

Uzupełnienie informacji dotyczącej komponentu sztucznej inteligencji (AI) w systemie przedsięwzięcia eOperat

1. Przedmiot i podstawa wyjaśnienia

Niniejsze wyjaśnienie stanowi uzupełnienie Opisu założeń przedsięwzięcia informatycznego „eOperat” w zakresie planowanego komponentu sztucznej inteligencji (AI). Dokument sporządzono w odpowiedzi na uwagę w karcie oceny wskazującą na potrzebę rozszerzenia opisu komponentu AI, w tym określenia zastosowanych modeli lub klas modeli (OCR, HTR, NLP czy RAG, wyszukiwanie semantyczne), zasad trenowania modeli i wykorzystywanych danych, a także omówienia kluczowych aspektów wykorzystania AI: zgodności regulacyjnej (compliance), bezpieczeństwa, ochrony danych, jakości wyników oraz odpowiedzialności za rezultaty generowane przez system. Zgodnie z treścią uwagi, wobec ograniczeń objętości poszczególnych sekcji wzoru OZPI, wyjaśnienie przedstawiono w formie odrębnego dokumentu pisemnego.

2. Ogólny opis komponentu AI

Komponent AI jest wbudowanym elementem systemu eOperat, budowanym w infrastrukturze GUGiK. Jego zadaniem jest wsparcie wykonawcy prac geodezyjnych w kompletowaniu i kontroli operatu technicznego oraz ograniczenie czynności manualnych i redakcyjnych. Komponent nie zastępuje decyzji człowieka ani organu — pełni funkcję pomocniczą (wsparcia decyzyjnego), a rozstrzygnięcia formalne pozostają po stronie geodety uprawnionego.

Zakres funkcjonalny komponentu obejmuje dwie zasadnicze grupy zdolności: (1) automatyczne rozpoznawanie i analizę treści dokumentów wchodzących w skład operatu (OCR/NLP) na potrzeby kontroli kompletności i spójności oraz (2) wyszukiwanie semantyczne w ustrukturyzowanej bazie wiedzy opartej na obowiązujących przepisach i standardach technicznych. Na tych zdolnościach opiera się warstwa wsparcia użytkownika (Asystent AI), która prowadzi geodetę przez procedurę i wskazuje wymagane działania. Rozpoznawanie pisma odręcznego (HTR) nie jest przewidziane na obecnym etapie i pozostaje poza zakresem przedsięwzięcia.

3. Funkcje realizowane przez komponent AI

3.1. Rozpoznawanie i analiza treści dokumentów (OCR/NLP)

Operat techniczny jest przekazywany w postaci składowych o zróżnicowanej strukturze: plików współrzędnych (TXT), plików geometryczno-opisowych (GML) oraz dokumentów tekstowo-graficznych (PDF, w tym skany). Komponent OCR (optyczne rozpoznawanie znaków) umożliwia pozyskanie warstwy tekstowej z dokumentów rastrowych i PDF, natomiast komponent NLP (przetwarzanie języka naturalnego) realizuje ekstrakcję informacji, klasyfikację typów dokumentów oraz porównanie danych opisowych z danymi w plikach TXT i GML. Celem jest wykrywanie braków formalnych, niespójności między składowymi operatu a danymi źródłowymi oraz przeciwdziałanie zjawisku „cyfrowego papieru”, w którym dokumentacja elektroniczna nie jest maszynowo czytelna.

Wyniki analizy AI mają charakter podpowiedzi i wskazań dla użytkownika oraz danych wejściowych dla deterministycznych reguł walidacyjnych (m.in. modułu Walidacja GML). Rozstrzygnięcie o kompletności i poprawności operatu następuje na podstawie reguł zdefiniowanych zgodnie z przepisami, a nie wyłącznie na podstawie oceny modelu.

3.2. Wyszukiwanie semantyczne i baza wiedzy

Komponent wyszukiwania semantycznego opiera się na reprezentacjach wektorowych (embeddingach) tekstu oraz na wektorowej bazie wiedzy zbudowanej z obowiązujących aktów

prawnych (m.in. ustawa Prawo geodezyjne i kartograficzne oraz rozporządzenia wykonawcze), standardów technicznych i zatwierdzonych materiałów metodycznych. Umożliwia to dopasowanie wymagań formalnych do konkretnego rodzaju sprawy geodezyjnej oraz odnajdywanie właściwych podstaw prawnych na podstawie znaczenia zapytania, a nie wyłącznie zgodności słów kluczowych.

Warstwa generowania odpowiedzi działa w modelu ugruntowanym na źródłach (podejście typu RAG – *retrieval-augmented generation*): odpowiedzi są tworzone wyłącznie na podstawie odnalezionych, zweryfikowanych fragmentów bazy wiedzy, z podaniem odwołań do źródeł. Ogranicza to ryzyko halucynacji i zapewnia rozliczalność treści.

