ТЗ системы сбора информации о состоянии сайтов

# Введение

Система позволяет отслеживать состояние и характеристики сайтов. Их технические показатели, запросы по статистике, запросы и позиции на продвижении, состояние поисковых систем, массы внешних ссылок с урлами и их индексацией.

# Общие принципы представления и хранения информации

Система позволяет сохранять данные в БД. Данные разделяются на 2 типа первичные и системные. Первичные – это данные, которые требуют дальнейшей обработки.

Вторичные – данные которые могут использоваться для обзора состояния интернета, сайтов, внешних ссылок на него.

Интерфейс представлен набором сервиса, который принимает параметры или отдает результаты по № заказа, который выдается после запуска.

Отдельно пишется представление и формы работы с сервисом. Есть статистика обработки и состояния обработки заказов. Все заказы отправляются в очередь и потом из нее заказ забирает на обработку сервис по определенным правилам.

# Вторичные данные

Раздел содержит перечисление всех видов обработанной информации, которая может быть собрана из разных источников и использованная в работе менеджером.

## Состояние апдейтов поисковых систем

Надо найти сайт где они публикуются и собирать просто —

<http://www.seolib.ru/script/updatestat/>

тут есть информация по апам

Параметры функции:

1. Работает раз в сутки
2. Собирает поля
   1. Поисковая система
   2. Дата анализа
   3. Предполагаема дата апа
   4. Дата апа
   5. Шторм (изменение к предыдущему апу)

## Проверка и сохранение позиций сайта по запросам

Есть запросы надо видеть состояние по ним в перспективе или графиках. Использовать сеолиб

Логин mastersitov

Пароль QAZwsxEDC

Сайт [http://www.seolib.ru](http://www.seolib.ru/)

<http://www.seolib.ru/analytics/projects/actives/positions/detail/40225/12/>

тут есть информация по проведенным проверкам, ее надо будет собирать

Параметры функции:

1. Работает раз в сутки
2. Собирает поля:
   1. Номер запроса
   2. Дата проверки
   3. Проект
   4. Урл в рейтинге
   5. Позиция
   6. Поисковая система

## Получение числа бэклинков на сайт

Определение по yahoo число беков на сайт. Делается раз в сутки по каждому проекту.

<http://webmaster.yandex.ru/>

<http://poisk-ssilok.ru/>

## Ссылки на запрос

На период с параметрами кол-во тиц PR индексация, когда закуплена, вид ссылки. Создается функция под общий сервис, которая

1. Запускается
2. Авторизуется в сервисе
3. Вызывает функцию сбора
4. Фиксирует результат вызова и исключение

Функция по сбору каждая отдельно собирает поля:

1. Запрос для определения ссылки на запрос по проекту
2. Дата сбора
3. Урл ссылки
4. Вид ссылки
5. Ссылка на сервис где куплена
6. Штрафные баллы (по умолчанию 0)

<http://www.sape.ru/> сервис ссылок, через авторизацию

<http://gogetlinks.net/> сервис ссылок, через авторизацию

<http://www.miralinks.ru/> сервис ссылок, через авторизацию

## Параметры УРЛ для ссылок

Функция определяет параметры для ссылок из собранных. Вызов раз в АП

Поля:

1. ТИЦ
2. PR
3. Индекс Я
4. Индекс G
5. Кеш G
6. Кеш Я

Определение тиц и ПР <http://www.seolib.ru/script/ticpr/>

<http://www.yandex.ru/> индекс кеш, заголовок мета, титл

## Релевантные страницы по запросу

Для разных поисковых систем. Определение их параметров. Поля:

1. Ссылка на запрос
2. Ссылка на проект
3. Урл релевантной страницы
4. Титл
5. H1
6. Keywords
7. Desc
8. Сниппет поисковой системы
9. Поисковая система

Поиск по ограничению для сайта в поисковой системе. Потом запрос страницы и получение информации из ее содержимого.

## Оценка продвижения по запросам в директ

Запросы задаешь , профиль региона и др и получаешь оценку

Поля :

1. Запросы через разделитель
2. Стоимость
3. Профиль гео

mastersitov@yandex.ru

QAZwsxEDC

<http://direct.yandex.ru/registered/main.pl?authredirlevel=1342438316.0&cmd=ForecastByWords&ncrnd=724948>

## Сбор частотности по запросам продвижения

Хранение их в БД помесячно. Запросы из проекта или запросы через разделитель получаешь частоту по запросу в сервисе. Сервис Я, G, R. Поля:

1. Запрос
2. Частотность
3. Сервис
4. Дата

<http://wordstat.yandex.ru/> по Я

<http://adstat.rambler.ru/wrds/> по R

<https://adwords.google.com/o/Targeting/Explorer?__c=1000000000&__u=1000000000&ideaRequestType=KEYWORD_IDEAS> по G

## Сбор стоимости по сеопульт

Запросы и урл сайта на входе на выходе запросы и стоимость

http://seopult.ru/item.html

mastersitov

QAZwsxEDC

<http://seopult.ru/item.html?item_id=3797550&step=3>

# Справочники

Система позволяет редактировать данные справочники менеджеру

## Справочник проекта

Поля:

1. Название проекта
2. Дата начала работ
3. Адрес сайта
4. Ссылка на запросы от клиента
5. Ссылка на запросы дополнительные (добавленные для проекта но не оплаченные)
6. УРЛ на проект в директ
7. УРЛ на проект в сеолиб
8. УРЛ на проект в сеопульт
9. УРЛ на проект в сайп

## Справочник запросов продвижения

Поля:

1. Ссылка на проект
2. ИД запроса
3. Запрос
4. Тип запроса (заказчик, менеджер)
5. Дата начала работы с ним

## Справочник интерграционных систем и паремтров доступа к ним

Директ, сеолиб, сеопульт, сейп, миралинк

Поля:

1. Название
2. УРЛ
3. Логин
4. пароль

## Справочник ЛОГ запуска сбора данных

Поля:

1. название запуска
2. дата время старта
3. дата время финиша
4. параметры данных
5. код результата
6. исключение сбоя

НЕ ВАЖНО

## Сбор атрибутов страниц по ключевикам для первой выдачи поисковика (титл, метатеги, h1)

Анализ выдачи поисковика, и сбор информации по выдаче

## Сбор поискового ядра по выдаче поисковика

Есть выдача, пишешь запросы, потом получаешь все запросы из страниц выдачи по сайтам что попали в выдачу

## Новостной поток лент по продвижению с разных источников

Определить список сайтов и завести новости с них в БД или ленту выдачи с указанием источника

## Параметры сайтов на продвижении

Показатели сайта, тиц, pr, наличие в поиске, в каталоге Я, возраст

## Фиксация ссылок и сервисов для сайтов на продвижении.

Оценка параметров закупленных страниц.

анализ ссылочной массы, прирост, убыль по проектам, стоимостные характеристики

## Формирование списка хороших сайтов-доноров

Возможно список сайтов на которых не надо ничего покупать.

со темой которая там закуплена и типом которая закуплена