Инв. № подл. Подп. и дата Взам. инв. № Инв. № дубл. Подп. и дата

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук Департамент программной инженерии

СОГЛАСОВАНО Профессор департамента программной инженерии факультета компьютерных наук, к.т.н	УТВЕРЖДАЮ Академический руководитель образовательной программы «Программная инженерия» профессор департамента программной инженерии, канд. техн. наук		
Е. М. Гринкруг «» 2020 г.	В. В. Шилов «» 2020 г.		
СИСТЕМА УПРАВЛЕН АВТОМАТИЧЕСКОМУ СБ ИНТЕ	ОРУ ДАННЫХ ИЗ СЕТИ		
Программа и мето	одика испытаний		
ЛИСТ УТВЕ	РЖДЕНИЯ		
RU.17701729.04	1.13-01 51 01-1		
	Исполнитель: студент группы БПИ 174 Д. Ю. Редникина «» 2020 г.		

СИСТЕМА УПРАВЛЕНИЯ ЗАДАНИЯМИ ПО АВТОМАТИЧЕСКОМУ СБОРУ ДАННЫХ ИЗ СЕТИ ИНТЕРНЕТ

Программа и методика испытаний

RU.17701729.04.13-01 51 01-1

Листов 29

Содержание

1	Обт	ьекты испытаний	4
	1.1	Наименование программы	4
	1.2	Область применения программы	4
		1.2.1 Функциональное назначение	4
		1.2.2 Эксплуатационное назначение	4
		1.2.3 Область применения	4
2	Цел	ь испытаний	5
3	Тре	бования к программе	6
	3.1	Требования к функциональным характеристикам	6
	3.2	Требования к интерфейсу	8
	3.3	Требования к надежности	8
4	Тре	бования к программной документации	9
	4.1	Состав программной документации	9
5	Сре	едства и порядок испытаний	10
	5.1^{-}	Технические средства, используемые во время испытаний	10
	5.2	Программные средства, используемые во время испытаний	10
	5.3	Порядок проведения испытаний	10
6	Me	годы испытаний	11
	6.1	Проверка требований к функциональным характеристикам	11
		6.1.1 Регистрация	11
		6.1.2 Авторизация	12
		6.1.3 Проекты (6.2)	13
		6.1.4 Краулеры (6.2)	16
		6.1.5 Запуски (6.2)	17
		6.1.6 Периодические задачи (6.2)	20
		6.1.7 Результаты тестирования	25
	6.2	Проверка требований к программной документации	26
П	рило	жение А	27
Cı	писо	к источников	28
.П	ист т	регистрации изменений	29

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

1 Объекты испытаний

1.1 Наименование программы

Система управления заданиями по автоматическому сбору данных из сети Интернет (System for managing tasks of collecting data from the Internet)

1.2 Область применения программы

1.2.1 Функциональное назначение

Система будет применяться как средство управления проектами по созданию, редактированию и запуску веб краулеров для сбора данных в сети интернет. Продукт позволит следить за запусками в режиме реального времени, а также создавать периодические запуски по расписанию.

1.2.2 Эксплуатационное назначение

Программа будет использоваться как инструмент для самостоятельной или совместной работы над проектами для запуска, управления сбора данных с помощью веб краулеров в сети интернет.

Таким образом, программный продукт позволит создавать, запускать образы пауков (см. 6.2) для сбора, управления, логирования и дальнейшего экспорта данных в целях сбора, изучения и мониторинга данных (см. 6.2).

1.2.3 Область применения

Технологии web-scraping используются как в науке, так и в бизнесе - многие люди чувствуют потребность в извлечении данных из HTML разметки интернет страниц. Существующие аналоги реализуют базовый функционал(сбор данных), но не предоставляют такие дополнительные возможности как периодический запуск или совместное редактирование. Многие аналоги (прим. scrapyd [1]) имеют ограниченный функционал. Главные возможности, которыми продукт обеспечит предполагаемых пользователей:

- Совместное управление запусками клаулеров
- Периодический запуск задач
- Сбор логов, ошибок
- Группировка краулеров, а также их запусков в проект
- Бесплатная функциональность

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

2 Цель испытаний

Проверка на соответсвие требованиям, указанным в документе «Система управления заданиями по автоматическому сбору данных из сети Интернет. Техническое задание».

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

3 Требования к программе

3.1 Требования к функциональным характеристикам

Программа выполняется в рамках темы курсовой работы в соответствии с учебным планом подготовки бакалавров по направлению 09.03.04 «Программная инженерия» Национального исследовательского университета «Высшая школа экономики», факультет компьютерных наук.

