Техническое задание для реализации парсера сайта <https://www.walter-tools.cn/ru-ru>

С целью получения данных с сайта https://www.walter-tools.cn/ru-ru, необходимо реализовать парсер электронного каталога продукции с сайта компании walter.

Данный сайт в данный момент из русскоязычного сегмента выдает 403 ошибку, поэтому для захода следует использовать VPN.

**Выбор инструментария:**

В целом можно использовать любой инструмент для забора данных, приоритет в реализации отводятся к стеку PHP + CURL. Необходимо чтобы была возможность использовать инструмент без привлечения разработчика, просто запуская парсер.

**Какую информацию необходимо парсить:**

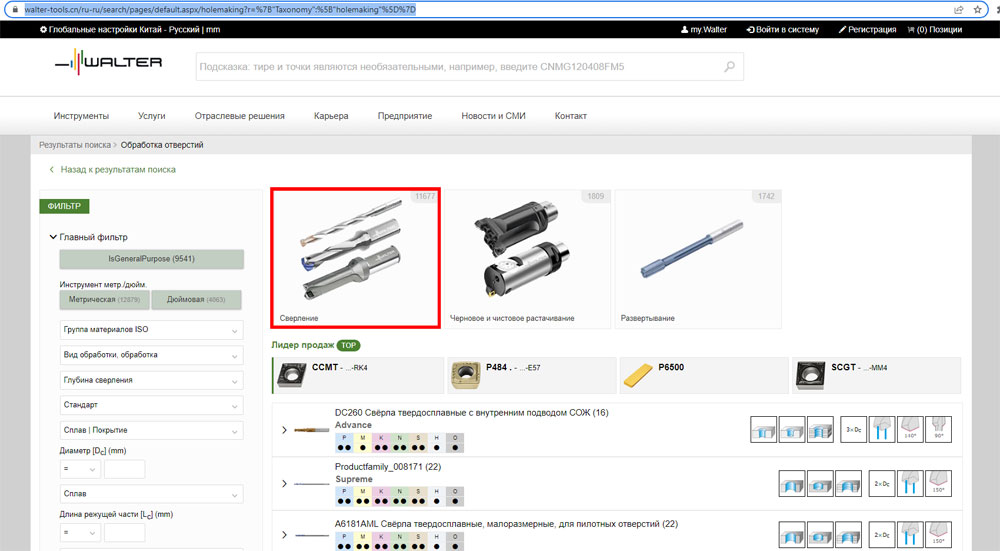
Требуется получить иерархическую структуру Категорий и карточек товара, находящихся в данной категории, с их свойствами.

Первостепенная задача получить информацию по категории – Обработка отверстий - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking?r=%7B%22Taxonomy%22:%5B%22holemaking%22%5D%7D>

Что нужно получить с точки зрения иерархической структуры категорий на примере категории Сверление - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking/drilling-from-solid?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%22%5D%7D>:

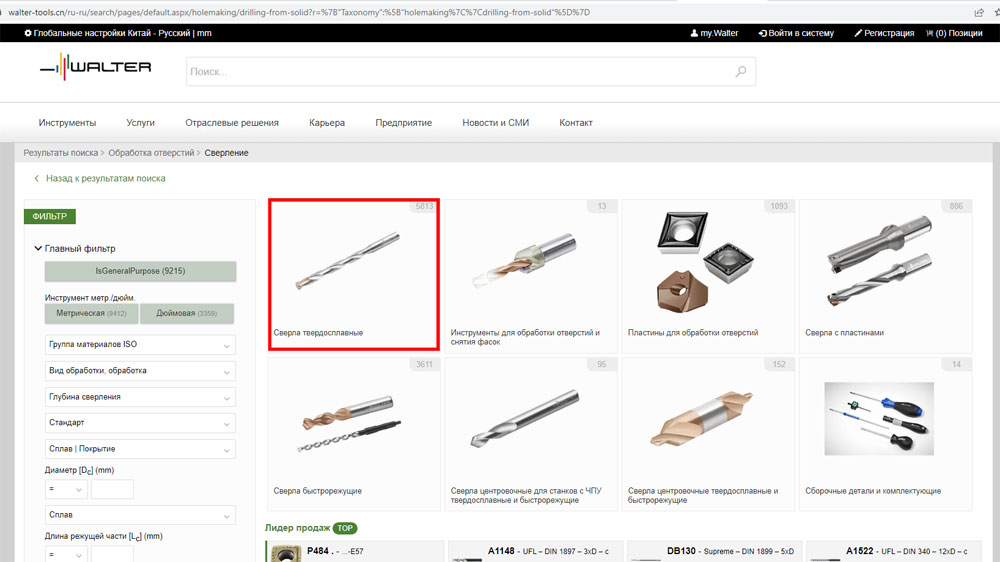
Уровень 1 категории

Сверление - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking/drilling-from-solid?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%22%5D%7D>



