# Модуль Парсер-менеджер товаров с торговых площадок

Предназначен для обновления и добавления товаров с торговых площадок Яндекс-маркет(категории электроника, бытовая техника, компьютеры и некоторые дугие), Hotline.ua, Onliner.by . В дальнейшем возможно добавление других сайтов.

Возможности:

1. Добавление, обновление товаров .
2. Автоматический поиск товара на сайте-доноре по названию товара или модели.
3. Сохранение постоянных ссылок на товары с сайта донора, для дальнейшего использования ссылки.
4. Гибкое использование фильтров для поиска редактируемых товаров, по отсутствию характеристик, описания, цены, изображения или целой категории товаров.
5. Возможность выбора данных для обновления, данные указываются для каждого источника, с одного можно парсить атрибуты с другого, цену, описание и тд.
6. Возможность дополнять или заменять данные в товарах.
7. Использование прокси-сервера или указание целого листа с прокси-серверами. С последующей проверкой доступности сайта –донора (Яндекс.Маркет ). После проверки прокси сортируются по времени доступа.
8. Контроль при парсинге нескольких товаров, если произошло прерывание парсинга из-за обрыва, таймаута, их можно допарсить позже.
9. Подробный лог парсинга.

Внимание! Необходима библиотека CURL в PHP.

Установка:

Копирование в корень магазина. После запуска модуля создастся две таблицы в БД, одна для хранения ссылок на товары, другая для хранения ссылок при парсинге нескольких товаров.

Желательно добавить в файл admin/model/catalog/product.php в функции deleteProduct() строку

$this->db->query("DELETE FROM " . DB\_PREFIX . "parser\_urls WHERE product\_id = '" . (int)$product\_id . "'");

Для удаления ссылок при удалении товара.

Если модуль удаляется эту строку тоже нужно удалять.

Принцип работы модуля:

После установки модуля необходимо сделать первоначальные настройки для парсера и для сайтов-доноров, указывается сайт источник.

Обязательно нажать «Сохранить». После настроек модуль готов к работе. Будут показаны товары, указанные в фильтре, или все если фильтр не используется.

Обновление товаров.

Для поиска товара необходимо нажать «Поиск» и будет произведен поиск товара по названию или модели. В появившемся списке товаров выбирается нужный, после чего ссылка на товар постоянно запомнится в парсере. Если поиск не дал результатов, можно ввести прямую ссылку на товар, скопировав ее с сайта источника. Для этого нужно нажать «+». Также есть возможность поиска товара по произвольному запросу(не по названию товара или модели). Для это нужно нажать «Ввести запрос» и ввести в поле, что необходимо найти.

После появления ссылки с строке товара, товар можно парсить, есть возможность указать новую цену для товара, предварительно убрав галочку «Заменять цену» в настройках .

Также возможен парсинг нескольких товаров, для этого нужно выбрать товары и нажать «Парсить все». Будут спарсены товары у которых есть ссылки.

ВАЖНО! Парсинг нескольких товаров может быстро привести к бану, в особенности Яндексом, поэтому используйте эту функцию, если имеете прокси-сервера.

При использовании фильтров по пустым атрибутам, изображениям и тд., после парсинга данных значений они не будут отображаться в парсере.

Добавление новых товаров.

Товары будут добавляться с выбранного сайта-источника. Для поиска товара надо нажать «Найти товар», можно сразу ввести прямую ссылку, нажав «+». После появления ссылки, можно парсить товар, нажав «Добавить товар». Для новых товаров можно указать категорию для отображения.

# Парсинг с яндекса через прокси с куками

# Парсинг через прокси с Яндекса может не работать без кук. Поэтому следует добавить куки из браузера Firefox в парсер:

# Устанавливаем дополнение менеджер кук для Firefox <https://addons.mozilla.org/en-US/firefox/addon/cookie-manager/> .

# Выбираем на сайте яндекса регион, с которого будем парсить и ходим по страницам.

# Находим значок менеджера в правом нижнем углу и жмем его.

# Находим экспорт кук, сохраняем в файл cookie.txt.

# 

# Заливаем файл себе на хостинг в папку магазина system/logs/