

Index Copernicus Sp. z o.o.

ul. Żurawia 6/12, 00-503 Warszawa

REGON 145900211, NIP 7010326698

KRS 0000403736
Pieczęć Zamawiającego

ZAPYTANIE OFERTOWE NR 1/2017

W związku z planowaną realizacją Projektu pn. "Utworzenie Centrum Badawczo-Rozwojowego w przedsiębiorstwie Index Copernicus Sp. z o.o. w celu prowadzenia prac badawczo-rozwojowych nad innowacyjnymi technologiami informatycznymi" w ramach Programu Operacyjnego Inteligentny Rozwój, 2014-2020 działanie 2.1 „Wsparcie inwestycji w infrastrukturę przedsiębiorstw” firma

INDEX COPERNICUS SP. Z O. O.

zaprasza do złożenia ofert na wykonanie zamówienia pod nazwą: Zakup 6 (sześciu) modułów programistycznych stanowiących bazę Centrum Badawczo – Rozwojowego dla prowadzenia prac badawczo – rozwojowych.

I. Określenie przedmiotu zamówienia:

Przedmiotem zamówienia jest wytworzenie, przeniesienie praw autorskich oraz uruchomienie specjalistycznego systemu informatycznego (zwanego dalej systemem informatycznym laboratorium) stanowiącego szkielet infrastruktury badawczej, która umożliwi prowadzenie prac badawczo-rozwojowych z zakresu zastosowania sieci neuronowych do ekstrakcji wiedzy z dużych zbiorów danych, wyznaczania Potencjału Zasobu Wiedzy oraz tworzenia profili semantycznych dla zasobów wiedzy. Wykonawca dostarczy również kompletne, opisane kody źródłowe gotowe do skompilowania wraz z dokumentacją w języku UML lub równoważnym.

Dostarczony w ramach przedmiotu zamówienia system informatyczny ma umożliwić odtwarzanie w warunkach laboratoryjnych ekosystemu, jakim jest sektor nauki – w formie stale rozwijanych oraz udoskonalanych baz danych zawierających informacje o kluczowych obiektach biznesowych typu podmiotowego (instytucje ekosystemu nauki, ludzie - naukowcy); typu przedmiotowego (nośniki Zasobów Wiedzy – czasopisma naukowe, książki, projekty, komponenty infrastruktury badawczej); zdarzeniach ewaluacyjnych (częstkowe aktywności wykonywane na co dzień przez obiekty biznesowe typu podmiotowego, rozpięte na obiektach biznesowych.

Dostarczony w ramach przedmiotu zamówienia system informatyczny powinien charakteryzować się modelem danych możliwym do ciągłego rozwijania i modyfikacji, w szczególności poprzez możliwość łatwego definiowania oraz implementacji reprezentacji cyfrowej nowych obiektów biznesowych i zdarzeń ewaluacyjnych – zachowując integralność rozwijanego modelu danych w ramach poszczególnych modułów.

Dostarczony w ramach przedmiotu zamówienia system informatyczny laboratorium będzie zasilany z dużej liczby źródeł danych o różnym poziomie ufności i kompletności, zarówno z wewnętrznych baz danych Zamawiającego, naukowych baz danych i repozytoriów krajowych oraz innych źródeł. System informatyczny laboratorium powinien radzić sobie z agregacją danych z różnych źródeł, ich wartościowaniem, deduplikacją oraz walidacją.

System informatyczny laboratorium, powinien umożliwić utrzymywanie oraz stałe rozwijanie referencyjnych baz danych poszczególnych obiektów biznesowych (instytucje, naukowcy, czasopisma naukowe, projekty) oraz zdarzeń ewaluacyjnych (zgodnie z definicją z *Rozporządzenia w sprawie przyznawania kategorii naukowej jednostkom naukowym i uczelniom, w których zgodnie z ich statutami nie wyodrębniono podstawowych*

jednostek organizacyjnych), nadawać poszczególnym obiektom biznesowym i zdarzeniom ewaluacyjnym unikalne numery identyfikacyjne oraz przechowywać analogi tych identyfikatorów z zewnętrznych baz danych (źródeł), a przez to umożliwiać synchronizację danych o obiektach.

System powinien umożliwić prowadzenie badań nad wybranymi porcjami danych, a w odpowiedzi na tę potrzebę stworzyć dedykowane do poszczególnych badań bazy danych oraz bazy danych wynikowych.

System powinien realizować procesy big data w zakresie klasyfikacji treści, ekstrakcji obiektów biznesowych i metadanych oraz wskazywania powiązań między obiektami biznesowymi i zdarzeniami ewaluacyjnymi w koncepcji profili zasobów wiedzy.

KODY CPV:

48000000-8 Pakiety oprogramowania i systemy informatyczne

48517000-5 Pakiety oprogramowania informatycznego

72000000-5 Usługi informatyczne: konsultacyjne, oporacowywania oprogramowania, internetowe i wsparcia

72212460-1 Usługi opracowywania oprogramowania analitycznego, naukowego, matematycznego lub prognozującego

Opis przedmiotu zamówienia oraz oczekiwane rezultaty przedstawione zostały w Załączniku nr 1.

Szczegółowy harmonogram prac został zawarty w Załączniku nr 2.

Termin realizacji zamówienia - do 31.12.2018 r.

II. Istotne warunki zamówienia:

- 2.1. Oferenci przed zawarciem umowy realizacji Zamówienia, zawrą z Zamawiającym klauzulę NDA (umowa o poufności) celem zabezpieczenia tajemnic przedsiębiorstwa Zamawiającego. Zawarcie umowy o poufności jest warunkiem do zapoznania się ze szczegółową specyfikacją zamówienia i podpisania umowy realizacji zamówienia.
- 2.2. Wykonawca przeniesie na Zamawiającego majątkowe prawa autorskie do wytworzonych dokumentów, analiz i zestawień oraz oprogramowania wraz z prawami pokrewnymi na wszystkich znanych na dzień zawarcia umowy polach eksploatacji.
- 2.3. Dokumenty przygotowane i przekazywane przez Wykonawcę będą w formie elektronicznej, w formatach obsługiwanych przez powszechnie dostępne aplikacje biurowe (Office, Open Office lub równoważne), a w przypadku konieczności korzystania przy sporządzaniu dokumentacji z oprogramowania specjalistycznego – w formie elektronicznych wydruków .pdf. Wdrożenie oprogramowania odbędzie się w biurze Zamawiającego na nośnikach wskazanych przez Zamawiającego.
- 2.4. Wykonane prace zostaną przeanalizowane przez zespół Zamawiającego. W przypadku stwierdzenia uchybień i niezgodności ze specyfikacją, Wykonawca zostanie wezwany do usunięcia wad, a po ich usunięciu, nastąpi ponowny odbiór (w razie konieczności czynność będzie powtarzana do uzyskania zadowalającego wyniku).
- 2.5. Wykonanie zamówienia niezgodnie ze szczegółową specyfikacją będzie podstawą dla Zamawiającego do egzekwowania kar umownych określonych w Umowie.
- 2.6. Wykonawca, w ramach umowy zapewni uczestnictwo w konsultacjach/wdrożeniu systemu (w ramach swojego wynagrodzenia (cena)) w siedzibie Zamawiającego do dnia zakończenia projektu i uruchomienia pełnej funkcjonalności systemu tj. do dnia 31 grudnia 2018 r.
- 2.7. Zaproponowane przez Oferentów rozwiązanie musi odpowiadać funkcjonalnościom i parametrom określonym w punktach dokumentacji projektowej dotyczących technologii informatycznej oraz przebiegowi realizacji – Załącznik nr 1. oraz Załącznik nr 2.