Программа должна удовлетворять следующим требованиям:

1. Авторизация

Чтобы использовать сервис, клиентская программа должна иметь возможность авторизоваться в системе с помощью REST API

- (а) Для регистрации пользователю нужно указать следующие данные
 - і. Почта уникальна для каждого зарегистрированного пользователя;
 - іі. Имя длина больше 1 символ;
 - ііі. Логин длина больше 2 символов;
 - iv. Пароль длина больше 2 символов;
- (b) Для авторизации пользователя в системе должны быть указаны следующие данные
 - і. Почта;
 - іі. Пароль;

2. Проекты

Должны быть реализованы запросы REST API для предоставления клиенту следующей функциональности

- (а) Создание проекта со следующей информацией
 - і. Имя проекта;
 - іі. Описание проекта опциональное поле;
- (b) Обновление метаданных о проекте (редактирование) могут быть обновлены только участником с минимальным уровнем дотупа ReadAndWrite. Следующие данные могут быть обновлены:
 - і. Имя проекта;
 - іі. Описание проекта;
 - ііі. Настройки проекта для запуска краулеров;
 - іv. Аргументы для запуска краулеров проекта;
- (c) Обновление egg файла проекта (редактирование) минимальный уровень доступа участника, обновляющий данные о проекте ReadAndWrite.
- (d) Удаление данных о проекте. Удалить проект может только владелец Owner.
- (e) Просмотр списка проектов (с пагинацией), к которым у пользователя есть как минимум ReadOnly доступ.

3. Участники проектов

Должны быть реализованы запросы REST API для предоставления клиенту следующей функциональности

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

7 RU.17701729.04.13-01 51 01-1

- (а) Просмотр информации об участниках проекта;
 - і. Имя, почта, логин участника;
 - ii. Статус участника в проекте (ReadOnly, ReadAndWrite или Owner);
- (b) Обновление статуса участника проекта. Это действие совершать может только владелец проекта;
- (c) Удаление участника из проекта. Данное действие может совершать только владелец проекта;
- (d) Добавление нового участника с указанными правами на редактирование. Данное действие может совершать только владелец проекта;

4. Краулеры

Должны быть реализованы запросы REST API для предоставления клиенту следующей функциональности

- (а) Просмотр списка краулеров проекта;
- (b) Редактирование информации о краулере для последующих запусков. Следующая информация может быть изменена
 - і. Настройки краулера для запуска;
 - іі. Аргументы для запуска;

5. Запуски краулеров

Должны быть реализованы запросы REST API для предоставления клиенту следующей функциональности

- (a) Просмотр списка запусков в определенном статусе (Pending, Running или Finished) с пагинацией, совершенных в проектах, к которым у пользователя есть как минимум ReadOnly доступ;
- (b) Редактирование запуска остановка запуска, перевод его в состояние Finished. Операция может быть применена только к запускам в состоянии Running или Pending;
- (c) Удаление запуска удаление всех данных о запуске из базы данных. Операция может быть применена только к запускам в состоянии Finished;
- (d) Создание запуска со следующей информацией
 - і. Краулер, с которым происходит запуск;
 - іі. Настройки запуска это могут быть как и предопределенные настроки на $scrapyd^{1}$, так и собственные настройки;
 - ііі. Аргументы запуска аргументы для запуска краулера, которые передаются через командную строку;
 - іv. Описание запуска;

6. Периодические запуски

Должны быть реализованы запросы REST API для предоставления клиенту следующей функциональности

(а) Просмотр списка периодических запусков с пагинацией;

http://doc.scrapy.org/en/latest/topics/settings.html

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

- (b) Редактирование следующей информации о периодическом запуске
 - i. Настройки будущих запусков это могут быть как и предопределенные настроки на scrapyd, так и собственные настройки;
 - іі. Аргументы будущих запусков аргументы для запуска краулера, которые передаются через командную строку;
 - ііі. Краулер, с помощью которого будет совершен запуск;
 - iv. cron-expression расписания запуска;
- (с) Удаление периодического запуска;
- (d) Отмена последующих запусков перевод периодической задачи в состояние Disabled;
- (e) Возобновление запусков перевод периодической задачи в состояние Enabled;
- (f) Создание периодического запуска со следующими данными
 - і. Название;
 - іі. Описание опциональное;
 - ііі. Краулер;
 - iv. Приоритетность, влияющая на очередь запусков (Low, Normal или High);
 - v. Статус (Enabled или Disabled);
 - vi. Настройки будущих запусков это могут быть как и предопределенные настроки на scrapyd, так и собственные настройки;
 - vii. Аргументы будущих запусков аргументы для запуска краулера, которые передаются через командную строку;

3.2 Требования к интерфейсу

Доступ к программе должен быть предоставлен посредством интерфейса консольной программы.

3.3 Требования к надежности

- Программа должна осуществлять проверку корректности посылаемого текста в кодировке utf-8 и возвращать ошибку в случае наличия невалидных символов.
- Программа должна работать непрерывно (не должно быть ситуации, в которой сервер возвращает status code 500).
- В базу данных не должны заноситься некорректные данные.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

4 Требования к программной документации

4.1 Состав программной документации

Состав программной документации должен включать в себя следующие компоненты:

- 1. Техническое задание «Система управления заданиями по автоматическому сбору данных из сети Интернет » (ГОСТ 19.201-78)
- 2. Программа и методика испытаний «Система управления заданиями по автоматическому сбору данных из сети Интернет » (ГОСТ 19.301-78)
- 3. Пояснительная записка «Система управления заданиями по автоматическому сбору данных из сети Интернет » (ГОСТ 19.404-79)
- 4. Руководство оператора «Система управления заданиями по автоматическому сбору данных из сети Интернет » (ГОСТ 19.505-79)
- 5. Текст программы «Система управления заданиями по автоматическому сбору данных из сети Интернет » (ГОСТ 19.401-78)

Вся документация должна быть составлена согласно ЕСПД (ГОСТ 19.101-77, 19.104-78, 19.105-78, 19.106-78 и ГОСТ к соответствующим документам (см. выше)) [2]. Все документы сдаются в электронном виде в составе курсовой работы LMS НИУ ВШЭ.

