# 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 |
| ID | | Идентификатор личности | Int |
| Name | | Наименование личности | nvarchar(2048) |

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 |
| ID | | Идентификатор ключевого слова | Int |
| Name | | Название ключевого слова | nvarchar(2048) |
| PersonID | | Идентификатор личности, которой соответствует данное ключевое слово. Является внешним ключом к таблице Persons. | int |

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

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

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 |
| ID | | Идентификатор страницы сайта | Int |
| Url | | Полный URL адрес страницы | nvarchar(2048) |
| SiteID | | Идентификатор сайта (ресурса), который предоставлен администратором для анализа. Является внешним ключом к таблице Sites | int |
| FoundDateTime | | Дата и время обнаружения страницы системой | datetime |
| LastScanDate | | Дата и время последней проверки на упоминания | datetime |

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

|  |
| --- |
| **Таблица Sites** |
| Имя столбца | | Описание | Тип |
| ID | | Идентификатор сайта | Int |
| Name | | Наименование сайта | nvarchar(2048) |