

Warszawa, dnia 25 lipca 2023 r.

Protokół

z przeprowadzenia Wstępnych konsultacji rynkowych poprzedzających wszczęcie postępowania o udzielenie zamówienia publicznego dotyczącego przygotowania i wykonania wyszukiwarki orzeczeń Krajowej Izby Odwoławczej w oparciu o istniejące narzędzie oraz usług hostingu

I. Zamawiający

Urząd Zamówień Publicznych
ul. Postępu 17 A
02-676 Warszawa

II. Podstawa prawna

Wstępne konsultacje rynkowe przeprowadzono na podstawie art. 84 ustawy z dnia 11 września 2019 r. – Prawo zamówień publicznych (t.j. Dz. U. z 2022 r. poz. 1710 z późn. zm.).

III. Cel, przedmiot konsultacji i data ogłoszenia Wstępnych konsultacji rynkowych:

Zamawiający w dniu 19 czerwca 2023 r. ogłosił Wstępne konsultacje rynkowe poprzedzające wszczęcie postępowania o udzielenie zamówienia publicznego dotyczącego przygotowania i wykonania wyszukiwarki orzeczeń Krajowej Izby Odwoławczej w oparciu o istniejące narzędzie oraz usług hostingu.

Zamawiający wskazał, że celem konsultacji jest poinformowanie podmiotów zainteresowanych uzyskaniem przyszłego zamówienia o swoich planach i wymaganiach dotyczących zamówienia, oraz pozyskanie od podmiotów posiadających wiedzę specjalistyczną informacji w zakresie wymagań dotyczących architektury, infrastruktury, zasobów, narzędzi, wydajności, hostingu i sposobów/metod testowania wyszukiwarki orzeczeń, wpływu tych wymagań na koszty realizacji, a także określenia ram czasowych (terminu) na przygotowanie i wykonanie wyszukiwarki.

Zamawiający, w dniu 19 czerwca 2023 r. upublicznił informację o wstępnych konsultacjach rynkowych, poprzez publikację *Ogłoszenia o wstępnych konsultacjach rynkowych* na swojej stronie internetowej. Link do ogłoszenia: <https://www.gov.pl/web/uzp/wstepne-konsultacje-rynkowe3>

Wraz z ogłoszeniem, Zamawiający na ww. stronie internetowej umieścił również *Regulamin przeprowadzania wstępnych konsultacji rynkowych* wraz ze wzorem

Zgłoszenia do udziału we Wstępnych konsultacjach rynkowych. Ponadto, poza publikacją ogłoszenia o wstępnych konsultacjach rynkowych na stronie internetowej, Zamawiający w dniu 19 czerwca 2023 r., za pośrednictwem poczty elektronicznej, skierował informację do wybranych wykonawców (wg poniższej listy), informując o fakcie upublicznienia ogłoszenia o Wstępnych konsultacjach rynkowych na swojej stronie internetowej oraz zachęcając do udziału w konsultacjach. W ww. korespondencji Zamawiający poinformował o przedmiocie planowanego zamówienia publicznego oraz przekazał dokumenty upublicznione na stronie internetowej.

Lista podmiotów, do których Zamawiający skierował korespondencję mailową informującą o Ogłoszeniu o wstępnych konsultacjach rynkowych:

- 1) Sii sp. z o.o.
Al. Niepodległości 69
02-626 Warszawa
- 2) Billennium S.A.
ul. Koszykowa 61
00-667 Warszawa
- 3) Persooa
Cambridge Innovation Campus
ul. Chmielna 73
00-801 Warszawa
- 4) WeNet Group S.A.
Postępu 14A
02-676 Warszawa
- 5) SeSentiOne S.A.
ul. Lęborska 3B
80-386 Gdańsk (w Warszawie na Postępu 15)

Czas przeprowadzenia Wstępnych konsultacji rynkowych został określony *Ogłoszeniu o wstępnych konsultacjach rynkowych* na 14 dni licząc od dnia 26 czerwca 2023 r., tj. od dnia w którym upływał termin na przesyłanie Zgłoszeń do udziału w konsultacjach. Termin prowadzenia negocjacji został przedłużony o 7 dni, o czym w dniu 7 lipca 2023 r. Zamawiający powiadomił podmioty uczestniczące w konsultacjach oraz zamieścił taką informację na stronie <https://www.gov.pl/web/uzp/wstepne-konsultacje-rynkowe3>