Пояснительная записка «Система управления заданиями по автоматическому сбору данных из сети Интернет » должна быть проверена на плагиат (< 40% заимствований). Документ, подтвержадющий проверку Пояснительной записки сдается в печатном виде вместе с подписанным отзывом от научного руководителя.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

5 Средства и порядок испытаний

5.1 Технические средства, используемые во время испытаний

- 1. Компьютер оснащенный процессором Intel Core i5 с тактовой частотой 2,3 ГГц;
- 2. 16 Гб ОЗУ;
- 3. Жесткий диск с объемом свободной памяти более чем 50 ГБ;
- 4. Клавиатура и мышь;
- 5. Доступ в интернет.

5.2 Программные средства, используемые во время испытаний

- 1. macOS 10.15.2;
- 2. scrapyd [1];
- 3. Scala 2.12.6;
- 4. Play-framework 2.6.13;
- 5. PostgreSQL 11 [5];

5.3 Порядок проведения испытаний

В процессе разработки программы были написаны unit-тесты, а также автотесты, тестирующие API сервера.

Автотесты исполняются на отдельной тестовой базе данных, которая перед началом каждого теста заполняется тестовыми данными.

Для запуска автотестов пользователю надо ввести команду sbt test в директории проекта cw_backend_scala, перед этим в отдельной директории запустить сервер scrapyd [1] командой scapyd.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6 Методы испытаний

6.1 Проверка требований к функциональным характеристикам

Для тестирования API запросов были написаны автотесты с помощью библиотеки scralatestplus [3]. Данная библиотека позволяет подменять зависимости основной программы и использовать FakeRequest как упрощенный способ выполнять запросы к контроллерам приложения.

Также для тестирования приложения используется библиотека play-silhouette-testkit [4]. Эта библиотека позволяет авторизовываться в приложении без использования cookies.

Для тестирования АРІ запросов были написаны следующие тест-кейсы:

6.1.1 Регистрация

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Регистрация	Пользователь	Пользователь отправ-	Status code 200.
	пользовате-	неавторизован	ляет запрос POST	Cookies в хэдере
	ЛЯ		/api/signup с персональ-	
			ными данными в формате	
			app.forms.SignUpForm	
2	Регистрация	Пользователь	Пользователь отправляет	Status code 409.
	пользова-	неавторизован.	запрос POST /api/signup	Сообщение об
	теля с уже	В БД системы	с персональными данны-	ошибке «User
	зарегистри-	уже существует	ми, которые совпадают с	Already Exists
	рованным	пользователь с	login и/или email другого	Message»
	email или	вводимым ло-	пользователя в базе дан-	
	ЛОГИНОМ	гином и/или	ных системы.	
		email.		

Таблица 1 — Тест кейсы: регистрация

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.2 Авторизация

Nº	Название	Предусловие	Описание	Ожидаемый
				результат
1	Авторизация	Пользователь	Пользователь отправ-	Status code 200.
	пользовате-	неавторизован в	ляет запрос на POST	Cookies в хэдере
	ЛЯ	системе. Пользо-	/api/signin с персональ-	
		ватель был ранее	ными данными в формате	
		зарегистрирован	app.forms.SignInForm	
		в системе		
2	Авторизация	Пользователь	Пользователь отправ-	Status code 403.
	пользовате-	неавторизован в	ляет запрос на POST	
	ля с непра-	системе. Пользо-	/api/signin с персональ-	
	вильным	ватель был ранее	ными данными (email,	
	паролем	зарегистрирован	пароль) так, что пароль	
		в системе.	не совпадает с уже вве-	
			денным при регистрации.	
3	Авторизация	Пользователь	Пользователь отправ-	Status code 409.
	пользовате-	неавторизован в	ляет запрос на POST	
	ля с несуще-	системе. Пользо-	/api/signin с персональ-	
	ствующими	ватель был ранее	ными данными (email,	
	в системе	зарегистрирован	пароль), которые не сов-	
	credentials	в системе.	падают ни с одной парой	
			credentials, хранящихся в	
			БД.	

Таблица 2 — Тест кейсы: авторизация

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.3 Проекты (6.2)

Nº	Название	Предусловие	Описание	Ожидаемый
				результат
1	Запрос на	Пользователь ав-	Пользователь отправ-	Status code 200.
	список про-	торизован в си-	ляет запрос GET	Json со спис-
	ектов без	стеме. У поль-	/api/projects/10 для	ком проектов, к
	пагинации	зователя имеют-	получения первых 10-и	которым у поль-
		ся 20 проектов.	из списка проектов, к	зователя есть
			которому у него есть	как минимум
			доступ.	readonly access,
				в хронологиче-
				ском порядке их
				создания.
2	Запрос на	Пользователь ав-	Пользователь отправля-	Status code 200.
	список про-	торизован в си-	ет запрос GET с query	Json со спис-
	ектов с	стеме. У поль-	argument limit=10 и id	ком проектов, к
	пагинацией	зователя имеют-	для получения 10-и про-	которым у поль-
		ся 20 проектов.	ектов из списка, которые	зователя есть
			в хронологическом по-	как минимум
			рядке (по дате создания)	readonly access,
			располагаются раньше	id которых >=
			чем указанный проект id.	idProject

