


ИНСТРУКЦИЯ ПО УСТАНОВКЕ АР РАСПОЗНАВАНИЕ РЕЧИ (AprSpeechToolkit)

Automated Process | 09.02.2026

Установка пакета

Чтобы установить пакет **«AprSpeechToolkit»**, выполните следующие шаги:

Перейдите в «Дизайнер системы» по кнопке  и в блоке «Приложения» откройте раздел «Установка и удаление приложений» (рис. 1).

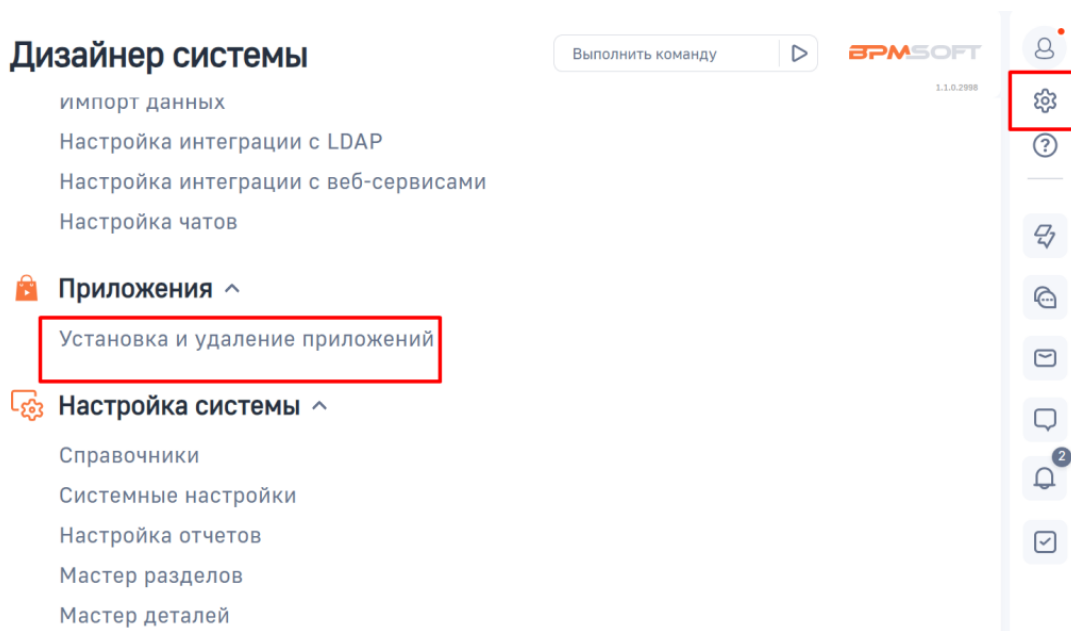


Рисунок 1 – Переход в раздел «Установка и удаление приложений».

Нажмите на кнопку «Добавить приложение» и выберите действие «Установить из файла» (рис. 2).

