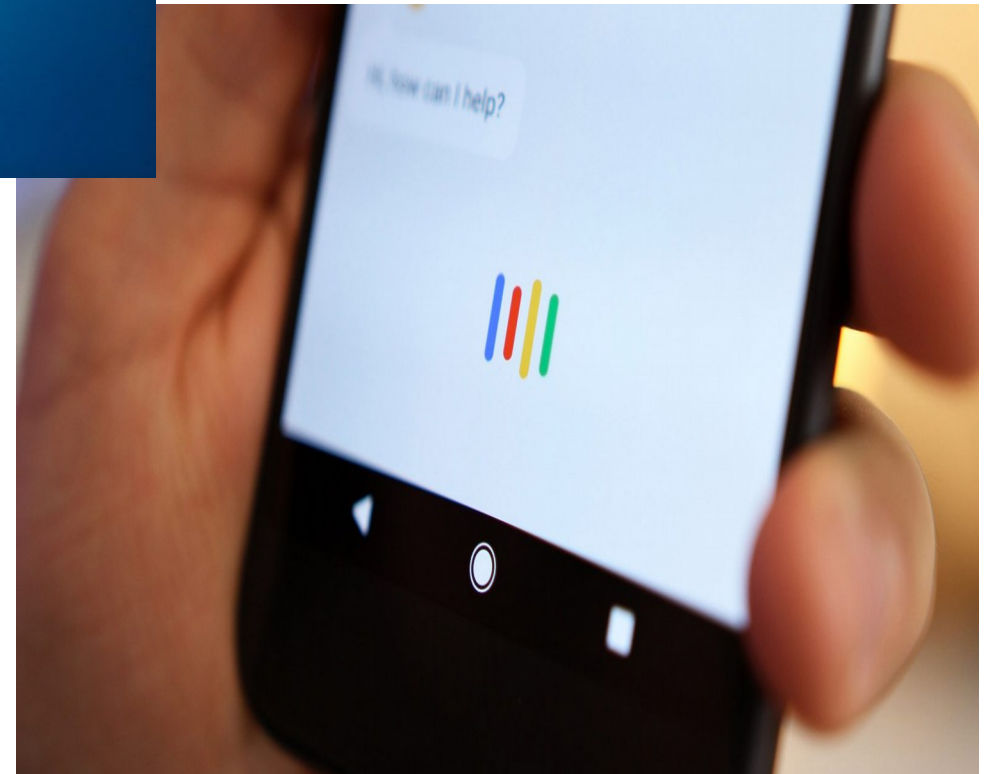
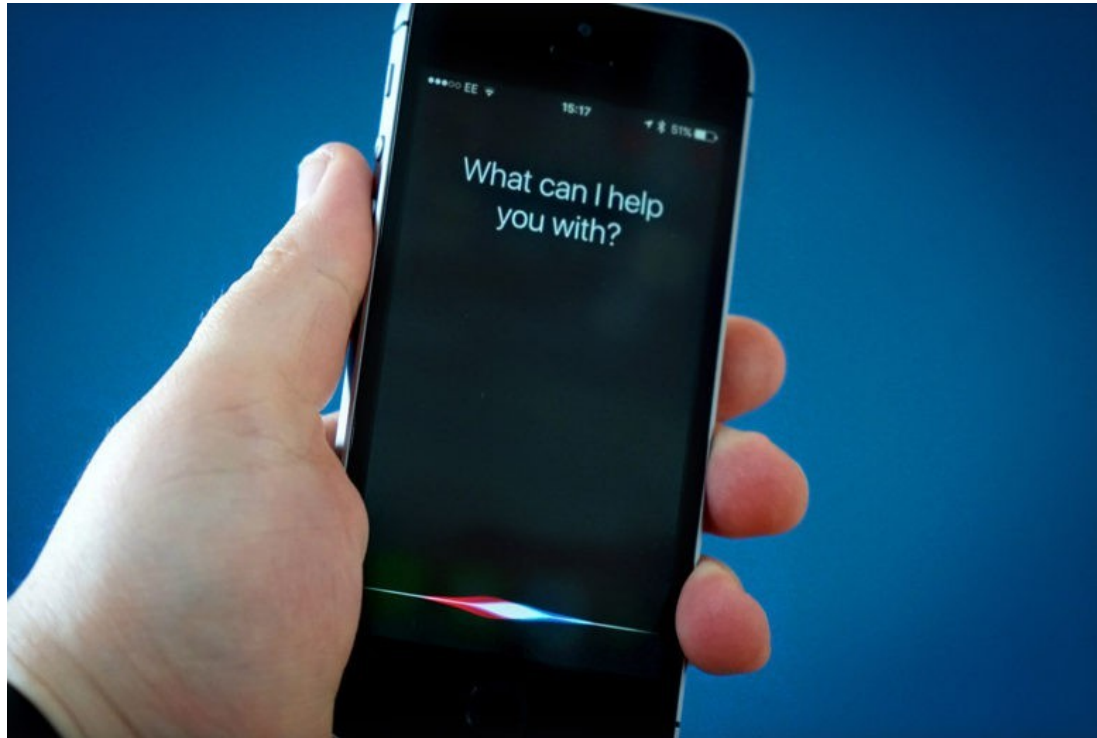


Text-to-Speech:

A gentle introduction

Josh Meyer + Fran Tyers
@_josh_meyer_ + @ezesanlasai





1



Переводчик

Отключить моментальный перевод



киргизский английский русский Определить язык ▼



английский киргизский русский ▼

Перевести

yolo



4/5000

Yolo





1



Переводчик

Отключить моментальный перевод



киргизский английский русский Определить язык ▼



английский киргизский русский ▼

Перевести

yolo



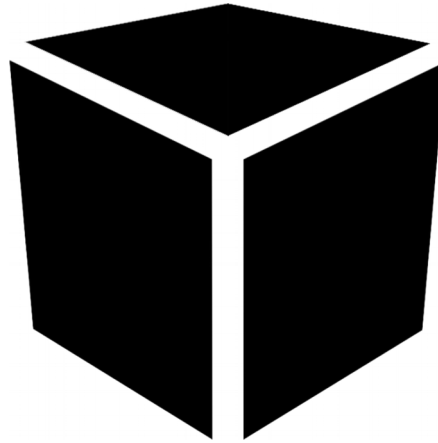
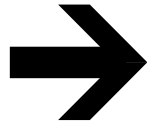
4/5000

Yolo



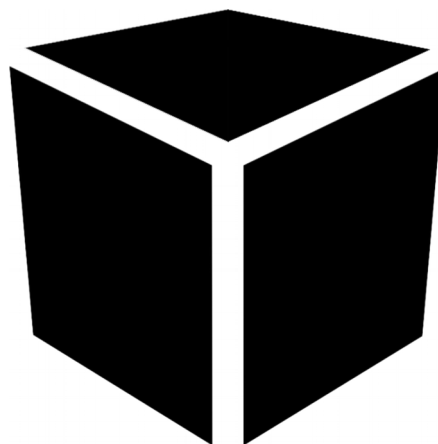


Moscow always has
traffic jams.



RAW TEXT

Moscow always has
traffic jams.



