

Глава 5

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ СОЗДАНИЯ И ОБРАБОТКИ ИНФОРМАЦИОННЫХ ОБЪЕКТОВ

Традиционно понятие «технология» трактуется как совокупность методов и инструментов для достижения желаемого результата в некоторой области человеческой деятельности.



Информационные технологии (ИТ) — это совокупность методов, производственных процессов, программно-технических и лингвистических средств, объединённых с целью сбора, обработки, хранения, распространения, отображения и использования информации, представленной в цифровой форме.

Синонимом этого понятия в русском языке выступает понятие «информационно-коммуникационные технологии» (ИКТ). Используемые англоязычные аббревиатуры — ИТ, ИСТ.

В отличие от технологий «материальных» и исходным материалом, и результатом применения информационных технологий всегда являются данные. Информационная технология — это процедура автоматизированного преобразования данных, формирования на их основе новых данных. Целью разработки и применения всех информационных технологий является максимальная автоматизация тех информационных процессов, которые ранее требовали ручного человеческого труда, зачастую рутинного, предполагавшего значительные временные затраты.

Информационные технологии как отдельная отрасль деятельности получили наибольшее развитие с появлением и распространением компьютеров — универсальных автоматических средств обработки данных.

Существует большое количество оснований для классификации информационных технологий. Их делят на индивидуальные

и коллективные, локальные и сетевые, технологии управления данными и процессами, защиты информации, разработки программного обеспечения и т. д.

В курсе информатики основной школы информационные технологии были представлены (классифицированы) по видам обрабатываемой информации. Вы знакомы и, скорее всего, используете в учебной деятельности, технологии обработки текста и графической информации, мультимедийные технологии, электронные таблицы, базы данных. Знание базовых принципов обработки информации, владение наиболее распространенными технологиями — необходимый навык для любого современного человека.

§ 23

Текстовые документы

23.1. Виды текстовых документов

В различных словарях можно найти следующие толкования понятия «текст»:

- 1) упорядоченный набор слов, предназначенный для того, чтобы выразить некий смысл;
- 2) всякая записанная речь (литературное произведение, сочинение, документ и т. п., а также часть, отрывок из них);
- 3) последовательность языковых и иных знаков, образующая единое целое, служащее объектом изучения.

С позиции информатики, текст — это последовательность знаков некоторого алфавита.



Вам известно, что в памяти компьютера тексты представляются в двоичном коде: 1) за каждым символом алфавита закрепляется определённый двоичный код; 2) в двоичном коде представляется и информация о типе и размере используемого шрифта, положении строк, полей, отступов и прочая дополнительная информация.

Практически в любой профессиональной деятельности работник сталкивается с необходимостью подготовки текстовых документов

различного назначения и объёма: от заявления о приёме на работу до составления отчёта по результатам проделанной работы.

Можно выделить следующие виды текстовых документов:

- художественный текст;
- научный текст;
- деловой документ;
- рекламный документ;
- личный документ.

Для каждой из перечисленных разновидностей текстовых документов существует определённый набор правил, которых следует придерживаться при работе над ними. Личное письмо отличается по стилистике от официального документа, а художественное произведение — от научного текста. Различаются также словари наиболее употребляемых слов и терминов для перечисленных разновидностей документов.

Дайте характеристику каждому из видов текстовых документов — художественному тексту, научному тексту, деловому документу, рекламному документу, личному документу.

23.2. Виды программного обеспечения для обработки текстовой информации

Существует множество программных продуктов, предназначенных для работы с текстовой информацией. Представим классификацию этой разновидности прикладного программного обеспечения по его назначению.

Текстовые редакторы — это программы, которые помогают подготовить текст простой структуры, но не обладают необходимыми средствами оформления его для печати. Типичный пример — редактор Блокнот (в ОС Windows).

Текстовые процессоры — более сложные программные комплексы, позволяющие выполнить оформление текста, точно задать его расположение, включить в него графические материалы. Примеры — Microsoft Word, OpenOffice Writer.

Специальные программные средства для подготовки научных текстов, содержащих математические, химические или другие формулы, сложные схемы и специфические обозначения, используемые в научных, учебных и технических публикациях и документах. При подготовке научных, технических и учебных текстов часто используется свободно доступная система подготовки публикаций T_EX.

