



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



 **PARP**
Grupa PFR

OPIS PRZEDMIOTU ZAMÓWIENIA

zapytanie ofertowe nr - 1/ON/2024

Crmos - innowacyjny system rekomendacji szans sprzedaży w branży zamówień publicznych

Rodzaj prac: Prace rozwojowe

Cel modułu

Celem modułu jest oprac. innowacyjnego w skali kraju prototypu systemu automatycznej rekomendacji przetargów (postępowań przetargowych) opartego na analizie zmiennych zw. z działalnością wykonawcy, jego preferencjami i historią podejmowanych działań (w zakresie udziału w przetargach) oraz zmiennych w przetargach (danych opisujących postęp. przetarg.). System pozwoli na wyekstrahowanie krytycznych dla wykonawcy zmiennych w przetargach w postaci syntetycznego raportu.

Efektom projektu będzie oprac. mechanizmu rekomendacji postępowań przetargowych dla MŚP, uczestniczących w przetargach publicznych. Mechanizm ten uwzględni charakterystykę ofert składanych przez MŚP oraz indyw. preferencje wykonawców.

Automatyzacja systemu rekomendacji przetargów, czyli z jednej strony wymagań określonych przez zamawiającego, a z drugiej strony możliwości i preferencji wykonawcy, wymaga opracowania innow. algorytmu w celu integracji i agregacji zmiennych z rozproszonych i niestandardyzowanych źródeł danych.

Problem badawczy/technologiczny

Nadrzędnym probl. badawczym (PB) jest uzyskanie rekomendacji generowanych przez system na poziomie 90% dopasowanych rekomendacji przetargu do danego wykonawcy w oparciu o zmienne postępowania, jak i zmienne wykonawcy. Problemem jest rozproszenie, nierzadko niestandardyzowanych źródeł informacji o przetargach, co utrudnia podejmowanie zamówień, szczególnie w sektorze MŚP (gdzie działa się w oparciu o ograniczone zasoby). Zatem wyzw. technolog. (WT) systemu będzie takie zastos. algorytmu rekomendującego, który uwzględni odmienne postępowania w poszczególnych branżach, gdzie kluczowe mogą być zupełnie inne cechy.

Następnie WT jest oprac. mechanizmu, który będzie w stanie w zbiorze tekstów pisanych oraz plików graficznych (zw. łączne materiały) rozpoznać i wyszukać fragmenty treści, będące zmiennymi, opisującymi postępowania przetargowe (przetargi) oraz zmiennymi opisującymi wykonawców (podmioty składające oferty w post. przetargowym/profil wykonawcy). Sukcesem będzie uzyskanie skuteczności 90% dla rozpoznawania cech w tekście pisany lub pliku graficznym na poziomie min. 80% skutecznie odczytanych fragmentów.



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



 PARP
Grupa PFR

Wyzwaniem będzie zastosowanie jednorodnej modyfikacji metody pozyskiwania danych z tego rodzaju materiałów, bowiem przybierają różny format edycyjny oraz sposób zapisu treści.

Automatyzacja systemu rekomendacji przetargów, czyli z jednej strony wymagań określonych przez zamawiającego, a z drugiej strony możliwości i preferencji wykonawcy, wymaga wykorzystania innowacyjnego podejścia w celu integracji i agregacji informacji z rozproszonych i niestandardizowanych źródeł danych. Statystyczne modele językowe (SLM - Statistical Language Models) nie działają dobrze dla wielu języków, szczególnie w przypadku słownictwa technicznego i branżowego [13]. Dlatego dla tak postawionego problemu wymagane jest wykorzystanie od strony technologicznej rozwiązań. Poszukiwana będzie metoda pozyskiwania zmiennych danych opierających się na ekstrakcji danych, systemach OCR (optical character recognition), DL (deep learning) oraz ML (machine learning).

Pierwszą próbą uzyskania metody rozw. wyzwań będzie zbudowanie wydajnego systemu rozpoznawania treści dla dokumentów związanych z przetargiem i wykonawcą. W tym obszarze można wydzielić następujące wyzwania [4], [13]:

>analiza układu strony

>optyczne rozpoznawanie znaków

>łączenie znaków w słowa.

