

# Текстовая информация на компьютере: - программы и форматы

Руководство **не** только для **чайников**



© 1999 Артемий Лебедев

Автор: Билялов Р.Р.

Симферополь 2008

# Зарождение письменности

Со времён неандертальцев человечество пыталось сохранить информацию в виде рисунков и пиктограмм.

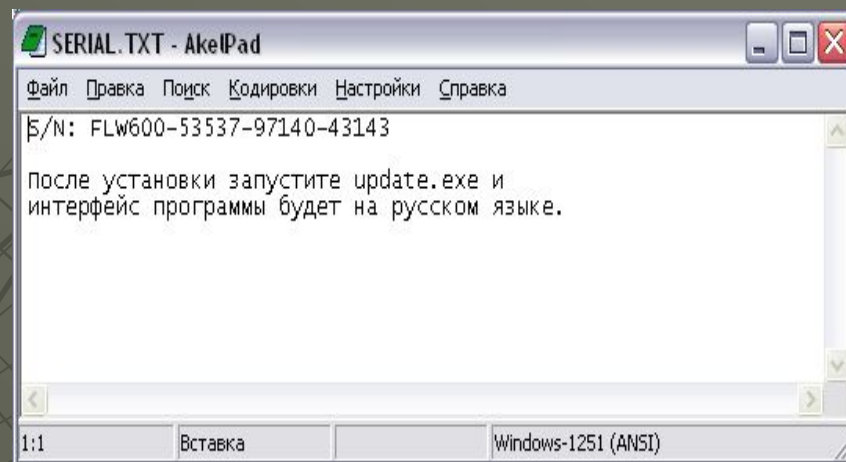
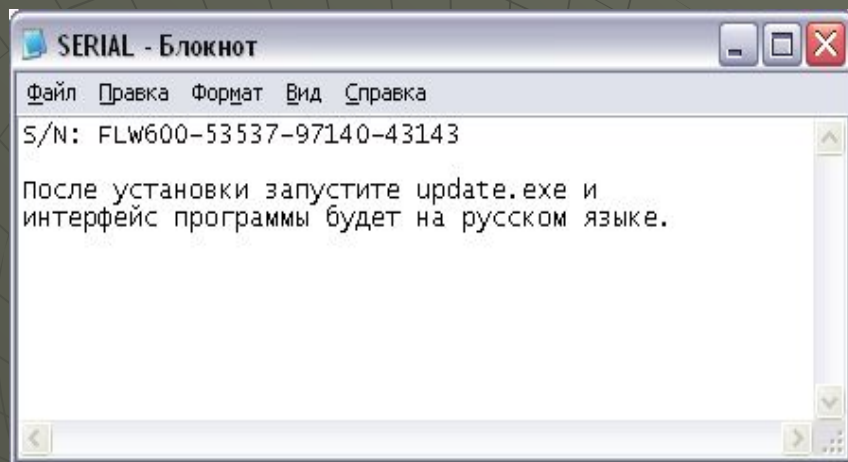




С развитием технологий, науки и религии количество  
накопленной информации  
возрастает лавинообразно и передавать её только рисунком уже  
невозможно ,  
поэтому на смену пиктограммам приходит **клинопись**



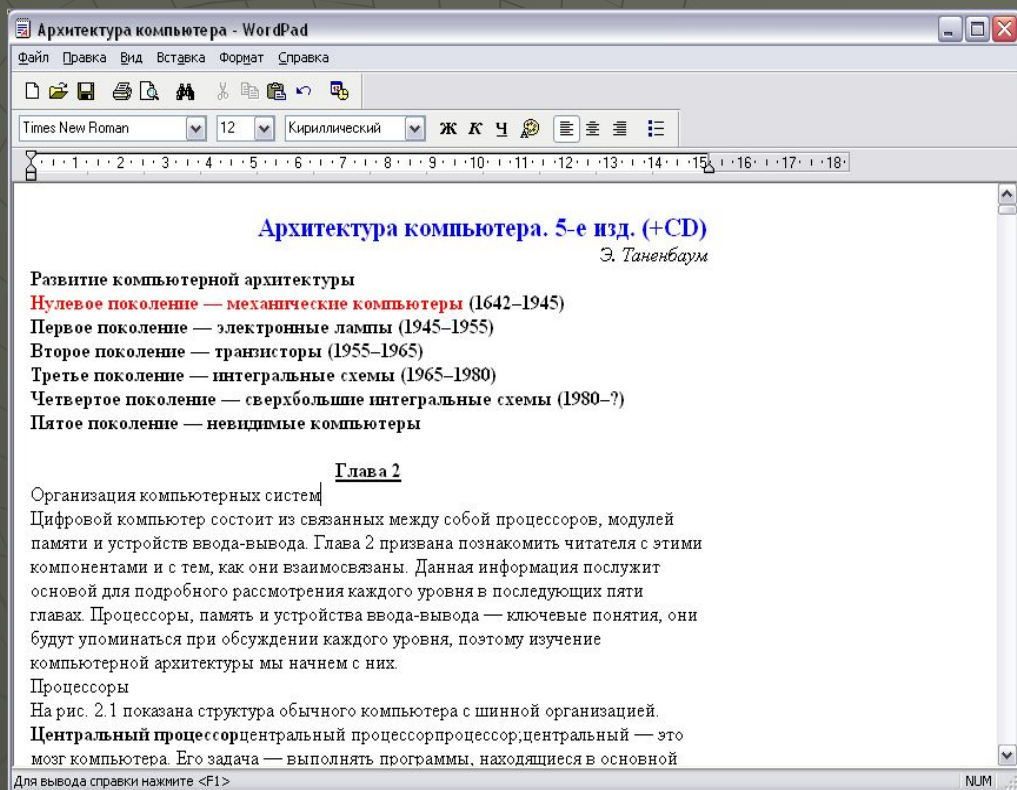
# Простейшие текстовые редакторы



- ♦ Вы видите два редактора – БЛОКНОТ и AkeIPad. На первый взгляд даже трудно найти различия в этих программах. Таких редакторов великое множество, но их объединяет общий формат - \*.txt (правда бывают случаи, когда про это расширение забывают, например файл - **readme**).
- ♦ Основным достоинством данного формата является минимальный размер файла.
- ♦ Удобен при отправке коротких сообщений, SMS
- ♦ Из недостатков – никаких украшательств – голый текст, хотя при доле фантазии можно ... 😊 [ :) 😞 :- ( :- )



