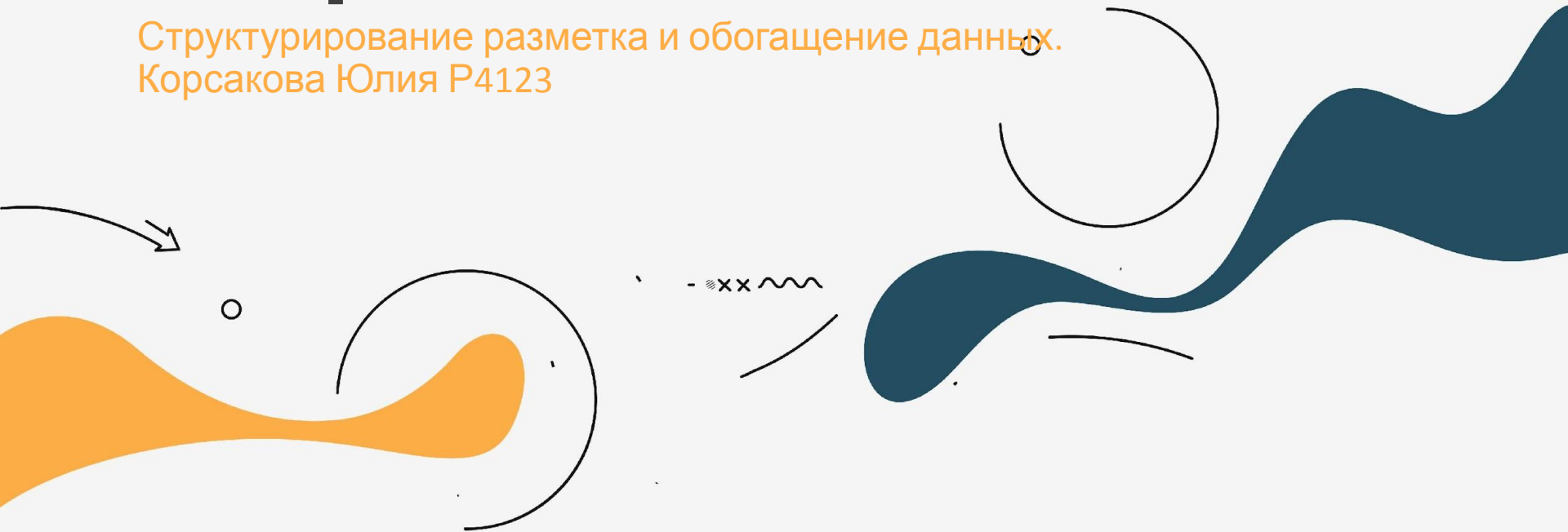


Диагностика ментального здоровья

Структурирование разметка и обогащение данных.
Корсакова Юлия Р4123



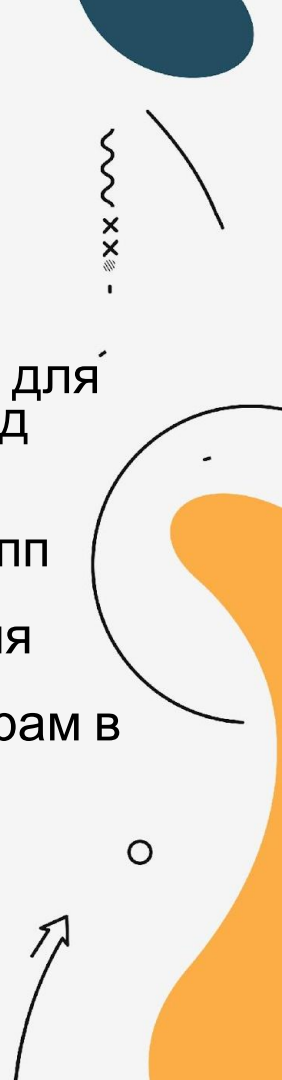
Задача

Система:

- ИИ для распознавания эмоций человека по тексту и аудио для дальнейшего внедрения в голосовые помощники, боты и тд

