

Написать программу (парсер), осуществляющую сбор данных о компаниях с ресурса wikipedia.org.

Табличные данные, включая изображения (логотипы организации), отображаемые в правом верхнем углу сайта после перехода на соответствующую страницу организации, представить в виде xml документа, по одному документу на каждую организацию. Изображения предлагается сохранять в том же документе XML в кодировке base64.

Будет предоставлен файл со списком компаний, по которым необходимо собрать полезную информацию.

Информация, которую необходимо собрать: Название, логотип, Год основания, Расположение, Ключевые фигуры, Уставной капитал. Некоторая информация может отсутствовать. И информация по некоторым компаниям может отсутствовать вообще(Страница не будет найдена).

Таким образом, общий алгоритм работы следующий:

- 1) Получаем очередное название компании из входного файла (input.txt)
- 2) Загружаем веб страницу и парсим необходимую информацию
Ссылка для получения страницы компании:
<https://ru.wikipedia.org/wiki/Special:Search?search=COMPANY&go=Go>
где вместо COMPANY должна быть подставлено наименование компании.
- 3) Сохраняем информацию в виде xml файла.

Желательно (будет преимуществом, но не обязательно) удовлетворить следующие дополнительные требования:

- 1) Программа должна быть реализована многопоточно
- 2) Использовать очередь поставщик-потребители: поставщик непрерывно формирует объекты с нужной информацией, помещая в очередь, в то время как потребители (параллельно) извлекают их из нее и сохраняют в виде xml документа

Результат работы должен быть запакован в ZIP архив и отправлен ответным письмом не позднее 48 часов с момента получения задания.

Оцениваться будет:

- Полнота и правильность решения
- Качество кода
- Выполнение дополнительных требований


Альфа-банк	
 Альфа-Банк	
Тип	Акционерное общество
Лицензия	№ 1326 от 16.01.2015 ↗
Участие в ССВ	№ 309 от 16.12.2004 ↗
Год основания	1990
Расположение	 Москва (Россия)
Ключевые фигуры	Пётр Авен (президент), Андрей Соколов (председатель правления), Йоханн Йонах (председатель совета директоров)
Уставный капитал	 62,7 млрд руб. (31.12.2014, МСФО) ^[1]
Собственный капитал	 216,6 млрд (31.12.2014, МСФО)
Активы	 2,4 трлн руб. (31.12.2014, МСФО)
Чистая прибыль	 53,0 млн (2014, МСФО)
Число сотрудников	24 990 ^[2]
Основное общество	ОАО «АБ Холдинг»
Аудитор	БДО Юникон (РСБУ) ПрайсуотерхаусКуперс Аудит (МСФО)
Источники данных о банке	200 крупнейших российских банков по размеру чистых активов (на 1 января 2010 года) ↗
Сайт	alfabank.ru ↗

Рисунок 1. Пример компании Альфа Банк