## [Обработка текстовой информации.](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage2.html)

 Особенности обработки текстовой информации заключаются в том, что при вводе и редактировании формируется содержание текстового документа. Оформление документа задают операциями форматирования. Команды форматирования позволяют оформить документ в том виде, в котором он будет выглядеть на экране монитора или на бумаге после печати на принтере. Наименьшим форматируемым элементом документа является символ, затем – абзац, страница и раздел.

Создание и оформление документов основано на использовании так называемых шаблонов документов и стилей оформления.

***Стиль*** – это совокупность всех параметров оформления, определяющих вид символа, абзаца, страницы или раздела. Как правило, стили хранятся в самом документе. Вместе с тем, их удобно хранить в компактном виде в отдельном файле, называемом шаблоном.

***Шаблон*** содержит стили и текстовые заготовки, используемые в документах определенного типа.

Операции, используемые при подготовке текстовых документов

Основной набор типовых операций, осуществляемых с помощью современных программных средств обработки текстовых документов, включает операции, производимые над документом в целом, над абзацами документа и над его фрагментами.

К операциям, производимым над документом, относятся:

* *создание нового документа (присвоение документу уникального имени и ввод текста документа с помощью клавиатуры);*
* *загрузка имеющегося на внешнем носителе документа в оперативную память;*
* *сохранение документа (копирование документа из оперативной памяти во внешнюю);*
* *распечатка документа (создание бумажной копии документа).*

Отметим, что операция удаления документа в программных средствах обработки текстовых документов отсутствует, в силу того, что редактор работает с копией документа, находящейся в оперативной памяти, а сам документ может находиться либо во внешней оперативной памяти, так и на другом компьютере.

Вся работа в среде редактора осуществляется с копией этого документа, загруженной в оперативную память с внешнего носителя, поэтому все изменения в ней необходимо периодически сохранять.

***Операции, производимые над абзацем*** изменяют его параметры как структурного элемента текста. Параметры абзаца задаются с помощью опций:

* *выравнивания строк абзаца (влево, вправо, по центру, по ширине);*
* *задания величины отступа в красной строке абзаца;*
* *задания ширины и положения абзаца на странице;*
* *задания межстрочного расстояния внутри абзаца и др.*

***Операции, производимые с фрагментом текста*** включают его выделение, перемещение, копирование или удаление. Фрагментами текста могут быть отдельный символ, слово, группа слов, абзац и т. д. Поэтому возможно удаление этих объектов в отличие от удаления документа в целом. Выделенный фрагмент текста можно напечатать, произвести контекстный поиск и замену символов в нем, применить шрифтовое выделение и ряд других операций.

Далее перечислим другие наиболее часто используемые операции при обработке документа в целом или выделенных фрагментов.

***Контекстный поиск и замена***. Предварительно разметив текст, требующий многократного обращения к определенным местам документа можно использовать быстрый поиск нужных мест документа по аналогии с закладками в обычной книге (например, при подготовке статьи путем многократного редактирования документа). Можно также задать некоторый образец (символ, слово, группу слов или цепочку символов) и подать команду поиска. Поиск по образцу удобен, например, при замене термина в документе (например, имеется документ, в тексте которого встречается год - 2000, с помощью контекстного поиска и замены можно быстро изменить его на 2007 год).

***Режим автоматического переноса*** слов используют для улучшения внешнего вида текста. При выключенном режиме автоматического переноса слово, не поместившееся на строке, полностью переносится на следующую строку, но в этом случае края текста остаются неровным. При включенном режиме автоматического переноса реализуется мягкий вариант переноса: слово автоматически переносится в соответствии с правилами переноса слов русского или другого языка.

***Проверка правописания слов и синтаксиса*** выполняется специальными программами, которые могут быть автономными, например Орфо, или встроенными в текстовый процессор. Эти программы значительно различаются по своим возможностям. Наиболее мощные из них проверяют не только правописание, но и склонение, спряжение, пунктуацию и даже стиль. Все указанные ошибки выявляются на основе сравнения, разработанного текста с хранящимися в памяти основными правилами. Заметим, что используемый в рамках текстового процессора набор таких правил существенно ограничен. Словарь синонимов поможет избежать повторений и сделает элегантным ваш стиль изложения.

***Установка общих параметров страницы*** предусматривает размер полей, размер и ориентацию бумаги, нумерацию страниц, колонтитулы. Существует также команда запрета разрыва страниц, которая используется, когда вы хотите, чтобы определенная часть документа (например, таблица) находилась на одной странице, если таблица не умещается на странице, то она переносится целиком на следующую страницу. Для введения ***нумерации страниц*** в создаваемом вами документе текстовый процессор предложит специальное меню, в котором вы сможете указать все интересующие вас условия нумерации: месторасположение на листе номера страницы, отказ от нумерации первой страницы, использование колонтитулов и другие. Номера страниц проставляются в колонтитуле. ***Колонтитулом*** называется заголовочное данное, помещаемое в начале или конце каждой страницы документа. Колонтитулы обычно содержат номера страниц, название глав и параграфов, название и адрес фирмы и т. п. Колонтитулы могут различаться для четных и нечетных страниц, а также для первой страницы и последующих.