Издательские системы — комплексы программных средств, позволяющих выполнить весь цикл допечатной подготовки издания: импорт или набор текста, его оформление и расположение на листах, вставку иллюстраций и сложных объектов, и в итоге — вывод издания на печать. Примерами таких программ могут быть пакеты Adobe InDesign, Scribus, QuarkXPress. Процесс и результат создания страниц издания называют вёрсткой, а точную копию самого издания — оригинал-макетом. Использование издательских систем позволило значительно сократить срок подготовки печатных изданий, снизить трудоёмкость этого процесса, значительно расширить творческие возможности дизайнеров печатных изданий.

Электронные переводчики и словари предназначены для автоматического перевода текстов с одного языка на другой, проверки правописания текстов на разных языках. Особым видом словарей являются тезаурусы — словари, в которых слова связываются на основе каких-либо лексических отношений (например, слова являющиеся синонимами, антонимами и т. п.). Примеры — PROMT, ABBYY Lingvo.

Системы оптического распознавания текстов (например, ABBYY FineReader) предназначены для преобразования отсканированного графического изображения текстового документа в текстовый формат.

Кроме того, программы для работы с текстовой информацией интегрированы в системы программирования, а также являются частью HTML-редакторов, предназначенных для создания веб-страниц.

23.3. Создание текстовых документов на компьютере

При подготовке текстовых документов на компьютере используются три основные группы операций: ввод, редактирование, форматирование.

Операции ввода позволяют сформировать содержимое и первоначальный вид текстового документа и сохранить его в памяти компьютера. Ввод может осуществляться не только набором с помощью клавиатуры, но и путём сканирования бумажного оригинала и последующего перевода документа из графического формата в текстовый (расознавания).



Напомним основные правила ввода текстовых документов с помощью клавиатуры.

- Между словами ставится только один пробел.
- Переход на новую строку в процессе набора текста происходит автоматически, не требуя ввода специального символа.
- Окончание абзаца маркируется нажатием клавиши Enter, позволяющей перейти на новую строку — первую строку нового абзаца.
- Перед знаками препинания (такими как ; : . , ! ?) пробел не ставится. Перед тире вводится пробел. После любого знака препинания вводится один пробел или символ конца абзаца.
- Знак «дефис» в словах вводится без пробелов.
- После открывающих и перед закрывающими скобками ({ } () []) и кавычками пробел не вводится.
- Для ввода римских цифр используются прописные латинские буквы I, V, X, L, C, D, M.
- Знак «неразрывный (нерастяжимый) пробел» препятствует символам, между которыми он поставлен, располагаться на разных строчках, и сохраняется фиксированным при любом выравнивании абзаца (не может увеличиваться, в отличие от обычного пробела). Этот знак очень удобно применять при вводе дат (которые не принято располагать на двух строчках), фамилий с инициалами и т. п. Например: А. С. Пушкин. Ставится знак «неразрывный пробел» с помощью одновременного нажатия комбинации клавиш Ctrl + Shift + пробел.

При вводе и редактировании текста полезно включать режим отображения скрытых символов — символов, которые вводятся пользователем при наборе текста, но при печати не выводятся на бумагу, а на экране отображаются только при включении соответствующего режима (табл. 5.1). Режим отображения скрытых символов даёт возможность лучше понять структуру документа.

Для автоматизации ввода в современных текстовых процессорах существуют инструменты **Автозамена** и **Автотекст**.

Бывает, что при вводе текста с клавиатуры пользователь допускает опечатки: вместо нужной клавиши нажимает соседнюю, пропускает букву, меняет две буквы местами. Такие опечатки исправляются автоматически инструментом **Автозамена**, имеющим встроенный словарь наиболее типичных опечаток и ошибочных написаний.

Таблица 5.1

Примеры скрытых символов

Название символа	Клавиша (комбинация клавиш)	Изображение
Пробел	Пробел	•
Конец абзаца	Enter	¶
Неразрывный пробел	Ctrl + Shift + пробел	○
Разрыв строки	Shift + Enter	↵
Разрыв страницы	Ctrl + Enter Разрыв страницы

Для быстрого ввода стандартных фраз по нескольким первым буквам можно использовать инструмент **Автотекст**. Он автоматически предлагает вставить короткую фразу из списка элементов автотекста, как только будут набраны несколько первых букв этой фразы.

