

Warszawa, 2018-12-11

CA.511.1.2018
2018-22395

**Do wszystkich Wykonawców
uczestniczących w konkursie**

Dotyczy: Konkursu dwuetapowego na Moduł Wykrywania Nadużyć w ramach pilotażu Programu GovTech Polska prowadzonego przez Zamawiającego Centrum Systemów Informatycznych Ochrony Zdrowia, **znak sprawy: CA.511.1.2018**

- I. Zamawiający Centrum Systemów Informatycznych Ochrony Zdrowia zgodnie z pkt 9.1. Regulaminu konkursu dwuetapowego na Moduł Wykrywania Nadużyć w ramach pilotażu Programu GovTech Polska (dalej „Regulaminu”) przesyła w załączeniu dokumenty dot. wymagań Zamawiającego odnoszących się do zakresu i formy opracowania oraz sposobu prezentacji Rozwiązania. Dodatkowo Zamawiający załącza plik klucze_publiczne.zip zawierające dwa pliki niezbędne do anonimizacji prac konkursowych. W celu wyjaśnienia sposobu anonimizacji Zamawiający udostępni krótki film instruktażowy dostępny pod adresem: <https://youtu.be/ZLG2OCsT7iM>
- II. Zamawiający zgodnie z pkt 9.1. Regulaminu określa zakres oferty, którą winien złożyć Uczestnik Konkursu – wzór oferty stanowi Załącznik nr 3 do niniejszego pisma.

III. Zamawiający dokonuje poniższych zmian Regulaminu:

1. Rozdział 3 pkt 2 regulaminu otrzymuje brzmienie:

Na udostępnionej próbce danych Uczestnik wykonuje analizy przy pomocy oprogramowania własnego Uczestnika, których wynikiem jest tabela indywidualnych nadużyć.

Tabela indywidualnych nadużyć zostanie przekazana w pliku CSV i składać się winna z maksymalnie 100 pojedynczych rekordów. W przypadku, gdy tabela nadużyć zawierać będzie więcej niż 100 rekordów ocenie przez Sąd Konkursowy podlegać będzie pierwsze 100 rekordów przedstawionych przez Uczestnika. Jeżeli w wygenerowanym pliku wyjściowym ten sam schemat powtarza się więcej niż 25 razy jego kolejne wystąpienia są ignorowane – w miejsce zignorowanego rekordu Sąd Konkursowy dobiera kolejny rekord z pliku wyjściowego. Przez określenie „ten sam schemat” Organizator rozumie tą samą kombinację „Kod Sprawozdany”, „Kod Poprawny”. W uzasadnionych przypadkach sąd konkursowy jako „ten sam schemat” może uznać taką samą kombinację „Kod Sprawozdany”, „Kod Poprawny”, „Uzasadnienie”, pod warunkiem, że różnice w „Uzasadnieniach” wskazują na istotną merytorycznie różnicę pomiędzy schematami o tej samej kombinacji „Kod Sprawozdany”, „Kod Poprawny”.

Pojedyncze nadużycie powinno być raportowane w formie rekordu, wg następujących pól:

Nazwa Pola	Opis	Obowiązkowe/nieobowiązkowe
Id Produktu	Id produktu odpowiadający id z próbki (ocenie podlegają jedynie rekordy z próbki o niepustej zawartości pola ID_PRODUKTU)	Pole obowiązkowe
Id Pozycji Rozliczeniowej	Id pozycji rozliczeniowej (zgodny z danymi z próbki) w przypadku, którego w opinii Uczestnika doszło do nadużycia	Pole obowiązkowe
Kod Sprawozdany	Sprawozdany (zgodny z danymi z próbki) kod produktu jednostkowego	Pole obowiązkowe
Kod Poprawny	Kod produktu jednostkowego, który wg Uczestnika powinien być sprawozdany w przypadku tego produktu	Pole obowiązkowe
Uzasadnienie	Uzasadnienie	Pole opcjonalne w I etapie / Pole obowiązkowe w II etapie

2. Rozdział 10 Punkty 1-2 Regulaminu otrzymują nowe brzmienie:

- 10.1. Rozwiązania należy złożyć od 7 grudnia 2018 do dnia 7 stycznia 2019 roku do godziny 10:00 w formie zgodnej z instrukcjami przekazanymi przez Zamawiającego. Ostatnie Rozwiązanie zgłoszone przez danego Uczestnika przed wskazaną godziną w dniu zakończenia Etapu I traktowane jest jako ostateczne. Rozwiązania złożone po tym terminie nie będą rozpatrywane.
- 10.2. Rozwiązanie Zadań I Etapu podlega ocenie Sądu Konkursowego, który dokonuje jej do dnia 21 stycznia 2019 roku.

3. Rozdział 10 Punkty 4-8 Regulaminu otrzymują nowe brzmienie:

- 10.4. Dla uczestników, zakwalifikowanych do II Etapu Konkursu zostanie zorganizowane 3-dniowe seminarium (dedykowane spotkanie informacyjne), w którym winni uczestniczyć wszyscy Uczestnicy zaproszeni przez Zamawiającego. Planowany termin seminarium to 29-31 stycznia 2019 roku. W czasie seminarium Zamawiający odpowie na pytania Uczestników II Etapu oraz przedstawi materiały dedykowane II Etapowi i inne niezbędne informacje. Wnioski z seminarium zostaną spisane w protokole, który zostanie przesłany do wszystkich Uczestników Konkursu zakwalifikowanych do Etapu II. Seminarium to jest częścią Etapu II Konkursu.
- 10.5. Etap składania Rozwiązań przez Uczestników prowadzony w ramach Etapu II Konkursu, będzie trwał od 5 lutego do 5 marca 2019 roku do godziny 10.00. Ostatnie Rozwiązanie zgłoszone przez

-
- danego Uczestnika przed godziną 10.00 w dniu zakończenia Etapu II traktowane jest jako ostateczne
- 10.6. Etap składania Rozwiązań przez Uczestników prowadzony w ramach Etapu II Konkursu, będzie trwał od 5 lutego do 5 marca 2019 roku do godziny 10.00. Ostatnie Rozwiązanie zgłoszone przez danego Uczestnika przed godziną 10.00 w dniu zakończenia Etapu II traktowane jest jako ostateczne.
- 10.7. Szczegółowy opis Zadań Etapu II Konkursu zostanie przekazany Uczestnikom pocztą elektroniczną w terminie do dnia 5 lutego 2019 roku.
- 10.8. Rozwiązania składane w Etapie II Konkursu podlegają ocenie Sądu Konkursowego, który dokonuje jej do dnia 27 kwietnia 2019 roku, w oparciu o kryteria określone w Regulaminie.

IV. Odpowiedzi udzielone na pytania zadane przez Uczestników:

1. Chciałbym doprecyzować - jak rozumiem zależałoby Państwu na wykrywaniu przypadków niewłaściwego zaklasyfikowania do diagnozy głównej (?) Przy czym domyślam się, że w sensie kosztów interesują Państwa przypadki up-codingu - czyli klasyfikowania do kodu, za którym idą większe płatności. Jeśli tak zdefiniujemy zadanie trzeba by odnieść się do wyceny za dany kod ICD - tak by nie uwzględniać przypadków down-coding.
Odp. Zadanie polega na wykrywaniu nadużyć wywołujących negatywne skutki finansowe dla publicznego płatnika. W szczególności nadużycia te mogą polegać na tzw. upcodingu czyli zmianie parametrów wpływających na rozliczenie z płatnikiem (np. rozpoznanie, wykonane procedury) w celu podwyższenia płatności.
2. Czy w przypadku wszystkich pacjentów leczenie odbywało się tylko w publicznej służbie zdrowia czy także w prywatnej? Jeśli w obydwu -będziemy mieli do czynienia z różnymi populacjami - uzasadnionym byłoby posiadanie pola wskazującego na rodzaj miejsca (inne rodzaje usług - publiczny sektor - dominują procedury zabiegów, prywatny - konsultacje i czasem zabiegi)
Odp. Dane dotyczą jedynie kontaktów z publiczną służbą zdrowia.
3. Co dokładnie powinna zawierać oferta w Etapie I?
Odp. Oferta powinna zawierać: a) maksymalny planowanym łączny koszt wykonania prac realizowanych na podstawie Rozwiązania (kwota brutto), b) informacje o wykorzystywanych licencjach Open Source (zgodnie z Regulaminem, pkt 14.3 i 14.4), c) opis Rozwiązania (maksymalnie 5 stron A4). Opis Rozwiązania może w szczególności zawierać wstępną koncepcję IT rozwijaną i podlegającą ocenie w II etapie konkursu obejmującą i) koncepcję wdrożenia i współpracy z Zamawiającym, ii) koncepcję wymiany informacji z innymi systemami (np. bazami danych), standaryzacji interfejsów oraz obsługi formatów plików, iii) koncepcje organizacji interakcji użytkownik-oprogramowanie, iv) koncepcję możliwości rozwoju Rozwiązania przez Zamawiającego.