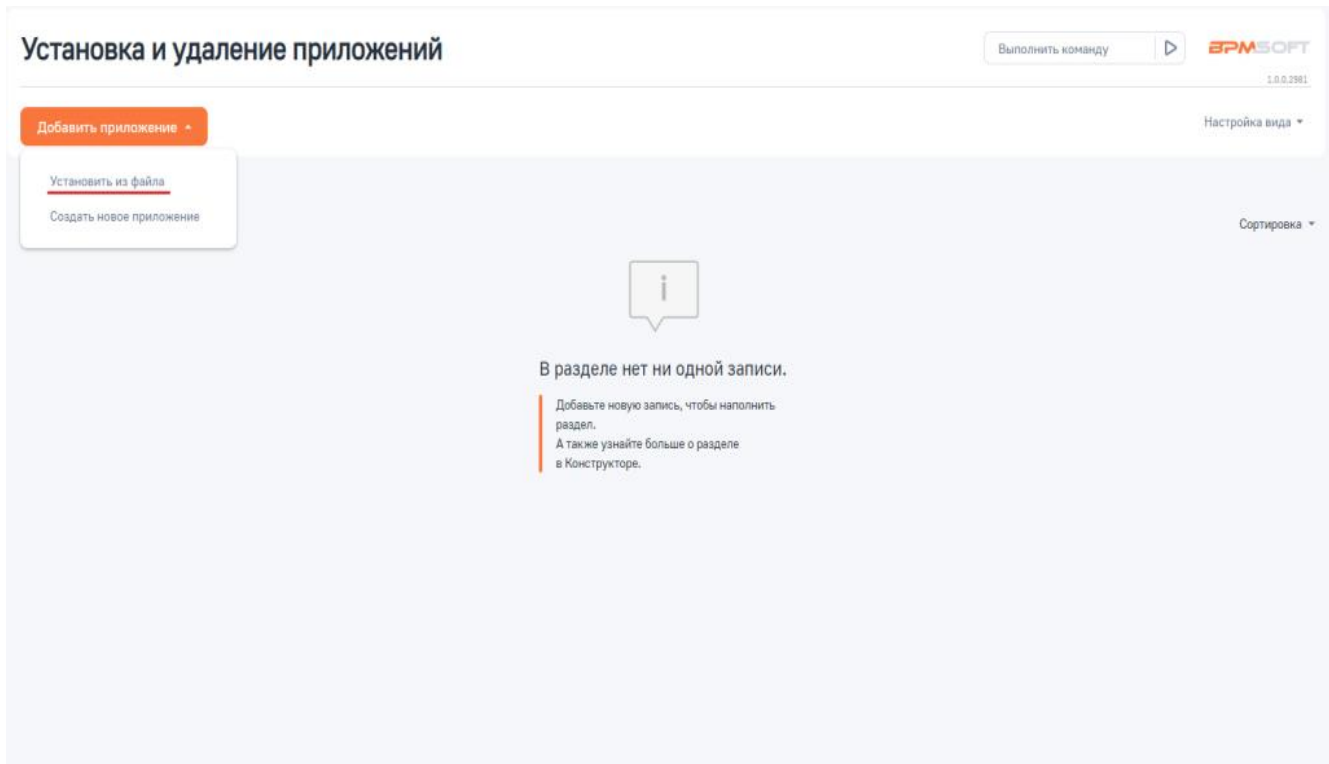


Рисунок 2 – Страница «Установка и удаление приложений».

Перетащите установочный файл пакета **«AprSpeechToolkit»** в область на странице установки или нажмите кнопку «Выбрать файл» и в стандартном проводнике вашей операционной системы выберите пакет **«AprSpeechToolkit»** (рис. 3).

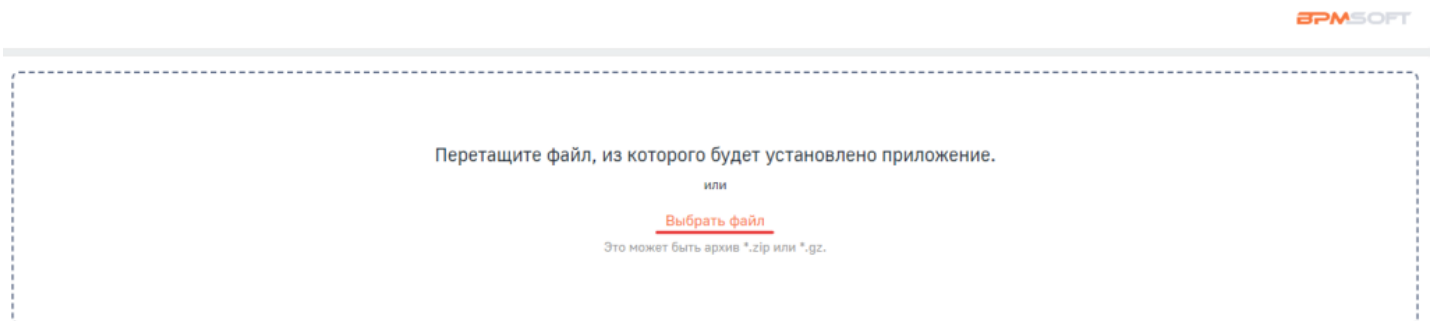


Рисунок 3 – Выбор установочного файла пакета.

После добавления установочного файла запустится процесс установки (рис. 4).

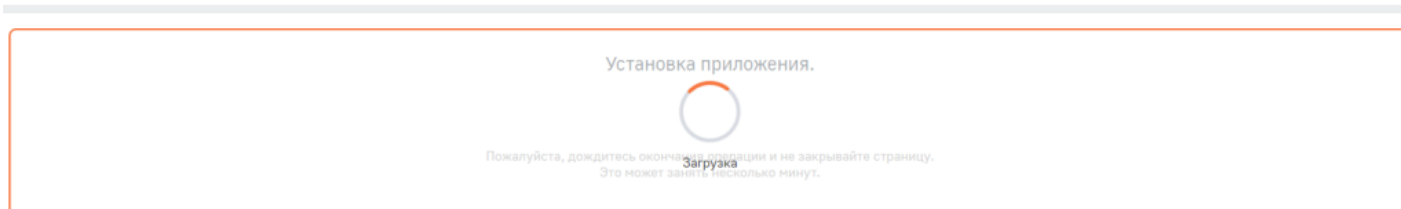


Рисунок 4 – Процесс установки пакета.

При успешной установке пакета, на странице отобразиться сообщение «Приложение успешно установлено» (рис. 5). В случае возникновения ошибок обратитесь в службу технической поддержки BPMSoft.