Операции редактирования (правки) позволяют изменить уже существующий электронный документ путём добавления или удаления фрагментов, перестановки частей документа, слияния нескольких файлов, разбиения единого документа на несколько более мелких и т. д. (рис. 5.1)

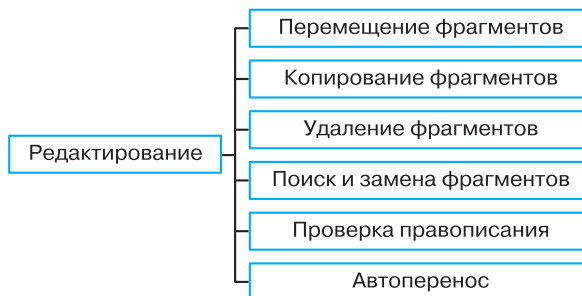


Рис. 5.1. Основные операции редактирования документов



На протяжении многих веков для внесения изменений в текст нужно было заново переписывать его. Основное преимущество компьютерной технологии создания текстовых документов заключается именно в удобстве его редактирования. Возможность быстро исправлять ошибки является одной из основных причин повсеместного перевода подготовки текстовой информации с бумажной на компьютерную основу.

Ввод и редактирование при работе над текстом часто выполняются параллельно. При вводе и редактировании формируется содержание текстового документа.



Совокупность значений свойств объекта называют форматом объекта, а изменение этих значений — форматированием объекта.

Операции форматирования позволяют точно определить, как будет выглядеть текст на экране монитора или на бумаге после печати на принтере. Операции форматирования могут применяться как к отдельным объектам текстового документа (табл. 5.2), так и ко всему документу в целом. В первом случае говорят о прямом форматировании, во втором — о стилевом.

Таблица 5.2

Объекты текстового документа и их свойства

Объект	Свойства объекта
Символ	Шрифт, размер, цвет, начертание, видоизменение и др.
Слово	Межзнаковый интервал и др.
Абзац	Отступ от границ левого и правого полей, отступ первой строки, выравнивание, междустрочный интервал, расстояние между абзацами и др.
Страница	Размеры страницы, её ориентация, размеры полей, колонтитулы и др.
Список	Вид списка (маркированный, нумерованный, многоуровневый) и др.
Таблица	Количество строк и столбцов, их высота и ширина, границы, заливка и др.
Графическое изображение	Тип, размер, положение, обтекание текстом и др.

Такие действия по оформлению документа как выравнивание абзацев, установка абзацных отступов и интервалов между абзацами, строками в абзацах и символами в словах и т. п. выполняются специальными средствами текстовых процессоров, а не вставкой пробелов и пустых строк.

Для облегчения анализа и последующего преобразования текста очень важно соблюдать основные правила его ввода, редактирования и форматирования.

В современных текстовых процессорах есть специальные инструменты, обеспечивающие автоматическую нумерацию страниц, таблиц и рисунков.

При работе с большими текстами, как правило, применяют **стилевое форматирование**. Смысл этой операции заключается в том, что структурным элементам, несущим одну и ту же функциональную нагрузку (например, заголовкам одного уровня, основному тексту, примерам и т. д.), назначается определённый стиль форматирования — набор параметров форматирования (шрифт, его начертание и размер, абзацные отступы, междустрочный интервал и др.).

Стиль — это имеющий имя набор значений свойств объектов каждого типа, входящих в текстовый документ.

В заключение приведём основные правила оформления текстов:

- основной текст документа желательно оформлять в одном формате, другой формат использовать для выделения заголовков, отдельных смысловых фрагментов;
- количество разных цветов и шрифтов в документе не должно превышать трёх;
- размер символов и междустрочный интервал следует подбирать такими, чтобы текст читался легко, без дополнительной нагрузки для глаз;
- цветовая гамма должна отвечать назначению документа — поздравительная открытка может иметь яркий, насыщенный окрас, а простое письмо стоит оформлять в более спокойных тонах;
- однотипную информацию целесообразно представлять в списках и таблицах;

- графические изображения в документе (рисунки, диаграммы, схемы) должны дополнять содержание текста, разъяснять или иллюстрировать его отдельные моменты; графические изображения следует оформлять в едином стиле;
- на всех страницах документа рекомендуется делать одинаковый фон и поля, если иное не требуется содержанием документа.