RAW AUDIO



RAW TEXT

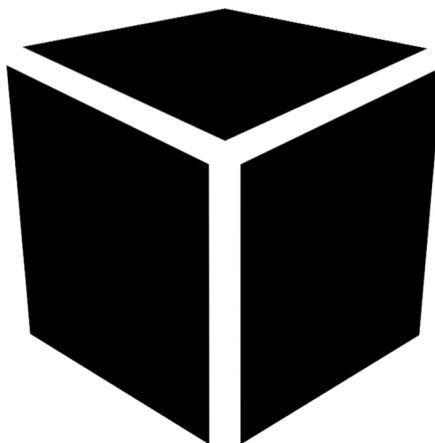


**SPEECH
SYNTHESIZER**

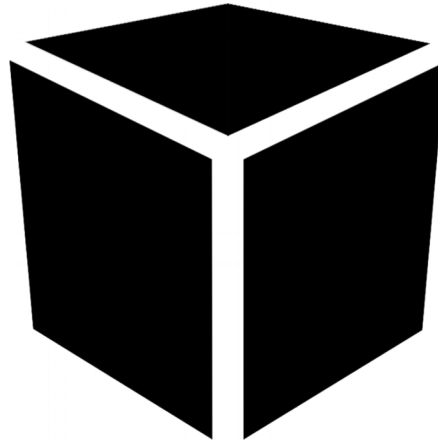


RAW AUDIO

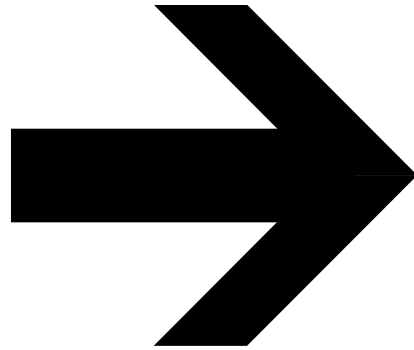
Moscow always has
traffic jams.



Moscow always has
traffic jams.



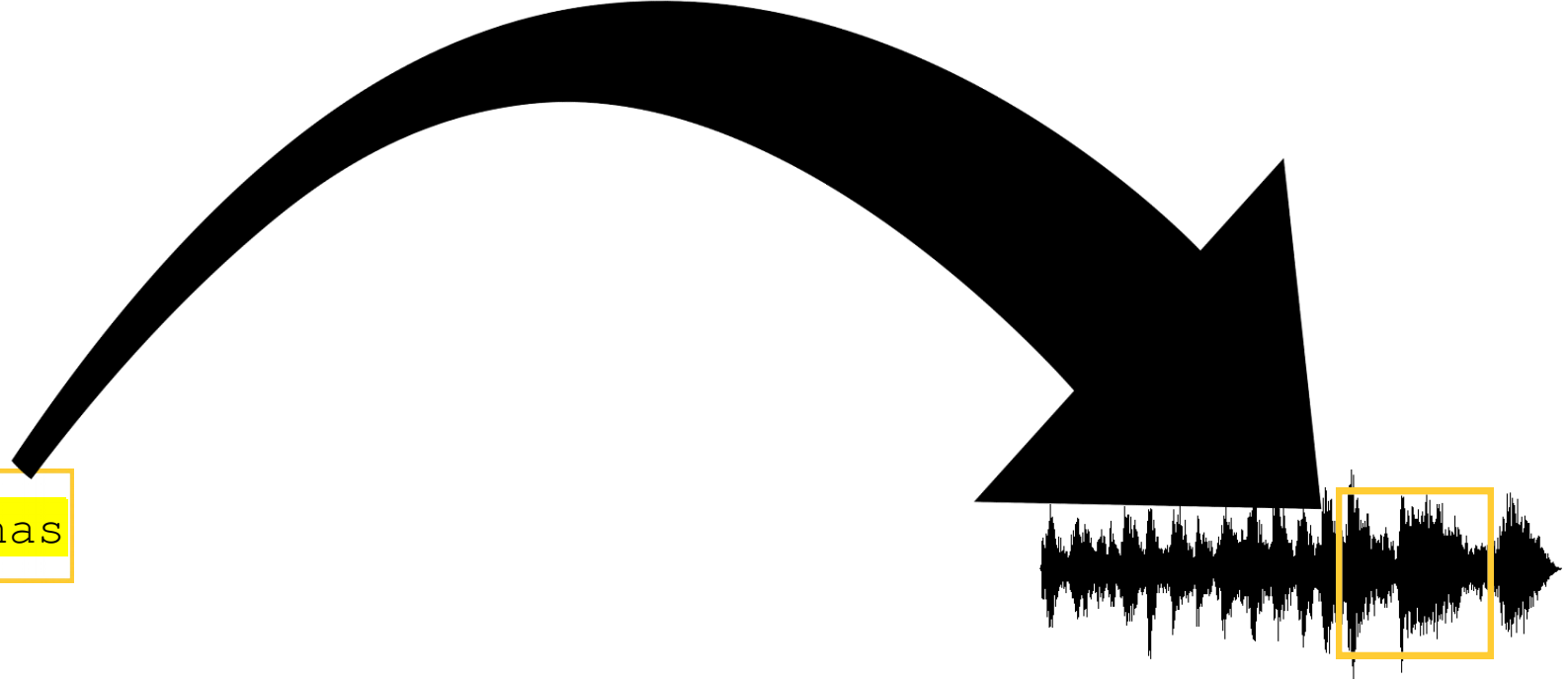
Moscow always has
traffic jams.



