

TECHMO

Voice Technologies

RAPORT RYNKOWY #8

stan na 30.11.2020 r.

Lista Odbiorców Raportu Rynkowego Techmo jest ograniczona. Z założenia otrzymuje go jedna osoba w danej organizacji, z którą Techmo ma podpisaną Umowę Partnerską. Ewentualną dalszą dystrybucję raportu pozostawiamy do decyzji Odbiorcy.

Informacje z Polski

Najbardziej cyfrowy ubezpieczyciel



Bezpieczny głos tylko dla wojska



Voicebot w wydawnictwie



Analiza voicebota



Brak wzajemnych roszczeń



Google dla polskich start-up'ów



Informacje ze Świata

COVID znaleziony w głosie



Higieniczny głos



Popularne, ale nie doskonałe



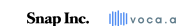
Kosztowny AI w Europie



Komputer lepszy w mowie spontanicznej



Snapchat bardziej głosowy?



Publikacje naukowe

Najbardziej cyfrowy ubezpieczyciel

TUZ Ubezpieczenia cyfryzuje się i inwestuje w IT. Wprowadzone rozwiązania to nowoczesne i elastyczne API. Równocześnie znacznie wzrosła wydajność systemu taryfikacyjnego: z 9 tys. zapytań dziennie do 25 tys. Średni czas odpowiedzi to teraz 4,5 sekundy, wobec wcześniejszych ponad 20. Cyfryzacja dotyczy także backoffice. Dziś TUZ to druga, najszybciej rosnąca firma ubezpieczeniowa. Niedawno została uznana za „Najbardziej cyfrowego ubezpieczyciela” podczas Insurance Forum.

źródło: gu.com.pl ↗

Bezpieczny głos tylko dla wojska

Zespół naukowców z Wydziału Elektroniki WAT opracował mikrotelefon, który potwierdzi tożsamość rozmówcy. Urządzenie działa podczas rozmów przez telefon, komunikator internetowy i radiostację. Uzupełnieniem technologii jest token znaku wodnego, który rozpoznaje tożsamość w mediach cyfrowych poprzez unikalny PIN. Aparaturę wyprodukowano w ograniczonej liczbie egzemplarzy i nie jest dostępna na rynku usług komercyjnych ani sprzętu elektronicznego.

źródło: naukawpolsce.pap.pl ↗

Voicebot w wydawnictwie

Wolters Kluwer Polska, które specjalizuje się w publikacjach prawniczych i biznesowych, do obsługi klientów w infolinii zaangażowało voicebota. Piotr, bo tak ma na imię, specjalizuje się w rozliczeniach zakupów i zwrotach towarów. Jak podkreśla Joanna Konieczna Sielska, Dyrektor Logistyki i Obsługi Klienta, głównym celem wydawnictwa było podniesienie komfortu kontaktów z CC dla klientów. Autorem wdrożenia jest InteliWISE, partner Techmo. Wolters Kluwer planuje poszerzać zakres usług opartych na AI.

źródło: inteliwise.pl ↗

Analiza voicebota

Interesujące case study dla branży głosowej przedstawiła firma Iteo z Katowic. Na blogu zaprezentowała i opisała rozwiązania zastosowane w swoim voicebocie Hellobot. Są to ASR, AI, TTS i scenariusz rozmowy. Iteo przedstawiło również metodykę testowania. Szczególną uwagę zwrócono na weryfikację scenariuszy rozmów, dążąc do tego aby każda pierwsza konwersacja z botem, zakończyła się zgodnie z zaplanowanym schematem. Z informacji uzyskanych u źródła wiemy, że Hellobot ma już za sobą kilka komercyjnych wdrożeń.

źródło: iteo.com ↗

Brak wzajemnych roszczeń

Biznes.pap.pl informuje o porozumieniu między Summa Linguae Technologies S.A. z Krakowa, a funduszem inwestycyjnym LSP Investments S. à r.l. z Luksemburga. Treść dokumentu nawiązuje do wskaźnika EBITDA za rok 2019, który wyniósł ponad 12 mln złotych. Osiągnięto więc docelowy poziom z Umowy Inwestycyjnej. Oznacza to, że nie będzie miała miejsca emisja nowych akcji przez Summa, ani zbycie przez LSP.

źródło: infostrefa.com ↗

Google dla polskich start-up'ów

Setki firm z Europy i Izraela zgłosiły się do akceleratora Google for Startups. Wybrano dziewięć, w tym aż trzy z Polski: Infermedica, Sundose i Yosh.AI. Wsparcie to udział w 3-miesięcznym programie obejmującym m.in. spotkania i warsztaty z inżynierami Google, dostęp do ich technologii, konsultacje z ekspertami zewnętrznymi. Przedsiębiorstwa będą rozwijać koncepty, które w różny sposób odpowiadają na wyzwania współczesnego świata.

źródło: fintek.pl ↗

COVID znaleziony w głosie

W ostatnich miesiącach na świecie prowadzone są prace nad możliwością wykrywania koronawirusa z nagrania audio z smartphona. Pracownicy MIT jako jedni z pierwszych podzielili się wynikami. Jak czytamy we wpisie, do wytrenowania modelu wykorzystano 4000 nagrań kaszlnięć, w tym około połowa osób z pozytywnym wynikiem COVID-19. W rezultacie algorytm wykazał 98,5% skuteczności potwierdzenia pacjentów zakażonych. Jednocześnie prawidłowo wskazał 100% osób zakażonych, lecz bez widocznych objawów.

źródło: news.mit.edu →

Higieniczny głos

Firma *Adobe* przedstawiła wyniki ankiety przeprowadzonej na 1000 użytkowników technologii głosowych w USA. Pytania dotyczyły doświadczeń i zachowań z aplikacjami wykorzystującymi mowę. Wykazano, że użytkownicy korzystają chętniej z technologii głosowej na więcej sposobów. Istotną kwestią okazała się również możliwość utrzymania restrykcji sanitarnych. 86% respondentów zgodziło się ze stwierdzeniem, że technologia głosowa może przyczynić się do bardziej higienicznego uczestnictwa w wydarzeniach lub wizyt w firmach.

źródło: venturebeat.com →

Popularne, ale nie doskonałe

Dokonano oceny współczynnika błędów (WER) dla trzech najpopularniejszych systemów automatycznego rozpoznawania mowy. Deklarowany WER na poziomie 2-3% odbiega od wyników przeprowadzonych badań. W zależności od domeny wyniósł od 14 do 23%. Autorzy badania wskazują na różnice w metodologii, bowiem do przetestowania systemów ASR podeszli bardziej holistycznie w porównaniu do powszechnie stosowanych metod weryfikacji. W swojej pracy wykorzystali nagrania autentycznych rozmów pomiędzy pracownikami a klientami z 50 call center.

źródło: venturebeat.com →

Kosztowny AI w Europie

Amazon w ciągu ostatniego roku rozbudował szereg usług działających w ramach ekosystemu *Alexa*. Nowości to między innymi: dokładniejsze rozumienie języka naturalnego i głębsze pozyskiwanie informacji z otoczenia w celu lepszej interakcji z użytkownikiem. Nowe kompetencje AI pozwalają wykorzystywać poszerzone dane w celu budowania nieetycznej przewagi firmy na rynku. W rezultacie Komisja Europejska oskarżyła *Amazon* o zdobycie nielegalnej przewagi na europejskim rynku, co grozi grzywną w wysokości do 10% rocznych światowych przychodów firmy.

źródło: aitrends.com →

Komputer lepszy w mowie spontanicznej

Naukowcom z *Karlsruhe Institute of Technology* (KIT) udało się opracować algorytm, który przewyższa umiejętności ludzi w rozpoznawaniu mowy spontanicznej w czasie rzeczywistym. System ten osiąga współczynnik błędów na poziomie 5%, czyli o 0,5% mniej niż człowiek. Badacze przyznają, że rozwój tej "umiejętności" dla AI ma szansę zapewnić lepszą interakcję głosową z maszynami w postaci swobodnej komunikacji czy precyzyjnego tłumaczenia symultanicznego.

źródło: kit.edu →

Snapchat bardziej głosowy?

Snap Inc., właściciel marki *Snapchat*, który jest jedną z najpopularniejszych platform społecznościowych, przejął izraelski startup *Voca.ai*. Nabyty podmiot sukcesywnie od kilku lat rozwija systemy voicebotów do call center. Dodatkowo w tym roku jest bardzo zaangażowany w wykrywanie koronawirusa w głosie. Nowy właściciel już wcześniej testował rozwiązania głosowe w *Snapchacie* z wykorzystaniem innych technologii. Zakup start-upa może wynikać z planu rozwijania tychże funkcji w *Snapchat* w oparciu o własny system.

źródło: techcrunch.com →

PUBLIKACJE NAUKOWE

Augmentation adversarial training for self-supervised speaker recognition

źródło: arxiv.org →

Automated accurate speech emotion recognition system using twine shuffle pattern and iterative neighborhood component analysis techniques

źródło: sciencedirect.com →

CopyPaste: An Augmentation Method for Speech Emotion Recognition

źródło: arxiv.org →

Group-level Speech Emotion Recognition Utilising Deep Spectrum Features

źródło: dl.acm.org →

Deep variational metric learning for transfer of expressivity in multispeaker text to Speech

źródło: inria.fr →

AudVowelConsNet: A Phoneme-Level Based Deep CNN Architecture for Clinical Depression Diagnosis

źródło: sciencedirect.com →

Transformer in action: a comparative study of transformer-based acoustic models for large scale speech recognition applications

źródło: arxiv.org →

Improving pronunciation assessment via ordinal regression with anchored reference samples

źródło: arxiv.org →

A Low footprint Automatic Speech Recognition System For Resource Constrained Edge Devices

źródło: dl.acm.org →

Speaker Anonymization for Personal Information Protection Using Voice Conversion Techniques

źródło: ieeexplore.ieee.org →

Communication with Desired Voice

źródło: ntt-review.jp →