

TECHMO DEVS' DIARY # 2

06.2022



- ▶ W nowej wersji Techmo Dictation 2.16.5 wdrożyliśmy mechanizm zarządzania plikami logów, z uwzględnieniem opcji zachowania archiwalnych logów przy starcie serwisu.
- ▶ Techmo ASR MRCP 1.12.9 - dodaliśmy konfigurowalny typ zwracanego statusu w nagłówku 'completion-cause' oraz poprawiliśmy format odpowiedzi w przypadku zbyt długiej wypowiedzi klienta (poprzednio 'success-maxtime').
- ▶ Udostępniliśmy nową wersję modelu medycznego - 202208_medical, który został uzupełniony o ogólne słownictwo medyczne.
- ▶ Przygotowaliśmy nową wersję modelu akustycznego dla języka angielskiego. Dzięki odpowiednim zmianom w procesie treningu modelu oraz zastosowaniu metod wzbogacania danych treningowych model charakteryzuje się lepszym działaniem w warunkach szumu i hałasu tła.
- ▶ Dla modułów dodatkowych ASR: detekcji wieku, płci i rozpoznawania języka opracowaliśmy nowy, wspólny format danych testowo-treningowych, dzięki czemu jesteśmy gotowi na szybką reakcję, lepszą obsługę zgłoszeń oraz dalsze usprawnianie tych modułów.
- ▶ Opracowaliśmy testowe modele dla języków angielskiego i hiszpańskiego z detekcją słów spoza leksykonu (OOV – Out-of-Vocabulary). Dzięki zastosowaniu mechanizmu OOV, system nie będzie wymuszał rozpoznania najbardziej podobnego słowa.

[WORK IN PROGRESS]

- ▶ Przeprowadzamy aktualizację modułów detekcji wieku, płci i rozpoznawania języka dodając wybrane algorytmy detekcji mowy VAD oraz wspólny mechanizm przetwarzania i parametryzowania sygnału audio.



- ▶ Udostępniliśmy wersję TTS 2.5.3. Poprawiliśmy obsługę efektów postprocessingu głosów w celu dostarczenia jak najlepszych profili generowanej mowy do wybranego zastosowania.
- ▶ Stworzyliśmy nowy silnik generowania transkrypcji fonetycznej dla języka polskiego, oparty o międzynarodowy alfabet fonetyczny IPA.
- ▶ Zaadaptowaliśmy aktualne wersje modeli głosów do nowej wersji systemu TTS 3.0, nad którą pracujemy.

[WORK IN PROGRESS]

- ▶ Rozpoczęliśmy prace nad stworzeniem środowiska TTS Studio, które w łatwy sposób pozwoli tworzyć treści, kontrolować i dostosowywać charakter generowanych wypowiedzi, wykorzystując Techmo TTS.



- ▶ Techmo Voice Reporter czyli system rozpoznawania mowy do zastosowań przemysłowych został rozszerzony o możliwość instalacji i uruchamiania w środowisku Windows.
- ▶ Udostępniliśmy nową wersję systemu zarządzania licencjami typu High Availability (HA) czyli Techmo LMS. Wersja 2.0.0, uwzględnia rozbudowaną obsługę raportów, zarządzanie instancjami serwisów, konfigurowalne mechanizmy logowania oraz poprawioną obsługę baz MSSQL.