



НАЦИОНАЛЬНЫЙ ЦЕНТР  
КОГНИТИВНЫХ РАЗРАБОТОК  
УНИВЕРСИТЕТА ИТМО

**НЦКР**



# **Платформа поддержки жизненного цикла интеллектуальных объектов на основе промышленных больших данных**

руководство оператора

Санкт-Петербург, 2020 г.



## ОГЛАВЛЕНИЕ

1. Назначение программного комплекса.....	3
2. Условия выполнения программы.....	4
3. Выполнение программы.....	4
3.1. Основной сценарий применения .....	4
3.2. Интерфейс управления приложением.....	5
4. Сообщения оператору.....	12



## 1. Назначение программного комплекса

Программный комплекс платформы поддержки жизненного цикла интеллектуальных объектов на основе промышленных больших данных (ПК) обеспечивает формирование процесса организации выполнения задач, решаемых на основе машинного обучения с использованием интегрированных больших данных и знаний (реализуемый в форме интеллектуальных моделей, основанных на этих данных).

Унифицированное сопровождение цикла проектирования, разработки и обучения цифровых объектов ИИ на основе платформы достигается за счет объединения на одной платформе:

- бирж, позволяющих спроектировать модель для решения ожидаемых задач;
- области рабочих проектов и пространства пользователя для разработки, обучения и представления результатов.

ПК предназначен для решения следующих задач:

- формирование проектных условий согласования работы, используемых ресурсов, сроков, цен между всеми участниками цифрового рынка;
- запуск и выполнение работ проекта по согласованным условиям в единой среде.

Формирование проектных условий происходит на создаваемом цифровом рынке, представленном тремя цифровыми биржами: биржей задач, биржей данных и биржей аналитики. Запуск работ и выполнение проекта реализуется через предоставляемую область рабочих проектов и зону администрирования.

**Биржа задач** представляет собой аукцион, на который попадают задачи, описанные в определенной форме, с выделенными ключевыми характеристиками: целью, ожидаемыми результатами, сроками и бюджетом.

**Биржа данных** также является аукционом, на котором отображаются витрины данных с информацией об их структуре и ценовой политике, мета-описаниями и прочими характеристиками, а также доступна возможность просмотра среза ожидаемых данных.

**Биржа аналитики** является пространством предложений на оказание услуг обработки больших данных и построения интеллектуальных моделей поверх них.

**Область рабочих проектов или рабочее пространство пользователя** представлено проектами, в которых он участвует в различных ролях: заказчик, владелец данных, аналитик, административное лицо.

**Заказчик** осуществляет согласование условий запуска проекта (финансы, ресурсы, задачи), а также просматривает текущий процесс выполнения проекта, корректируя ход исследования путем уточнения результатов и целей.

**Владелец данных** определяет политику работы с собственными данными: какие данные и какими срезами (стратегии) можно реализовывать на цифровом рынке или какие данные могут быть обработаны и какими функциями. Также владелец данных следит за своевременным обновлением представленных массивов данных.

**Аналитик** — это специалист в области решения задач с использованием машинного обучения (МО), который владеет информацией о текущем рынке данных, доступных в платформе ИИ моделях и методах обработки.



## 2. Условия выполнения программы

Работа оператора осуществляется путем взаимодействия с интерфейсом, доступ к которому производится посредством web-браузера (Google Chrome, Safari, Firefox или Opera), установленного на персональный ЭВМ оператора. Для корректного выполнения приложения персональная ЭВМ оператора должна обладать следующими минимальными характеристиками:

- оперативная память – не менее 2.0 ГБ;
- дисковая подсистема – не менее 1 ГБ.

Персональная ЭВМ оператора должна работать под управлением одной из следующих ОС:

- Windows (версии не ниже Windows 7);
- Linux (на ядре версии не ниже 2.6);
- Mac OS.

Рабочая станция оператора должна быть подключена к сети Интернет.

## 3. Выполнение программы

### 3.1. Основной сценарий применения

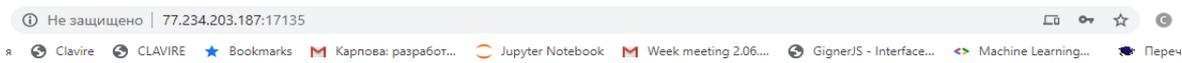
Основной сценарий содержит следующие шаги:

- владельцы данных регистрируют их для аналитики и продажи, аналитики регистрируют свои команды, специализирующиеся на выполнении МО-проектов (проектов, создаваемых на основе машинного обучения);
- заказчики публикуют свои задачи, описывая цели, ресурсы, сроки;
- группы аналитиков просматривают опубликованные задачи, витрины данных и формируют свои предложения по возможностям реализации проектов;
- заказчики фильтруют и подтверждают в рамках своей задачи наиболее интересное предложение по времени, финансам, данным и ожидаемому результату, запуская проект в реализацию;
- проект создается под задачу с сопутствующим окружением для аналитической группы, включая данные, ресурсы и пространство для разработки на основе Jupyter Hub;
- разработка ведется поэтапно с генерацией сценариев для реализации процессинга и анализа данных, с дальнейшим построением моделей и получением финального результата.

Производные сценарии формируются путем группирования отдельных блоков.

### 3.2. Интерфейс управления приложением

Взаимодействие оператора с ПК осуществляется посредством графического пользовательского интерфейса, доступ к которому осуществляется по адресу <http://77.234.203.187:17135/> с помощью индивидуального логина и пароля (рисунок 3.2.1). Логин и пароль необходимо получить у администраторов сервиса, отправив запрос по адресу: [nasonov@itmo.ru](mailto:nasonov@itmo.ru).



