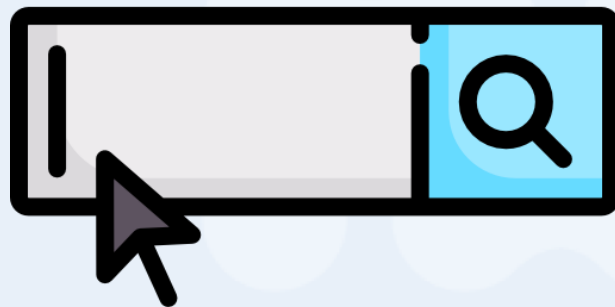




Поисковые системы и работа с ними

Поисковая система

это компьютерная система, предназначенная для поиска информации. Одно из наиболее известных применений поисковых систем — веб-сервисы для поиска текстовой или графической информации во Всемирной паутине.



Типы поисковых систем

Краулер («бот», «робот» или «паук»), индекс и программное обеспечение поисковой системы.

Системы, управляемые человеком (каталоги ресурсов) Эти поисковые системы получают списки веб-страниц.

Гибридные системы Такие поисковые системы, как Yahoo, Google, MSN, сочетают в себе функции систем, использующие поисковых роботов, и систем, управляемых человеком.

Мета-системы Метапоисковые системы объединяют и ранжируют результаты сразу нескольких поисковиков.

Популярные ошибки

400 - ошибка запроса серверу

401 - неправильные имя/пароль

403 - ошибка авторизации

404 - ресурс не найден

500 - внутренняя ошибка сервера

503 - ошибка маршрутизации запроса

Протоколы поисковых систем

HTTP соединение – это прикладной протокол для переноса информации, применяемый во время получения данных с интернет-сайтов.

HTTPS соединение – дополнительное расширение HTTP, которое поддерживает кодирование по протоколам SSL и TLS.

Техническое отличие – HTTP обычно использует порт соединения 80, тогда как HTTPS – порт 443

HTTPS используется для обеспечения конфиденциальности в процессе обмена данными между документами или клиентами сети гипертекстовых документов.

Изначально защищенный протокол разрабатывался конкретно для защиты персональных данных пользователей веб-браузеров. В общем плане может использоваться для шифрования любых сетевых трансферов по глобальной или локальной сети.

SSL

SSL (secure sockets layer — уровень защищённых сокетов) представляет собой криптографический протокол для безопасной связи. С версии 3.0 SSL заменили на TLS (transport layer security — безопасность транспортного уровня), но название предыдущей версии прижилось, поэтому сегодня под SSL чаще всего подразумевают TLS.

Цель протокола — обеспечить защищенную передачу данных. При этом для аутентификации используются асимметричные алгоритмы шифрования (пара открытый — закрытый ключ), а для сохранения конфиденциальности — симметричные (секретный ключ). Первый тип шифрования более ресурсоемкий, поэтому его комбинация с симметричным алгоритмом помогает сохранить высокую скорость обработки данных.

Шифрование, аутентификация и целостность

Протокол TLS предназначен для предоставления трёх услуг всем приложениям, работающим над ним, а именно: шифрование, аутентификацию и целостность. Технически, не все три могут использоваться, однако на практике, для обеспечения безопасности, как правило используются все три:

- **Шифрование** – сокрытие информации, передаваемой от одного компьютера к другому;
- **Аутентификация** – проверка авторства передаваемой информации;
- **Целостность** – обнаружение подмены информации подделкой.

Принцип работы

Запрос, который идет от пользователя подвергается морфологическому анализу. Информационное окружение любого документа, имеющегося в базе, генерируется (оно и будет в дальнейшем отображаться как сниппет, т.е. информационное поле текста соответствующего данному запросу).

Полученные данные передают как входные параметры специализированному модулю ранжирования. Они обрабатываются по всем документам, и в итоге для каждого такого документа рассчитывается свой рейтинг, который характеризует релевантность такого документа запросу пользователя, и иных составляющих.

