

Разработка агрегатора комментариев для администрирования массовых открытых онлайн курсов

Выполнил:

Фокин Константин Сергеевич, гр. 7303

Руководитель:

Заславский Марк Маркович, к.т.н., доцент

Актуальность

Обусловлена следующими причинами:

- Важность роли своевременной обратной связи между преподавателем и студентом в сфере онлайн-образования;
- Увеличение количества учащихся массовых открытых онлайн курсов (далее - MOOK) в связи с переходом образовательных организаций на дистанционный формат обучения;
- Необходимость снижения временных затрат преподавателей на администрирование MOOK.

Цель и задачи

Цель: разработать web-приложение для агрегирования комментариев MOOK выбранных платформ дистанционного обучения.

Задачи:

1. Исследовать существующие подходы к агрегированию информации для администрирования MOOK.
2. Сформировать набор требований для разработки web-приложения.
3. Разработать web-приложение и интегрировать его с выбранными платформами MOOK.
4. Оценить эффективность разработанного решения.

Задача 1: Исследование существующих подходов к агрегированию информации для администрирования MOOK.

Таблица 1: Результаты сравнения аналогов

Критерий / Аналог	Stepik	StepikBot	Moodle	Coursera	edX
Фильтрация комментариев	Да	Нет	Да	Да	Да
Сортировка комментариев	Нет	Нет	Нет	Да	Да
Информация о студенте	Да	Да	Да	Нет	Нет

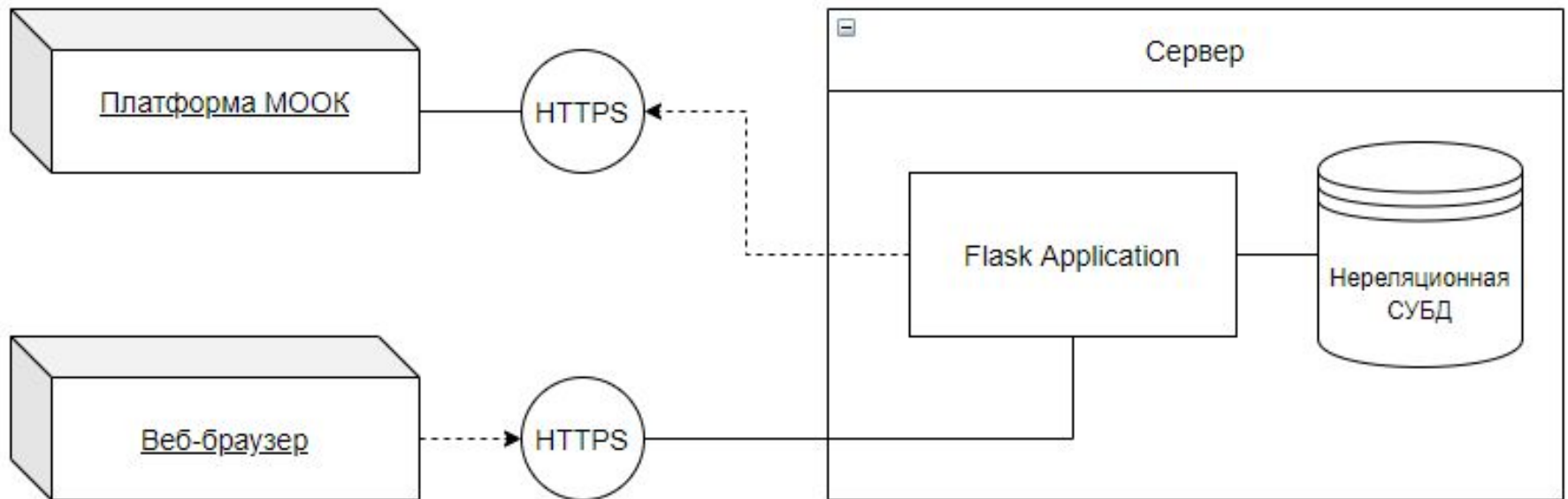
Платформы Stepik и Moodle являются прогрессирующими и широко используемыми для дистанционного обучения, занимая на данный момент лидирующие позиции по числу университетов-партнеров не только в России, но и во всем мире. Основываясь на их популярности и солидной функциональной базе, принято решение использовать форумы курсов данных платформ в качестве исходной базы комментариев для разрабатываемого web-приложения.