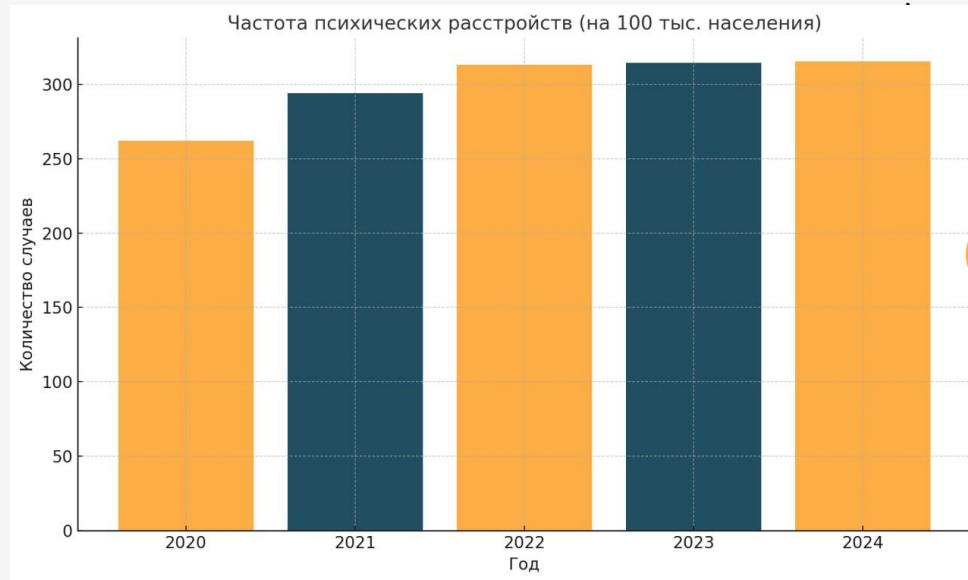
Назначение системы:

- Раннее выявление проблем: автоматический скрининг групп риска.
- Мониторинг динамики: отслеживание изменений состояния пациента между визитами к врачу.
- Поддержка телемедицины: помощь психологам и психиатрам в дистанционной диагностике.
- Массовые исследования: анализ психического здоровья населения (например, после пандемии или кризиса).

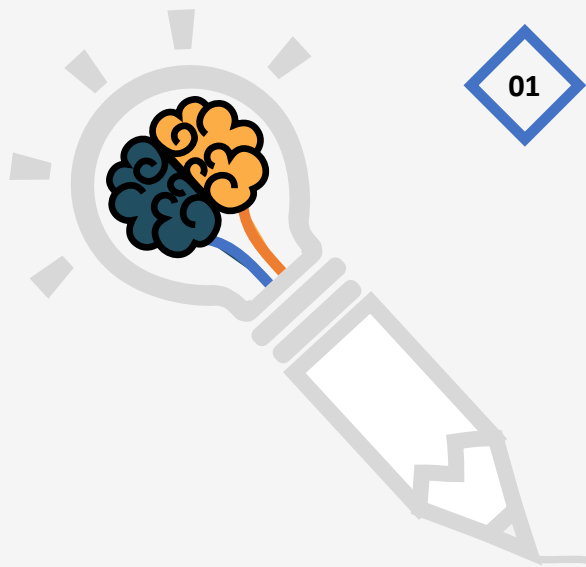


Актуальность

- Рост ментальных расстройств
- Нехватка специалистов
- Цифровизация общения



Зачем нужны данные



01

Обучение

Обучение ИИ на данных с эмоциональной речью

02

Настройка под конкретную задачу

На случайно взятых текстах и аудио тяжело выделить отличительные признаки эмоций

03

Разнообразие данных

Для лучших результатов работы модели

04

Проверка и улучшение

Множество данных также необходимо для дальнейшей проверки и оценки модели

Данные. Модальности

Текст

Имеется множество источников текстовых данных, таких как соц сети и форумы, отзывы на маркетплейсах

Признаки текстовых данных



Для создания системы планируется использовать морфологические, лингвистические и векторные признаки



Аудио

Аудио данные собрать сложнее. Их можно извлечь из видео роликов или специально записанных датасетов



