Отчет 3-4

Слайды 1-3. Архитектура программного ПО.

На первом слайде представлена. Диаграмма пакетов (модулей). Модуль Main связующий, который будет соединять, util модуль для вспомогательных компонентов. И 3 основных модуля:

Parsers модуль, отвечающий за сбор источников и текстов из интернета. На данный момент этот модуль уже по большей части завершён.

Следующий модуль это processing, модуль отвевающий за извлечение информации из полученных ранее текстов. Самый объемный по содержанию модуль. На данный момент основная работа ведётся именно над ним.

И последний модуль, это review. Модуль, в котором будет формироваться конечный отчёт.

На слайдах 2 и 3 показаны диаграммы объектов модуля parsing и processing.

Сбор источников происходит в интернете с помощью поисковых систем Гугл и Яндекс. Список поисковых запросов перечислен на слайде на слайде 4. Всего для начала было собранно 527 источников. Мы отсеяли магазины и маркетплейсы, так как нам не нужна информация, содержащаяся там. Так же убрали нерелевантные ссылки, в которых нет значимой для нас информации. Осталось около 310 источников.

На слайдах 5-8 показаны пару примеров собранной информации.

На слайде 5 представлена часть информации о русских организациях, занимающихся производством кабельной продукции. Она пока представлена в виде таблицы. Здесь только основная информация об организациях, так как многое недоступно физическим лицам в открытом доступе. Из недоступной информации - юридические дела компании, их долги, обороты и тд. Либо даже если и можно собрать всю эту информацию, то разработка модуля, который отвечал бы за это, заняло бы у нас очень много времени, так как скорее всего такая информация содержится только в государственных реестрах и не факт, что она легко доступна и на это нет времени. Но можно рассмотреть вариант с использованием спец сайтов, которые агрегируют и дают доступ к этой информации, это было бы гораздо проще, а стоимость таких ресурсов в районе 500-1000 рублей в месяц.

На слайде 6 информация о зарубежных компаниях. Название и адрес по аналогии с эталонной презентацией.

На слайде 7. Пример информации, которая была получена из гостов, это классификация кабелей и определение кабельной продукции.

На слайде 8 пример с графиками, тут представлены котировки алюминия. По этим графикам можно отследить динамику стоимости алюминия по годам. Так же можно провести какой-то анализ, рассчитать базовую стоимость кабельной продукции.