- IV. Podmioty, które złożyły zgłoszenie do udziału we wstępnych konsultacjach rynkowych**
W przewidzianym w Ogłoszeniu o wstępnych konsultacjach rynkowych terminie wpłynęły dwa zgłoszenie do udziału we wstępnych konsultacjach rynkowych od:

- 1) Sii sp. z o.o.
Al. Niepodległości 69
02-627 Warszawa
- 2) Billennium S.A.
ul. Koszykowa 61
00-667 Warszawa

Zamawiający w dniu 26 czerwca br., po upływie terminu przewidzianego na przesyłanie Zgłoszeń do udziału w Konsultacjach, przekazał firmom Sii sp. z o.o. i Billennium S.A. zaproszenie do udziału w konsultacjach, w którym zaznaczył, że w trakcie konsultacji Zamawiający będzie miał na celu uzyskanie informacji i wyjaśnień w następującym zakresie:

1. Jakie są koszty i czas wykonania projektu wyszukiwarki przy założeniu określonego zakresu i funkcjonalności?
2. Jakie są przewidywane koszty związane z hostingiem wyszukiwarki opartej o sztuczną inteligencję? Czy istnieje preferowany model opłat miesięcznych, opłat za zużycie zasobów, czy inny model?
3. Jakie są wymagania dotyczące infrastruktury i zasobów, które należy uwzględnić przy hostowaniu wyszukiwarki?
4. Czy istnieją preferowane platformy chmurowe lub dostawcy usług hostingowych do hostowania wyszukiwarek?
5. Jaki rodzaj architektury sztucznej inteligencji jest najbardziej odpowiedni do implementacji wyszukiwarki?
6. Czy istnieją specyficzne języki programowania, frameworki lub narzędzia, które preferujecie do implementacji wyszukiwarki opartej o sztuczną inteligencję?
7. Jak określić główne wymagania dotyczące wydajności wyszukiwarki? Jak szybko powinna być w stanie przetwarzać zapytania i zwracać wyniki? Jak to przekłada się na koszty utrzymania wyszukiwarki.
8. Jakie dodatkowe funkcje w wyszukiwarce uzyskamy dzięki zastosowaniu SI, takie jak podpowiedzi, rekomendacje, personalizacja, wykrywanie anomalii itp.?
9. Opis mechanizmu aktualizowania danych z których korzysta wyszukiwarka i ponownego przetrenowanie modelu z nowymi danymi.
10. Wymagania dotyczące redundancji, odtwarzania po awarii lub planu kontynuacji działania w przypadku awarii?
11. Metody testowania wyszukiwarki (testy funkcjonalne, wydajnościowe, jakościowe itp.), testy iteracyjne oraz czas potrzebny do osiągnięcia pełnej funkcjonalności i spełnienie wszystkich określonych wymagań projektowych.

W dniu 28 czerwca 2023 r. firma Billennium S.A. skierowała do Zamawiającego pytanie dotyczące możliwości wspólnego uczestniczenia we Wstępnych konsultacjach rynkowych

razem z ich siostrzaną firmą eOne sp. z o.o., z uwagi na fakt, że większość specjalistów adresujących obszar AI/SI (Sztucznej Inteligencji) reprezentuje podmiot eOne. W odpowiedzi Zamawiający poinformował firmę Billennium S.A., iż Regulamin Konsultacji przewiduje możliwość zaproszenia do udziału w Konsultacjach rynkowych również wykonawców, którzy złożą zgłoszenie do udziału w Konsultacjach po wyznaczonym terminie. W związku z powyższym firma eOne sp. z o.o. w dniu 30 czerwca 2023 r. przekazała Zamawiającemu Zgłoszenie do udziału w Konsultacjach, w których będzie uczestniczyła wspólnie z Billennium S.A.

