

Rynek Polska 2

Rynek świat 3

Techmo Developers' Diary 4

Publikacje naukowe 5



Zapraszam do zapoznania się z nowym Raportem Techmo. Tworzymy go po to, by dzielić się wiedzą i inspirować naszych Partnerów i Klientów. Mamy świadomość, że młody rynek rozwiązań głosowych wymaga stymulacji i kolejnych świeżych pomysłów. Jednocześnie głupotą byłoby nie korzystać z doświadczeń innych branż czy też spostrzeżeń mądrych ludzi, którzy widzieli niejedno. Tego typu treść dostarcza na przykład artykuł Forbes'a, który przytaczamy w tym raporcie. 12 ekspertów wskazuje najważniejsze kierunki rozwoju aplikacji głosowych. Jak zwykle bywa jest to jakaś optyka. Pewnie wymagać będzie w przyszłości korekty. Niemniej trudno się nie zgodzić ze stawianymi tam тезami. Przykład?

Narzędzia głosowe pozwolą medykom szybciej i bardziej bezpiecznie zajmować

się pacjentami. Branża medyczna jest gotowa i złaźniona rozwiązań głosowych. Doświadczenia pandemii tylko zwiększyły ten apetyt i gotowość. W dobie rosnących oczekiwań, starzejącego się społeczeństwa oraz w obliczu nieprzewidywalnych okoliczności - czas na zmianę utartych ścieżek. Głos pozwoli szybciej diagnozować, obsługiwać wzmożony ruch pacjentów, nawigować wśród masy danych czy sterować aparaturą medyczną. Na tym tle wręcz banalne wydaje się dyktowanie recept czy epikryzy przez lekarza. A jednak wciąż czekamy na polskie rozwiązanie głosowe z prawdziwego zdarzenia. A teraz uwaga, ono nadciąga. Czy też powinienem powiedzieć: Nadciągamy!

Miłej lektury jak zawsze.

Piotr Stankiewicz

CBDO, Techmo

tel.: +48 535 215 200

e-mail: piotr.stankiewicz@techmo.pl



P.S. Zachęcam do kontaktu ze mną, jestem zawsze ciekaw nowych spojrzeń, przemyśleń czy pomysłów.

INFORMACJE Z POLSKI

Zobaczyć i usłyszeć

Wsparcie dla niedowidzących dzięki zastosowaniu AI - urządzenie, które obserwuje otoczenie i udziela informacji.

źródło: biznes.trojmiasto.pl

Zmiany w Unimie 2000

Sprzedaż, fuzja, przejęcie? Zarząd rozważa różne opcje.

źródło: infostrefa.com

IKO z nowymi funkcjami

PKO Bank Polski rozwija aplikację mobilną. Także asystent głosowy zyskał kolejne tematy i użyteczności.

źródło: money.pl

Unijne strategie AI dla Polski

Cyfryzacja przemysłu i administracji, stabilne regulacje, zaangażowanie naukowców - to obszary, które w naszym kraju powinny najbardziej zmienić się dzięki rozwiązaniom AI.

źródło: calg.pl

Demokratyzacja robotyzacji

Czy specjaliści biznesowi mogą zastąpić informatyków, a boty - specjalistów?

źródło: dfe.org.pl

Automatyzacja administracji

O voicebotach w polskich i zagranicznych samorządach, które pomogą nie tylko interesantom.

źródło: filarybiznesu.pl

Kup pan cegłę

Dwujęzyczny voice-chat-bot od KodaBots obsługuje klientów w real estate.

źródło: kodabots.com

Voicebot dla wygranych

W agencji UNIQUE ONE wirtualna asystentka weryfikuje zwycięzców konkursów promocyjnych.

źródło: smb.pl

Polisa ma głos

Dronn i kolejny asystent głosowy w branży ubezpieczeniowej - tym razem w Prévoir.

źródło: prevoir.pl

Anty-koronabot

W Zakopanem voicebot rejestruje na szczepienia p/COVID-19.

źródło: podhale24.pl

INFORMACJE ZE ŚWIATA

Co trzeba wiedzieć o technologii głosowej?

12 osób z branży głosowej - członków Forbes Technology Council - i 12 różnych spojrzeń na technologię głosową. Wskazówki przydatne dla konsumentów jak i twórców technologii.

źródło: forbes.com

UE zadba o jakość narzędzi AI

Komisja Europejska dąży do stworzenia ram prawnych oraz strategii europejskiej dla spopularyzowania AI w gospodarce.

źródło: ec.europa.eu

Quo vadis RPA?

Analitycy z firmy Gartner przedstawili prognozy rozwoju technologii na najbliższe dwa lata: 24-procentowy wzrost oraz wymuszenie na organizacjach używania min. 3 produktów do automatyzacji procesów.