Задача 2: Формирование требований для разрабатываемого web-приложения

- Интеграция с платформами MOOK Stepik и Moodle;
- Импорт данных об администрируемых курсах (комментарии, студенты);
- Возможность осуществлять выборки (фильтрацию) комментариев сразу по нескольким параметрам (дата, количество ответов, курс, прогресс прохождения курса автора);
- Возможность выбора порядка отображения (сортировки) комментариев по параметрам из пункта выше;
- Импорт / экспорт настроек поиска для каждого пользователя;
- Время импорта 500 комментариев для выбранных платформ MOOK не должно превышать 120 секунд, чтобы выполнять условия высокой доступности системы (99% - 23 часа 45 минут в сутки), т.к. процесс загрузки данных блокирует систему;
- Время поиска 500 комментариев по заданным параметрам для комфортного взаимодействия с системой не должно превышать 3 секунд;
- Возможность осуществлять поиск комментариев одновременно с нескольких MOOK, расположенных на одной платформе.

Задача 3: Разработка web-приложения

Рисунок 1: Схема архитектуры web-приложения



Задача 3: Разработка web-приложения

Рисунок 2: Страница со списком курсов

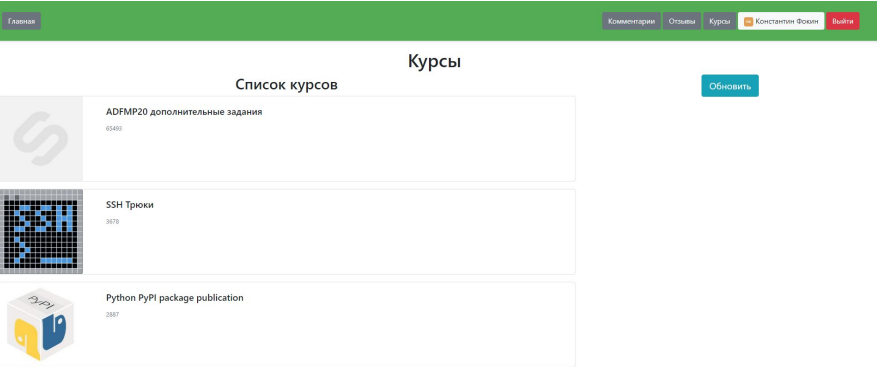


Рисунок 3: Страница со списком комментариев

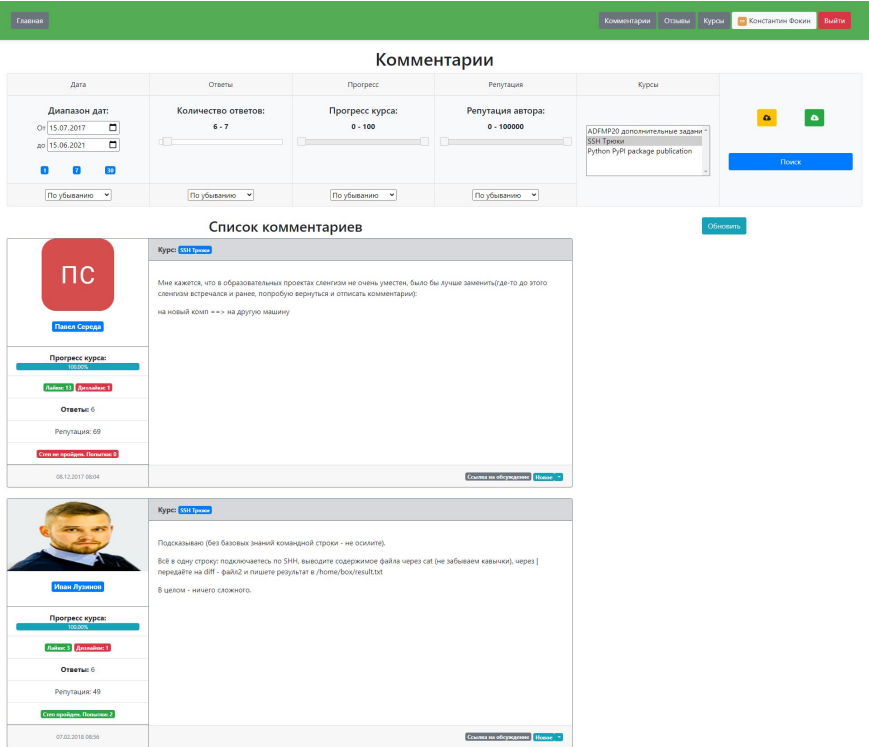
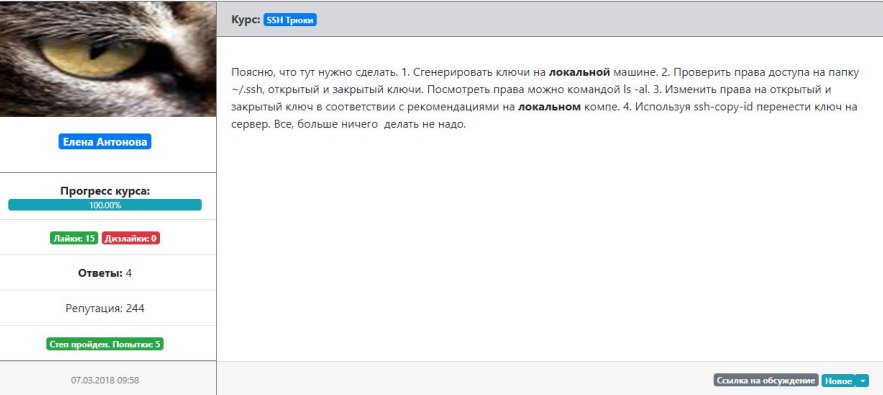


