Warszawa, dn. 2 kwietnia 2020 r.

**PROTOKÓŁ**

**z przeprowadzenia dialogu technicznego związanego z postępowaniem o udzielenie zamówienia publicznego, którego przedmiotem jest stworzenie programu do monitoringu radiowego**

**umożliwiającego automatyzację analizy treści programów radiowych**

1. Do dialogu technicznego przystąpiło czterech Uczestników:

* VoiceLab.AI Spółka z o.o.;
* AGB Nielsen Media Research Spółka z o.o.;
* Tibigen Arkadiusz Węgłowski wraz z Instytutem Chemii Bioorganicznej PAN, Poznańskim Centrum Superkomputerowo-Sieciowym oraz TechInnowacje Spółka z o.o.;
* Krzysztof Gurba - Gurba 5.0.

Zespół ds. dialogu technicznego spotkał się indywidualnie z każdym z ww. Uczestników. W wyniku spotkań zebrane zostały następujące informacje na temat uczestników dialogu oraz wstępnej koncepcji Uczestnika co do sposobu realizacji zamówienia:

* **VoiceLab.AI Spółka z o.o.**

Zaproponowane rozwiązanie bazuje na rozpoznawaniu i transkrypcji mowy. Nagrania audio są analizowane i przetwarzane na tekst, z podziałem na mówców. Na podstawie tych tekstów możliwe jest wielopoziomowe wyszukiwanie konkretnych treści/haseł w nagraniach. Możliwe jest też rozpoznawanie zdefiniowanych głosów (voice-matching) i na tej podstawie wyszukiwanie wypowiedzi tych osób (tzw. „setek”). Do voice-matchingu potrzebne jest odpowiednie przygotowanie próbek głosowych.

Oprogramowanie umożliwia jednoczesne monitorowanie na bieżąco stacji radiowych (i innych treści ze ścieżką audio) na podstawie streamingu. Można je stosować nie tylko do nagrań programów radiowych, ale także do wszelkich plików ze ścieżką audio (np. VoD).

Warunkiem prawidłowego działania tego rozwiązania jest dobra (bez szumów) jakość nagrania.

Do spełnienia części wymagań Zamawiającego Uczestnik musi (i twierdzi, że ma do tego odpowiednie zaplecze oraz wiedzę) przygotować dedykowane oprogramowanie. Wymóg dotyczący rozpoznawania piosenek wymagałby po stronie Uczestnika kupna tego rodzaju rozwiązania, bądź stworzenia bazy referencyjnej i odpowiedniego jej oprogramowania.

Uczestnik realizuje zamówienia związane z rozpoznawaniem tekstu dla: banków do kontaktu z klientami (wdrożone), Senatu i dla sądów (w trakcie przygotowań) .

Oferuje dostęp do strony testowej, z prowadzonymi na bieżąco nagraniami programów radiowych dostępnych w streamingu internetowym i wybranymi programami VoD. Testowanie możliwe jest po zarejestrowaniu konta w serwisie [newsense.ai](http://newsense.ai).

* **AGB Nielsen Media Research**

AGB prowadzi ciągły monitoring stacji telewizyjnych na potrzeby badania telemetrycznego. W niektórych krajach prowadzi również ciągły monitoring stacji radiowych. Do realizacji tych monitoringów dysponuje narzędziem *MediaPad*, które umożliwia tagowanie i opisywanie nagrań, a także wprowadzanie elementów nagrań do bazy, dzięki czemu są one później automatycznie rozpoznawane (np. reklamy). Taki monitoring automatyzuje część zadań, ale wymaga też pracy operatora/użytkownika.

