

Тема: “Разработка программных средств анализа информационного контента для формирования запросов к поисковым системам”



Данилычев Кирилл Павлович
группа ИЭС 166-14

Научный руководитель

к.т.н., доцент

Артюхов О. И.

Цель работы:

Целью данной выпускной квалификационной работы является разработка системы, позволяющей автоматизировать деятельность SEO-специалистов.

Выбор автоматизируемых задач

1. Составление семантического ядра
2. Оптимизация контента сайта

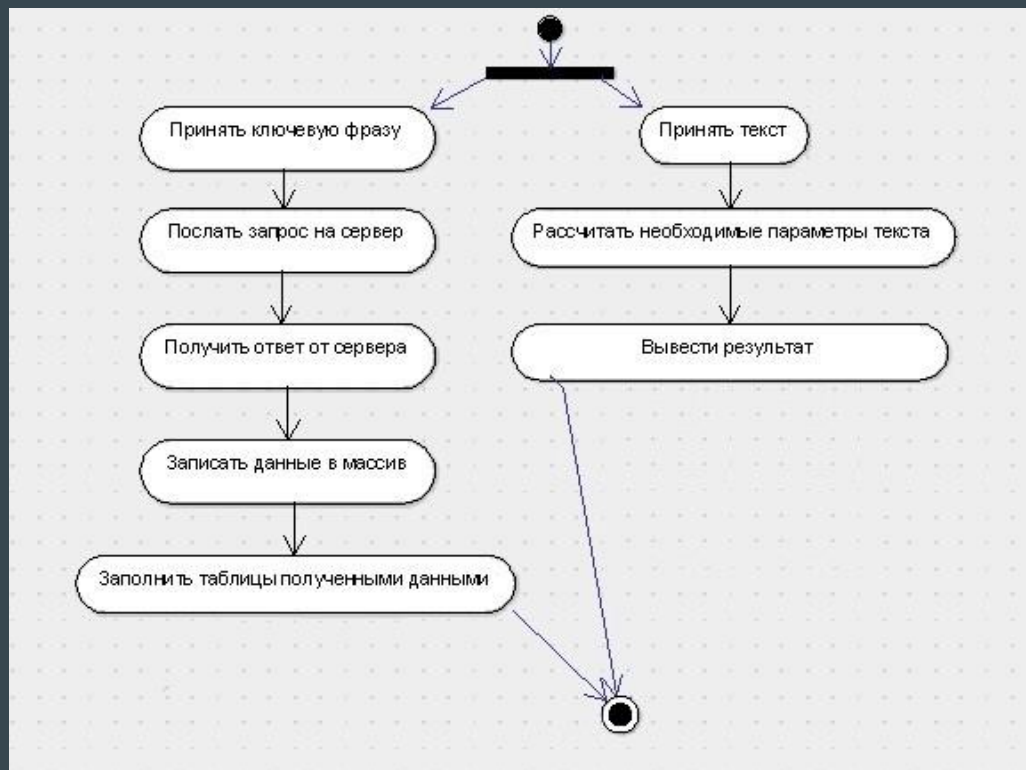
Работы по продвижению, использующие семантическое ядро



Оптимизация контента сайта

Оптимизация контента сайта также является одной из важнейших задач в деятельности SEO-специалистов. Для успешного продвижения сайта необходимо, чтобы контент на нем был **релевантным**. Кроме того нужно придерживаться оптимальной **плотности ключевых слов**. Чтобы соблюдать данные требования, необходимо регулярно проводить анализ текстового содержимого сайта.

Диаграмма деятельности системы



Стадии разработки:

1. Анализ предметной области и постановка задачи разработки системы
2. Проектирование системы
3. Реализация системы

Среда и средство разработки

В качестве среды разработки была выбрана бесплатная IDE Microsoft Visual Studio Express 2015. Языком разработки мною был выбран C#.

Интерфейс формы анализа слов

Данилычев К.П. 166-14

Анализ слов Анализ текста

Параметры

Ключевая фраза: Поисковая система: Яндекс

Стоп-слова

Результаты

По ключевой фразе		В пределах одной сессии		Синонимы	
Слово	Запросов	Слово	Запросов	Слово	
▶ мзи	80495	▶ московский эко...	7057	▶ институт официальный сайт	
сайт мзи	11148	московский эне...	2939	институт москва	
мзи официальн...	10875	бауманский уни...	4875	московский университет	
мзи официальн...	10465	московский инс...	102180	институт московский	
ниу мзи	4274	ричард хаммонд	3550	вуз москва	
мзи москва	3428	джеремии кларк...	8687	институт москва официальны	

Группа элементов “Параметры”

Параметры

Ключевая фраза:

Поисковая система: Яндекс

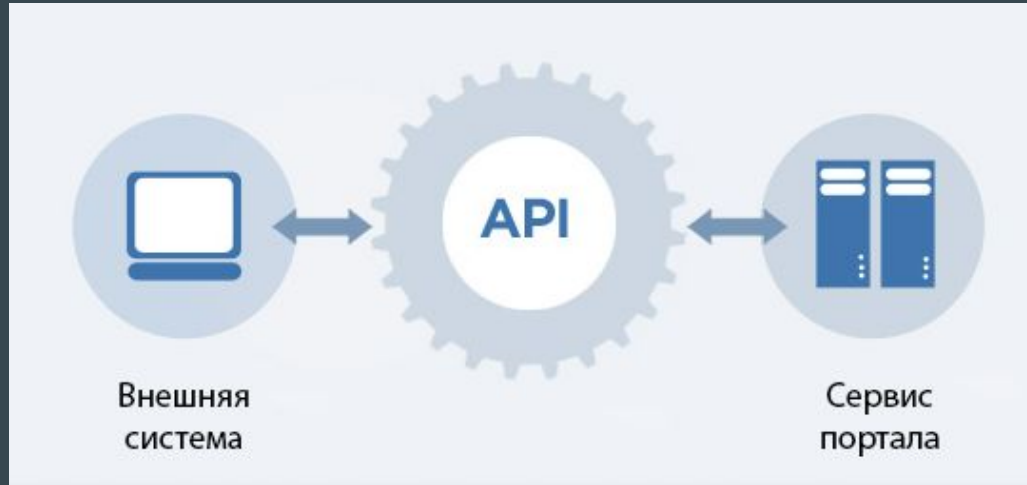
Стоп-слова

Группа элементов “Результаты”

Результаты						
По ключевой фразе		В пределах одной сессии		Синонимы		
	Слово	Запросов		Слово	Запросов	
*			*			
[Grayed out content]			[Grayed out content]		[Grayed out content]	

При помощи модуля API приложение может устанавливать связь с Яндекс сервером и получать от него необходимые данные, а именно: статистику по запросу, введенному пользователем и запросам, похожим на него, количество показов в месяц, список запросов, которые являются синонимами введенного запроса пользователем.

Работа по парсингу ключевых слов будет происходить с помощью API YANDEX



Вторая часть реализует анализатор текста. В нее включается модуль расчета параметров и определения семантического ядра текста. К данному модулю подключена внешняя библиотека, в которой реализован алгоритм выделения словарных основ (стеммер)

Интерфейса формы анализа текста

Данилычев К.П. 166-14

Анализ слов | Анализ текста

Параметры

Московский энергетический институт сегодня - один из крупнейших технических университетов России в области энергетики, электротехники, электроники, информатики.

МЭИ готовит инженерные и научные кадры для иностранных государств начиная с 1946 года.

В настоящее время в МЭИ обучаются студенты и аспиранты из 68 стран мира.

За успехи в подготовке инженеров и научных кадров награжден двумя государственными орденами и шестью орденами зарубежных государств.

Начать | Сохранить

Результаты

Статистика текста

СТАТИСТИКА ТЕКСТА
Количество символов 662
Количество символов без пробелов 589
Количество слов 75
Количество уникальных слов 55
Количество стоп-слов 5
Классическая тошнота документа 2,24
Академическая тошнота документа 18,67%

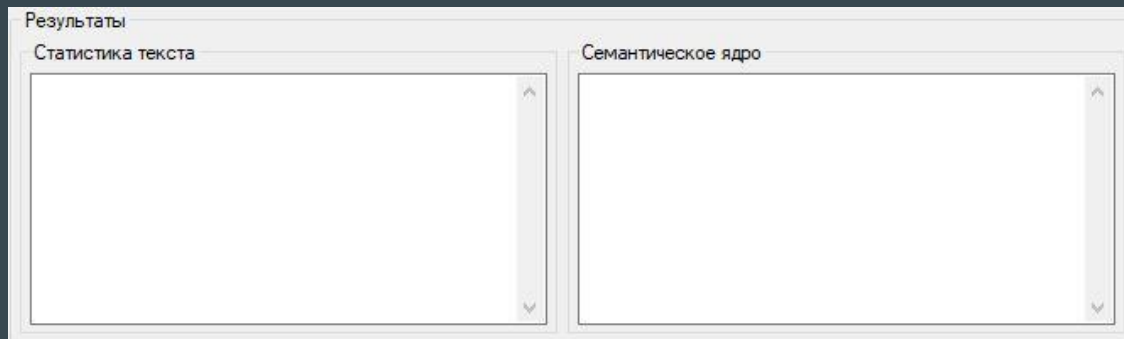
Семантическое ядро

СЕМАНТИЧЕСКОЕ ЯДРО
Международной 3 4,00
ассоциации 3 4,00
университетов 2 2,67
МЭИ 2 2,67
государств 2 2,67
орденами 2 2,67
СПОВА

Группа элементов “Параметры”

Параметры

Группа элементов “Результаты”



Разработанная система автоматизации деятельности SEO-специалистов позволяет:

1. Проводить в поисковой системе Яндекс сбор статистики запросов
2. Проводить сбор статистики по запросам
3. Выводить список синонимов к заданному ключевому слову
4. Проводить анализ текста

Все поставленные задачи были выполнены. Цель работы была достигнута.

Спасибо за внимание.

Тема: “Разработка программных средств анализа информационного контента для формирования запросов к поисковым системам”



Данилычев Кирилл Павлович
группа ИЭС 166-14

Научный руководитель

к.т.н., доцент

Артюхов О. И.