznego (tj. tekstu jednolitego) Występująca w Polsce sytuacja powoduje, że działające bazy o publicznoprawnym charakterze nie potrafią w pełni zaspokoić potrzeb użytkowników zwłaszcza tych którzy są profesjonalistami. Świadczy o tym choćby to że sędziowie w Polsce informację prawną uzyskują głównie poprzez korzystanie z baz komercyjnych takich jak LEX i LEGALIS (97 % sędziów) Systemy te, co istotne dla profesjonalnych prawników zapewniają



również dostęp do orzecznictwa sądowego i poglądów doktryny prawa, czego nie oferują ani ISAP ani RCL. Powoduje to, że dominującą rolę w zakresie wyszukiwania informacji o prawie odgrywają systemy o charakterze komercyjnym. Stan taki, występujący zresztą w wielu innych krajach, nie pozwala na realizację idei równości w dostępie do informacji o prawie.

Rozdział V pracy ma charakter *stricte* informatyczno-prawny i omawia typy systemów informacji prawnej stosowane przy budowie systemów wyszukiwawczych. Przedstawiona w nim została typologia systemów wyszukiwawczych wśród których tradycyjnie wyodrębnia się systemy bezpośredniego i pośredniego udostępniania informacji. Na wstępie dokonana została analiza możliwości stosowania na gruncie prawa systemów typu bezpośredniego (Question Answering systems). Koniecznym było przy tym odrębne zanalizowanie możliwości ich stosowania na gruncie prawa kontynentalnego i precedensowego, a to ze względu na fakt, że ich rozumienie jest na gruncie każdego z tych systemów prawnych odmienne. Na gruncie prawa kontynentalnego za system typu bezpośredniego uznać można (zgodnie z poglądami M. Zielińskiego) jedynie taki system, który na wyrażone w instrukcji wyszukiwawczej pytanie prawne potrafi by udzielić odpowiedzi w postaci podania tekstu normy prawnej bądź norm prawnych wyinterpretowanych w sposób automatyczny z tekstów przepisów prawa. Na gruncie prawa precedensowego rozumienie systemu typu bezpośredniego jest inne i uznaje się, że systemem takim jest również taki system, który pozwala na udzielenie użytkownikowi odpowiedzi na zadane przez niego, sformułowane w języku naturalnym pytanie prawne w postaci podania definicji pojęcia, o które użytkownik pyta, uporządkowanej ze względu na stopień relewancji listy orzeczeń sądowych wiążących się z tą problematyką oraz odwołań do literatury prawniczej. Przeprowadzona w tym rozdziale analiza o charakterze teoretyczno-prawnym wykazała, że budowa systemów typu bezpośredniego na gruncie prawa kontynentalnego nie jest możliwa, a mające realizować to zadanie projekty badawcze mimo wielu prób nie doprowadziły do osiągnięcia zadowalających rezultatów. Jako główną przeszkodę w tworzeniu systemów o tym charakterze uznano problemy wiążące się z wykładnią prawa, a także to, że zależności między przetwarzanymi jednostkami tekstowymi jakie mają miejsce w tekstach pisanych w dyskursie normatywnym w porównaniu z zależnościami, jakie występują w tekstach pisanych w dyskursie opisowym, mają dużo bardziej skomplikowany charakter z czego często próbujący tworzyć systemy o tym charakterze informatycy nie zdają sobie sprawy. Za istotną przeszkodę w realizacji systemów typu bezpośredniego na gruncie systemów prawa kontynentalnego uznać należy w szczególności to, że wykładnia prawa wymaga w wielu wypadkach nie tylko posługiwania się językowymi dyrektywami wykładni

lecz także dyrektywami o charakterze funkcjonalnym, co wymagałoby, stworzenia programów komputerowych, które potrafiłyby uwzględniać przy interpretacji prawa elementy ocen, które musiałyby być w pewien sposób sformalizowane. Istotnym jest przy tym to, że nawet gdyby takie metody zostały odnalezione, to ich stosowanie byłoby sprzeczne z powszechnie akceptowaną na gruncie teorii prawa zasadą konkurencyjności metod wykładni. W tym stanie rzeczy sformułowano konkluzję stwierdzającą, iż budowa systemów typu bezpośredniego na gruncie prawa kontynentalnego nie jest na obecnym etapie wiedzy możliwa. Co więcej wątpliwym jest, czy będzie możliwa w przyszłości.