- 2.8. Zasady płatności: Wynagrodzenie będzie wypłacane w ratach za każdy prawidłowo zrealizowany i oddany bez zastrzeżeń etap prac z zastrzeżeniem, iż faktura końcowa (po zdaniu i odbiorze przez Zamawiającego) nie może być mniejsza niż 10% wartości zamówienia. Wykonawca dopuszcza płatności zaliczkowe.
- 2.9. Realizacja prac: Wykonawca zapewni sposób realizacji prac w metodyce zwinnej SCRUM, spotkania SCRUM odbywać się będą nie rzadziej, niż raz na dwa tygodnie w siedzibie Zamawiającego lub w innym wskazanym przez Zamawiającego miejscu. Wykonawca, w ramach swojego wynagrodzenia (cena) zapewni obecność osób wykonujących oprogramowanie na spotkaniach SCRUM. Całość prac w zakresie planowania, rejestrowania postępów oraz deponowania kodu odbywać się będzie we wskazanym przez Zamawiającego środowisku Atlassian (Confluence, JIRA, Bitubcket).
- 2.10. Odbiór i realizacja prac: Zamawiający wyznaczy koordynatora współpracy, którego zadaniem będzie usprawnienie procesu przekazywania informacji pomiędzy Wykonawcą, Zamawiającym, a partnerami projektu. Rolą koordynatora będzie także ocena poziomu zaawansowania prac i wyrażanie zgody na wystawianie faktur częściowych. Odbiór końcowy nastąpi na podstawie protokołu zdawczo-odbiorczego, sporządzonego na piśmie.
- 2.11. Sposób obliczenia ceny: Wykonawca zobowiązany jest do wskazania ceny netto za każde z zadań w złotych polskich (lub innej walucie) wraz z podaniem należnego podatku VAT za kompleksową realizację każdego z zadań w formularzu ofertowym wg wzoru Zamawiającego. Cena musi obejmować wszystkie koszty jakie Wykonawca musi ponieść w związku z realizacją Zamówienia tj. podatki, koszty dojazdu, sporządzenie dokumentacji itd. Cenę należy podać z dokładnością do dwóch miejsc po przecinku.

2.12. Kryteria wyboru ofert:

Przy wyborze najkorzystniejszej oferty Zamawiający będzie się kierował następującymi kryteriami i ich wagami:

Cena oferty	–	50%
Kryteria funkcjonalne	–	50%

Wybrana zostanie oferta, która otrzyma maksymalną liczbę punktów przyznawanych według wzoru:

$$Pc = PcOn + PfOn$$

Gdzie:

Pc – punktacja całkowita

PcOn – punktacja za cenę oferenta

PfOn – punktacja za kryteria funkcjonalne deklarowane przez Oferenta

Punktacja za cenę będzie liczona według następującego wzoru:

$$PcOn = (Cmin/CO_n) * 50$$

Gdzie:

CO_n – cena n-tego Oferenta

Cmin – najniższa oferowana cena

Punktacja za kryteria funkcjonalne będzie liczona według następującego wzoru:

$$PfOn = \text{suma}(PF_n)$$

Gdzie:

PFn – punkty za poszczególne kryterium funkcjonalne od F1 do F12 na podstawie deklaracji Oferenta w formularzu oferty

Zasady punktacji w kryteriach funkcjonalnych

Kryteria funkcjonalne oferty przedstawiają kluczowe parametry przedmiotu zamówienia, które Zamawiający chce osiągnąć w celu zapewnienia maksymalnej wydajności pracy centrum badawczo-rozwojowego. Dla każdego z kryteriów funkcjonalnych F1-F12 Oferent deklaruje (wybiera) tylko jedną z predefiniowanych wartości (lub predefiniowanego zakresu) z kolumny Punkty (PFn), którą osiągnie system informatyczny będący przedmiotem zamówienia, dostarczony przez Oferenta.

Numer kryterium	Kryterium	Jednostka	Wartość	Punkty (PFn)
F1	Ekstrakcja obiektów typu artykuł ze zbiorów liczących 6.000.000 rekordów pozycji bibliografii załącznikowej, w których artykuły naukowe występują w ciągu znaków jako tekst	liczba unikalnych artykułów referencyjnych z nadanym unikalnym nr GICID (z poprawnością powyżej 98%)	> 300 000	5,00
			200 000 - 300 000	3,00
			100 000 - 200 000	1,50
			< 100 000	0,00
F2	Ekstrakcja obiektów typu autor ze zbiorów liczących 6.000.000 rekordów pozycji bibliografii załącznikowej, w których autorzy występują w ciągu znaków jako tekst	liczba referencji z wykrytymi autorami ze wskazanym przynajmniej jednym imieniem i nazwiskiem (o poprawności powyżej 99%)	> 5 000 000	5,00
			4 500 000 - 5 000 000	3,00
			3 000 000 - 4 500 000	1,50
			< 3 000 000	0,00
F3	Ekstrakcja rodzaju nośnika (książka, czasopismo, monografia naukowa, konferencja, strona www itp.) ze zbiorów liczących 6.000.000 rekordów pozycji bibliografii załącznikowej, w której nośnik występuje w ciągu znaków jako tekst	liczba referencji z wykrytym nośnikiem bibliograficznym (o poprawności powyżej 99%)	> 3 000 000	3,00
			2 500 000 - 3 000 000	2,00
			2 000 000 - 2 500 000	1,00
			1 500 000 - 2 000 000	0,50
			<1 500 000	0,00
F4	Ekstrakcja metadanych z danych źródłowych pochodzących z importu ciągłego oraz wsadowego na zadanej próbie 6.000.000	liczba poprawnie wskazanych typów obiektów w 90%	\geq 3	5,00
			2	2,00

	rekordów	rekordów (poprawność powyżej 95%)	1	0,50
			BRAK	0,00
F5	Baza profili unikalnych (zdeduplikowanych) autorów z przypisanymi im powiązanymi unikalnymi (zdeduplikowanymi) obiektami typu artykuł wyekstrahowanych ze zbiorów liczących 6.000.000 rekordów pozycji bibliografii załącznikowej, w których zarówno autorzy, jak i artykuły występują w ciągu znaków jako tekst	liczba unikalnych profili naukowców z nadanym unikalnym numerem GICID, zawierających przynajmniej 10 powiązanych publikacji	> 100 000	5,00
			80 000 - 100 000	3,00
			60 000 - 80 000	1,50
			20 000 - 60 000	0,50
			<20 000	0,00
F6	Baza unikalnych obiektów typu monografia z przypisanymi powiązaniem obiektów typu autor oraz redaktor ze zbioru liczącego 650.000 rekordów zdarzeń ewaluacyjnych typu monografia i rozdział w monografii	liczba unikalnych monografii z nadanym unikalnym numerem GICID (poprawność powyżej 95%)	> 100 000	5,00
			80 000 - 100 000	3,00
			60 000 - 80 000	1,00
			40 000 - 60 000	0,50
			<40 000	0,00
F7	Efektywność klasyfikacji wykonywana na oddzielnych zbiorach testowych, które nie będą w sobie zawierały próbek ze zbiorów treningowych (jakość klasyfikacji dla 1 000 000 próbek, wartość uśredniona dla 10 testów)	poprawne odpowiedzi	> 98%	6,00
			97% - 98%	5,00
			96% - 97%	4,00
			95% - 96%	3,00
			90% - 95%	1,00
			<90%	0,00
F8	Efektywność klasyfikacji wykonywana na oddzielnych zbiorach testowych, które nie będą w sobie zawierały próbek ze zbiorów treningowych (dopasowanie się sieci do nieznannej funkcji za pomocą błędu MSE (błąd średnio-kwadratowy) dla 1.000.000 próbek, wartość uśredniona dla 10 testów)	wartość z ostatnich iteracji	< 0.05	4,00
			0.05 - 0.1	3,00
			0.1 - 0.2	2,00
			0.2 - 0.3	0,75
			0.3 - 0.4	0,50
			> 0.4	0,00
F9	Liczba zagranicznych repozytoriów danych o	liczba repozytoriów	4	4,00

	osiągnięciach naukowych będących zdarzeniami ewaluacyjnymi, zawierających przynajmniej 1.000.000 zwalidowanych rekordów z lat 2010-2015, z którymi system synchronizuje dane w procesie ciągłym	międzynarodowych	3	2,00
			2	0,25
			<2	0,00
F10	Liczba krajowych repozytoriów danych o osiągnięciach naukowych będących zdarzeniami ewaluacyjnymi, zawierających przynajmniej 1.000.000 zwalidowanych rekordów z lat 2010-2015, z którymi system synchronizuje dane w procesie ciągłym	liczba repozytoriów krajowych	3	4,00
			2	2,00
			1	0,25
			BRAK	0,00
F11	Gwarancja - Maksymalny czas naprawy błędu krytycznego	dni robocze	do 1 dnia	2,00
			do 2 dni	1,50
			do 3 dni	0,50
			> 3 dni	0,00
F12	Gwarancja - Maksymalny czas naprawy błędu niekrytycznego (pozostałe)	dni robocze	do 2 dnia	2,00
			do 3 dni	1,50
			do 5 dni	0,50
			> 5 dni	0,00

Zamawiający określa minimalny wymóg punktowy dla oceny kryteriów funkcjonalnych. Oferty, które w ocenie kryteriów funkcjonalnych nie osiągną minimalnego wymogu punktowego sumy punktów PfOn = 20 zostaną odrzucone jako niespełniające wymagań zamówienia (niezgodne ze specyfikacją).