В зависимости от условий заданных пользователем этот рейтинг вполне может быть подкорректирован дополнительными.

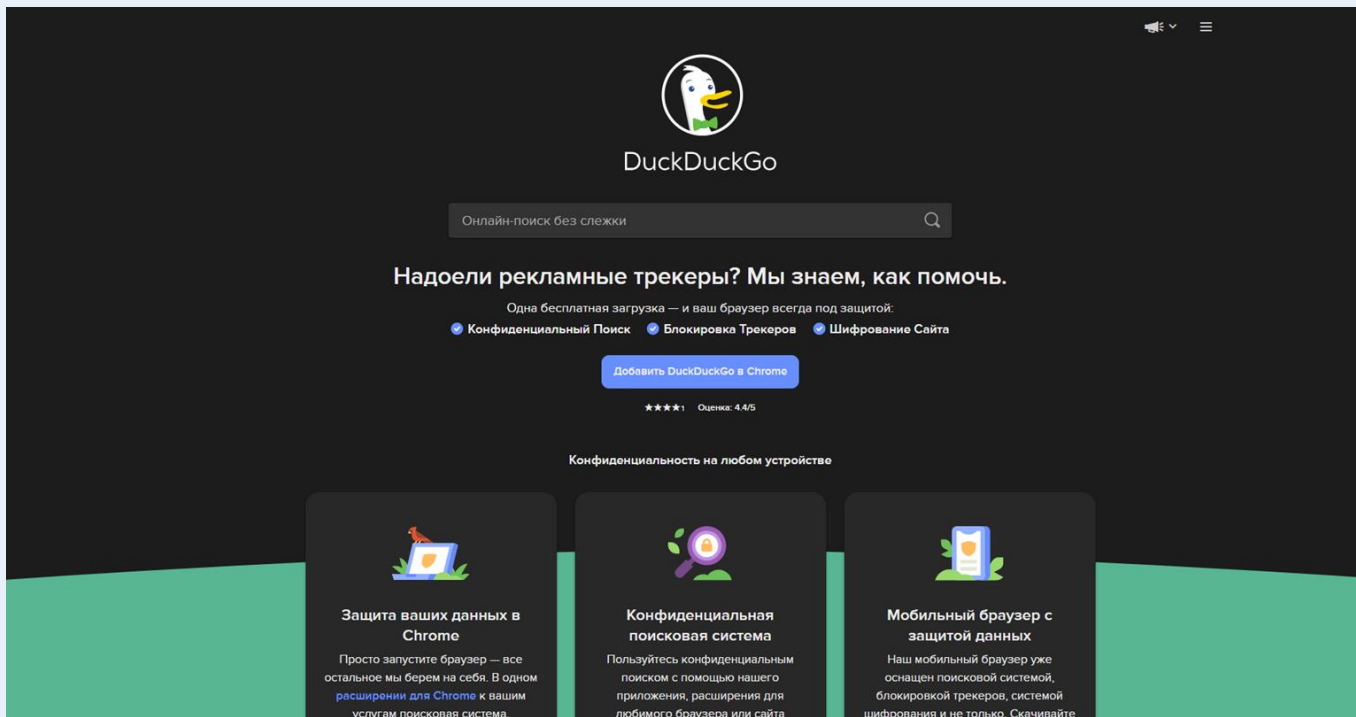
Затем генерируется сам сниппет, т.е. для любого найденного документа из соответствующей таблицы извлекают заголовок, аннотацию, наиболее отвечающую запросу, и ссылка на этот документ, при этом найденные словоформы и слова подсвечивают.

Результаты полученного поиска передаются осуществившему его человеку в виде страницы, на которую выдают поисковые результаты (SERP).