4. Czy dane wejściowe, na podstawie których Komisja podda ocenie rozwiązanie dla Etapu I, będą spójne strukturą i kontekstem urologicznym względem danych otrzymanych od organizatora?

Odp. Tak

5. Czy oferta (Regulamin 3.2, pkt 3) dotycząca proponowanego oprogramowania do wykrywania nadużyć, uwzględniająca maksymalny koszt wykonania prac, stanowić będzie jedno z kryteriów oceny Uczestników, którzy przystąpią do II Etapu konkursu?

Odp. W I etapie oferta ma głównie charakter informacyjny i będzie podlegała jedynie ocenie pod względem spełnienia wymagań określonych w regulaminie (w szczególności w punktach 3.2.3, 3.2.4, 14.3, 14.4).

6. Czy wytyczne dla rozwiązania (Regulamin 3.3) obowiązują również oprogramowanie wytworzone podczas trwania I Etapu?

Odp. Tak, przy czym należy pamiętać, że Rozwiązanie I etapu ma charakter pracy studialnej i może być rozwijane w dalszym toku Konkursu.

7. Czy wskazanie licencji Open Source, wykorzystanych elementów do stworzenia rozwiązania, jest konieczne już na I Etapie konkursu?

Odp. Jeżeli Uczestnik wykorzystuje w Rozwiązaniu elementy na licencji typu Open Source to zgodnie z 14.3 Regulaminu jest on zobowiązany wskazać te elementy Rozwiązania, poinformować o tym fakcie Zamawiającego oraz dostarczyć treść licencji określającej dokładne warunki korzystania. Punkt ten dotyczy Rozwiązań składanych na każdym etapie konkursu.

8. Pytanie odnośnie formatu pliku wyjściowego, w regulaminie jest:

"Id Produktu: Id produktu (odpowiadający id z próbki), w przypadku którego w opinii Uczestnika doszło do nadużycia"

Dla danych z pliku: GOVTECH_SWIADCZENIA, kolumna: ID_PRODUKTU w tabeli wypełniona jest jedynie w 8517 rekordach, przy czym jest w nich tylko 56 unikalnych wartości ID_PRODUKTU, kolumna: ID_KODU_PRODUKTU_JEDN jest wypełniona dla wszystkich rekordów (93 822), a jest w niej: 721 unikalnych wartości. Prosiłbym o dokładne uściślenie o dokładnej kolumnie którą mamy traktować jako: "id z próbki" (czy jest to kolumna: ID_POZYCJI_ROZLICZENIOWEJ, która jest zawsze wypełniona i unikalna dla każdego rekordu?)

Aby usunąć niejasności prosiłbym o podanie przykładowego rekordu nadużycia jaki może zostać zaraportowany ze wskazaniem z których plików wartości są pobrane (oczywiście przykład może nie mieć sensu biznesowego, a jedynie sens walidacyjny).

W regulaminie jest mowa o następujących polach:

- Id Produktu
- Kod Sprawozdany
- Kod Poprawny

- Uzasadnienie

Mamy wątpliwości odnośnie pola Id Produktu.

W regulaminie jest mowa, że odpowiada ono "Id z próbki".

Nie jesteśmy w stanie zidentyfikować do jakiej kolumnie w danych się ono odnosi.

