

## Установка зависимостей

Для использования Whisper требуется:

1) установить набор ffmpeg:

```
sudo apt install ffmpeg
```

2) установить зависимости:

```
pip install -r requirements.txt
```

## Результаты запуска

Size	Parameters	English-only model	Multilingual model	Required VRAM	Relative speed
tiny	39 M	<code>tiny.en</code>	<code>tiny</code>	~1 GB	~32x
base	74 M	<code>base.en</code>	<code>base</code>	~1 GB	~16x
small	244 M	<code>small.en</code>	<code>small</code>	~2 GB	~6x
medium	769 M	<code>medium.en</code>	<code>medium</code>	~5 GB	~2x
large	1550 M	N/A	<code>large</code>	~10 GB	1x

Существует 5 видов моделей, чем модель тяжелее, тем качественнее получается расшифровываемое сообщение и тем больше она весит, работает и требует оперативной памяти.

В запусках участвовали первые 3 вида. Образцом для перевода служили аудиосообщения из телеграма, общей длительностью 2 минуты 13 секунд.

Результаты запуска находятся в папке *result*. Образцы находятся в папке *samples*.

Модель **tiny** отработала за 16 секунд, полученный текст низкого качества, слова разбиты на куски.

Модель **base** отработала за 24 секунды, получен текст более высокого качества количество разорванных слов значительно уменьшилось, но имена собственные не воспринимаются моделью, также страдают склонения глаголов.

Модель **small** отработала за 73 секунды, получен текст приемлемого качества, имена собственные (такие как Черчилль) воспринимаются моделью. Также модель начала распознавать знаки препинания.