

*AI dla prawników :*  
*sztuczna inteligencja w praktyce zawodów prawniczych*

Przemysław Chojecki,  
Andrzej Pałys

Spis treści

Wstęp	11
Rozdział I. Analiza umów	13
1. Aspekty technologiczne – Natural Language Processing (NLP)	14
1.1. Tokenizacja i embeddingi – jak komputery zamieniają słowa w liczby	15
1.2. Kluczowe techniki: rozpoznawanie nazwanych jednostek, klasyfikacja tekstu	17
1.3. Ewolucja: modele statystyczne vs. sieci neuronowe	18
1.4. Duże modele językowe: GPT, BERT i LLaMA	19
2. Retrieval-augmented generation (RAG)	20
2.1. Łączenie wyszukiwania dokumentów z generacją tekstu	21
2.2. Architektura RAG: baza wiedzy z modelem językowym	22
3. Wyzwania – halucynacje i błędne interpretacje	24
3.1. Dlaczego modele „zmyślają” fakty	24
3.2. Przykłady halucynacji	25
3.3. Strategie ograniczania halucynacji	27
3.4. Konieczność weryfikowania faktów przez użytkownika	28
Rozdział II. Research podatkowy	33
1. Technologie – uczenie modelu, zestawy danych treningowych	34
1.1. Dane czyste vs. dane zaszumione	34
1.2. Proces uczenia modelu	37
1.3. Wpływ różnorodności i wielkości danych na model	39
2. Rodzaje baz danych	40
2.1. Relacyjne i nierelacyjne bazy danych	40
3. Wyzwania – jakość danych	45
3.1. Typowe wady danych prawnych	45
3.2. Techniczne strategie poprawy jakości danych	47
Rozdział III. Analiza due diligence	51
1. Technologie – Optical Character Recognition (OCR)	52

1.1. Rozpoznawanie znaków w obrazach i skanach	53
1.2. Rola sieci konwolucyjnych (CNN)	54
1.3. Wyzwania: czcionki, jakość skanów, układ dokumentu	55
2. Application Programming Interface (API) i Graphical User Interface (GUI)	58
2.1. GUI	59
2.2. API	60
2.3. Znaczenie User Experience (UX) w obszarze sztucznej inteligencji	62
3. Wyzwania – zarządzanie zmianą technologiczną	63
3.1. Opór wobec nowych narzędzi	63
3.2. Kultura organizacyjna a adopcja AI	65
3.3. Rola liderów wdrożenia	67

#### Rozdział IV. Tłumaczenie na rozprawie 69

1. Technologie – Automatic Speech Recognition (ASR) i Text-to-Speech (TTS)	70
1.1. Jak działa ASR – analiza sygnału akustycznego i dopasowanie do słów	71
1.2. Nowoczesne architektury ASR	72
1.3. Jak działa TTS – od tekstu do naturalnie brzmiącego głosu	73
1.4. Wyzwania w ASR i TTS	75
2. Modele multimodalne	77
2.1. Łączenie różnych typów danych (tekst, dźwięk, obraz)	77
2.2. Przykłady architektur multimodalnych (CLIP, Flamingo)	79
2.3. Praktyczne wyzwania w multimodalnych systemach	80
3. Wyzwania – obsługa wielu języków	82
3.1. Modele jednolingwistyczne vs. modele wielojęzyczne	83
3.2. Języki zapisywane w innym alfabecie niż łaciński	84
3.3. Granice automatycznych tłumaczeń w prawie	86

#### Rozdział V. Tłumaczenie i tworzenie tekstu 89

1. Technologie – Large Language Models (LLM) – ciąg dalszy	90
1.1. Architektura Transformer i mechanizm self-attention	90
1.2. Embeddingi i reprezentacja znaczeń	91
1.3. Skalowanie modelu a jakość wyników	92
2. Sztuka zadawania pytań, czyli inżynieria promptów	93
2.1. Typy promptów: zero-shot, few-shot, chain-of-thought	94
2.2. Rola długości i kontekstu promptu	99
3. Wyzwania – role w projektach IT i AI	101
3.1. Role w metodyce Agile – Scrum	102
3.2. Role analityczne i architektoniczne	103
3.3. Role inżynierskie i testerskie	103
3.4. Dodatkowe role w projektach AI	104
3.5. Wyzwania związane z rolami projektowymi	105

