Отчет 5

Возможно получить структурированную информацию и слабо структурированную информацию. Это:

* Информация о российских организациях.
* Информация об иностранных организациях.
* Котировки металлов.
* Курсы валют.
* Определение кабельной продукции (ГОСТ).
* Классификация кабельной продукции (ОКПД2).

Также можно дополнительно получить на основе этих данных:

* Оценка себестоимости кабеля.
* Прогнозирование котировок металлов.
* Оценка благонадежности контрагентов.

Что не удается полностью или частично извлечь из открытых источников:

* Мировой спрос на кабельную продукцию.
* Спрос кабельной продукции по странам.
* Динамика производства силового кабеля в России.
* Баланс спроса и предложения силового кабеля в России.
* Экспорт и импорт силового кабеля в России.
* Отраслевая структура потребления силового кабеля в России.
* Закупки силового кабеля в Газпром Нефть.
* Контрафакт силового кабеля в России.

Причины, по которым возникли проблемы с получением данной информации из отрытых источников.

Во-первых, большую часть информации невозможно найти в открытых источниках. Многие данные представленные в презентации, вроде мирового спроса на кабельную продукцию не ищутся с помощью различных комбинаций поисковых запросов. Либо данной информации нет в открытом доступе, либо она очень хорошо запрятана.

Во-вторых, информация по большей части закрыта, она содержится в первоисточниках, к которым нет доступа, вроде отчётных документов организаций.

В-третьих, информация из открытых источников сильно не структурирована. Учитывая этот и предыдущие факторы, мы получаем то, что ключевую информацию трудно извлечь из открытых источников. Мы можем собрать несколько десятков текстов, из них только в одном будет содержаться интересующий нас факт, и при извлечении информации из текста, мы либо соберем очень много лишней информации вместе с нужным фактов, либо пропустим все и не соберем ничего.

В-четвертых, даже та небольшая часть информации что возможно найти не актуальна на текущий день. Это старые архивные данные, которые по той или иной причине попали в открытый доступ и некоторые новые статьи публикуются на основе этих исторических данных.