Tradycyjne systemy OCR w przypadku analizy dokumentów mogą popełniać błędy, co wynika z braku standaryzacji zapisu i różnorodności źródeł [4]. Dokładność odwzorowania podaje się w lit. na poziomie 45%-80% [8], [9], [13], [19]. Dlatego też szczególnym wyzwaniem staje się wykrywanie i segmentowanie słów w takie, które występują w języku polskim oraz są jednoznacznie do zrozumienia w kontekście przetargów.

Wyzwanie stanowi sposób pełnej automatyzacji systemu rozpoznawania treści. Metody, które pozwalają na zmniejszenie liczby błędów w systemie OCR opierać się muszą na metodach DL, a zwłaszcza rekurencyjnych sieciach neuronowych [4]. Jak pokazują badania [9], [14], uczenie się systemu opartego na OCR jest obecnie czasochłonne. Jest to proces nietrywialny i wymagający wiedzy eksperckiej. Idealne, wydajne i skuteczne podejście do uczenia powinno obejmować zintegrowany system, w którym sekwencja danych na etapie przygotowania i szkolenia przebiegają płynnie, a postęp i skuteczność uczenia się systemu jest oceniana obiektywnie [9].

Wyzwaniem technologicznym algorytmu rekomendującego będzie uwzględnienie różnorodności postępowań w poszczególnych branżach, gdzie kluczowe mogą być zupełnie inne zmienne. Według lit. [5], [6], [7], [11], [14], [15], [16], [20], [21], [25], [26] najlepsze wyniki w zakresie segmentacji tekstu oferują metody hybrydowe sieci FCN (Fully Convolutional Networks) do ekstrakcji cech. Rozwinięciem tych metod są algorytmy Transformers, których użycie i modyfikacja stają uzasadnione technologicznie w kontekście rozwoju systemów sztucznej inteligencji [27], [28], [29]. Poszukiwany będzie sposób wiązania zmiennych w różnych zakresach i branżach. Poszukiwane będzie rozwinięcie ww. algorytmu, który to opiera się na mechanizmie uwagi. Pozwala na lepsze zrozumienie kontekstu zdania, w

odróżnieniu do innych sekwencyjnych metod, które np. widzą tylko część zdania do obecnie przetwarzanego. Algorytm transformer określa cały kontekst zdania na każdym etapie przetwarzania, a to umożliwi podjęcie próby skutecznej modyfikacji działania ekstrakcji danych, ponieważ zmniejsza prawdopodobieństwo pomyłki.

W końcu poszukiwany będzie sposób weryfikacji skuteczności dopasowania rekomendacji przetargów do wykonawców o różnym profilu.

Rozwiązaniem problemu jest nietrywialne i dot. będzie uzyskania efektywnego sposobu rekomendacji postępowań przetargowych dla MŚP. Docelowy system powinien uwzględniać konsensus oraz powiązanie między zmiennymi przetargów a zmiennymi ofert składanych przez MŚP oraz ich preferencjami.

Pozycje lit. dla stanu nauki i techniki wskazano numer. w HRF.

Metoda badawcza

W celu rozwiązania PB oraz WT wykorzyst. zostanie spójna metodyka hybrydowa i zastos. zostaną kolejno:

- >metody rozpoznawania tekstów i obrazów dla rozpoznania zawartości z dokumentów przetargu (w tym plików graf.) oraz zapisanie ich w bazie danych. Metody będą modyf. technik i opierać się będą na segmentacji, zastos. sieci neuronowej i rozpoznawaniu wzorców, aby docelowo osiągać skuteczność rozpoz.; następnie wykorzyst. zostaną:
- >metody przetwarzania danych (data mining) pozwolą na uzyskanie wiedzy z baz danych oraz poszukiwanie prawidłowości w zbiorach danych za pomocą metod statyst. np. analizą skupień (np. K-means, DBSCAN, hierarch.). Pozwoli to na klastrowanie danych i oprac. modelu referenc. karty przetargu; równoległe do ww. metod stos. będą:
- >metody uczenia maszynowego (ML) w celu uczenia algorytmu przetwarzania zbiorów dokumentów, od ich skanowania przez wykrywanie wzorców słów i wyrażeń w nich zawartych (w połąc. z metodami rozpoznawania tekstów i obrazów) aż po automatyczne grupowanie słów i podobnych wyrażeń (w połąc. z metodami przetwarzania danych), które najlepiej charakteryzują ów zbiór, by uzyskać dopasowanie kart przetargu do profilu wykonawcy. Planowana metoda ML to połączenie indukcyjnego (na podstawie reguł) i analitycznego uczenia (wykorz. sieci neuron.). Zastos. będzie np. architektura Transformer, na bazie której powstanie kastomizowany model BERT oraz algorytm Apriori pozwalający na zdef. reguł dla modelu sieci zależności między zmiennymi przetargu (zbiory elementów). Ww. czynności stanowią oś badawczą projektu i pozw. na uzyskanie prototypu systemu. Horyzontalnie stosow. będą metody uzup. proces:
- >metoda symulacyjna wykorzyst. do badania działania prototypu algorytmu w warunkach testowych
- >metody eksperckie weryfikacji skuteczności dla oceny modelu karty przetargu, profilu wykonawcy oraz sieci zależności między zmiennymi przetargu wsparte metodyką myślenia sieciowego (z wykorzyst. bad. ankietowych/FGI m.in. dla skuteczności algorytmu rekomend.).