  
Sign in



**SIGN IN**

Рисунок 3.2.1 – Начало работы с системой, стартовый экран приложения

После авторизации в системе пользователю будет доступен главный экран приложения (рисунок 3.2.2), на котором представлена биржа задач. С главного экрана биржи также осуществляются переходы на экраны бирж данных (Data) и аналитики (Analytics), а также вкладку ресурсов (Resources). В демонстрационных целях, данные вкладки имеют ограниченное использование.

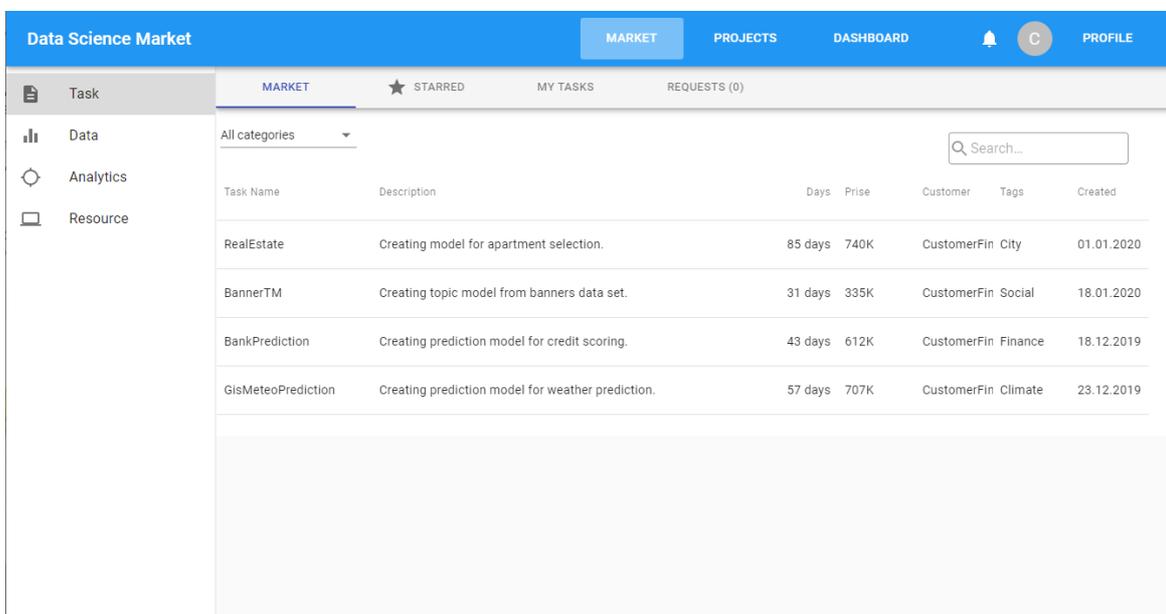


Рисунок 3.2.2 – Главный экран приложения

С вкладки Task можно перейти на внутреннюю вкладку My Tasks, которая отвечает за задачи, создаваемые пользователем (рисунок 3.2.3). Здесь можно выбрать опцию запуска новой задачи (Add Task), где можно зафиксировать стоимость, сроки и описание проблемы (рисунок 3.2.4).

