

# Text to Speech

Brian Arnesto Sitorus, Filip Jatelnicki,  
Sudarshna Patel, Tushar Pal



**Universiteit  
Leiden**  
The Netherlands

# Problem Formulation

- Experimenting with multiple architectures of neural networks (pretrained and own models)
- Make a report on the process as well as results
- We will start with English language audio dataset from Mozilla and adapt it to another language from the same dataset



# Challenges

- Scaling up training process to the vast dataset(s).
- Finding and adapting pretrained models for the task
- Capturing all required metrics for a good comparison of models



# Goal/ End Products

- Best model that takes speech recording and translates it to text
- Report of comparisons between the models, as well as the architecture descriptions and other details



# Task Division

	Brian	Filip	Sudarshna	Tushar
Literature research	yes			
Academic Writing	yes	yes	yes	yes
Presentation		yes		yes
Engineering	yes	yes	yes	yes
Communication	yes			

# Thank You



Universiteit  
Leiden  
The Netherlands