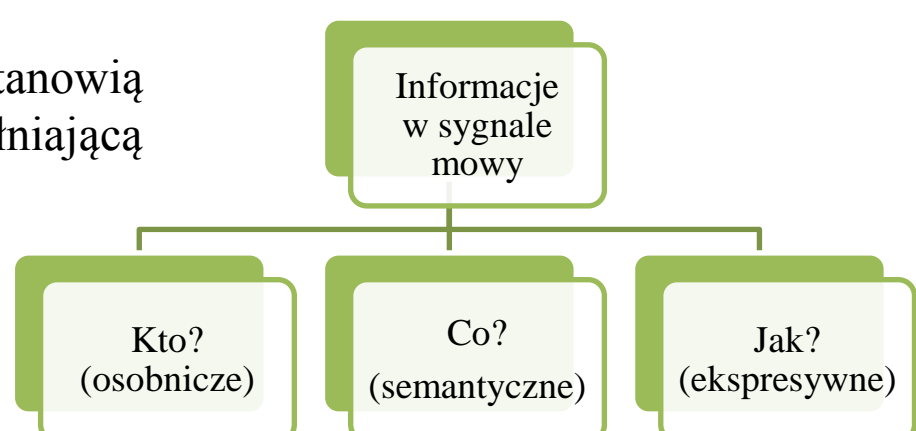


W ramach prac AGH powstała baza danych nagrań mowy emocjonalnej (emocje odgrywane) zgromadzona w celu badań nad zawartością afektywną sygnału mowy.

## MOTYWACJA

Emocje wyrażane w mowie stanowią ważną część komunikatu, uzupełniając komunikat słowny:



Rozwój systemów technologii mowy realizujących detekcję emocji wymaga odpowiednich zasobów nagrań. Jednym z największych wyzwań stojących przed analizą mowy emocjonalnej jest posiadanie odpowiednio dużej i różnorodnej pod względem treści i mówców bazy nagrań mowy emocjonalnej;

Emocje podstawowe mogą być scharakteryzowane w sposób uniwersalny przez jednoznaczny wzorzec lub konfigurację parametrów akustycznych [Scherer]

## PARAMETRY AUDIO

częstotliwość próbkowania: **44,1 kHz**  
rozdzielczość: **16 bit**  
ścieżki: **mono**  
sygnał/szum: **SNR > 25 dB**



Rejestrator dźwięku Zoom H4N



Mikrofon pojemnościowy AKG C5 VOCAL

## STRUKTURA I ROZMIAR BAZY

- podział na części tematyczne (osobno: zdania, teksty ciągłe, cyfry, polecenia) znajdujące się w osobnych plikach
- warstwa metadanych: każde nagranie jest opisane przy pomocy akronimu mówcy, oznaczenia jego płci oraz zawartości nagranej treści
  - plik audio (w formacie .wav)
  - łączna wielkość zarchiwizowanych danych wszystkich mówców to **1.5 GB**.

## DODATKOWE ASPEKTY

### PROBLEMY

- nieautentyczność emocji
- interkorelacje rzeczywistych emocji
- indywidualizm ekspresji i percepcji emocji

### ZALETY

- Dobra jakość nagrania
- Możliwość uzyskania gamy emocji dla dowolnej treści
- Możliwość dokładnego zaplanowania struktury korpusu
- Brak wątpliwości etycznych

## MÓWCY

- 6 kobiet, 6 mężczyzn w wieku 20-30 lat
- aktorzy, studenci aktorstwa, amatorzy

## TREŚĆ

### ZDANIA

- Dzień dobry.
- Witam serdecznie.
- Miło cię widzieć.
- Dlaczego nie?
- Muszę z tobą porozmawiać.
- OK. Do zobaczenia!
- Pomóż mi.
- Wzajemnie.
- Nie chcę.
- U mnie super.
- Święty pomysł!
- Chyba Ci się nudzi.
- Bardzo się cieszę.
- On jest najlepszy na świecie.
- Czy naprawdę tak myślisz?
- Która jest teraz godzina?
- Przecież ty nic nie wiesz.
- To był żart.
- No oczywiście.
- Dziękuję, rozumiem.
- Mam dość.
- Skąd mnie znasz?
- Myszę, że tak.
- Masz rację.
- Na pewno tak jest.
- Dziś jest bardzo ładna pogoda.
- To bardzo dobrze.
- Proszę bardzo.
- Nic nie rozumiesz.
- Mam rację.
- Przepraszam cię.
- Ależ skąd!
- A jednak.
- Dziękuję za rozmowę.
- Wcale nie trzeba.
- Nie, dziękuję.
- Nie wiem.
- Nie trzeba.
- Nie denerwuj się tak.
- Pokaż, co potrafisz.
- O czym Ty mówisz?
- Dziękuję za pomoc.
- Nieźle.
- Uśmiechnij się.
- Idę do domu.
- Milego dnia!

