

*Егоров Александр Николаевич
студент 2 курса аспирантуры, электро-энергетический факультет
Вологодский государственный университет,
Россия, г. Вологда
e-mail: aleks.egorow2010@yandex.ru*

*Научный руководитель: Суконищikov А. А., канд. тех. наук, доцент,
Вологодский государственный университет,
Россия, г. Вологда*

МЕТОД ПОЛУЧЕНИЯ ПОИСКОВЫХ ЗАПРОСОВ С ПОМОЩЬЮ АГЕНТА ПОИСКА

***Аннотация:** В статье рассматривается метод получения поисковых запросов из поисковой системы Яндекс с помощью API. Приводится алгоритм работы агента поиска с подробным описанием его функций и методов, необходимых для корректной работы.*

***Ключевые слова:** директ, агент поиска, методы API, Яндекс директ, Интеллектуальная система.*

*Egorov Alexander Nikolaevich
2nd year postgraduate student, faculty of electrical energy
Vologda state University,
Russia, Vologda*

*Scientific supervisor: A. A. Clothiers., candidate of technical sciences,
associate professor
Vologda state University,
Russia, Vologda*

A METHOD OF OBTAINING SEARCH QUERIES WITH THE SEARCH AGENT

***Abstract:** The article discusses the method of obtaining search queries from the Yandex search engine using the API. The algorithm of the search agent with a detailed description of its functions and methods required for correct operation is given.*

***Keywords:** direct, search agent, API methods, Yandex direct, Intelligent system.*

В статье «Интеллектуальная система для создания контекстной информации на основе многоагентной системы» я описал общий алгоритм

работы интеллектуальной системы и отобразил общую схему ее работы с описанием каждого агента.

В данной статье более подробно опишу принцип работы агента поиска и его реализацию по сбору запросов. Задачей этого агента является получить все возможные запросы из поисковой системы и записать их в свою базу данных, чтобы в дальнейшем отслеживать все изменения и пополнять ее.

Для того, чтобы агент смог выполнить данную задачу ему требуется три массива данных:

1. Массив фраз для поиска. Это фразы, которые набирают потенциальные покупатели в поисковой системе, чтобы найти информацию о нашем товаре или услуге. Например, «Купить игровой компьютер с доставкой по всей России».

2. Массив минус слов или фраз. Данный массив содержит слова или фразы, запросы с которыми мы хотим исключить из нашей выборки. Например, «Возможно ли купить игровой компьютер с доставкой по всей России». В этом примере минус фразой будет «возможно ли», потому что она делает поисковую фразу информационной, а мы хотим исключить такого рода запросы.

3. Регионы, по которым требуется осуществить поиск. Здесь можем выбрать один или несколько регионов, таких как: Санкт-Петербург, Москва, или по регионам внутри областей: Волховский район, Сосновый Бор, или по другим странам: Болгария, Австрия. Зависит от нашей задачи

После того, как мы составим три массива нам надо воспользоваться API Яндекс Директ, чтобы осуществить сбор поисковых запросов из реальной базы данных поисковой системы.

Напомню, что Яндекс.Директ — это единая платформа для размещения контекстной и медийной рекламы. Директ покажет объявления целевой аудитории, заинтересованной в ваших товарах или услугах. API — это конструктор, в котором есть набор деталей (функций, методов) и правил их использования. При работе через API приложение отправляет запрос к сервису и получает ответ, содержащий запрошенные данные. Если рассматривать API

Директа в целом, то это мощный инструмент для управления рекламными кампаниями. С его помощью можно автоматизировать рутинные операции, использовать собственные алгоритмы управления ставками, выгружать статистику и решать множество других задач. В рамках данной статьи мы коснемся лишь маленькой части, связанной с работой Яндекс Wordstat.

На момент написания статьи существует две актуальные версии API: 4 версия и 5 версия. Большим функционалом и гибкостью обладает 5 версия, но она не умеет работать с сервисом Wordstat, а 4 версия умеет, поэтому мы будем использовать именно ее.

Из особенностей API Яндекс Директ можно выделить:

- бесплатное использование;
- подробную документацию с описанием и примерами для каждой функции;
- авторизацию с использованием протокола OAuth 2.0 (это значит, что в запросах необходимо указывать авторизационный токен)
- запросы выполняются по протоколу https методом post;
- входные и выходные структуры данных передаются в теле запроса и ответа;
- два формата взаимодействия: JSON и SOAP/XML.

Перед тем, как начать работу с API необходимо пройти процедуру регистрации приложения на сервисе Яндекс.OAuth. После прохождения модерации выдадут отладочный OAuth-токен. Разработку и отладку будем производить на тестовых данных в специальной «Песочнице», так как, если это сделать на реальной кампании в Яндекс.Директ, то может привести к ее некорректной работе.

Все подготовительные процедуры мы прошли, приступим к разработке приложения, которая будет производиться на языке программирования PHP версии 7.2 с использованием формата взаимодействия JSON.

Напомню, что в рамках данной статьи я опишу работу агента поиска по сбору запросов. Для того, чтобы получить необходимый нам отчет о статистике поисковых запросов в API Яндекс Директ предусмотрен метод «GetWordstatReport», но перед тем как получить отчет, нам нужно его создать и дождаться его формирования, для выполнения этих операций существует два метода: «CreateNewWordstatReport» и «GetWordstatReportList».