23.4. Средства автоматизации процесса создания документов

Мы рассмотрели основные операции ввода, редактирования и форматирования документов. Многие из них в той или иной мере направлены на автоматизацию процесса создания текстовых документов. Ещё больше возможностей в этом направлении обеспечивает использование шаблонов, макросов и средств, обеспечивающих работу со структурными компонентами документа.

Многие типовые документы должны иметь стандартный вид, который определяет, что и где размещается в создаваемом тексте, например: кому адресован документ, от кого он, дата создания документа и другие реквизиты.



Требования к оформлению, структуре и содержанию многих документов устанавливаются стандартами. Все они находятся в открытом доступе в Интернете.

Найдите в Интернете и познакомьтесь со следующими стандартами:

- ГОСТ Р 6.30-2003 «Унифицированные системы документации. Унифицированная система организационно-распорядительной документации. Требования к оформлению документов»;
- ГОСТ 2.105-95 «ЕСКД. Общие требования к текстовым документам»;
- ГОСТ Р 7.0.5-2008 «СИБИД. Библиографическая ссылка. Общие требования и правила составления»;
- ГОСТ 7.32-2001 «СИБИД. Отчёт о научно-исследовательской работе. Структура и правила оформления».

Какие из них имеют статус государственных, а какие — межгосударственных? Как это отражено в названии стандартов?

Что означают аббревиатуры ГОСТ, ЕСКД, СИБИД?

Какие из этих стандартов могут быть полезны в вашей учебной деятельности?

В текстовых процессорах есть шаблоны для создания документов разного типа.

Шаблон — это отформатированный определённым образом документ-заготовка, который хранится в отдельном файле и используется в качестве основы для создания новых документов определённого типа.

Пользователю достаточно ввести свою информацию в отдельные блоки шаблона, и она автоматически приобретёт заранее заданное оформление. В недалёком будущем каждому из вас для поиска подходящей работы придётся составить и разослать резюме. Подготовить его лучше всего с использованием соответствующего шаблона (рис. 5.2).

▶ **[Введите свое имя]**

[Введите свой адрес]

Телефон: [Введите свой номер телефона]

Электронная почта: [Введите свой адрес электронной почты]

Веб-узел: [Введите название своего веб-узла]

Цели

[Введите цели]

Образование

[Введите ученую степень] ([Введите дату завершения])

▶ [Введите список поощрений]

Опыт работы

[Введите должность] ([Введите дату начала] – [Введите дату окончания])

[Введите название организации] ([Введите адрес организации])

[Введите служебные обязанности]

Умения и навыки

▶ [Введите список умений и навыков]

Рис. 5.2. Шаблон резюме

В текстовом процессоре Microsoft Word все шаблоны распределены на три группы:

- 1) установленные — шаблоны документов определённых типов (писем, факсов, отчётов и др.), которые установлены на компьютере в составе пакета Microsoft Office;
- 2) Microsoft Office Online — шаблоны документов разнообразных типов (поздравительных открыток, визиток, бюллетеней, сертификатов, грамот, приглашений, заявлений, календарей и др.), которые расположены на веб-сайте Microsoft Office Online;
- 3) шаблоны пользователя — шаблоны, которые созданы пользователем.

При запуске программы Microsoft Word автоматически открывается шаблон **Новый документ** (файл Normal.dotm). При этом по умолчанию устанавливается формат (значения свойств) основных объектов документа — страницы, абзаца, символа, а также задаётся стилевое форматирование заголовков, списков, таблиц и др.

Запустите имеющийся в вашем распоряжении текстовый процессор. Исследуйте формат основных объектов документа в шаблоне **Новый документ**, выяснив для них значения свойств, приведённых ниже:

Страница	Абзац	Символ
<ul style="list-style-type: none"> • Ориентация • Размер листа • Верхнее поле • Нижнее поле • Левое поле • Правое поле 	<ul style="list-style-type: none"> • Междустрочный интервал • Интервал после абзаца • Выравнивание • Отступы 	<ul style="list-style-type: none"> • Шрифт основного текста • Размер • Цвет • Интервал

В процессе работы над документом в программе Microsoft Word часто приходится выполнять задания по некоторому алгоритму, состоящему из определённой последовательности действий. Например, подчеркнуть слова, написанные латинскими буквами, отформатировать какое-то слово во всём документе определённым образом и др.

Макрос — это последовательность команд, сгруппированных в одну макрокоманду, для автоматического выполнения определённого задания.

Основное назначение макроса состоит в том, чтобы освободить пользователя от многократного повторения однообразных действий во время обработки текстового документа, выполнить за него рутинную работу. Макрос создаётся один раз, сохраняется в шаблоне или документе и может многократно выполняться для автоматизации и ускорения обработки текстового документа.

Многостраничные документы (рефераты, брошюры, книги и т. п.) принято делить на структурные части — главы, параграфы, пункты и т. п., создавая таким образом иерархическую структуру документа. Рассмотрим в качестве примера структуру этого учебника. На верхнем (нулевом) уровне иерархии находится название документа («Информатика. 10 класс»); на первом уровне — названия глав, второй уровень составляют названия параграфов, третий — названия пунктов в параграфах, дальше размещается основной текст учебника.

Структура документа — это иерархическая схема размещения составных частей документа.

Использование в текстовом процессоре Microsoft Word специальных стилей с именами Заголовок 1, Заголовок 2 и т. д. даёт возможность автоматизировать создание иерархической структуры документа. В текстовом процессоре Microsoft Word для просмотра структуры документа используется режим просмотра **Структура**. В нём удобно редактировать иерархическую схему документа, изменяя с помощью специальных инструментов уровень текстовых фрагментов и последовательность их размещения.

Современные текстовые процессоры позволяют автоматически создавать оглавления документов, в которых к заголовкам разделов разных уровней применено стилевое форматирование. С помощью специальной команды пользователь указывает, заголовки каких уровней следует включить в оглавление, и абзацы указанных стилей автоматически выбираются из текста документа и помещаются с указанием номеров страниц, с которых они были взяты, в новый раздел «Оглавление».

Оглавление документа — это перечень названий структурных частей документа, упорядоченных в соответствии с его иерархической схемой, с указанием соответствующих номеров страниц.

23.5. Совместная работа над документом



Под совместной (коллективной) работой над документом принято понимать участие нескольких человек в создании одного текстового документа, при котором у каждого из них есть возможность отслеживать все изменения, сделанные в документе другими разработчиками, а также осуществлять возврат к одной из предыдущих версий документа.

На протяжении долгого времени процесс совместной работы над документом был устроен так: кто-то создавал текст на заданную тему, распечатывал его и отдавал другому человеку. Этот человек дописывал «от руки» свои вставки и комментарии и возвращал черновики на доработку. Такие итерации могли происходить неоднократно, в результате работа шла медленно, к тому же отдельные ценные мысли и идеи могли быть утеряны безвозвратно.

Современные инструменты создания текстовой информации предоставляют принципиально иные возможности для совместной работы над документом, поддерживая следующие варианты её организации.

1. Документ, над которым ведётся работа, можно сделать составным, объединяющим несколько других документов. Каждый разработчик создаёт и редактирует свою часть составного документа независимо от других, при этом в процессе работы он может просматривать текущую версию общего документа. Такой вариант организации совместной работы основывается на возможности создания мастер-документа, к которому могут прикрепляться другие документы.
2. Каждая из частей составного документа может редактироваться несколькими людьми. Для этого в современных текстовых процессорах предусмотрена возможность отслеживать и протоколировать сделанные изменения, делать примечания — пометки на полях.

Для отслеживания собственных исправлений, а также чтобы другие соавторы могли вносить изменения в документ, текстовый процессор предлагает использовать маркеры исправлений. Маркеры исправлений помогают увидеть, какие изменения были внесены в документ по сравнению с его последней версией. Для изображения исправлений используется специаль-

ный формат, например подчёркивание. С помощью маркеров исправлений можно сохранить запись о каждом сделанном исправлении и в дальнейшем либо принять его, либо отказаться. Исправление помечается полным именем его автора, датой и временем создания.

При совместной работе над документом важно, чтобы в настройках текстового процессора были указаны корректные данные о пользователе, т. к. именно они останутся в редактируемом документе.



Примечания — это обозначенные инициалами и пронумерованные комментарии, которые записываются и отображаются в специальном окне примечаний и не затрагивают текст документа. Перед тем как вставить примечание, имеет смысл выделить фрагмент текста, который следует прокомментировать. В этом случае при просмотре примечания текст, к которому оно относится, будет подсвечен.