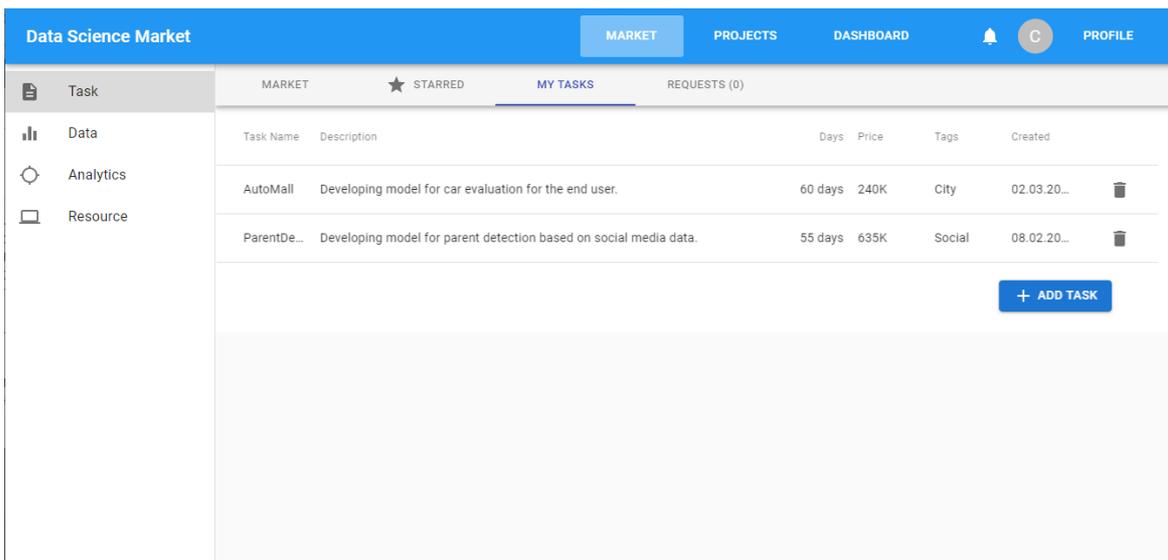


Рисунок 3.2.3 – Экран текущих задач

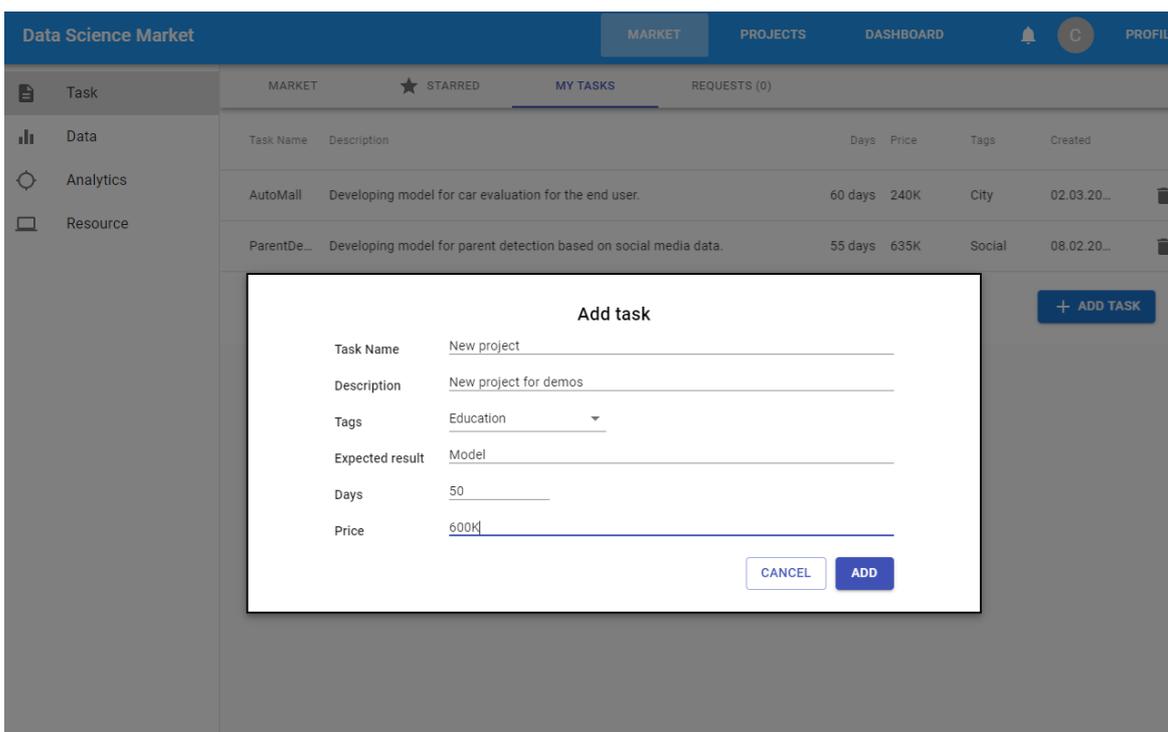


Рисунок 3.2.4 – Создание задачи

После этого задача будет доступна всем командам на бирже, которые готовы сделать предложение по реализации данной задачи (рисунок 3.2.5).



Task Name	Description	Days	Price	Customer	Tags	Created
AutoMall	Developing model for car evaluation for the end user.	60 days	240K	CustomerSpbGov	City	02.03.2020
RealEstate	Creating model for apartment selection.	85 days	740K	CustomerFinBank	City	01.01.2020
ParentDetection	Developing model for parent detection based on social media data.	55 days	635K	CustomerSpbGov	Social	08.02.2020
BannerTM	Creating topic model from banners data set.	31 days	335K	CustomerFinBank	Social	18.01.2020
BankPrediction	Creating prediction model for credit scoring.	43 days	612K	CustomerFinBank	Finance	18.12.2019
GisMeteoPrediction	Creating prediction model for weather prediction.	57 days	707K	CustomerFinBank	Climate	23.12.2019
New project	New project for demos	50	600K	CustomerSpbGov	Education	04.04.2020

Рисунок 3.2.5 – В данном случае созданный проект доступен последней строкой в списке

Предложение на задачу формируется путем просмотра подробной информации о задаче (двойным щелчком мыши по записи) и нажатием кнопки Send Request, где указываются: используемые данные из имеющихся одобренных заявок (в примере, vk + facebook), цена и ожидаемые результаты (рисунок 3.2.7). Этот подход реализованный в задачах, бирже данных и бирже ресурсов позволяет строить новые бизнес-процессы, присущие цифровым рынкам создания / продажи / анализа больших данных, а также моделей и сервисов на их основе.

Для того чтобы получить одобренные заявки на данные, необходимо сформировать по примеру выше Request на бирже данных вкладка (Data), выбрав соответствующий pack на предполагаемую покупку (рисунок 3.2.6).

Data name	Description	Access options	Customer	Category	Price
Finance_bank	This data contains anonymized user data from SPB bank transactions	2	CustomerFinBank		60K
HR_data	Anonymized data from finance company.	2	CustomerFinBank		60K

Pack 1	Pack 11
>100k >100k 1 RUB per record 0.9 RUB per record Data type Quantity Price: 110K RUB	>100k >100k 1 RUB per record 0.9 RUB per record Data type Quantity Price: 10.1K RUB

Рисунок 3.2.6 – Пример предложений на бирже данных

Как только предложение сделано, во вкладке Requests (у владельца задачи) появляется новый Request (предложение), в которое можно зайти и одобрить его, нажав Start project (рисунок 3.2.8).