# Редакторы среднего класса



- ♦ В данном случае на экране WordPad – редактор **интегрированный** (встроенный) в состав операционной системы Window's.
- ♦ Имеет собственный формат \*.rtf
- ♦ По сравнению с простейшими редакторами более функционален: может форматировать текст, вставлять объекты из других программ (рисунки, диаграммы, таблицы). Встроена функция поиска и замены **заданного по маске текста**.

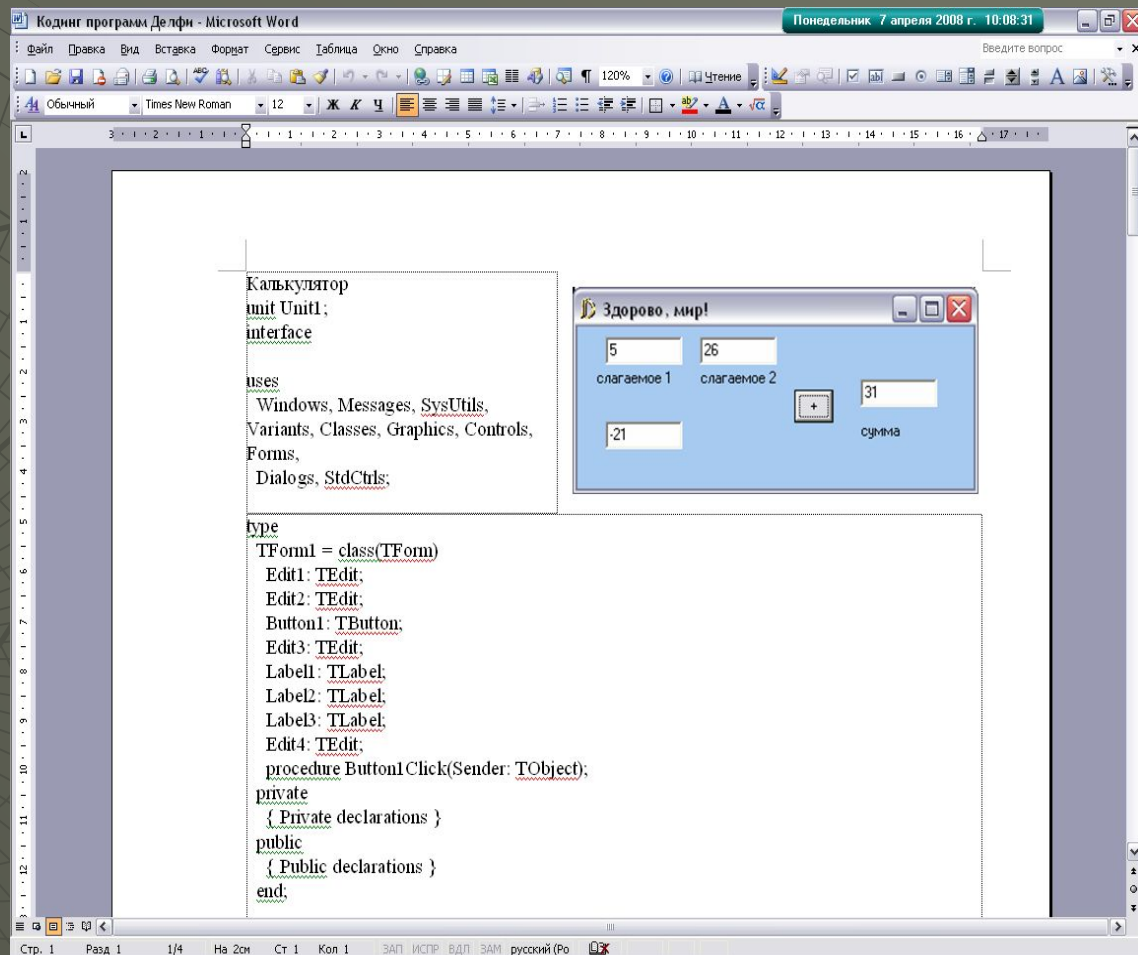
- ♦ Вот несколько побратимов:  
AbiWord 2.29, Atlantis Okean Mind Bred 3.0.3, Hieroglif 3.7, Patriot 2.1, и т.д.

# Текстовый процессор Word

- Один из наиболее распространённых текстовых редакторов планеты – входит в состав пакета Microsoft Office. Имеет собственный формат - \*.doc. В версии 2007 года - \*.docx