3.6. AI Act a definicje: model AI ogólnego przeznaczenia (art. 63) oraz operator zmiennoprzecinkowy (art. 67) 107

Rozdział VI. Korekta wiadomości e-mail	111
1. Technologie – analiza semantyczna i analiza sentymentu	112
1.1. Embeddingi w praktyce analizy wiadomości e-mail	112
1.2. Analiza kontekstu zdania	113
1.3. Klasyfikacja emocji	115
2. Wyzwania – prywatność treści użytkownika	118
2.1. Uczenie modeli a prywatność danych użytkowników	118
2.2. RAG i dane w infrastrukturze użytkownika lub dostawcy	120
2.3. Secondary use	122
3. Uprzedzenia (biases) w modelach AI	124
3.1. Skąd się biorą uprzedzenia w danych?	124
3.2. Metody ograniczania uprzedzeń	126

Rozdział VII. Predykcyjna analiza w ściganiu przestępstw	129
1. Technologia – infrastruktura chmurowa (IaaS, PaaS, SaaS)	131
1.1. Podstawowe rodzaje środowisk chmurowych	131
1.2. Zalety i wady środowisk chmurowych	133
1.3. Zrozumienie kosztów środowisk chmurowych	135
2. Big data analytics	138
2.1. Właściwości big data (objętość, prędkość, różnorodność)	138
2.2. Algorytmy do wykrywania anomalii i wzorców	139
2.3. Przetwarzanie rozproszone	142
2.4. Big data analytics a AI	144
3. Wyzwania – czas procesowania zapytań	145
3.1. Psychologia opóźnień	146

Rozdział VIII. Pasywne rozpoznawanie naruszeń prawa własności	149
1. Technologia – analiza obrazu	150
1.1. Algorytmy rozpoznawania kształtów i obiektów	150
1.2. Wyzwania dla analizy obrazów	152
2. Pozyskiwanie i integracja danych	153
2.1. Mechanizmy pozyskiwania i łączenia różnych źródeł	154
2.2. Standardy wymiany danych	157
2.3. Wyzwania interoperacyjności	158
3. Wyzwania – fałszywe alarmy	159
3.1. Skąd biorą się fałszywe alarmy?	160
3.2. Jak radzić sobie z fałszywymi alarmami?	161

Rozdział IX. Zgłaszanie naruszeń (sygnaliści)	163
1. Technologie – chatboty i systemy konwersacyjne	164

1.1. Technologia wspierająca konwersację	165
1.2. Współczesne chatboty	167
2. Autoryzacja i autentykacja	168
2.1. Różnica między autentykacją a autoryzacją	169
2.2. Czynniki autentykacji, standardy bezpieczeństwa i SSO	169
2.3. Wyzwania dla autoryzacji i autentykacji	172
3. Wyzwania – zagadnienia cyberbezpieczeństwa	173
3.1. Ogólne zagadnienia cyberbezpieczeństwa	173
3.2. Zero trust security	177
3.3. Cyberbezpieczeństwo w AI	178

## Rozdział X. Wykrywanie zagrożeń (miasto/policja) 183

1. Technologia – systemy predykcyjne w czasie rzeczywistym	184
1.1. Algorytmy analizy strumieni danych	185
1.2. Edge computing	188
2. Wyzwania – koszty korzystania z rozwiązań AI	190
2.1. Źródła kosztów rozwiązań AI	190
2.2. Aspekty hardware’owe i energetyczne – koszty środowiskowe i zrównoważony rozwój	192
3. Explainability i audytowalność AI	194
3.1. Podstawowe techniki explainability (XAI)	194
3.2. Audyt algorytmiczny	