### POLECENIA

Nowy, Otwórz, Usui, Prawo, Lewo, Góra, Dół, Cofnij, Zapisz, Start, Stop, OK

### CYFRY

Jeden, Dwa, Trzy, Cztery, Pięć, Sześć, Siedem, Osiem, Dziewięć, Zero

### TEKST

Gabinet zebrał się na oczekiwany z wielkim zainteresowaniem nadzwyczajnym posiedzeniu. Równolegle do zabiegów premiera, prezydent prowadzi nieoficjalne konsultacje w sprawie wyjścia z sytuacji patowej, w jakiej znalazła się nie tylko gospodarka i finanse, ale również polityka. Nie będzie zapowiadanych wcześniej inicjatyw, jak jednorazowe opodatkowanie wszystkich kont bankowych bez względu na ich stan, ani przywrócenia podatku od nieruchomości, nieobowiązującego mieszkań, w których zameldowany jest sam właściciel.



12 mówców x 5 min mowy x 6 emocji

## ZASTOSOWANIA

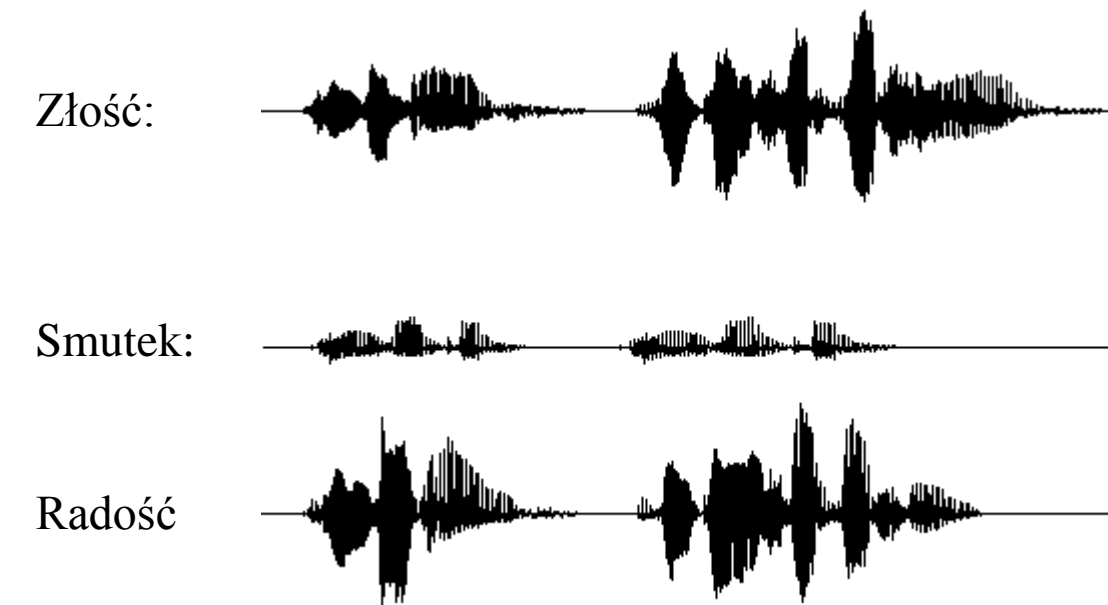
- afektywne interfejsy człowiek – komputer
- synteza mowy emocjonalnej (np. dla wirtualnych awatarów)
- normalizacja mowy pod kątem zawartości emocjonalnej
- badania nad ludzką percepcją emocji w mowie
- analiza wpływu cech głosu ludzkiego na efektywność rozpoznawania mowy
- zidentyfikowanie wzorców reakcji emocjonalnych w sygnale mowy i zweryfikowanie możliwej jakości automatycznego wnioskowania na temat nastawienia emocjonalnego mówcy (cechy wysokopoziomowe mowy)
- w dziedzinach medycznych: wspomaganie diagnozowania stanu chorych bądź skuteczności leczenia w przypadkach zaburzeń psychicznych i neurologicznych (autyzm, schizofrenia, ADHD, choroba afektywna dwubiegunowa).
- aplikacje komercyjne, np. wspomaganie pracy call centers celem podwyższenia komfortu użytkownika aplikacji i efektywności interfejsów głosowych

## DOSTĘPNOŚĆ W RAMACH LICENCJI

- Przetwarzanie:** Zgoda mówców na przetwarzanie nagrań w systemach informatycznych technologii mowy
- Prezentowanie publiczne:** Zgoda na anonimowe (bez podawania imienia i nazwiska nagranej osoby) odtwarzanie na konferencjach, wykładach i prezentacjach systemów technologii mowy
- Komercjalizacja:** możliwość sprzedaży przez AGH innym uczelniom i firmom do wykorzystania w ich systemach informatycznych i pracach badawczych

## PRZYKŁAD

- Zdania: 'Dzień dobry. Witam serdecznie' - przebieg czasowy sygnału akustycznego:



## PRZYKŁADY ANALIZ

Zdanie: 'O czym ty mówisz' - ten sam mówca (mężczyzna)

- a) Sygnał w dziedzinie czasu
- b) Spektrogram wraz z przebiegiem parametrów:

