

# Сканирование



# Содержание

- Введение
- Типы сканеров
- Устройство сканера
- Параметры сканеров
- Как сканировать
- Обработка изображений
- Распознавание текста. FineReader
- Глоссарий

# Введение

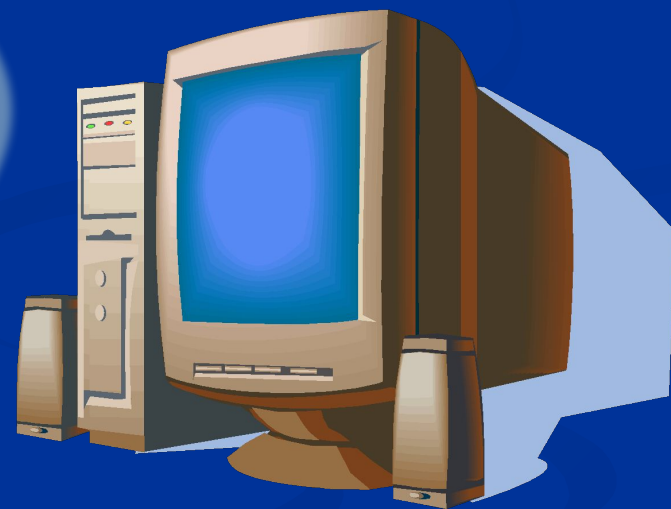
Каждый пользователь постоянно сталкивается с проблемой преобразования документов из бумажной формы в электронную



Конечно, можно набрать текст на клавиатуре вручную, НО ...



**Кроме того, вручную нельзя  
вводить изображения**

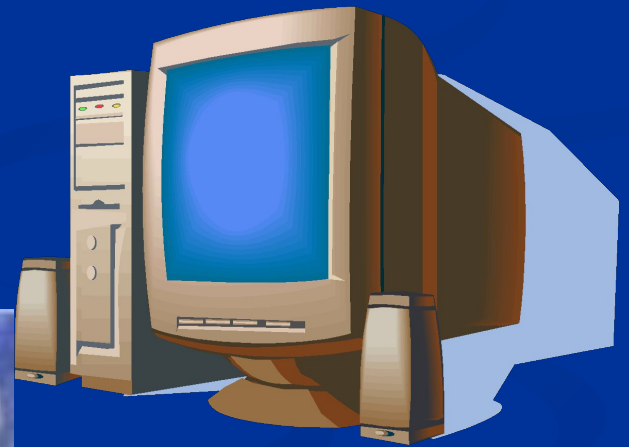


**Выходом из положения является**

***сканер***



**СКАНЕР — устройство, позволяющее  
вводить в компьютер рисунки,  
слайды, фотографии или другую  
графическую информацию**



# *ТИПЫ СКАНЕРОВ*

- 1. Ручные**
- 2. Настольные**
- 3. Барабанные**
- 4. Другие**  
(например, предназначенные для сканирования книг, в форме пирамиды)

# Ручные

Применяются в основном  
в комплекте с ноутбуками,

Часто используются  
для считывания штрих-кода



Для сканирования надо плавно провести сканирующей головкой по изображению

## **ДОСТОИНСТВА**

### **1. Мобильность**

(Небольшие размеры,  
масса)

### **2. Низкая цена**

## **НЕДОСТАТКИ**

**1.** Узкая полоса сканирования (<10 см)

**2.** Низкое разрешение

**3.** Малая скорость работы

**4.** Двигать устройство нужно мягко,  
без резких движений, иначе возможны  
перекосы изображения



# *Настольные*

1. Планшетные
2. Рулонные
3. Проекционные



# 1. Планшетные

## Наиболее распространенные

Для сканирования открыть крышку, положить сканируемый лист на сканер изображением вниз, закрыть крышку



Дальнейшее управление сканированием осуществляется с помощью программы, поставляемой вместе со сканером

## ДОСТОИНСТВА

1. Позволяют работать с большими форматами (до А3) документов
2. Сканирование автоматическое

## НЕДОСТАТКИ

1. Дороже ручных сканеров
2. Большие габариты

## 2. Рулонные

- Часто используются при сканировании текста
- Бывают двухсторонние



**Отдельные листы документов протягиваются через сканер, при этом и осуществляется их сканирование.**

### **ДОСТОИНСТВО**

Автоматическая подача страниц

### **НЕДОСТАТКИ**

1. Дорогие
2. Нельзя сканировать книги, журналы и т.п.