Рисунок 4: Карточка комментария



Задача 4: Оценка эффективности разработанного решения

Тестирование проведено на основе курсов:

- “SSH Трюки” и “Python PyPI package publication” - платформа Stepik;
- 3 тестовых курса - платформа Moodle.

Выполнены:

- Измерение количества действий преподавателя для всех сценариев использования;
- Измерение временных показателей загрузки и поиска комментариев для выбранных платформ MOOK (рис. 5).
- Оценка объема памяти, требуемого для хранения данных одного загруженного курса (комментарии, статистика студентов) в БД.

Рисунок 5: Измерение временных показателей

	Загрузка комментариев	Поиск комментариев
Требование	500 коммент. / 120 сек.	500 коммент. / 3 сек.
Stepik	504 коммент. / 104.3 сек.	504 коммент. / 2.52 сек.
Moodle	699 коммент. / 36.2 сек.	579 коммент. / 1.59 сек.

Задача 4: Оценка эффективности разработанного решения

Результаты:

- При выполнении любого сценария использования от пользователя требуется не более 4 действий;
- Количество действий для выполнения поиска комментариев в разработанном web-приложении (3) не превышает количество действий для выполнения той же операции на самих платформах (по 4).
- Время импорта комментариев с платформ MOOK и дальнейшего поиска в web-приложении соответствует выдвинутым при проектировании требованиям.
- Для хранения данных одного загруженного курса в среднем необходимо не более 4 МБ, а каждый следующий импортируемый курс будет занимать меньший объем на величину

$$V = N \times V_{User}, \text{ где:}$$

N - количество пользователей, чьи данные уже загружены в БД в рамках других курсов;

V_{User} - средний объем документа с данными пользователя для конкретной платформы MOOK.

Заключение

- Проведенный анализ используемых подходов к администрированию комментариев MOOK выявил необходимость разработки web-приложения, выполняющего функцию агрегатора комментариев для сокращения временных затрат на сопровождение MOOK;
- Выдвинут ряд требований для разрабатываемого web-приложения, а именно:
 - интеграция с платформами MOOK Stepik и Moodle,
 - импорт комментариев доступных пользователю MOOK,
 - расширенные возможности фильтрации и сортировки комментариев,
 - уменьшение числа действий пользователя по сравнению с инструментом платформы MOOK;
- Разработано и опубликовано web-приложение для администрирования комментариев с форумов и обсуждений платформ Stepik и Moodle;
- Исследование эффективности работы web-приложения показало, что его использование требует от преподавателя меньшего числа действий, а также оно предоставляет больше возможностей для направленного поиска комментариев;
- Дальнейшая работа по улучшению системы связана с обеспечением интеграции с другими образовательными платформами и выделением процесса импорта данных с платформ MOOK в отдельную подсистему.

Апробация работы

- Репозиторий проекта
https://github.com/moevm/bsc_fokin.
- Развернутое приложение в сети Интернет
<http://bsc-fokin.rc.robotbull.com>



DEMO

Запасные слайды

Оценка объема Stepik

Параметры курса:

- количество пользователей курса - $N = 50\,000$;
- пользователи, оставившие отзыв после прохождения курса - $R = 1\%$;
- количество комментариев к курсу относительно числа записанных пользователей - $C = 10\%$;
- среднее количество комментариев, оставленных одним автором - $U_c = 3$;