Признаки аудио данных

В качестве признаков аудио данных будут использованы спектрограммы и MFCC коэффициенты



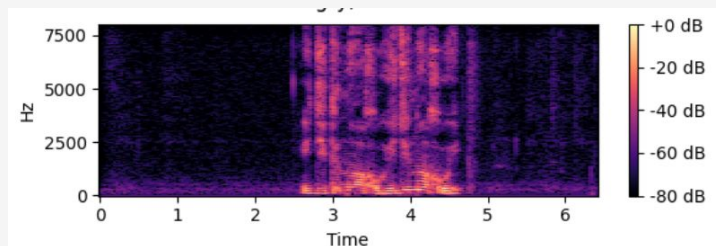
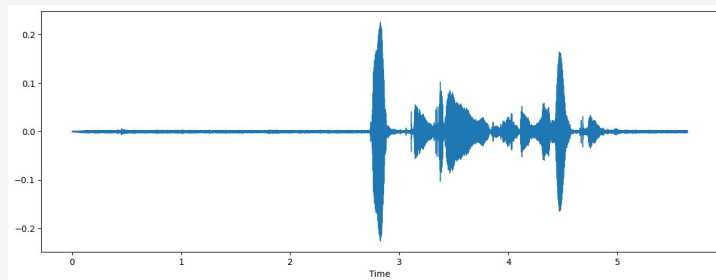
Данные.Примеры

Мультимодальный датасет Dusha

Текст

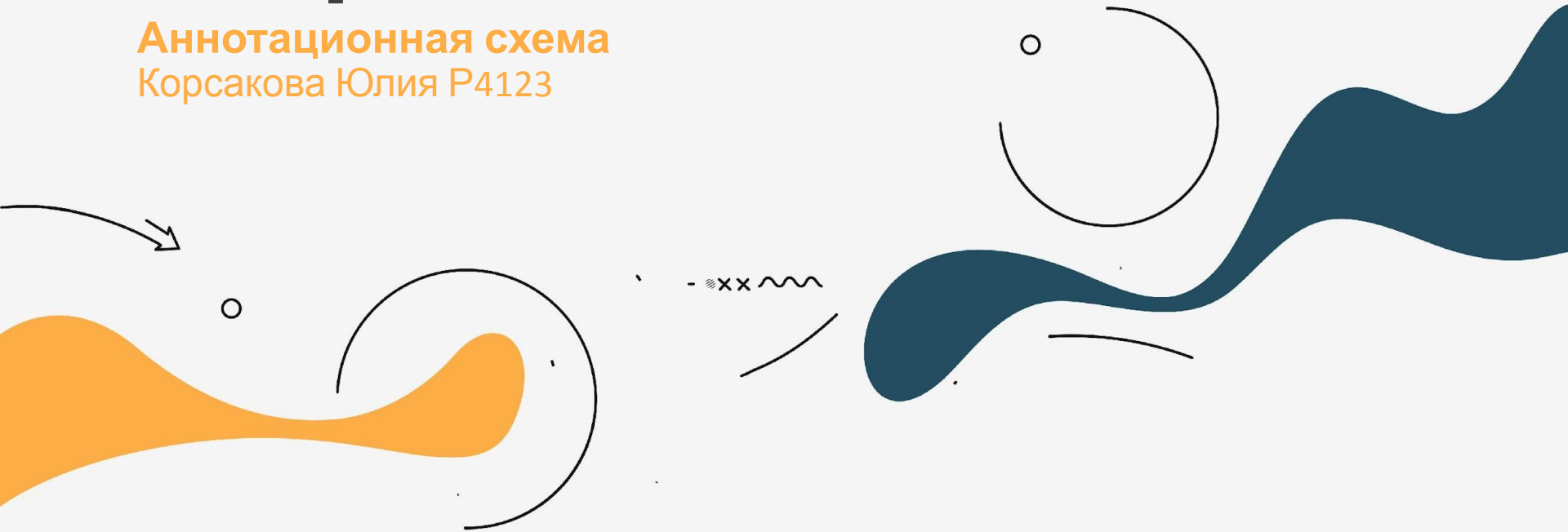
"Как можно было так поступить?! Это же откровенное предательство!"

