



i2 TextChart

i2 TextChart –это расширение для Analyst's Notebook, позволяющее аналитикам импортировать структурированные данные из больших объемов неструктурированных текстовых документов. Программное обеспечение может интерпретировать контент, написанный более сотней языков, и преобразовывать его в сущности, ссылки и свойства, знакомые пользователям Analyst's Notebook.

TextChart поставляется как готовый к использованию пакет. Он содержит стандартные правила извлечения терминов и, если вы используете встроенный вариант хранилища терминов, требует очень мало настроек.

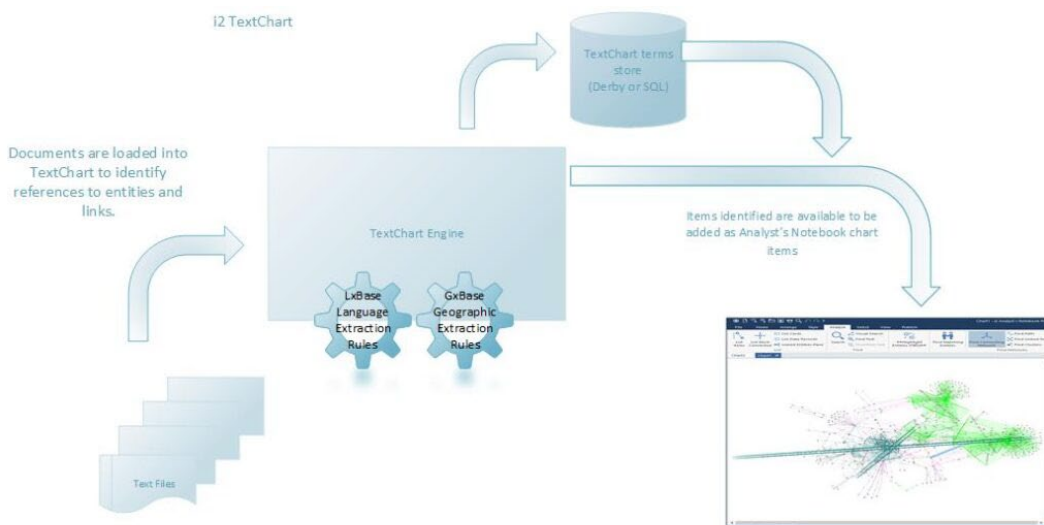
TextChartPremium значительно увеличивает масштаб вашего развертывания:

- Документы из разных источников могут обрабатываться с использованием разных правил получения.
- Сроки могут быть загружены одновременно из нескольких источников.
- Полученные сроки хранятся в базе данных, которую аналитики могут запросить как внешний источник в Analyst's Notebook Premium.

Ключевые компоненты i2 TextChart:

- Расположение файлов: пути файловой системы, содержащие документы для обработки
- Механизм извлечения: где документы обрабатываются для получения содержащихся в них сроков
- LxBase: определение лексикона, представляющего собой список правил, определяющих, как термины идентифицируются в исходном материале
- GxBase: географическое определение, являющее собой перечень правил, определяющих, как географические места идентифицируются в исходном материале

В стандартном развертывании TextChart можно также настроить репозиторий для полученных сроков и внешние картографические программы для работы с географическими данными.



The diagram illustrates the i2 TextChart Enterprise architecture. It shows two clusters of 'Text' and 'Text Files' being processed by 'TextChart worker' nodes. These workers use 'LxBase Language Extraction Rules' and 'GxBase Geographic Extraction Rules'. The extracted data is sent to a 'TextChart Master'. The master then loads documents to identify references to entities and links, storing 'TextChart SQL Extracted terms' in a database. These terms are then available to be added as Analyst's Notebook chart items via the 'i2 Connect Gateway'. The final output is a visualization of the extracted terms and links, showing a network of entities and relationships.