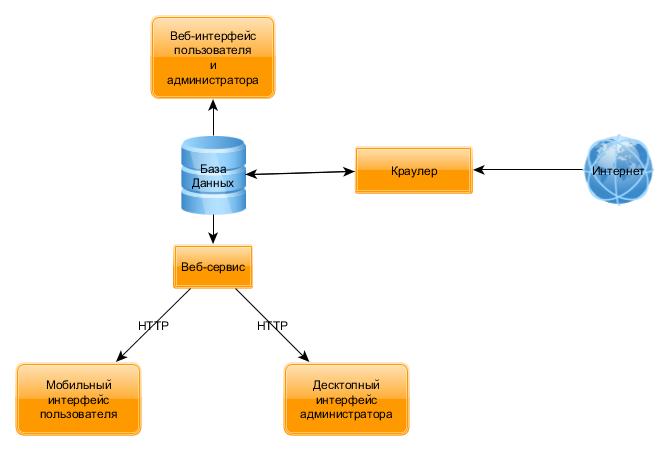
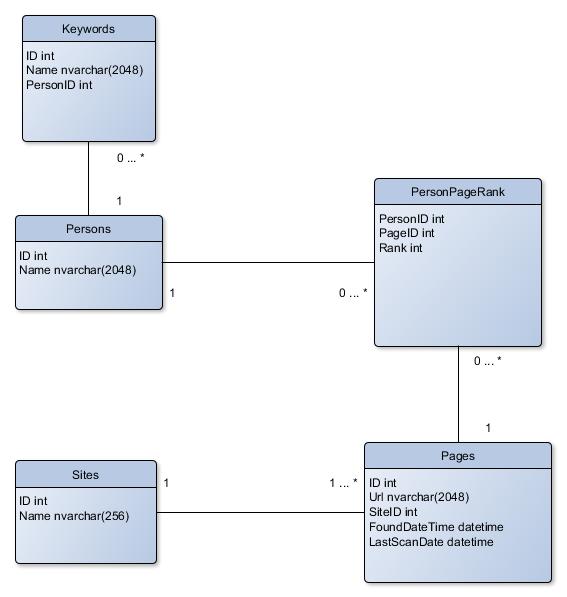
# Компоненты



1. **Веб-интерфейс пользователя и администратор** – веб-сайт, являющийся интерфейсом для работы пользователя и администратора. Пользователь имеет возможность просматривать общую и ежедневную статистику.
2. **Краулер –** это служба, которая по определенному расписанию, или принудительно, обращается к ресурсам в интернете. Анализирует их на наличие новых страниц. Загружает новые страницы, после чего обрабатывает контент новой страницы. Результаты обработки и ссылку на новую страницу записывает в Базу Данных.
3. **Веб-сервис –** транслирует запросы от Мобильного и Десктопного интерфейсов пользователя, в SQL запросы к базе данных. Результат, полученный от запроса, передает в XML формате, обратно Мобильному и Десктопному интерфейсу.
4. **Мобильный интерфейс пользователя –** приложение, разработанное для мобильной платформы Android, для работы пользователя. Пользователь получает возможность просматривать общую и ежедневную статистику.
5. **Десктопный интерфейс администратор –** приложение разработанное для деcктопной версии Windows. Администратор получает возможность настраивать работу системы, заполнять справочники, а также просматривать общую и ежедневную статистику.

# Модель данных



1. **Persons (личности)–** таблица базы данных, отвечающая за хранение имен личностей. Каждой личности соответствует от 0 до бесконечности ключевых слов.

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблица Persons** | | | |
| Имя столбца | Описание | Тип | Пример |
| ID | Идентификатор личности | Int | 1 |
| Name | Наименование личности | nvarchar(2048) | «Путин» |

1. **Keywords (ключевые слова)–** таблица базы данных, отвечающая за хранение ключевых слов, соответствующих каждой конкретной личности. Встретив какое-либо из ключевых слов на HTML-странице веб-сайта, мы считаем, что встретили упоминание соответствующей ему личности на данной HTML-странице. Каждому ключевому слову соответствует ровно одна личность.

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблицы Keywords** | | | |
| Имя столбца | Описание | Тип | Пример |
| ID | Идентификатор ключевого слова | Int | 1 |
| Name | Название ключевого слова | nvarchar(2048) | «Путиным» |
| PersonID | Идентификатор личности, которой соответствует данное ключевое слово. Является внешним ключом к таблице Persons. | int | 1 |

1. **PersonPageRank (рейтинг личности на странице сайта) –** таблица базы данных, отвечающая за хранение количества упоминаний каждой персоны, на обработанной странице.

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблицы PersonPageRank** | | | |
| Имя столбца | Описание | Тип | Пример |
| PersonID | Идентификатор личности, которой соответствует данное ключевое слово. Является внешним ключом к таблице Persons. | Int | 1 |
| PageID | Идентификатор страницы сайта, на которой найдены упоминания о персонах. Является внешним ключом к таблице Pages | int | 1 |
| Rank | Количество упоминаний персоны на странице | int | 2 |

1. **Pages (страница сайта) –** таблица базы данных, содержит страницы сайта, которые были найдены при анализе сайтов из таблицы Sites.

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблицы Pages** | | | |
| Имя столбца | Описание | Тип | Пример |
| ID | Идентификатор страницы сайта | Int | 1 |
| Url | Полный URL адрес страницы | nvarchar(2048) | «https://news.mail.ru/politics/» |
| SiteID | Идентификатор сайта (ресурса), который предоставлен администратором для анализа. Является внешним ключом к таблице Sites | int | 1 |
| FoundDateTime | Дата и время обнаружения страницы системой | datetime | “10-10-2015 15:23” |
| LastScanDate | Дата и время последней проверки на упоминания | datetime | “03-04-2015 15:23” |

1. **Sites (сайт) –** таблица базы данных, содержит названия сайтов для анализа на упоминания.

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблица Sites** | | | |
| Имя столбца | Описание | Тип | Пример |
| ID | Идентификатор сайта | Int | 1 |
| Name | Наименование сайта | nvarchar(2048) | «Новости Политики - Новости Mail.Ru» |
| Url | Полный URL адрес страницы | nvarchar(2048) | «https://news.mail.ru/politics/» |