V. Terminy i forma przeprowadzenia przez Zamawiającego Wstępnych konsultacji rynkowych

Wstępne konsultacje rynkowe z udziałem przedstawicieli firmy Sii sp. z o.o., odbyły się w formie wideokonferencji przy użyciu aplikacji Teams w dniu 3 lipca 2023 roku w godzinach: 11:00 – 12:00.

Z kolei wstępne konsultacje rynkowe z udziałem przedstawicieli firmy Billennium S.A. i eOne sp. z o.o. odbyły się w formie wideokonferencji przy użyciu aplikacji Teams w dniu 3 lipca 2023 roku w godzinach: 12:30 – 13:30.

Następnie, do dnia zakończenia Konsultacji, tj. do 17 lipca br., Konsultacje były prowadzone w formie wymiany korespondencji elektronicznej.

VI. Przebieg Wstępnych konsultacji rynkowych

1. Dyrektor Generalny Urzędu Zamówień Publicznych w dniu 16 czerwca 2023 r. wyznaczył niżej wskazanych pracowników Urzędu do przeprowadzenia Konsultacji:

- pracownicy merytoryczni z Wydziału Informatyki i Bezpieczeństwa Systemów Teleinformatycznych:
Mariusz Gac
Jarosław Antosiewicz
- pracownik z Zespołu ds. Zamówień publicznych:
Aneta Florczak

Ze strony Zamawiającego, w Konsultacjach uczestniczył również Andrzej Umiński - Dyrektor Biura Organizacyjno-Finansowego.

2. W trakcie prowadzonych spotkań z uczestnikami Konsultacji oraz wymiany korespondencji pozyskano następujące informacje:

Pyt. 1 Jakie są koszty i czas wykonania projektu wyszukiwarki przy założeniu określonego zakresu i funkcjonalności?

Odp. Żadna z firm nie przedstawiła kosztów budowy samego narzędzia - wyszukiwarki. Czas niezbędny na wykonanie wyszukiwarki to około 2-3 miesięcy.

Pyt. 2. Jakie są przewidywane koszty związane z hostingiem wyszukiwarki opartej o sztuczną inteligencję? Czy istnieje preferowany model opłat miesięcznych, opłat za zużycie zasobów, czy inny model?

Odp. Ustalono, że przewidywany miesięczny koszt utrzymania wyszukiwarki może mieścić się w przedziale od 77 zł do 1570 zł netto i zależy od przyjętego przez Zamawiającego rozwiązania. Estymacje miesięcznych kosztów utrzymania Narzędzia można oprzeć na wykorzystaniu wzorca wyszukiwania retrieve-rerank.

Pyt. 3. Jakie są wymagania dotyczące infrastruktury i zasobów, które należy uwzględnić przy hostowaniu wyszukiwarki?

Odp. Obecne na konsultacjach firmy wskazywały na rozwiązanie chmurowe, które to można odpowiednio wyskalować w zależności od potrzeb Zamawiającego.

Pyt. 4. Czy istnieją preferowane platformy chmurowe lub dostawcy usług hostingowych do hostowania wyszukiwarek?

Odp. Uczestnicy Konsultacji wskazali, że przy wykonaniu narzędzia - wyszukiwarki swoje rozwiązanie zbudowaliby na chmurze Azure.

Pyt. 5. Jaki rodzaj architektury sztucznej inteligencji jest najbardziej odpowiedni do implementacji wyszukiwarki?

Odp. Technologia OpenAI (Embedding) opartej na chmurze Azure – wskazanie każdego z uczestników.

Pyt. 6. Czy istnieją specyficzne języki programowania, frameworki lub narzędzia, które preferujecie do implementacji wyszukiwarki opartej o sztuczną inteligencję?

Odp. Tak istnieją takie języki, preferowany jest w tym przypadku Python, Elasticsearch, C#

Pyt. 7. Jak określić główne wymagania dotyczące wydajności wyszukiwarki? Jak szybko powinna być w stanie przetwarzać zapytania i zwracać wyniki? Jak to przekłada się na koszty utrzymania wyszukiwarki.