Алгоритм работы имеет следующий вид:

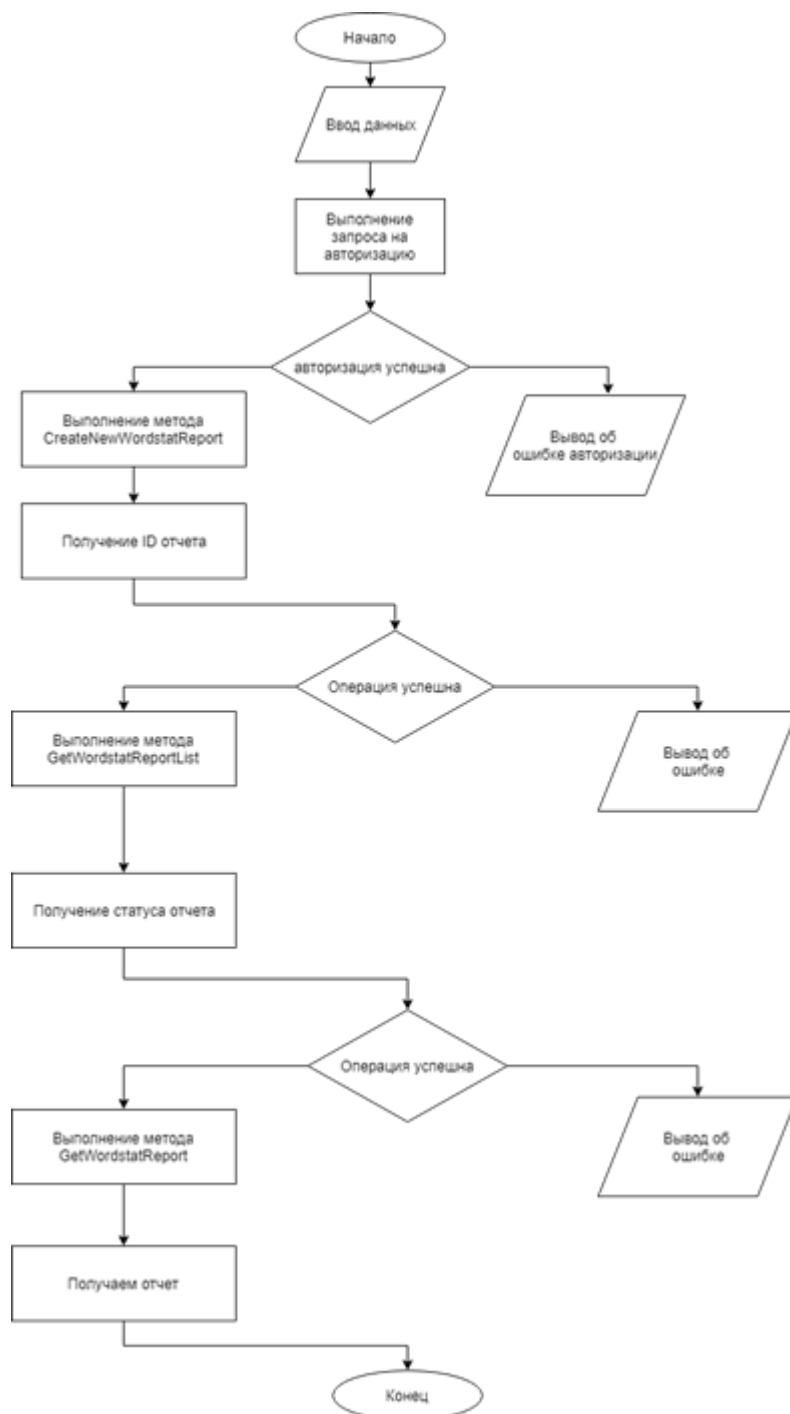


Рисунок 1. Алгоритм работы

Сначала мы вводим данные для авторизации (логин от почты Яндекс, OAuth-токен, и адрес для отправки запросов в Песочницу), три массива данных (фразы для поиска, минус слова или фразы, регионы, по которым требуется осуществить поиск). Далее отправляем запрос на авторизацию в кодировке UTF-8 и, при успешном выполнении операции, формируем запрос на создание отчета. Для этого воспользуемся методом «CreateNewWordstatReport», которой, для создания отчета, требуется следующие параметры: массив ключевых фраз с минус-фразами, массив идентификаторов регионов. Метод вернет идентификатор будущего отчета.

Дальше нам необходимо дождаться окончания формирования отчета. Для этого будем проверять статус отчета раз в десять минут. Окончанием формирования отчета будет свидетельствовать статус Done.

Сам отчет мы получаем с помощью вызова метода «GetWordstatReport». Из обязательных параметров требуется идентификатор отчета, который мы получили при использовании метода «CreateNewWordstatReport».

В результате у нас будет массив реальных запросов с их количеством в течение 28 последних дней. В дальнейшем планируется настроить сохранение этих данных в базу данных и связать с модулем генерации текста по данным запросам.

Список литературы:

1. GS.SEO-AUDITOR.COM.RU - Глобальная статистика Интернета // Глобальная статистика Интернета, URL: <http://gs.seo-auditor.com.ru/sep/2017/> (дата обращения: 11.11.2018).
2. YANDEX.RU - О Яндекс.Директе // О Яндекс.Директе, URL: <https://yandex.ru/support/direct/> (дата обращения: 26.10.2017).

3. TOCHNO-TOCHNO.RU – Агентство интернет рекламы // Агентство интернет рекламы, URL: <https://tochno-tochno.ru/info/press/ctr-v-poiske-yandex/> (дата обращения: 21.10.2018).

4. NEWS.NETCRAFT.COM – Internet Security and Data Mining // Internet Security and Data Mining, URL: https://news.netcraft.com/archives/2009/02/01/one_million_ssl_sites_on_the_web.html (дата обращения: 09.11.2018).

5. YANDEX.RU - Правила и требования к рекламе // Правила и требования к рекламе, URL: <https://yandex.ru/support/direct/moderation/adv-rules.html> (дата обращения: 03.10.2018).