

ООО «БИЭЙ»

**TAMARA**  
**(Text Adorable Modular Awesome Reliable Analysis)**  
**Инструкция по установке**

## Оглавление

<b>Общая информация.....</b>	<b>3</b>
Функциональные возможности Системы.....	3
Эксплуатационное назначение Системы .....	3
<b>Установка .....</b>	<b>3</b>
Минимальные требования к аппаратному обеспечению .....	3
Требования к программному обеспечению .....	3
Порядок установки .....	4
<b>Описание настроек .....</b>	<b>4</b>
/apps/biei/tamara/docker-compose.yml .....	4
/apps/biei/tamara/config.properties.....	4
/apps/biei/fscrawler/test_job/_settings.yml.....	5
<b>Проверка работоспособности .....</b>	<b>5</b>

## Общая информация.

Основное назначение Системы в анализе текстовой информации и соотнесении документов на основании их контента к словарному значению. Архитектура системы предполагает возможность дальнейших шагов в анализе документов их кластеризации, группировки и т.д.

### Функциональные возможности Системы

- Генерация n-gramm для документа
- Изменение документов в ElasticSearch
- Создание словарей
- Добавление существующих словарей
- Подключение индексов с документами
- Создание «привязок» документов к словарям
- Тэгирование документов на основе анализа их n-gramm
- Работа в фоновом режиме
- Сбор и текстовая визуализация статистики по работе Системы
- Логирование

### Эксплуатационное назначение Системы

Система предназначена для следующих категорий конечных пользователей:

- Специалист по анализу данных
- Аналитик данных
- ГИС инженер

### Установка

Дистрибутив имеет следующий состав:

- docker-образ приложения<sup>1</sup>
- архив с демонстрационными данными и первичной настройкой

Данную информацию можно получить и скачать с сайта [www.bieidb.com](http://www.bieidb.com). Раздел <https://bieidb.com/instructions/tamara>

### Минимальные требования к аппаратному обеспечению

Для минимальной установки и работы аппаратные требования указаны ниже:

Свободное место: от 4ГБ

Кол-во процессоров/ядер: от 4 ядер

Доступный объем оперативной памяти: 6ГБ

При увеличении количества обрабатываемых документов, необходима дополнительная доступная оперативная память.

### Требования к программному обеспечению

Необходимо следующее программное обеспечение

---

<sup>1</sup> Кроме того, доступна установка непосредственно в инфраструктуру. Для этого необходимо связаться с нами, за получением дополнительных инструкций и дистрибутива.

- Операционная система: Linux (CentOS 7+, RHEL 7+)
- Установленные пакеты docker (20.10.2) и docker-compose (1.28.2). Для установки необходимо обратиться к инструкции, расположенной на сайте производителя <https://docs.docker.com/get-docker/>

#### Порядок установки

#### Порядок установки следующий:

1. Необходимо скачать архив с данными и настройками.
2. Создать каталог /apps/.
3. Скопировать содержимое архива в каталог /apps/.
4. Необходимо предоставить права на эту папку пользователю, под которым будут запускаться docker контейнеры.
5. В каталоге /apps/biei/tamara находится файл docker-compose.yml. Он содержит конфигурацию для запуска докер контейнеров.
6. Запускаем контейнеры из каталога /apps/biei/tamara/. Команда запуска docker-compose up -d<sup>2</sup>
7. Если приложение необходимо остановить, то для этого нужно остановить все запущенные контейнеры, выполнив команду: docker-compose down

#### Описание настроек

[/apps/biei/tamara/docker-compose.yml](#)

Основной файл с описанием docker-образов.

Особые настройки образов:

Tamara:

- Конфигурация для Java через переменную окружения JAVA\_OPTS
- Подключаемый каталог /apps/biei/tamara с конфигурацией для приложения
- Мэпинг портов для доступа к приложению. По умолчанию 8080:8080

Kibana:

- Мэпинг портов для доступа к приложению (5656:5601)
- Адрес Elasticsearch

FSCrawler:

- Подключаемые каталоги с конфигураций и файлами для обработки
- Специальная команда запуска, ожидающая успешный запуск Elasticsearch

Postgresql:

- Мэпинг портов 5432:5432
- Подключаемый каталог с данными для БД
- Пользователь и пароль для подключения к БД

ElasticSearch:

- Настройки запуска ElasticSearch через определение окружения
- Мэпинг портов 9292:9200
- Подключаемый каталог с данными

[/apps/biei/tamara/config.properties](#)

Конфигурационный файл для приложения

---

<sup>2</sup> С более подробной информацией по запуску docker-compose, можно ознакомиться на ресурсе <https://docs.docker.com/compose/>

ext.url – путь к БД для хранения внутренней конфигурации приложения (UI + точки хранения)

org.quartz.jobStore.\* - набор настроек для планировщика

org.quartz.dataSource.tamaraDS.\* - набор настроек для хранения данных планировщика

/apps/biei/fscrawler/test\_job/\_settings.yml

Преднастроенная конфигурация для сканирования файлов и извлечения контента.

