

# Документация к API «Голос Атома»

v.1.0



Команда ИИрокезы для кейса «Росатом»

Улучшение представлений результатов в сервисе "Мой голос"

2023

## Оглавление

1. Краткое описание .....	3
2. Кластеризация, эмоциональность, проверка орфографии, схожесть фраз ...	4
2.1 Схожесть фраз в сравнении с целевой фразой.....	4
2.2 Эмоциональная окраска для массива входящих фраз.....	5
2.3 Объединение фраз с учетом проблематики открытых вопросов .....	6
2.4 Объединение фраз с учетом проблематики открытых вопросов (быстрое)	7
2.5 Проверка орфографии.....	7
2.6 Изменение неправильной раскладки.....	8
3. Аналитика и статистика, работа с загрузкой файлов опросов .....	9
3.1 Загрузка файла.....	9
3.3 Топ позитивных ответов в вопросах во всем опросе (в загруженном файле) .....	9
3.4 Эмоциональное разделение ответов в вопросах .....	10
3.5 Вопросы, в которых ответы не связаны с тематикой вопроса .....	12
4. Работа со стоп словами, нецензурные выражения .....	14
4.1 Проверка одного слова .....	14
4.2 Проверка фразы .....	14
4.3 Проверка массива фраз.....	15
4.4 Количество ответов с нецензурными словами в опросе или вопросе..	16

## 1. Краткое описание

Голос Атома – программный модуль обработки данных опросов сервиса «Мой голос» с собственным API для возможностью интеграции с другими системами. Ответы на открытые вопросы проходят проверку на орфографию, раскладку и блок нецензурных слов.

Использование предобученной модели позволяет быстро кластеризовать ответы с учетом тематика вопроса, выделять эмоциональную составляющую и визуализировать статистику опроса более чем по 20 параметрам.

Посмотреть реализованные сервис можно по ссылкам ниже (при невыполнении запроса нужно смотреть коды ошибок):

<https://myvote-analytics.kovalev.team/api/docs#/>

<https://myvote-api.kovalev.team/swagger-ui/index.html#>

<https://myvote-profanity.kovalev.team/docs#>

## 2. Кластеризация, эмоциональность, проверка орфографии, схожесть фраз

### 2.1 Схожесть фраз в сравнении с целевой фразой

/api/similarity-sentences/ – схожесть фраз в сравнении с целевой фразой

На вход подается целевая фраза и список морфем для сравнения

Пример запроса:

```
curl -X 'POST' \
  'https://myvote-analytics.kovalev.team/api/similarity-sentences/' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "source_sentence": "Погода хорошая и теплая",
    "sentences": [
      "Погода хорошая",
      "Пушкин русский поэт",
      "Сегодня не холодно"
    ]
  }'
```

Пример ответа:

```
{
  "source_sentence": "Погода хорошая и теплая",
  "sentences": [
    [
      0.9376865029335022,
      "Погода хорошая"
    ],
    [
      0.4102936387062073,
      "Пушкин русский поэт"
    ],
    [
      0.7974203824996948,
      "Сегодня не холодно"
    ]
  ]
}
```

## 2.2 Эмоциональная окраска для массива входящих фраз

/api/emotions/ – эмоциональная окраска для массива входящих фраз

На вход подается список морфем

Пример запроса:

```
curl -X 'POST' \
  'https://myvote-analytics.kovalev.team/api/emotions/' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "texts": [
      "очень круто",
      "круто",
      "очень"
    ]
  }'
```

Пример ответа:

```
{
  "emotions": [
    [
      "positive",
      "очень круто",
      {
        "neutral": 0.006071935407817364,
        "positive": 0.9889675974845886,
        "negative": 0.0049604689702391624
      }
    ],
    [
      "positive",
      "круто",
      {
        "neutral": 0.0049170246347785,
        "positive": 0.9922296404838562,
        "negative": 0.002853366080671549
      }
    ],
    [
      "neutral",
      "очень",
      {
```

```

        "neutral": 0.726581335067749,
        "positive": 0.031829312443733215,
        "negative": 0.24158942699432373
    }
]
]
}

```

## 2.3 Объединение фраз с учетом проблематики открытых вопросов

/api/clustering/ – объединение фраз с учетом проблематики открытых вопросов

На вход подается список фраз, результатов будет соотношения между каждой из них

Привет запроса:

```

curl -X 'POST' \
  'https://myvote-analytics.kovalev.team/api/clustering/' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "texts": [
      "Веселая погода играла красками",
      "Пушкин русский поет",
      "Погода была хорошая",
      "Поэты России - лучшие",
      "Направление ветра менялось очень быстро",
      "Пушкин жил в Петербурге",
      "Самолет летит в зал",
      "Виталия все еще не спал"
    ]
  }'

```

Пример ответа:

```

{
  "clusters": {
    "0": [
      "Пушкин русский поет",
      "Поэты России - лучшие",
      "Пушкин жил в Петербурге"
    ]
  }
}

```

```

],
"1": [
    "Виталья все еще не спал"
],
"2": [
    "Веселая погода играла красками",
    "Погода была хорошая"
],
"3": [
    "Самолет летит в зал"
],
"4": [
    "Направление ветра менялось очень быстро"
]
}
}

```

## 2.4 Объединение фраз с учетом проблематики открытых вопросов (быстрое)

/api/clustering/fast - быстрая обработка, минимум 25 элементов, максимум – 176. Запрос аналогичен пункту 2.3.

