

Nr postępowania: ZP/PN/01/20

Wykonawcy w postępowaniu

Dotyczy postępowania o udzielenie zamówienia publicznego w trybie przetargu nieograniczonego na Opracowanie elementów studium techniczno – użytkowego oraz dostarczenie i wdrożenie oprogramowania do integracji danych

Zamawiający, na podstawie art. 38 ust. 2 ustawy z dnia 29.01.2014 r. Prawo Zamówień Publicznych (tekst jednolity Dz. U. z 2019 r., poz. 1843 z późn. zm.), zwanej dalej ustawą Pzp, poniżej przedstawia treść pytań wraz z odpowiedziami do przedmiotowego postępowania:

1. Pragniemy zwrócić uwagę, że profesjonalnie działający na rynku Wykonawcy nie są w stanie oszacować wolumenu danych tylko na podstawie informacji o systemach czy w oparciu o dane publikowane w serwisach Otwarte Dane Gdańsk czy Otwarte Dane Gdynia. W szczególności wskazane przez Zamawiającego serwisy nie odnoszą się w żaden sposób do zbiorów danych wskazanych w grupie „Koleje” czy „Innobaltica”.

W szczególności na wolumen danych składają się takie elementy jak - sposób implementacji systemu, format danych w jakim będą one przekazywane oraz same wartości w bazie danych.

W związku z powyższym, prosimy o potwierdzenie, że Zamawiający pokryje wszystkie koszty związane z potrzebą skalowania środowiska w wersji produkcyjnej systemu, to jest po zakończeniu 6-miesięcznego Etapu wdrożeniowego.

Odpowiedź

Analogicznie, Zamawiający nie jest w stanie oszacować narzutu nakładanego na rozwiązanie Wykonawcy w zakresie przechowywania danych ani potrzeb sprzętowych. Jeżeli narzędzie będzie skalowalne to koszty związane ze skalowaniem środowiska w okresie powdrożeniowym są po stronie Zamawiającego.

2. Część z systemów wskazanych przez Zamawiającego może nie posiadać odpowiednich interfejsów do eksportu danych lub pobrania danych przez narzędzia ETL. Prosimy o potwierdzenie, że w przypadku kiedy pobrane danych za pomocą API lub dostępu bezpośredniego do bazy danych za pomocą ODBC nie będzie możliwe - odpowiedzialność za eksport danych do plików płaskich, CSV lub Excel będzie leżeć po stronie Zamawiającego i partnera będącego właścicielem/autorem systemu źródłowego.

Odpowiedź

Zamawiający przewiduje również inne metody dostępu - zdefiniowane zostały w definicji "Dostęp do danych". Jeśli żadna z wymienionych metod w definicji "Dostępu do danych" nie będzie mogła zostać użyta do przekazania danych Wykonawcy to Zamawiający przewiduje iż będzie generował pliki płaskie lub będzie musiał zrezygnować z pobierania i analizowania danych ze wskazanego systemu. Niemniej Zamawiający będzie dążył do wypracowania wspólnie z partnerami metod które pozwolą na pozyskanie danych. Jedną z możliwości które przewiduje Zamawiający jest możliwość wybudowania nowych interfejsów do istniejących systemów i koszt tych nowych interfejsów będzie leżał po stronie Zamawiającego.

3. W Załączniku nr.2 do OPZ podają Państwo listę systemów źródłowych, ale bez specyfikacji technicznej tj. bazy danych z których korzysta aplikacja, jej wersji oraz sposobu dostępu (API, pliki płaskie, webservice). Bez tych informacji nie jest możliwe odpowiednie oszacowanie pracochłonności pobrania informacji z tych źródeł. Dodatkowo zwracamy uwagę, że niektóre z tych rozwiązań mogą działać na różnych bazach danych, albo mogły być zmodyfikowane na potrzeby konkretnego wdrożenia. Uwzględniając powyższe, prosimy, aby Zamawiający działając zgodnie z art. 26 ust. 1 PZP przedstawił podstawowe informacje dot. systemów źródłowych obejmujące w szczególności (poza nazwą systemu)

co najmniej informację nt. bazy danych wykorzystywanej przez dany system, jej wersji, sposobu dostępu do danych systemu/ formatu danych.

