



# Эмоциональный синтез речи

Автор идеи: Евгений Шуранов

Менторы: Евгений Шуранов, Юлия Матвеева

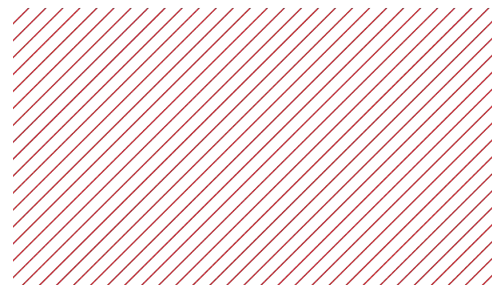
Команда:

Анна Горемыкина

Антон Нестеренко

Дмитрий Астанков

Руслан Ахмеров



# Критерии оценки на защите

---

1. Оценка «**удовлетворительно**» 🤔 — нейтральный синтез речи:
  - оценка MOS 3,5.
2. Оценка «**хорошо**» 😮 — базовый эмоциональный синтез речи:
  - датасеты MSP-Podcast/MSP-Improv/MSP-Face;
  - оценка MOS 3,2;
  - F1-score эмоций — 60%.
3. Оценка «**отлично**» 🤯 — продвинутый эмоциональный синтез речи:
  - GST-based (global style tokens) prosody transfer model;
  - модификации поверх GST;
  - оценка MOS 3,5;
  - F1-score эмоций — 70%.

[illegible]

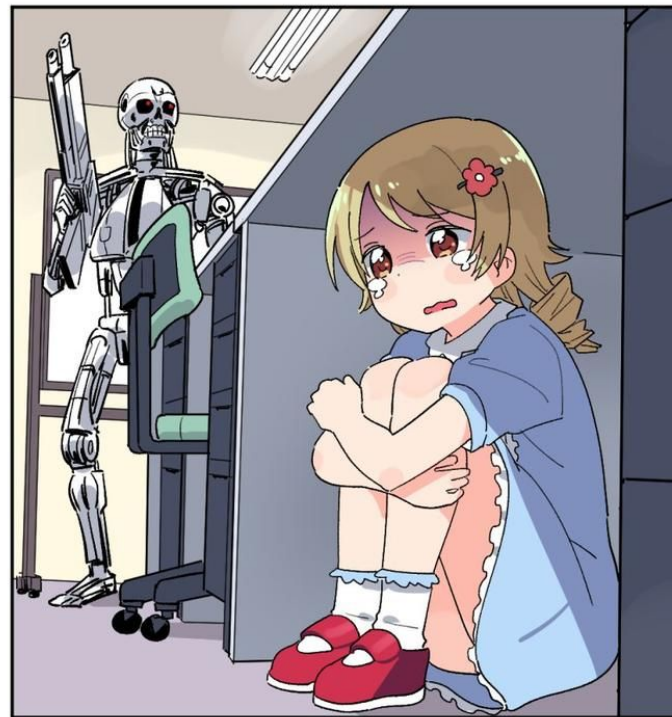
# Описание проекта

---

**Проблема:** безэмоциональные чат-боты отпугивают\* пользователей.

**Задача:** увеличение продолжительности пользовательской сессии с голосовыми чат-ботами.

\* <https://aclanthology.org/W19-5935.pdf>





# Описание проекта

---

**Область применения:** голосовые чат-боты.

**Пользователи:** компании, интегрирующие чат-ботов в свой бизнес (B2B).

**Актуальность:** неудовлетворительное качество существующих универсальных решений.

**Интеграция в существующие сервисы:** Алиса, Маруся, Олег, Игорь.





# Описание проекта

---

**Инструменты и технологии:** Python, PyTorch, torchaudio, inflect, textgridtools, FastAPI, pytest, flake8, mypy, Git.

## Данные:

- **открытые:** VCTK (CC), Emo-V-DB (MIT), MSP-Face (MIT);
- **закрытые:** MSP-Improv (Academic License), MSP-Podcast (Academic License).

## Модели:

- **VAD:** Silero-VAD (better than webrtc-vad);
- **phonemizer & aligner:** Montreal Forced Aligner (SOTA for TTS);
- **feature extractor:** Non-Attentive Tacotron (SOTA for TTS);
- **vocoder:** HiFi-GAN (SOTA for TTS).