1. Nazwa zamówienia:

Opracowanie i dostarczenie innowacyjnego systemu rekomendacji przetargów w branży zamówień publicznych.

Przedmiotem zamówienia jest opracowanie i dostarczenie innowacyjnego systemu rekomendacji przetargów w branży zamówień publicznych. Celem projektu jest stworzenie prototypu systemu automatycznej rekomendacji, który będzie oparty na analizie zmiennych związanych z działalnością wykonawcy oraz przetargów.

System ma umożliwić wykonawcom precyzyjne i skuteczne dobieranie przetargów, uwzględniając preferencje i historię działań wykonawcy. Ma również dostarczać informacji dotyczących kluczowych aspektów przetargów, takich jak nazwa zamawiającego, wartość zamówienia, termin składania ofert, wymagania techniczne itp.

W ramach projektu wykonawca będzie odpowiedzialny za opracowanie efektywnego systemu wykorzystującego metody przetwarzania języka naturalnego (NLP), deep learning (DL) i machine learning (ML). System będzie zdolny do analizowania i klasyfikowania informacji z różnych źródeł danych, takich jak portale przetargowe, bazy danych, dokumenty zamówienia, opisy wykonawców itd.

Opracowany prototyp systemu powinien generować rekomendacje przetargów dostosowane do preferencji i historii działań wykonawcy. Rekomendacje powinny być czytelne i zrozumiałe, prezentując kluczowe informacje dotyczące przetargów w sposób łatwy do interpretacji.

Projekt obejmuje także dostarczenie pełni działającego prototypu systemu oraz dokumentacji kodu, umożliwiającej jego rozwijanie i dostosowywanie w przyszłości.

2. Rodzaj prac:

Prace badawczo-rozwojowe mające na celu opracowanie prototypu systemu automatycznej rekomendacji przetargów opartego na analizie zmiennych związanych z działalnością wykonawcy oraz przetargów.

3. Cel zamówienia:

Celem zamówienia jest opracowanie prototypu systemu rekomendacji przetargów, który będzie uwzględniał preferencje i historię działalności wykonawców oraz zmienne w przetargach. System ma umożliwić wyekstrahowanie kluczowych zmiennych związanych z przetargami i dostarczyć syntetyczny raport dla wykonawców.

4. Zakres zamówienia:

a) Opracowanie i dostarczenie prototypu systemu automatycznej rekomendacji przetargów opartego na analizie zmiennych związanych z działalnością wykonawcy oraz przetargów.



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



 **PARP**
Grupa PFR

- b) Zintegrowanie i agregacja informacji z różnych źródeł danych dotyczących przetargów.
- c) Opracowanie wydajnego systemu rozpoznawania treści dokumentów przetargowych i wykonawców, w tym rozpoznawanie słów i segmentacji tekstu.
- d) Uwzględnienie różnorodności postępowań przetargowych w różnych branżach oraz opracowanie algorytmów rekomendacji uwzględniających preferencje i profile wykonawców.
- e) Dostarczenie w pełni działającego prototypu systemu rekomendacji przetargów.

5. Okres realizacji:

Projekt będzie realizowany przez okres 24 miesięcy od momentu rozpoczęcia prac.

6. Wymagane narzędzia i licencje:

Wykonawca powinien dysponować odpowiednimi narzędziami, licencjami i technologiami niezbędnymi do przeprowadzenia prac zgodnie z zakresem zamówienia.

7. Prawa autorskie:

Po przyjęciu przedmiotu umowy przez Zamawiającego, wszystkie prawa autorskie przechodzą na Zamawiającego wraz z prawem własności nosników danych.