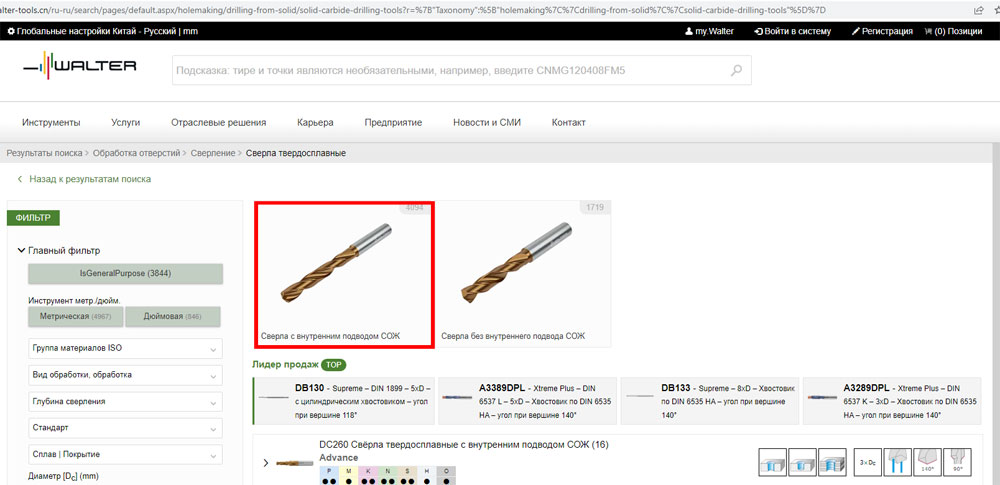
Уровень 2 категории

Сверла твердосплавные - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking/drilling-from-solid/solid-carbide-drilling-tools?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%7C%7Csolid-carbide-drilling-tools%22%5D%7D>



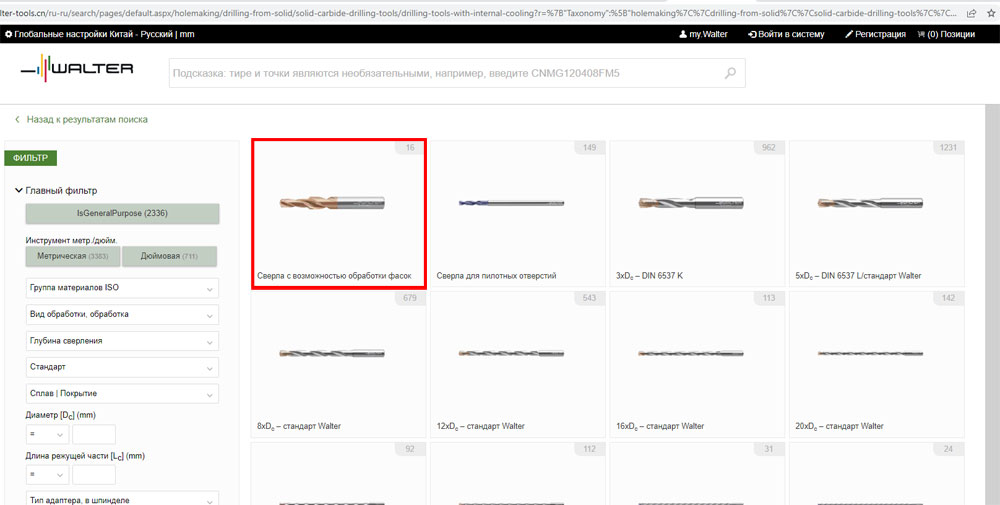
Уровень 3 категории

Сверла с внутренним подводом СОЖ - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking/drilling-from-solid/solid-carbide-drilling-tools?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%7C%7Csolid-carbide-drilling-tools%22%5D%7D>



Уровень 4 категории

Сверла с внутренней обработкой фасок: - <https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/holemaking/drilling-from-solid/solid-carbide-drilling-tools/drilling-tools-with-internal-cooling?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%7C%7Csolid-carbide-drilling-tools%7C%7Cdrilling-tools-with-internal-cooling%22%5D%7D>