W dalszej części rozdziału omówiona jest problematyka wiążąca się z budową systemów Question-Answering na gruncie prawa precedensowego. Przedstawione zostały w szczególności architektura takiego systemu oraz ostatnio realizowana próba jego stworzenia w postaci systemu LEXIS ANSWERS. Omówiono też ostatnio stworzony system ROSS. W rezultacie przeprowadzonych rozważań jest wniosek, iż na gruncie prawa precedensowego budowa systemów Question-Answering jest możliwa, choć pamiętać należy o tym, iż ich zdefiniowanie jest tu inne niż na gruncie prawa kontynentalnego. W dalszej części rozdziału przedstawiono typologię systemów pośredniego wyszukiwania informacji prawnej tj. systemów, które w odpowiedzi na pytanie użytkownika systemu, nie udzielają mu bezpośredniej odpowiedzi lecz dostarczają dokumenty, z których odpowiedź ta może zostać wyinterpretowana. Stosowanie tego typu systemów jest możliwe zarówno na gruncie prawa kontynentalnego jak i precedensowego. Wskazano również na problemy wiążące się z ich budową. W ramach typologii wyodrębniono systemy klasyfikacji hierarchicznej, systemy haseł przedmiotowych, systemy indeksowe i systemy pełnotekstowe. Omówiono tradycyjne systemy wyszukiwania informacji, do których należą systemy klasyfikacji hierarchicznej oraz systemy haseł przedmiotowych. Opis tych systemów powiązany był z przykładami ich wykorzystania w systemach komputerowych, a także w systemach tradycyjnych. Przeprowadzona też została analiza przydatności ich stosowania w wyszukiwaniu informacji prawnej. W rezultacie uznano, że wady tych systemów powodują, iż mogą one być wykorzystywane w współczesnych komputerowych systemach wyszukiwania informacji prawnej tylko jako jeden ze sposobów przeszukiwania bazy danych, będący subsydiarnym w stosunku do przeszukiwań pełnotekstowych bądź opartych o indeksy. Dalsza część rozdziału przedstawia zasady budowy systemów indeksowych i pełnotekstowych w oparciu o prezentację występujących w nich w nich zapisów. Omówione również zostały wady i zalety tych systemów.

Rozdział VI książki zatytułowany „Nieklasyczne koncepcje wyszukiwania informacji prawnej” omawia niestandardowe sposoby wyszukiwania informacji, za które uznano te które nie polegają na wykorzystaniu indeksów, a zwłaszcza strategii pełnotekstowej. Bezkrityczne wykorzystanie w systemach komercyjnych ostatniej z wymienionych spowodowało wystąpienie zjawiska, które uznać można za drugą fazę kryzysu informacyjnego w prawie. Pierwsza faza kryzysu, która miała miejsce w latach 60 – tych ubiegłego stulecia miała charakter ilościowy i była wynikiem znacznego zwiększenia liczby unormowań prawnych, a w przypadku krajów anglosaskich również liczby case’ów. Przewyciężenie tego stanu rzeczy było stosunkowo proste i dokonało się poprzez zastosowanie komputerów do wyszukiwania informacji prawnej. Natomiast jak chodzi o drugą fazę kryzysu, która rozpoczęła się w połowie lat 80 – tych i, jak się wydaje, trwa nadal, to jej przewyciężenie jest dużo trudniejsze ze względu na jej jakościowy charakter. Użytkownicy systemów otrzymują wprawdzie w wyniku przeprowadzonych wyszukiwań olbrzymia liczbę informacji, ale informacja ta ma często charakter nieustrukturyzowany, nieuporządkowany, nieukazujący powiązań pomiędzy dokumentami i niewskazujący na odniesienia do pojęć stworzonych w ramach doktryny prawa. W tym stanie rzeczy pojawiły się koncepcje wyszukiwania informacji prawnej oparte o wykorzystanie nowych metod.