W tabeli SWIADCZENIA ID_PRODUKTU ma taką samą wartość jak ID_KODU_PRODUKTU_JEDN.

To samo ID_PRODUKTU pojawia się w wielu wierszach - również dla różnych pacjentów.

Nie identyfikuje ono, więc jednoznacznie świadczenia.

Odp. Ocenie pod kątem wystąpienia nadużyć podlegają pozycje rozliczeniowe posiadające ID_PRODUKTU.

Przykładowe raportowanie:

Id Produktu	Id Pozycji Rozliczeniowej	Kod sprawozdany	Kod poprawny	Uzasadnienie
2500	2000001	5.51.01.0003042	5.51.01.0003044	Wykonane procedury, historia pacjenta i rozpoznania wskazują na to, że sprawozdana procedura '21.841 Plastyka przegrody nosowej' była w rzeczywistości procedurą '21.13 Nacięcie przegrody nosowej'

ZMIANA REGULAMINU

Pojedyncze nadużycie powinno być raportowane w formie rekordu, wg następujących pól:

Nazwa Pola	Opis	Obowiązkowe/nieobowiązkowe
Id Produktu	Id produktu odpowiadający id z próbki (ocenie podlegają jedynie rekordy z próbki o niepustej zawartości pola ID_PRODUKTU)	Pole obowiązkowe
Id Pozycji Rozliczeniowej	Id pozycji rozliczeniowej (zgodny z danymi z próbki) w przypadku, którego w opinii Uczestnika doszło do nadużycia	Pole obowiązkowe
Kod Sprawozdany	Sprawozdany (zgodny z danymi z próbki) kod produktu jednostkowego	Pole obowiązkowe
Kod Poprawny	Kod produktu jednostkowego, który wg Uczestnika powinien być sprawozdany w przypadku tego produktu	Pole obowiązkowe
Uzasadnienie	Uzasadnienie	Pole opcjonalne w I etapie / Pole obowiązkowe w II etapie

9. W regulaminie jest napisane, że w rozwiązaniu należy podać Id Produktu i Kod Sprawozdany. W pliku GOVTECH_SWIADCZENIA.csv pole ID_PRODUKTU jest tożsame z polem ID_KODU_PRODUKTU_JEDNOSTKOWEGO, które prowadzi do kodu sprawozdanego. Czy w regulaminie nie ma błędu i nie chodziło o ID_PACJ z pliku GOVTECH_SWIADCZENIA.csv?

Odp. Patrz pytanie 8

10. Jakiego pola użyć jako identyfikatora usługi? W regulaminie jest ID_PRODUKTU, ale to pole nie jest unikalne w zbiorze danych.

Odp. Unikalnym identyfikatorem rekordów podlegających ocenie jest para ID_POZYCJI_ROZLICZENIOWEJ, ID_PRODUKTU

11. Czy interesują nas też nadużycia typu zgłoszenie usługi, która nie miała miejsca w rzeczywistości (tzn. nie ma świadczenia którym należy je zastąpić)

Odp. Tak, w tym przypadku kwota nadużycia jest całkowita wartość danego produktu. Informacje o tym należy umieścić w uzasadnieniu.

12. Ponad połowa świadczeń posiada identyfikator "ID_KOD_SWIADCZENIODAWCY" = 999999. Jak interpretować tą wartość? Czy jest to po prostu nieznan kod świadczeniodawcy?

Odp. Tak, jest to nieznan kod świadczeniodawcy.

13. Prosimy o sprecyzowanie opisu danych: zgadujemy, że pacjenci podczas jednego pobytu/kontakt (czyli w ramach jednego ID_KONT) mogą mieć wiele "produktów" czyli usług rozliczanych przez NFZ. Jakie znaczenie mają wobec tego procedury? Czy w ramach jednego produktu może być wiele procedur? Jak interpretować to, że w pliku z procedurami jest dużo mniej ID_KONT niż w pliku ze świadczeniami (8142 vs 61635)?