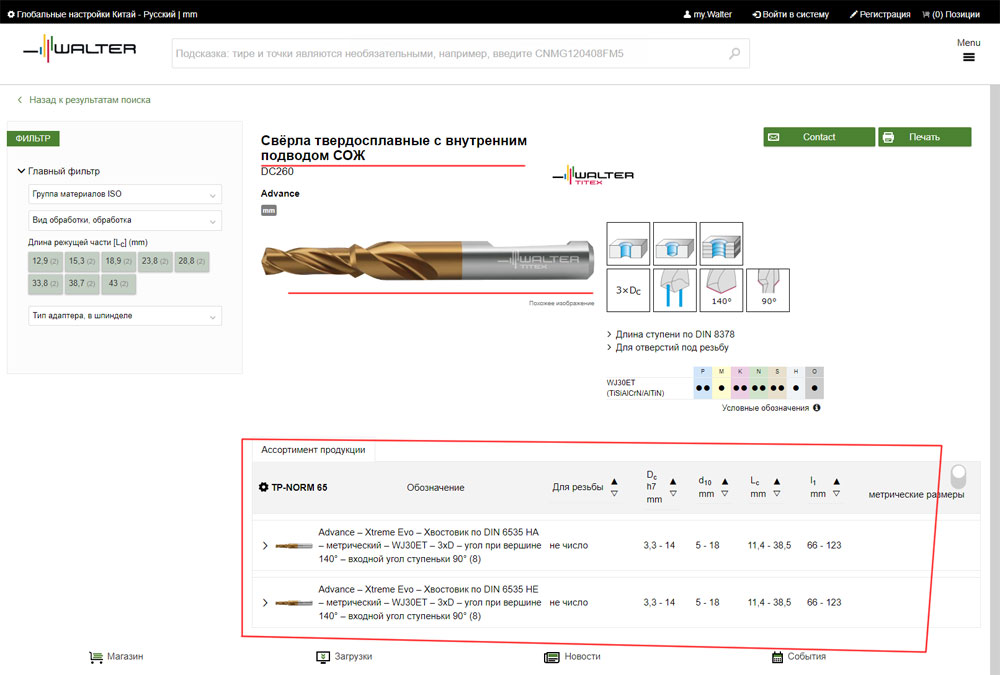


***Для иерархии категорий, необходимо сохранить Наименование категории и ее изображение.***

При достижении уровня карточки товара, необходимо сохранить следующие данные:

Пример карточки товара:

<https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/family/006644_120998_T62?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%7C%7Csolid-carbide-drilling-tools%7C%7Cdrilling-tools-with-internal-cooling%7C%7Cchamfer-drills%22%5D%7D>



Карточка товара состоит из информации об изделии и ассортимента данного изделия.

Требуется сохранить Наименование товара и Изображение товара

Для ассортимента продукции потребуется сохранять следующие данные:

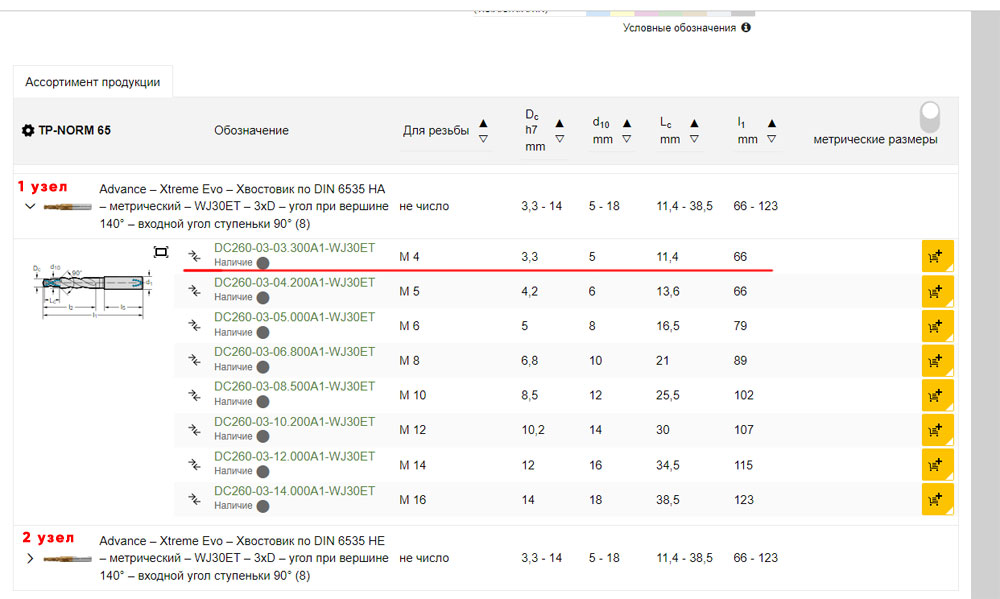
Каждая позиция ассортимента является отдельным узлом иерархии:

Пример карточки товара – Сверла твердосплавные с внутренним подводом СОЖ:

<https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/family/006644_120998_t62?r=%7B%22Taxonomy%22:%5B%22holemaking%7C%7Cdrilling-from-solid%7C%7Csolid-carbide-drilling-tools%7C%7Cdrilling-tools-with-internal-cooling%7C%7Cchamfer-drills%22%5D%7D>

Каждый узел нужно сохранить следующим образом:

Наименование узла и его свойства:



**1 узел**

Advance – Xtreme Evo – Хвостовик по DIN 6535 HA – метрический – WJ30ET – 3xD – угол при вершине 140° – входной угол ступеньки 90° (8)

Свойства (<https://www.walter-tools.cn/ru-ru/search/pages/default.aspx/product/dc260-03-03.300a1-wj30et>): (**с данной страницы, необходимо сохранить артикул узла!!!**)

|  |  |  |
| --- | --- | --- |
| Описание | Условное обозначение | **Значение** |
| Для резьбы | Для резьбы | **M 4** |
| Диаметр резания (h7) | Dc | **3,3**mm |
| DC01WKPS01000000 | d10 | **5**mm |
| SDL | Lc | **11,4**mm |
| Общая длина | l1 | **66**mm |
| Длина стружечной канавки | l2 | **28**mm |
| Длина хвостовика | l5 | **36**mm |
| Диаметр хвостовика (h6) | d1 | **6**mm |

**Варианты предоставления полученных данных:**

1. **SQL (желательный формат)**

**Представляет собой dump базы данных в формате SQL. Содержит следующие таблицы и поля:**

**Таблица категорий:**

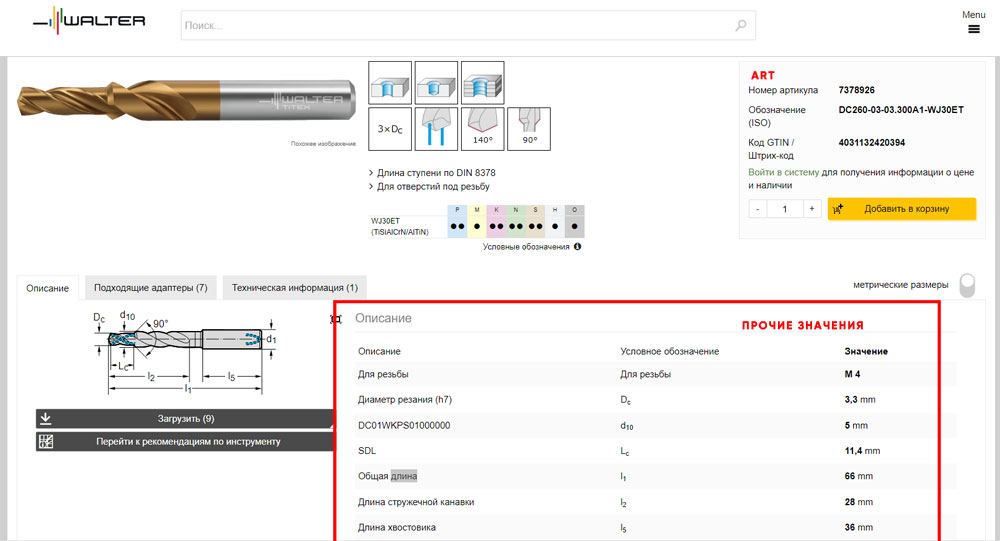
|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **PID** | **NAME** | **IMG** |
| **Автоинкрементное поле** | **ID родителя** | **Наименование категории** | **Изображение категории** |

**Таблица карточек товара:**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | CID | NAME | **IMG** |
| **Автоинкрементное поле** | ID категории | Наименование карточки товара | **Изображение карточки товара** |

**Таблица ассортимента:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | IID | NAME | ART | Прочие значения |
| **Автоинкрементное поле** | ID карточки товара | Наименование ассортимента | Артикул | Значения свойств – каждое в своем поле |



**Так же допускаются и другие форматы выходного файла (XLSX, XML) при соблюдении иерархии и отношений сущностей, описанных в SQL реализации.**