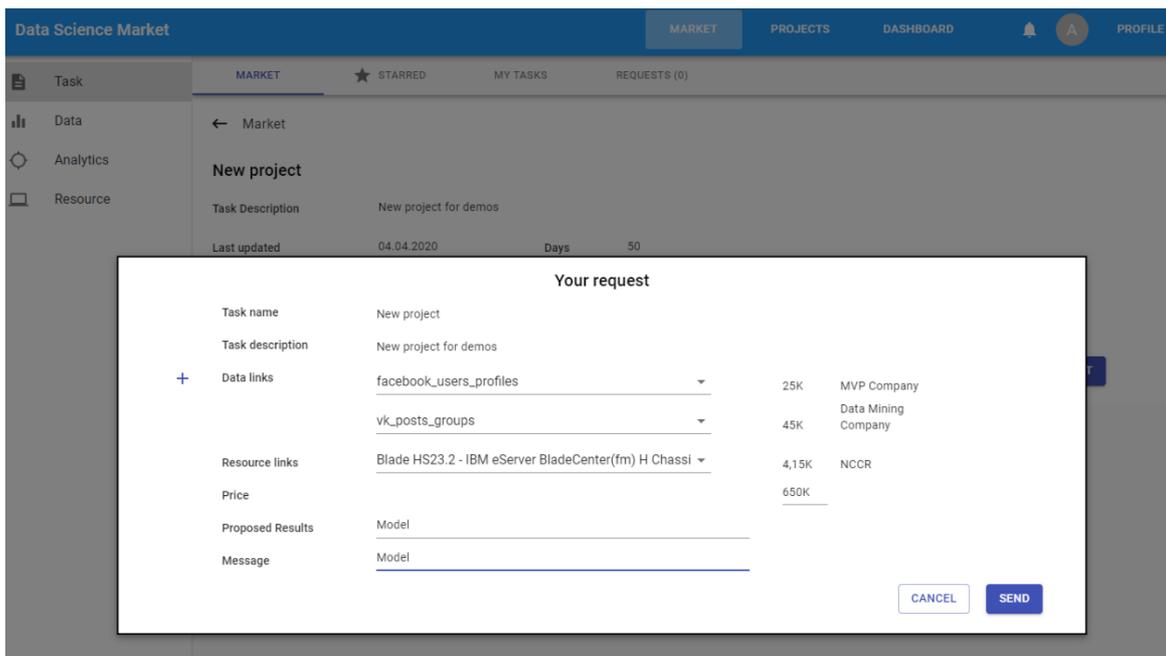


Рисунок 3.2.7 – Формирование предложение на задачу

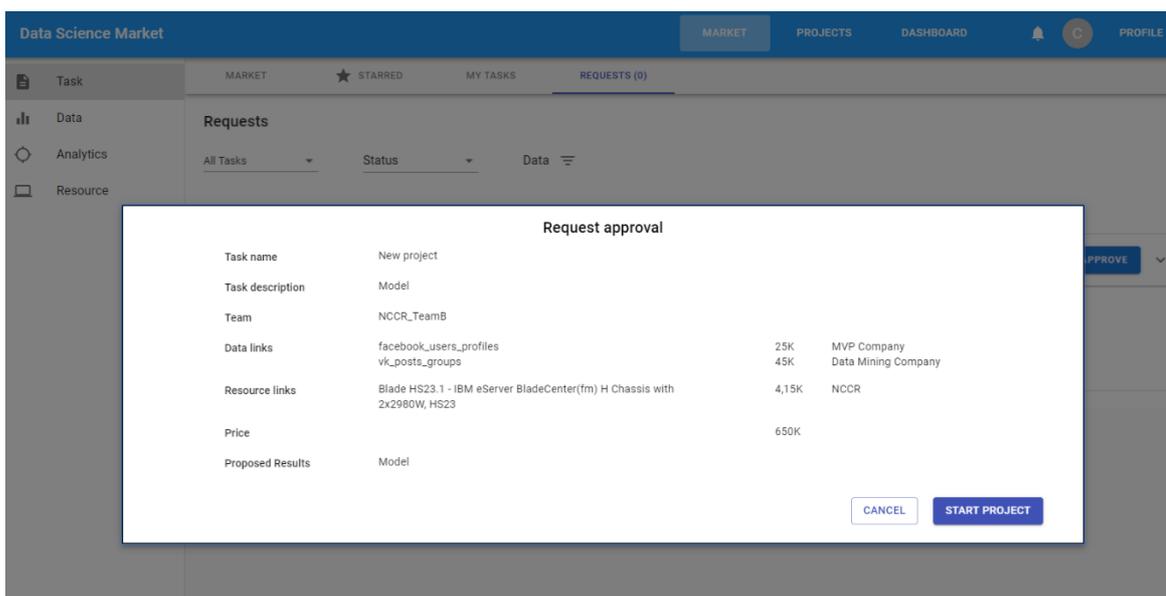


Рисунок 3.2.8 – Просмотр предложение с одобрением через кнопку Start Project

После запуска проекта, можно увидеть во вкладке Projects (основного горизонтального меню) список проектов и созданный New project (рисунок 3.2.9). Зайдя в него, можно обнаружить вкладки со сценариями проекта (Project scenarios), результатами (results), а также вкладки используемых ресурсов и конфигурации с ограниченными функциями в рамках демонстрации. Необходимо отметить, что, в зависимости от используемой учетной записи и доступа к приложению, будут доступны разные списки (на рисунке 3.2.9 пример окна владельца задачи), это связано с тем, что у разных пользователей разная история реализации проектов. Во вкладке Project scenarios можно создать первый сценарий по шаблону City\_Template, для демонстрационных целей данный шаблон содержит уже реализованную модель оценки удовлетворенности жителей городом (рисунок 3.2.10). Механизм шаблонизации, тиражирования и единой точки разработки с формированием интерактивных результатов, позволяют обеспечивать быстрое конструирование необходимых моделей.