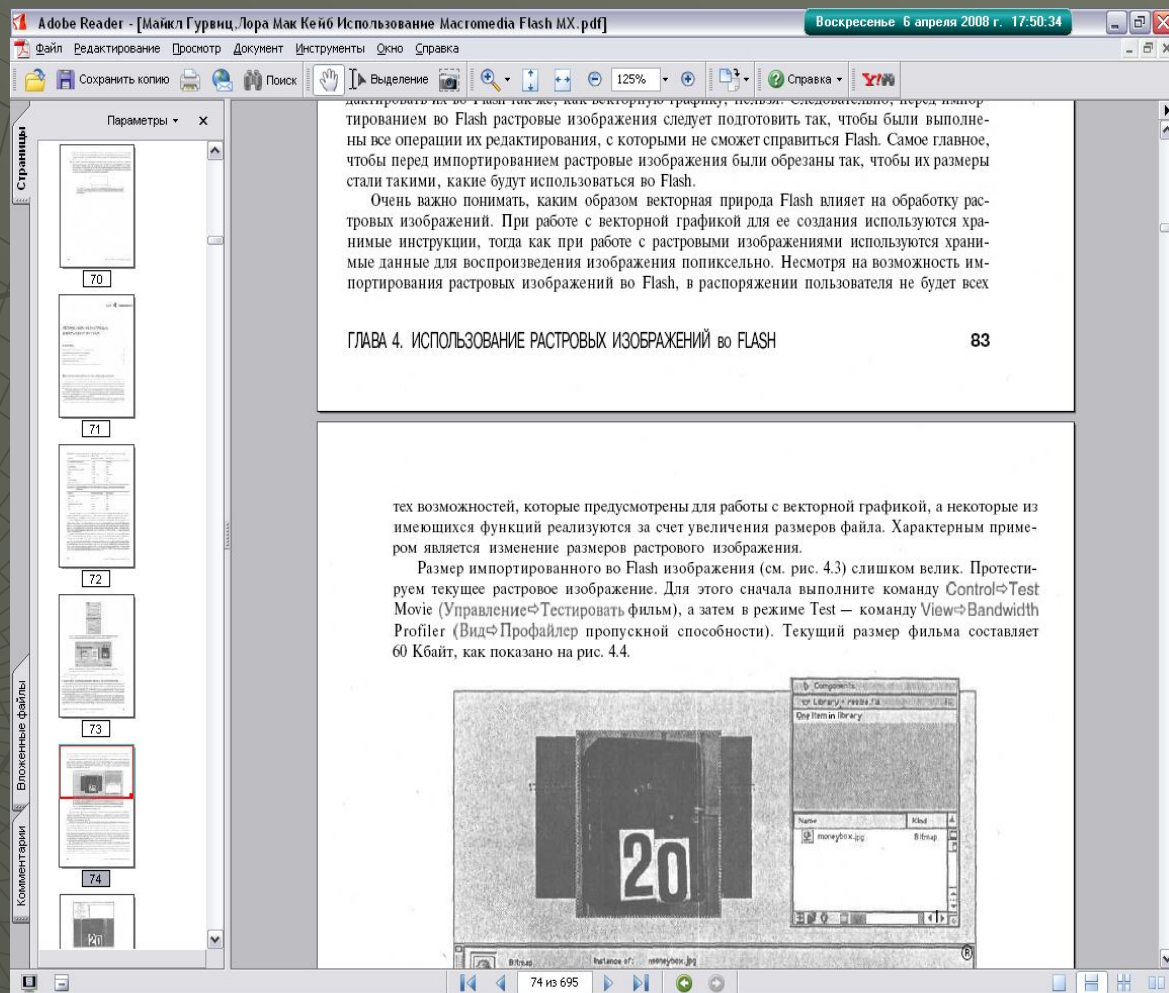
- В возможности редактора входит не только набор и редактирование текста, но и проверка орфографии, вставка графических объектов, форматирование текста(работа с различными шрифтами, их размерами и выделениями), вставка и обработка таблиц, поддержка макросов и даже встроенный язык программирования.

- Одним словом – с помощью Word можно верстать книги и журналы, красиво оформлять письма и даже отправлять их по электронной почте



# Графическое хранение текста

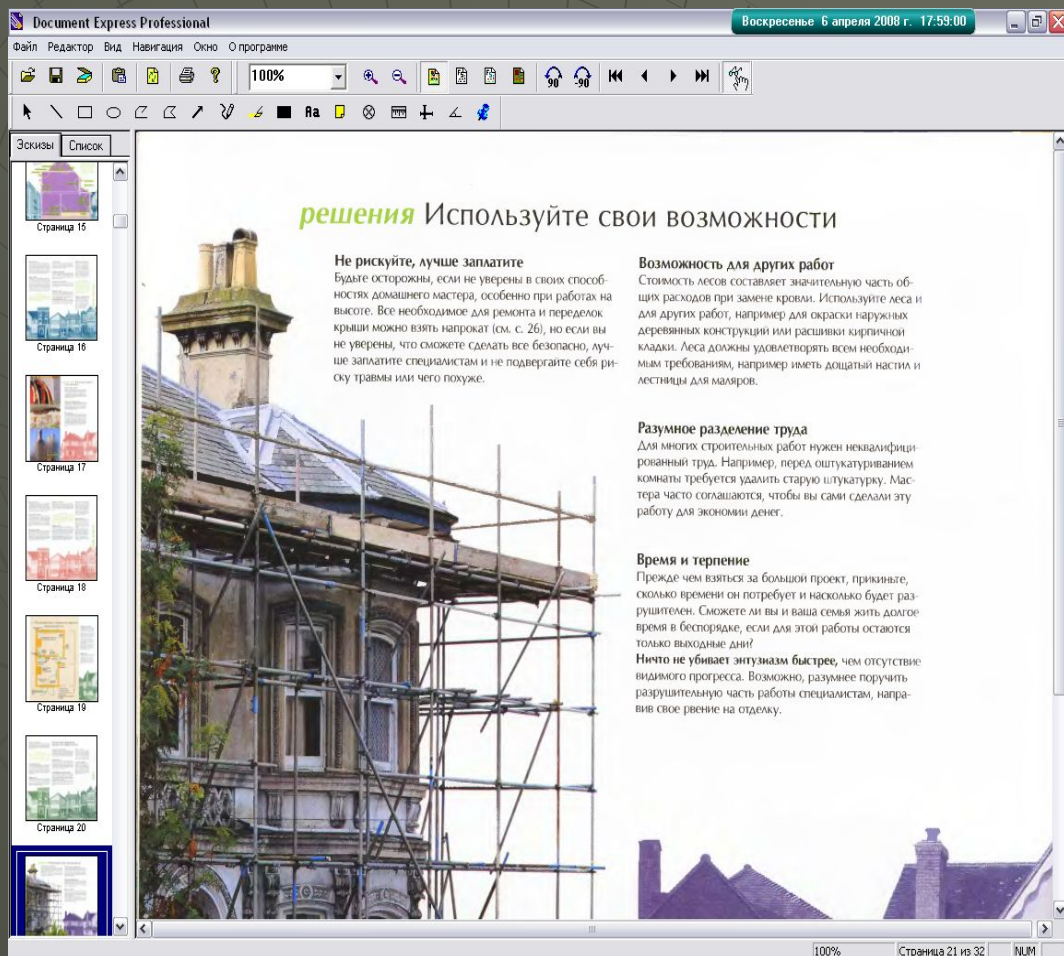
- Вы видите окно программы Adobe Acrobat Reader.
- Формат открываемого файла - \*.pdf
- Это графический вид хранения текстовой информации, то есть книгу сканировали на сканере и сохранили в данном формате, что значительно уменьшило объём файла.
- Книга в черно-белом варианте будет «**весить**» намного меньше цветной
- Программы – побратимы: Foxit PDF Reader 2.3, Jaws PDF Editor 2.5 (не только читает, но и редактирует файл)