Odpowiedź

Wg Zamawiającego to właśnie jest jednym z zadań Wykonawcy: zidentyfikowanie możliwości pobrania danych z danego Systemu jako podmiotu profesjonalnego działającego w obszarze pozyskiwania danych. Przy czym Zamawiający zakłada tak jak w odpowiedzi na pytanie 2 albo Zamawiający dostarczy pliki płaskie w ramach tych systemów lub będzie dążył do opracowania sposobu pozyskania danych lub zrezygnuje z wykorzystania danego systemu źródłowego.

4. Prosimy o wyjaśnienia jak należy rozumieć dostęp do źródła danych na poziomie użytkownika aplikacyjnego?

Prosimy o potwierdzenie, że będzie możliwe pobranie danych w sposób zautomatyzowany ze źródeł danych bez konieczności budowy dedykowanych konektorów do źródeł danych (np. w języku C czy C++) tzn. poprzez interfejs API.

Prosimy także o wskazanie dla każdego źródła rodzaju interfejsu (WebService, REST, SDK w danym języku).

Odpowiedź

Zamawiający nie może potwierdzić tak postawionego stwierdzenia. Jako użytkownika aplikacyjnego rozumie użytkownika, który z poziomu aplikacji jest w stanie generować pliki płaskie.

5. Prosimy o potwierdzenie, że przez kopię i lustra baz danych systemów źródłowych rozumiemy takie rozwiązania technologiczne do których można się dostać za pomocą konektora ODBC, tzn. nie zachodzi potrzeba uruchamiania środowiska bazy danych w środowisku chmurowym i odtwarzania bazy danych z pliku backupu (np. Odtworzenie bazy danych postgresql w chmurze)?

Jeżeli zaistnieje taka potrzeba to prosimy o wskazanie baz danych oraz ich wersji, które mają być odzyskiwane (np. baza Isoke, postgresql, v.11)

Odpowiedź

W wybranych wypadkach będzie zachodziła potrzeba odtwarzania kopii baz danych.

6. Zamawiający wskazuje jako systemy źródłowe - system Interconnet oraz PZUM, szynę danych.

Prosimy o:

a) informację jak będzie wyglądać integracja z systemem Interconnect, jakie dane będą z tego systemu pobierane

b) potwierdzenie, że integracja z systemem PZUM będzie realizowana tylko i wyłącznie poprzez szynę danych i usługi realizowane po stronie systemu PZUM

c) potwierdzenie, że integracja z systemem PZUM jest potrzebna do sprawdzenia jakości danych w systemie PZUM podczas migracji oraz podczas pracy produkcyjnej systemu PZUM

d) potwierdzenie, że rozwiązanie dostarczone przez zamawiającego nie będzie służyło w celu ciągłej aktualizacji danych z innych źródeł danych do PZUM podczas pracy produkcyjnej systemu PZUM tzn. jedyne przeniesienie danych z hurtowni danych do PZUM będzie się odbywać podczas migracji (nie będzie ciągłej integracji)

Odpowiedź

a) Dane w formacie GTFS oraz NeTEX. Integracja poprzez API.

b) Tak. Poprzez szynę i/lub usługi.

c) Tak.

d) Nie potwierdzamy, jeżeli narzędzie będzie wydajne rozważamy ich permanentne użycie również w trakcie produkcyjnego użytkowania Systemu.