Аудио



Диагностика ментального здоровья

Аннотационная схема
Корсакова Юлия Р4123



Аннотационная схема

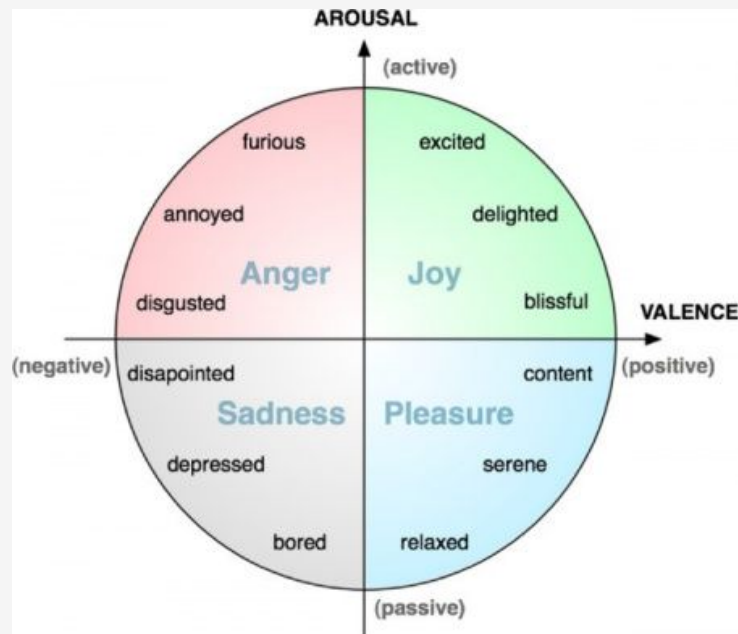
| Уровень разметки | Тип разметки | Модальность | Формат | Визуализация | ID |
|-------------------|---------------------------------------|-------------|----------|--------------|--------------------------------|
| Морфология | Морфологические признаки (части речи) | Текст | CoNNL-U | inception | word_id, phrase_id, emotion_id |
| Леммы | Леммы | Текст | CoNNL-U | inception | word_id, phrase_id, emotion_id |
| Тональность | rise/fall/high/low/mid | Аудио | TextGrid | Praat | phrase_id, emotion_id |
| Словоупотребления | Словоупотребления | Аудио | TextGrid | Praat | word_id, phrase_id, emotion_id |
| Громкость речи | low/mid/high | Аудио | TextGrid | Praat | phrase_id, emotion_id |

Структура данных

| speaker_id | audio_path | duration | emotion_id | text |
|------------|---|----------|------------|------------------------------|
| 0 | wavs/857b7099a4f5766105d166e2283066fa.wav | 4.4 | 5 | “Да, я сегодня был на улице” |