III. Warunki udziału w postępowaniu:

O udzielenie Zamówienia mogą ubiegać się Oferenci, którzy:

1. Posiadają wymagane przepisami uprawnienia do wykonania zamówienia.
2. Posiadają niezbędną wiedzę i doświadczenie oraz dysponują potencjałem technicznym niezbędnym do wykonania zamówienia.
3. Znajdują się w sytuacji ekonomicznej i finansowej zapewniającej wykonanie zamówienia, w tym w szczególności nie posiadają zaległości w regulowaniu zobowiązań podatkowych oraz opłaceniu składek ZUS.
4. W okresie ostatnich 5 lat wykonali oraz wdrożyli przynajmniej jeden system teleinformatyczny realizujący proces gromadzenia, przetwarzania danych o osiągnięciach naukowych oraz parametryzacji osiągnięć naukowych na podstawie Rozporządzenia Ministra Nauki i Szkolnictwa Wyższego w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym z dnia 13 lipca 2012 roku lub (lub nowszym).

IV. Zakres wykluczenia:

Oferentami nie mogą być podmioty powiązane z Zamawiającym osobowo lub kapitałowo. Przez powiązania kapitałowe lub osobowe rozumie się wzajemne powiązania między beneficjentem lub

osobami upoważnionymi do zaciągania zobowiązań w imieniu beneficjenta lub osobami wykonującymi w imieniu beneficjenta czynności związane z przeprowadzeniem procedury wyboru wykonawcy a wykonawcą, polegające w szczególności na:

- a) uczestniczeniu w spółce jako wspólnik spółki cywilnej lub spółki osobowej,
- b) posiadaniu co najmniej 10% udziałów lub akcji, o ile niższy próg nie wynika z przepisów prawa lub nie został określony przez IZ w wytycznych programowych,
- c) pełnieniu funkcji członka organu nadzorczego lub zarządzającego, prokurenta, pełnomocnika,
- d) pozostawaniu w związku małżeńskim, w stosunku pokrewieństwa lub powinowactwa w linii prostej, pokrewieństwa drugiego stopnia lub powinowactwa drugiego stopnia w linii bocznej lub w stosunku przysposobienia, opieki lub kurateli.

V. Miejsce i termin składania ofert:

1. Ofertę należy dostarczyć do siedziby Zamawiającego lub przesłać pocztą tradycyjną/kurierską na adres: ul. Sienna 86/74, 00-815 Warszawa lub przesłać pocztą elektroniczną na adres: agata.przytucka@indexcopernicus.com
2. Ofertę należy złożyć w nieprzekraczalnym terminie do dnia 18.12.2017 r. W przypadku wysłania oferty za pośrednictwem poczty tradycyjnej lub kurierskiej liczy się data dostarczenia przesyłki do biura Zamawiającego.
3. Oferty złożone po terminie nie będą rozpatrywane.
4. Ocena ofert zostanie dokonana w siedzibie Zamawiającego do dnia 19.12.2017 r.
5. Wyniki dotyczące wyboru najkorzystniejszej oferty zostaną zamieszczone na stronie internetowej Zamawiającego.
6. Zamawiający w terminie do 21.12.2017 r. zawiadomi wybranego Wykonawcę o wynikach przeprowadzonego postępowania.
7. Dodatkowych informacji na temat Projektu udziela Pani Agata Przytucka, tel. +48 22 487 53 93, mail: agata.przytucka@indexcopernicus.com
8. Zapytanie ofertowe zamieszczono na stronie www.indexcopernicus.com

VI. Sposób przygotowania ofert:

1. Każdy Oferent może złożyć tylko jedną ofertę.
2. Ofertę należy sporządzić na formularzu ofertowym, którego wzór stanowi Załącznik nr 3 do Zapytania ofertowego.
3. Oferta musi być napisana w języku polskim i podpisana przez osobę upoważnioną do reprezentowania Oferenta.
4. Oferta musi zawierać: cenę ofertową, cenę netto oraz należny podatek VAT, za realizację przedmiotu zamówienia zgodnie z warunkami określonymi w zapytaniu.
5. Oferta musi być opatrzona pieczętą firmową Oferenta (a w razie braku – czytelnym podpisem), posiadać datę sporządzenia, zawierać adres lub siedzibę Oferenta, numer telefonu, adres e-mail, numer NIP.
6. Oferta musi być ważna przez **okres co najmniej 30 dni**, liczonych od dnia upływu terminu składania ofert.
7. Do oferty należy załączyć Oświadczenie potwierdzające posiadanie niezbędnej wiedzy i doświadczenia oraz dysponowania potencjałem technicznym niezbędnym do wykonania zamówienia (Załącznik nr 4). Oświadczenie musi być podpisane przez osobę upoważnioną do reprezentacji Oferenta oraz opatrzone pieczętą firmową (a w razie braku – dodatkowo czytelnym podpisem).

8. Do oferty należy załączyć oświadczenie potwierdzające sytuację ekonomiczną i finansową Oferenta zapewniającą wykonanie zamówienia (Załącznik nr 5). Oświadczenie musi zawierać informacje dotyczące Oferenta w zakresie, że:
 - nie jest zagrożony upadłością, nie otwarto jego likwidacji ani nie ogłoszono upadłości,
 - nie zalega z uiszczeniem należności wobec Skarbu Państwa.Oświadczenie musi być podpisane przez osobę upoważnioną do reprezentacji Oferenta oraz opatrzone pieczętką firmowa (a w razie braku – dodatkowo czytelnym podpisem).
9. Do oferty należy załączyć oświadczenie O spełnieniu warunków udziału w postępowaniu (Załącznik nr 6). Oświadczenie musi być podpisane przez osobę upoważnioną do reprezentacji Oferenta oraz opatrzone pieczętką firmowa (a w razie braku – dodatkowo czytelnym podpisem).
10. Do oferty należy załączyć oświadczenie O braku powiązań z Zamawiającym (Załącznik nr 7). Oświadczenie musi być podpisane przez osobę upoważnioną do reprezentacji Oferenta oraz opatrzone pieczętką firmowa (a w razie braku – dodatkowo czytelnym podpisem).
11. Koszty przygotowania oferty obciążają Oferenta.

VII. Zastrzeżenia:

1. Zamawiający udzieli zamówienia Oferentowi, którego oferta odpowiada wszystkim wymaganiom przedstawionym w zapytaniu ofertowym i przedstawi najkorzystniejszą ofertę w oparciu o kryteria wyboru określone w zapytaniu ofertowym.
2. W uzasadnionych przypadkach Zamawiający dopuszcza unieważnienie postępowania.
3. Zamawiający zastrzega możliwość zakończenia postępowania bez dokonania wyboru Wykonawcy bez podania przyczyn.
4. W toku badania i oceny ofert Zamawiający może żądać od Oferentów wyjaśnień dotyczących treści złożonych ofert.
5. Oferent może przed upływem terminu składania ofert zmienić lub wycofać swoją ofertę.
6. Oferty złożone po terminie nie będą rozpatrywane.
7. Oferty niespełniające wymogów formalnych, zawartych w niniejszym Zapytaniu ofertowym, nie będą rozpatrywane.
8. Zamawiający nie dopuszcza możliwości składania ofert częściowych.
9. Zamawiający nie dopuszcza możliwości składania ofert wariantowych.
10. Zamawiający nie dopuszcza możliwości powierzenia całości zamówienia podwykonawcom.
11. Powierzenie wykonania części zamówienia podwykonawcom wymaga każdorazowej zgody Zamawiającego.
12. Zamawiający może zażądać od Oferenta przedstawienia odpowiednich dokumentów potwierdzających złożone w toku postępowania oświadczenia.

Załączniki:

Załącznik 1 – Specyfikacja przedmiotu zamówienia

Załącznik 2 – Harmonogram realizacji zamówienia

Załącznik 3 – Formularz ofertowy

Załącznik 4 – Oświadczenie niezbędna wiedza

Załącznik 5 – Oświadczenie sytuacja finansowa

Załącznik 6 – Oświadczenie spełnienia warunków udziału w postępowaniu

Załącznik 7 – Oświadczenie o braku powiązań

Zatwierdzam

Leszek Stypułkowski

Prezes Zarządu

Specyfikacja przedmiotu zamówienia

Przedmiotem zamówienia jest wykonanie oraz przeniesienie praw autorskich do 6 modułów programistycznych systemu teleinformatycznego stanowiącego szkielet infrastruktury badawczej, która umożliwi efektywne prowadzenie prac badawczo rozwojowych z zakresu zastosowania uczenia maszynowego i głębokiego uczenia (sieci neuronowe) do ekstrakcji wiedzy z dużych zbiorów danych, wyznaczania Potencjału Zasobu Wiedzy oraz tworzenia profili semantycznych dla zasobów wiedzy.