HARD

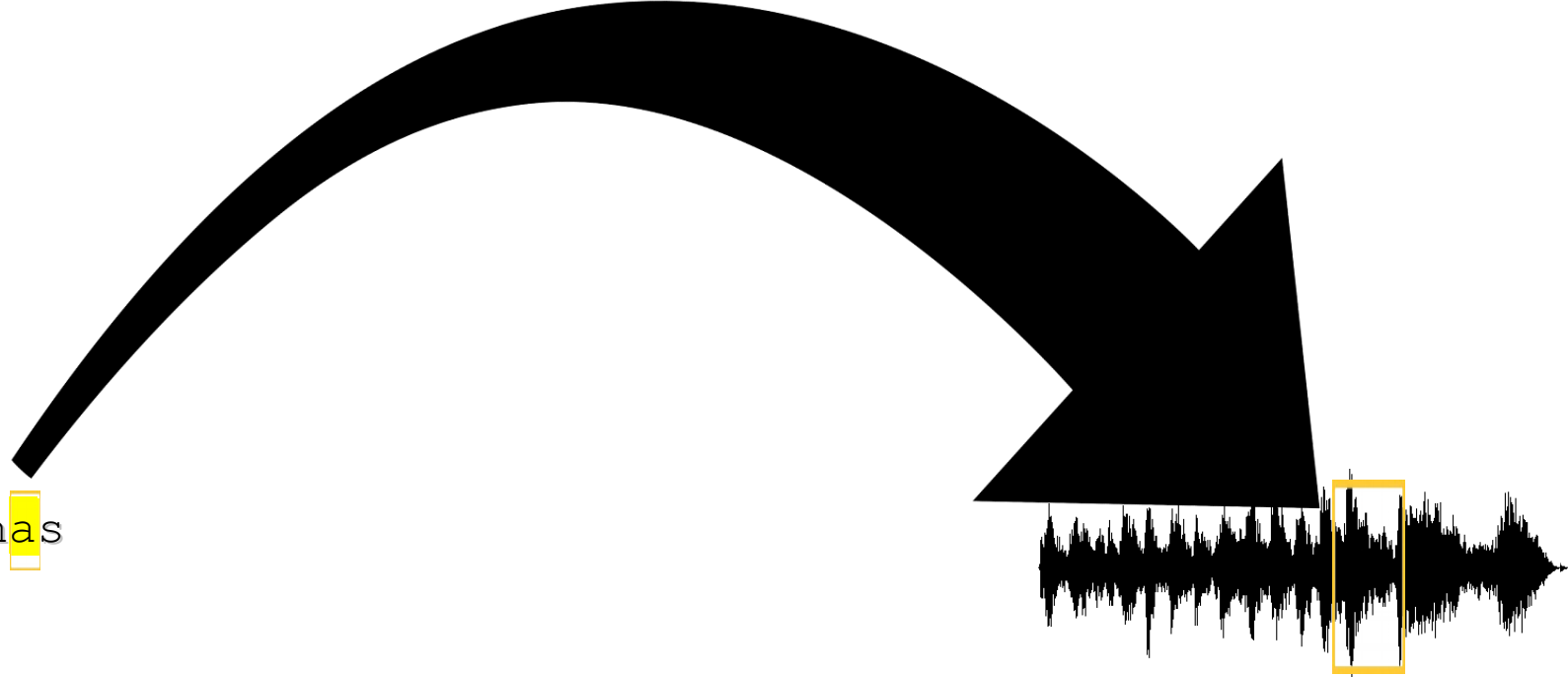


Moscow always has
traffic jams.



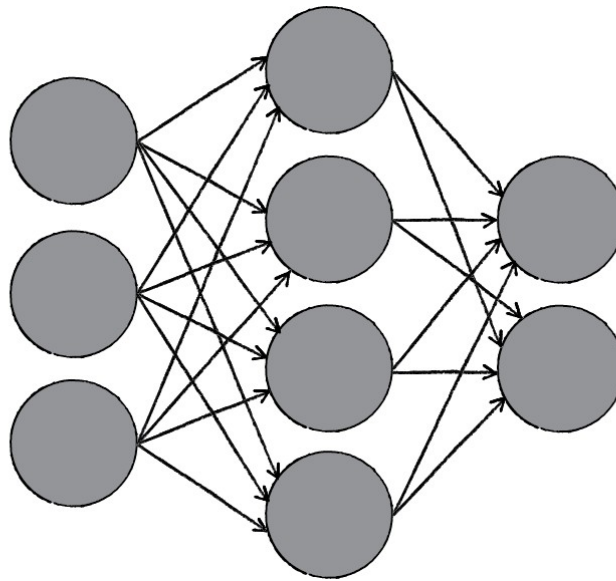
HARD

Moscow always has
traffic jams.



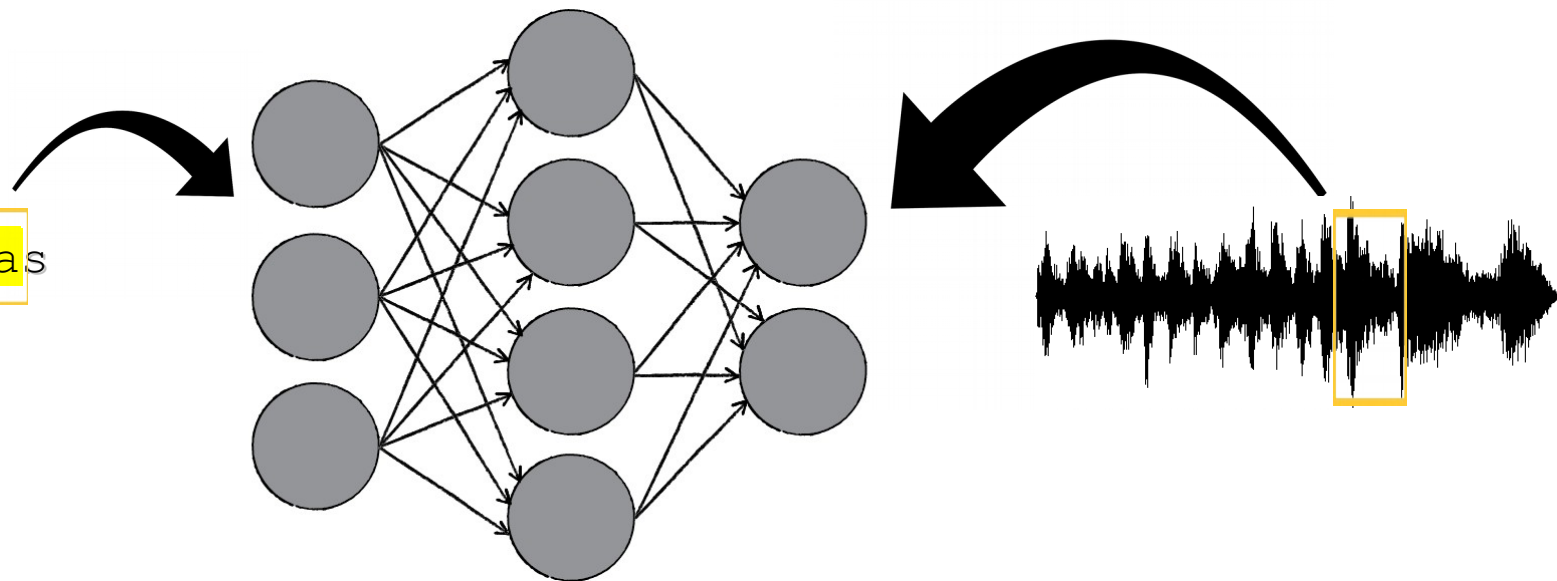
EASIER

Moscow always has
traffic jams.



Neural
Networks

Moscow always has
traffic jams.

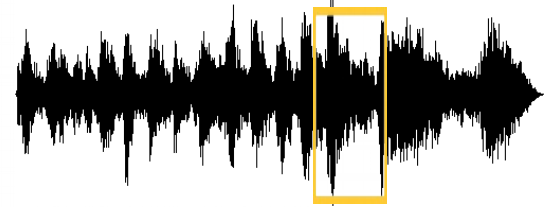
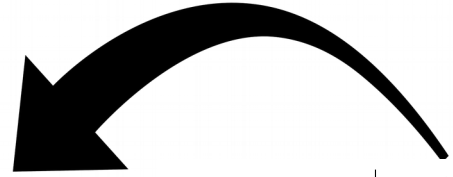


Moscow always has
traffic jams.