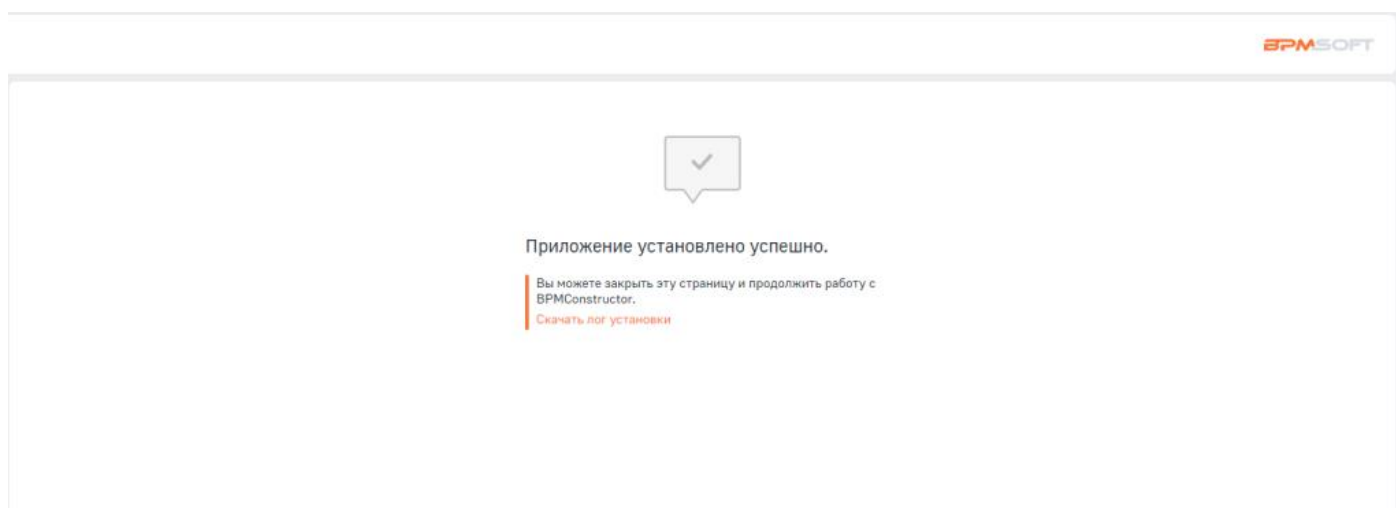


Рисунок 5 – Уведомление об успешной установке пакета.

Настройка докера (Разворачивание docker образа)

Для скачивания образа в командной строке нужно выполнить:

```
docker pull automatedprocess/speech-toolkit-worker
```

Или в приложении Docker Desktop в строке поиска (рис. 6) ввести:

automatedprocess затем скачать нажатием кнопки «Pull» на образе «automatedprocess/speech-toolkit-worker».

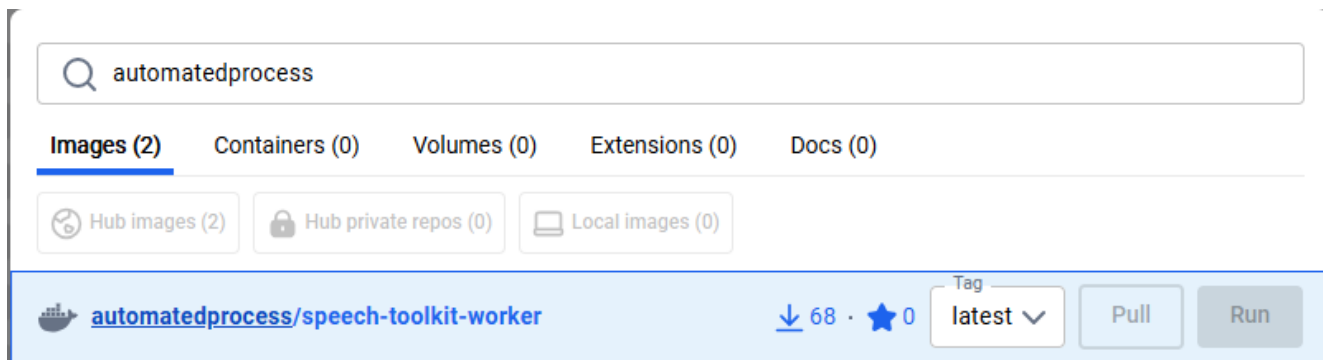


Рисунок 6 – Поиск в Docker Desktop.

Запуск docker контейнера

Для запуска контейнера в приложении Docker Desktop выполните следующие шаги:

- 1) перейти в пункт меню «Images», нажать на кнопку запуска в меню «Actions» (рис. 7) скачанного образа;

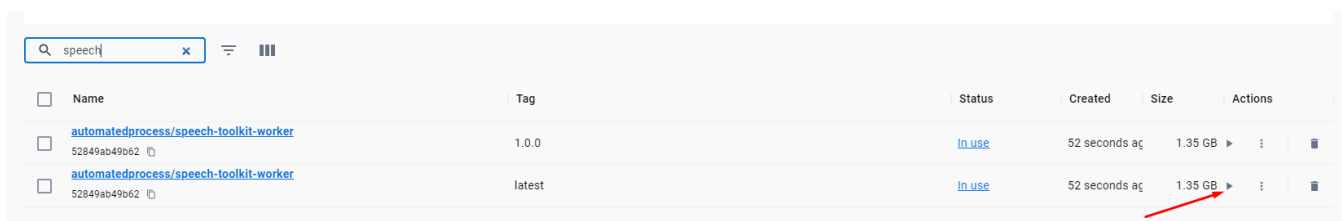


Рисунок 7 – Запуск контейнера.