W zakres zamówienia wchodzi zaprojektowanie, wytworzenie, rozruch, wdrożenie, świadczenie asysty technicznej, przeniesienie praw autorskich oraz uruchomienie specjalistycznego oprogramowania informatycznego stanowiącego szkielet infrastruktury badawczej zgodnie z poniższą specyfikacją:

1. Moduły

1.1. Specyfikacja ogólna modułów

1.1.1. Moduł nr 1. „Importer danych” - Zadaniem modułu jest zasilanie hurtowni analitycznej o nowe dane pozyskane z zewnętrznych źródeł (DarkData). Zakładamy, że moduł ten będzie w ramach prac badawczych zasilany o nowe „wtyczki” pozyskujące wartościowe dane, wytypowane w oparciu o wyniki analiz przeprowadzone na gotowym systemie.

1.1.1.1. Importer danych ciągłych (DataImporter1) - Zadaniem komponentu importu danych ciągłych jest zasilanie systemu o dane pozyskane ze źródeł „ciągłych”, czyli dostarczających dane w sposób stały i przyrostowy (dane typu DarkData). Będą to takie źródła danych, które są stale zasilane i które warto pozyskać. Zakładamy, że dla każdego ze źródeł danych będzie możliwe przygotowanie oddzielnej „wtyczki”, pozyskującej przyrostowo dane w określonych przedziałach czasu, z określonego źródła. Zakładamy, że komponent w ramach prac badawczych będzie rozbudowywany o nowe „wtyczki” w razie potrzeb i pozyskiwania nowych źródeł danych.

1.1.1.2. Importer wsadowy (DataImporter2) - Zadaniem komponentu importu wsadowego, jest zasilanie systemu o dane pozyskane ze źródeł danych serwujących dane nie w sposób ciągły (dane typu DarkData). Często będą to dane „jednorazowe”. Zakładamy, że dla każdego źródła danych, będzie możliwe przygotowanie oddzielnej „wtyczki”, która będzie potrafiła przyjąć dane do systemu. Dodatkowo, komponent będzie realizował także możliwość wprowadzania danych w trybie „ręcznym”. Dane będą importowane do systemu na żądanie. Zakładamy, że komponent w ramach prac badawczych będzie rozbudowywany o nowe „wtyczki” w razie potrzeb i pozyskiwania nowych, wartościowych źródeł danych.

- 1.1.1.3.** Importer danych – konwerter danych (DataConverter) - Zadaniem komponentu konwertera danych jest konwersja danych z formatu źródłowego (DarkData) danego źródła danych, na opracowany wcześniej, zunifikowany format danych. Przygotowane przez kolejne „wtyczki” komponentu dane, będą gotowe do oczyszczenia i dalszego przetwarzania przez kolejne moduły systemu. Ze względu na efektywność, komponent ten może być zintegrowany z konkretnymi wtyczkami dla poszczególnych źródeł danych. Zakładamy, że komponent w ramach prac badawczych będzie rozbudowywany o nowe „wtyczki” w razie potrzeb, ewentualnych zmian w zunifikowanym modelu i pozyskiwania nowych źródeł danych.
- 1.1.1.4.** Importer danych – czyszczenie (DataCleaner) - Zadaniem komponentu jest wstępne oczyszczanie zunifikowanych danych. Zakładamy, że dane wychodzące z komponentów importera i konwertera będą wymagały wstępnego oczyszczenia i konwersji bezstratnej (czyli takiej, gdzie konwersja jest jednoznaczna). Komponent będzie uruchamiał na zunifikowanych danych kolejne „wtyczki”, których zadaniem będzie oczyszczanie danych.
- 1.1.1.5.** Importer danych – filtrowanie (DataFilter) - Zadaniem komponentu jest odpowiednie oznaczenie danych zakwalifikowanych jako „błąd oczywisty” tak, aby dalsze moduły systemu pominęły je.
- 1.1.2. Moduł nr 2 „Ekstrakcja danych”** - Zadania modułu Ekstrakcji danych to transformacja, czyszczenie i ekstrakcja danych ze zunifikowanego formatu źródłowego na dane w postaci obiektów i metadanych obiektów, które następnie zasila główne repozytorium danych systemu, a więc będą stanowiły dane źródłowe dla systemów prezentacyjnych i laboratoryjnych (analitycznych). Wszystkie dane powstałe w wyniku działania modułu docelowo trafią do głównego repozytorium danych systemu będącego głównym repozytorium danych do dalszych prac analityczno-badawczych. Każdy kwant informacji powinien mieć oznaczenie importu i źródła. Działanie poszczególnych komponentów będzie polegało na ekstrakcji odpowiednich danych, a następnie ich dokładnym czyszczeniu i opisaniu wyznaczonym wskaźnikiem poziomu ufności. Całość będzie opierała się na systemie „wtyczek”, czyli mniejszych komponentów, które można przypinać, modyfikować, dodawać i usuwać z poszczególnych komponentów, a których zadaniem będzie wyczyszczenie danych.
- 1.1.2.1.** Ekstraktor obiektów (ObjectExtractor) - Głównym zadaniem komponentu jest ekstrakcja obiektów ze zunifikowanego źródła danych, ich analiza i czyszczenie w oparciu o kolejne, testowane „wtyczki”.
- 1.1.2.2.** Ekstraktor metadanych (MetadataExtractor) - Głównym zadaniem komponentu jest ekstrakcja metadanych dla obiektów i powiązań z zunifikowanego źródła danych w oparciu o kolejne, testowane „wtyczki”.

1.1.2.3. Ekstraktor powiązań (AsociacionExtractor) - Głównym zadaniem komponentu jest ekstrakcja powiązań pomiędzy obiektami z zunifikowanego źródła danych w oparciu o kolejne, testowane „wtyczki”. Główne zadania to ekstrakcja powiązań, ich oczyszczenie, mapowanie ich cech na słowniki oraz wyznaczanie dla nich poziomów ufności.

1.1.3. Moduł nr 3 „DataIndexer” - Zadania modułu inteligentnego indeksera, to identyfikacja, nazwanie i numerowanie danych jednakowych. Do głównych zadań modułu inteligentnego indeksera należy właśnie wykrywanie potencjalnych duplikatów (zarówno obiektów, jak i ich powiązań) i ich odpowiednie opisanie i oznaczenie. W ten sposób zbudowana zostanie baza unikalnych obiektów i ich powiązań.

1.1.3.1. Baza parametrów deduplikacji (DeduplicationParameterHandler) - Komponent będzie odpowiedzialny za możliwość tworzenia własnych grup ważonych parametrów pozwalających na określenie duplikatu. Dana grupa parametrów będzie przypisana dla danej klasy obiektu, jedna klasa obiektu może posiadać więcej niż jedną grupę parametrów – jest to zależne od liczby wykorzystywanych algorytmów. Zakładamy, że parametry powstaną w wyniku pracy analitycznej. Komponent będzie stanowił bazę wiedzy oraz podstawę pracy algorytmów deduplikacji danych. Odpowiednia definicja parametrów umożliwi wykorzystanie zaawansowanych algorytmów deduplikacji danych, takich jak sieci neuronowe, drzewa decyzyjne, itp.

1.1.3.2. Konwerter parametrów (DeduplicationParameterConverter) - Komponent będzie odpowiedzialny za konwersję danych pochodzących z głównej bazy danych zgodnie z przypisanymi algorytmami. Oprócz konwersji zgodnie z parametrami komponent jest również odpowiedzialny za normalizację danych wejściowych do algorytmów (np. normalizacja za pomocą standardowego odchylenia średniokwadratowego).

1.1.3.3. Deduplikator danych (DeduplicationAlgorithms) - Komponent będzie odpowiedzialny za przetwarzanie danych w celu wykrycia duplikatów. Będzie on bazował na przygotowanych testowanych „wtyczkach”, między innymi na algorytmach prostych opartych na porównaniu pól, ewentualnie z uprzednim oczyszczeniem, po algorytmie uczenia maszynowego (sieci neuronowe).