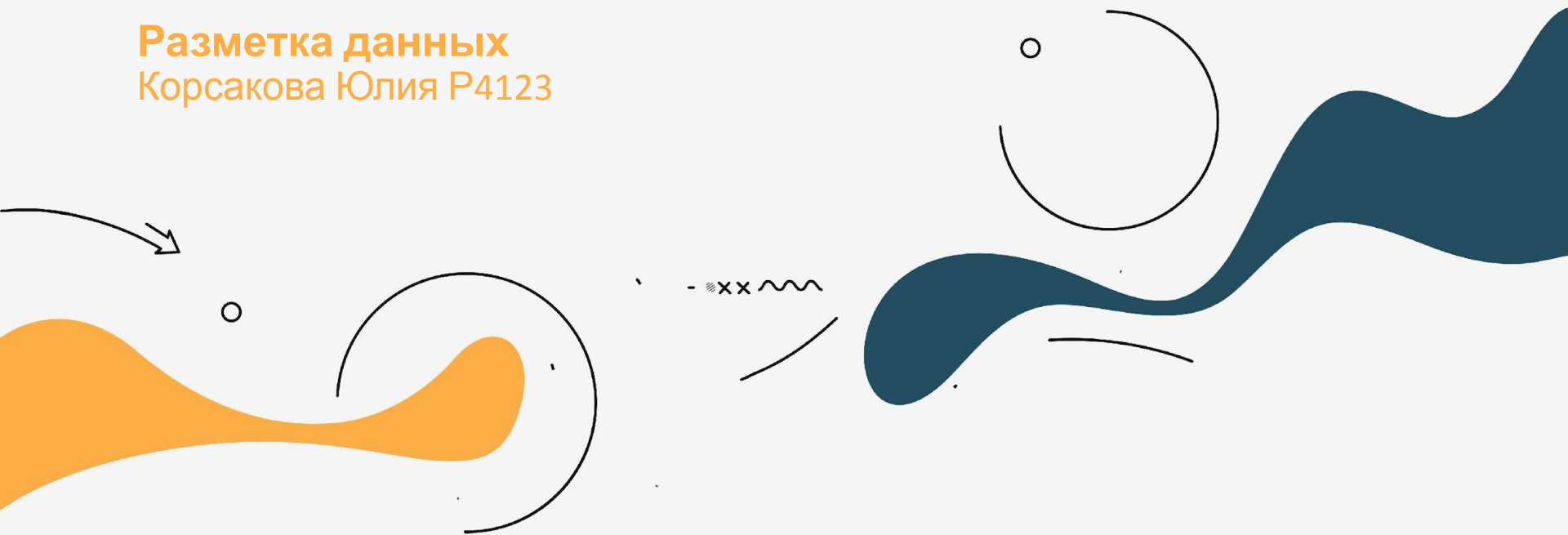
| emotion_id | emotion | valence | aurosal |
|------------|---------|---------|---------|
| 0 | joy | 1 | 1 |

Выделение эмоций



Распознавание эмоций

Разметка данных
Корсакова Юлия Р4123



Аннотационная схема

| Уровень разметки | Тип разметки | Модальность | Формат | Визуализация | ID |
|-------------------|---------------------------------------|-------------|----------|--------------|--------------------------------|
| Морфология | Морфологические признаки (части речи) | Текст | CoNNL-U | inception | word_id, phrase_id, emotion_id |
| Леммы | Леммы | Текст | CoNNL-U | inception | word_id, phrase_id, emotion_id |
| Тональность | rise/fall/high/low/mid | Аудио | TextGrid | Praat | phrase_id, emotion_id |
| Словоупотребления | Словоупотребления | Аудио | TextGrid | Praat | word_id, phrase_id, emotion_id |
| Громкость речи | low/mid/high/silence | Аудио | TextGrid | Praat | phrase_id, emotion_id |

