

Задание #1 к модулю HTTP fundamentals

Задание

Необходимо реализовать библиотеку и использующую её консольную программу для создания локальной копии сайта («аналог» программы [wget](#)).

Работа с программой выглядит так: пользователь указывает стартовую точку (URL) и папку куда надо сохранять, а программа проходит по всем доступным ссылкам и рекурсивно выкачивает сайт(ы).

Опции программы/библиотеки:

- ограничение на глубину анализа ссылок (т.е. если вы скачали страницу, которую указал пользователь, это уровень 0, все страницы на которые введут ссылки с неё, это уровень 1 и т.д.)
- ограничение на переход на другие домены (без ограничений/только внутри текущего домена/не выше пути в исходном URL)
- ограничение на «расширение» скачиваемых ресурсов (можно задавать списком, например так: gif,jpeg,jpg,pdf)
- трассировка (verbose режим): показ на экране текущей обрабатываемой страницы/документа

Рекомендации по реализации

В качестве основы можно взять следующие библиотеки:

- работа с HTTP
 - System.Net.Http.HttpClient – рекомендуемый вариант
 - Если вы работаете с .Net 4.5 + он включен в сам фреймворк. В более ранних версиях и для прочих платформ получите через [NuGet](#)
 - Введение в работу с ним можно найти тут <https://blogs.msdn.microsoft.com/henrikn/2012/02/16/httpclient-is-here/>
 - Обратите внимание – он весь построен на асинхронных операциях (но мы можем работать в синхронном режиме!)
 - System.Net.HttpWebRequest – legacy
- Работа с HTML
 - Можно воспользоваться одной из библиотек, перечисленных [тут](#)
 - Самый популярный вариант HtmlAgilityPack, хотя он достаточно и старый и имеет свои проблемы.