Odp. Wskazano, że jakość działania systemu może być ustalana poprzez miarę MRR na podstawie zbioru testowego skonstruowanego przez zamawiającego oraz wykonawcę projektu, a także za pomocą ewaluacji jakościowej na grupie docelowych użytkowników systemu. Krytyczne dla akceptacji systemu jest spełnienie kryteriów dotyczących czasu przetwarzania zapytania, który musi być na tyle zredukowany, aby nie wpływać na pracę użytkowników.

Pyt. 8. Jakie dodatkowe funkcje w wyszukiwarce uzyskamy dzięki zastosowaniu SI, takie jak podpowiedzi, rekomendacje, personalizacja, wykrywanie anomalii itp.?

Odp. W trakcie Konsultacji ustalono, że działanie na konkretnych dokumentach ograniczy m.in. halucynacje SI, dzięki czemu wyniki wyszukiwania będą bardziej adekwatne.

Pyt. 9. Opis mechanizmu aktualizowania danych, z których korzysta wyszukiwarka i ponownego przetrenowanie modelu z nowymi danymi.

Odp. Do aktualizowania danych nie potrzeba dodatkowego mechanizmu, za aktualizację odpowiada technologia embeddingu.

Pyt. 10. Wymagania dotyczące redundancji, odtwarzania po awarii lub planu kontynuacji działania w przypadku awarii?

Odp. Należy wykonywać kopie zapasowe bazy danych oraz kopie zapasowe aplikacji.

Pyt. 11. Metody testowania wyszukiwarki (testy funkcjonalne, wydajnościowe, jakościowe itp.), testy iteracyjne oraz czas potrzebny do osiągnięcia pełnej funkcjonalności i spełnienie wszystkich określonych wymagań projektowych.

Odp. Należy wykonać testy integracyjne, a nie tylko funkcjonalne czy jakościowe. Informacje dotyczące testowania - patrz odpowiedź na pyt. 7.

W trakcie Konsultacji uzyskano również informację nt. możliwości podziału kosztów realizacji zamówienia na trzy kategorie:

- a) opłata za utworzenie samego Narzędzia – pojedyncza płatność
- b) koszty jego utrzymania związane z chmurą – rozliczenie miesięczne, w zależności od przyjętych założeń co do spodziewanego wolumenu dokumentów przetwarzanych w skali roku oraz zakładanej wydajności
- c) wsparcie techniczne i utrzymanie – rozliczenie miesięczne

W trakcie Konsultacji zwrócono uwagę Zamawiającego na zasadność posiadania przez wykonawców doświadczenia w konstrukcji, ewaluacji i wdrożeniu systemów wyszukiwawczych, które są zasilane przez zaawansowane technologie sztucznej inteligencji, ze szczególnym naciskiem na głębokie uczenie, doświadczenie w konstrukcji i wdrożeniu tzw. modeli językowych dedykowanych dla języka polskiego.

Zwrócono również uwagę, że Narzędzie będzie stale się uczyło na podstawie interakcji użytkowników, co umożliwi adaptację do indywidualnych potrzeb i preferencji.

3. W trakcie Konsultacji firma Billennium S.A. zwróciła się do Zamawiającego z pytaniami pomocniczymi dotyczącymi przedmiotu Konsultacji – nowej wyszukiwarki oraz odnoszącymi się do aktualnie udostępnianego przez Zamawiającego narzędzia. W dniu 5 lipca br. Zamawiający pocztą elektroniczną przekazał ww. firmie odpowiedzi na pytania. Poniżej Zamawiający zamieszcza treść ww. pytań i udzielone odpowiedzi:

- 1) Czy Państwa obecna wyszukiwarka orzeczeń ma jakieś konkretne ograniczenia lub problemy, które chcielibyście Państwo rozwiązać poprzez wprowadzenie wyszukiwania semantycznego?

Odp. W założeniu nowa wyszukiwarka powinna być intuicyjna, tak żeby umożliwić nawet osobom, które nie mają doświadczenia w zamówieniach publicznych łatwe

wyszukiwanie orzeczeń czy też zagadnień. Obecna wyszukiwarka całkiem dobrze wyszukuje orzeczenia po sygnaturze, dacie czy też członkach składu orzekającego. Pewnym problemem jest wyszukiwanie po frazach czy też słowach kluczowych. W tej kwestii bardzo liczymy na to, że wyszukiwanie semantyczne spowoduje, że nowa wyszukiwarka będzie jeszcze bardziej przydatnym i użytecznym narzędziem.

