



УСТАНОВКА И ЭКСПЛУАТАЦИЯ
ПРОГРАММНОГО КОМПЛЕКСА «МИП»



ОГЛАВЛЕНИЕ

УСТАНОВКА	3
ЭКСПЛУАТАЦИЯ	7

Структура проекта

```

├── docker-compose.yml
├── README.md
├── services
│   ├── backend
│   │   ├── db_checker.py
│   │   ├── Dockerfile
│   │   ├── manage.py
│   │   ├── project
│   │   │   ├── api
│   │   │   │   ├── bp_v1.py
│   │   │   │   ├── models.py
│   │   │   │   ├── resources
│   │   │   │   └── utils.py
│   │   │   ├── config.py
│   │   │   ├── db
│   │   │   ├── __init__.py
│   │   │   ├── libs
│   │   │   ├── tasks.py
│   │   │   └── templates
│   │   ├── requirements.txt
│   │   └── scripts
│   ├── consumer
│   │   ├── app
│   │   │   ├── config.py
│   │   │   ├── requirements.dev.txt
│   │   │   ├── schemas.py
│   │   │   └── utils
│   │   └── databus

```

```
| | | | └── consumer.py
| | | └── storage
| | └── __init__.py
| └── Dockerfile
| └── entrypoint.sh
| └── requirements.txt
| └── run_consumer.py
| └── start
└── nginx
└── pgadmin
└── postgresql
└── rabbitmq
└── redis
└── tests
```

Выполните в bash следующие команды

```
git clone https://git.rdecs.org/rdecs/analyzer2.git
cd analyzer
cat .env-example > .env
docker-compose up -d --build
```

Проверка статуса установки

```
user@local ~/analyzer # docker-compose ps
```

Примерный ответ

NAME	COMMAND	SERVICE	STATUS	PORTS
analyzer-analyzer_celery-1	"run_celery"	analyzer_celery	running	5000/tcp
analyzer-analyzer_consumer-1	"/usr/src/app/entryp..."	analyzer_consumer	running	
analyzer-analyzer_db-1	"docker-entrypoint.s..."	analyzer_db	running	0.0.0.0:6532->5432/tcp,
:::6532->5432/tcp				
analyzer-analyzer_nginx-1	"/docker-entrypoint..."	analyzer_nginx	running	0.0.0.0:9098->80/tcp,
:::9098->80/tcp				
analyzer-analyzer_redis-1	"docker-entrypoint.s..."	analyzer_redis	running	6379/tcp
analyzer-analyzer_web_service-1	"run_app"	analyzer_web_service	running	5000/tcp
analyzer-pgadmin-1	"/entrypoint.sh"	pgadmin	running	443/tcp, 0.0.0.0:5052->
80/tcp, :::5052->80/tcp				
analyzer-rabbit-1	"docker-entrypoint.s..."	rabbit	running	4369/tcp, 5671-5672/tc
p, 15671/tcp, 15691-15692/tcp, 25672/tcp, 0.0.0.0:15672->15672/tcp, :::15672->15672/tcp				

Установка СУБД PGAdmin

1. Открыть в браузере localhost:5052 или <адрес сервера>:5052
 Логин пароль из .env-example
 Tools -> Import/Export Servers
 Загрузить (upload) файл services/pgadmin/servers.json
2. Указать загруженный файл для импорта (затем Next)
3. Отметить галкой Servers (затем Next)
4. Нажать Finish
5. Ввести пароль из .env-example (значение переменной POSTGRES_PASSWORD)