- 2) развернуть группу полей «Optional settings» (рис. 8);



Run a new container

automatedprocess/speech-toolkit-worker:latest

Optional settings

Container name

Speech2text

A random name is generated if you do not provide one.

Ports

Enter "0" to assign randomly generated host ports.

Host port

8443

:443/tcp

Host port

7016

:8080/tcp

Volumes

Host path

C:\ProgramData\Speer ...

Container path

/data

+

Environment variables

Variable

Value

+

Cancel

Run

Рисунок 8 – Параметры запуска.

- 3) в поле «Container name» вписать любое название, например, «Speech2text» или оставить пустым;
- 4) в «Host port (:443)» - любой, например, 8443;
- 5) в «Host port (:8080)» - любой, например, 7016;

- 6) важно! в «Host path» указать папку хост-машины, в которой будут все данные контейнера, например, C:\ProgramData\Speech2text;
- 7) важно! в «Container path» - текст /data;
- 8) нажать кнопку «Run», дождаться запуска контейнера;
Запуск занимает до 10 минут в зависимости от ресурсов машины. При первом запуске скачиваются модели и необходимые зависимости. Информация в логах следующего вида:
info: Microsoft.Hosting.Lifetime[0]
Content root path: /app
говорит о том, что контейнер запущен.
Полный лог см. в Приложении 2.
- 9) для проверки перейти по адресу в браузере `http://localhost:7016/swagger/index.html`.

Для запуска нового контейнера в командной строке выполнить:

```
docker run -d --name speech2txt -p 7016:8080 -p 8443:443 -v /C/ProgramData/Speech2text:/data automatedprocess/speech-toolkit-worker
```

Если контейнер уже существует, но остановлен, то запускать командой:

```
docker start speech2txt
```

Для Linux(Ubuntu) см. Приложение 1.

Как настроить приложение

После успешной установки пакета нужно добавить в справочник «Воркеры ST»(AprSTWorker) ссылки на развернутые контейнеры(один или несколько). Ссылка – `http(s)://url-сервера:port(указанный при настройке контейнера в свойствах «Host port (:443)»/ «Host port (:8080)»)`. (Рис 8).

Использование продукта

При появлении нового аудио звонка («Звонок» -> «Файлы и примечания») запись отправляется (по сигналу, через 10 секунд после добавления) в очередь на распознавание.

Обработка файлов происходит согласно очереди приоритизации. Очередь настраивается в справочнике «Роль в приоритете ST», где необходимо указать «Роль» - Роль/Пользователь системы, значение «Приоритет» и «Название схемы файла» - имя объекта на английском в котором хранятся аудио для распознавания, например, «CallFile». При обработке звонка происходит поиск в данном справочнике по полю «Создал» указанному в звонке. Если для звонка найдено несколько значений приоритета, будет использоваться минимальный из них.

Чем меньше значение приоритета, тем раньше обработается звонок.

Если в справочнике не найдена ни одна запись, звонок обработан не будет.

Далее необходимо заполнить системную настройку **«Максимальный приоритет распознавания для аудио файлов ST»**. Если все найденные для звонка записи в справочнике «Роль в приоритете ST» имеют приоритет больше значения данной настройки, звонок обработан не будет.

Раз в час запускается бизнес процесс **«Обработка новых аудио сегментов»**, который импортирует обработанный текст. Процесс так же можно запустить вручную, либо по кнопке «Обновить» на вкладке «Распознавание текста».

Если в объекте «Файлы и ссылки звонка» (CallFile) несколько файлов, будет обработан только один из них.

В результате на странице объекта «Звонок» на вкладке «Распознавание текста» в поле «Результат» (рис. 9) будет отображаться расшифрованный текст аудио звонка:

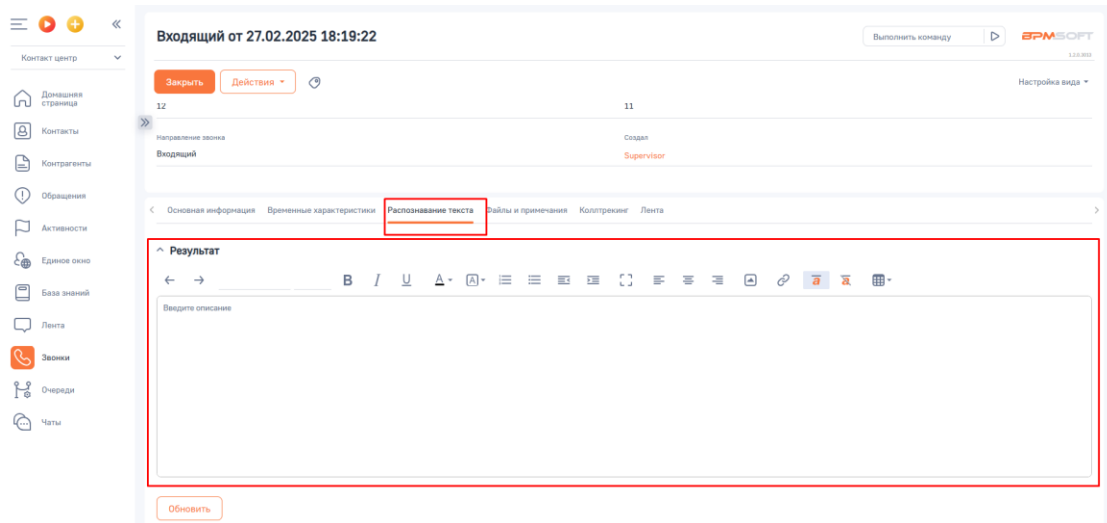


Рисунок 9 – Результат распознавания.

Приложение 1.

Набор команд для установки образа на Linux(Ubuntu).

Если на вашем сервере уже установлен docker, начните с пункта «9».

1. Обновление системы

```
sudo apt update
```

2. Установка зависимостей

```
sudo apt install -y ca-certificates curl gnupg lsb-release
```

3. Добавление GPG ключа

```
sudo install -m 0755 -d /etc/apt/keyrings
```

```
curl -fsSL https://download.docker.com/linux/ubuntu/gpg | sudo gpg --dearmor -o /etc/apt/keyrings/docker.gpg
```

```
sudo chmod a+r /etc/apt/keyrings/docker.gpg
```

4. Добавление репозитория

```
echo "deb [arch=$(dpkg --print-architecture) signed-by=/etc/apt/keyrings/docker.gpg] https://download.docker.com/linux/ubuntu $(. /etc/os-release && echo "$VERSION_CODENAME") stable" | sudo tee /etc/apt/sources.list.d/docker.list > /dev/null
```

5. Установка Docker

```
sudo apt update
```

```
sudo apt install -y docker-ce docker-ce-cli containerd.io docker-buildx-plugin docker-compose-plugin
```

6. Запуск Docker

```
sudo systemctl start docker
```

```
sudo systemctl enable docker
```

7. Добавление пользователя в группу (чтобы не использовать sudo)

```
sudo usermod -aG docker $USER
```

```
newgrp docker
```

8. Проверка

```
docker --version
```

```
docker run hello-world
```

9. Создать папку /data в корне

```
sudo mkdir -p data
```

```
sudo chown $USER:$USER data
```

```
chmod 755 data
```

10. Загрузка и запуск образа

```
docker pull automatedprocess/speech-toolkit-worker
```

```
docker run -d --name speech2txt -p 80:8080 -p 8443:443 -v /data:/data  
automatedprocess/speech-toolkit-worker
```

```
docker logs -f speech2txt
```

Примечание: в случае запуска контейнера на сервер Linux с параметрами `-p 80:8080 -p 8443:443` (указаны в команде `docker run...`) в справочнике «Воркеры ST» в колонке «Url» ссылка указывается без порта.

Приложение 2. Пример лога успешного запуска docker контейнера.

```
2026-01-14 13:03:39 [entrypoint] Установка модели пунктуации...
```

```
2026-01-14 13:03:39 Collecting git+https://huggingface.co/kontur-ai/sbert_punc_case_ru
```

```
2026-01-14 13:03:39 Cloning https://huggingface.co/kontur-ai/sbert_punc_case_ru to /tmp/pip-req-build-mub1u7tk
```

```
2026-01-14 13:03:39 Running command git clone --filter=blob:none --quiet  
https://huggingface.co/kontur-ai/sbert_punc_case_ru /tmp/pip-req-build-mub1u7tk
```

```
2026-01-14 13:03:41 Resolved https://huggingface.co/kontur-ai/sbert_punc_case_ru to commit  
f778dc6c63bb0ec235a220488862509810e54583
```

```
2026-01-14 13:03:41 Preparing metadata (setup.py): started
```

2026-01-14 13:03:41 Preparing metadata (setup.py): finished with status 'done'

2026-01-14 13:03:41 Requirement already satisfied: numpy in /venv/lib/python3.10/site-packages (from sbert-punc-case-ru==0.2) (1.26.4)

2026-01-14 13:03:42 Collecting torch

2026-01-14 13:03:42 Downloading torch-2.9.1-cp310-cp310-manylinux_2_28_x86_64.whl (899.8 MB)

2026-01-14 13:05:41

— 899.8/899.8 MB 7.6 MB/s eta 0:00:00

2026-01-14 13:05:44 Collecting transformers>=4.36.2

2026-01-14 13:05:44 Downloading transformers-4.57.5-py3-none-any.whl (12.0 MB)