Некоторые текстовые процессоры позволяют использовать ***макросы***. **Макросом** называют файл, в котором хранится программа часто повторяющейся последовательности действий, заданная пользователем. [**Макрос**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)имеет уникальное имя. С помощью **макросов** можно автоматизировать многие типовые технологические этапы при работе с документами, например,**макрос**, выполняющий последовательность команд по созданию стиля для каждого абзаца документа. После вызова **макроса** записанная в нем последовательность действий или команд будет в точности воспроизведена автоматически.

Операции сохранения записывают отредактированный документ или его фрагмент, находящийся в оперативной памяти, на внешний носитель для постоянного хранения. Тип сохраненного документа обычно присваивается текстовым процессором автоматически. Например, в текстовом процессоре [**Word**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)документу присваивается расширение .doc. Возможны режимы “*Сохранить и продолжить редактирование*”, “*Сохранить и выйти*”. “*Выйти без сохранения*”.

Текстовые процессоры с помощью [***резервного копирования***](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)обеспечивают защиту созданных документов от возможной утраты. Для этого специальной командой сохранения обеспечивается режим, когда одновременно хранятся два файла одного и того же документа - текущий и резервный. После внесения изменений в документ и его сохранения предыдущая его версия автоматически сохраняется как резервный файл с именем копия+имя документа и расширением [**.wbk**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html), а отредактированная версия рассматривается как текущий файл. В редакторе **Word** можно установить автоматическое создание резервной копии при каждом сохранении документа. Резервная копия представляет собой предыдущую версию документа, что позволяет сохранить как текущие сведения, так и сведения, сохраненные ранее. Сохранение резервной копии помогает сохранить данные, если случайно были сохранены изменения, которые вы не хотели сохранять, или если был удален исходный файл.

Важным фактором защиты создаваемых документов является функция [***автосохранения***](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html), которая может выполняться как обычная операция сохранения или как специальная операция сохранения текущего состояния текстового процессора в специальном файле. В последнем случае при аварийном прекращении работы это состояние может быть восстановлено, включая содержимое всех окон, положение курсоров в окнах и т.п. Это дает возможность восстановить документ в случае, если выполнение программы будет прервано по причине сбоя или при внезапном отключении электричества. Если включено автосохранение, внесенные в документ изменения сохраняются во временный файл с указанной частотой. Использование автосохранения не избавляет от необходимости сохранять открытый документ обычным способом. При сохранении файл восстановления заменяет собой исходный документ (если не было задано новое имя файла). следует заметить, что включенный режим автосохранения замедляет работу с документом, особенно, если он достаточно большой.

Существуют универсальные форматы текстовых файлов, которые могут быть прочитаны большинством текстовых редакторов и оригинальные форматы, которые используются отдельными текстовыми редакторами. Для преобразования текстового файла из одного формата в другой используются специальные программы – конверторы. В текстовых процессорах такие конверторы входят в состав системы. Перечислим некоторые форматы текстовых файлов:

только текст ( .[**txt**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)) – наиболее универсальный формат. Сохраняет текст без форматирования, в текст вставляются только управляющие символы конца абзаца.

текст в формате( .[**rtf**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)) – универсальный формат, который сохраняет все форматирование. Текст в данном формате может быть прочитан и интерпретирован многими приложениями (например, предыдущими версиями **Word**)

документ **Word**( [.**doc**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)) – оригинальный формат используемой версии Word. Полностью сохраняет форматирование, использую 16-битную кодировку символов, что требует использование шрифтов **Unicode**.

текстовый документ **Works** для **Windows**( .[**wps**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)) - оригинальный формат интегрированной системы **Works**. Преобразование в формат документа **Word** возможно при наличии конвертора, при этом форматирование **Works** сохраняется не полностью.

**HTML** документ ( [.**htm**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html), [.**html**](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html)) – формат хранения **Web**-страниц. Содержит управляющие коды (теги) языка разметки гипертекста.

Преобразование формата текстового документа можно производить в процессе его сохранения или открытия. Данное преобразование производится с помощью конверторов, входящих в состав текстовых редакторов, позволяющих импортировать и экспортировать документы из одного приложения в другое.

Например, благодаря наличию в текстовом процессоре **Word** компонент **Graph**, **Equation** и **WordArt** возможно вставлять в документ различные диаграммы, математические формулы и текстовые эффекты. Эти компоненты не всегда доступны, что зависит от вида установки **Word**(обычная или выборочная установка). Так компонента **Equation**, как правило, при обычной установке недоступна.

Текстовые процессоры позволяют создать [***оглавление документа***](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html), применяя соответствующие стили к каждому заголовку, который будет в него включен. Для облегчения ориентации в документе можно также создать [***предметный указатель***](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html), в котором приводится список тем, обсуждаемых в данном документе, вместе с номерами страниц, содержащих соответствующую информацию. Возможно добавление примечаний, закладок, а также [***перекрестных ссылок***](file:///C%3A%5CUsers%5C%D0%9F%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%5CDesktop%5C%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%20%D0%B0%D1%80%D1%85%D0%B8%D0%B2%5Cpublic_html%5Cpage37.html).

При работе с большими документами проще располагать отдельные разделы в различных файлах, а затем компоновать весь документ в режиме структуры документа. Такой подход облегчает подготовку отдельных частей документа и позволяет производить общие операции со всем большим документом (формировать общее оглавление, проводить общую нумерацию страниц и т. д.).