1.1.3.4. Numerator danych (GICIDHandler) - Komponent ten będzie zarządzał numeracją obiektów. Zakładamy, że każdy nowy, zdeduplikowany obiekt zostanie oznaczony nowym, unikalnym identyfikatorem (w celu jego jednoznacznej identyfikacji).

1.1.3.5. Rozstrzyganie sporów (DataStewardComponent) - Ze względu na wysoką staranność o jakość wyników przewidywana jest w wyjątkowych

przypadkach pomoc człowieka rozumianego jako DataStewarda, który ostatecznie rozstrzygnie o wyniku.

1.1.4. Moduł nr 4 „DataProfiler” - Zadaniem modułu Inteligentnego Profilera jest przygotowanie bazy profili obiektów (najlepszych wersji). Przyjmujemy, że istnieje możliwość aktualizacji najlepszych wersji obiektów wraz z rosnącą liczbą ich wersji pochodzących z różnych źródeł zewnętrznych. Moduł Inteligentnego Profilera zostanie zrealizowany przez kilka głównych komponentów, opartych o testowane w procesie badawczym „wtyczki” wyłuskujące wartościowe dane i wyznaczające wartości parametrów.

1.1.4.1. Baza parametrów profili (ProfileParameters) - Komponent ten będzie pozwalał na budowanie charakterystycznych grup parametrów ilościowych oraz jakościowych dla danej klasy obiektów. Parametry w głównej mierze będą wydobywane dzięki pracy analitycznej oraz z biegiem czasu aktualizowane. Celem jest nie tylko wyłonienie najlepszej wersji danych opisujących obiekt, ale również określenie jej jakości na podstawie źródła pochodzenia oraz metadanych i ich poziomu ufności.

1.1.4.2. Konwerter parametrów (ProfileParameterConverter) - Głównym zadaniem będzie konwertowanie obiektu zgodnie z przypisanymi do jego typu parametrami. Będzie stanowił więc podstawę do pracy „wtyczek” algorytmów Profilera.

1.1.4.3. Profiler danych (ProfilerAlgorithms) - Komponent będzie odpowiedzialny za odnajdowanie, budowanie i ocenę gotowych profili obiektów. Gotowe profile, jak i wszystkie wytworzone przez komponent dane, będą składowane w głównej bazie profili (MasterDB).

1.1.5. Moduł nr 5 „Portal prezentacyjny” - Moduł ten będzie odpowiedzialny za prezentację danych pochodzących ze wszystkich podsystemów laboratorium. Moduł zostanie zrealizowany jako portal internetowy, serwujący i pozwalający na analizę zgromadzonych danych. Dzięki niemu, możliwe będzie zarządzanie procesem badawczym i porównywanie wyników działań poszczególnych algorytmów. Komponent będzie prezentował informacje o wyekstrahowanych przez system danych o obiektach, ich powiązaniach i profilach. Portal będzie opierał się o dane zgromadzone w bazie MasterDB, a więc będzie prezentował informacje zakwalifikowane jako najlepsze i najpewniejsze. Będzie umożliwiał ich przeszukiwanie, prezentowanie i porównywanie zgromadzonych informacji, wraz z historią ich „utworzenia”.

1.1.5.1. Baza obiektów, powiązań i profili - Komponent będzie prezentował informacje o wyekstrahowanych przez system danych o obiektach, ich powiązaniach i profilach. Portal będzie opierał się o dane zgromadzone w bazie MasterDB, a więc będzie prezentował informacje zakwalifikowane jako najlepsze i najpewniejsze. Będzie umożliwiał ich przeszukiwanie,

ZAŁĄCZNIK NR 1 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

prezentowanie i porównywanie zgromadzonych informacji, wraz z historią ich „utworzenia”.

1.1.5.2. Baza analityczna - Komponent będzie pozwalał na analizę danych zgromadzonych w systemie.

1.1.5.3. Użytkownicy i portal typu SocialMedia - Komponent będzie umożliwiał tworzenie i zarządzanie kontami użytkowników w module prezentacyjnym. Użytkownicy będą mogli zgłaszać informacje o nieprawidłowych danych, bądź propozycje ulepszeń i modyfikacji kolejnych iteracji badań. Zgłoszenia takie trafią do operatora, który będzie mógł (bądź nie) wprowadzić stosowne poprawki. Każdy użytkownik będzie przypisany do odpowiednich ról, np. aktywni użytkownicy będą mogli dostawać od systemu dane do potwierdzenia.

1.1.6. Moduł nr 6 „DataAnalyser” - Moduł będzie stanowił główną bazę analityczną, pozwalającą na dogłębną analizę pozyskanych, zgromadzonych i przetworzonych przez system danych. Zakładamy, że wszystkie dane zgromadzone w systemie, w tym dane zgromadzone lub wytworzone w kolejnych modułach trafią, bądź zostaną skonsolidowane w dużej bazie badawczej pozwalającej na skuteczną analizę dużych baz danych. Wykorzystane mogą tu zostać takie narzędzia jak Apache Hadoop, Apache Spark, MS HDInsight i tym podobne. Umożliwi to dowolne i dogłębne prowadzenie analiz i weryfikacji jakości przygotowanych danych. Wynikiem prac analitycznych będą także nowe i poprawione „wtyczki” algorytmów do przebadania, między innymi wtyczki czyszczące dane, algorytmy deduplikacyjne, profilujące itp. Dane zawarte w bazie będą mogły posłużyć np. do uczenia sieci neuronowych czy modeli uczenia maszynowego.

1.2. W ramach powyższych modułów powinny zostać zrealizowane również podmoduły wspólne

1.2.1. „CAS” – Moduł ten będzie odpowiedzialny za zarządzanie kontami oraz uprawnieniami użytkowników,

1.2.2. „System powiadomień” – Moduł ten będzie odpowiedzialny za informowanie użytkowników o podejmowanych działaniach w systemie oraz ich wynikach, w szczególności w przypadku niepowodzenia,

1.2.3. „Zarządzanie procesem” – Moduł ten będzie odpowiedzialny za udostępnienie funkcjonalności umożliwiających konfigurację procesów realizowanych w Centrum Badawczo – Rozwojowym na przestrzeni modułów: Importera Danych, Ekstrakcji Danych, DataIndexera, DataProfiler, Portalu Prezentacyjnego oraz Analizy Danych.

1.2.4. Budowa interfejsów wymiany danych między modułami.

1.2.5. Budowa i obsługa procesów utajniania danych.

ZAŁĄCZNIK NR 1 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

1.3. Algorytmy – w ramach przedmiotu zamówienia należy przygotować zestaw podstawowych algorytmów umożliwiających przetestowanie poprawności działania laboratorium i podsystemów wchodzących w jego skład.

1.3.1. Algorytmy ekstrakcji obiektów poprzez dekompozycję przypisów bibliograficznych będących ciągami znaków.

1.3.2. Algorytmy asocjacji obiektów.

Wymagania dodatkowe w ramach przedmiotu zamówienia:

Przedmiot zamówienia obejmuje następujące elementy:

2. Przeprowadzenie analizy wymagań (weryfikacja i uszczegółowienie wymagań funkcjonalnych i нефункциональных oraz opracowanie koncepcji modułów).

3. Stworzenie projektu technicznego:

3.1. Rozwiązania infrastruktury teleinformatycznej (sprzętowej, sieciowej i programowej) modułów dla wymaganych środowisk pracy - uwzględniający dedykowaną przez Zamawiającego infrastrukturę na potrzeby tych modułów.

3.2. Referencyjnej bazy danych (zawierający m.in. opis modułów funkcjonalnych, model danych, opis procesów przetwarzania danych, opis interfejsów, opis bezpieczeństwa systemu itp.).

3.3. Środowiska informatycznego umożliwiającego uruchamianie, zarządzanie i konfigurację algorytmów w wydzielonych obszarach, będących docelowym zadaniem badań w laboratorium.

4. Budowa, rozruch oraz wdrożenie modułów, implementacja algorytmów

4.1. Budowa i konfiguracja infrastruktury teleinformatycznej modułów.

4.1.1. Wdrożenie infrastruktury teleinformatycznej (przeznaczonej w tym celu przez Zamawiającego).

4.1.2. Instalacja, konfiguracja i wdrożenie oprogramowania.

4.1.3. Konfiguracja środowisk.

5. Dodatkowe założenia dotyczące realizacji przedmiotu zamówienia

5.1. Infrastruktura teleinformatyczna i moduły powinny:

5.1.1. Zostać zoptymalizowane do przetwarzania dużych zbiorów danych.

5.1.2. Umożliwiać zrównoleglenie wykonywania algorytmów.

5.1.3. Umożliwiać sterowanie przebiegiem całościowego procesu przetwarzania danych za pomocą panelu zarządczego z możliwością wyboru źródeł danych oraz zakresu danych.