Разметка данных. Текст

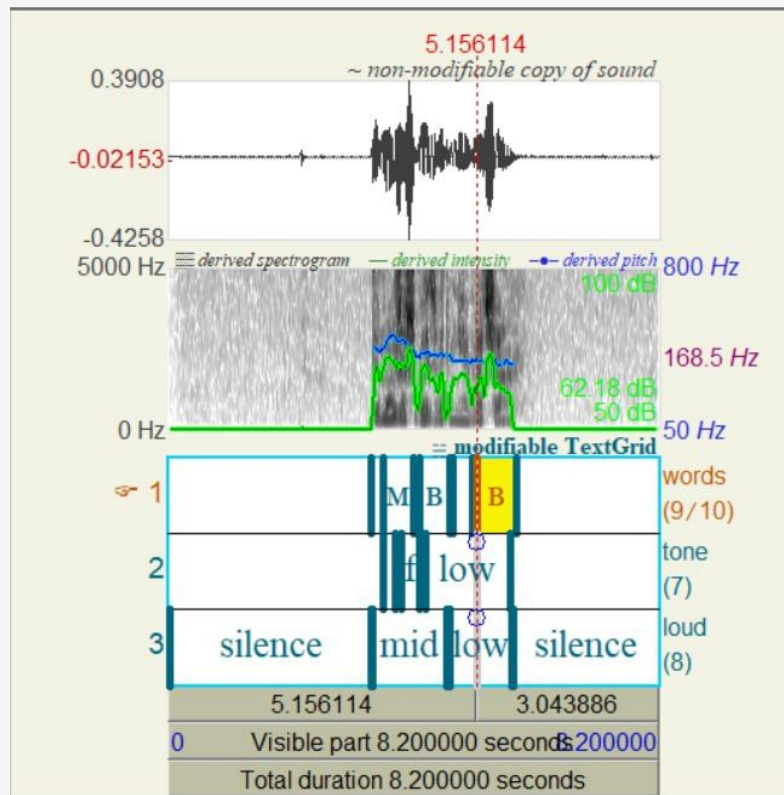
пример разметки в **INCEpTION**

| | | | | | | |
|---|-------------|-------------|-------------|-------------|-------------|-------------|
| | (MorFea...) | (MorFea...) | (MorFea...) | (MorFea...) | (MorFea...) | (MorFea...) |
| | да | у | нас | день | пройти | удачно |
| 1 | да | у | нас | день | прошёл | удачно |

результат разметки в формате **CoNLL-U**

| | | | | | | | | | |
|--------------------------------------|--------|--------|---|---|------|---|---|---|---|
| # text = да у нас день прошёл удачно | | | | | | | | | |
| 1 | да | да | - | - | PART | - | - | - | - |
| 2 | у | у | - | - | ADP | - | - | - | - |
| 3 | нас | нас | - | - | PRON | - | - | - | - |
| 4 | день | день | - | - | NOUN | - | - | - | - |
| 5 | прошёл | пройти | - | - | VERB | - | - | - | - |
| 6 | удачно | удачно | - | - | ADV | - | - | - | - |

Разметка данных. Аудио



Разметка данных в **Praat**.

3 уровня разметки:

- слова
- тональность
- громкость

Разметка данных. Аудио

результат разметки в формате **TextGrid**

```
item [1]:
  class = "IntervalTier"
  name = "words"
  intervals: size = 7
  intervals [1]:
    xmin = 3.388911375
    xmax = 3.5908773750000003
    text = "ты"
  intervals [2]:
    xmin = 3.5908773750000003
    xmax = 4.095792375
    text = "можешь"
  intervals [3]:
    xmin = 4.095792375
    xmax = 4.1967753750000005
    text = ""
  intervals [4]:
    xmin = 4.1967753750000005
    xmax = 4.701690375
    text = "вызвать"
  intervals [5]:
    xmin = 4.701690375
    xmax = 4.752181875000001
    text = ""
```

```
item [2]:
  class = "IntervalTier"
  name = "tone"
  intervals: size = 7
  intervals [1]:
    xmin = 0
    xmax = 3.5908773750000003
    text = ""
  intervals [2]:
    xmin = 3.5908773750000003
    xmax = 3.7928433750000003
    text = "rise"
  intervals [3]:
    xmin = 3.7928433750000003
    xmax = 3.893826375
    text = ""
  intervals [4]:
    xmin = 3.893826375
    xmax = 4.1967753750000005
    text = "fall"
  intervals [5]:
    xmin = 4.1967753750000005
    xmax = 4.297758375000001
    text = ""
```

```
item [3]:
  class = "IntervalTier"
  name = "loud"
  intervals: size = 4
  intervals [1]:
    xmin = 0.031226624999999838
    xmax = 3.388911375
    text = "silence"
  intervals [2]:
    xmin = 3.4141571250000005
    xmax = 4.651198875
    text = "mid"
  intervals [3]:
    xmin = 4.701690375
    xmax = 5.736766125000001
    text = "low"
  intervals [4]:
    xmin = 5.762011875000001
    xmax = 8.2
    text = "silence"
```