Odp. W pliku z procedurami jest mniej kontaktów ponieważ dotyczą one tylko produktów podlegających badaniu (niepuste ID_PRODUKTU). W ramach jednego produktu rozliczeniowego może być wykonanych wiele procedur.

14. Czy jest możliwość otrzymania próbki uczącej (czyli próbki dla której mamy zdarzenia typu świadczenia itp. oraz zdarzenia "nadużycie")

Odp. Organizator nie zapewnia takiej próbki w ramach procedury konkursowej.

15. Czym różni się ICD9 od "produktu"? Co dokładnie zawierają tabele: spec_kom oraz ICD_10 współlistniejące? Gdzie można znaleźć wycenę w ICD9 oraz produktu?

Odp. Informacje na ten temat można znaleźć: <http://www.nfz.gov.pl/dla-swiadczeniodawcy/jednorodne-grupy-pacjentow/komunikaty-jgp/informacja-dotyczaca-publicacji-opisu-algorytmu-jgp,43.html> . Cennik produktów rozliczeniowych znajduje się w Zarządzeniach Prezesa NFZ, przy czym należy pamiętać, że na potrzeby oceny Rozwiązania konkursowego jako obowiązujący przyjęty zostanie stan prawny wynikły z wprowadzenia Zarządzenia Prezesa NFZ nr 88/2015/DSOZ bez późniejszych zmian.

16. Gdzie można znaleźć znacznik geograficzny (dla pacjenta oraz dla jednostki, która dostarczyła usługę)?

Odp. W próbce danych nie ma takiej informacji.

17. Czy problemem, który należy zidentyfikować jest to, czy danemu pacjentowi należycie przypisano dany produkt, tj. czy chcemy identyfikować anomalie w przypisywaniu produktów?

Odp. Celem jest identyfikacja rozliczeń niekorzystnych dla płatnika.

18. Czy też chcemy wykrywać, czy szczegółowe wykonanie poszczególnych badań jest odpowiednio przypisane do produktu?

Odp. Jeżeli według uczestnika może mieć to wpływ na niekorzystne dla płatnika rozliczenie świadczenia, wtedy można uznać to za nieprawidłowość.

19. Czy są jakiegokolwiek wskazówki dot. oferty na proponowane rozwiązanie, które należy przesłać wraz z tabelą wyników? Czy jest jakiś konkretny format (docx., pptx?), czy jest jakieś ograniczenie co do ilości stron, slajdów? Czy musi być to tylko opis tekstowy, czy mogą być ilustracje?

Odp. Informacje zawarte w punkcie 3.

20. Czy są jakieś wymogi jakie proponowane rozwiązanie powinno spełniać? Czy są jakieś ograniczenia, wymogi jakie oprogramowanie powinno spełniać?

Odp. Informacje o istotnych wymogach jakie powinno spełniać Rozwiązanie zawiera Regulamin (pkt 3.3.).

21. Na jaki adres należy wysłać plik z ofertą oraz plik z tabelą?

Odp. Rozwiązania po anonimizacji (zaszyfrowane zgodnie z treścią Załącznik nr 1 Instrukcja szyfrowania prac konkursowych) należy kierować do Sekretarza Konkursu (m.czak@csioz.gov.pl).

22. „Czy Zamawiający dopuszcza, aby w przekazywanym rozwiązaniu w etapie I, pole „Kod Poprawny” było puste w przypadku rekordów, które nie powinny być sprawozdane? Czyli takich, dla których nie ma prawidłowego kodu. Zgodnie z regulaminem pole „Kod Poprawny” jest obowiązkowe. Stąd pytanie czy dopuszczona jest wartość pusta, a jeśli nie to jaka jest oczekiwana wartość tego pola w ww. przypadku.”

Odp. Tak. Jest to rozwiązanie dopuszczalne, jednak powinno być odpowiednio wyjaśnione w uzasadnieniu.

*Zastępca Dyrektora ds. Kluczowych
Systemów Ochrony Zdrowia*

(-) Wiktor Rynowiecki

Załącznik nr 1 – Instrukcja szyfrowania prac konkursowych
Załącznik nr 2 – Klucz publiczny
Załącznik nr 3 – Wzór oferty

Sporządził: Pędzińska Aleksandra