Коллекция	Объем документа
StepikCourse	1288 B
StepikComment	514 B
StepikReview	460 B
StepikUser	330 B

$$\begin{aligned} V &= V_{StepikCourse} + C \times N \times V_{StepikComment} + O \times N \times V_{StepikReview} + \frac{C \times N}{U_c} \times V_{StepikUser} = \\ &= 1288\,B + 0.1 \times 50\,000 \times 514\,B + 0.01 \times 50\,000 \times 460\,B + \frac{0.1 \times 50\,000}{3} \times 330\,B = \\ &= 3\,351\,288\,B \approx 3\,273\,KB \approx 3.2\,MB \end{aligned}$$

Оценка объема Moodle

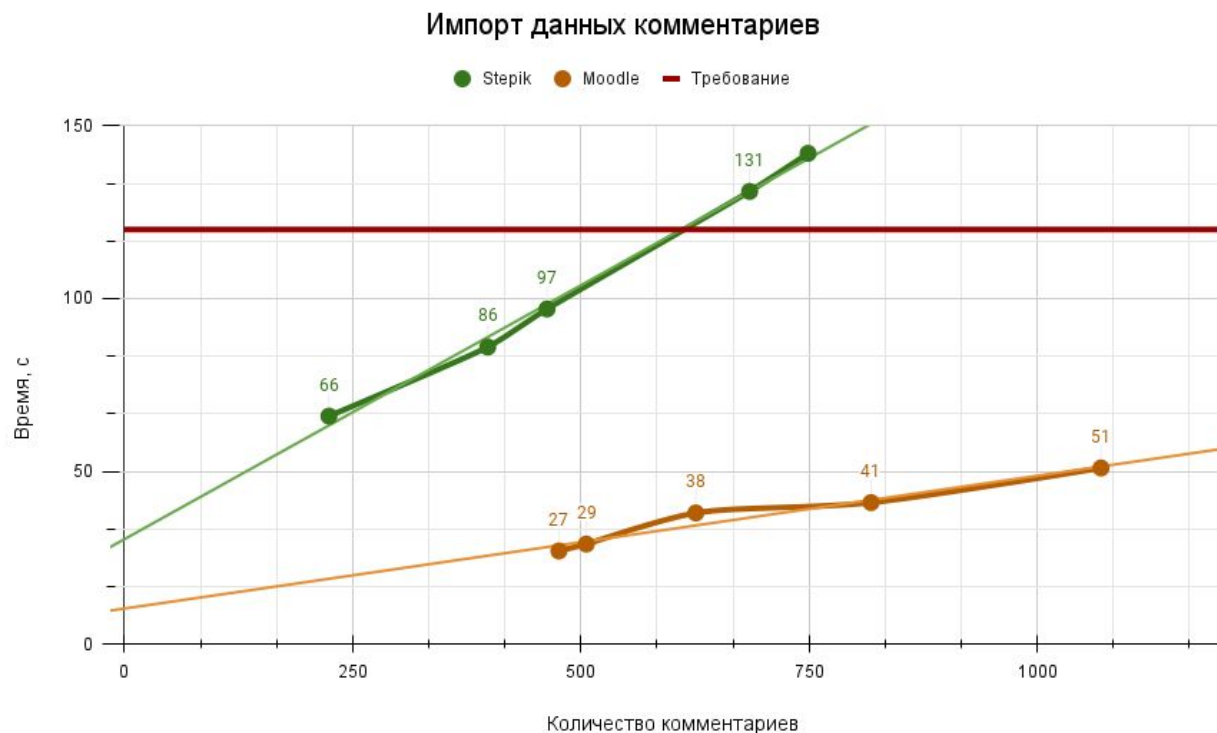
Параметры курса:

- количество форумов в курсе - $F = 10$;
- количество пользователей курса - $N = 50\,000$;
- количество комментариев к курсу относительно числа записанных пользователей - $C = 10\%$;
- среднее количество комментариев (обсуждения и посты), оставленных одним автором - $U_c = 3$;
- среднее количество постов, оставленных под одним обсуждением - 2, значит:
 - доля обсуждений среди всех комментариев - $D = 1/3$;
 - доля постов среди всех комментариев - $P = 2/3$;
- количество используемых тегов - $T = 20$;