Полное описание конфигурации можно найти

[https://fscrawler.readthedocs.io/en/latest/user/getting\\_started.html](https://fscrawler.readthedocs.io/en/latest/user/getting_started.html)

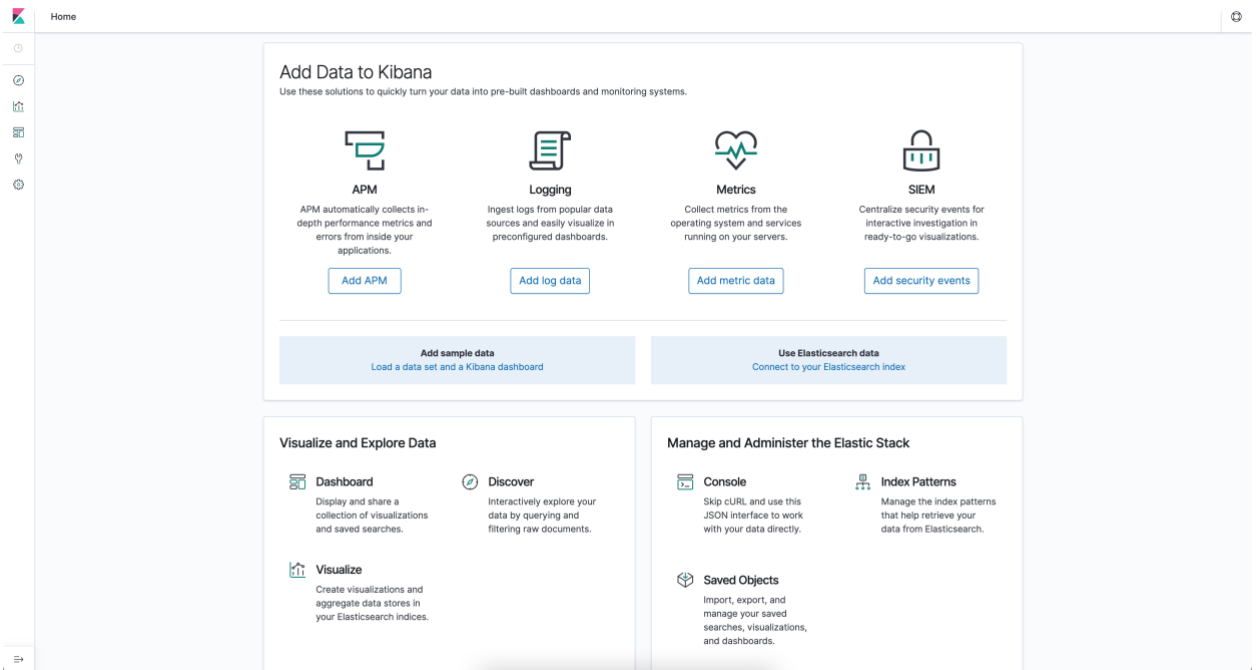
### Проверка работоспособности

После успешного запуска контейнеров с приложением на вашем СЕРВЕРЕ (адрес сервера – <SRV\_ADDR>), должны быть доступны следующие ссылки

[http://<SRV\\_ADDR>:9292](http://<SRV_ADDR>:9292) - адрес Elasticsearch

```
{
  "name" : "esearch",
  "cluster_name" : "es-tamara-docker-cluster",
  "cluster_uuid" : "sRZXOFNaTvCPzLa2qTvw2Q",
  "version" : {
    "number" : "7.3.1",
    "build_flavor" : "oss",
    "build_type" : "docker",
    "build_hash" : "4749ba6",
    "build_date" : "2019-08-19T20:19:25.651794Z",
    "build_snapshot" : false,
    "lucene_version" : "8.1.0",
    "minimum_wire_compatibility_version" : "6.8.0",
    "minimum_index_compatibility_version" : "6.0.0-beta1"
  },
  "tagline" : "You Know, for Search"
}
```

[http://<SRV\\_ADDR>:5656](http://<SRV_ADDR>:5656) - адрес Kibana



Проверка доступа к БД PostgreSQL:  
telnet <SRV\_ADDR> 5432

[http://<SRV\\_ADDR>:8080](http://<SRV_ADDR>:8080) - адрес Тамара (имя пользователя/пароль - user/password)

Первый запуск Тамара потребует ввода лицензионного ключа. При открытии приложения появится форма с информацией о способе его получения.

First, copy the code and send it to [info@bieidb.com](mailto:info@bieidb.com) for getting the serial number for application. Then insert this serial number to the next text field and press Activate.

Code  
bfebfbff00040661

Serial number  
[Empty field]

Activate

Сформированный код необходимо отправить на указанный адрес электронной почты: [info@bieidb.com](mailto:info@bieidb.com). В ответ вы получите сгенерированную лицензию в виде серийного номера. Его необходимо вставить в соответствующий элемент формы, нажать кнопку Activate и обновить страничку.

First, copy the code and send it to [info@bieidb.com](mailto:info@bieidb.com) for getting the serial number for application. Then insert this serial number to the next text field and press Activate.

Code  
bfebfbff00040661

Serial number  
QsZXYXpKfb6lzN/HouI

Activate

В результате, вы перейдете на форму ввода имени пользователя и пароля. Последующий ввод лицензии не потребуется.

Логин/пароль по-умолчанию: user/password

# TAMARA

(Text Adorable Modular Awesome  
Reliable Analysis)

## Log in

Username •

Password •



Log in

По все вопросам или ошибкам необходимо обращаться по адресу [info@bieidb.com](mailto:info@bieidb.com).