xx
||

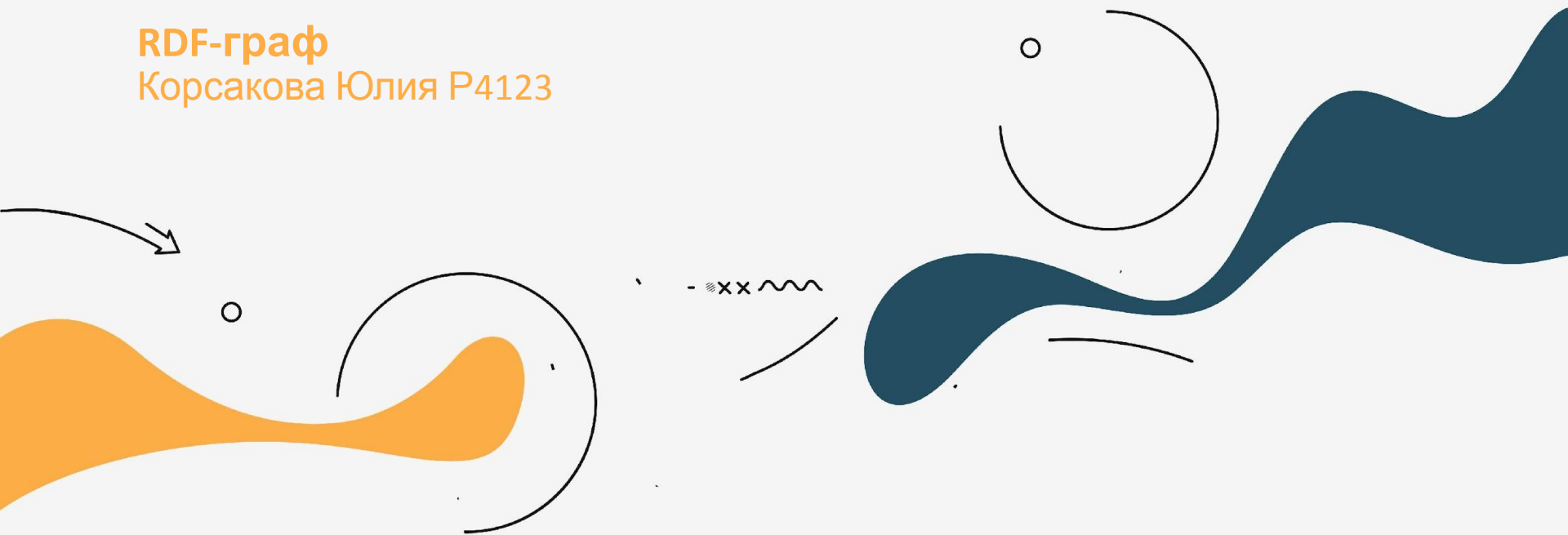
o

2

Распознавание эмоций

RDF-граф

Корсакова Юлия Р4123



Объединение данных в JSON

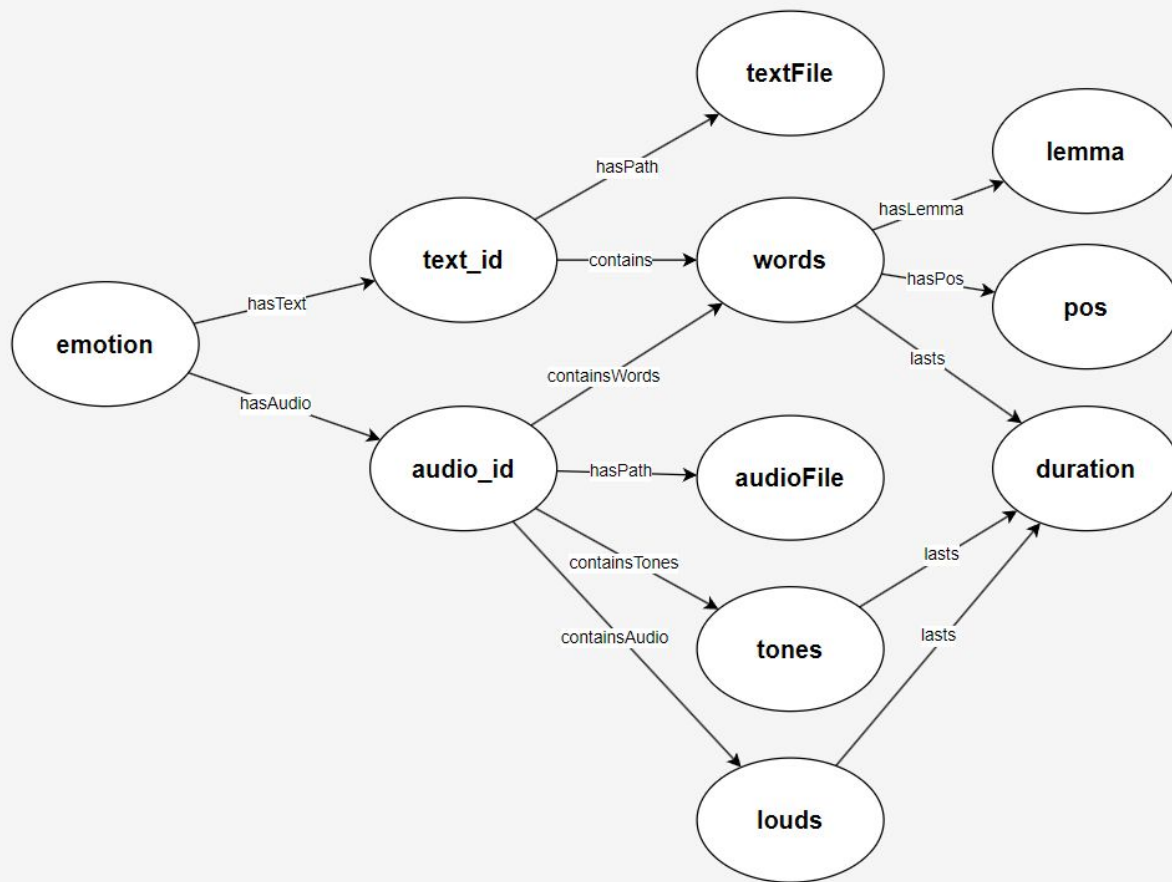
```
{
  "id": "pos_3",
  "emotion": "pos",
  "text_path": "data/data/pos_3.txt",
  "text_words": [
    {
      "form": "ну",
      "lemma": "ну",
      "pos": "INTJ"
    },
    {
      "form": "конечно",
      "lemma": "конечно",
      "pos": "ADV"
    },
    {
      "form": "умница",
      "lemma": "умница",
      "pos": "NOUN"
    },
    {
      "form": "моя",
      "lemma": "мой",
      "pos": "PRON"
    }
  ]
},
```

```
"audio_path": "data/data/pos_3.wav",
"audio_words": [
  {
    "label": "ну",
    "start": 1.67453,
    "end": 1.90384,
    "duration": 0.229
  },
  {
    "label": "конечно",
    "start": 1.90384,
    "end": 2.85382,
    "duration": 0.95
  },
  {
    "label": "умница",
    "start": 2.85382,
    "end": 3.60725,
    "duration": 0.753
  },
  {
    "label": "моя",
    "start": 3.60725,
    "end": 4.52447,
    "duration": 0.917