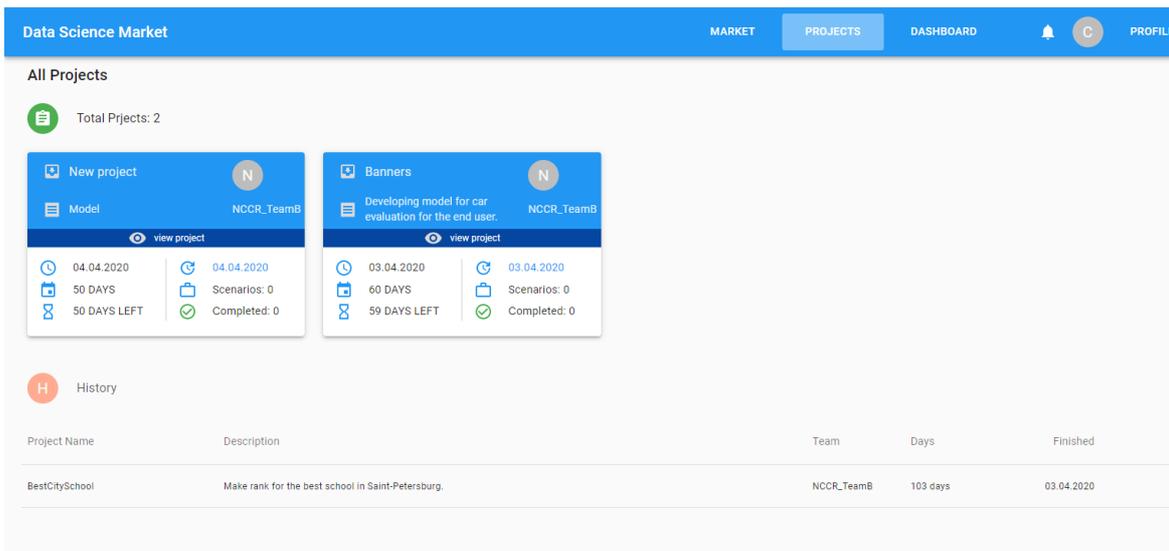


Рисунок 3.2.9 – Просмотр вкладки проектов

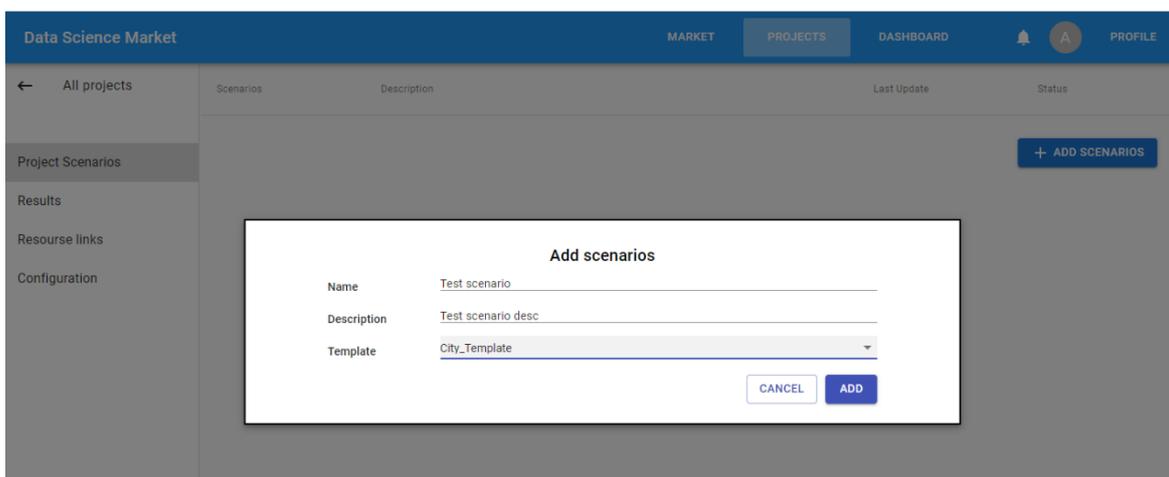


Рисунок 3.2.10 – Создание нового сценария

После того как сценарий создан, он появляется в списке сценариев, и при необходимости можно попасть в его рабочую область, нажав на сценарий двойным щелчком (рисунок 3.2.11). Для просмотра демонстрационной задачи необходимо сгенерировать сценарий, запустив его выполнение на сервере (Меню → Cell → Run All), как показано на рисунке 3.2.11. Благодаря встроенной DSL библиотеке платформы, будет сохранен первый результат проекта (код генерации можно просмотреть, прокрутив весь сценарий до конца). Поддержка доступа к сценариям различных ролей пользователей одновременно позволяет обеспечить совместное выполнение проектов распределенными командами, наличие DSL и регламента процесса позволяет привлекать членов команд с ограниченными компетенциями в области ИИ. Зайдя во вкладку Results, можно увидеть результаты работы данного сценария, а в них сгенерированный результат последнего запуска (рисунок 3.2.12).