- 2) Czy istnieje konkretny format lub standard, w jakim są przechowywane orzeczenia w bazie danych, i czy istnieją wymagania dotyczące obsługi tego formatu przez nową wyszukiwarkę?

Odp. Silnik bazy PostgreSQL odpowiada za przechowywanie wszystkich danych systemu oraz przetwarzanie wprowadzonych tekstów orzeczeń i ich metryk na potrzeby wyszukiwania pełnotekstowego w tym z uwzględnieniem odmiany słów (fleksji). Nie istnieją szczególne wymagania dotyczące obsługi tego formatu przez nową wyszukiwarkę.

- 3) Czy istnieje konkretny format lub standard, w jakim są przechowywane orzeczenia w bazie danych, i czy istnieją wymagania dotyczące obsługi tego formatu przez nową wyszukiwarkę?

Odp. Ilość specyficznych terminów lub fraz właściwych dla dziedziny prawa jaką jest Prawo zamówień publicznych jest duża. W celach poglądowych w załączniku przekazuję indeks tematyczny zagadnień w ramach orzeczeń w porządku alfabetycznym, który stosowany jest przez nasz zespół do analizy orzecznictwa. Oczywiście nie chcielibyśmy zamykać się tylko na te zagadnienia, tak żeby umożliwić wyszukiwanie nawet tych orzeczeń, które dotyczą zagadnień dość rzadko się pojawiających lub nawet jednostkowych. *(patrz również pkt VI.4)*

- 4) Czy istnieje potrzeba wykorzystania meta-danych związanych z orzeczeniami, takich jak data wydania, sąd, czy rodzaj sprawy, w wynikach wyszukiwania?

Odp. Tak, przede wszystkim data ogłoszenia, sygnatura, przewodniczący składu orzekającego. Wykorzystanie takich meta-danych z powodzeniem funkcjonuje w obecnie działającej wyszukiwarce. W dalszej kolejności dobrze byłoby wykorzystać rozróżnienie na: rodzaj orzeczenia Izby (wyrok lub postanowienie), uchwały, postanowienie w sprawie wniosków o uchylenie zakazu zawarcia umowy oraz orzeczenia sądu okręgowego (wyrok lub postanowienia).

- 5) Jakie są oczekiwania dotyczące precyzji i dokładności wyników wyszukiwania semantycznego w kontekście orzeczeń?

Odp. Rezultaty powinny odpowiadać dokładnie temu, czego użytkownik szukał. System powinien rozumieć zapytanie użytkownika na wystarczającym poziomie, aby dostarczyć wyniki, które spełniają jego oczekiwania. Oczekuje się, że system będzie prezentował informacje, które są wiarygodne, aktualne i zgodne z intencjami

użytkownika. Wyniki powinny być nie tylko poprawne w sensie gramatycznym, ale również merytoryczne i zgodne z kontekstem pytania.

- 6) Czy wyszukiwarka powinna mieć możliwość wykonywania złożonych zapytań, takich jak łączenie różnych kryteriów wyszukiwania lub filtracji?

Odp. Tak, wyszukiwarka powinna mieć możliwość wykonywania złożonych zapytań, które umożliwiają łączenie różnych kryteriów wyszukiwania lub filtracji. Potrzeba ta jest spowodowana specyfiką wyszukiwanych treści i chęcią zawężenia wyników do określonych parametrów.

- 7) Czy wyszukiwarka powinna mieć możliwość wykonywania złożonych zapytań, takich jak łączenie różnych kryteriów wyszukiwania lub filtracji?

Odp. Najlepiej byłoby i tak i tak tj. wyszukujący po wyszukiwaniu powinien mieć dostęp do fragmentów lub streszczenia (jeśli narzędzie to umożliwi), ale gdyby wyszukującego taki fragment nie zadowolił, to powinien mieć szybki dostęp do pełnego dokumentu.

- 8) Czy są Państwo zainteresowani implementacją mechanizmu uczenia się maszynowego, który pozwoli na ciągłe doskonalenie wyników wyszukiwania w oparciu o interakcje użytkowników?