Mechanizm RAG nie modyfikuje treści dokumentów źródłowych ani nie tworzy własnej bazy wiedzy. Model językowy wykorzystuje wyłącznie odnalezione fragmenty zweryfikowanych dokumentów jako kontekst do wygenerowania odpowiedzi, co zwiększa wiarygodność informacji oraz umożliwia ich weryfikację przez użytkownika.

3.3. Warstwa wsparcia użytkownika (Asystent AI)

Asystent AI jest warstwą prezentacji zdolności opisanych w pkt 3.1 i 3.2. Prowadzi użytkownika przez proces realizacji sprawy, udziela spersonalizowanych informacji o procedurach i wymaganiach formalnych oraz wskazuje elementy operatu wymagające uzupełnienia lub korekty. Asystent nie podejmuje czynności władczych, nie generuje decyzji administracyjnych i nie zastępuje weryfikacji dokonywanej przez organ.

4. Modele i klasy modeli

W rozwiązaniu przewiduje się wykorzystanie następujących klas modeli. Ze względu na etap przedsięwzięcia wskazano klasy modeli oraz wymagania, a nie konkretnych dostawców; ostateczny wybór modeli nastąpi na etapie projektowania architektury z zachowaniem zasady hostingu we własnej infrastrukturze.

Klasa modelu	Zastosowanie	Charakter	Trenowanie
Model OCR	Pozyskanie warstwy tekstowej z dokumentów PDF/skanów wchodzących w skład operatu	Model gotowy (open-source), tekst drukowany	Bez trenowania na danych osobowych
Model językowy (LLM/NLP)	Ekstrakcja informacji, klasyfikacja dokumentów, streszczenia, generowanie odpowiedzi ugruntowanych (RAG)	Model gotowy o otwartych wagach, hostowany on-premise	Bez trenowania; ewentualny prompt-engineering i kontekstowanie
Model embeddingowy	Reprezentacje wektorowe tekstu na potrzeby wyszukiwania semantycznego	Model gotowy, dostrajalny	Możliwe dostrajanie na danych referencyjnych (nieosobowych)
Klasyfikatory pomocnicze	Wspomaganie kontroli kompletności i typowania składowych operatu	Model uczony metodą nadzorowaną	Trenowane na danych referencyjnych/syntetycznych

Rozpoznawanie pisma odręcznego (HTR) nie jest objęte zakresem przedsięwzięcia.

Ze względu na etap przygotowania przedsięwzięcia wskazano klasy modeli AI oraz wymagania funkcjonalne wobec komponentów sztucznej inteligencji. Wybór konkretnych modeli, ich wersji oraz

parametrów implementacyjnych nastąpi na etapie projektowania architektury rozwiązania oraz realizacji zamówienia, zgodnie z wymaganiami funkcjonalnymi, bezpieczeństwa i wydajności określonymi dla systemu eOperat.

5. Zasady trenowania modeli i dane treningowe

Przyjęto podejście hybrydowe. Modele podstawowe (OCR oraz model językowy) wykorzystywane są w postaci gotowej, wstępnie wytrenowanej (pretrained), bez trenowania na danych produkcyjnych i bez trenowania na danych osobowych pochodzących z operatów lub z PZGiK. Dodatkowe dostrajanie ograniczone jest do komponentów pomocniczych — modeli embeddingowych i klasyfikatorów wspierających walidację — i prowadzone jest wyłącznie na danych referencyjnych o charakterze nieosobowym.

Zasady dotyczące danych wykorzystywanych do dostrajania i budowy bazy wiedzy:

- źródłami są obowiązujące akty prawne, standardy techniczne oraz zatwierdzone materiały metodyczne; baza wiedzy jest wersjonowana i aktualizowana wraz ze zmianami przepisów,
- do dostrajania komponentów pomocniczych wykorzystuje się dane zanonimizowane lub syntetyczne odwzorowujące strukturę operatu, bez danych osobowych,
- dane produkcyjne przekazywane przez użytkowników nie są wykorzystywane do trenowania modeli (wykorzystywane są wyłącznie do realizacji konkretnej usługi świadczonej przez system eOperat i nie stanowią zbioru treningowego wykorzystywanego do dalszego uczenia modeli AI),
- stosuje się zasadę minimalizacji danych oraz maskowanie danych wrażliwych przed przekazaniem treści do modeli,
- procesy dostrajania i aktualizacji bazy wiedzy podlegają wersjonowaniu, testom akceptacyjnym i rozliczalności (rejestr zmian).

6. Infrastruktura i hosting

Modele AI wykorzystywane są jako modele o otwartych wagach (open-weight) hostowane we własnej infrastrukturze GUGiK. Rozwiązanie nie zakłada przekazywania treści operatów ani danych osobowych do zewnętrznych, komercyjnych usług AI. Takie podejście zapewnia suwerenność danych, kontrolę nad cyklem życia modeli oraz ogranicza ryzyko wycieku danych. Zasoby obliczeniowe niezbędne do uruchomienia modeli (w tym serwery na potrzeby AI) ujęto w pozycji kosztowej „Infrastruktura” Opisu założeń.