W rozdziale omówione zostały najważniejsze kierunki rozwoju koncepcji wyszukiwania informacji prawnej o nieklasycznym charakterze. Najpierw omówione zostały systemy oparte o wykorzystanie języka naturalnego, ze względu na fakt, że znalazły one zastosowanie praktyczne w ramach istniejących już systemów komercyjnych. W dalszej kolejności przedstawione zostały założenia systemów opartych o wykorzystanie analizy wektorowej, a także metod analizy statystycznej, używanych dla celu ustalania podobieństwa dokumentów. Omówiono również systemy określane jako konceptualne. Wskazane zostały także perspektywy zastosowania dla wyszukiwania informacji prawnej sztucznej inteligencji.

Omówienie systemów opartych o wykorzystanie języka naturalnego zaprezentowano na przykładzie systemów WESTLAW IS NATURAL (WIN), LEXIS FREESTYLE SEARCHING, a także ostatnio powstałych WESTLAWNEXT oraz LEXIS ADVANCED. Z analizy działania tych systemów wynika, że systemy takie nie są w stanie zaspokoić potrzeb profesjonalnych użytkowników, nie pozwalają bowiem na uzyskanie informacji o odpowiedniej, potrzebnej zawodowym prawnikom jakości. Istotną ich zaletą jest jednak to, że próbują one zastąpić realizowaną w systemach opartych o wyszukiwanie boolowskie koncepcję „exact match” koncepcją „best match” umożliwiającą uzyskanie w wyniku

wyszukania uporządkowanej ze względu na domniemany stopień relewancji w stosunku do potrzeb użytkowników listy dokumentów. W dalszej części rozdziału dokonano wprowadzenia teoretycznego poprzez wprowadzenie i omówienie takich pojęć jak „przestrzeń semantyczna terminów” i „przestrzeń semantyczna dokumentów”. Pozwoliło to w dalszej kolejności na ukazanie metod wyodrębniania podprzestrzeni w przestrzeni dokumentów, polegających na wykorzystaniu analizy wektorowej oraz metod analizy statystycznej zmierzających do ustalenia podobieństwa dokumentów. Ostatnią z tych metod zaprezentowano na przykładzie rozwiązań stosowanych w projekcie ERASMUS. W dalszej kolejności zaprezentowano metody wyodrębniania podprzestrzeni w przestrzeni terminów, prezentując założenia budowy tezauryusa statystycznego, a także map semantycznych.

W dalszej części omówione zostały tzw. konceptualne systemy wyszukiwania informacji prawnej. Zwolennicy tworzenia tego typu systemów podkreślają, że pojęcia w lepszy sposób reprezentują potrzeby użytkowników niż słowa, stąd też wyszukiwanie w oparciu o pojęcia jest wyszukiwaniem o lepszym charakterze i pozwala na pełniejsze zaspokojenie potrzeb użytkowników. Przedstawione zostały projekty wiążące się z wyszukiwaniem konceptualnym, a także analiza stosowanych w ramach systemów konceptualnych metod reprezentacji dokumentów. Za najbardziej obiecujące uznano stosowanie metod reprezentacji automatycznej o charakterze interpretującym, które zresztą wykorzystano w komercyjnym systemie FLEXLEX. Ukazanie podstawowych metod stosowanych w systemach konceptualnych wiązało się ze wskazaniem wad i zalet każdej z nich. Omówione również zostały znajdujące zastosowanie bądź teoretycznie możliwe do zastosowania nowe techniki wyszukiwawcze, w szczególności techniki opierające się o wykorzystanie sieci neuronowych, algorytmów genetycznych, koncepcji Latent Semantic Analysis i hipertekstu. Szczególną uwagę poświęcono w ramach tego rozdziału omówieniu koncepcji wiążących się z zastosowaniem w systemach wyszukiwania informacji prawnej sztucznej inteligencji. W opinii wielu ekspertów przyszłość rozwoju systemów wyszukiwania informacji wiąże się z integrowaniem tego, co jest uznawane za AI (Artificial Intelligence) z IR (Information Retrieval). Zdaniem E. Schweighofera dziedzinami AI&Law są: 1) logic based knowledge systems, 2) concept based knowledge systems, 3) integracja AI i systemów wyszukiwawczych przy pomocy sieci neuronowych 4) case based reasoning, 5) natural language processing. W monografii omówione zostały jedynie te z dziedzin, które wiążą się z zastosowaniem rozwiązań możliwych do zaaplikowania w systemach wyszukiwawczych. Szczególną uwagę poświęcono problematyce ontologii prawniczych, co wiąże się z tym, że stanowią one nowy dział