Коллекция	Объем документа
MoodleCourse	1516 B
MoodleForum	2190 B
MoodleDiscussion	718 B
MoodlePost	737 B
MoodleUser	300 B
MoodleTag	31B

$$\begin{aligned}
 V &= V_{MoodleCourse} + F \times V_{MoodleForum} + C \times N \times D \times V_{MoodleDiscussion} + C \times N \times P \times V_{MoodlePost} + \\
 &+ \frac{C \times N}{U_c} \times V_{MoodleUser} + T \times V_{MoodleTag} = 1516 B + 10 \times 2190 + 0.1 \times 50\,000 \times \frac{1}{3} \times 718 B + \\
 &+ 0.1 \times 50\,000 \times \frac{2}{3} \times 737 B + \frac{0.1 \times 50\,000}{3} \times 300 B + 20 \times 31 B = \\
 &= 4\,177\,369 B \approx 4\,079 KB \approx 3.98 MB
 \end{aligned}$$

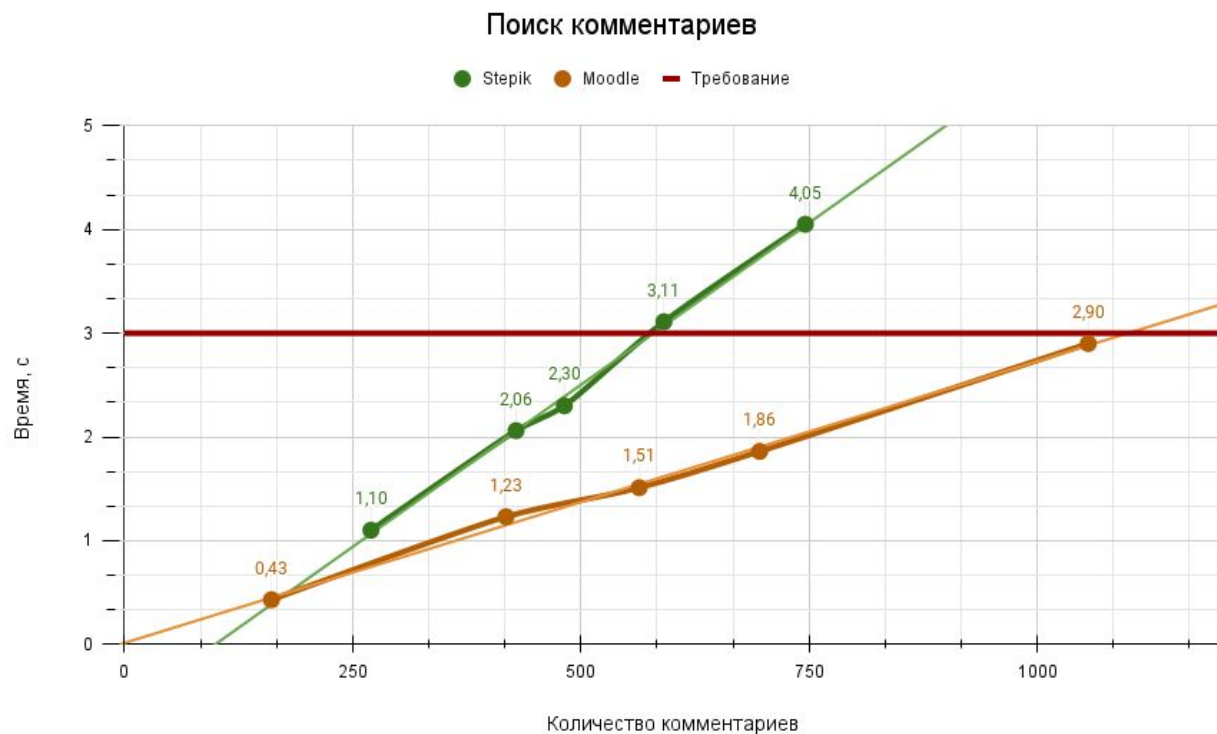
Измерение времени загрузки комментариев



Линии тренда:

- **Stepik** - $f(x) = 147 * 10^{-3}x + 30.4$;
- **Moodle** - $f(x) = 38.4 * 10^{-3}x + 10.4$.

Измерение времени поиска комментариев



Линии тренда:

- Stepik - $f(x) = (6.25x - 621) * 10^{-3}$;
- Moodle - $f(x) = (2.72x + 13.7) * 10^{-3}$.

Сценарии использования

Сценарии	Авторизация	Импорт курсов	Импорт комментариев	Поиск комментариев	Импорт настроек поиска	Экспорт настроек поиска
Кол-во действий	2-4	1	2	3	4	4

Используемые технологии

- **Stepik API**
- **Moodle API** - 3.8.2+ (Build: 20200312)
- **Python** - 3.8
- **Flask** - 1.0.2
- **MongoDB**
- **Gunicorn** - 19.9.0
- **Bootstrap** - 4.5.0
- **JavaScript**
- **JQuery** - 3.5.1
- **Docker** - 19.03.13