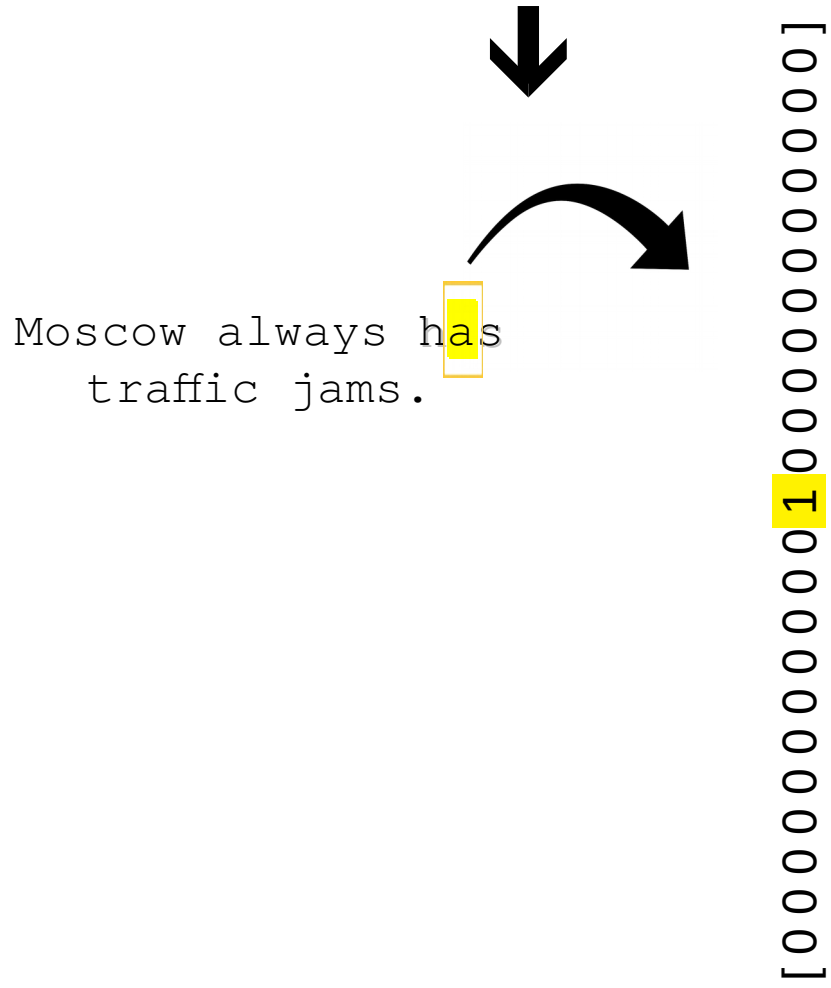


[00000000000010000000000000]

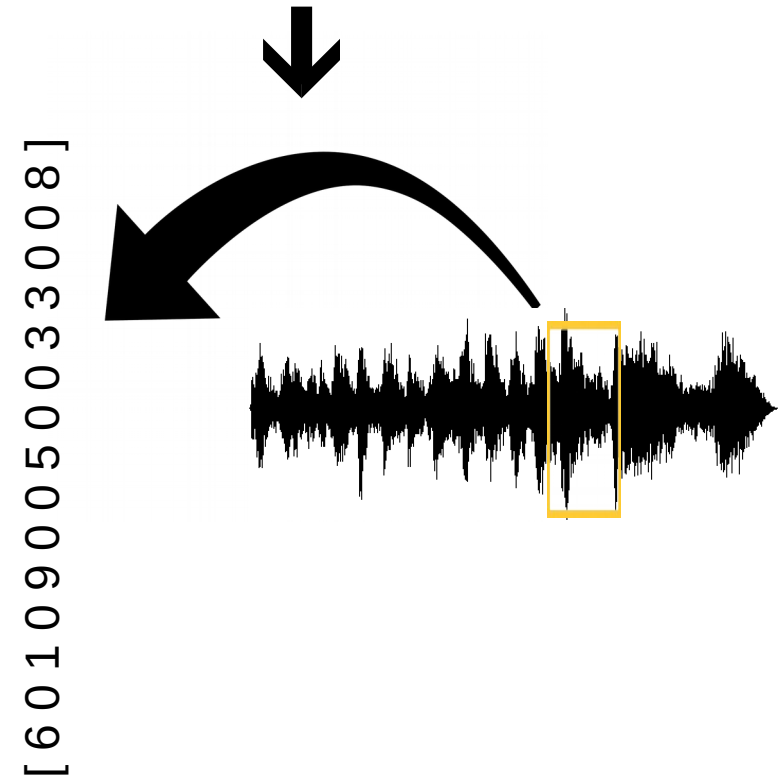
[601090050033008]



TEXT FEATURE EXTRACTION



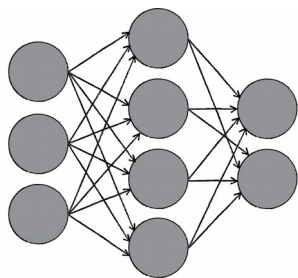
AUDIO FEATURE EXTRACTION



Moscow always has
traffic jams.

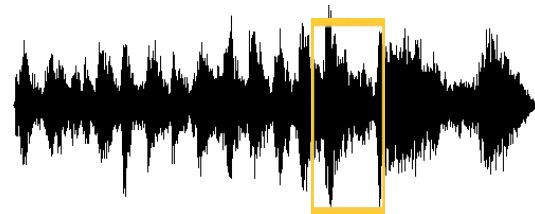
INPUT

[00000000000010000000000000]



OUTPUT

[601090050033008]



Text Feature Extraction

Moscow always has
traffic jams.

Moscow always has traffic jams .



WORDS

Moscow always has
traffic jams.

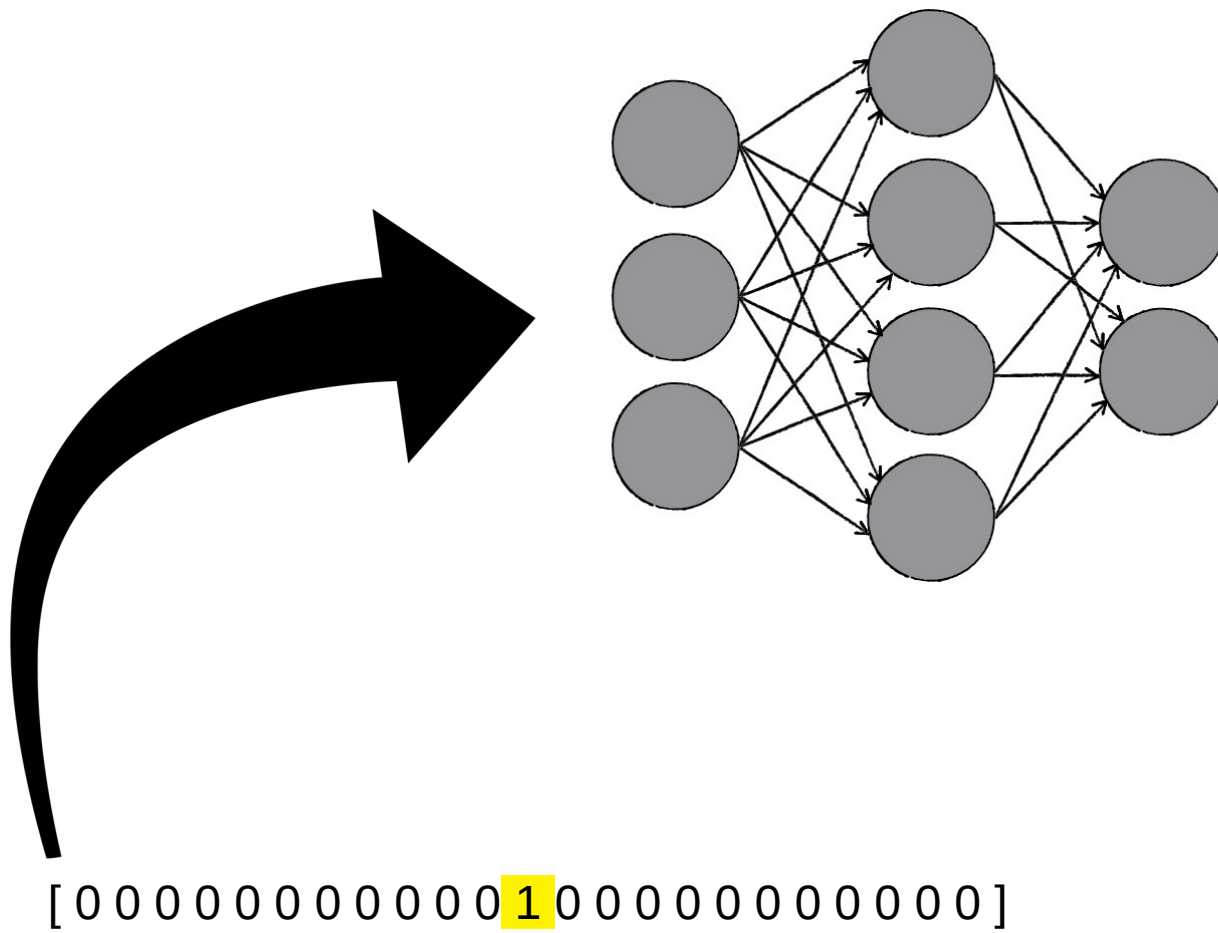
M o s c o w a l w a y s h a s t r a f f i c j a m s .

