

Лабораторная работа № 8

Тема работы: «Распознавание информации. Форматирование текста в текстовом редакторе MS Word. Стилизовое оформление документа»

Цели работы

1. Обучить приемам распознавания информации; выработать умение использовать средства автоматизации для разработки и оформления документа, в результате чего учащиеся выполняют сканирование и распознавание информации; используют возможности текстового редактора для стилизового оформления документа.

Задание

1. Распознать pdf документ в ABBYY FineReader
2. Применить к текстовой информации ниже перечисленное форматирование в текстовом редакторе MSWord (п.5).

Основные теоретические сведения

Распознавание текста это процесс перевода графического изображения символов (букв) в компьютерные текстовые символы. Сделать это можно имея качественную цифровую копию оригинального текста и набор современных компьютерных программ для распознавания текста.

ABBYY FineReader —это система оптического распознавания текстов (OCR—Optical Character Recognition). Она предназначена для конвертирования в редактируемые форматы отсканированных документов, PDF-документов и файлов изображений документов, включая цифровые фотографии.

Преимущества программы ABBYY FineReader:

- Скорость и высокая точность распознавания
- Используемая в ABBYY FineReader система оптического распознавания быстро и точно распознает и максимально полно сохраняет исходное оформление документа.
- Благодаря технологии адаптивного распознавания документов ADRT® (Adaptive Document Recognition Technology) ABBYY FineReader позволяет анализировать и обрабатывать документ целиком, а не постранично. В результате восстанавливается исходная структура документа, включая форматирование, уровни заголовков, гиперссылки, а также колонтитулы, номера страниц и сноски.
- Еще одной особенностью программы ABBYY FineReader является малая чувствительность к дефектам печати и способность распознавать тексты, набранные практически любыми шрифтами.
- Программа также позволяет распознавать текст, полученный с фотоаппарата или встроенной камеры мобильного телефона. Дополнительные настройки предобработки изображений позволяют существенно улучшить качество изображений и получить более качественные результаты распознавания.
- Программа эффективно использует несколько ядер компьютера, чтобы обеспечить максимальную скорость обработки. Кроме того, возможна быстрая обработка в специальном черно-белом режиме.

Приступаем к распознаванию текста.

Для начала с помощью специализированных программ анализируется структура документов. На этом этапе важен контроль специалиста, так как возможные ошибки на этом этапе потребуют серьезных усилий по их устранению в последствии.

Следующий этап -непосредственное распознавание текста. Этот процесс так же доверяем компьютерным программам. Однако для успешного распознавания и минимизации ошибок, специалист должен, прежде всего, правильно подготовить документ для распознавания, настроить оборудование для получения максимально качественной цифровой копии оригинала.

Проверка результата и устранение ошибок

После окончания процесса распознавания, не смотря на постоянно совершенствующееся программное обеспечение, получившийся результат необходимо проверить и откорректировать вручную. Чем хуже состояние оригинала, тем больше возникает ошибок и только человек может провести коррекцию и исправить все недочеты.

Далее происходит проверка на предмет синтаксических и орфографических ошибок, расставляются, по необходимости, знаки препинания и специальные символы.

И последний этап это форматирование документа. Проверяется и если нужно корректируется размер шрифта, стили заголовков и текста, разбивается на абзацы, главы, проверяется нумерация страниц и оглавление документа. Также проверяется общая структура документа и верстка.

Графические элементы, картинки, графики, схемы и другие иллюстрации так же проходят проверку на предмет соответствия.

После проведения всех вышеописанных процедур, мы получаем точную и самое главное редактируемую копию оригинала. Теперь мы с легкостью можем вносить изменения, копировать и делиться этим документом, получить необходимое количество твердых копий. На основании этого документа можно создавать свои собственные документы. При этом документ будет храниться, без риска быть испорченным, потерянным или украденным.

Порядок выполнения работы

1. Распознать текст из Приложения 1.
- 2.Применить к тексту следующее форматирование:
 - 2.1. Установить для всего документа, кроме заголовков, красную строку 1,5см
 - 2.2. Убрать пустые строки, разделяющие абзацы.
 - 2.3. Установить для всего документа, кроме заголовков, шрифт TimesNewRoman, размер шрифта 12.
 - 2.4. Применить ко всему документу, за исключением заголовков, форматирование по ширине.
 - 2.5. Разместить заголовки и подзаголовки по центру.
 - 2.6. Внимательно просмотреть текст и исправить ошибки.
 - 2.7. Применить к заголовку стиль Заголовков 1.

2.8. Применить к остальным заголовкам стиль Заголовок 2.

2.9. Начинать каждый раздел с новой страницы; для этого использовать соответствующий переключатель в меню окна Вставка | Разрыв... .

2.10. Добавить в конце документа заголовок Список литературы, используя стиль Заголовок 1