2026-01-14 13:05:46

— 12.0/12.0 MB 4.8 MB/s eta 0:00:00

2026-01-14 13:05:46 Requirement already satisfied: tqdm>=4.27 in /venv/lib/python3.10/site-packages (from transformers>=4.36.2->sbert-punc-case-ru==0.2) (4.67.1)

2026-01-14 13:05:46 Requirement already satisfied: pyyaml>=5.1 in /venv/lib/python3.10/site-packages (from transformers>=4.36.2->sbert-punc-case-ru==0.2) (6.0.2)

2026-01-14 13:05:47 Collecting tokenizers<=0.23.0,>=0.22.0

2026-01-14 13:05:47 Downloading tokenizers-0.22.2-cp39-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.3 MB)

2026-01-14 13:05:48

— 3.3/3.3 MB 2.3 MB/s eta 0:00:00

2026-01-14 13:05:49 Collecting safetensors>=0.4.3

2026-01-14 13:05:49 Downloading safetensors-0.7.0-cp38-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (507 kB)

2026-01-14 13:05:49

— 507.2/507.2 KB 2.6 MB/s eta 0:00:00

2026-01-14 13:05:50 Collecting regex!=2019.12.17

2026-01-14 13:05:50 Downloading regex-2025.11.3-cp310-cp310-manylinux2014_x86_64.manylinux_2_17_x86_64.manylinux_2_28_x86_64.whl (791 kB)

2026-01-14 13:05:50

— 791.7/791.7 KB 2.6 MB/s eta 0:00:00

2026-01-14 13:05:50 Requirement already satisfied: filelock in /venv/lib/python3.10/site-packages (from transformers>=4.36.2->sbert-punc-case-ru==0.2) (3.18.0)

2026-01-14 13:05:50 Requirement already satisfied: packaging>=20.0 in /venv/lib/python3.10/site-packages (from transformers>=4.36.2->sbert-punc-case-ru==0.2) (25.0)

2026-01-14 13:05:50 Collecting huggingface-hub<1.0,>=0.34.0

2026-01-14 13:05:50 Downloading huggingface_hub-0.36.0-py3-none-any.whl (566 kB)

2026-01-14 13:05:51

— 566.1/566.1 KB 2.6 MB/s eta 0:00:00

2026-01-14 13:05:51 Requirement already satisfied: requests in /venv/lib/python3.10/site-packages (from transformers>=4.36.2->sbert-punc-case-ru==0.2) (2.32.4)

2026-01-14 13:05:51 Requirement already satisfied: sympy>=1.13.3 in /venv/lib/python3.10/site-packages (from torch->sbert-punc-case-ru==0.2) (1.14.0)

2026-01-14 13:05:51 Collecting nvidia-cusolver-cu12==11.7.3.90

2026-01-14 13:05:51 Downloading nvidia_cusolver_cu12-11.7.3.90-py3-none-manylinux_2_27_x86_64.whl (267.5 MB)

2026-01-14 13:07:08

— 267.5/267.5 MB 4.2 MB/s eta 0:00:00

2026-01-14 13:07:09 Collecting nvidia-cuda-cupti-cu12==12.8.90

2026-01-14 13:07:09 Downloading nvidia_cuda_cupti_cu12-12.8.90-py3-none-manylinux2014_x86_64.whl (10.2 MB)

2026-01-14 13:07:11

— 10.2/10.2 MB 4.6 MB/s eta 0:00:00

2026-01-14 13:07:11 Collecting triton==3.5.1

2026-01-14 13:07:12 Downloading triton-3.5.1-cp310-cp310-manylinux_2_27_x86_64.whl (170.3 MB)

2026-01-14 13:07:42

— 170.3/170.3 MB 5.1 MB/s eta 0:00:00

2026-01-14 13:07:42 Collecting nvidia-cuda-nvrtc-cu12==12.8.93

2026-01-14 13:07:42 Downloading nvidia_cuda_nvrtc_cu12-12.8.93-py3-none-manylinux2010_x86_64.manylinux_2_12_x86_64.whl (88.0 MB)

2026-01-14 13:07:59

— 88.0/88.0 MB 4.3 MB/s eta 0:00:00

2026-01-14 13:08:00 Collecting nvidia-cufft-cu12==11.3.3.83

2026-01-14 13:08:00 Downloading nvidia_cufft_cu12-11.3.3.83-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (193.1 MB)

2026-01-14 13:08:41

— 193.1/193.1 MB 7.6 MB/s eta 0:00:00

2026-01-14 13:08:41 Collecting jinja2

2026-01-14 13:08:41 Downloading jinja2-3.1.6-py3-none-any.whl (134 kB)

2026-01-14 13:08:41

134.9/134.9 KB 12.6 MB/s eta 0:00:00

2026-01-14 13:08:42 Collecting nvidia-nvshmem-cu12==3.3.20

2026-01-14 13:08:42 Downloading nvidia_nvshmem_cu12-3.3.20-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (124.7 MB)