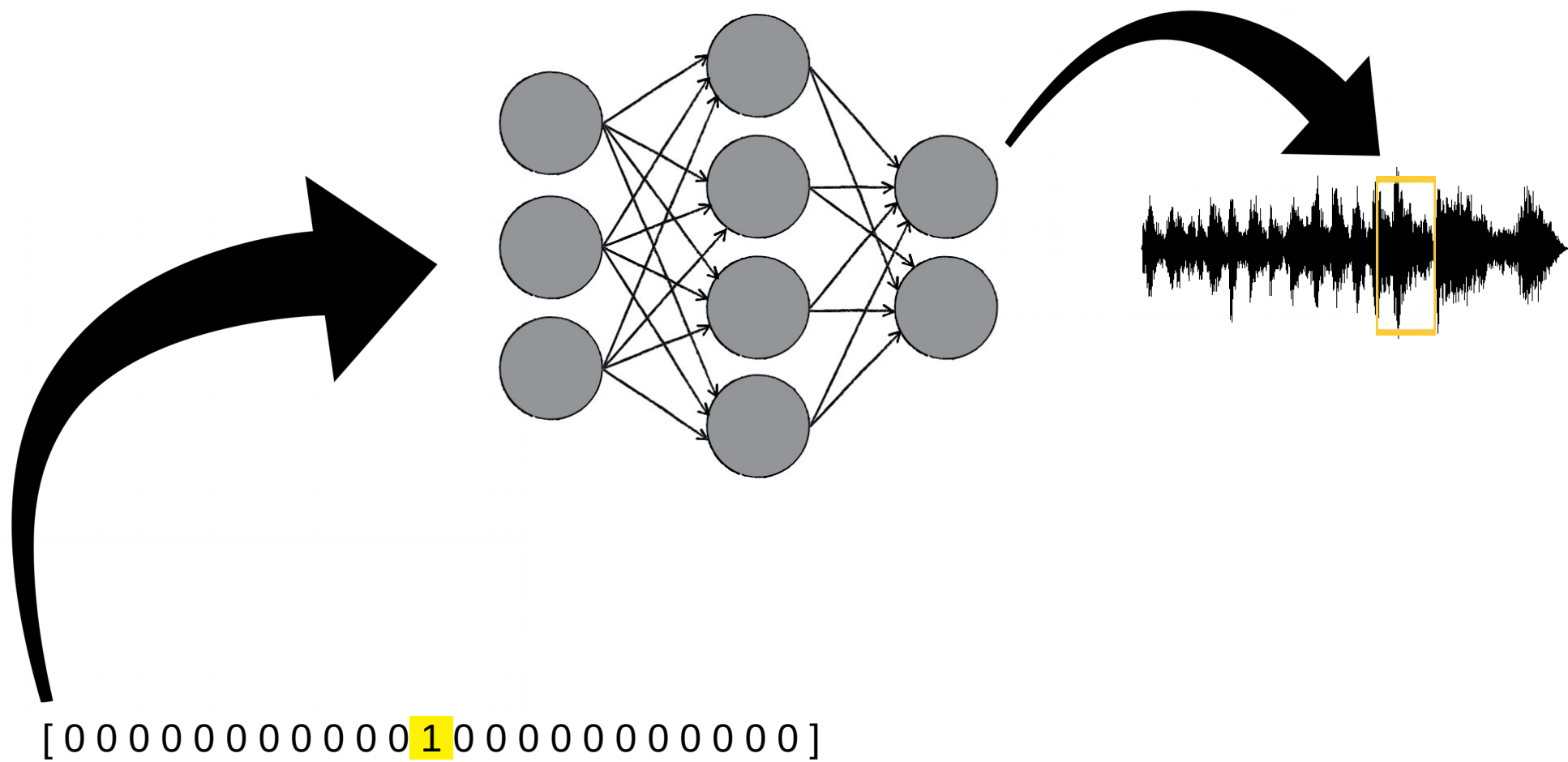


LETTERS

Moscow always has
traffic jams.

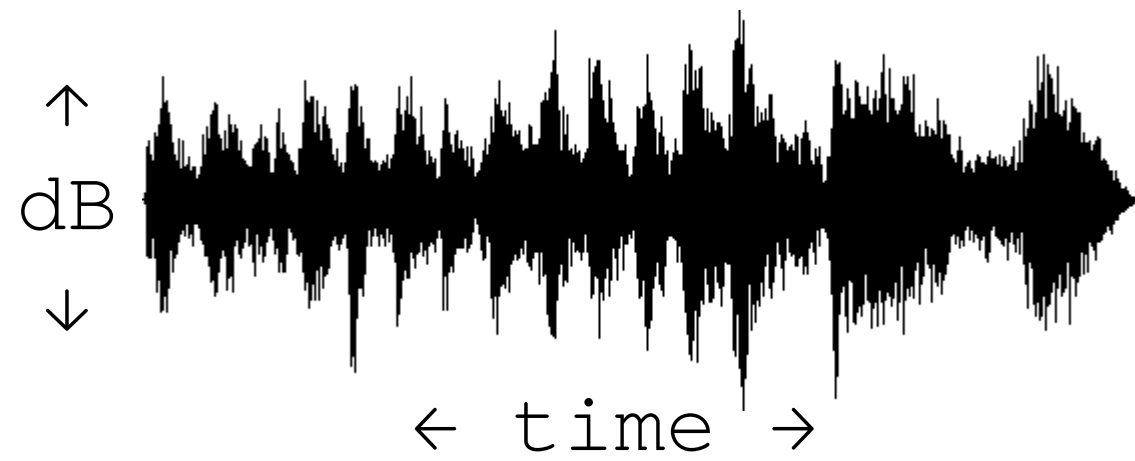
[0000000000000010000000000000]