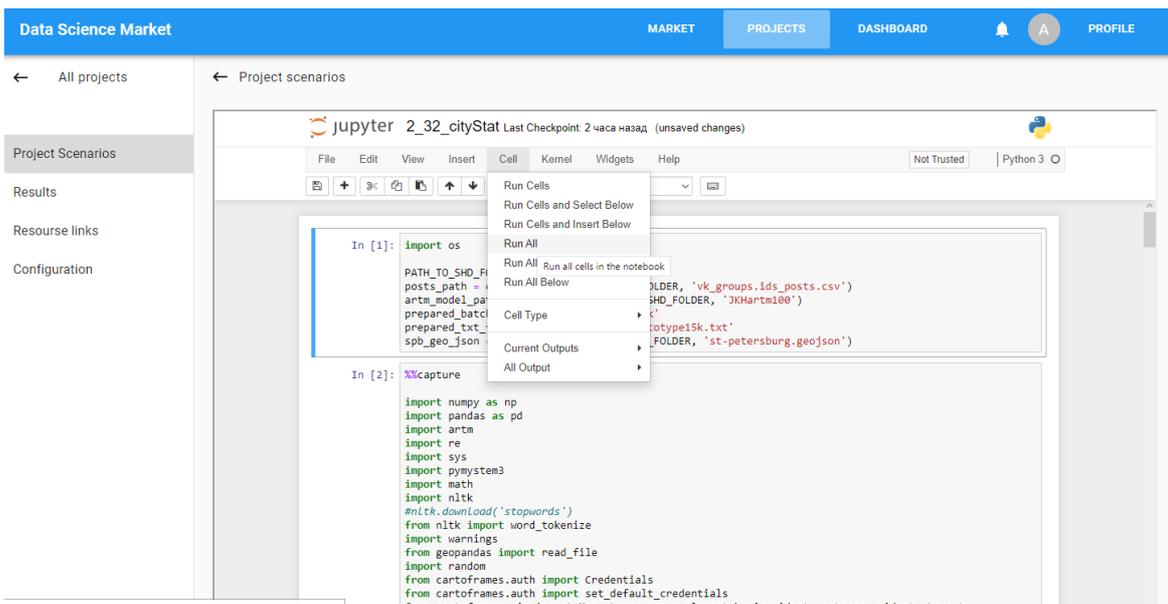


Рисунок 3.2.11 – Выполнение сценария на стороне сервера

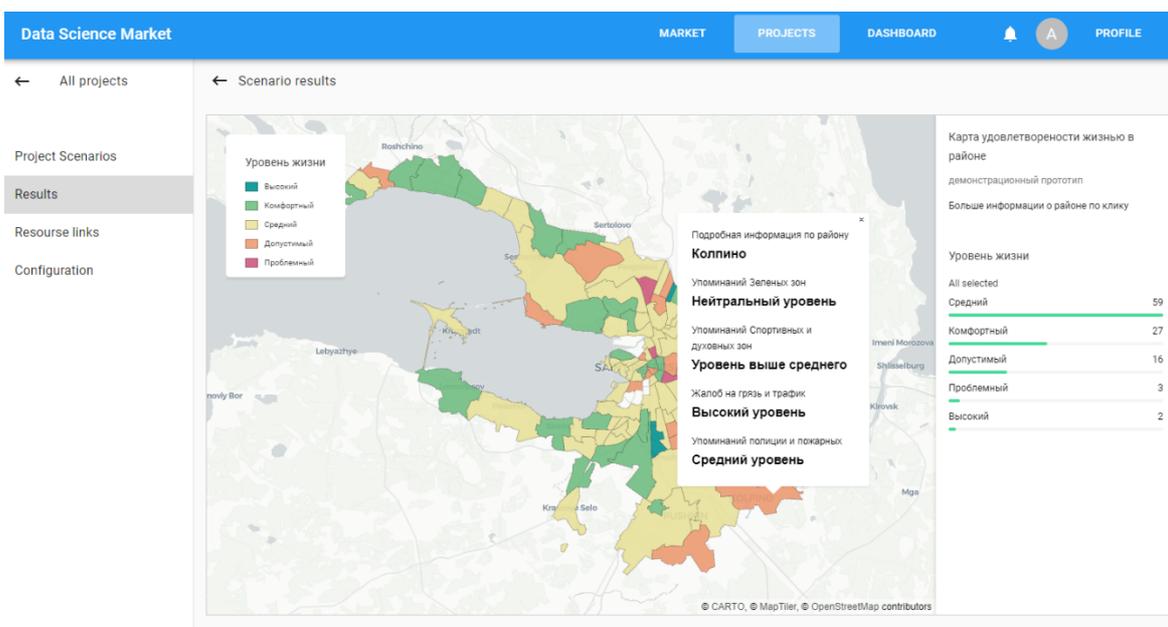


Рисунок 3.2.12 – Результат выполнения сценария

Другим примером демонстрации использования являются интерактивные результаты работы системы с использованием механизмов обмена данными с серверной частью через встроенные элементы Jupiter Server-а. Которые позволяют управлять процессом.

Также можно посмотреть агрегированную информацию по всем сущностям через элемент Dashboard на рисунке 3.2.13.