Obecnie Biuro KRRiT korzysta z danych telemetrycznych Uczestnika i jego oprogramowania do analizy danych *Arianna*. Uczestnik zaproponował podczas pierwszego spotkania trzy ogólne rozwiązania:

1. Monitoring programów radiowych z wykorzystaniem narzędzia *MediaPad* w ramach usługi Uczestnika;
2. Udostępnienie, w ramach umowy, narzędzia *MediaPad* (z ewentualnymi dostosowaniami do naszych potrzeb), aby Biuro KRRiT mogło samodzielnie prowadzić monitoring;
3. Przygotowanie nowego narzędzia, dedykowanego dla KRRiT (co jednak wymagałoby przeznaczenia większej ilości czasu na stworzenie oprogramowania i wdrożenie; Uczestnik szacuje ten czas na ok. 18 miesięcy).

Uczestnik nie dysponuje narzędziem do rozróżniania w nagraniu muzyki od słowa. Atutem jest wygodny i sprawdzony w praktyce interfejs do nawigowania po nagraniach, a także możliwość wczytania danych z monitoringu do narzędzia analitycznego, jakim jest *Arianna*.

* **Tibigen Arkadiusz Węgłowski wraz z Instytutem Chemii Bioorganicznej PAN, Poznańskim Centrum Superkomputerowo-Sieciowym (PCSS) oraz TechInnowacje Spółka z o.o.**

Firma Tibigen dysponuje narzędziami do zarządzania materiałami audiowizualnymi. PCSS posiada laboratoria, w tym przetwarzania mowy i audiowizualnym (analiza zdjęć), a także infrastrukturę techniczną, za pomocą której pozyskuje sygnały z dowolnych źródeł (ukf, internet etc.) i przekazuje je do archiwum audio i video. TechInnowacje jest spółką powołaną przy PCSS do komercjalizacji projektów.

Przykłady zrealizowanych projektów:

* iTVP – platforma udostępniania archiwów audiowizualnych (2007);
* Net Radio – system nadawania kanałów radiowych w internecie;
* Dla UAM – platforma do analizy mediów – nagrywanie, archiwizowane i analiza treści programów informacyjnych (transkrypcja);
* System dla NiNA;
* Projekt dla MSWiA.

KRRiT może otrzymać system zaprojektowany do analizy plików audio, w którym możliwe jest automatyczne rozróżnianie muzyki i słowa, jingli i innych stałych elementów programu oraz transkrypcja mowy; w przyszłości może on być rozwijany dla innych typów mediów.

* **Krzysztof Gurba - Gurba 5.0**

Uczestnik deklaruje, że jest w stanie dostarczyć produkt, który umożliwi rozróżnienie mowy od muzyki, rozpoznanie słowa w czasie rzeczywistym i jego kontekstowe powiązanie. Produkt ma zostać stworzony na bazie stosowanego przez Uczestnika systemu NLP (Natural Language Processing) opartego na głębokich sieciach neuronowych. Ta technologia pozwala na bardziej precyzyjne wyszukiwanie haseł, czy fraz, niż zwykła transkrypcja. System może sugerować użytkownikowi tematykę audycji, np. regionalną, sportową etc. Na dalszym etapie prac Uczestnik może zaimplementować również wspomnianą na wstępie metodę oceny bezstronności przekazu.

Uczestnik jest w stanie dostarczyć również interfejs zaprojektowany dokładnie pod kątem potrzeb KRRiT (np. sterowanie głosem, gestami etc.). Wdrożenie systemu wymaga stworzenia specjalnego słownika (na podstawie próbnych nagrań). Słownik ten musiałby być raz na jakiś czas (np. co roku) aktualizowany.

Uczestnik sugerował, że automatyzacja analizy nagrań byłaby znacznie łatwiejsza i dokładniejsza, gdyby nadawcy stosowali tak zwany „watermarking” do oznaczania elementów programu (np.  spotów reklamowych, audycji, muzyki itp.). KRRiT, jako regulator, mogłaby promować takie rozwiązanie, jako korzystne również dla nadawców.

1. Zestawienie funkcjonalności rozwiązań proponowanych przez Uczestników dialogu zawarto w tabelach stanowiących Załącznik do niniejszego protokołu.
2. Dialog zakończył się w dniu 15 marca 2020 roku.

Dyrektor

Biura KRRiT

*Maciej Myszka*