Zamówienie obejmuje następujące obszary działania:

1. Opracowanie metod budowy reprezentacji wektorowej dla treści przetargów oraz wykonawców. W obrębie tego obszaru należy przeprowadzić testy empiryczne modeli językowych, dostroić je na korpusie opisów przetargów oraz wykonawców, a także opracować warianty modeli z osadzeniem wybranych metadanych. Dodatkowo, w ramach obszaru, wymagane jest stworzenie algorytmów rzutowania opisów przetargów i wykonawców na słowniki kluczowych cech oraz algorytmów rozszerzania tych słowników na podstawie analizy zbiorowej kolekcji opisów.
2. Wytworzenie metod grupowania opisów przetargów i wykonawców. Wymaga to przeprowadzenia badań empirycznych wybranych algorytmów grupowania, stworzenia algorytmów wykorzystujących reprezentację wektorową, metadane i kluczowe cechy oraz przeprowadzenia analizy cech różnicujących.
3. Opracowanie zestawu miar oceny podobieństwa między wykonawcami a przetargami, metody semantycznego wyszukiwania opisów przetargów z opisem wykonawcy jako zapytaniem oraz miary dopasowania wykonawcy do przetargu na podstawie analizy własności symbolicznych. Wymagane jest także stworzenie algorytmu rekomendacji oparty



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



PARP
Grupa PFR

na złożonych, wieloaspektowych reprezentacjach wykonawców i przetargów oraz modułu wyjaśnialności dla generowanego rankingu.

KRYTERIA WYBORU OFERTY

- CENOWE

$$CC = C_{min} / CO * 40$$

Gdzie: CC – liczba punktów za kryterium cena CO – cena netto oferty C_{min} – najniższa cena netto oferty, która wpłynęła w odpowiedzi na zapytanie ofertowe.

- POZACENOWE

Kryterium „Doświadczenie kadry w udziale w projektach”

- 2 punkty za udział pracownika/ osoby oddelegowanej do realizacji zamówienia w projektach o tematyce przetwarzania języka oraz modeli językowych.

wykonawca uzyska:

0 pkt - za brak osoby

2 pkt - za 1 osobę

4 pkt- za 2 osoby

6 pkt- za 3 osoby

8 pkt- za 4 osoby

10 pkt- za 5 i więcej osób

- 1 punkt za udział pracownika/ osoby oddelegowanej do realizacji zamówienia w pozostałych projektach informatycznych.

wykonawca uzyska:

0 pkt - za brak osoby

1 pkt - za 1 osobę



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



PARP
Grupa PFR

2 pkt- za 2 osoby

3 pkt- za 3 osoby

4 pkt- za 4 osoby

5 pkt- za 5 i więcej osób

Wykonawca w tym kryterium może uzyskać maksymalnie 15 pkt, których waga stanowi 30%

$DS = DSo / DSmax * 30$

Gdzie: DSo - suma punktów za udział w projektach, DSmax - najwyższy wskaźnik DSo z przesłanych ofert

Kryterium „Doświadczenie kadry w tworzeniu narzędzi NLP”:

- 2 punkty za udział pracownika/ osoby oddelegowanej do realizacji zamówienia w tworzeniu narzędzia NLP z wykorzystaniem głębokich sieci neuronowych

wykonawca uzyska:

0 pkt - za brak osoby

2 pkt - za 1 osobę

4 pkt- za 2 osoby

6 pkt- za 3 osoby

8 pkt- za 4 osoby

10 pkt- za 5 i więcej osób

- 1 punkt za udział pracownika/ osoby oddelegowanej do realizacji zamówienia w tworzeniu innego narzędzia NLP

wykonawca uzyska:

0 pkt - za brak osoby



Fundusze Europejskie
dla Nowoczesnej Gospodarki



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



 **PARP**
Grupa PFR

1 pkt - za 1 osobę

2 pkt- za 2 osoby

3 pkt- za 3 osoby

4 pkt- za 4 osoby

5 pkt- za 5 i więcej osób

Wykonawca w tym kryterium może uzyskać maksymalnie 15 pkt, których waga stanowi 30%

$NR = NRo / NRmax * 30$

Gdzie: NRo - suma punktów dla tworzonych narzędzi NLP, NRmax - najwyższy wskaźnik NRo z przesłanych ofert