Каждая страница – это отдельный рисунок!



# Графическое хранение текста

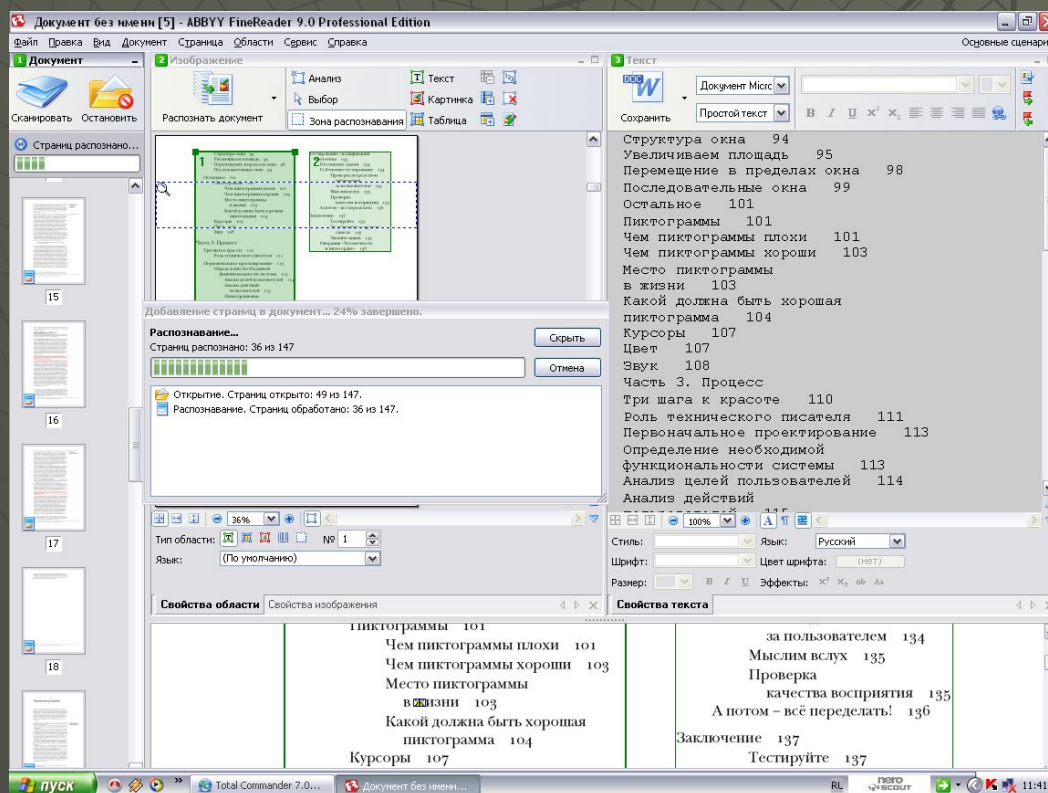


- ◆ Новый формат графического отображения книг и журналов \*.djvu.
- ◆ Он более уплотнён по сравнению с \*.pdf, то есть файл имеет меньший объём
- ◆ Позволяет в некоторых версиях маркировать текст, выделять цветом и линиями.
- ◆ Программы: Djvubookmark, Djvue Editor, DjVu Reader, WinDjView



# Перевод текста из графического формата в текстовый редактируемый

То есть превращение (если это возможно) графики в текст!  
Это такие программы как ABBYY FineReader, Qunei Form и т.д.



Принцип работы этой программы довольно сложен.

Первоначально отделяются графические элементы от текста.

Затем в текстовом блоке к каждому изображению буквы прикладываются контуры букв – распознавание

