Оркестратор (RBT 1002)

Руководство пользователя

Содержание

[Термины и определения 3](#_Toc136293464)

[1. Общие сведения 4](#_Toc136293465)

[1.1. Назначение 4](#_Toc136293466)

[1.2. Системные требования 4](#_Toc136293467)

[1.3. Общие сведения о Сервисе 4](#_Toc136293468)

[2. Функциональные характеристики 4](#_Toc136293469)

[3. Как работать с Оркестратором 5](#_Toc136293470)

Термины и определения

|  |  |
| --- | --- |
| **Термин/Сокращение** | **Определение/расшифровка** |
| Потенциальный партнер | Юридическое лицо, заявившее о намерении присоединиться к партнерской программе компании |
| Запрещенные слова | Слова и/или выражения, которые могут напрямую или косвенно указывать на высокие риски и нелегитимную деятельность потенциального партнера. |
| ОЗУ | Оперативное запоминающее устройство |
| CPU | Central processing unit, процессор |
| JS | Java Script (язык программирования) |

# Общие сведения

## Назначение

Программное обеспечение «Оркестратор (RBT 1002)» (далее – Сервис) предназначено для выполнения проверки репутации потенциального партнера на открытых ресурсах.

## Функциональные характеристики

* Выполняет поиск ссылок и информации на внешнем ресурсе по указанным параметрам. Максимальное количество анализируемых ссылок не превышает 75 для одного домена партнера; для поиска выполняется запуск JS-скриптов на сайтах потенциальных партнеров.
* Выполняет лексический анализ найденных запрещенных слов и выражений согласно словарю и выполняет подсчет степени риска по результатам анализа.
* Выполняет поиск запрещенных выражений и слов согласно словарю.
* По результатам генерирует отчет с запрещенными словами и выражениями, найденными в ходе анализа. Отчет сохраняется в файле pdf.
* Выполняет проверку сайта на редирект, просроченный SSL сертификат, наполненность сайта, если на сайте меньше 150 слов и меньше 2 рабочих ссылок на свой домен.

## Технические требования

К оборудованию для работы с Сервисом предъявляются следующие требования:

* Оборудование:
* 1 CPU – Intel Pentium4 и выше или AMD Athlon II и выше.
* ОЗУ - 2 Гбайт и выше (для 64-разрядных систем — 4 Гбайт).
* Сетевое подключение 10 Мбит.
* Компьютер под управлением ОС Windows версии 7 и выше, или актуальной версии MacOS, или ОС Linux (любого из современных дистрибутивов, например AstraLinux, Ubuntu, Debian);

Приложения:

* Docker.
* Интернет-браузер: Яндекс.Браузер, Edge, Mozilla FireFox, Google Chrome, Safari актуальных версий.

# Как работать с Оркестратором

Для поиска запрещенных слов и выражений выполните следующие действия:

1. Загрузите docker-образ из архива. Для этого в терминале наберите команду:

docker load -i orchestrator.tar/

1. Запустите терминал.
2. Запустите Оркестратор. Для этого в терминале наберите команду:

docker run -it -v [data:/data rbt1002:v1](http://data/data%20rbt1002:v1)/

python "1002Gos.py" "Orchestrator\_terms.xlsx" "<https://site.ru/>" "2"/

**Параметры команды:**

* [data:/data rbt1002:v1](http://data/data%20rbt1002:v1)/python

где

* + - * + data:/ – наименование папки, в которую будет сохранен отчет;
        + data - название папки в файловой системе контейнера
        + rbt1002:v1 – название контейнера для запуска
* 1002Gos.py - наименование скрипта. Файл скрипта расположен по адресу D:\rbt-1002Government\1002Gos.py.
* Orchestrator\_terms.xlsx - имя файла с терминами для поиска. Файл с терминами расположен по адресу D:\rbt-1002Government\ Orchestrator\_terms.xlsx.
* https://site.ru - сайт, который необходимо проверить.
* Цифра (здесь 2) - количество параллельных процессов. Рекомендуется использовать количество ядер CPU компьютера, на котором работает ПО.

1. Откройте файл отчета для просмотра результатов. Отчет о результатах будет сохранен в папке Data:

D:\rbt-1002Government\data – путь к файлу отчета

## Структура файла с терминами

Файл содержит следующие поля данных:

* + - * + Термины – запрещенные слова (рекомендуется указывать слово или фразу в именительном падеже).
        + Тематика - группа слов.
        + ЗВД - номер группы.
        + Приоритет - приоритет группы, указывается численное значение.
        + Коэффициент - коэффициент слова (количество найденных повторений слова на странице сайта).
        + Комментарий - комментарий (опционально).