zastosowania sztucznej inteligencji w prawie, a budowa odpowiedniej ontologii jest warunkiem rozwoju systemów wyszukiwawczych opartych o sztuczną inteligencję. W pracy przedstawiono dwa systemy ontologii prawnej. Pierwszy z nich to FOL (Functional Ontology for Law) stworzona przez A. Valente oraz ontologia FBO (Frame Based Ontology) stworzona przez R. van Kralingena. Budowanie wskazanych wyżej ontologii służy tworzeniu systemów wyszukiwania informacji o charakterze konceptualnym, opartych concept based knowledge system. Przeszkodą w ich szerszym zastosowaniu są problemy wiążące się z akwizycją wiedzy, które pozwalają na ich budowę jedynie w wąskich dziedzinach prawa czy też fragmentach tych dziedzin. W monografii omówione zostały w dalszej kolejności ontologie leksykalne, takie jak Word Net, Euro World Net, a także LOIS (Lexical Ontologies for Legal Information Sharing). Omówiono też nowe koncepcje określane jako CLO (Comprehensive Legal Ontology). Propozycje wykorzystania ontologii leksykalnych dla budowy systemów wyszukiwawczych uznane zostały za kierunek właściwy, który w przyszłości, doprowadzić może do znacznego zwiększenia efektywności systemów wyszukiwania informacji prawnej. Problemem jest jednak to, że budowa takich ontologii ukazujących powiązania w ramach języka jest niezwykle pracochłonna ze względu na fakt, iż tworzone są one w sposób manualny. Przedmiotem opisu i analizy były również w tym rozdziale systemy oparte o wykorzystanie sieci neuronowych takie jak AIR i SCALIR. Ukazane zostały również podstawy metod analizy lingwistycznej (NLP). Szczególną uwagę poświęcono ostatnio pojawiający się systemów o charakterze hybrydowym tj. systemom wykorzystującym wiele nowych choć często mających różny charakter metod, z których część zlicza się do metod charakterystycznych dla sztucznej inteligencji. W szczególności omówiono system FLEXICON (obecnie FLEXLEX). Cechą tego systemu stworzonego na Uniwersytecie British Columbia w Kanadzie przez D. Gelbarta i J. C. Smitha jest połączenie ze sobą różnych metod wyrosłych na gruncie sztucznej inteligencji, a także metod statystycznej analizy języka. Istotnym jest to, że system ten jako jedyny z systemów opartych na wykorzystaniu sztucznej inteligencji ma charakter komercyjny i jest wykorzystywany w praktyce. Omówienie systemów opartych o zastosowanie sztucznej inteligencji wskazuje, iż mimo osiągnięcia pewnych sukcesów jak dotychczas nie osiągnięto na tym polu w pełni zadowalających rezultatów. Jako przyczynę takiego stanu rzeczy wskazano istotne różnice występujące pomiędzy badaniami prowadzonymi na gruncie tworzenia systemów wyszukiwawczych a badaniami dotyczącymi sztucznej inteligencji. Po pierwsze, badania prowadzone na gruncie systemów wyszukiwawczych zmierzają do stworzenia technik możliwych do stosowania przy przetwarzaniu obszernych zbiorów tekstowych, podczas gdy badania prowadzone na gruncie AI skupiają się na sposobach