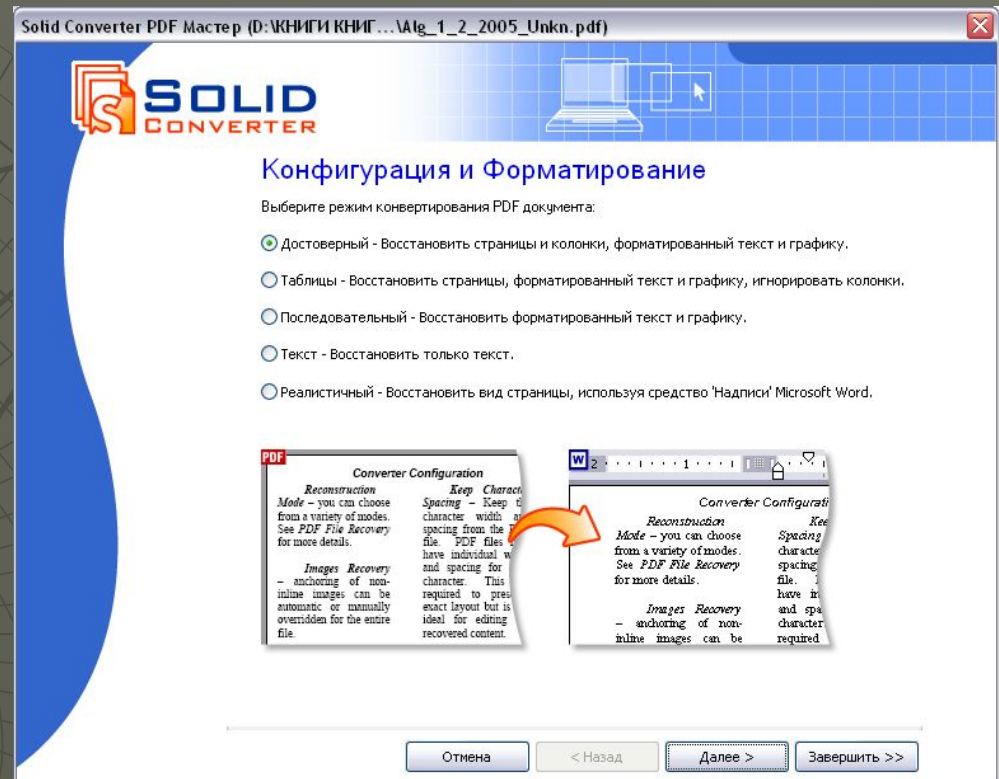
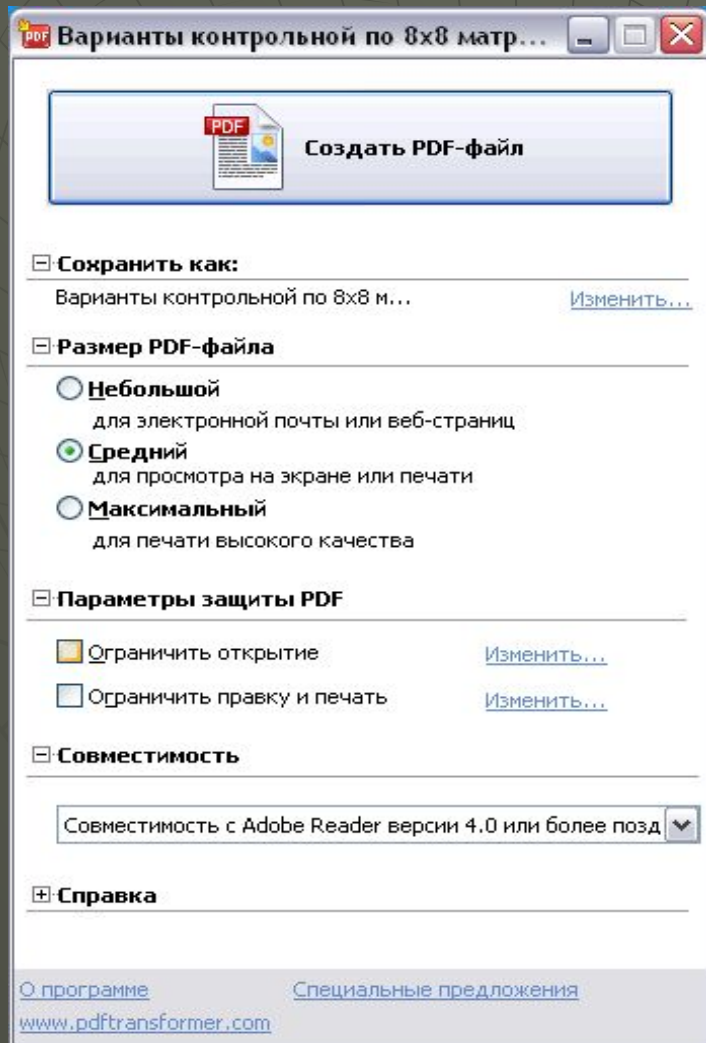
На последнем этапе анализируется архитектура документа – всё расставляется по своим местам, графика, текст, таблицы и т.д.

Из таких систем распознавания можно назвать, следующие:

Qunei Form

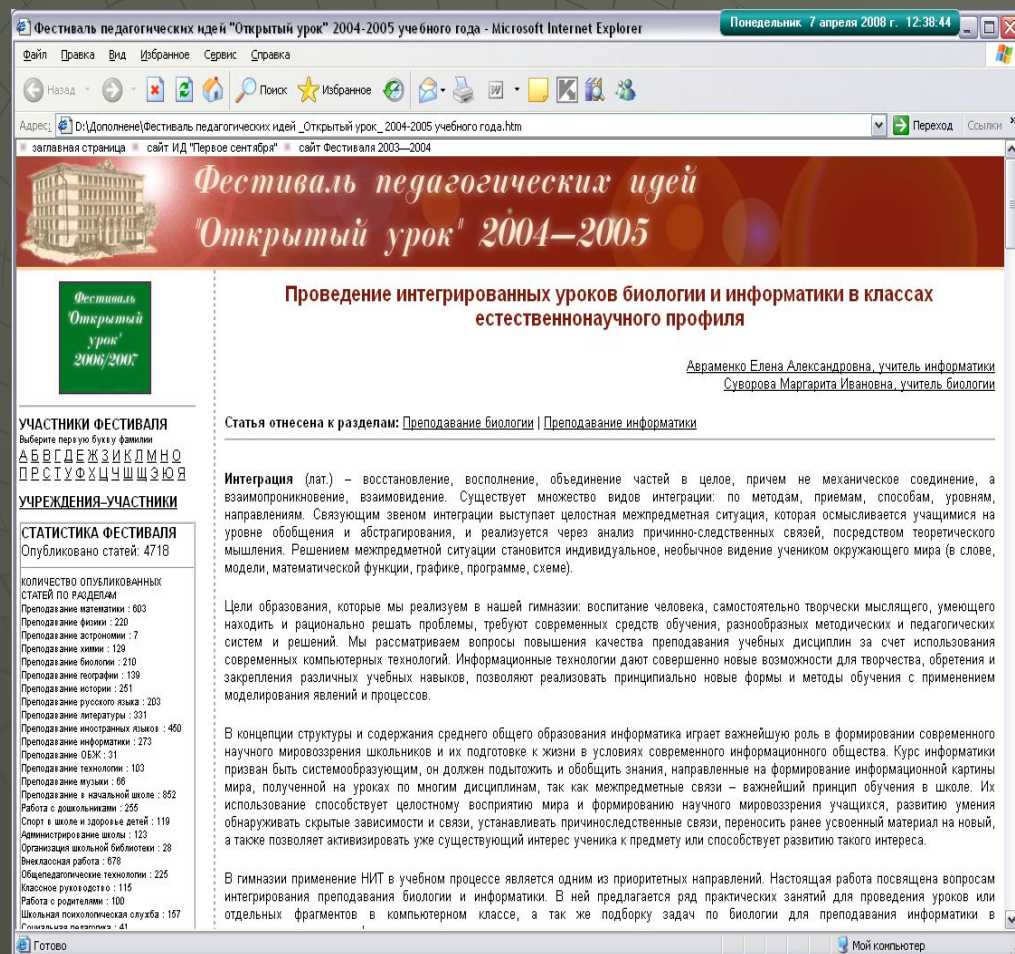
AfterScan Version 5.0

# Конвертеры текстовых форматов



Существует также масса конвертеров из одного формата в другой. (из графики в текст и из текста в графику)

# Текст на страницах Internet

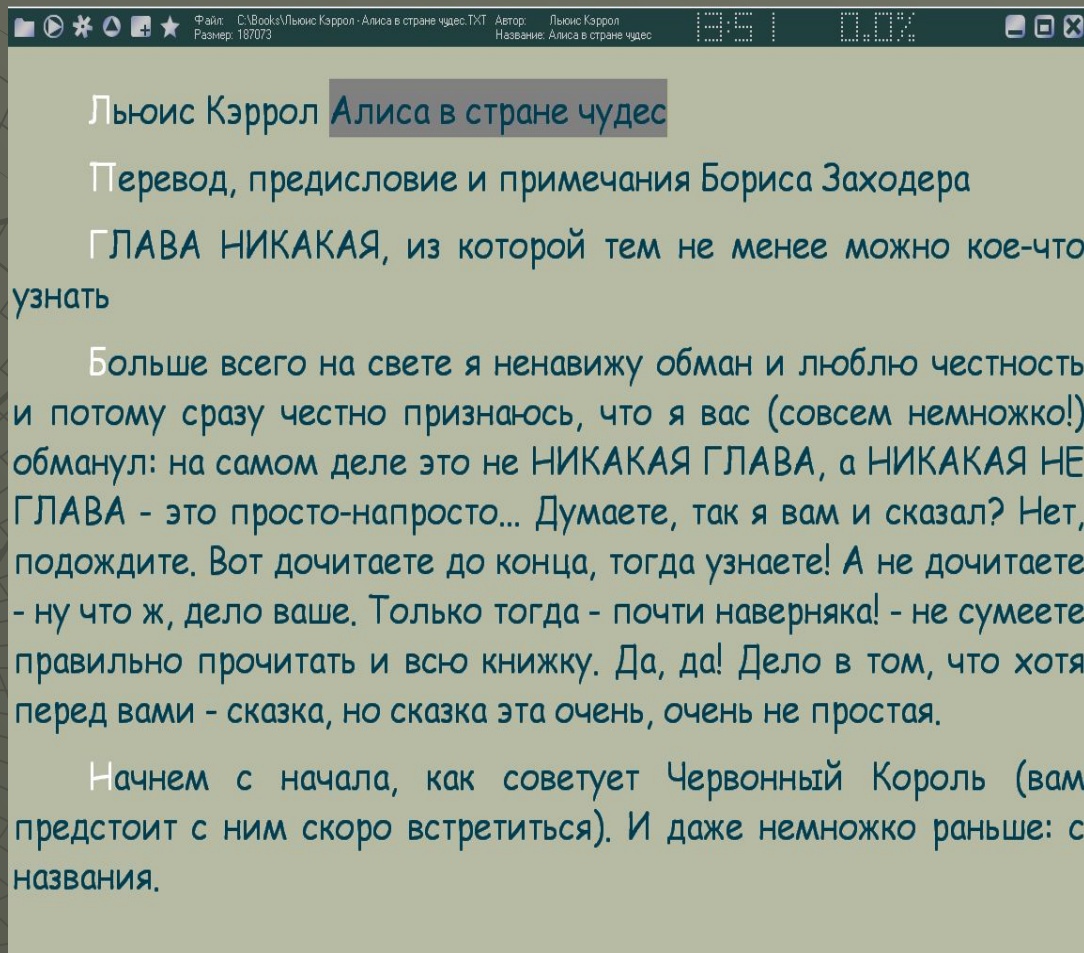


- Мы видим страничку сайта полученную из сети
- Текстовая информация представлена в формате \*.html
- HTML – язык разметки информации на форме, то есть задаются поля ввода, управляющие кнопки, цвет фона, картинки и т.д.
- Вёрстка Web – сайта это отдельная профессия – Web - дизайнер