Имеются возможности сравнения двух версий документов или объединения всех исправлений и примечаний в один документ, для просмотра их всех сразу. Сделанные в тексте изменения перед сохранением окончательной версии документа можно принять или отклонить.

Оба рассмотренных варианта организации совместной работы поддерживаются и в OpenOffice Writer, и в Microsoft Word.

3. Можно редактировать документ непосредственно в сети, так же отслеживая все изменения, сделанные другими пользователями. Этот вариант поддерживается Google Docs — сетевым приложением, доступным через веб-браузер. Поскольку редактируемый документ сохраняется на сервере, он доступен всем пользователям, между которыми он разделяется. Каждый раз при сохранении документа сохраняется его новая версия и информация о том, кто этот документ редактировал. Команда просмотра изменений открывает список сделанных изменений с указанием, кто и когда их внёс. Различные версии документов можно сравнить между собой.
4. Можно использовать и промежуточный вариант, когда документ редактируется локально на компьютере, а храниться в сети и там же отслеживаются все изменения.

23.6. Оформление реферата как пример автоматизации процесса создания документов

Старшеклассники, студенты, курсанты в процессе своей учебной деятельности готовят рефераты по различным предметам.



Реферат — это самостоятельная исследовательская работа, в которой автор раскрывает суть исследуемой проблемы, приводит различные точки зрения, делает собственные выводы.

Выбрав тему реферата, необходимо определить цель работы, составить план (поставить задачи, определить порядок и сроки выполнения задач), найти и изучить материалы различных информационных источников, собрать и обработать информацию, сделать выводы, оценить полученные результаты.

Содержание реферата должно быть логичным, изложенным ясным языком. Основные положения реферата желательно подкреплять цитатами и ссылками на информационные источники.

Есть определённые требования и к оформлению реферата.

Реферат должен быть выполнен на одной стороне листов белой бумаги формата А4 (210 × 297 мм).

Размеры полей страницы (не менее):

- правое — 30 мм (для замечаний преподавателя);
- верхнее, нижнее, левое по 20 мм.

Отступ первой строки: 8–12 мм, одинаковый по всему тексту.

Интервал междустрочный: полуторный.

Выравнивание абзаца: по ширине.

Гарнитура шрифта основного текста — Times New Roman или аналогичная.

Кегль (размер): 12–14 пунктов.

Цвет шрифта: чёрный.

Заголовки разделов и подразделов следует печатать на отдельной строке с прописной буквы без точки в конце, не подчёркивая. Если заголовок состоит из нескольких предложений, их разделяют точкой. Выравнивание по центру или по левому краю. Интервал: перед заголовком — 12 пунктов, после — 6 пунктов.

Страницы следует нумеровать арабскими цифрами, соблюдая сквозную нумерацию по всему тексту (титальный лист и оглавление).

ние включают в общую нумерацию). На титульном листе номер не проставляют.

В верхней части титульного листа пишется, в каком образовательном учреждении выполняется работа, далее буквами увеличенного кегля указывается тип («Реферат») и тема работы, ниже в правой половине листа — информация о тех, кто выполнил и кто проверит работу. В центре нижней части титульного листа пишется название населённого пункта и год выполнения работы.

Из курса русского языка вам известно, что цитата — это приведённое полностью или частично высказывание из авторского текста (научной, художественной, публицистической и др. литературы или доклада). Цитаты оформляются как прямая речь или как продолжение предложения.

Правовой статус цитирования определяется Гражданским кодексом РФ, согласно которому цитирование «допускается без согласия автора или иного правообладателя и без выплаты вознаграждения, но с обязательным указанием имени автора, произведение которого используется, и источника заимствования» (статья 1274 части 4 Гражданского кодекса РФ).

Допустим, при работе над рефератом вы взяли (заимствовали) информацию из источника «Андреева, Е. В. Математические основы информатики. Элективный курс: учебное пособие / Е. В. Андреева, Л. Л. Босова, И. Н. Фалина. — М.: БИНОМ. Лаборатория знаний, 2005.». Необходимо правильно оформить ссылку на этот источник и внести его в список литературы.

Правила оформления ссылок регулируются ГОСТ Р 7.0.5-2008 «Система стандартов по информации, библиотечному и издательскому делу. Библиографическая ссылка. Общие требования и правила составления».

