

IS	105
Gruppe 3 NoName	ICA06

Gruppemedlemmer:

Ali Al Musawi
 Tor Borgen
 Ann Margrethe Ly Pedersen
 Brage Fosso
 Adrian Lorentzen
 Arne Bastian Wiik
 Morten Schibbye

Alle kode henvisninger ligger i README.md
<https://github.com/GB-Noname/is105-ica06>

Hva skal programmet vårt gjøre?

Programmet vårt skal kunne gjøre tekst om til tale og tale om til tekst

Refleksjon:

Alternativ løsning til eks.1:

Hvordan hadde dere gått frem for å implementere et system for gjenkjennelse av språk?

Enten gi mulighet for å laste opp to bilder som vises parallelt, slik at bruker kan sammenligne de to spektrene. Eller så skal programmet hente inn to tekstfiler som er outputen til Praat, slik at en kan sammenligne ulike verdier, som for eksempel pitch, volum osv.

For å kjenne igjen språk kan vi analysere selve språket ved å passe lydfilen fra praat inn i google-speech for så å sende den videre inn til Google translate (API) som kan kjenne igjen hvilket språk som snakkes. For eksempel å analysere hele setningen med "detect language" aktivert, for så å bruke retur JSON fra API for å hente antagelige språket.

Man kan også analysere selve språket ved å passe lydfilen fra praat inn i google-speech for så å sende den videre til google.translate og gjøre en basis språk analyse, for eksempel bruke vokaler, gjentatte bokstaver etc. for å detektere hvilket språk det er. I et program ville dette vært verdier vi sjekker teksten opp mot.

Løsning til Eks.2

Her har vi portere espeakbox over til vårt eksisterende program fra ICA05, ved å bruke text input fra HTML form vil vi sette stringen til et variabel for så å parse den inn i JSON request stringen.

Deretter skrive responsen til HTTP og spille av filen.

Her har vi eksisterende kode som er veldig tilsvarende det vi skal lage, så vi trenger kun noen nye metoder og parametere i eksisterende kode.

Alternativ løsning til Eks.3

Vi prøvde å kjøre et program lokalt som real time tar opp det vi sier gjennom mikrofon og sender til google Speech cloud som da skulle returnerer teksten og vise den i kommandolinjen. Men etter å ha eksperimentert rundt dette så kom vi fram at det ikke går (tar mange prosesser) da mikrofon/høytaler ikke ser ut til å kunne aksesseres på de virtuelle serverne.

Istedenfor, så brukte vi et Web Speech Google API som passet bedre med strukturen vi har på serveren.

Hvilke byggeklosser skal programmet bestå av?

- GB-NoName API site
- Espeakbox
- Web Speech API (<https://developers.google.com/web/updates/2013/01/Voice-Driven-Web-Apps-Introduction-to-the-Web-Speech-API>)
- Implementere Espeakbox og Web Speech API i GB-NoName API site (fra ICA05).