



AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

**Najnowsze innowacje oparte o technologię
mowy - co to jest prawdziwa innowacja, w
jaki sposób jest ona realizowana w
przeglądzie omawianych innowacji**

Kacper Tondera, Piotr Marcowski

Nazwa wydziału: WIMIR

Miejsce i data prezentacji: Kraków, 26.01.2015

Innowacja – definicja: czy takie oczywiste?

- **nowość mająca zastosowanie w działalności gospodarczej, (Wikipedia),**
- **wprowadzenie czegoś nowego; też: rzecz nowo wprowadzona, (SJP PWN),**
- **zastosowanie nowej wiedzy w procesie produkcji, (wg Davida Begga),**
- **Wdrożenie nowego lub istotnie ulepszanego produktu (wyrobu lub usługi) lub procesu, nowej metody organizacyjnej lub nowej metody marketingowej w praktyce gospodarczej, organizacji miejsca pracy lub stosunkach z otoczeniem.**
- **praktyczny wynik zastosowania nauki czy wiedzy, może być postrzegany jako uczenie się z różnych źródeł. Jeśli dopiero nowy produkt zostanie wdrożony w praktyce to stanie się innowacją.**
- **„ze starego zrób nowe”,**

Rodzaje innowacji:

- **Produktowa** – nowość lub udoskonalenie wyrobów,
- **Procesowa** - zmiana w stosowanych przez przedsiębiorstwo metodach wytwarzania / metodach świadczenia usług,
- **Organizacyjna** - wprowadzanie nowej metody organizacji w biznesowych praktykach przedsiębiorstwa, organizacji miejsca pracy lub też w relacjach zewnętrznych.
- **Marketingowa** - zastosowanie nowej metody marketingowej obejmującej znaczące zmiany w wyglądzie produktu, jego opakowaniu, pozycjonowaniu, promocji, polityce cenowej lub modelu biznesowym, wynikającej z nowej strategii marketingowej przedsiębiorstwa

Vocelab

- Spółka z Trójmiasta, która powstała w wyniku konkursu Innowator organizowanego przez Fundację na rzecz Nauki Polskiej.
- Produkty firmy to m.in. VoiceBanking, VoiceInsurance, VoiceTelecom, VoiceMed, VoiceLegal, VoiceSecurity, VoiceAutomotive, VoiceInDevice.

VoiceLab



- pozwala klientom banku na wygodne i bezpieczne logowanie się do systemów bankowych,
- klient może zlecać przelewy głosem w bankowej aplikacji mobilnej,
- sprawdzać saldo na rachunku,
- wygodny i skuteczny sposób komunikacji z osobą niedowidzącą.



- Obejmują biometrię głosową = brak tradycyjnej autoryzacji,
- bezpieczniejsza od klasycznych, opartych często na danych z dowodu osobistego,
- wdrożone rozwiązania już w dwóch bankach w Polsce i negocjacje kolejnych wdrożeń.
- Głosowy proces weryfikacji i autoryzacji jest w pełni automatyczny = oszczędność czasu doradców klienta.



- za pomocą głosu kierowca uruchomi radio, zapyta o stan pojazdu, wyznaczy temperaturę w samochodzie, oraz wytyczy trasę w nawigacji.
- pojazd można zaprogramować i stworzyć profil własnych funkcji,
- odpowiednie ustawienie lusterek, optymalną wysokość fotela kierowcy, ulubiona temperatura w samochodzie, albo automatycznie uruchomiona ulubioną stacją radiową,

GrammarTrainer

GrammarTrainer jest innowacyjną metodą, wykorzystującą oprogramowanie komputerowe do nauki gramatyki przy pomocy kolorów i kształtów. Służy osobom, które z uwarunkowań zdrowotnych nie są w stanie nauczyć się gramatyki przy pomocy konwencjonalnych metod.

Z programu może korzystać osoba, która:

- Zna słownictwo na poziomie 2-latka
- Rozpoznaje te słowa na piśmie
- Jest w stanie obsługiwać komputer myszką lub klawiaturą
- Preferuje „wzrokowe” metody nauczania

Obecnie funkcjonuje jedynie program dla gramatyki języka angielskiego, jednak nic nie stoi na przeszkodzie dostosowania go do potrzeb innych języków.

Przykładowe zagadnienia pierwszej części programu:

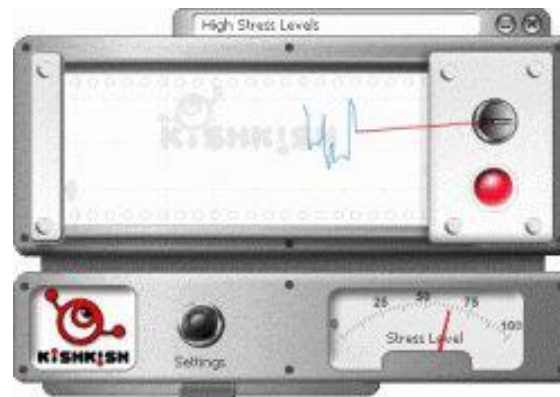
- How to form singular versus plural
- How to combine articles and adjectives with nouns
- The different forms of the verb *to be* or *to have* (*is* versus *are*; *has* versus *have*)
- Simple prepositional phrases (*on the square*, *in the circle*, etc.)
- How to use number words and the word *not* in sentences
- Complex adjectives like *bigger than* and *biggest*
- How to join phrases and sentences with *and*
- The pronouns *it* and *they*
- Possessive phrases, like *the triangle's dots* or *the square's stripes*
- Complex phrases like *more triangles than squares*
- Quantifiers like *many*, *most*, *few*, *all*, *some*, *none*, *only*, *except*, and *neither*
- How to *ask* a variety of questions, including *yes/no*, *what*, *where*, *which*, and *how many*.

Oprogramowanie dostosowane jest do ciągłej ewaluacji jej użytkownika, co pozwala na bieżąco monitorować postępy i w odpowiednim czasie wprowadzać kolejne tematy, a następnie łączyć je ze sobą w dłuższe wypowiedzi.



Plug-in: KishKish Lie Detector

- Plug-in, wtyczka do Skype'a,
- wykrywacz kłamstw działający na zasadzie analizy napięcia w głosie rozmówcy,
- używa technologii analizy napięcia głosowego VSA (Voice Stress Analysis) umożliwiającej pomiar i prezentację w czasie rzeczywistym poziomu napięcia u rozmówcy,



Ideę TALK wraz z urządzeniem wymyślił Arsh Shah Dilbagi - 16-letni uczeń z Indii. Nastolatek przygotował swój pomysł na konkurs Google Global Science Fair, w którym osoby od 13 do 18 roku życia mogą prezentować swoje projekty naukowe związane z programowaniem i elektroniką.

Urządzenie nazwane TALK może przywrócić mowę osobom chorym na stwardnienie boczne zanikowe, osobom z jakąkolwiek formą paraliżu, a także cierpiącym na zaburzenia mowy.