Правила оформления библиографических сведений в списке использованной литературы должны отвечать требованиям ГОСТ 7.1-2003 «Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Библиографическое описание. Общие требования и правила составления».

Для оформления ссылки в текстовом процессоре Microsoft Word можно:

- 1) установить курсор после заимствованного текста и выполнить команду **Ссылки → Вставить ссылку → Добавить новый источник...**;
- 2) заполнить поля диалогового окна **Создать источник** (рис. 5.3);



Рис. 5.3. Окно Создать источник

3) затем щёлкнуть на кнопке **ОК** и после заимствованного фрагмента текста появится ссылка на источник в виде: «(Е. В. Андреева, 2005)».

Для оформления списка литературы в текстовом процессоре Microsoft Word можно:

- 1) выбрать стиль отображения списка литературы (библиографического списка), выполнив команду **Ссылки → Стиль: → ГОСТ – сортировка по именам 2003**;
- 2) выполнить команду **Ссылки → Список литературы → Вставить список литературы**.

23.7. Другие возможности автоматизации обработки текстовой информации

Компьютер помогает не только автоматизировать процесс создания текстовых документов, но и решить множество других задач, связанных с обработкой текстовой информации. Вот некоторые из них:

- **поиск текста** в общем массиве по заданным нечётким признакам;
- **рубрицирование текста** — разбиение поступающего потока текстов на тематические подпотоки в соответствии с заранее заданными рубриками;
- **реферирование текста** — подготовка его сокращённой версии;
- **перевод текста** с одного языка на другой;
- **анализ текста** на предмет выявления заимствований.

Область информатики, решающая эти и другие задачи, связанные с обработкой информации на естественном языке, называется компьютерной лингвистикой.

Рассмотрим более подробно задачу поиска текста в общем массиве. Существует несколько подходов к её решению.

Первый подход опирается на поиск фрагмента текста, соответствующего некоторому образцу. Таким способом в большом текстовом массиве можно находить упоминания тех или иных слов, адреса, номера телефонов и другие элементы. Основное достоинство такого подхода — возможность применять его к массиву текста без предварительной обработки (например, сразу при посимвольном получении текста). Применение рассматриваемого способа бывает затруднено, если текст хранится в разных местах.

Второй подход предусматривает предварительную обработку текста с целью получения его преобразованного, сокращённого вида (индекса). Получив запрос, поисковая система выделяет список слов и составляет список документов, в которых они содержатся. При этом рассчитывается релевантность — мера соответствия документа запросу, зависящая от наличия искомых слов, близости их друг к другу и других параметров. Документы с высокой релевантностью помещаются в начало списка, с низкой — в конец.

Одно из интересных применений автоматического анализа текстов — выявление заимствований.

Антиплагиат (antiplagiat.ru) — российский интернет-проект, программно-аппаратный комплекс для проверки текстовых документов на наличие заимствований из страниц сети Интернет и других источников. Проект доступен для всех пользователей.



САМОЕ ГЛАВНОЕ

Информационные технологии (ИТ) — это совокупность методов, производственных процессов, программно-технических и лингвистических средств, объединённых с целью сбора, обработки, хранения, распространения, отображения и использования информации, представленной в цифровой форме.

С позиции информатики, текст — это последовательность знаков некоторого алфавита. Существует множество программных продуктов, предназначенных для работы с текстовой информацией.

При подготовке текстовых документов на компьютере используются три основные группы операций: ввод, редактирование, форматирование.

Операции ввода позволяют сформировать содержимое и первоначальный вид текстового документа и сохранить его в памяти компьютера.

Операции редактирования (правки) позволяют изменить уже существующий электронный документ путём добавления, удаления, перестановки фрагментов, слияния нескольких файлов, разбиения единого документа на несколько более мелких и т. д.

Операции форматирования позволяют точно определить, как будет выглядеть текст на экране монитора или на бумаге после печати на принтере. Операции форматирования могут применяться как к отдельным объектам текстового документа, так и ко всему документу в целом.

Автоматизация процесса создания текстовых документов обеспечивается за счёт возможности работы с фрагментами, проверки правописания, стилевого форматирования, а также использования шаблонов, макросов и средств, обеспечивающих работу со структурными компонентами документа.

