

OPIS PRZEDMIOTU ZAMÓWIENIA

Dialog techniczny ma na celu uzyskanie informacji niezbędnych do przeprowadzenia zamówienia publicznego polegającego na opracowaniu systemu automatycznego klasyfikowania krajowych zgłoszeń wynalazków i wzorów użytkowych w oparciu o międzynarodową klasyfikację patentową (MKP), w ramach tegorocznej edycji GovTech Polska.

I WPROWADZENIE: Czym jest wyzwanie, dlaczego jest takie ważne dla organizacji i obywateli?

Realizacja wyzwania ma na celu przyspieszenie procesu klasyfikowania zgłoszeń, optymalizację zaangażowania zasobów i rozkładu pracy oraz efektywniejsze monitorowanie pracy w toku.

Obecny proces obsługi klasyfikacji krajowych zgłoszeń o udzielenie ochrony dla wynalazków i wzorów użytkowych rozpoczyna się od zapoznania się z nimi przez pracowników (ekspertów) urzędu i przypisania im klasyfikacji pozwalającej na dalsze przekazanie ich do właściwych ekspertów merytorycznych (powiązanych z określoną klasą w MKP). Proces ten jest czynnością bazującą na słowach kluczowych pojawiających się w treści zgłoszeń i doświadczeniu pracowników. Do urzędu wpływa ponad 100 ww. zgłoszeń tygodniowo - skala wpływa na czas ich procedowania. Celem urzędu jest skrócenie czasu wykonywanej wstępnej analizy i klasyfikacji zgłoszeń przed przekazaniem ich do merytorycznej oceny (badania) przez poszczególne grupy ekspertów.

II WIZJA: Jaki jest oczekiwany długoterminowy rezultat? Co zostanie wdrożone? Co to zmieni?

Wprowadzenie automatycznego systemu informatycznego klasyfikowania krajowych zgłoszeń wynalazków i wzorów użytkowych w oparciu o międzynarodową klasyfikację patentową, z wykorzystaniem sztucznej inteligencji i algorytmów uczenia maszynowego, pozwoli na sprawniejszą, szybszą i efektywniejszą obsługę zgłoszeń, a także zmniejszenie nakładów administracyjnych.

III ROZWIĄZANIE: Co chciałby zobaczyć Zamawiający?

Celem postępowania jest wytworzenie systemu, który na bazie dotychczasowych zgłoszeń i przypisanej im klasyfikacji, przydzieli właściwą klasyfikację dla nowo składanych zgłoszeń wynalazków i wzorów użytkowych oraz przypisze zgłoszenie do odpowiedniego eksperta do dalszych ocen merytorycznych.

Kluczowe funkcje systemu

- przyjmowanie na wejściu dokumentów w formacie PDF (z warstwą tekstową w języku polskim) i zwracanie wyniku klasyfikacji wg międzynarodowej klasyfikacji patentowej (MKP – <http://mkp.uprp.gov.pl>),
- klasyfikowanie dokumentów do klasy MKP przy poprawności dopasowania powyżej 90%, wskazując przy tym prawdopodobieństwo poprawności dopasowania dla trzech najwyżej punktowanych wyników (najwyższy wynik jest uznawany za wiążący),
- przypisanie badanego, sklasyfikowanego dokumentu do eksperta zajmującego się daną klasą względem aktualnego obciążenia pracą (wielu ekspertów zajmuje się tym samymi klasami),
- posiadanie funkcji uczenia się na podstawie nowo gromadzonych danych (aktualizacja modelu na podstawie poprawnych i błędnych klasyfikacji – błędne klasyfikacje korygowane ręcznie przez ekspertów),
- system musi udostępniać interfejsy do integracji z systemami urzędu (API).

Dodatkowe funkcje systemu

- funkcja OCR dokumentów dostarczanych w formacie bez warstwy tekstowej.

IV Wstępne mierniki skuteczności wdrożenia <do uszczegółowienia w dialogu>

- zmniejszenie średniego czasu i pracochłonności obsługi procesu klasyfikacji zgłoszeń,
- poziom skuteczności automatycznej klasyfikacji powyżej 90% dla IV poziomu MKP.

V Wstępne kryteria I Etapu <do uszczegółowienia w dialogu>: Jak weryfikujemy propozycje aplikacje wpływające w I etapie?

Ocenie podlega:

- Skuteczność klasyfikacji na danych wzorcowych – poprawność przypisania do właściwych działów i poddziałów MKP (II poziom klasyfikacji).

Dane w I Etapie : Jakie dane udostępniamy? Jaki mają charakter? Skąd będą pobierane?

Klasyfikacja MKP oraz próbka danych 5000 sklasyfikowanych opisów patentowych oraz opisów wzorów użytkowych w formacie PDF.

Zamawiający na próbce 100 nierozpoznanych dokumentów zgłoszeniowych potwierdzi trafność wyuczenia modelu.

Istotne jest, iż dane do wyuczenia modelu są wzbogacone względem dokumentów zgłoszeniowych o dane bibliograficzne (metadane danego opisu wraz z ujętą klasyfikacją).

VI Wstępne kryteria II Etapu <do uszczegółowienia w dialogu>: Jak weryfikujemy rozwinięcie aplikacji I Etapu i jej zbieżności do głównego celu?

Ocenie podlega:

- Skuteczność klasyfikacji na danych wzorcowych – poprawność przypisania do właściwych działów, poddziałów, klas i podklas (liści drzewa klasyfikacji MKP – IV poziom).

Dane w II Etapie:

Klasyfikacja MKP oraz próbka danych 100 000 sklasyfikowanych opisów patentowych oraz opisów wzorów użytkowych w formacie PDF.

Zamawiający na próbce 1 000 nierozpoznanych dokumentów potwierdzi trafność wyuczenia modelu.

Istotne jest, iż dane do wyuczenia modelu są wzbogacone względem dokumentów zgłoszeniowych o dane bibliograficzne (metadane danego opisu wraz z ujętą klasyfikacją).

VII Podsumowanie:

Nagrody I etap: <do uszczegółowienia w dialogu>

Zwrot kosztów II etap: <do uszczegółowienia w dialogu>

Oczekiwany budżet przeznaczony na cały konkurs: <do uszczegółowienia w dialogu>

Oczekiwany czas wdrożenia: <do uszczegółowienia w dialogu>