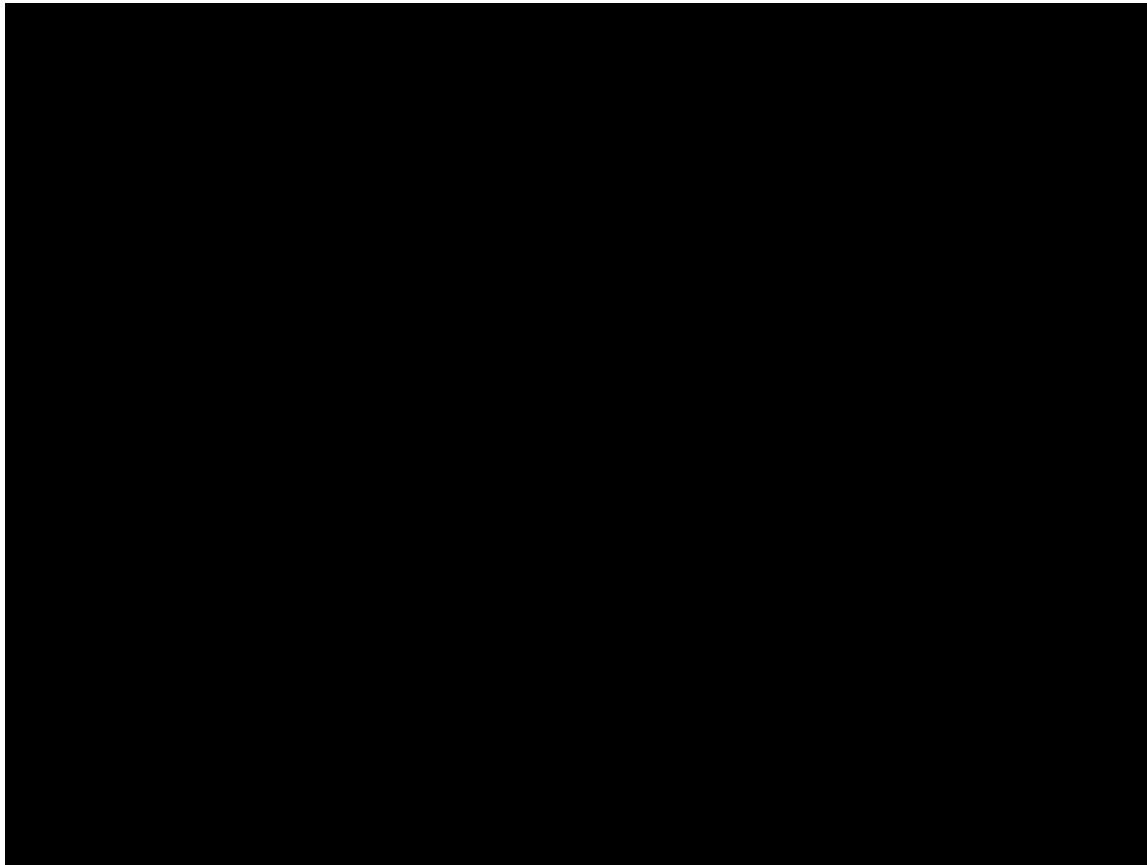
TALK za pomocą mikrofonu MEMS przekształca oddech w impulsy elektryczne, które odbierane są przez mikroprocesor i tłumaczone na słowa, które odczytuje syntezytor mowy. Dzięki temu, korzystając z alfabetu Morse'a, użytkownik może porozumieć się ze światem zewnętrznym tylko za pomocą oddechu.

Ogromną zaletą systemu jest niski koszt, samo urządzenie kosztuje zaledwie kilkadziesiąt dolarów, ponad stukrotnie mniej niż dotychczas stosowane metody o zbliżonym przeznaczeniu.





<https://www.youtube.com/watch?v=kwJ3uTXDpnU>



ZTE Star 2 - smartfon w całości sterowany mową

- Chińska firma ZTE to światowy lider innowacyjności, w 2011 i 2012 w rankingu World Intellectual Property Organization w zestawieniu liczby złożonych wniosków patentowych uplasował się na pierwszym miejscu, zaś w 2013 na drugim,
- W ubiegłym roku na największych targach elektroniki użytkowej i nowych technologii CES przedstawił m. in. Nowego smartfona: ZTE Star 2, w całości sterowanego mową ludzką.



ZTE Star 2 – smartfon w całości sterowany mową

- Stworzony w oparciu o ponad 1000 nowych patentów, z czego 158 związanych jest ze sterowaniem głosem,
- czas reakcji sprzętu na polecenia wynosi maksymalnie 1,2 sekundy, a system zapewnia skuteczność na poziomie 90% nawet przy otaczającym hałasie,
- Do jego działania nie jest wymagane połączenie z internetem,
- Za pomocą odpowiednich komend można w nim po prostu włączać aplikacje i przełączać się między nimi, fotografować, robić notatki czy obsługiwać odtwarzacz muzyki, ale też odblokować telefon niepowtarzalnym głosem użytkownika.

Rozwój technologii syntezy mowy (Amazon)

Amazon otwiera w Gdańsku centrum rozwoju technologii. Rekrutacja specjalistów z branży IT już ruszyła. Tym razem Amazon chce zostać liderem rynku syntezy mowy. Nad czym dokładnie pracować będą ci specjaliści nie wiadomo, gdyż objęte jest to ścisłą tajemnicą. Wiadomo natomiast, że w gdańskim centrum Amazona pracują już osoby związane z najnowszym produktem spółki, czyli Amazon Echo. Najnowszy gadżet firmy to domowy asystent, któremu zadać można normalne pytania, np. o kursy walut, temperaturę, czy rozkład jazdy autobusu, a na wszystkie powinniśmy uzyskać odpowiedź. W Gdańsku Amazon chce pracować na podobnych technologiach.