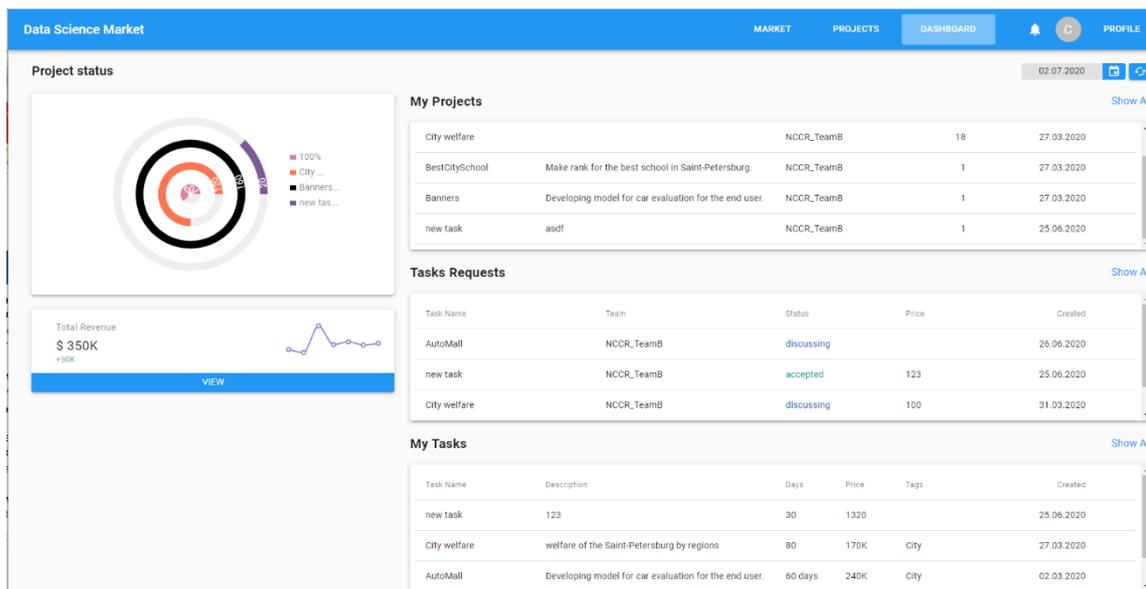


Рисунок 3.2.13 – Просмотр агрегированной информации по профилю пользователя

При необходимости можно просматривать текущие запущенные сценарии под контролируруемыми проектами. Эта опция доступна во вкладке проектов и позволяет для экономии ресурсов останавливать те сценарии, что более не являются необходимыми (рисунок 3.2.14).

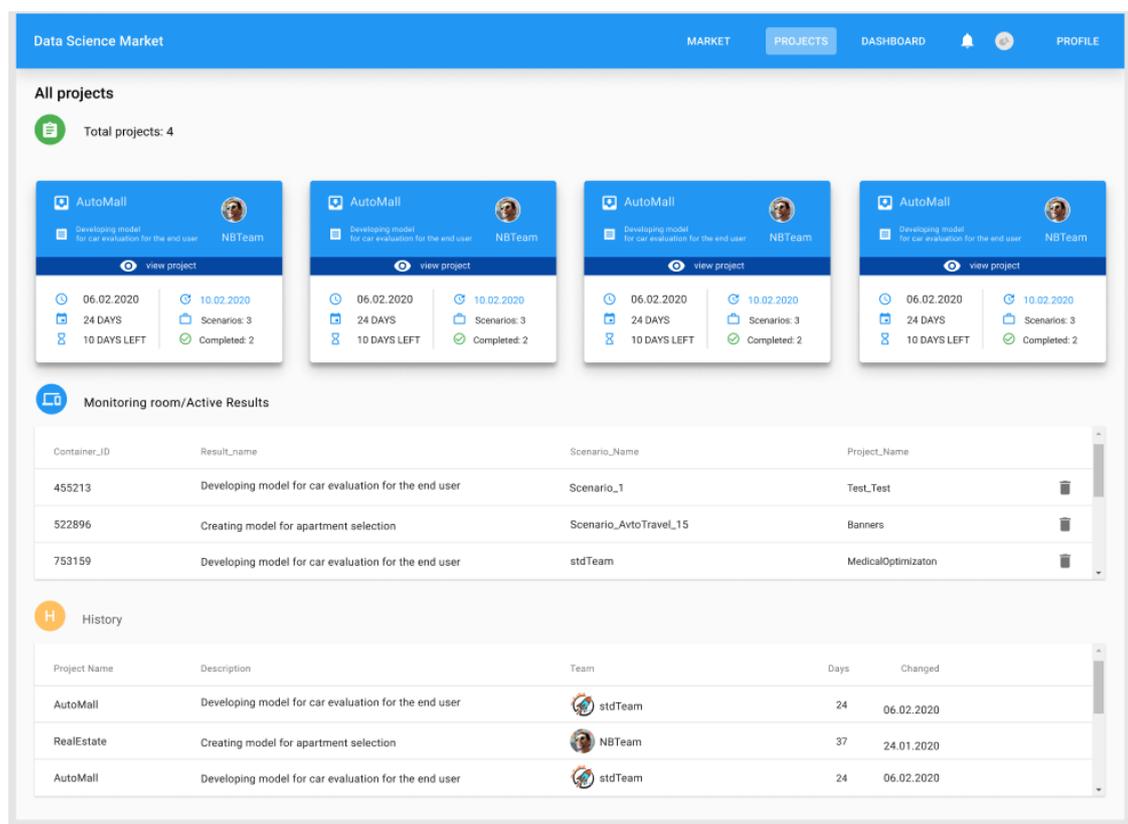


Рисунок 3.2.14 – Просмотр запущенных сценариев в текущих проектах

Дополнительными опциями является возможность управления собственными командами при наличии соответствующих прав. Доступны функции создания, изменения и удаления составов команд под разные проекты (рисунок 3.2.15)