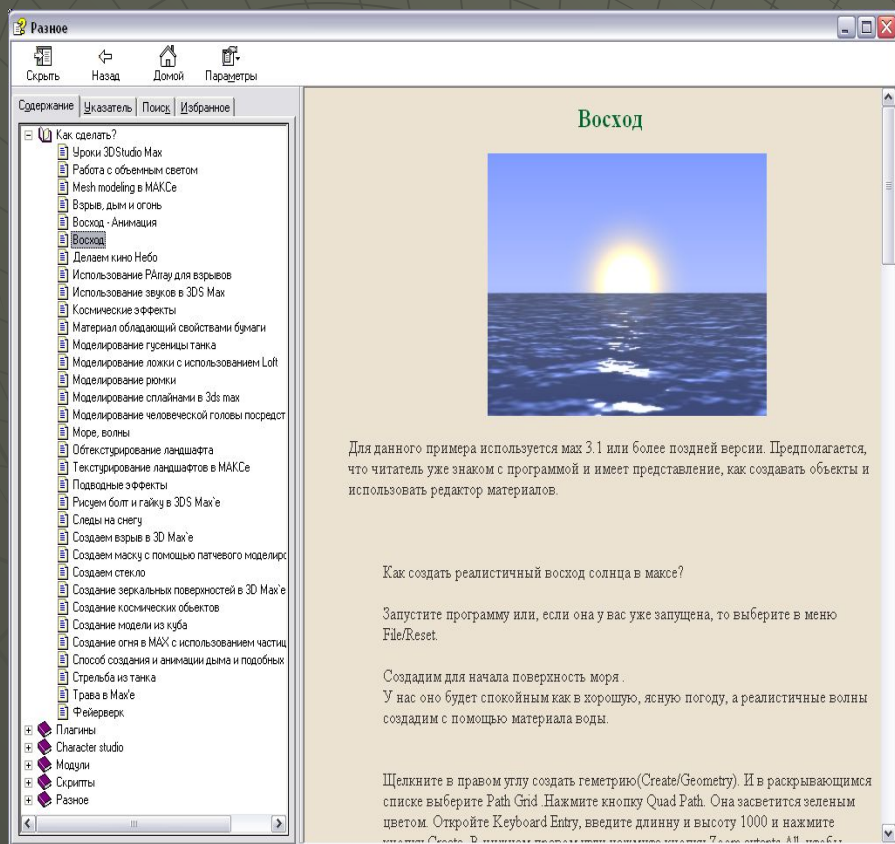


# Программы облегчающие чтение

- Здесь мы видим программу ICE Book Reader
- Может читать форматы \*.lit, \*.pdb, \*.prc но может так же **импортировать** текст из файлов формата \*.txt, из буфера обмена
- Удобство этой программы – текст в режиме **скроллинга** (медленно движется по экрану в верх), что позволяет читать книгу не прикасаясь к компьютеру и с далёкого расстояния.
- Открою небольшую тайну – есть программы, которые сами читают текст **синтезированным** (компьютерным) голосом



# Файлы помощи (help)



- ◆ Имеют собственное расширение (\*.chm)
- ◆ Разветвлённая структура проводника позволяет быстро обнаружить необходимый материал
- ◆ Текст можно копировать через буфер обмена
- ◆ Редактировать такой файл можно только через специализированные программы

# Компьютерные сверх технологии

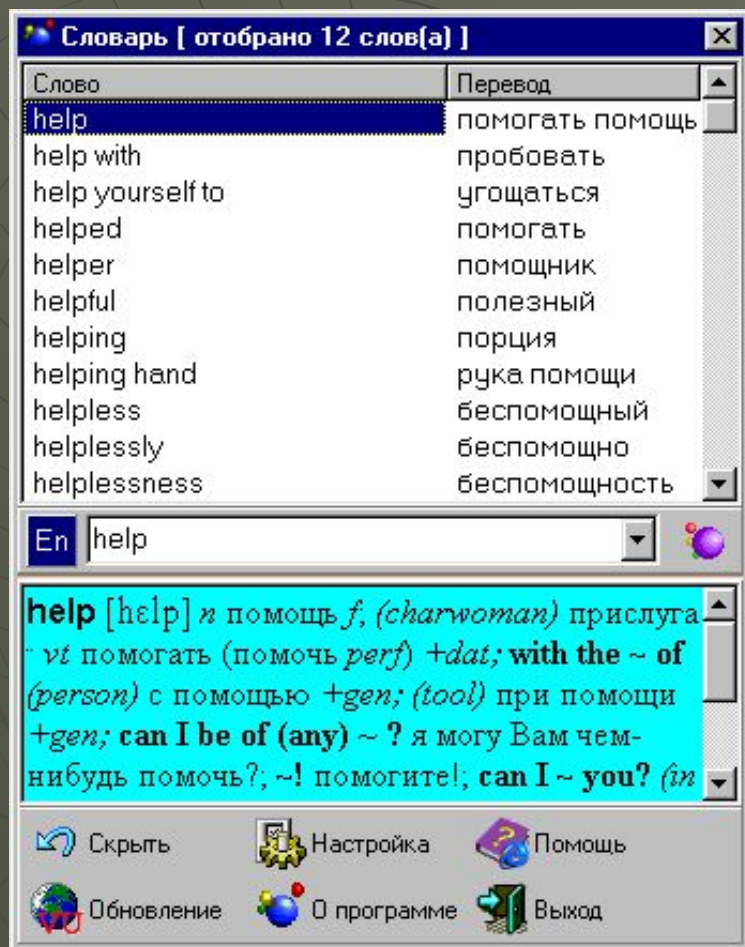
Программа AceReaderPro ставит перед собой невероятную задачу – научить человека сверхбыстрому чтению! AceReader делает следующее: быстро показывает на экране слова или фразы из читаемого текста. Зачем? Любой читатель, знакомый с основами быстрого чтения, знает, что главный враг скорости – проговаривание читаемого текста про себя (есть такое умное слово – субвокализация).