Ekspansja Amazona na rynek syntezy mowy rozpoczęła się od przejęcia w 2013 roku gdyńskiej spółki Ivona Software.

ARM1 – projekt realizowany przez Poznańskie Centrum Superkomputerowo-Sieciowe

Zaawansowany system automatycznego rozpoznawania i przetwarzania mowy polskiej na tekst.

Dedykowany dla służb odpowiedzialnych za bezpieczeństwo państwa.

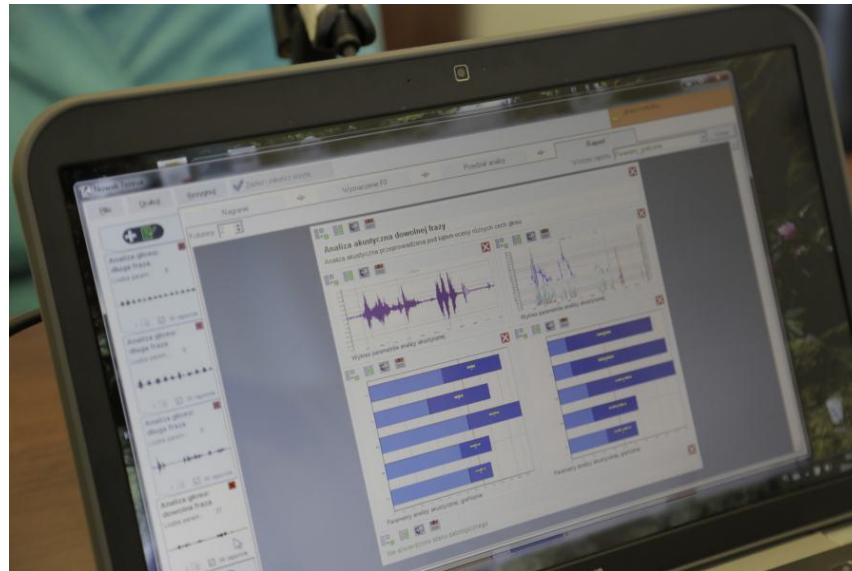
System będzie wyposażony w szereg funkcji, które umożliwią użytkownikom końcowym automatyczne zapisywanie głosowych wypowiedzi w postaci tekstu, przy czym rozpoznawanie mowy będzie niezależne od głosu mówcy oraz stosowanego słownictwa. Możliwe będzie przetwarzanie mowy na tekst z nagrań dokonanych wcześniej w różnych warunkach akustycznych. Kolejną funkcją systemu będzie przeszukiwanie dużych baz nagrań, pod kątem wskazanych fraz słownych, co umożliwi efektywne korzystanie z tego typu zasobów.

Działanie systemu będzie ukierunkowane na ostatecznych odbiorców projektu tj. służby odpowiedzialne za bezpieczeństwo państwa, tak pod względem wymienionych funkcji, jak i charakterystycznego słownictwa oraz kontekstu wypowiedzi, co pozwoli na zwiększenie trafności procesu rozpoznawania.

Data zakończenia prac: 2016-04-25.

Co jest w planach? Sfera badań i testów

- we Wrocławiu lekarze z Uniwersyteckiego Szpitala Klinicznego (USK) we Wrocławiu z pomocą inżynierów z Politechniki Wrocławskiej pracują nad prototypem urządzenia, które non stop przez 24 godziny będzie monitorować mowę pacjentów cierpiących na Alzheimera.



Co jest w planach? Sfera badań i testów

- wstępne wyniki badań wskazują, że analiza głosu może być nie tylko obiektywnym wskaźnikiem początku choroby, ale też wykładnikiem postępów jej leczenia,
- dzięki wrocławskim badaniom Alzheimerera będzie można rozpoznać nawet o kilka lat wcześniej niż teraz i od razu zacząć leczenie, które opóźni nadejście dolegliwości związanych z tą chorobą.
- Badanie polega na rejestracji wypowiedzianych samogłosek przez pacjenta konkretną ilość razy, a następnie także rejestracji samogłosek dowolną ilość razy,
- Próbki zostają poddane analizie, która ma za zadanie wy badać neurodegenerację charakteryzującą właśnie osoby cierpiące na Alzheimerera.

DZIĘKUJEMY ZA UWAGĘ!