## 2.5 Проверка орфографии

/api/spellcheck/ - проверка орфографии

На вход подается список фраз, если в них встречаются орфографические ошибки, они будут исправлены

```

curl -X 'POST' \
  'https://myvote-analytics.kovalev.team/api/spellcheck/' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '
{
  "texts": [
    "дела",
    "дила",
    "не хвтило тплива",

```

```

        "солнце",
        "нрмально"
    ]
}

```

Пример ответа:

```

{
  "spells": [
    "дела",
    "Дела",
    "Не хватило топлива",
    "солнце",
    "Нормально"
  ]
}

```

## 2.6 Изменение неправильной раскладки

api/v1/lang-checker – изменение неправильной раскладки

Пример запроса:

```

curl -X 'POST' \
  'https://myvote-api.kovalev.team/api/v1/lang-checker' \
  -H 'accept: */*' \
  -H 'Content-Type: application/json' \
  -d '{
    "message": "hello цщкдв руддщ world"
  }'

```

Пример ответа:

```

{
  "tokens": [
    "hello",
    "world",
    "hello",
    "world"
  ],
  "message": "hello world hello world"
}

```



### 3. Аналитика и статистика, работа с загрузкой файлов опросов

#### 3.1 Загрузка файла

/api/v1/data/upload – загрузка файла в папке .zip, содержащего набор json с результатами опроса

Пример запроса:

```
curl --location --request POST 'https://myvote-analytics.kovalev.team/api/v1/data/upload' \
--form 'file=@"/Users/User/Desktop/Архив.zip"'
```

#### 3.3 Топ позитивных ответов в вопросах во всем опросе (в загруженном файле)

/api/v1/analytics/question/top-positive - топ позитивных ответов в вопросах во всем опросе (в загруженном файле), передается id файла, требуемая эмоция и топ в лимите

**analytic-controller**

**GET** /api/v1/analytics/question/top-positive

**Parameters**

Name	Description
<b>file_id</b> * required integer(\$int64) (query)	<input type="text" value="5"/>
<b>emotional</b> * required string (query)	<input type="text" value="positive"/>
<b>limit</b> * required integer(\$int32) (query)	<input type="text" value="10"/>

**Execute**

Пример запроса:

```
curl -X 'GET' \
```

```
'https://myvote-  
api.kovalev.team/api/v1/analytics/question/top-  
positive?file_id=5&emotional=positive&limit=10' \  
-H 'accept: */*'
```

Пример ответа:

```
[  
  {  
    "name": "Как ваше настроение?",  
    "count": 578  
  },  
  {  
    "name": "Какое настроение у Вас сегодня?",  
    "count": 324  
  },  
  {  
    "count": 304,  
    "name": "Как ваше настроение сегодня?"  
  },  
  {  
    "name": "Что позволяет вам лично поддерживать уверенность и рабочий  
настрой?",  
    "count": 284  
  },  
  {  
    "name": "Какой лидер актуален сегодня для вас?",  
    "count": 224  
  },  
  {  
    "count": 158,  
    "name": "Ваши ожидания от предстоящего мероприятия одним словом"  
  },  
  {  
    "name": "За счет каких моих качеств и качеств моей команды нам удалось  
достичь побед?&nbsp;",  
    "count": 156  
  },  
  {  
    "name": "Как настроение?",  
    "count": 150  
  },  
  {  
    "name": "Какие большие цели сейчас вдохновляют лично вас?",  
    "count": 148  
  },  
  {  
    "count": 120,  
    "name": "Ваше настроение - с какими эмоциями пришли на конференцию.  
Одним-двумя словами"  
  }  
]
```

### 3.4 Эмоциональное разделение ответов в вопросах

/api/v1/analytics/answer/emotional-distribution - эмоциональное разделение  
ответов в вопросах

Передается id файла, лимит вывода топа вопросов и целевая эмоция.

В ответ генерируется топ вопросов, в которых больше всего было негативных ответов и выводится процентное соотношение негативных, позитивных и нейтральных отзывов в этом вопросе к общему количеству негативных, позитивных и нейтральных отзывов.