Таблица 3 — Тест кейсы: просмотр списка проектов

Nº	Название	Предусловие	Описание	Ожидаемый
				результат
3	Создание	Пользователь ав-	Пользователь отправ-	Status code 200 и
	проекта	торизован.	ляет запрос POST	json с id создан-
			/api/projects c re-	ного проекта.
			лом запроса в виде	
			app.forms.ProjectCreForm	

Таблица 4 — Тест кейсы: создание проекта

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

№	Название	Предусловие	Описание	Ожидаемый
				результат
4	Редактирова-	Пользователь	Пользователь отправ-	Status code 200 и
	ние проекта	авторизован. У	ляет запрос PUT	json c id создан-
		пользователя	/api/projects/id с ука-	ного проекта. В
		есть хотя бы	занием id проекта для	таблице projects
		1 созданный	редактирования, с те-	должна обно-
		проект. Поль-	лом запроса в виде	виться запись
		зователь имеет	app.forms.ProjectForm.	с id созданного
		owner доступ к		проекта, где
		проекту.		name, description,
				spiderSettings
				совпадают с
				переданными
				пользователем.
5	Редактирова-	Пользователь	Пользователь с отправля-	Status code 403
	ние проекта:	авторизован. У	ет PUT /api/projects/id с	
	доступ поль-	пользователя	корректным телом запро-	
	зователя	доступ к проекту	са (см. п. 4)	
	readonly	id - readonly		
6	Редактирова-	Пользователь	Пользователь с отправля-	Status code 200.
	ние проекта:	авторизован. У	ет PUT /api/projects/id с	Информация о
	доступ поль-	пользователя	корректным телом запро-	проекте успешно
	зователя	доступ к проекту	са (см. п. 4)	обновлена
	writeandread	id - writeandread		

Таблица 5 — Тест кейсы: проекты - редактирование метаданных

№	Название	Предусловие	Описание	Ожидаемый
				результат
7	Удаление	Пользователь	Пользователь от-	Status code
	проекта	авторизован.	правляет DELETE	200, в таблице
	владельцем	У пользова-	/api/projects/:projectId	membership нет
		теля есть до-	запрос.	никаких записей
		ступ к проекту		о проекте с id, в
		id с правами		таблице project
		owner. (таблица		тоже.
		membership)		
8	Удаление	Пользователь	Пользователь от-	Status code 403
	проекта:	авторизован.	правляет DELETE	
	неверный id	Id проекта	/api/projects/:projectId	
		нет в таблице	запрос.	
		membership.		
9	Удаление	Пользователь	Пользователь от-	Status code 403
	проекта:	авторизован. У	правляет DELETE	
	права на	пользователя	/api/projects/:projectId	
	проект НЕ	права на проект	запрос.	
	owner	HE owner.		

Таблица 6 — Тест кейсы: проект DELETE

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

№	Название	Предусловие	Описание	Ожидаемый
				результат
10	Deploy архи-	Пользователь	Пользователь делает	Status code 200. B
	ва проекта	авторизован. У	запрос по адресу PUT	таблице проектов
		пользователя	$/\mathrm{api}/\mathrm{projects}/\mathrm{:projectId}/$	у проекта с id по-
		есть права на	deploy с телом eggFile:	явился bytes в ко-
		редактирование	<файл с расширением	лонке eggFile
		проекта (ли-	egg $>$	
		бо owner, либо		
		writeandread)		
11	Deploy	Пользователь	Пользователь делает	Status code 422
	архива про-	авторизован. У	запрос по адресу PUT	
	екта: файл	пользователя	$/\mathrm{api/projects/:projectId}/$	
	неверного	есть права на	deploy с телом eggFile:	
	формата	редактирование	файл не того формата (не	
		проекта (ли-	egg).	
		бо owner, либо		
		writeandread)		
12	Deploy архи-	Пользователь	Пользователь делает	Status code 403
	ва проекта:	авторизован. У	запрос по адресу PUT	
	права до-	пользователя	$/\mathrm{api}/\mathrm{projects}/\mathrm{:projectId}/$	
	ступа не	есть права на	deploy с телом eggFile:	
	позволя-	редактирование	<файл с расширением	
	ит менять	проекта readonly	egg $>$	
	проект			

Таблица 7 — Тест кейсы: проект PUT - deploy

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.4 Краулеры (6.2)

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Отображения	Пользователь ав-	Пользователь отправ-	Status code 200,
	списка крау-	торизован и име-	ляет запрос GET	список пауков в
	леров	ет доступ к про-	/api/projects/:projectId	проекте из БД
		екту id минимум	/crawlers с id проекта.	
		readonly.		
2	Отображение	Пользователь ав-	Пользователь отправ-	Status code 403
	списка крау-	торизован и не	ляет запрос GET	
	леров: нет	имеет доступ к	/api/projects/:projectId	
	доступа к	проекту id	/crawlers с id проекта.	
	проекту			

Таблица 8 — Тест кейсы: краулеры GET

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Изменение	Пользователь	Пользователь отправ-	Status code 200
	настроек	авторизован. У	ляет запрос PUT	
	паука	пользователя	/api/project/:projectId	
		есть write или	/crawlers.	
		owner доступ к		
		проекту.		
2	Изменение	Пользователь	Пользователь отправ-	Status code 403
	настроек	авторизован. У	ляет запрос PUT	
	паука: нет	пользователя	$/\mathrm{api}/\mathrm{project}/\mathrm{:projectId}$	
	прав доступа	есть readonly к	/crawlers.	
	к редактиро-	проекту.		
	ванию			
3	Изменение	Пользователь ав-	Пользователь отправ-	Status code 403
	настроек	торизован.	ляет запрос PUT	
	паука: не		/api/project/:projectId	
	совпадает id		/crawlers.	
	проекта с id			
	паука			

Таблица 9 — Тест кейсы: краулеры PUT

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.5 Запуски (6.2)

No	Название	Предусловие	Описание	Ожидаемый
				результат
1	Список джо-	Пользователь	Пользователь отправ-	Status code 200.
	бов	авторизован. У	ляет запрос GET	Возвращается
		пользователя	/api/projects/jobs/:limit,	массив джобов
		есть список из	где указывает limit=10	со статусами
		20 запущенных	и тип джобы finished.	finished и длиной
		јоb-ов в ста-	После этого отправляет	10. Отсортирован
		тусе finished,	тот же запрос со статусом	по id.
		запущенных	running	
		с пауками из		
		разных проектов.		
2	Список джо-	Пользователь	Пользователь отправ-	Status code 200.
	бов: пагина-	авторизован. У	ляет запрос GET	Возвращается
	ция	пользователя	/api/projects/jobs/:limit,	массив джобов
		есть список из	где указывает limit=10,	со статусами
		20 запущенных	excludiveFrom=3 и тип	finished и длиной
		јоb-ов в ста-	джобы finished	2. Отсортирован
		тусе finished,		по id.
		запущенных		
		с пауками из		
		разных проектов.		

Таблица 10 — Тест кейсы: список запусков GET

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

№	Название	Предусловие	Описание	Ожидаемый
				результат
3	Создание за-	Пользователь ав-	Клиент отправля-	Status 200. Pe-
	пуска	торизован. Име-	ет запрос POST	зультат - id co-
		ет ReadAndWrite	/api/projects/:projectId/	зданной задачи.
		или Owner до-	jobs	При запросе GET
		ступ к проекту		$/\mathrm{api/jobs\ status} =$
		projectId.		Running задача с
				id возвращается.
4	Создание	Пользователь ав-	Клиент отправля-	Status 403. Coo6-
	запуска:	торизован. Име-	ет запрос POST	щение об ошибке
	ReadOnly	ет ReadOnly до-	/api/projects/:projectId/	${\rm «NoPermission} {\rm »}.$
	доступ	ступ к проекту	jobs	
		projectId.		
5	Создание	Пользователь ав-	Клиент отправля-	Status 403. Coo6-
	запуска: id	торизован. Име-	ет запрос POST	щение об ошибке
	краулера не	ет ReadAndWrite	/api/projects/:projectId/	«Crawler Doesnt
	соответсвует	доступ к проекту	jobs	Correspond To
	id проекта	projectId. Id пау-		Project»
		ка не соответсву-		
		ет id проекта.		

Таблица 11 — Тест кейсы: создание запуска POST

№	Название	Предусловие	Описание	Ожидаемый
				результат
6	Отмена джо-	Пользователь	Клиент отправ-	Status 200. Bo3-
	бы	запустил джобу.	ляет запрос PUT	вращается объ-
		Работа находится	/api/projects/:projectId/	ект JobExecution
		в статусе pending	jobs/:jobScrapydId/:jobId,	с обновленными
		или running.	где jobScrapydId, jobId - id	данными: status
			джобы в статусе running.	= finished.
7	Отмена джо-	Пользователь ав-	Клиент отправ-	Status 403. Coo6-
	бы: id проек-	торизован в при-	ляет запрос PUT	щение об ошиб-
	та не совпа-	ложении. В про-	/api/projects/:projectId/	ке «No Project
	дает с id дж-	екте с id нет за-	jobs/:jobId.	Found».
	обы	пущенных джоб.		
8	Отмена	Пользователь	Клиент отправ-	Status 403. Coo6-
	джобы: у	авторизован в	ляет запрос PUT	щение об ошибке
	пользовате-	приложении.	$/\mathrm{api/projects/:projectId}/$	«No Permission».
	ля readonly	Имеет readonly	jobs/:jobId.	
	права	access к проекту		
		projectId.		