5.1.4. Umożliwiać łatwą integrację z systemami zewnętrznymi.

5.1.5. Umożliwiać prowadzenie jednocześnie kilku niezależnych badań na wydzielonych obszarach badawczych.

5.2. Realizacja podstawowych algorytmów dekompozycji, ekstrakcji, asocjacji powinna odbywać się w jak największym stopniu poprzez mechanizmy uczenia maszynowego, głębokiego uczenia z wykorzystaniem sieci neuronowych.

5.3. Realizacja podstawowych algorytmów deduplikacji oraz indeksacji danych powinna odbywać się w jak największym stopniu poprzez mechanizmy uczenia maszynowego, głębokiego uczenia z wykorzystaniem sieci neuronowych.

5.4. Realizacja podstawowych algorytmów profilowania danych powinna odbyć się w jak największym stopniu poprzez mechanizmy uczenia maszynowego, głębokiego uczenia z wykorzystaniem sieci neuronowych w oparciu o wskazane narzędzia analityczne.

6. Testowanie i odbiór modułów:

6.1. Import i ekstrakcja obiektów biznesowych ze wskazanego źródła danych, w tym czyszczenie, konwersja, filtrowanie oraz walidacja danych, w szczególności skompletowanie danych danego obiektu z różnych zdarzeń, które go dotyczą. Liczba rekordów opisujących zdarzenia - ok. 6.000.000, liczba ekstrahowanych typów obiektów - minimum 3, liczba obiektów, które powinny zostać wyekstrahowane - około 1.000.000 obiektów referencyjnych.

6.1.1. Konfiguracja procesu importu (w trybie ciągłym i wsadowym) i ekstrakcji.

6.1.2. Import zadanego źródła danych o zdarzeniach.

6.1.3. Czyszczenie, konwersja, filtrowanie i walidacja danych wg skonfigurowanego wcześniej procesu.

6.1.4. Przeprowadzenie procesu ekstrakcji obiektów wraz z ich metadanymi i powiązaniem z zaimportowanym źródłem wg skonfigurowanego wcześniej procesu.

6.1.5. Testy ilościowe i jakościowe wyekstrahowanych obiektów wraz z ich metadanymi i powiązaniem wg zadanych kryteriów. Ocena poprawności procesu na podstawie bazy wzorcowej.

6.1.6. Testy odporności na awarie (utrata lub desynchronizacja przetwarzanych danych).

6.2. Import, ekstrakcja oraz indeksowanie obiektów biznesowych ze wskazanego źródła danych, w tym czyszczenie, konwersja, filtrowanie, walidacja, deduplikacja i numerowanie danych, pochodzących z kilku niezależnych źródeł danych w celu sporządzenia referencyjnej bazy danych. Liczba niezależnych źródeł danych - minimum 3, łączna liczba rekordów we wszystkich źródłach danych - minimum 10.000.000 rekordów, liczba typów obiektów - minimum 3, liczba obiektów, które powinny zostać wyekstrahowane - około 5.000.000 obiektów, liczba obiektów referencyjnych - minimum 2.000.000.

- 6.2.1. Konfiguracja procesu importu (w trybie ciągłym i wsadowym), ekstrakcji, indeksowania, deduplikacji i profilowania danych.
- 6.2.2. Import zadanych źródeł danych.
- 6.2.3. Czyszczenie, konwersja, filtrowanie i walidacja danych wraz z określeniem poziomu ufności dla każdego źródła wg skonfigurowanego wcześniej procesu.
- 6.2.4. Przeprowadzenie procesu ekstrakcji obiektów wraz z ich metadanymi i powiązaniem z zaimportowanymi źródłami wg skonfigurowanego wcześniej procesu.
- 6.2.5. Przeprowadzenie procesu indeksowania obiektów, w tym deduplikacja, nazwanie i indeksacja danych oraz wskazanie rekordów do obsługi ręcznej (DataSteward) wg skonfigurowanych wcześniej parametrów procesu.
- 6.2.6. Przeprowadzenie procesu tworzenia referencyjnej bazy danych obiektów wg skonfigurowanych wcześniej parametrów procesu.
- 6.2.7. Testy ilościowe i jakościowe obiektów referencyjnych wraz z ich metadanymi i powiązaniem wg zadanych kryteriów. Ocena poprawności procesu na podstawie bazy wzorcowej.
- 6.2.8. Testy odporności na awarie (utrata lub desynchronizacja przetwarzanych danych).
- 6.3. Testy portalu prezentacyjnego w zakresie prezentacji wyników poszczególnych etapów importu, ekstrakcji, indeksowania i profilowania obiektów pochodzących z testowych zbiorów danych. Liczba niezależnych źródeł danych - minimum 3, łączna liczba rekordów we wszystkich źródłach danych - minimum 10.000.000 rekordów, liczba typów obiektów - minimum 3, liczba obiektów, które powinny zostać wyekstrahowane - około 5.000.000 obiektów, liczba obiektów referencyjnych - minimum 2.000.000, liczba użytkowników jednoczesnych - minimum 100.
 - 6.3.1. Przeprowadzenie procesu tworzenia oraz zarządzania kontami użytkowników na zadanym źródle danych.
 - 6.3.2. Prezentacja bazy obiektów, profili i powiązań wraz z historią ich pozyskania w celu analizy przeprowadzanych testowych procesów.
 - 6.3.3. Graficzna wizualizacja danych umożliwiająca analizę procesowanych obiektów (np. grafy, wykresy) w tym testy integracji modułu portalu prezentacyjnego z platformą Apache Spark.
 - 6.3.4. Testy integracji modułu portalu prezentacyjnego z modułem inteligentnego indeksera w celu analizy funkcjonalności DataStewarda.
 - 6.3.5. Testy integracji modułu portalu prezentacyjnego z modułem analizy danych.
 - 6.3.6. Ocena interfejsów użytkownika w pod kątem intuicyjności, przejrzystości oraz czasu odpowiedzi systemu na przesyłane przez użytkowników żądania (w tym weryfikacja średniego, minimalnego, maksymalnego czasu odpowiedzi).
 - 6.3.7. Testy ilościowe i jakościowe zebranych informacji o zachowaniu użytkowników.

ZAŁĄCZNIK NR 1 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

6.3.8. Testy odporności na awarie (utrata lub desynchronizacja przetwarzanych danych).

6.4. Testy modułu analizy danych. Liczba niezależnych źródeł danych – minimum 3, łączna liczba rekordów we wszystkich źródłach danych – minimum 30.000.000, liczba typów obiektów – minimum 3.

6.4.1. Testy integracji modułu z narzędziami analitycznymi (Apache Spark/Apache Hadoop/ MS HDInsight).

6.4.2. Testy dostarczonych bibliotek zawierających algorytmy uczenia maszynowego na zadanych zbiorach testowych oraz wzorcowych, w tym testy funkcjonalności opisanych w szczegółowej specyfikacji projektu.

6.4.3. Ocena mechanizmów przetwarzania dużych zbiorów danych pod kątem optymalizacji czasu działania (wyniki poniżej określonego progu czasu działania algorytmów).