# Что сделали

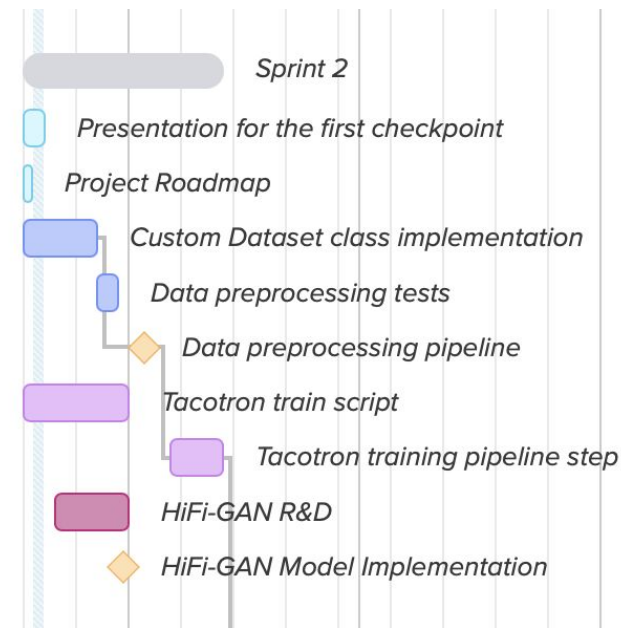
---

1. Выбрали архитектуру модели.
2. Создали GitHub-репозиторий, оформили доску задач, распределили первоначальные задачи, спланировали спринты.
3. Реализовали Feature Extractor Model (Non-Attentive Tacotron), покрыли модель тестами.
4. Реализовали шаги предобработки для датасета VCTK.
5. Добавили механизм автоматической проверки качества кода.



# Планы к следующему рубежному контролю

▼ Sprint 2			0%
Presentation for the first checkpoint	All	0%	
Project Roadmap	Anya	0%	
Custom Dataset class implementation	Dima, Ruslan	0%	
Data preprocessing tests	Dima, Ruslan	0%	
Data preprocessing pipeline	Dima, Ruslan	<input type="checkbox"/>	
Tacotron train script	Anton	0%	
Tacotron training pipeline step	Anton	0%	
HiFi-GAN R&D	Anya	0%	
HiFi-GAN Model Implementation	Anya	<input type="checkbox"/>	

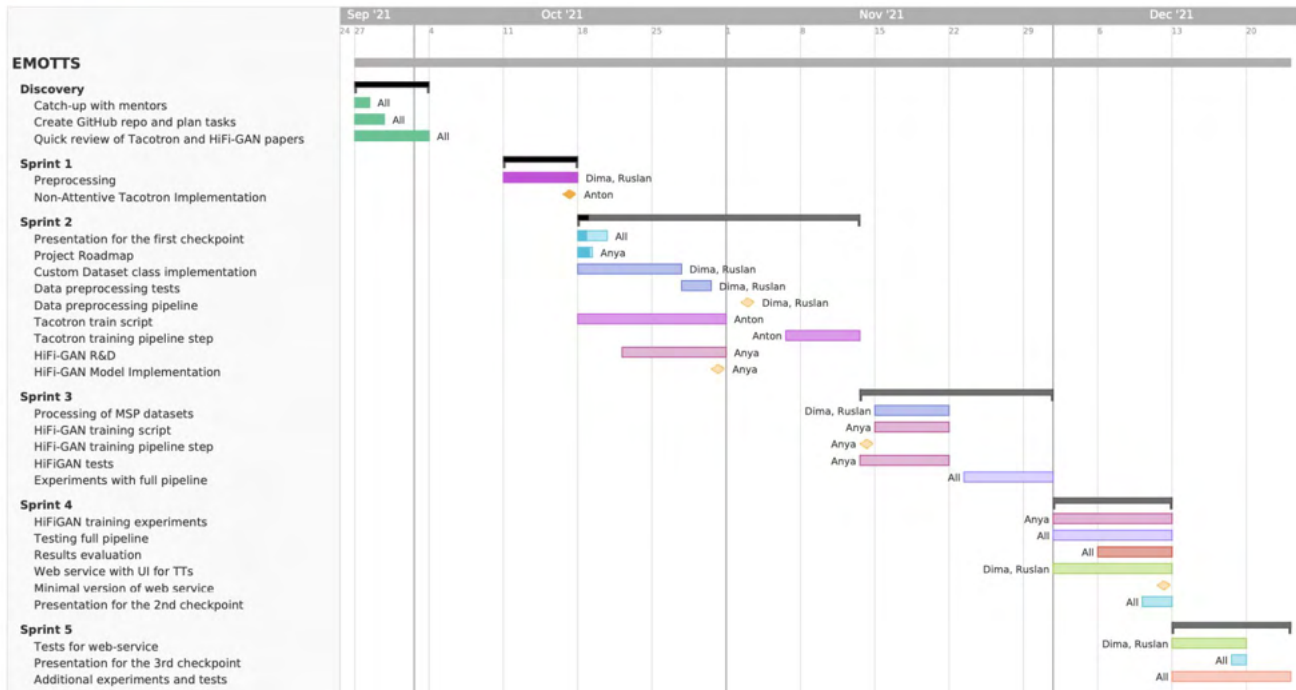




# Roadmap



Created with Free Edition





Команда



# Команда

---

- ✕ Анна Горемыкина: general research, testing, vocoder.
- Антон Нестеренко: code review, testing, feature extractor.
- Дмитрий Астанков: data preprocessing, CI/CD.
- △ Руслан Ахмеров: data preprocessing, code review, web-services.