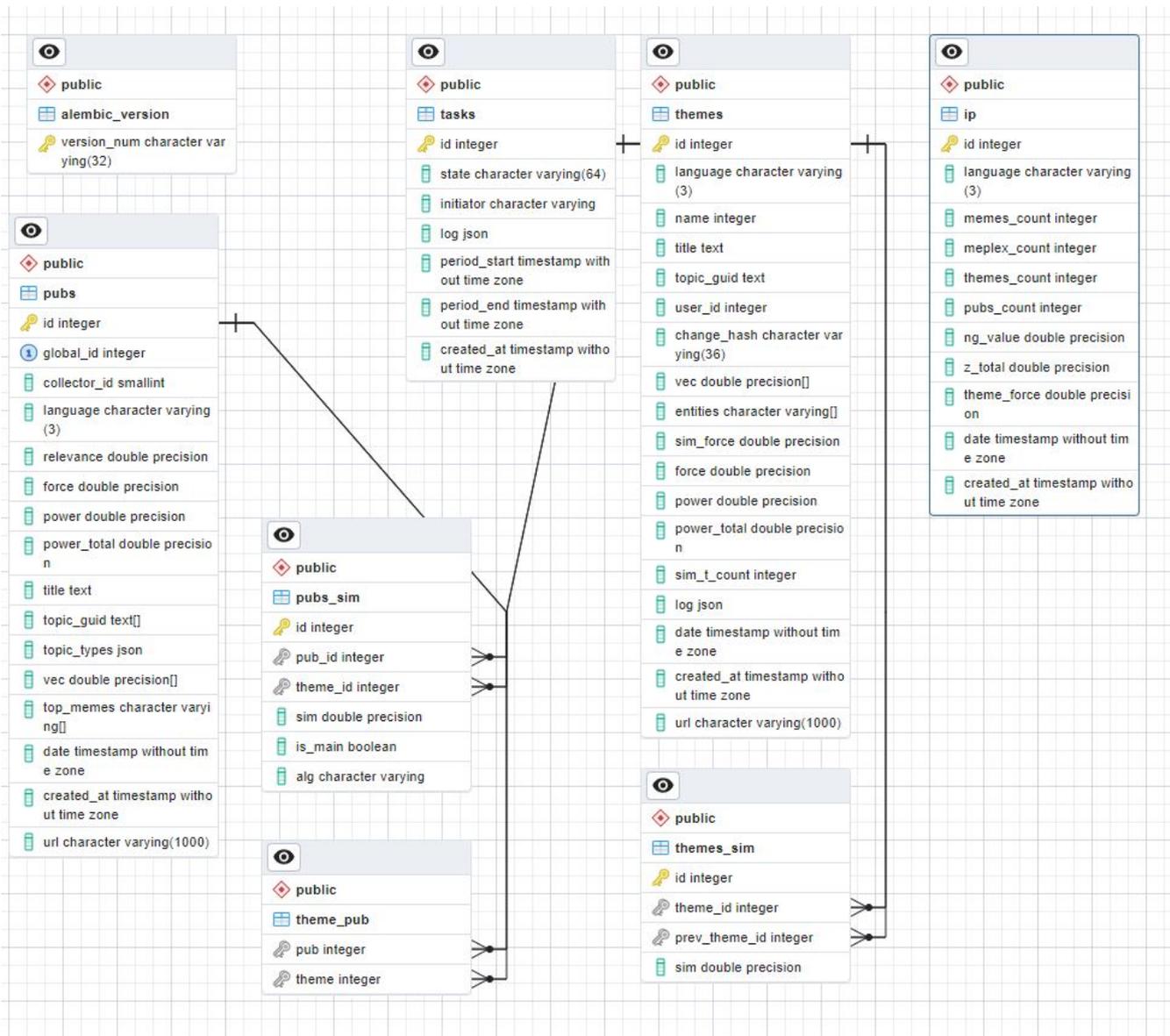


Рисунок. 1 Структура базы данных

Эксплуатация

Все необходимые настройки для введения ПО в эксплуатацию обсуждаются с пользователем индивидуально.

Ниже представлен список возможных функций программы

- Сбор (получение) публикаций от различных поставщиков данных
- Структурирование исходных данных с помощью набора NLPмоделей и обогащение их дополнительной информацией
- Классификация публикаций по пользовательским тематикам
- Динамическая кластеризация похожих публикаций в темы
- Прогноз популярности тем публикаций

В результате работы программы, подсчитанные значение силы инф. воздействия, агрегированные данные по темам и тематикам предоставляются как единая структура формата JSON. Каждый блок хранит свой список, содержащий «ключ: значение». Например, блок pubs содержит информацию о публикациях, поля списка их описывают более подробно. Детально значение полей и типового JSON приведено ниже.

```
{
  "theme_id": 12376979, # id темы
  "title": "...", Заголовок темы
  "species": {}, # класс тематики к которой относится данная тема (таблица species) структуру выбрать исходя из удобства, отображает пиктограмму класса и название класса (таблица species с заранее определенными пиктограммами и классами)
  "reach": 3, # охват темы (поле reach таблица themes),
  "thematic": {"thematic_uuid": "uuid", "name": "name 1"}, # тематика (таблица thematics)
  "relevance": 0.1, # релевантность темы
  "estimation": 1.0 # оценка по индикаторам (определяет цвет иконки класса),
  "geo": ['Россия', 'Москва'], # география (параметр geo, таблица themes)
  "user_thematics": [{"thematic_uuid": "uuid", "name": "name 1"}, ... ], # список тематик пользователя (таблица thematics)
  "main_source": {
    "url": "https://..." # первоисточник (url основной публикации темы, themes_pubs -> is_main),
    "rate": 5,
    "subscribers": 1200000,
  }
}
```

```

"mean_pubs_per_day": 100
},
"entities": {
  "org": [
    {"name": "org name", "emo": 0.1, "indicator": 1.0},
    {"name": "org name", "emo": 0.1, "indicator": -1.0},
  ],
  "per": [
    {"name": "per name", "emo": 0.5, "indicator": 0},
    {"name": "per name", "emo": 0.3, "indicator": 1.0},
  ],
  "num": [
    {"name": "2014 год", "emo": 0.0, "indicator": 0},
    {"name": "27 руб", "emo": 0.0, "indicator": -1.0},
  ]
},
"pubs": [
  "date": "2022-10-11 00:00:00", # время публикации
  "title": "Заголовок",
  "text": "Текст",
  "language": "RU",
  "source": "tsn.com/ru", # источник публикации (имя сайта или тг канала, выделяется из url)
  "url": "https://...", # ссылка на публикацию (исходная)
  "media": [ # картинки к публикации (прямая ссылка на глобальное хранилище с картинками. Тип
задаем на перспективу, в дальнейшем там может быть еще и видео или аудио)
    {"type": "img", "src": "https://..."},
    {"type": "img", "src": "https://..."},
    {"type": "img", "src": "https://..."}
  ],
  "highlights": [ # Координаты ключевых слов в тексте публикации, которые необходимо подсветить
при клике на "Показать ключевые слова"
    {"start": 15, "end": "20", "name": "phrase 1"}, {"start": 125, "end": "132", "name": "phrase 2"}
  ]
]
}

```