Таблица 12 — Тест кейсы: отмена запуска PUT

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

No	Название	Предусловие	Описание	Ожидаемый
				результат
9	Удаление за-	Пользователь авторизо-	Клиент от-	Status 200. Pe-
	пуска	ван. Имеет ReadAndWrite	правляет за-	зультат - id уда-
		доступ к проекту	прос DELETE	ленной джобы.
		projectId. Джоба на-	/api/projects/	
		ходится в статусе finished.	:projectId $/$	
			jobs/:jobId	
10	Удаление за-	Пользователь авторизо-	Клиент от-	Status 422. Coo6-
	пуска: джоба	ван. Имеет ReadAndWrite	правляет за-	щение об ошибке
	в статусе	доступ к проекту	прос DELETE	«Couldn't delete
	$\operatorname{running}/$	projectId. Джоба на-	/api/projects/	job execution».
	pending	ходится в статусе	:projectId/	
		running/pending.	jobs/:jobId	
11	Удаление	Пользователь авторизо-	Клиент от-	Status 403. Cooб-
	запуска:	ван. Имеет ReadOnly до-	правляет за-	щение об ошибке
	ReadOnly	ступ к проекту projectId.	прос DELETE	«No access».
	доступ к	Джоба находится в стату-	/api/projects/	
	проекту	ce finished.	:projectId/	
			jobs/:jobId	
12	Удаление за-	Пользователь авторизо-	Клиент от-	Status 403. Co-
	пуска: id pa-	ван. Имеет ReadAndWrite	правляет за-	общение об
	боты не со-	доступ к проекту	прос DELETE	ошибке «Job
	ответсвует id	projectId. Id джобы не	/api/projects/	Execution Doesnt
	проекта	соответсвует id проекта.	:projectId/	Correspond To
			jobs/:jobId	Project»

Таблица 13 — Тест кейсы: удаление запуска DELETE

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.6 Периодические задачи (6.2)

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Список пе-	Пользователь автори-	Пользователь от-	Статус-код 200.
	риодических	зован и имеет Owner	правляет запрос GET	Json с массивом
	задач	доступ к проекту с	api/projects/:projectId/	из 5 периоди-
		номером projectId.	periodicJobs/:limit,	ческих задач
			где $limit = 5$	проекта в по-
				рядке убывания
				id.
2	Список пе-	Пользователь автори-	Пользователь от-	Статус-код 200.
	риодических	зован и имеет Owner	правляет запрос GET	Json с массивом
	задач с	доступ к проекту с	api/projects/:projectId/	из 2 периоди-
	пагинацией	номером pId. Всего	periodicJobs/:limit/	ческих задач
		периодических задач	:exclusiveFrom,	проекта в поряд-
		в проекте 10 штук.	ho где $ ho$ limit $=$ 5,	ке убывания id
			exclusiveFrom =	(2, 1).
			2.	
3	Список	Пользователь автори-	Пользователь от-	Статус-код
	периодиче-	зован и не имеет до-	правляет запрос GET	403. Сообщение
	ских задач:	ступа к проекту с но-	api/projects/:projectId/	"NoPermission"
	доступа нет	мером pId.	periodicJobs/ :limit c	
			любым параметром	
			limit.	

Таблица 14 — Тест кейсы: список периодических задач GET

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

Nº	Название	Предусловие	Описание	Ожидаемый
				результат
1	Создание	Пользователь	Пользователь отправ-	Статус-код 200.
	периоди-	авторизован и	ляет запрос POST	Возвращает-
	ческого	имеет доступ	$/\mathrm{api/projects/:projectId/}$	ся JSON с id
	запуска	ReadAndWrite	periodicJobs c	созданной пе-
		к проекту	телом запроса	риодической
		projectId.	PeriodicJobCreateForm.	задачи.
2	Создание	Пользователь	Пользователь отправ-	Статус-код 420.
	периоди-	авторизован и	ляет запрос POST	Возвращается
	ческого	имеет доступ	$/\mathrm{api/projects/:projectId/}$	сообщение с
	запуска:	ReadAndWrite	periodicJobs c	ошибкой «invalid
	некоррект-	к проекту	телом запроса	cron expression».
	ный cron-	projectId.	PeriodicJobCreateForm,	
	expression		но cron-expression ука-	
			зывает в теле запроса	
			некорректный (прим.	
			«text»).	
3	Создание	Пользователь	Пользователь отправ-	Статус-код
	периодиче-	авторизован и	ляет запрос POST	403. Сообще-
	ского запус-	имеет доступ	$/\mathrm{api/projects/:projectId/}$	ние об ошибке
	ка: права	ReadOnly к про-	periodicJobs c	«NoPermission»
	ReadOnly	екту projectId.	телом запроса	
			PeriodicJobCreateForm.	
4	Создание	Пользователь	Пользователь отправ-	Статус-код 403.
	периоди-	авторизован и	ляет запрос POST	Сообщение об
	ческого	имеет доступ	/api/projects/:projectId	ошибке «No
	запуска:	ReadAndWrite	/periodicJobs c	Corresponding
	паук не со-	к проекту	телом запроса	Crawler»
	ответсвует	projectId. Id пау-	PeriodicJobCreateForm.	
	проекту	ка в теле запроса		
		не соответствует		
		projectId.		