Audio Feature Extraction





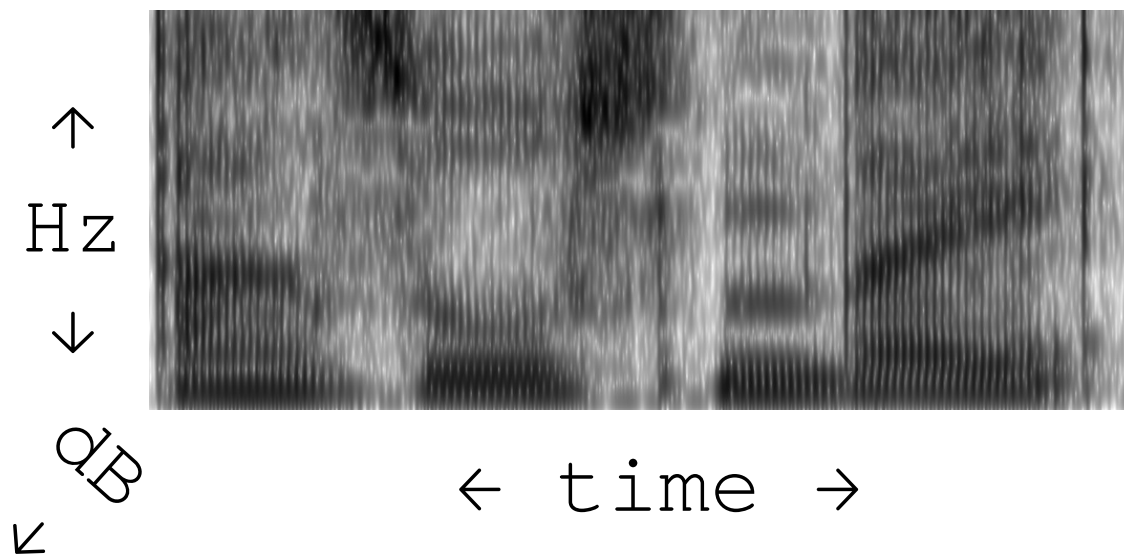
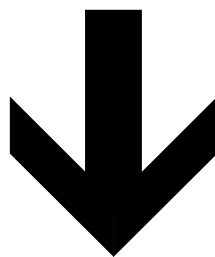


