

Presentation to the master's qualification work on the topic

«Development of an intelligent system to support the switching of modes of reading the electronic books with audio data»

Fulfilled: 2nd year student

of the group 3ACIT-18m

Andrii Henri

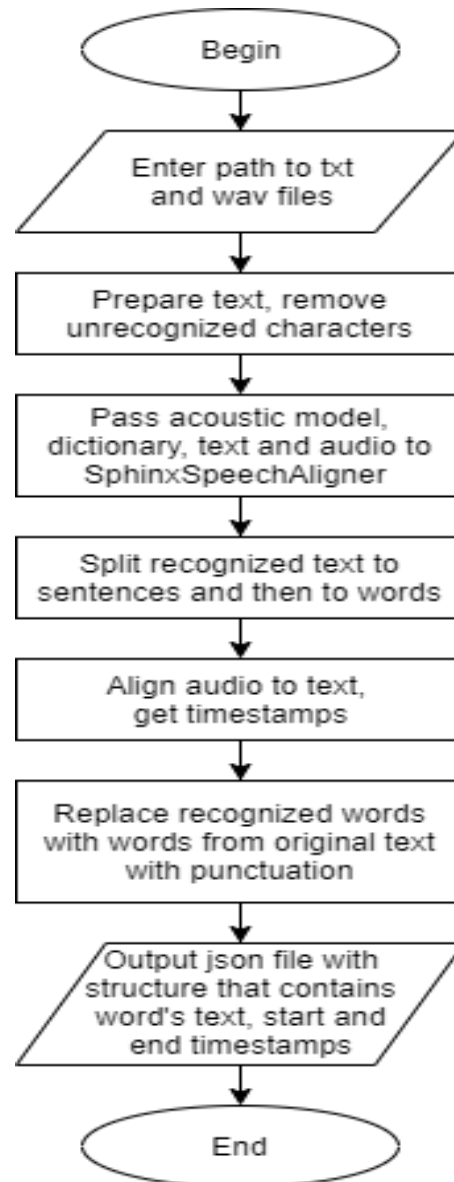
Formulation of the problem

- Develop an algorithm for aligning text to audio
- Develop a system that uses the algorithm for automatically aligning
- Develop a mobile application that uses the above developments

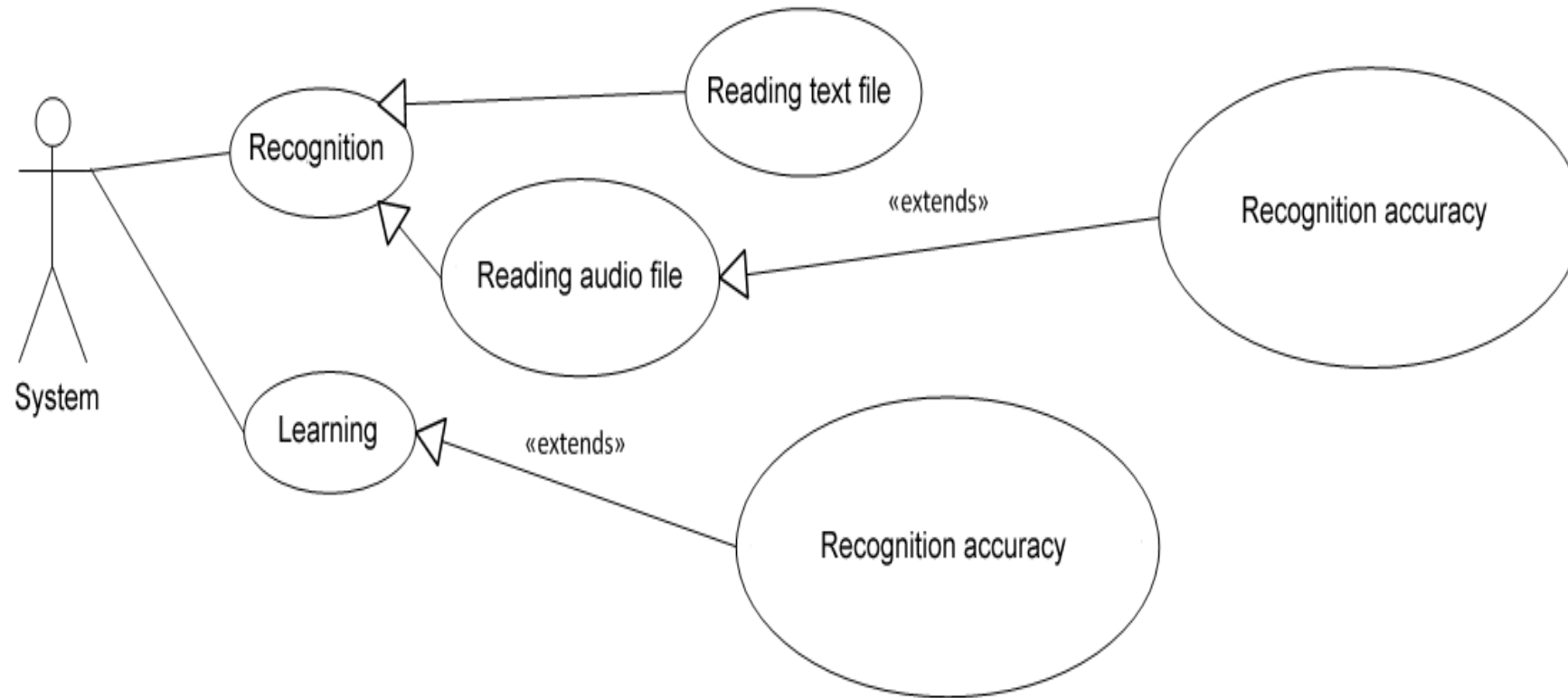
Comparison of speech recognition systems