  }
],
```

```
"tones": [
  {
    "label": "high",
    "start": 1.69091,
    "end": 2.70641,
    "duration": 1.016
  },
  {
    "label": "rise",
    "start": 2.98485,
    "end": 3.19778,
    "duration": 0.213
  },
  {
    "label": "mid",
    "start": 3.19778,
    "end": 4.37706,
    "duration": 1.179
  }
],
```

```
"louds": [
  {
    "label": "silence",
    "start": 0.02026,
    "end": 1.64178,
    "duration": 1.622
  },
  {
    "label": "high",
    "start": 1.69091,
    "end": 4.44258,
    "duration": 2.752
  },
  {
    "label": "silence",
    "start": 4.44258,
    "end": 5.32,
    "duration": 0.877
  }
],
```

Структура графа



Пример запроса

Количество тональностей класса *high* в каждой эмоции

```
PREFIX ex: <http://example.org/schema#>
```

```
SELECT ?emotion (COUNT(?tone) AS ?high_count)
WHERE {
    ?emotion ex:hasAudio ?audio .
    ?audio ex:containsTones ?tone .
    ?tone ex:value "high" .
}
GROUP BY ?emotion
```

| | neg | neut | pos |
|-------|-----|------|-----|
| count | 5 | 3 | 6 |

GitHub



Dataset