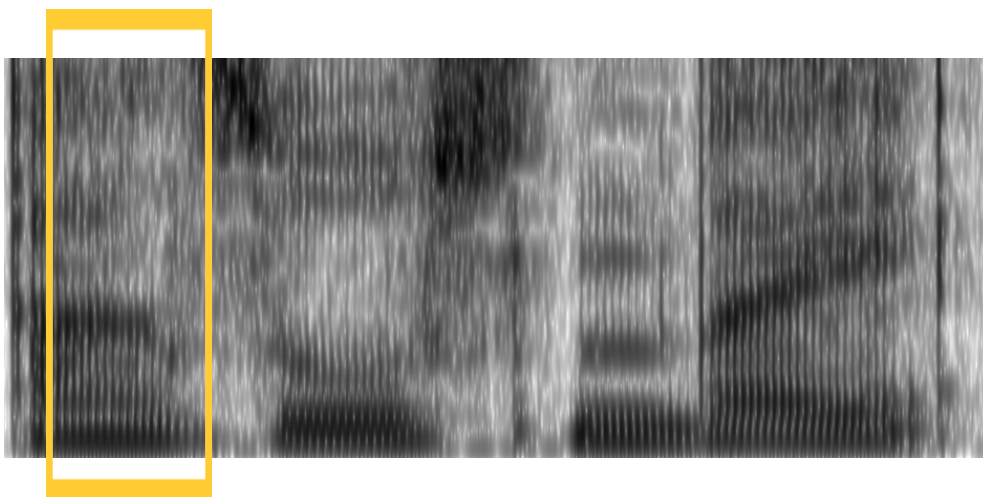
↑
dB
↓

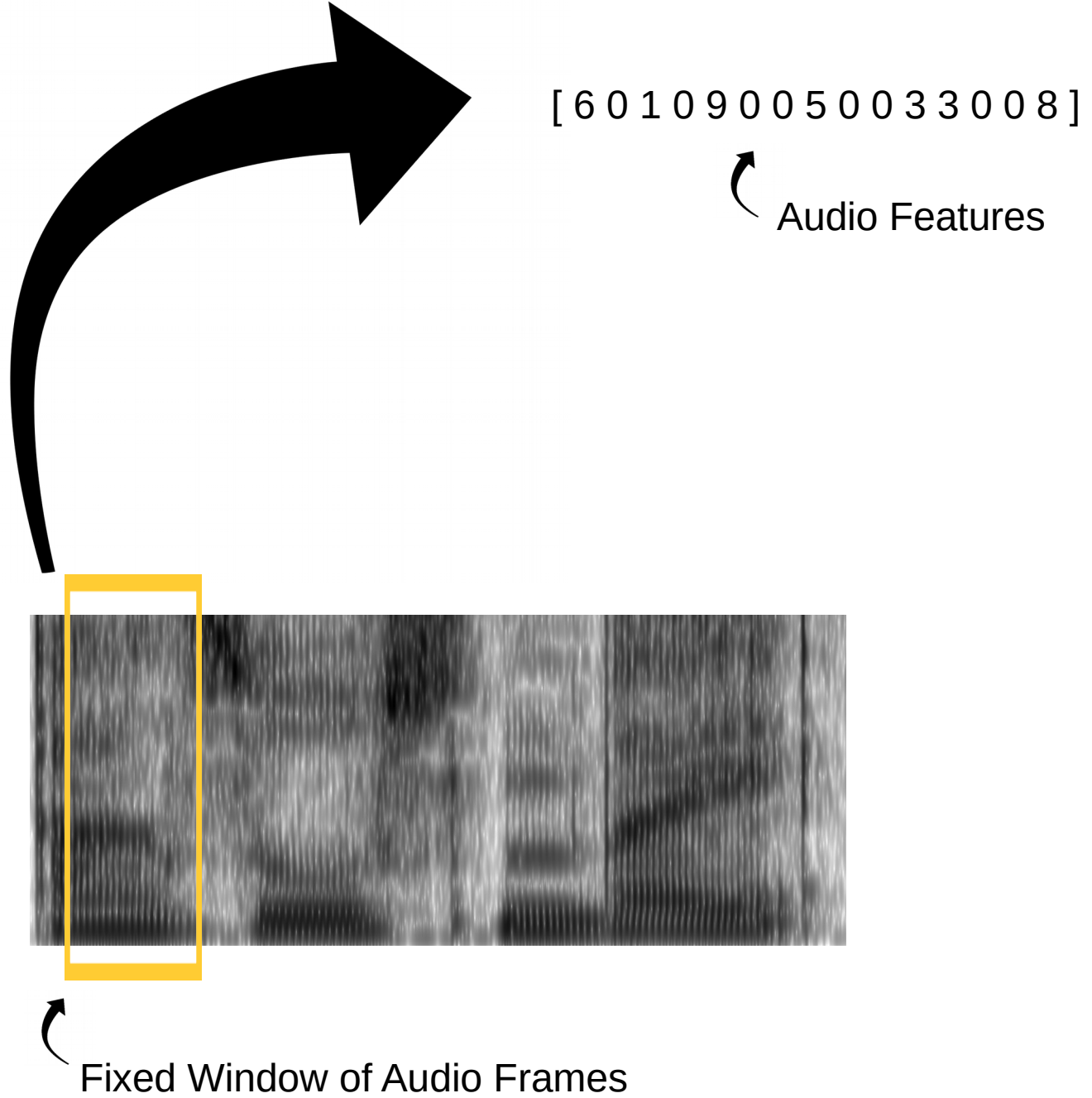
↑
Hz
↓

↙ Q/B ↘

← time →





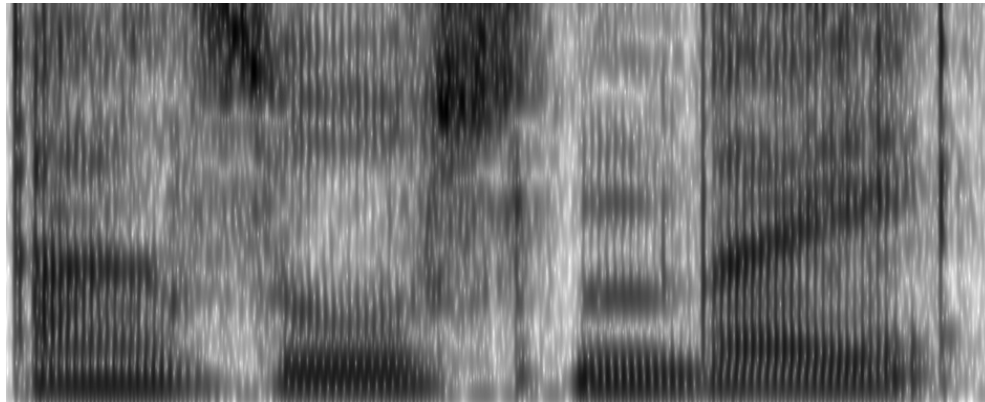


AN EXAMPLE

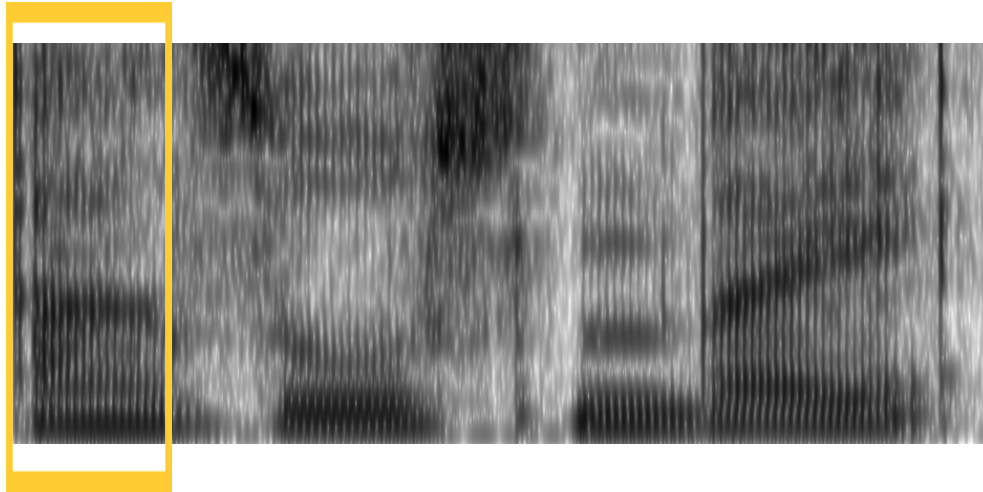
Moscow always has traffic jams.



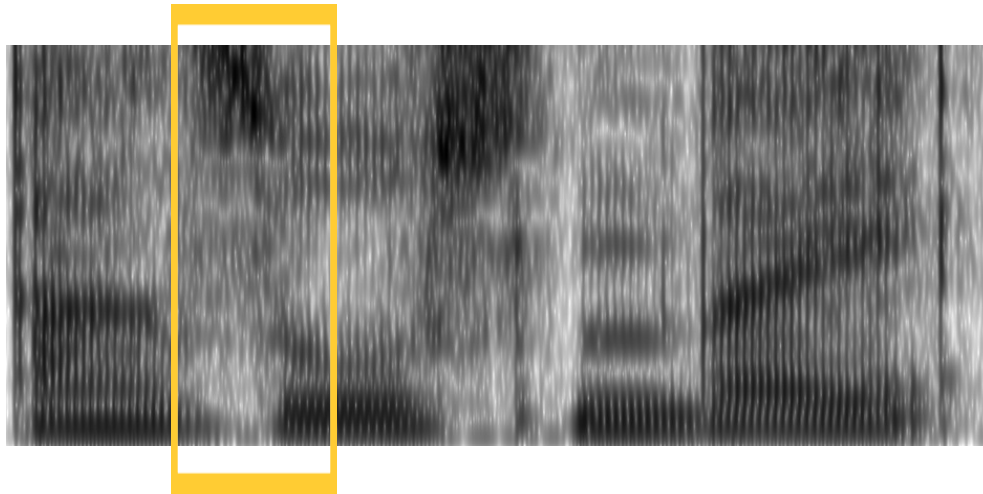
Moscow always has traffic jams.



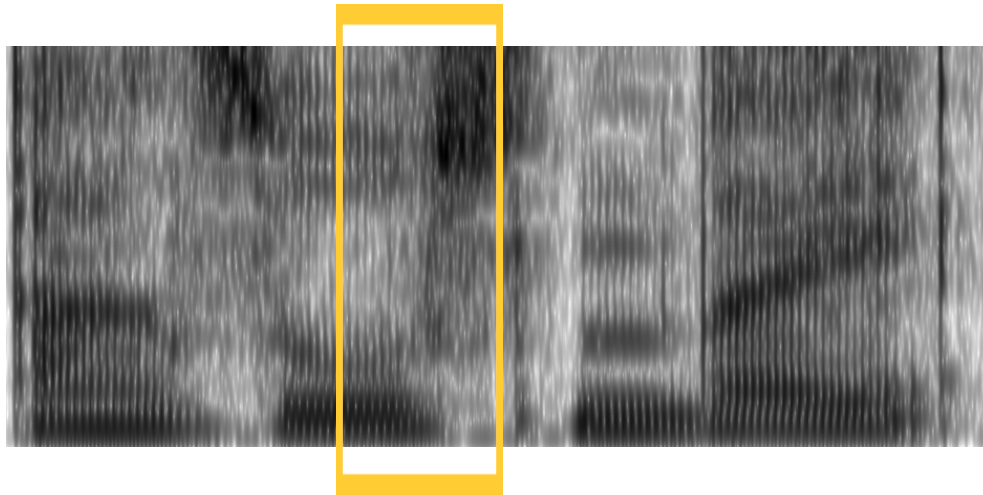
Moscow always has traffic jams.