7. W celu umożliwienia Wykonawcy wyceny pracochłonności implementacji reguł weryfikacji prosimy o wskazanie ile maksymalnie reguł walidacji wymagają Państwo do realizacji w modelu referencyjnym w

INNOBALTICA SP. Z O.O. 80-067 GDAŃSK, UL. RÓWNA 19/21

NIP 957-10-03-404, REGON 220639884, BANK PKO BP SA 63 1020 1811 0000 0102 0311 6969

KRS 0000311943, VII WYDZIAŁ GOSPODARCZY KRAJOWEGO REJESTRU SĄDOWEGO KAPITAŁ ZAKŁADOWY 21 755 000,00 PLN



ramach przedmiotu zamówienia. Prosimy o podanie sumarycznej liczby reguł dla wszystkich 1005 atrybutów (z uwzględnieniem, że być może niektóre atrybuty mogą mieć więcej niż 1 regułę). Brak informacji o maksymalnej liczbie reguł czyni opis przedmiotu zamówienia niejednoznaczny i nieprecyzyjny uniemożliwiając przygotowanie rzetelnej i porównywalnej z innymi Wykonawcami oferty.

Odpowiedź

Wykonawca jako podmiot profesjonalny powinien oszacować pracochłonność implementacji reguł weryfikacji bazując na swoim doświadczeniu oraz na możliwości dostępu do danych otwartych.

8. Prosimy o określenie stopnia skomplikowania reguł walidacyjnych koniecznych do implementacji w ramach przedmiotu zamówienia. Prosimy o podanie kilku przykładowych najbardziej skomplikowanych reguł jakich implementacji oczekuje Zamawiający.

Brak informacji o stopniu skomplikowania reguł walidacyjnych przy ich potencjalnie dużej liczbie czyni opis przedmiotu zamówienia niejednoznaczny i nieprecyzyjny uniemożliwiając przygotowanie rzetelnej i porównywalnej z innymi Wykonawcami oferty.

Odpowiedź

Przykładowe kryterium weryfikacji:

Data aktualizacji pliku źródłowego musi wskazywać dzień bieżący, aby pobrać go do systemu docelowego.

ALGORYTM DZIAŁANIA:

Plik źródłowy: 'Rozkład jazdy GTFS' dostępny jest w postaci archiwum .zip. i znajduje się do pobrania pod nw. adresem:

<https://ckan.multimediasgdansk.pl/dataset/c24aa637-3619-4dc2-a171-a23eec8f2172/resource/30e783e4-2bec-4a7d-bb22-ee3e3b26ca96/download/gtfsgoogle.zip>

Data aktualizacji jest wskazana w polu: Ostatnia modyfikacja na Platformie CKAN.

Zakłada się, że źródło danych w CKAN będzie sprawdzane pod kątem daty aktualizacji cyklicznie od północy do 7:30.

Ze wskazanego wyżej zasobu pobieramy listę słupków przystankowych znajdującą się w pliku: „stops.txt”.

Plik zawiera dane na 14 dni wprzód licząc od dnia bieżącego.

Każdy wiersz z pobranego pliku reprezentuje pojedynczy słupek przystankowy.

Uzyskujemy następujące atrybuty:

- a) stop_id,
- b) stop_code,
- c) stop_name,
- d) stop_lat,
- e) stop_lon,

zgodnie z tabelą „Standaryzacja danych2”.

Przykładowy preferowany sposób naprawy niezgodności:

Jeśli data aktualizacji pliku źródłowego nie wskazuje na dzień bieżący, należy:

1. skontaktować się z ZTM w Gdańsku;
2. przejść do KJD001.1.

Przykładowy alternatywny sposób naprawy niezgodności:

Skorzystanie z danych pobranych w dniu, w którym ostatnio spełniono kryterium weryfikacji.