Таблица 15 — Тест кейсы: создание периодических запусков POST

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Редактирование периодического запуска	Пользователь авторизован и имеет ReadAndWrite доступ к проекту projectId, в котором есть периодическая задача periodicJobId, а также паук crawlerId.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/:periodicJobId с телом запроса PeriodicJobChangeForm.	Статус-код 200.
2	Редактирование периодическо-го запуска: некорректный cron-expression	Пользователь авторизован и имеет ReadAndWrite доступ к проекту projectId, в котором есть периодическая задача periodicJobId, а также паук crawlerId.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/ :periodicJobId с телом запроса PeriodicJobChangeForm, но cron-expression указывает в теле запроса некорректный (прим. "text").	Статус-код 422. Со- общение об ошибке «invalid cron expression»
3	Редактирование периодического запуска: в проекте нет паука	Пользователь авторизован и имеет ReadAndWrite доступ к проекту projectId, в котором есть периодическая задача periodicJobId. В проекте нет паука с crawlerId.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/:periodicJobId с телом запроса PeriodicJobChangeForm с некорректным стаwlerId в теле.	Статус-код 403 и сооб- щение об ошибке.
4	Редактирование периодического запуска: пользователь имеет ReadOnly доступ к проекту	Пользователь авторизован и имеет ReadOnly доступ к проекту projectId.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/:periodicJobId с телом запроса PeriodicJobChangeForm.	Статус-код 403 и сооб- щение об ошибке.
5	Редактирование периодическо- го запуска: в проекте нет редактируемой периодической задачи	Пользователь авторизован и имеет ReadAndWrite доступ к проекту projectId. В проекте нет редактируемой задачи periodicJobId.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/:periodicJobId стелом запроса PeriodicJobChangeForm.	Статус-код 403 и сооб- щение об ошибке.

Таблица 16 — Тест кейсы: редактирование информации о периодическом запуске PUT

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Отмена запус-	Пользователь ав-	Пользователь отправ- ляет запрос PUT	Статус-код 200.
	ков периодиче-	торизован и имеет ReadAndWrite доступ к проекту projectId,	ляет запрос PUT /api/projects/:projectId/ periodicJobs/	200.
		у которого есть периодическая за-	:periodicJobId/disable.	
		дача periodicJobId в статусе Enabled.		
2	Отмена запусков периодической задачи: задача уже отменена	Пользователь авторизован и имеет ReadAndWrite доступ к проекту projectId, у которого есть периодическая задача periodicJobId в статусе Disabled.	Пользователь отправляет запрос PUT /api/projects/:projectId/periodicJobs/:periodicJobId/disable.	Статус-код 422.

Таблица 17 — Тест кейсы: отмена запуска периодических задач PUT

№	Название	Предусловие	Описание	Ожидаемый
				результат
1	Возобновление	Пользователь ав-	Пользователь отправ-	Статус-код
	запусков пе-	торизован и имеет	ляет запрос PUT	200.
	риодической	ReadAndWrite доступ	/api/projects/:projectId/	
	задачи	к проекту projectId,	m periodic Jobs/	
		у которого есть	: periodic Job Id/enable.	
		периодическая за-		
		дача periodicJobId в		
		статусе Enabled.		
2	Возобновление	Пользователь ав-	Пользователь от-	Статус-код
	запусков пе-	торизован и имеет	правляет запрос	422.
	риодической	ReadAndWrite доступ	PUT /api/projects/	
	задачи: задача	к проекту projectId,	:projectId/ periodicJobs/	
	уже в статусе	у которого есть	: periodic Job Id/enable.	
	enabled	периодическая за-		
		дача periodicJobId в		
		статусе Disabled.		

Таблица 18 — Тест кейсы: возобновление запусков периодических задач по расписанию PUT

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

Nº	Название	Предусловие	Описание	Ожидаемый
				результат
1	Удаление пе-	Пользователь ав-	Пользователь отправ-	Статус код
	риодической	торизован и имеет	ляет запрос DELETE	200.
	задачи	ReadAndWrite доступ	/api/projects/:projectId/	
		к проекту projectId.	periodicJobs/	
		У проекта есть пе-	:periodicJobId.	
		риодическая задача		
		periodicJobId в стату-		
		ce enabled.		
2	Удаление пе-	Пользователь ав-	Пользователь отправ-	Статус код
	риодической	торизован и имеет	ляет запрос DELETE	200. Возвра-
	задачи: зада-	ReadAndWrite доступ	/api/projects/:projectId/	щается id
	ча в статусе	к проекту projectId.	periodicJobs/	удаленной
	disabled	У проекта есть пе-	:periodicJobId.	задачи.
		риодическая задача		
		periodicJobId в стату-		
		ce disabled.		
3	Удаление пе-	Пользователь ав-	Пользователь отправ-	Статус код
	риодической	торизован и имеет	ляет запрос DELETE	403. Co-
	задачи: доступ	ReadOnly доступ к	/api/projects/:projectId/	общение
	Readonly	проекту projectId.	periodicJobs/	об ошибке
	, and the second	_	:periodicJobId.	«NoPermission

