

Документация, содержащая описание функциональных характеристик экземпляра программного обеспечения, предоставленного для проведения экспертной проверки

**Программа для ЭВМ
«Система анализа и мониторинга событий»**

Оглавление

1. ОБЩИЕ СВЕДЕНИЯ	3
1.1 Наименование программы	3
1.2 Используемые языки программирования	3
1.3 Системные требования	3
2. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ	4

1. ОБЩИЕ СВЕДЕНИЯ

1.1 Наименование программы

Программа для ЭВМ «Система анализа и мониторинга событий» (далее - Программа)

1.2 Используемые языки программирования

- Python 3.10/3.11
- JavaScript (Node.js – вспомогательные скрипты)

1.3 Системные требования

Программные требования:

Операционная система с поддержкой Docker не менее 20.10, docker-compose не менее 3.5, например ОС Linux Debian/Ubuntu/RHEL

Аппаратные требования:

Минимальные, необходимые для тестирования:

- 8 вычислительных ядер CPU, 16 ГБ ОЗУ, 50 ГБ SSD

Производственные:

- Вычислительный узел: 16 ядер ЦПУ, 32 ГБ ОЗУ, 100 ГБ SSD
- Система хранения: 8 ядер, 16 ГБ ОЗУ, 1 ТБ SSD (MongoDB, Elasticsearch)

2. ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ

Программа для ЭВМ «Система анализа и мониторинга событий» является системой автоматического сбора, обогащения и предоставления контента. Система собирает события из множества источников, обогащает их ML-моделями (регионы, NER-теги) и предоставляет доступ к результатам через API.

Программа представляет собой набор компонентов (микросервисов), которые позволяют осуществлять полный цикл обработки контента – от сбора до предоставления обогащенных данных потребителям.

Программа включает в себя следующие модули:

- **Модуль обогащения, сбора, систематизации и агрегации данных** – обеспечивает сбор событий из различных источников и их обогащение дополнительными признаками.
- **Модуль графического представления** – предоставляет сводку событий с возможностью фильтрации и список событий, визуализируя собранные и обогащенные данные в удобном интерфейсе.
- **Модуль редакции** – обеспечивает создание и редактирование контента (справочники, события, публикации) через графический интерфейс, а также административную панель для управления справочниками.

Программа обеспечивает выполнение следующих функций:

- Сбор новостей из RSS-лент, Telegram-каналов, групп ВКонтакте, Дзен.Новостей и произвольных веб-страниц
- Извлечение именованных сущностей (NER): персоны, локации, организации
- Классификация регионов публикаций на основе ML-моделей
- Предоставление доступа к собранным новостям через REST API с авторизацией по токenu
- Индексирование и полнотекстовый поиск новостей
- Архивирование новостей в отдельную БД
- Мониторинг состояния системы
- Создание и редактирование контента (справочники, события, публикации)

Программа состоит из следующих компонентов:

- **news-boiler** – веб-краулер на базе Celery, осуществляющий сбор новостей из разнородных источников и сохраняющий их в MongoDB
- **news-fermenter** – сервис ML-обработки новостей, выполняющий NER-разметку и классификацию регионов
- **boiler-proxy** – FastAPI-сервис, предоставляющий read-only REST API доступ к новостям из MongoDB
- **news-viewer** – веб-приложение, предоставляющее графический интерфейс для просмотра и фильтрации событий и публикаций, визуализации сводки событий, а также отображения результатов работы системы в удобном для пользователя виде
- **news-editor** – веб-приложение, обеспечивающее графический интерфейс для создания, редактирования и управления контентом системы: событиями, публикациями и справочниками

- **app-post** – бэкенд-сервис, предоставляющий REST API для создания, редактирования и удаления событий, публикаций и справочников, а также управления источниками и аккаунтами пользователей

Поддерживаемые типы источников:

Тип	Описание
rss	RSS-ленты (httpx/selenium/curl)
rss_follow_links	RSS + переход по ссылке для полного контента
tg	Telegram-каналы через Telethon
vk	Группы и пользователи ВКонтакте через VK API
yandex_news	Дзен.Новости через Selenium
web	Произвольные веб-страницы