

Методические материалы для участников хакатона «Tender Hack — Москва»

Цель и задачи

Создание функционала генерации наименования и характеристик СТЕ в универсальном виде.

Сервис разрабатывается для Портала поставщиков — <https://zakupki.mos.ru/>.

В рамках хакатона участникам предстоит реализовать задачи:

- Проанализировать фактические описания, заполненные характеристики и т.д. по исходным данным;
- Выявить формат и содержание унифицированного вида СТЕ для каждой категории справочника (наименование, категория, бренд, модель, обязательные характеристики, дополнительные характеристики и т.д.);
- Разработать подсистему определения категории СТЕ на основе введенного полного наименования;
- Разработать подсистему генерации полноценного структурированного описания: структурированного наименования, характеристик (в том числе дополнительных) на основе введенного полного наименования СТЕ;
- Реализовать проект;
- Провести тестирование системы на реальных кейсах;
- Продемонстрировать работоспособность проекта;
- Определить возможности по масштабированию решения и следующим его доработкам;
- Презентовать проект.

Необходимый реализуемый функционал:

1. Система генерации структурированного наименования, характеристик (в том числе дополнительных) на основе введенного полного наименования СТЕ

Ограничения:

1. Запрещается использование платных API и иностранных API.
2. Предельная скорость работы системы – 60 сек.

Этапы защиты:

Этап 1: Защита проектов (презентация до 5 минут). Параллельно второй участник команды демонстрирует реальный кейс. Максимальное количество баллов по критериям — 100.

Этап 2: Отправка ссылок на репозитории (или в архиве) с кодом и визуальной частью. Защита топ-5 решений (презентация до 10 минут, затем 10 минут на вопросы жюри). Максимальное количество баллов по критериям — 100. Баллы суммируются с баллами первого этапа.

Критерии оценки:

1. Работоспособность системы – 50 баллов.
2. Полнота и точность заполнения свойств – 30 баллов
3. Валидация и нормализация введенных пользователем данных – 10 баллов
4. Использование передовых Data Science подходов – 10 баллов

Исходный датасет:

<https://disk.yandex.ru/d/HLUuy6dVxZUaAg>

Полезные материалы:

<https://zakupki.mos.ru/> - Сайт Портал Поставщиков

<https://www.promptingguide.ai/ru> - Мануал по промпт-инжинирингу