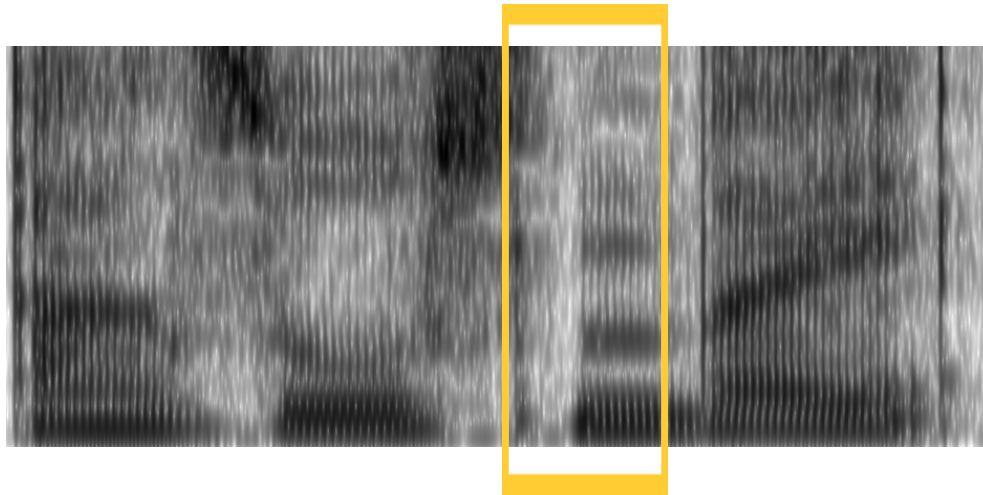
Moscow always has traffic jams.



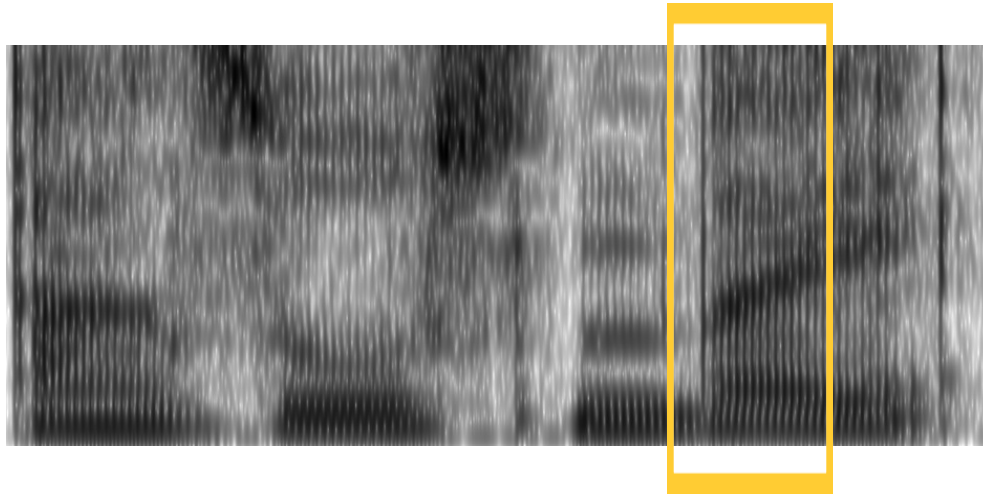
Moscow always has traffic jams.



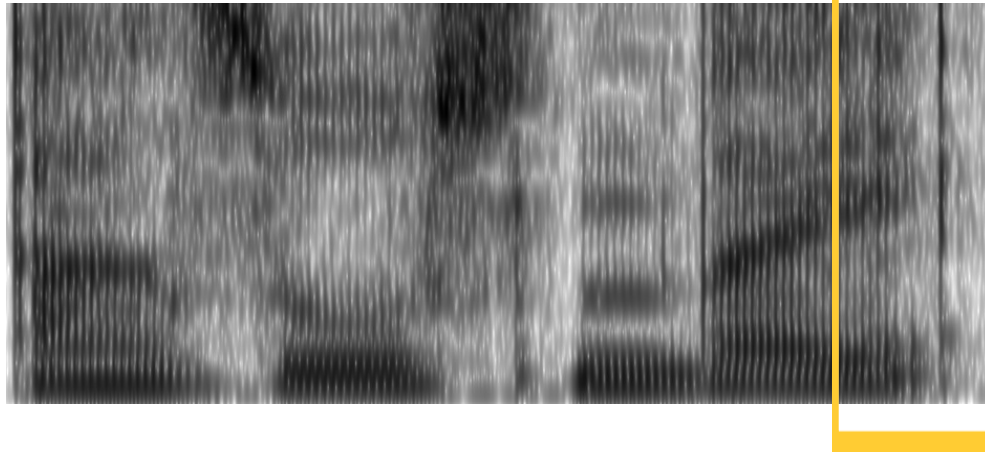
Moscow always has traffic jams.



Moscow always has traffic jams.



Moscow always has traffic jams.



Moscow always has traffic jams.

```
[001000001000000000]
[002000030000700002]
[0301000000000900]
[0050900010004000]
[060100000100000008]
[0020000300900000]
[0010000000000100]
[03000000100700002]
[0301000009000000]
[005090400004000]
[0601000000000008]
[0020000300900000]
[0000000010000000]
[00200000100700002]
[0301000009000000]
[005090410004000]
[060100000100000008]
[0020000010000000]
[0000000010000000]
[03000000100700002]
[0301000009000000]
[005090410004000]
[060100000100000008]
[0020000010000000]
[0000000010000000]
[03000000100700002]
[0301000009000000]
[005090410004000]
[060100000100000008]
[0020000310000000]
```

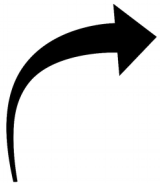
M o s c o w a l w a y s h a s t r a f f I c j a m s

[0010000100000000]
[002000300070002]
[3010000000900]
[005090010004000]
[601000010000008]
[002000300900000]
[001000000000100]
[030000010070002]
[301000090000000]
[005090400004000]
[60100000000008]
[002000300900000]
[000000010000000]
[002000010070002]
[301000090000000]
[005090410004000]
[601000010000008]
[002000010000000]
[000000010000000]
[030000010070002]
[301000090000000]
[005090410004000]
[601000010000008]
[002000310000000]

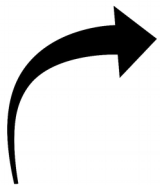
[illegible]

[0010000100000000]	[0020000300070002]	[3010000000009000]	[0050900100040000]	[6010000100000008]	[0020000300900000]	[0010000000001000]	[030000010070002]	[3010000900000000]	[0050904000040000]	[6010000000000008]	[0020000300900000]	[0000000100000000]	[002000010070002]	[3010000900000000]	[0050904100040000]	[6010000100000008]	[0020000100000000]	[0000000100000000]	[030000010070002]	[3010000900000000]	[0050904100040000]	[6010000100000008]	[0020000310000000]
--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------

TEXT FEATURES

[illegible]

AUDIO FEATURES



[001000010000000]	[0020000300070002]	[301000000000900]	[005090010004000]	[601000010000008]	[002000030090000]	[001000000000100]	[030000010070002]	[301000090000000]	[005090400004000]	[601000000000008]	[002000030090000]	[000000010000000]	[002000010070002]	[301000090000000]	[005090410004000]	[601000010000008]	[002000010000000]	[000000010000000]	[030000010070002]	[301000090000000]	[005090410004000]	[601000010000008]	[002000031000000]
-------------------	--------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------

References / Resources

- CSTR slides: `speech.zone`
- Ossian:
`github.com/CSTR-Edinburgh/Ossian`
- Merlin: `github.com/CSTR-Edinburgh`

Рахмат!