6.4.4. Testy odporności na awarie (utrata lub desynchronizacja przetwarzanych danych).

6.4.5. Wymagany zestaw testów jednostkowych pokrywających algorytmy oraz udostępniane funkcjonalności.

7. Technologie

Moduły programistyczne powinny zostać wykonane w technologii Java SE 9, Java EE 8 lub równoważnej, która jest podstawową technologią systemów informatycznych Zamawiającego. Dodatkowe preferowane technologie to:

- Relacyjne i nierelacyjne bazy danych
- Scala, Python, R
- Spring 5 (wraz z podprojektami takimi jak Spring MVC, Spring Web-Flow, Spring Data, Spring Security, itp..), Hibernate, JPA2
- Apache Spark, deeplearning4j, WEKA, H2O.io, jblas
- Apache Hadoop (wraz z elementami ekosystemu aplikacji)
- Automatyzacja budowy oprogramowania: Maven, Gradle
- System kontroli wersji GIT
- CSS, JavaScript, HTML5, Angular4
- JSF 2.x
- JUnit
- Solr, Apache Lucene

Leszek Stypułkowski

Przewodniczący Zarządu

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

str. 9

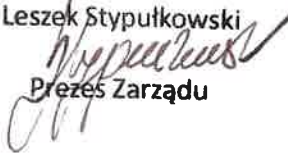
ZAŁĄCZNIK NR 2 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

HARMONOGRAM REALIZACJI ZAMÓWIENIA

L.p.	Moduł	Komponent	Data zakończenia testów w środowisku testowym	Data zakończenia prac i zdania wykonanej części
1	1.Importer danych (Importer)	1.1. Importer danych ciągłych (DataImporter1)	30 marzec 2018	15 kwiecień 2018
		1.2. Importer wsadowy (DataImporter2)	30 marzec 2018	15 kwiecień 2018
		1.3. Importer danych – czyszczenie (DataCleaner)	30 marzec 2018	15 kwiecień 2018
		1.4. Importer danych – filtrowanie (DataFilter)	30 marzec 2018	15 kwiecień 2018
		1.5. Importer danych – konwerter danych (DataConverter)	31 marzec 2018	15 kwiecień 2018
2	2.Ekstrakcja danych (DataExtractor)	2.1. Ekstraktor obiektów (ObjectExtractor)	30 kwiecień 2018	15 maj 2018
		2.2. Ekstraktor metadanych (MetadataExtractor)	30 kwiecień 2018	15 maj 2018
		2.3. Ekstraktor powiązań (AssociationExtractor)	30 kwiecień 2018	15 maj 2018
3	3. Inteligentny Indexer (DataIndexer)	3.1. Baza parametrów deduplikacji (DeduplicationParameterHandler)	30 czerwiec 2018	15 lipiec 2018
		3.2. Konwerter parametrów (DeduplicationParameterConverter)	30 czerwiec 2018	15 lipiec 2018
		3.3. Deduplikator danych (DeduplicationAlgorithms)	30 czerwiec 2018	15 lipiec 2018
		3.4. Numerator danych (GICIDHandler)	30 czerwiec 2018	15 lipiec 2018
		3.5. Rozstrzyganie sporów (DataStewardComponent)	30 czerwiec 2018	15 lipiec 2018
4	4. Inteligentny Profiler (DataProfiler)	4.1. Baza parametrów profili (ProfileParameters)	30 sierpień 2018	15 wrzesień 2018
		4.2. Konwerter parametrów (ProfileParameterConverter)	30 sierpień 2018	15 wrzesień 2018
		4.3. Profiler danych (ProfilerAlgorithms)	30 sierpień 2018	15 wrzesień 2018
5	5. Portal prezentacyjny	5.1. Baza obiektów, powiązań i profili	30 październik 2018	15 listopad 2018

ZAŁĄCZNIK NR 2 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

		5.2. Baza analityczna	30 październik 2018	15 listopad 2018
		5.3. Użytkownicy i portal typu SocialMedia	30 październik 2018	15 listopad 2018
6	6. Analizy danych (DataAnalyser		30 grudzień 2018	30 grudzień 2018

Leszek Stypułkowski

Prezes Zarządu

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

ZAŁĄCZNIK NR 3 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

FORMULARZ OFERTOWY

1. Wykonawca (należy wypełnić)

Nazwa firmy	
Adres	
NIP	
Regon	
Pesel <i>(dane wymagane w przypadku os. fizycznej)</i>	
Adres zamieszkania <i>(dane wymagane w przypadku os. fizycznej)</i>	
Nr telefonu	
Adres e-mail	
Osoba do kontaktu z Zamawiającym	

2. Zamawiający :

INDEX COPERNICUS SP. Z O. O.

3. Przedmiot zamówienia:

Przedmiotem zamówienia jest zakup 6 modułów programistycznych stanowiących bazę Centrum Badawczo-Rozwojowego dla prowadzenia prac badawczo-rozwojowych w związku z planowaną realizacją Projektu pn. "Utworzenie Centrum Badawczo-Rozwojowego w przedsiębiorstwie Index Copernicus Sp. z o.o. w celu prowadzenia prac badawczo-rozwojowych nad innowacyjnymi technologiami informatycznymi" w ramach Programu Operacyjnego Inteligentny Rozwój, 2014-2020 działanie 2.1 „Wsparcie inwestycji w infrastrukturę przedsiębiorstw”.

4. Cena oferty w złotych polskich (lub w innej walucie)

Cena oferty netto (liczbą)

Cena oferty netto (słownie)

Cena oferty brutto (liczbą)

Cena oferty brutto (słownie)

Należny podatek VAT



ZAŁĄCZNIK NR 3 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

5. Określenie spełnienia kryteriów funkcjonalnych oceny

Oferent powinien określić przy każdym z kryteriów funkcjonalnych wartość parametru, którą będzie w stanie zapewnić dostarczony przez niego system informatyczny. Deklaracja wartości kryterium następuje poprzez wpisanie „X” w tyło jednym wierszu danego kryterium. Niezaznaczenie żadnej opcji lub zaznaczenie więcej niż jednego znaku „X” dla danego kryterium będzie uznane za niezgodność oferty z warunkami zamówienia i będzie skutkowało odrzuceniem oferty. Oferty, które w kryteriach funkcjonalnych otrzymają w sumie mniej niż 20 punktów zostaną odrzucone jako nie spełniające warunków zamówienia. Punktacja z poszczególnych kryteriów funkcjonalnych będzie sumowa, a końcowa liczba punktów obliczona zgodnie ze wzorem przedstawionym w zapytaniu ofertowym nr 1/2017.

Numer kryterium	Kryterium	Jednostka	Wartość	Kolumna wyboru deklarowanej wartości
F1	Ekstrakcja obiektów typu artykuł ze zbiorów liczących 6 000 000 rekordów pozycji bibliografii załącznikowej, w których artykuły naukowe występują w ciągu znaków jako tekst	liczba unikalnych artykułów referencyjnych z nadanym unikalnym nr GICID (z poprawnością powyżej 98%)	> 300 000	
			200 000 - 300 000	
			100 000 - 200 000	
			< 100 000	
F2	Ekstrakcja obiektów typu autor ze zbiorów liczących 6 000 000 rekordów pozycji bibliografii załącznikowej, w których autorzy występują w ciągu znaków jako tekst	liczba referencji z wykrytymi autorami ze wskazanym przynajmniej jednym imieniem i nazwiskiem (o poprawności powyżej 99%)	> 5 000 000	
			4 500 000 - 5 000 000	
			3 000 000 - 4 500 000	
			< 3 000 000	
F3	Ekstrakcja rodzaju nośnika (książka, czasopismo, monografia naukowa, konferencja, strona www itp..) ze zbiorów liczących 6 000 000 rekordów pozycji bibliografii załącznikowej, w której nośnik występuje jako ciąg znaków	liczba referencji z wykrytym nośnikiem bibliograficznym (o poprawności powyżej 99%)	> 3 000 000	
			2 500 000 - 3 000 000	
			2 000 000 - 2 500 000	
			1 500 000 - 2 000 000	

ZAŁĄCZNIK NR 3 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

			< 1 500 000	
F4	Ekstrakcja metadanych z danych źródłowych pochodzących z importu ciągłego oraz wsadowego na zadanej próbkce 6 000 000 rekordów	liczba poprawie wskazanych typów obiektów w 90% rekordów (poprawność powyżej 95%)	>= 3	
			2	
			1	
			BRAK	
F5	Baza profili unikalnych (zdeduplikowanych) autorów z przypisanymi im powiązanymi unikalnymi (zdeduplikowanymi) obiektami typu artykuł wyekstrahowanych ze zbiorów liczących 6 000 000 rekordów pozycji bibliografii załącznikowej, w których zarówno autorzy, jak i artykuły występują w ciągu znaków jako tekst.	liczba unikalnych profili naukowców z nadanym unikalnym numerem GICID, zawierających przynajmniej 10 powiązanych publikacji	> 100 000	
			80 000 - 100 000	
			60 000 - 80 000	
			20 000 - 60 000	
			<20 000	
F6	Baza unikalnych obiektów typu monografia z przypisanymi powiązaniami obiektów typu autor oraz redaktor ze zbioru liczącego 650 000 rekordów zdarzeń ewaluacyjnych typu artykuł, monografia i rozdział w monografii	liczba unikalnych monografii z nadanym unikalnym numerem GICID	> 100 000	
			80 000 - 100 000	
			60 000 - 80 000	
			40 000- 60 000	
			<40 000	
F7	Efektywność klasyfikacji wykonywana na oddzielnych zbiorach testowych, które nie będą w sobie zawierały próbek ze zbiorów treningowych (jakość klasyfikacji dla	poprawne odpowiedzi	> 98%	
			97% - 98%	

