SimbirSoft

Что мы хотим от вас получить

Приложение, которое позволяет скачивать произвольную HTML-страницу посредством HTTP-запроса на жесткий диск компьютера и выдает статистику по количеству уникальных слов в консоль.

Требования к приложению

- 1. В качестве входных данных в приложение принимает строку с адресом web-страницы. Пример входной строки: https://www.simbirsoft.com/
- 2. Приложение разбивает текст страницы на отдельные слова с помощью списка разделителей. Пример списка: {' ', ',', '.', '!', '?','"', ';', ':', '[', ']', '(', ')', '\n', '\r', '\t'}
- 3. В качестве результата работы пользователь должен получить статистику по количеству уникальных слов в тексте. Пример: PA3PA6OTKA -1

ПРОГРАММНОГО - 2

05565545446

- ОБЕСПЕЧЕНИЯ 4
- 4. Приложение должно учитывать, что размер страницы может превышать размер доступной оперативной памяти. В таком случае вывод и подсчет статистики должен осуществляться корректно.
- 5. Приложение должно предусматривать возможность нештатных ситуаций и обеспечивать обработку исключений
- 6. Приложение должно быть реализовано с помощью стандартных классов (не стоит добавлять собственные реализации списков, словарей и т.п.)
- 7. Приложение написано в соответствии с принципами ООП
- 8. Приложение написано на языке выбранного направления (Java, C#)

Что нам понравится в приложении

Для зачисления на поток вполне достаточно, если ваше приложение работает и имеет более одного класса, но хорошим бонусом будет:

- 1. Хороший стиль кода, приближенный к общепринятым стандартам форматирования кода
- 2. Использование паттернов проектирования
- 3. Логирование ошибок в файл
- 4. Сохранение статистики в базу данных.
- 5. Возможность расширяемости проекта и многоуровневую архитектуру
- 6. Тесты

Как нам передать исходники?

Загрузите исходники вашего приложения в публичный репозиторий https://github.com/, снабдив его README-файлом. Данный файл обязательно должен содержать инструкцию по запуску и использованию приложения. Ссылку на репозиторий необходимо отправить на internship.backend@simbirsoft.com.