Таблица 19 — Тест кейсы: удаление периодических зада DELETE

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

6.1.7 Результаты тестирования

```
[info] EmailValidatorTest:
[info] - EmailString.wrongEmail
[info] - EmailString.validEmail
[info] AuthorizationSpec:
[info] Authorize person
[info] - should signUp: OK
[info] - should signUp: userAlreadyExists
[info] - should signUp: invalid email format
[info] - should signIn: wrong credentials
[info] JobTestCase:
[info] CrawlerSpec:
[info] - GET crawlers: OK
[info] - GET crawlers: no access
[info] - PUT crawlers
[info] - PUT crawlers: ReadOnly access
[info] - PUT crawlers: spider not found
[info] ProjectSpec:
[info] ProjectsController
[info] - should GET list of projects for user
[info] - should GET list of projects: with pagination
[info] - should CREATE project
[info] - should PUT project's metadata
[info] - should PUT project's metadata: Readonly access - no permission
[info] - should PUT project's metadata: ReadAndWrite access
[info] - should DELETE project: Owner access
[info] - should DELETE project: NOT Owner access
[info] - should DELETE project: doesn't exist
[info] - should PUT deploy
[info] - should PUT deploy: wrong format file
[info] - should PUT deploy: no access
[info] MembershipSpec:
[info] MembershipController
[info] - should GET members: ReadAndWrite access
[info] - should GET members: no access to project
[info] - should DELETE member: Owner access
[info] - should DELETE member: ReadAndWrite access
[info] - should PUT member: Owner access
[info] - should PUT member: ReadAndWrite access
[info] ApplicationSpec:
[info] ApplicationController Logout GET
[info] - should should be unauthorized error
[info] - should redicrect if user was found
[info] PeriodicJobSpec:
[info] PeriodicJobController
[info] - should GET periodic jobs: basic
[info] - should GET periodic jobs: pagination
[info] - should GET periodic jobs: no access to project
[info] - should POST periodic job: basic
[info] - should POST periodic job: invalid cron-expression
[info] - should POST periodic job: crawler doesn't correspond to project
[info] - should POST periodic job: ReadOnly access
[info] - should PUT periodic job: basic
[info] - should PUT periodic job: crawler does not correspond to project
[info] - should PUT periodic job: wrong cron-expression
[info] - should DELETE periodic job: enabled status
[info] - should DELETE periodic job: disabled status
```

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

```
[info] - should DELETE periodic job: readonly access
[info] - should DELETE periodic job: no existing job with id
[info] - should PUT cancel periodic job: basic
[info] - should PUT cancel periodic job: already disabled
\hbox{[info] - should PUT cancel periodic job: readonly access}
[info] - should PUT cancel periodic job: no existing job with id
[info] - should PUT enable periodic job: basic
[info] - should PUT enable periodic job: already enabled
[info] - should PUT enable periodic job: readonly access
[info] - should PUT enable periodic job: no existing job found with id
[info] SettingsMergerTest:
[info] - SettingsMerger.basic
[info] - SettingsMerger.wrongInput
[info] - SettingsMerger.testPriority
[info] JobExecutionSpec:
[info] JobsController
[info] - should GET jobs
[info] - should GET jobs: pagination
[info] - should POST schedule: ordinary
[info] - should POST schedule: ReadOnly access
[info] - should POST schedule: ProjectId doesn't match CrawlerId
[info] - should PUT cancel
[\verb|info|| - \verb|should|| \verb|PUT|| cancel: \verb|jobId|| doesn't match to projectId|
[info] - should PUT cancel: ReadOnly access
[info] - should DELETE job
[info] - should DELETE job: still running
[info] - should DELETE job: ReadOnly access
[info] - should DELETE job: jobId doesn't match projectId
[info] ScalaTest
[info] Run completed in 1 minute, 33 seconds.
[info] Total number of tests run: 68
[info] Suites: completed 10, aborted 0
[info] Tests: succeeded 68, failed 0, canceled 0, ignored 0, pending 0
[info] All tests passed.
[info] Passed: Total 68, Failed 0, Errors 0, Passed 68
[success] Total time: 107 s, completed 11-Apr-2020 16:44:40
```

Листинг 1 — Результаты тестирования в консоли

6.2 Проверка требований к программной документации

Вся документация, представленная в требованиях, готова.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

ПРИЛОЖЕНИЕ А

Используемые понятия и определения

- Web scraping это сбор данных с различных интернет-ресурсов. Общий принцип его работы можно объяснить следующим образом: некий автоматизированный код выполняет GET-запросы на целевой сайт и получая ответ, парсит HTML-документ, ищет данные и преобразует их в заданный формат.
- **Проект** сущность для объединения и предоставления доступа к запускам/краулерам/периодическим задачам.
- Веб краулер программа, являющаяся составной частью поисковой системы и предназначенная для перебора страниц Интернета с целью занесения информации о них в базу данных поисковика. Неотъемлемая часть проекта. Именно с помощью пауков пользователь может "краулить" сайты для сбора необходимой информации.
- Запуск единоразовый запуск краулера с настройками и аргументами, указанными для этого запуска.
- **Периодический запуск** запуск с множеством настроек, повторяющийся в определенные периоды времени (запуски по cron-expression).

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

28 RU.17701729.04.13-01 51 01-1

Список источников

- [1] Github scrapyd/scrapyd [Электронный ресурс] URL: https://github.com/scrapy/scrapyd (Дата обращения: 16.04.2020, режим доступа: свободный)
- [2] Единая система программной документации М.: ИПК, Издательство стандартов, 2000, 125 стр.
- [3] ScalaTest+Play [Электронный ресурс] URL:http://www.scalatest.org/plus(Дата обращения: 16.04.2020, режим доступа: свободный)
- [4] Testing silhouette [Электронный ресурс] URL:https://www.silhouette.rocks/docs/testing (Дата обращения: 16.04.2020, режим доступа: свободный)
- [5] Postgresql [Электронный ресурс] URL:https://www.postgresql.org (Дата обращения: 16.04.2020, режим доступа: свободный)

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.04.13-01 51 01-1				
Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

29 RU.17701729.04.13-01 51 01-1

	Лист регистрации изменений								
Изм.		Номера листов		Всего листов в документе	№ доку- мента	Входящий № сопро- водит. докум. и дата	Под- пись	Дата	
	из- ме- нен- ных	заме- нен- ных	но- вых	аннул- лиро- ван- ных					