Name	Description
<b>file_id</b> * required integer(\$int64) (query)	<input type="text" value="5"/>
<b>limit</b> * required integer(\$int32) (query)	<input type="text" value="5"/>
<b>number_order</b> * required string (query)	<input type="text" value="negative"/>

Execute

Пример запроса:

```
curl -X 'GET' \
  'https://myvote-
api.kovalev.team/api/v1/analytics/answer/emotional-
distribution?file_id=5&limit=5&number_order=negative' \
  -H 'accept: */*'
```

Пример ответа:

```
[
  {
    "negative": 8.42,
    "name": "Напишите свою ассоциацию со словом
\"бюрократ\"",
    "count": 938,
    "neutral": 3.67,
```

```

    "positive": 0.42
  },
  {
    "negative": 4.21,
    "name": "Вопрос 5. В чем, по Вашему мнению, причина
роста травматизма?",
    "count": 452,
    "neutral": 1.78,
    "positive": 0.13
  },
  {
    "negative": 2.98,
    "name": "Опишите одним словом ваше текущее
состояние",
    "count": 287,
    "neutral": 0.94,
    "positive": 0.88
  },
  {
    "negative": 2.81,
    "name": "5 слов, которые характеризуют произошедшее с
вами с прошедшей конференции?",
    "count": 164,
    "neutral": 0.58,
    "positive": 0.18
  },
  {
    "negative": 2.63,
    "name": "Опишите одним словом ваше эмоциональное
состояние в настоящий момент",
    "count": 149,
    "neutral": 0.52,
    "positive": 0.2
  }
]

```

### 3.5 Вопросы, в которых ответы не связаны с тематикой вопроса

/api/v1/analytics/answer/incorrect-answer-top - вопросы, в которых ответы не связаны с тематикой вопроса

На вход подается id загруженного файла с вопросами, процент несовпадения ответа с вопросом и лимит вывода

Пример запроса:

```
curl -X 'GET' \
  'https://myvote-
api.kovalev.team/api/v1/analytics/answer/incorrect-
answer-top?file_id=14&percent=0.7&limit=10' \
  -H 'accept: */*'
```

Пример ответа:

```
[
  {
    "question_name": "Ваши ожидания от Карьерного
трека?",
    "non_matching_answers": 2,
    "total_answers": 83
  },
  {
    "non_matching_answers": 12,
    "question_name": "Какой вклад вы готовы внести?",
    "total_answers": 68
  },
  {
    "question_name": "Как началось ваше утро?",
    "total_answers": 57,
    "non_matching_answers": 17
  },
  {
    "total_answers": 47,
    "question_name": "Кто такой амбассадор?",
    "non_matching_answers": 1
  },
  {
    "total_answers": 44,
    "question_name": "Какие дополнительные предложения от
столов необходимо включить в программу?",
    "non_matching_answers": 2
  }
]
```

## 4. Работа со стоп словами, нецензурные выражения

### 4.1 Проверка одного слова

/sensor-word/ - проверка одного слова

Пример запроса:

```
curl -X 'POST' \
  'https://myvote-profanity.kovalev.team/censor-
word/%D1%81%D0%B0%D0%BC%D0%BE%D0%BB%D0%B5%D1%82' \
  -H 'accept: application/json' \
  -d ''
```

Пример ответа:

```
{
  "uncensored": "самолет",
  "censored": "самолет",
  "original_profane_word": null,
  "bad_words_percent": 0
}
```

### 4.2 Проверка фразы

/sensor-words – проверка фразы

Пример запроса

```
curl -X 'POST' \
  'https://myvote-profanity.kovalev.team/censor-words' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "text": "погода плохая просто shit"
  }'
```

Пример ответа:

```
{
  "uncensored": "погода плохая просто shit",
  "censored": "погода плохая просто ****",
  "original_profane_word": null,
  "bad_words_percent": 0
}
```

### 4.3 Проверка массива фраз

/censor-words/batch – проверка массива фраз

Пример запроса:

```
curl -X 'POST' \
  'https://myvote-profanity.kovalev.team/censor-words/batch' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "texts": [
      "плохая погода",
      "погода просто shit",
      "было фигово, но все поправили"
    ]
  }'
```

Пример ответа:

```
[
  {
    "uncensored": "плохая погода",
    "censored": "плохая погода",
    "original_profane_word": null,
    "bad_words_percent": 0
  },
  {
    "uncensored": "погода просто shit",
    "censored": "погода просто ****",
    "original_profane_word": null,
    "bad_words_percent": 0.3333333333333333
  },
  {
    "uncensored": "было фигово, но все поправили",
    "censored": "было фигово, но все поправили",
    "original_profane_word": null,
    "bad_words_percent": 0
  }
]
```

#### 4.4 Количество ответов с нецензурными словами в опросе или вопросе

/api/v1/analytics/censored/stat - количество нецензурных ответов с нецензурными словами в опросе или вопросе. На вход подается id файла с опросов (загруженный .zip) и номер вопроса (не обязательный параметр).

Пример запроса:

```
curl -X 'GET' \
  'https://myvote-
api.kovalev.team/api/v1/analytics/censored/stat?file_id=5' \
  -H 'accept: */*'
```

Пример ответа:

17