Takie rozwiązanie zapewnia pełną kontrolę nad środowiskiem przetwarzania danych, umożliwia spełnienie wymagań bezpieczeństwa informacji obowiązujących w administracji publicznej oraz ogranicza ryzyko przekazywania danych do zewnętrznych dostawców usług AI.

7. Kluczowe aspekty wykorzystania AI

7.1. Zgodność regulacyjna (compliance)

Klasyfikacja systemu zgodnie z rozporządzeniem Parlamentu Europejskiego i Rady (UE) 2024/1689 (AI Act) zostanie przeprowadzona na etapie projektowania architektury rozwiązania. Na obecnym etapie przygotowania przedsięwzięcia nie zidentyfikowano funkcjonalności wskazujących na kwalifikację systemu jako systemu AI wysokiego ryzyka, jednak ostateczna klasyfikacja zostanie potwierdzona po opracowaniu architektury i zakresu funkcjonalnego rozwiązania. Niezależnie od wyniku klasyfikacji system zostanie zaprojektowany z uwzględnieniem zasad przejrzystości, zarządzania ryzykiem, nadzoru człowieka, rozliczalności oraz rejestrowania działań komponentów AI.

Rozwiązanie projektowane jest ponadto z uwzględnieniem RODO (rozporządzenie 2016/679), ustawy o krajowym systemie cyberbezpieczeństwa, ustawy o informatyzacji działalności podmiotów realizujących zadania publiczne oraz Krajowych Ram Interoperacyjności (KRI). W harmonogramie przedsięwzięcia przewidziano inicjalny oraz weryfikacyjny test prywatności.

7.2. Bezpieczeństwo

Bezpieczeństwo komponentu AI zapewniają: hosting modeli we własnej infrastrukturze GUGiK, uwierzytelnianie i autoryzacja dostępu z wykorzystaniem Węzła Krajowego, moduł Audytu i Logowania rejestrujący działania użytkowników, maskowanie danych wrażliwych przed przekazaniem do modeli oraz mechanizmy ograniczające podatność na manipulację treścią zapytań (m.in. wstrzykiwanie instrukcji – prompt injection). Model generowania odpowiedzi ugruntowany na zweryfikowanych źródłach ogranicza ryzyko odpowiedzi nieuprawnionych. Rozwiązanie objęte jest testami podatności, testami bezpieczeństwa i analizą statyczną kodu przewidzianymi w Opisie założeń.

7.3. Ochrona danych

Stosowane są zasady minimalizacji danych i ograniczenia celu. Dane osobowe nie są wykorzystywane do trenowania modeli, a treści przekazywane do modeli podlegają maskowaniu danych wrażliwych. Dostęp do danych podlega kontroli i rozliczalności (moduł Audytu i Logowania). Przetwarzanie danych osobowych podlega ocenie skutków dla ochrony danych (DPIA) zgodnie z RODO oraz istniejącym w GUGiK politykom bezpieczeństwa informacji.

7.4. Jakość wyników

Jakość wyników zapewniają: ugruntowanie odpowiedzi na zweryfikowanych źródłach z podaniem odwołań, oddzielenie deterministycznych reguł walidacyjnych od odpowiedzi generowanych przez AI, testy akceptacyjne i ewaluacja komponentu przed wdrożeniem, monitorowanie działania i przeciwdziałanie dryfowi modelu, a także mechanizm nadzoru człowieka (human-in-the-loop). Wskazania AI mają charakter pomocniczy i podlegają weryfikacji przez użytkownika oraz — na etapie przyjęcia do PZGiK — przez organ.

7.5. Odpowiedzialność za rezultaty

Odpowiedzialność za treść i kompletność operatu technicznego pozostaje po stronie geodety uprawnionego, który podpisuje operat, a odpowiedzialność za przyjęcie dokumentacji do PZGiK - po stronie właściwego organu. Komponent AI pełni funkcję wsparcia i nie generuje rozstrzygnięć władczych. Wyniki działania AI są rozliczalne (rejestrowanie zdarzeń, odwołania do źródeł), a użytkownik jest informowany o pomocniczym charakterze odpowiedzi i o interakcji z systemem AI.

8. Powiązanie z ryzykami przedsięwzięcia

Powyższe rozwiązania odpowiadają na ryzyka związane z AI wskazane w rejestrze ryzyk OZPI, w szczególności: niestabilność odpowiedzi asystenta i podatność na manipulację (ugruntowanie odpowiedzi na zweryfikowanej bazie wiedzy), dryf modelu i jakość źródeł (wersjonowanie bazy wiedzy, monitorowanie, ewaluacja), wyciek danych treningowych (hosting modeli we własnej infrastrukturze) oraz ujawnienie danych poufnych osobom nieuprawnionym (maskowanie danych wrażliwych, kontrola dostępu).