reprezentacji wiedzy o ograniczonym zasięgu. Po drugie, podstawową metodą reprezentacji wiedzy w wyszukiwaniu jest tekst, podczas gdy dla AI istotna jest reprezentacja wiedzy, która jest z tych tekstów wyinterpretowana. Po trzecie, wyszukiwanie informacji posiada silną tradycję opierania się o badania empiryczne, a nowe modele i teorie są testowane na dużych zbiorach dokumentów, natomiast sztuczna inteligencja pokazuje możliwości stosowania owych technik w ramach stosunkowo ograniczonych aplikacji. Po czwarte wreszcie nowe koncepcje w zakresie wyszukiwania informacji opierają się o wykorzystanie metod statystycznych, a także teorii prawdopodobieństwa, natomiast tradycyjnie rozumiana AI koncentruje się na rozumowaniach symbolicznych. W chwili obecnej na gruncie badań stosowanych w zakresie obu tych dyscyplin zaczynają być jednak stosowane te same metody. Należą do nich neural networks, case based reasoning i natural language processing, co pozwala domniemywać, że w przyszłości skala zastosowań sztucznej inteligencji w systemach wyszukiwawczych będzie większa. Rozdział kończy ocena nowych koncepcji wyszukiwania informacji prawnej, wiążąca się z twierdzeniem, iż dalszy rozwój systemów wyszukiwania informacji prawnej uzależniony jest od znalezienia metod, które uwzględniłyby w procesach wyszukiwania informacji elementy wiedzy prawniczej.

Ostatni rozdział monografii poświęcony jest przeglądowi najważniejszych systemów informacji prawnej na świecie. W kolejności omówione zostały systemy wyszukiwania informacji prawnej jakie stworzono na kontynencie północnoamerykańskim. Prezentacja tych systemów w części ma charakter historyczny, a prezentacja założeń systemów już niedziałających służyła ukazaniu historii rozwoju koncepcji systemów wyszukiwawczych. W szczególności omówiony został pierwszy system wyszukiwania informacji prawnej na świecie stworzony na Uniwersytecie w Pittsburgu, a także systemy ASPEN, LITE, LAW RESEARCH INCORPORATED, RIRA, JURIS, WRU, OBAR. Z systemów tych jedynie LITE przekształcony na FLITE jest systemem działającym do dziś. Opisanie zostały również najważniejsze systemy działające, do których należy LEXIS I WESTLAW, które są bazami dominującymi na rynku amerykańskim. Wskazano również na inne źródła informacji prawnej takie jak baza LOISLAW, BLOMBERG LAW, PACER COURTLINKS. Bazami tymi posługuje się jednak jedynie ok. 15% firm prawniczych w Stanach Zjednoczonych. W tym kontekście wskazano na niebezpieczeństwo duomonopolistycznej pozycji LEXIS i WESTLAW. Podane zostały również informacje na temat kanadyjskich systemów DATUM oraz Q/L (QUIC/LAW, obecnie QUICK/LAW). W sposób obszerny zostały omówione europejskie systemy wyszukiwania informacji prawnej. Zaprezentowano w kolejności systemy

działające w Belgii, Francji, Szwecji, Włoszech, Niemczech, Austrii, Norwegii, Wielkiej Brytanii, a także w szeregu innych krajów. Część opisów występujących w ramach poszczególnych państw systemów miała charakter historyczny. Mimo iż współcześnie systemy te już nie istnieją uznano, że przynajmniej niektóre z nich stosowały interesujące rozwiązania, co było motywem dla ich przedstawienia. Omówiono również polskie bazy wyszukiwania informacji prawnej ukazując historię ich rozwoju. Przedstawiono zawartość baz i sposoby wyszukiwania informacji we współcześnie dostępnych na rynku bazach komercyjnych LEX oraz LEGALIS. Podano również informacje o bazach publicznych, takich jak ISAP i RCL. W dalszej części rozdziału omówione również zostały europejskie bazy prawne wspólnot ponadnarodowych. W bardziej szczegółowy sposób przedstawiona została baza EUR-Lex, baza CURIA, a także baza HUDOC (**H**uman **R**ights **D**ocumentation). Podsumowując rozważania w tym rozdziale wskazano na postępującą monopolizację dostępu do informacji prawnej, zwłaszcza tej, która przeznaczona jest dla zawodowych prawników przy czym uznano, że niezbędnym jest wspieranie idei tworzenia baz o charakterze niekomercyjnym, których obowiązek tworzenia takich baz powinien spoczywać na organach władzy publicznej.

Monografię kończą załączniki, m.in. Deklaracja Wolnego Dostępu do Prawa, lista członków FALM czy Deklaracja Haska.



dr Jacek Petzel