9. W celu umożliwienia prawidłowej wyceny pracochłonności implementacji reguł walidacji i reguł czyszczenia danych prosimy o wyjaśnienie:

INNOBALTICA SP. Z O.O. 80-067 GDAŃSK, UL. RÓWNA 19/21

NIP 957-10-03-404, REGON 220639884, BANK PKO BP SA 63 1020 1811 0000 0102 0311 6969

KRS 0000311943, VII WYDZIAŁ GOSPODARCZY KRAJOWEGO REJESTRU SĄDOWEGO KAPITAŁ ZAKŁADOWY 21 755 000,00 PLN



Fundusze Europejskie
Infrastruktura i Środowisko



Rzeczpospolita
Polska

Unia Europejska
Fundusz Spójności



- a) czy istnieje znaczna różnica między regułami walidacji, a regułami czyszczenia danych tzn. czy możliwe jest wykorzystanie reguł walidacji do czyszczenia danych bez modyfikacji
- b) jeżeli konieczne są modyfikacje reguł walidacji do reguł czyszczenia danych, to prosimy o wskazanie liczby reguł które wymagają znacznej modyfikacji
- c) jeżeli konieczne są implementacje reguł czyszczenia danych niezależnie od reguł walidacji to prosimy o wskazanie liczby takich reguł.

Odpowiedź

- a) **Każdy przypadek należałoby rozpatrywać indywidualnie. Zamawiający przewiduje że istnieje taka możliwość ale nie we wszystkich przypadkach.**
- b) **Zamawiający nie potrafi jednoznacznie określić tej wartości bez uprzedniej analizy na poziomie atrybutu w systemie.**
- c) **W większości przypadków będą powiązane ze sobą.**

10. Prosimy o określenie stopnia skomplikowania reguł czyszczenia danych koniecznych do implementacji w ramach przedmiotu zamówienia. Prosimy o podanie kilku przykładowych najbardziej skomplikowanych reguł jakich implementacji oczekuje Zamawiający. Brak informacji o stopniu skomplikowania reguł czyszczenia danych przy ich potencjalnie dużej liczbie czyni opis przedmiotu zamówienia niejednoznaczny i nieprecyzyjny uniemożliwiając przygotowanie rzetelnej i porównywalnej z innymi Wykonawcami oferty.

Odpowiedź

Powtórzone pytanie. Odpowiedź ta sama co na pytanie 8.

11. Prosimy o potwierdzenie, że jeżeli w Karcie Jakości Danych jako metoda czyszczenia danych zostanie wskazana metoda ręczna to taka metoda nie podlega implementacji jako reguła czyszczenia danych i takie ręczne czyszczenie danych będzie wykonane przez Zamawiającego a nie przez Wykonawcę

Odpowiedź

Ręcznie jeżeli będzie to jedyna pozostała opcja i Wykonawca tego dowiedzie. Czyszczenie danych ręcznie, realizowane w systemach źródłowych będzie w gestii Zamawiającego i jego partnerów.

12. Prosimy o wskazanie ile sumarycznie Kart Jakości Danych Zamawiający wymaga do przygotowania dla artefaktów i atrybutów w ramach przedmiotu zamówienia.

Odpowiedź

Wykonawca jako podmiot profesjonalny powinien oszacować ilość Kart Jakości Danych bazując na swoim doświadczeniu oraz na możliwości dostępu do danych otwartych.

13. Prosimy o potwierdzenie, że Zamawiający uzgodni z Wykonawcą sposób dostarczenia Kart Jakości Danych w takim formacie , aby zagwarantować maksymalnie prosty sposób ich importu do systemu (np. w postaci plików Excel).

Odpowiedź

Tak

14. Prosimy o wskazanie ile maksymalnie sumarycznie wskaźników migracji Zamawiający wymaga do realizacji w ramach przedmiotu Zamówienia.

Odpowiedź

Maksymalnie 30.