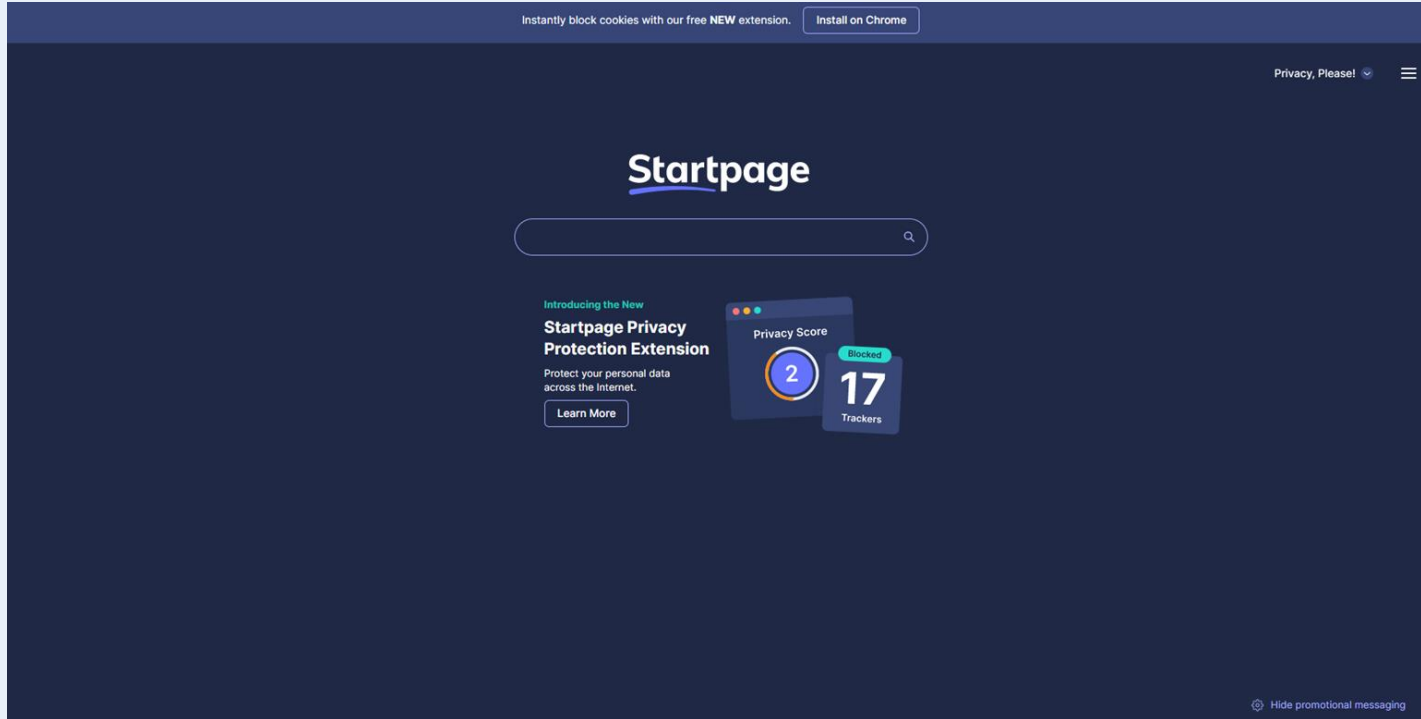
Анонимные

- DuckDuckGO
- StartPage
- TOR

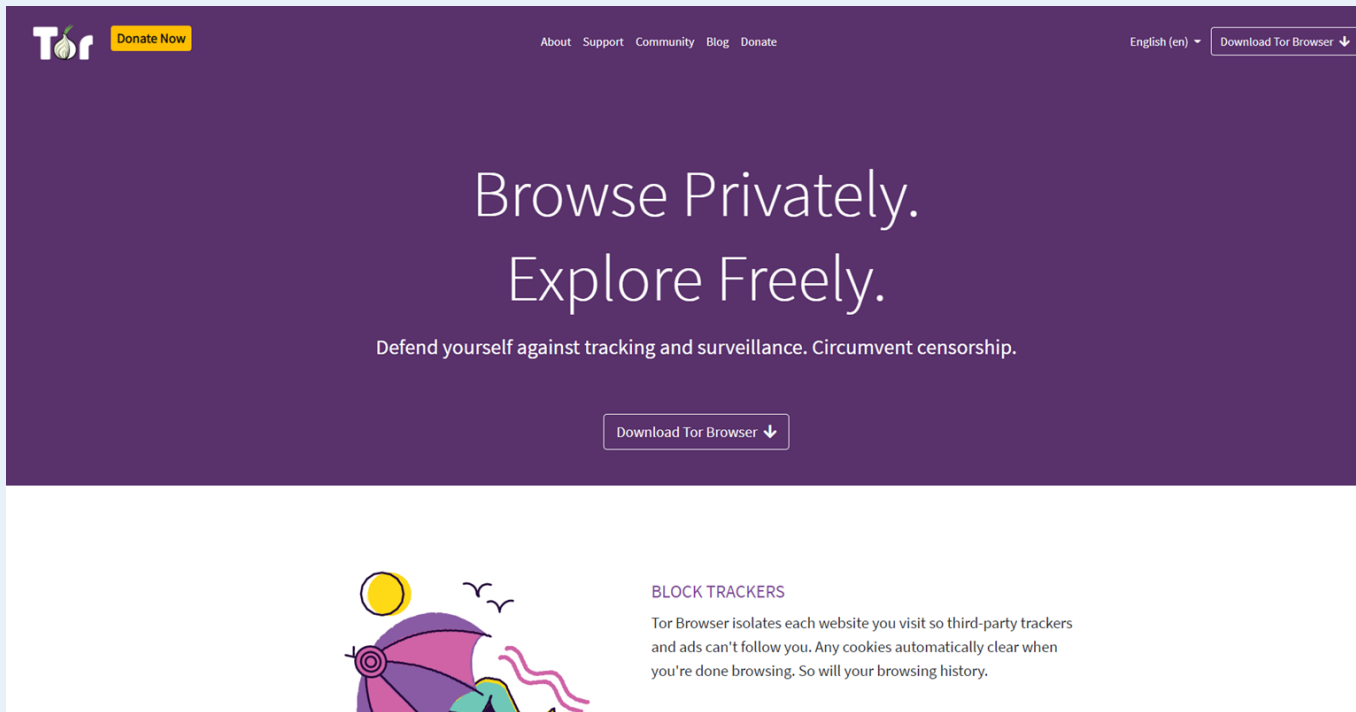
DuckDuckGo



Startpage



TOR



The screenshot shows the Tor Project's homepage. The top navigation bar is dark purple with the Tor logo, a 'Donate Now' button, and links for 'About', 'Support', 'Community', 'Blog', and 'Donate'. On the right, there's a language selector set to 'English (en)' and a 'Download Tor Browser' button. The main content area has a large purple background with the text 'Browse Privately. Explore Freely.' and 'Defend yourself against tracking and surveillance. Circumvent censorship.' Below this is another 'Download Tor Browser' button. The bottom section is white and features an illustration of a person under a purple umbrella with a yellow sun and birds in the background. To the right of the illustration is the heading 'BLOCK TRACKERS' and a paragraph explaining how Tor Browser isolates websites to prevent third-party trackers and ads from following the user.

Tor [Donate Now](#)

[About](#) [Support](#) [Community](#) [Blog](#) [Donate](#)

English (en) [Download Tor Browser](#)

Browse Privately. Explore Freely.

Defend yourself against tracking and surveillance. Circumvent censorship.

[Download Tor Browser](#)

BLOCK TRACKERS

Tor Browser isolates each website you visit so third-party trackers and ads can't follow you. Any cookies automatically clear when you're done browsing. So will your browsing history.

Практическое задание



Полезные ссылки

Поисковые системы <https://youtu.be/X8j65kvCEbk>

Спасибо за внимание!

Присоединяйтесь к нам



Telegra
m



Facebook



Instagram

academy@idomarketing.io
Юнусалиева 173/2
0(990)037-037