

Парсер (Parser)

Синонимы: Синтаксический анализатор

Парсер — программа, которая позволяет собирать, систематизировать, анализировать данные из определенных источников и объединять их в базы данных для представления в различной форме. Соответственно парсингом называют сам процесс извлечения и обработки информации.

В работе парсера выделяют следующие этапы:

1. В начале задаются необходимые настройки под конкретную задачу и выбирается область, в которой будет выполняться поиск;
2. Далее программа извлекает информацию в соответствии с заданными ранее параметрами;
3. Затем проводится анализ найденных данных и их систематизация;
4. В конце пользователь получает отчет в нужном ему виде.

К преимуществам синтаксического анализа можно отнести следующее:

- **Сбор большого объема данных.** С помощью парсинга можно собирать и анализировать данные по всей глобальной сети. Для человека такой объем работы не представляется возможным.
- **Автоматизация работы.** Программы-парсеры выполняют работу самостоятельно в автоматическом режиме, что снижает нагрузку на сотрудников и освобождает их от рутинных операций, оставляя таким образом больше времени на выполнение своих задач.
- **Поиск ошибок.** При определенной настройке парсеры могут выявлять ошибки на сайтах, в базах данных или на каких-либо веб-сервисах.

Основным недостатком можно считать то, что после сбора информации не всегда можно получить желаемый результат. В основном это зависит от того, как настроена программа и от ее возможностей. Также среди минусов — запрет некоторых сайтов на автоматический сбор данных.

Синтаксические анализаторы могут быть как компьютерным приложением, так и облачным сервисом. Также парсер может использоваться как расширение для браузера или даже как надстройка для Excel.

Чаще всего парсеры используют для решения следующих задач:

- **Преобразование данных.** Конвертация информации из одного формата в другой. Например, трансформация HTML-файла в JSON или преобразование JSON формата

для загрузки в хранилище данных.

- **Мониторинг цен.** Сбор информации о ценах, которые конкуренты устанавливают на какой-либо товар.
- **Технический аудит сайта.** Программы-парсеры могут отыскивать нерабочие ссылки или страницы на сайте.
- **Заполнение каталога товаров.** Консолидация информации по всем товарным позициям, представленным в открытых каталогах.
- **Сквозная аналитика.** Сбор информации из множества каналов рекламы и объединение в один набор данных.

Таким образом, парсер — это программа или сервис, позволяющий сэкономить много времени и сил сотрудников за счет автоматического выполнения большого объема однотипной работы. Но для эффективного сбора и изучения контента необходимо подбирать парсер под определенную задачу и задавать нужные параметры.