Вторая основная проблема при чтении – взгляд время от времени возвращается назад, и часть текста при этом прочитывается несколько раз подряд. AceReader как раз и позволяет с этим бороться. Собственно, работает он в двух режимах – может показывать слова в центре экрана или последовательно, так, как они были бы расположены на странице. В первом случае просто читаешь быстрее, поскольку не приходится переключать внимание и следить за текстом на странице. Второй метод отбивает привычку возвращаться к прочитанному. А регулятор скорости прокрутки текста (до 2000 слов в секунду) очень быстро отучивает от проговаривания слов. Однако AceReader – это не только инструмент для быстрого чтения, это, в первую очередь, обучающая программа.





# Программы переводчики



Основная задача этих программ – перевод текста с одного языка на другой (русско-английский)

Здесь можно выделить следующие программы: Lingvo, Pragma

Pragma - это многоязычная программа машинного перевода для перевода текстовых документов с одного языка на другой. Программа поддерживает семь языков: английский, русский, украинский, немецкий, латышский, польский и французский.

С помощью программы можно переводить текстовую информацию, представленную в виде документов MS Word, Интернет страниц, почтовых сообщений, справок, а также содержимое различных текстовых окон. К примеру, для трех языков - английский, русский и украинский существует 6 направлений перевода: англо-русский, русско-английский, англо-украинский, украино-английский, русско-украинский, украино-русский.

# Словарь

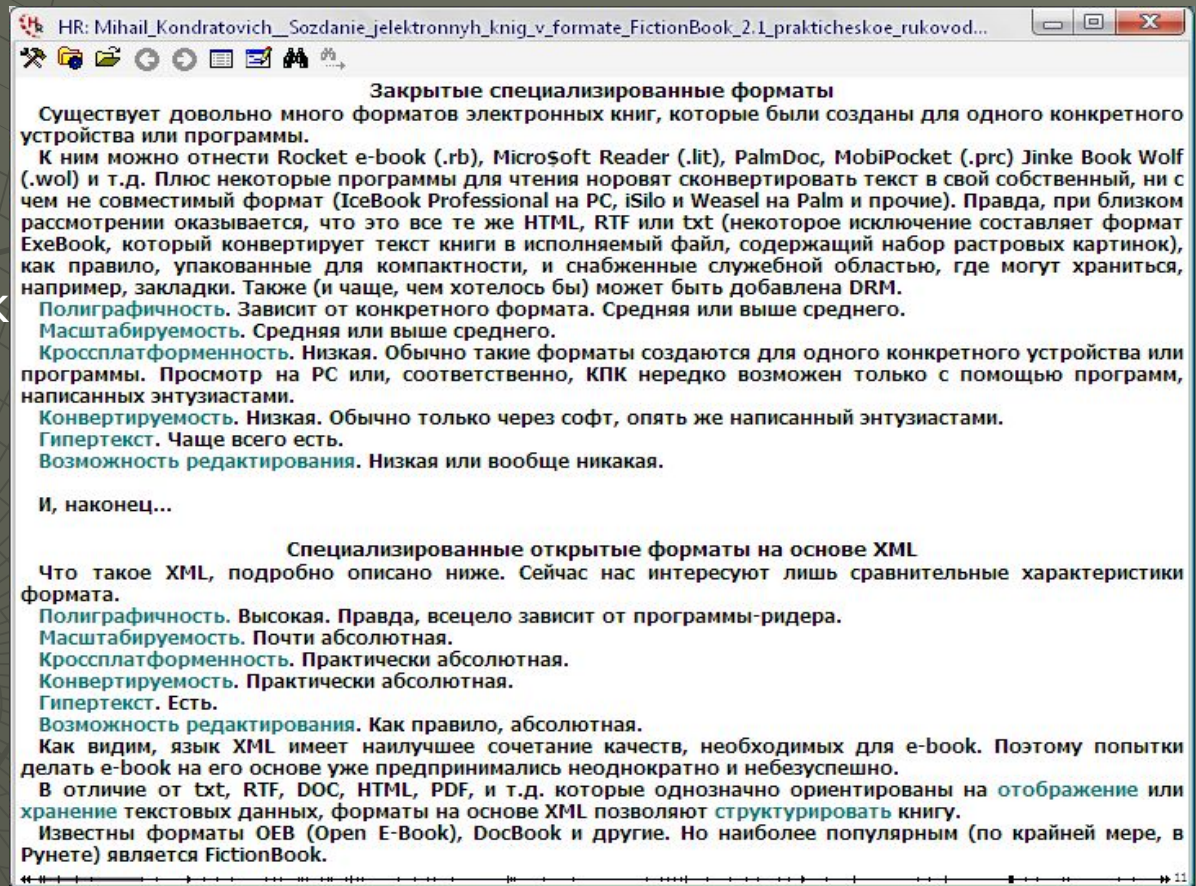
- ◆ Клинопись – текст нанесённый на глиняную табличку с помощью специальных палочек (древний Вавилон)
- ◆ Чайник – человек только начавший своё знакомство с компьютером
- ◆ Readme – прочитай меня(англ.), текстовый файл в котором зачастую содержится важная информация, например по установке программы на компьютер и метод её взлома
- ◆ Интегрированный – встроенный
- ◆ Текст заданный по маске – Например поиск – Доброе слово и скотине приятно  
выглядит так -\* слово ? скотине \*
- ◆ «весит» - жаргон. Речь идёт об объёме файла. Например файл имел объём 8Кб, а после сжатия архиватором стал «весить» 5Кб
- ◆ Импортировать – завезти(англ.), каким либо методом взять информацию из одной программы и вставить в другую
- ◆ Скроллинг – с правого боку некоторых окон имеется полоса прокрутки – скроллинга
- ◆ Синтез – создание чего либо искусственно, научными методами



# Книги формата \*.fp2

- ◆ Такие файлы открываются программой
- ◆ Naali Reader
- ◆ Общее название формата Fiction Book
- ◆ Полностью базируется на XML

Поддерживает  
структурирование и  
оформление







Спасибо за внимание!