Odp. Tak, jesteśmy zainteresowani implementacją mechanizmu uczenia się maszynowego by dostarczyć użytkownikom bardziej trafne i zrozumiałe wyniki wyszukiwania.

- 9) Jaka jest oczekiwana liczba jednoczesnych użytkowników wyszukiwarki, i jak to wpłynie na wymagania dotyczące skalowalności i wydajności systemu?

Odp. Wyszukiwarka musi zapewniać skalowalność w celu udostępnienia optymalnych zasobów wystarczających do spełnienia parametrów świadczenia usługi 5000 jednoczesnych użytkowników wyszukiwarki.

- 10) Czy możecie Państwo dostarczyć informacje na temat średniej liczby stron w dokumentach orzeczeń oraz czy istnieje znaczące zróżnicowanie w liczbie stron między dokumentami? Ta informacja może być kluczowa dla oszacowania wymagań związanych z przetwarzaniem przy wykorzystaniu technologii embeddings. Z wykorzystaniem technologii np. embeddings lub równoważnej.

Odp. Istnieje zróżnicowanie co do ilości stron w dokumentach, ale nie jest ono równomierne. Chodzi o to, że zdecydowana większość dokumentów nie przekracza 20-30 stron. Przy czym duża część materiałów to postanowienia, te zwykle są dość krótkie (do 5 stron). Kolejna grupa to orzeczenia, uchwały i postanowienia w sprawie wniosków o uchylenie zakazu zawarcia umowy, które średnio mają do 30 stron. Oczywiście to zależy od sprawy i zdarzają się orzeczenia dłuższe tj. powyżej 30 stron, a czasami nawet powyżej 100 stron, jednak pozostają one w mniejszości. Orzeczenia ponad stu stronicowe pojawiają się bardzo rzadko.

4. Zamawiający w toku konsultacji odpowiadając na pytanie z pkt VI. 3.3) udostępnił firmie Billennium S.A. „Indeks tematyczny orzeczeń – nowe Pzp”, który stanowi załącznik do niniejszego protokołu.

VII. Termin zakończenia Wstępnych konsultacji rynkowych

Zamawiający zakończył Wstępne konsultacje rynkowe w dniu 17 lipca 2023 roku, zgodnie z terminem przewidzianym w Informacji o przedłużeniu wstępnych konsultacji rynkowych.

VIII. Informacja na temat potencjalnego wpływu Konsultacji na Opis przedmiotu, specyfikację warunków zamówienia oraz warunki umowy, środki jakie Zamawiający podjął w celu zapobieżenia zakłóceniu konkurencji.

W celu zapobieżenia zakłóceniu konkurencji Zamawiający upublicznia powyższy protokół na stronie internetowej Zamawiającego z uwzględnieniem wszystkich zapytań kierowanych do uczestników, udzielonych odpowiedzi oraz innych ustaleń i informacji wynikających z przeprowadzonych Konsultacji.

Niniejszy protokół jest jawny i podlega upublicznieniu na stronie internetowej Zamawiającego.

Andrzej Umiński; UZP
Elektronicznie podpisany przez
Andrzej Umiński; UZP
DN: cn=Andrzej Umiński; UZP,
givenName=Andrzej, sn=Umiński,
serialNumber=PNOPL-75020415914
,2.5.4.97=VATPL-5262239325,
o=URZĄD ZAMÓWIEN
PUBLICZNYCH, street=Postępu 17A,
postalCode=02-676, l=Warszawa,
st=mazowieckie, c=PL
Data: 2023.07.25 13:31:47 +02'00'

Andrzej Umiński -

Mariusz Gac; UZP
Elektronicznie podpisany
przez Mariusz Gac; UZP
Data: 2023.07.25 15:27:30
+02'00'

Mariusz Gac -

Jarosław Antosiewicz; UZP
Elektronicznie podpisany
przez Jarosław
Antosiewicz UZP
Data: 2023.07.25 15:38:39
+02'00'

Jarosław Antosiewicz -

Florczak Aneta
Elektronicznie
podpisany przez
Florczak Aneta
Data: 2023.07.25
13:24:02 +02'00'

Aneta Florczak -