

программа
TEXT MINING

План презентации:

1. назначение программы;
2. скачивание программы;
3. работа с программой.

НАЗНАЧЕНИЕ TEXT MINING

Text Mining Tool - это бесплатная утилита, которая необходима для анализа pdf, doc, rtf, chm, html форматов и извлечения текста из них без использования других программ (Word, Acrobat и т.п).

Различие между форматами текста:

1. **pdf** - формат электронных документов, созданный фирмой Adobe Systems, предназначен для представления в электронном виде полиграфической продукции;
2. **doc** – формат используется для обозначения простых текстовых файлов без форматирования, однако позже стал использоваться для двоичных форматов с разметкой;
3. **rtf** - формат хранения размеченных текстовых документов, предложенный группами программистов, основавшими компании Microsoft и Adobe;
4. **html** - стандартный язык разметки документов во Всемирной паутине, используемый на многих веб-страницах.

КАК СКАЧАТЬ ПРОГРАММУ?

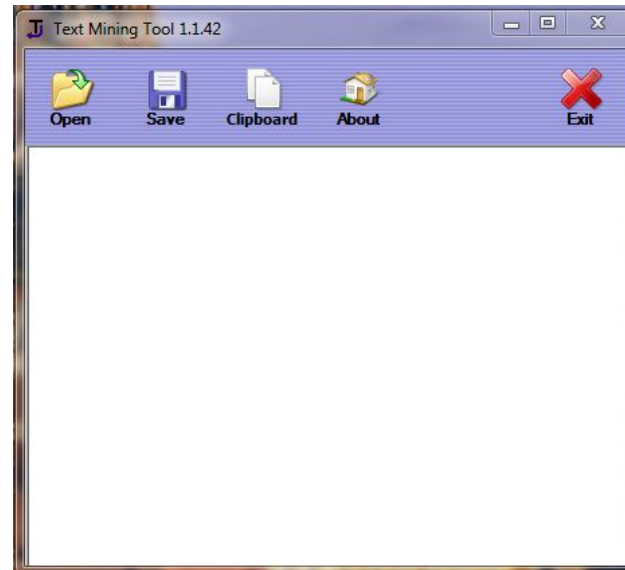
1). ссылка для скачивания программы -

http://www.freeware.ru/program_prog_id_12953.html

2). На сайте будет такая ссылка: Домашняя страница Скачать программу с Ware.ru

3). Сохранить архив.

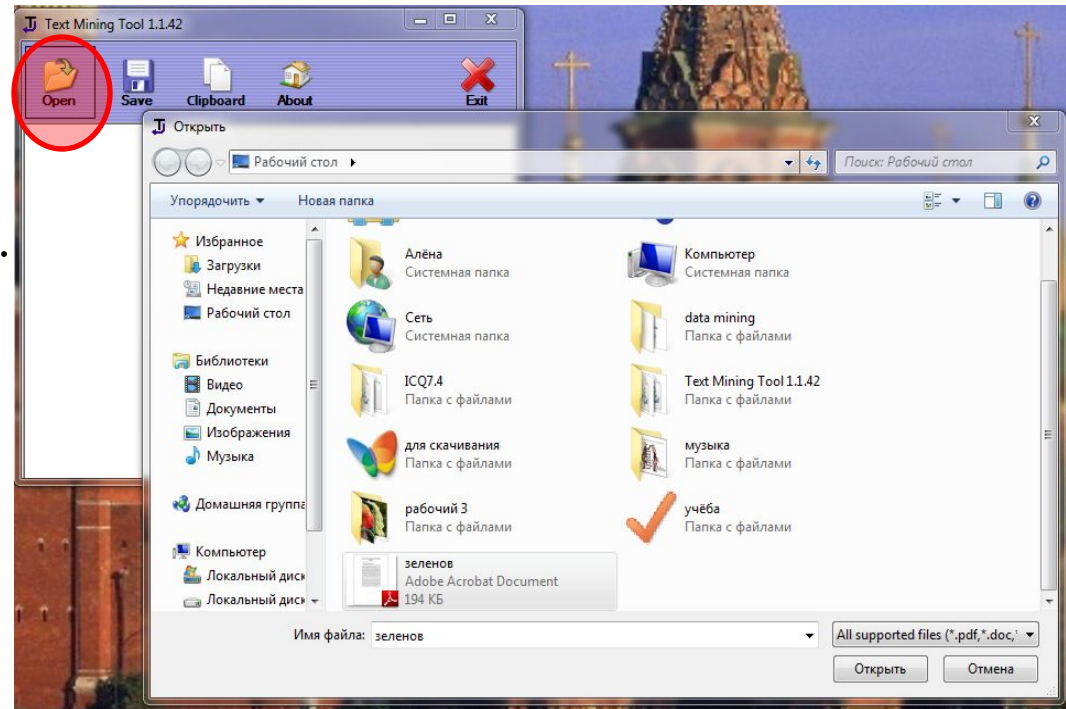
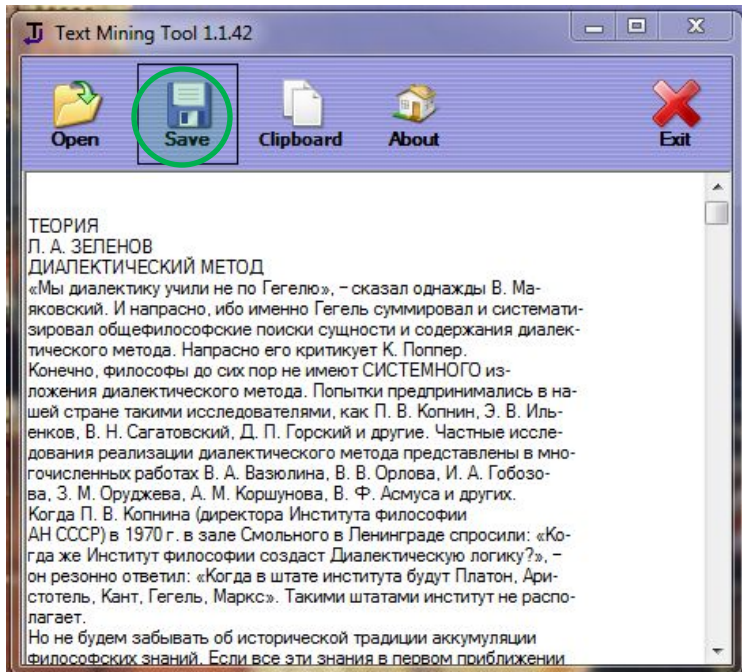
4). Скриншот программы



КАК РАБОТАТЬ С TEXT MINING?

1). Для того, чтобы загрузить файл в программу, необходимо нажать на **OPEN** и выбрать нужный файл.

2). После загрузки файла нажать на **SAVE** и сохранить файл.



3). Файл сохранится в формате txt.