INNOBALTICA SP. Z O.O. 80-067 GDAŃSK, UL. RÓWNA 19/21
NIP 957-10-03-404, REGON 220639884, BANK PKO BP SA 63 1020 1811 0000 0102 0311 6969
KRS 0000311943, VII WYDZIAŁ GOSPODARCZY KRAJOWEGO REJESTRU SĄDOWEGO KAPITAŁ ZAKŁADOWY 21 755 000,00 PLN



15. Prosimy o określenie stopnia skomplikowania wskaźników migracji koniecznych do implementacji w ramach przedmiotu zamówienia. Prosimy o podanie kilku przykładowych najbardziej skomplikowanych wskaźników jakich implementacji oczekuje Zamawiający. Brak informacji o stopniu skomplikowania wskaźników migracji czyni opis przedmiotu zamówienia niejednoznaczny i nieprecyzyjny uniemożliwiający przygotowanie rzetelnej i porównywalnej z innymi Wykonawcami oferty.

Odpowiedź

Przykładowy WM1:

Określenie liczby słupków przystankowych przeznaczonych do migracji powinno się odbywać poprzez:

Sprawdzenie, czy data aktualizacji zasobu na platformie CKAN zawierającego plik GTFS wskazuje na dzień bieżący

Jeśli pkt 1 jest spełniony: pobranie pliku GTFS zawierającego rozkłady jazdy

Wybranie pliku „stops.txt”

Sprawdzenie liczby słupków (suma wierszy w pliku GTFS), jaka tam występuje

Porównanie liczby słupków przeznaczonych do migracji z systemu źródłowego (ustalonej według wyżej wymienionego algorytmu) z liczbą słupków przeniesionych do systemu docelowego.

Miejsce: Otwarte dane ZTM w Gdańsku

Zasób: Rozkład jazdy GTFS

Plik: „stops.txt”

Przykładowy WM2:

Aby porównać pozycje GPS słupków przystankowych występujących między systemem źródłowym a systemem docelowym wykonać należy wykonać następujące kroki:

Uzyskanie 100% zgodności w WM1;

Wykorzystanie pliku „stops.txt” pobranego w WM1;

Sprawdzenie pól „lat” i „lon” występujących dla każdego stopID słupka występującego w ww. pliku źródłowego.

Sprawdzenie pól „lat” i „lon” występujących dla każdego ID słupka występującego w ww. pliku źródłowym z ich odpowiednikami znajdującymi się w systemie docelowym.

Miejsce: Otwarte dane ZTM w Gdańsku

Zasób: Rozkład jazdy GTFS

Plik: „stops.txt”

Pola: stopID, lat, lo

INNOBALTICA SP. Z O.O. 80-067 GDAŃSK, UL. RÓWNA 19/21

NIP 957-10-03-404, REGON 220639884, BANK PKO BP SA 63 1020 1811 0000 0102 0311 6969

KRS 0000311943, VII WYDZIAŁ GOSPODARCZY KRAJOWEGO REJESTRU SĄDOWEGO KAPITAŁ ZAKŁADOWY 21 755 000,00 PLN



**W związku z powyższym zamawiający przedłuża termin składania ofert do dnia 04.12.2020 r.
Godziny składania i otwarcia ofert pozostają bez zmian**

/-/ Grzegorz Beblowski

Przewodniczący Komisji Przetargowej

INNOBALTICA SP. Z O.O. 80-067 GDAŃSK, UL. RÓWNA 19/21
NIP 957-10-03-404, REGON 220639884, BANK PKO BP SA 63 1020 1811 0000 0102 0311 6969
KRS 0000311943, VII WYDZIAŁ GOSPODARCZY KRAJOWEGO REJESTRU SĄDOWEGO KAPITAŁ ZAKŁADOWY 21 755 000,00 PLN



**Rzeczpospolita
Polska**

Unia Europejska
Fundusz Spójności



Wylączna odpowiedzialność za treść niniejszej publikacji leży po stronie jej autorów. Jej treść nie musi odzwierciedlać opinii Unii Europejskiej. Ani Europejski Bank Inwestycyjny ani Komisja Europejska nie są odpowiedzialne za jakiegokolwiek wykorzystanie zawartych tu informacji.

Współfinansowane przez program ramowy Unii Europejskiej „Horyzont 2020”