2026-01-14 13:09:19

— 124.7/124.7 MB 2.8 MB/s eta 0:00:00

2026-01-14 13:09:20 Collecting nvidia-nvjitlink-cu12==12.8.93

2026-01-14 13:09:20 Downloading nvidia_nvjitlink_cu12-12.8.93-py3-none-manylinux2010_x86_64.manylinux_2_12_x86_64.whl (39.3 MB)

2026-01-14 13:09:31

— 39.3/39.3 MB 3.6 MB/s eta 0:00:00

2026-01-14 13:09:31 Collecting nvidia-cuspars-cu12==12.5.8.93

2026-01-14 13:09:31 Downloading nvidia_cuspars_cu12-12.5.8.93-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (288.2 MB)

2026-01-14 13:11:24

— 288.2/288.2 MB 4.8 MB/s eta 0:00:00

2026-01-14 13:11:25 Collecting nvidia-cudnn-cu12==9.10.2.21

2026-01-14 13:11:25 Downloading nvidia_cudnn_cu12-9.10.2.21-py3-none-manylinux_2_27_x86_64.whl (706.8 MB)

2026-01-14 13:13:25

— 706.8/706.8 MB 8.8 MB/s eta 0:00:00

2026-01-14 13:13:26 Collecting nvidia-cuda-runtime-cu12==12.8.90

2026-01-14 13:13:26 Downloading nvidia_cuda_runtime_cu12-12.8.90-py3-none-manylinux2014_x86_64.whl (954 kB)

2026-01-14 13:13:27

— 954.8/954.8 KB 8.9 MB/s eta 0:00:00

2026-01-14 13:13:27 Requirement already satisfied: typing-extensions>=4.10.0 in /venv/lib/python3.10/site-packages (from torch->sbert-punc-case-ru==0.2) (4.14.0)

2026-01-14 13:13:27 Collecting nvidia-cublas-cu12==12.8.4.1

2026-01-14 13:13:27 Downloading nvidia_cublas_cu12-12.8.4.1-py3-none-manylinux_2_27_x86_64.whl (594.3 MB)

2026-01-14 13:14:37

— 594.3/594.3 MB 7.2 MB/s eta 0:00:00

2026-01-14 13:14:38 Collecting nvidia-curand-cu12==10.3.9.90

2026-01-14 13:14:39 Downloading nvidia_curand_cu12-10.3.9.90-py3-none-manylinux_2_27_x86_64.whl (63.6 MB)

2026-01-14 13:14:46

— 63.6/63.6 MB 8.5 MB/s eta 0:00:00

2026-01-14 13:14:46 Collecting nvidia-cusparselt-cu12==0.7.1

2026-01-14 13:14:46 Downloading nvidia_cusparselt_cu12-0.7.1-py3-none-manylinux2014_x86_64.whl (287.2 MB)

2026-01-14 13:15:23

— 287.2/287.2 MB 8.3 MB/s eta 0:00:00

2026-01-14 13:15:24 Requirement already satisfied: fsspec>=0.8.5 in /venv/lib/python3.10/site-packages (from torch->sbert-punc-case-ru==0.2) (2025.5.1)

2026-01-14 13:15:24 Collecting nvidia-nvtx-cu12==12.8.90

2026-01-14 13:15:24 Downloading nvidia_nvtx_cu12-12.8.90-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (89 kB)

2026-01-14 13:15:24

— 90.0/90.0 KB 7.1 MB/s eta 0:00:00

2026-01-14 13:15:24 Collecting nvidia-cufile-cu12==1.13.1.3

2026-01-14 13:15:24 Downloading nvidia_cufile_cu12-1.13.1.3-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (1.2 MB)

2026-01-14 13:15:24

— 1.2/1.2 MB 9.3 MB/s eta 0:00:00

2026-01-14 13:15:25 Collecting networkx>=2.5.1

2026-01-14 13:15:25 Downloading networkx-3.4.2-py3-none-any.whl (1.7 MB)

2026-01-14 13:15:25

— 1.7/1.7 MB 8.9 MB/s eta 0:00:00

2026-01-14 13:15:25 Collecting nvidia-nccl-cu12==2.27.5

2026-01-14 13:15:25 Downloading nvidia_nccl_cu12-2.27.5-py3-none-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (322.3 MB)

2026-01-14 13:16:16

— 322.3/322.3 MB 7.3 MB/s eta 0:00:00

2026-01-14 13:16:17 Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /venv/lib/python3.10/site-packages (from huggingface-hub<1.0,>=0.34.0->transformers>=4.36.2->sbert-punc-case-ru==0.2) (1.1.4)

2026-01-14 13:16:17 Requirement already satisfied: mpmath<1.4,>=1.1.0 in /venv/lib/python3.10/site-packages (from sympy>=1.13.3->torch->sbert-punc-case-ru==0.2) (1.3.0)