źródło: venturebeat.com

Biometryczna tarcza na wyłudzenia

W roku 2020 biometria głosowa banku HSBC uchroniła ćwierć miliarda funtów od wyłudzenia. W stosunku do roku 2019 (o którym pisaliśmy rok temu), łączna kwota poddana próbie wyłudzenia spadła o 40 procent.

źródło: computerweekly.com

ASR uczy się z radia

Eksperyment w rozpoznawaniu mowy nigeryjsko-kongijskiej na podstawie archiwalnych nagrań radiowych. Efekt to 88-procentowa dokładność i załóżek dla niszowego asystenta głosowego.

źródło: venturebeat.com

Bot zrozumie sarkazm

Pomysł do wdrożenia w sekcjach komentarzy w social media. Kluczem do dobrego algorytmu jest wybranie prawidłowych słów-wskazówek oraz głębokie zrozumienie ich zależności.

źródło: neurosciencenews.com

Nowa jakość asystentów głosowych?

Google na swojej konferencji pokazał nowe narzędzie do tworzenia dialogów konwersacyjnych. Model pozwoli AI na generowanie bardziej precyzyjnych i różnorodnych odpowiedzi.

źródło: wccftech.com

Kto najlepszy z lekami

Trzej popularni asystenci głosowi i 50 najpopularniejszych leków w USA. Kto wygrywa, a kto próbuje nadrabiać? Wciąż duży problem stanowi obcy akcent.

źródło: businesswire.com

Więcej o Techmo:

TECHMO DEVELOPERS' DIARY

Serwis Licencyjny

Aktualna wersja: 2.0.0

- Uprościliśmy sposób instalacji serwisu w infrastrukturze klienta, zarówno dla rozwiązań typu HA jak i rozliczania sekundowego;
- Dodaliśmy możliwość jednoczesnej obsługi Dictation, TTS i TTS GPU w jednej instancji serwisu licencyjnego;
- Udostępniliśmy API pozwalające sprawdzić wykorzystanie serwisów w instalacjach on-premises;
- Zmniejszyliśmy ilość przechowywanych danych w bazie;

Techmo Dictation

- Opracowaliśmy prototyp bota mającego odbierać telefon zaufania dla osób szukających pomocy psychologicznej;
- Stworzyliśmy model ASR dla języka angielskiego wzbogacony o korpus tekstów z dziedziny zdrowia psychicznego;
- Aktualna skuteczność nowych modeli dla języka angielskiego w WER (Word Error Rate):
 - Zbiór LibriSpeechClean 100 – WER 4.62% (poprzednio 7.15%)
 - Zbiór LibriSpeechClean 360 – WER 4.57% (poprzednio 7.48%)
 - Zbiór LibriSpeechOther – WER 10.94% (poprzednio 16.35%)
 - Zbiór CommonVoice – WER 2.2% (poprzednio 18.54%)
- Aktualne wyniki nowych modeli dla języka hiszpańskiego:
 - Zbiór Common Voice – 14.17%WER (poprzednio 45.37%)

Techmo TTS

Aktualna wersja: 2.3.0

- Dodaliśmy możliwość definiowania słowników wymowy w trakcie działania serwisu, w sposób zgodny z PLS (Pronunciation Lexicon Specification);
- Dodane zostały metryki dla monitorowania serwisu TTS z wykorzystaniem systemu Prometheus, np.: czas życia serwisów, ilość przetworzonych zapytań, czas obsługi zapytań, szybkość syntezy (RTF – Real Time Factor);
- Uproszczony został sposób obsługi błędów – usunęliśmy dodatkowe pole z API i zastąpiliśmy obsługę kodami statusów GRPC;

Techmo TTS GPU

Aktualna wersja: 2.2.0

- Udostępniliśmy możliwość korzystania z TTS GPU w połączeniu z serwisem licencyjnym;
- Rozszerzona została konfiguracja mechanizmów logowania (sposób rotacji plików, rozmiar logów, format – json lub txt);
- Zoptymalizowaliśmy sposób budowania i dostarczania serwisu dla klienta, wygenerowane obrazy docker zajmują do 50% mniej przestrzeni dyskowej;

PUBLIKACJE NAUKOWE

Unsupervised Speech Recognition

źródło: ai.facebook.com

Masked Language Modeling and the Distributional Hypothesis: Order Word Matters Pre-training for Little

źródło: arxiv.org

PyTorch Tabular: A Framework for Deep Learning with Tabular Data

źródło: arxiv.org

AM2 ICO: Evaluating Word Meaning in Context across Low-Resource Languages with Adversarial Examples

źródło: arxiv.org

Parallel Tacotron 2: A Non-Autoregressive Neural TTS Model with Differentiable Duration Modeling

źródło: arxiv.org

Towards Multi-Scale Style Control for Expressive Speech Synthesis

źródło: arxiv.org

Grad-TTS: A Diffusion Probabilistic Model for Text-to-Speech

źródło: arxiv.org