	Google Speech API	Yandex SpeechKit	Microsoft Speech API	CMU Sphinx	HTK
Getting features	MFCC, PLP	MFCC, PLP	MFCC, PLP	MFCC, PLP. LPC	MFCC
Acousting modeling	Neural networks	HMM, Neural networks	HMM, Neural networks	HMM	HMM
Language modeling	WFST	WFST, a recurrent neural network	WFST	N-gramm, FST, CFG	N-gramm
WER, %	8,9	-	-	22,7	19,8
WRR, %	91,1	-	-	77,3	80,2

Aligning algorithm



Use case UML-diagram



Output JSON file

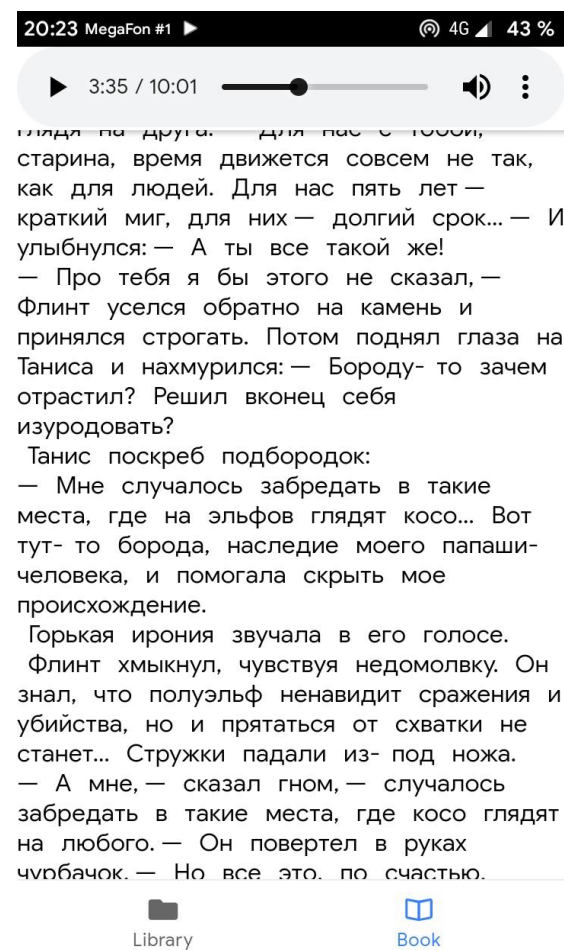
```
{
  "text": "никому ",
  "startTime": 7890,
  "endTime": 8290
},
{
  "text": "не ",
  "startTime": 8300,
  "endTime": 8430
},
{
  "text": "интересная ",
  "startTime": 8440,
  "endTime": 9080
},
{
  "text": "желтая звезда.\r\n\r\nВокруг нее, на ",
  "startTime": 12420,
  "endTime": 12550
},
{
  "text": "расстоянии ",
  "startTime": 12560,
  "endTime": 13200
},
{
  "text": "примерно ",
  "startTime": 13210,
  "endTime": 13730
},
{
  "text": "девяноста ",
  "startTime": 13740,
  "endTime": 14350
},
}
```

Mobile application modules structure

```
> e2e
> node_modules
> platforms
> plugins
> resources
▼ src
  ▼ app
    ▼ book
      TS book.module.ts
      <> book.page.html
      🌀 book.page.scss
      TS book.page.spec.ts
      TS book.page.ts
    ▼ library
      TS library.module.ts
      <> library.page.html
      🌀 library.page.scss
      TS library.page.spec.ts
      TS library.page.ts
    > tab3
    > tabs
    TS app-routing.module.ts
    <> app.component.html
    🌀 app.component.scss
    TS app.component.spec.ts
    TS app.component.ts
    TS app.module.ts
  ▼ assets
    > icon
```

Mobile application developing

- Screen views of application



Quality assessment results

Text	Words count	Deletions count	WER, %
The Hitchhiker's Guide to the Galaxy	1651	142	8.6
Dragons of Autumn Twilight	1587	177	11.15
The Little Prince	10705	2334	21.8

Quality assessment

$$WER = \frac{S+I+D}{N} * 100\%$$

Insertions

The quick **slick** brown fox

Deletions

The brown fox

Substitutions

The **slick** brown fox

Thank you!