2026-01-14 13:16:17 Collecting MarkupSafe>=2.0

2026-01-14 13:16:17 Downloading markupsafe-3.0.3-cp310-cp310-manylinux2014_x86_64.manylinux_2_17_x86_64.manylinux_2_28_x86_64.whl (20 kB)

2026-01-14 13:16:17 Requirement already satisfied: certifi>=2017.4.17 in /venv/lib/python3.10/site-packages (from requests->transformers>=4.36.2->sbert-punc-case-ru==0.2) (2025.6.15)

2026-01-14 13:16:17 Requirement already satisfied: urllib3<3,>=1.21.1 in /venv/lib/python3.10/site-packages (from requests->transformers>=4.36.2->sbert-punc-case-ru==0.2) (2.4.0)

2026-01-14 13:16:17 Requirement already satisfied: idna<4,>=2.5 in /venv/lib/python3.10/site-packages (from requests->transformers>=4.36.2->sbert-punc-case-ru==0.2) (3.10)

2026-01-14 13:16:17 Requirement already satisfied: charset_normalizer<4,>=2 in /venv/lib/python3.10/site-packages (from requests->transformers>=4.36.2->sbert-punc-case-ru==0.2) (3.4.2)

2026-01-14 13:16:17 Using legacy 'setup.py install' for sbert-punc-case-ru, since package 'wheel' is not installed.

2026-01-14 13:16:18 Installing collected packages: nvidia-cusparselt-cu12, triton, safetensors, regex, nvidia-nvtx-cu12, nvidia-nvshmem-cu12, nvidia-nvjitlink-cu12, nvidia-nccl-cu12, nvidia-curand-cu12, nvidia-cufile-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, networkx, MarkupSafe, nvidia-cusparse-cu12, nvidia-cufft-cu12, nvidia-cudnn-cu12, jinja2, huggingface-hub, tokenizers, nvidia-cusolver-cu12, transformers, torch, sbert-punc-case-ru

2026-01-14 13:16:52 Attempting uninstall: huggingface-hub

2026-01-14 13:16:52 Found existing installation: huggingface-hub 0.33.0

2026-01-14 13:16:52 Uninstalling huggingface-hub-0.33.0:

2026-01-14 13:16:52 Successfully uninstalled huggingface-hub-0.33.0

2026-01-14 13:17:18 Running setup.py install for sbert-punc-case-ru: started

2026-01-14 13:17:18 Running setup.py install for sbert-punc-case-ru: finished with status 'done'

2026-01-14 13:17:18 Successfully installed MarkupSafe-3.0.3 huggingface-hub-0.36.0 jinja2-3.1.6 networkx-3.4.2 nvidia-cublas-cu12-12.8.4.1 nvidia-cuda-cupti-cu12-12.8.90 nvidia-cuda-nvrtc-cu12-12.8.93 nvidia-cuda-runtime-cu12-12.8.90 nvidia-cudnn-cu12-9.10.2.21 nvidia-cufft-cu12-11.3.3.83 nvidia-cufile-cu12-1.13.1.3 nvidia-curand-cu12-10.3.9.90 nvidia-cusolver-cu12-11.7.3.90 nvidia-cusparse-cu12-12.5.8.93 nvidia-cusparselt-cu12-0.7.1 nvidia-nccl-cu12-2.27.5 nvidia-nvjitlink-cu12-12.8.93 nvidia-nvshmem-cu12-3.3.20 nvidia-nvtx-cu12-12.8.90 regex-2025.11.3 safetensors-0.7.0 sbert-punc-case-ru-0.2 tokenizers-0.22.2 torch-2.9.1 transformers-4.57.5 triton-3.5.1

2026-01-14 13:17:24 INFO: Started server process [60]

2026-01-14 13:17:24 INFO: Waiting for application startup.

2026-01-14 13:17:24 INFO: Application startup complete.

2026-01-14 13:17:24 INFO: Uvicorn running on http://0.0.0.0:8081 (Press CTRL+C to quit)

2026-01-14 13:17:28 % Total % Received % Xferd Average Speed Time Time Time
Current

2026-01-14 13:17:28 Dload Upload Total Spent Left Speed
100 74596 100 188 100 74408 7 3051 0:00:26 0:00:24 0:00:02 45

2026-01-14 13:17:54 [{"text":"Ну, сейчас к тебе придет бригада.
Давай.", "start_time":0.03, "end_time":2.91}, {"text":"Давай. Я
жду.", "start_time":5.76, "end_time":6.21}]info: Microsoft.Hosting.Lifetime[14]

2026-01-14 13:17:54 Now listening on: http://[::]:8080

2026-01-14 13:17:54 info: Microsoft.Hosting.Lifetime[0]

2026-01-14 13:17:54 Application started. Press Ctrl+C to shut down.

2026-01-14 13:17:54 info: Microsoft.Hosting.Lifetime[0]

2026-01-14 13:17:54 Hosting environment: Production

2026-01-14 13:17:54 info: Microsoft.Hosting.Lifetime[0]

2026-01-14 13:17:54 Content root path: /app