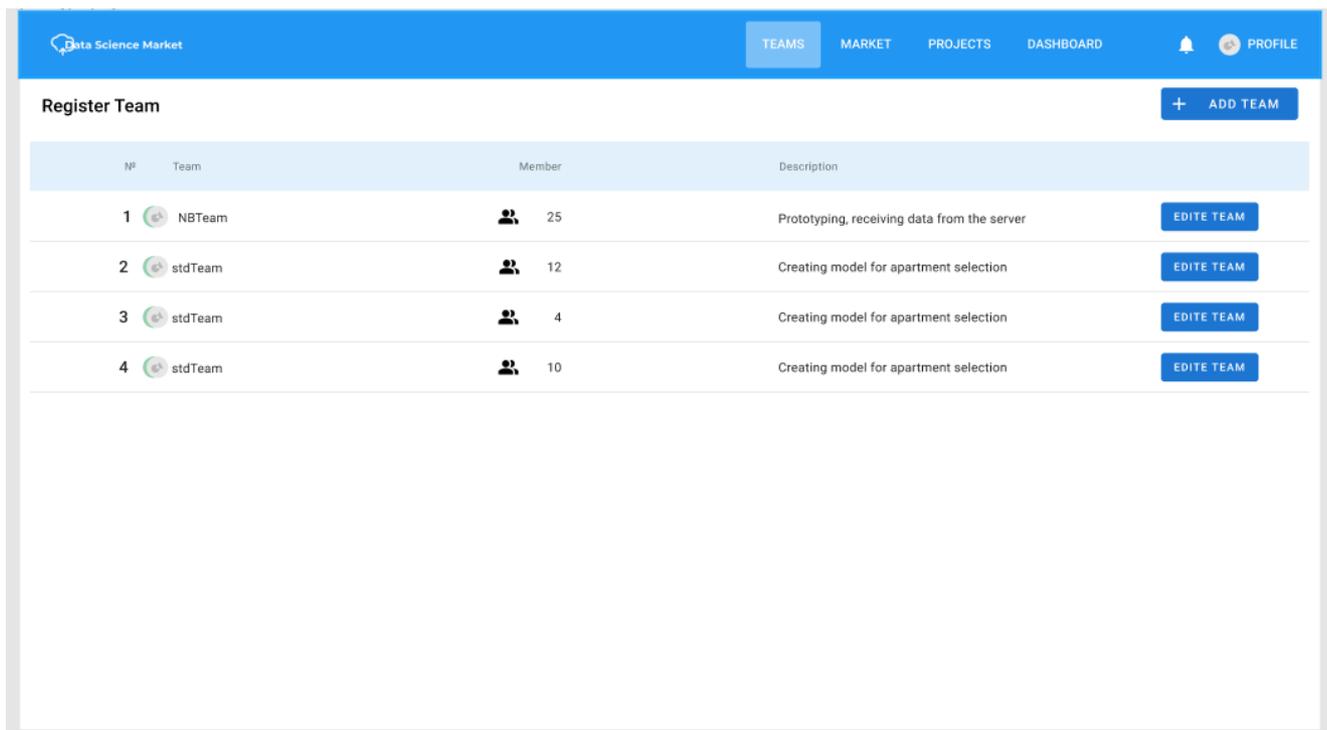


Рисунок 3.2.15 – Вкладка управления командами

Отдельной возможностью является работа в рамках своих зарегистрированных наборов данных, загруженных (зарегистрированных) пользователем извне (рисунок 3.2.16).

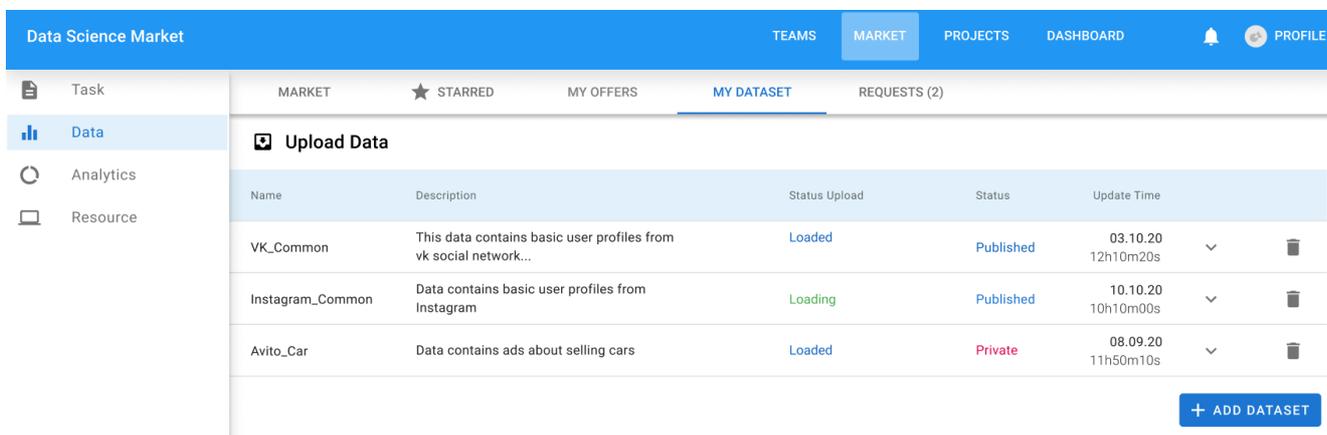


Рисунок 3.2.16 – Работа с собственными данными

#### 4. Сообщения оператору

Критичным для работы ПК является доступ к сети Интернет, без этого функционирование приложения невозможно. В случае наличия технических проблем со стороны используемого приложения просьба зафиксировать их и переслать по указанному адресу поддержки. Возможные ошибки, связанные с вводом и выводом данных, а также сопровождающие их сообщения, представлены в таблице:

Основные сообщения оператору

Сообщение	Экран	Типовые действия по выявлению и устранению ошибки
-----------	-------	---



Please, use another browser	Стартовый экран	Данный браузер не поддерживается, для взаимодействия с ЭО ПК необходимо использовать другой браузер.
Incorrect login/password, please try again	Стартовый экран	Неверный логин/пароль. Проверьте корректность ввода логина и пароля. Если у вас нет пароля, обратитесь к администратору сервиса.