Компьютер помогает автоматизировать не только процесс создания текстовых документов, но и решить множество других задач, связанных с обработкой текстовой информации, а именно:

- поиск текста в общем массиве по заданным нечётким признакам;
- рубрицирование текста — разбиение поступающего потока текстов на тематические подпотоки в соответствии с заранее заданными рубриками;
- реферирование текста — подготовка его сокращённой версии;
- перевод текста с одного языка на другой;
- анализ текста на предмет выявления заимствований.



Вопросы и задания

1. Что, по вашему мнению, стимулировало развитие технологий обработки текстовой информации? Как это связано с понятием «бесбумажные технологии»? Наступит ли такое время, когда все технологии обработки текста будут бесбумажными?
2. На страницах Википедии найдите информацию об истории создания программы Microsoft Word. Составьте хронологическую таблицу.
3. Каковы основные правила ввода текста с клавиатуры?
4. В чём отличие использования в тексте символа «тире» от символа «дефис», «пробела» от «неразрывного пробела»?



5. Найдите информацию о разных способах ввода текста в память компьютера (рукописном, голосовом, сканировании). Какое оборудование и программное обеспечение для этого используют?
6. Ваш одноклассник жалуется на то, что при вводе нового текста старый текст частично исчезает. Что вы ему посоветуете?
7. Перед вами текст на русском языке, содержащий отдельные английские слова:

Растровые изображения можно получить, сканируя рисунки или фотографии, фотографируя объекты цифровым фотоаппаратом, создавая рисунки с использованием графического планшета или разнообразных растровых графических редакторов (Paint, GIMP, Adobe Photoshop).

Почему правильно написанные слова оказались подчеркнуты так, как если бы в них были ошибки?

8. Почему иногда в тексте, написанном на одном языке, некоторые слова подчёркиваются волнистыми линиями, даже если в них нет ошибок? Как исправить эту ситуацию?
9. Как в документе осуществить автоматическую замену одного фрагмента текста на другой? Как, используя операцию автоматической замены, быстро объединить все абзацы текста?
10. Один ученик хотел заменить в тексте слово «мы» на слово «я», но получил текст, в котором появилось много ошибок. Объясните эту ситуацию. Как можно её исправить?
11. Списки каких типов вам известны? В каких ситуациях следует применять каждый из них?
12. В списке учеников вашего класса было пропущено несколько фамилий. Каким образом их можно вставить в текст?
13. Фамилии учеников вашего класса были введены в произвольном порядке. Каким образом можно расположить фамилии в алфавитном порядке?
14. В таблице выделено несколько строк. Выясните, что произойдёт при нажатии клавиши Delete; клавиши BackSpace.
15. Графические документы каких видов можно вставлять в текстовый документ?

16. В чём заключается процесс форматирования текста?
17. Автор некоторого документа ввёл заголовок текста, в котором символы отделил пробелами для увеличения расстояния между ними. Правильно ли он сделал? Что вы ему посоветуете?
18. Некоторое слово нужно обязательно разместить в данной строке текста, а оно там не помещается. Что вы предпримете в данной ситуации?
19. Как можно ввести в текст математические выражения, которые содержат верхний и нижний индексы, обычные дроби, буквы греческого алфавита?
20. Подумайте, какие преимущества обеспечивает стилевое форматирование по сравнению с прямым форматированием.
21. Как можно ускорить процесс форматирования текста, в котором часто встречаются фрагменты с одинаковым форматом?
22. Найдите информацию о правилах оформления деловых документов (заявления, справки, докладной записки и др.). Выясните, какие существуют требования к их оформлению.
23. Каковы общие правила стилевого оформления документов?
24. Выясните, что понимается под корпоративным (фирменным) стилем оформления документов. Кто и для чего его разрабатывает? Каковы его основные черты? Приведите примеры документов, оформленных в корпоративном стиле.
25. Исследуйте шаблоны документов текстового процессора, имеющегося в вашем распоряжении. Выясните их количество, тематику документов, которые могут быть созданы на их основе.
26. Что такое макрос? Для чего он используется?
27. Что такое структура документа? Приведите примеры структурированных документов.
28. Создайте кластер «Инструменты автоматизации создания текстовых документов» с учётом таких групп операций, как ввод, редактирование и форматирование.
29. Перечислите основные классы задач, связанных с автоматизацией обработки текстовой информации.
30. Охарактеризуйте наиболее распространённые форматы текстовых документов — TXT, RTF, PDF, DOC, DOCX, ODT.