ZAŁĄCZNIK NR 3 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

	1 000 000 próbek, wartość uśredniona dla 10 testów)		96% - 97%	
			95% - 96%	
			90% - 95%	
			<90%	
F8	Efektywność klasyfikacji wykonywana na oddzielnych zbiorach testowych, które nie będą w sobie zawierały próbek ze zbiorów treningowych (dopasowanie się sieci do nieznannej funkcji za pomocą błędu MSE (błąd średnio-kwadratowy) dla 1 000 000 próbek, wartość uśredniona dla 10 testów)	wartość z ostatnich iteracji	< 0.05	
			0.05 - 0.1	
			0.1 - 0.2	
			0.2 - 0.3	
			0.3 - 0.4	
			> 0.4	
F9	Liczba zagranicznych repozytoriów danych o osiągnięciach naukowych będących zdarzeniami ewaluacyjnymi, zawierających przynajmniej 1.000.000 zwalidowanych rekordów z lat 2010-2015, z którymi system synchronizuje dane w procesie ciągłym	liczba repozytoriów międzynarodowych	4	
			3	
			2	
			<2	
F10	Liczba krajowych repozytoriów danych o osiągnięciach naukowych będących zdarzeniami ewaluacyjnymi, zawierających przynajmniej 1.000.000 zwalidowanych rekordów z lat 2010-2015, z którymi system	liczba repozytoriów krajowych	3	
			2	
			1	

ZAŁĄCZNIK NR 3 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

	synchronizuje dane w procesie ciągłym		BRAK	
F11	Gwarancja - Maksymalny czas naprawy błędu krytycznego	dni robocze	do 1 dnia	
			do 2 dni	
			do 3 dni	
			> 3 dni	
F12	Gwarancja - Maksymalny czas naprawy błędu niekrytycznego (pozostałe)	dni robocze	do 2 dnia	
			do 3 dni	
			do 5 dni	
			> 5 dni	

6. Termin płatności faktur 30 dni.

7. Składający ofertę oświadcza, że:

- po zapoznaniu się z warunkami zamówienia przedstawionymi w zapytaniem ofertowym w pełni je akceptuje i nie wnosi do nich zastrzeżeń,
- akceptuje warunki, termin realizacji zamówienia, termin płatności faktur;
- zobowiązuje się do wykonania usługi, zgodnie ze wszystkimi wymaganiami Zamawiającego dotyczącymi realizacji przedmiotu zamówienia,
- akceptuje 30 dniowy termin związania ofertą liczony od daty ostatecznego składania ofert,
- wyraża zgodę na przetwarzanie danych osobowych w zakresie niezbędnym do przeprowadzenia postępowania o zamówienie publiczne zgodnie z ustawą z dnia 29.08.1997 r. o ochronie danych osobowych (Dz. U. 2002, nr 101, poz. 926 z późn. zmianami),

Data.....

.....
Podpis i pieczęć Oferenta

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

Leszek Stypułkowski
[Podpis]
Prezes Zarządu

ZAŁĄCZNIK NR 4 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

Oświadczenie o posiadaniu niezbędnej wiedzy i doświadczenia oraz dysponowaniu potencjałem technicznym do wykonania zamówienia

Składając ofertę w zapytaniu ofertowym nr 1/2017

W związku z planowaną realizacją Projektu pn. "Utworzenie Centrum Badawczo-Rozwojowego w przedsiębiorstwie Index Copernicus Sp. z o.o. w celu prowadzenia prac badawczo-rozwojowych nad innowacyjnymi technologiami informatycznymi" w ramach Programu Operacyjnego Inteligentny Rozwój, 2014-2020 działanie 2.1 „Wsparcie inwestycji w infrastrukturę przedsiębiorstw”

Dla firmy: **Index Copernicus Sp. z o.o.**

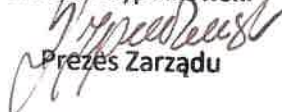
Oświadczam, że firma
posiada niezbędną wiedzę i doświadczenie oraz dysponuje potencjałem technicznym do wykonania zamówienia opisanego szczegółowo w zapytaniu ofertowym nr 1/2017

Dnia

.....

(Podpis i pieczęć oferenta)

Leszek Stypułkowski



Prezes Zarządu

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

ZAŁĄCZNIK NR 5 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

Oświadczenie sytuacja finansowa

Składając ofertę w zapytaniu ofertowym nr 1/2017

W związku z planowaną realizacją Projektu pn. "Utworzenie Centrum Badawczo-Rozwojowego w przedsiębiorstwie Index Copernicus Sp. z o.o. w celu prowadzenia prac badawczo-rozwojowych nad innowacyjnymi technologiami informatycznymi" w ramach Programu Operacyjnego Inteligentny Rozwój, 2014-2020 działanie 2.1 „Wsparcie inwestycji w infrastrukturę przedsiębiorstw”

Dla firmy: **Index Copernicus Sp. z o.o.**

Oświadczam, że firma


znajduje się w sytuacji ekonomicznej i finansowej zapewniającej wykonanie zamówienia:

- nie jest zagrożona upadłością, nie otwarto jej likwidacji ani nie ogłoszono upadłości,
- nie zalega z uiszczeniem należności wobec Skarbu Państwa.

Dnia

.....
(Podpis i pieczęć oferenta)

Leszek Stypułkowski



Prezes Zarządu

Index Copernicus Sp. z o.o.

ul. Żurawia 6/12, 00-503 Warszawa

REGON 145900211, NIP 7010326698

KRS 0000403736

ZAŁĄCZNIK NR 6 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

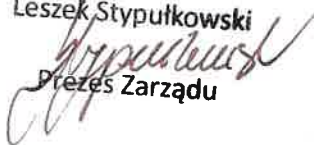
Oświadczenie o spełnieniu warunków udziału w postępowaniu

Będąc uprawnionym do reprezentacji oświadczam, że firma wykonała oraz wdrożyła przynajmniej jeden system teleinformatyczny realizujący proces gromadzenia, przetwarzania danych o osiągnięciach naukowych oraz parametryzacji osiągnięć naukowych na podstawie Rozporządzenia Ministra Nauki i Szkolnictwa Wyższego w Sprawie Kryteriów i Trybu Wyznaczania Kategorii Naukowej Jednostkom Naukowym z dnia 13 lipca 2012 rok (lub nowszym).

Dnia

.....

(Podpis i pieczęć oferenta)

Leszek Stypułkowski

Prezes Zarządu

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

ZAŁĄCZNIK NR 7 DO ZAPYTANIA OFERTOWEGO NR 1/2017 z dnia 17 listopada 2017 r. w sprawie zakupu 6 modułów programistycznych stanowiących bazę centrum badawczo – rozwojowego do prowadzenia prac badawczo – rozwojowych

Oświadczenie o braku powiązań z Zamawiającym

Będąc uprawnionym do reprezentacji oświadczam, że firma nie jest powiązana z Zamawiającym osobowo lub kapitałowo. Przez powiązania kapitałowe lub osobowe rozumie się wzajemne powiązania między beneficjentem lub osobami upoważnionymi do zaciągania zobowiązań w imieniu beneficjenta lub osobami wykonującymi w imieniu beneficjenta czynności związane z przeprowadzeniem procedury wyboru wykonawcy a wykonawcą, polegające w szczególności na:

- a) uczestniczeniu w spółce jako wspólnik spółki cywilnej lub spółki osobowej,
- b) posiadaniu co najmniej 10% udziałów lub akcji, o ile niższy próg nie wynika z przepisów prawa lub nie został określony przez IZ PO,
- c) pełnieniu funkcji członka organu nadzorczego lub zarządzającego, prokurenta, pełnomocnika,
- d) pozostawaniu w związku małżeńskim, w stosunku pokrewieństwa lub powinowactwa w linii prostej, pokrewieństwa drugiego stopnia lub powinowactwa drugiego stopnia w linii bocznej lub w stosunku przysposobienia, opieki lub kurateli.

Dnia

.....
(Podpis i pieczęć oferenta)

Leszek Stypułkowski

Prezes Zarządu

Index Copernicus Sp. z o.o.
ul. Żurawia 6/12, 00-503 Warszawa
REGON 145900211, NIP 7010326698
KRS 0000403736