# 3. Проекционные

Напоминают проекционный аппарат или фотоувеличитель



Документ кладется на сканер изображением вверх, блок сканирования находится тоже сверху. Перемещается только сканирующее устройство

## ДОСТОИНСТВО

Возможность сканирования проекций трехмерных предметов

## НЕДОСТАТОК

Дорогие

# Барабанные

Применяются в крупных типографиях



В качестве сканирующего элемента используется зеркальный барабан или система таких барабанов

## ДОСТОИНСТВА

1. Гигантское разрешение (до 18000 dpi);
2. Высокая скорость.

## НЕДОСТАТКИ

**ОЧЕНЬ ДОРОГИЕ**

# Принцип работы



Свет отражается от объекта, через систему зеркал попадает на чувствительную матрицу (CCD — Couple-Charged Device), далее на АЦП и передается в компьютер.

# ПАРАМЕТРЫ СКАНЕРА

1. Разрешающая способность
2. «Цветность»
3. Динамический диапазон
4. Наличие Twain-интерфейса

# 1. Разрешающая способность

- **Разрешение** (важнейший параметр сканера) показывает, сколько пикселей на дюйм может «увидеть» сканер
- Разрешение измеряется в **DPI** (**Dots per Inch**, **пиксели на дюйм**)
- Чем больше разрешение, тем выше качество сканирования.





## 2. «Цветность»

### СКАНЕРЫ

#### ЧЕРНО-БЕЛЫЕ

#### ЦВЕТНОЙ

24 бит (8бит на  
канал цвета)

ПОЛУТОНОВЫЙ  
(серый) 16, 64  
или 256  
оттенков серого

ШТРИХОВОЙ  
(Черно-белый)



### 3. Динамический диапазон (плотность)

измеряется от 0 D (абсолютная прозрачность) до 4,0 D (абсолютно черная поверхность) по **логарифмической** шкале, т.е. 3,0 D темнее 2,0 D в 10 раз.

Для работы вполне достаточно значения 2,0. Однако для получения высококачественных результатов при сканировании слайдов вам понадобится сканер с плотностью около 3,2.

0 D

1 D

2 D

3 D

4 D

# 4. Наличие Twain-интерфейса

## Twain

– это стандарт, введенный для получения возможности сканирования в любой программе, поддерживающей этот интерфейс:

- **CorelDRAW,**
- **ACDSee,**
- **ABBYY FineReader,**
- **Adobe PhotoShop и т.д.**

# СКАНИРОВАНИЕ

## Программа *FineReader*

*Что такое OCR*

- 0.1 Подготовка изображения
- 0.2 Подготовка к сканированию
- 0.5 Выбор параметров сканирования
  - 1<sup>а</sup>. Сканирование
  - 1<sup>б</sup>. Открытие файлов с изображениями
  - 2. Обработка изображения
  - 3. Выделение блоков.
  - 4. Распознавание.
  - 5. Проверка на ошибки.
  - 6. Сохранение

# Что такое OCR

**OCR (Optical Character Recognition)** –  
Оптическое **распознавание** символов  
– это преобразование отсканированного  
изображения в редактируемый текст.

## **OCR-системы:**

- **CuneiForm** фирмы *Cognitive Technologies*  
– бесплатная (freeware)
- **ABBYY FineReader**
- Microsoft Office **Document Scanning**

# 0.1 Подготовка изображения

**Для сканирования не подойдет  
любая картинка или текст**

**Изображения должны быть:**

- Хорошего качества
- Не мятые
- Без масляных и других пятен
- Четкими

## 0.2 Подготовка к сканированию

1. Включить сканер
2. Включить компьютер
3. Запустить программу FineReader через

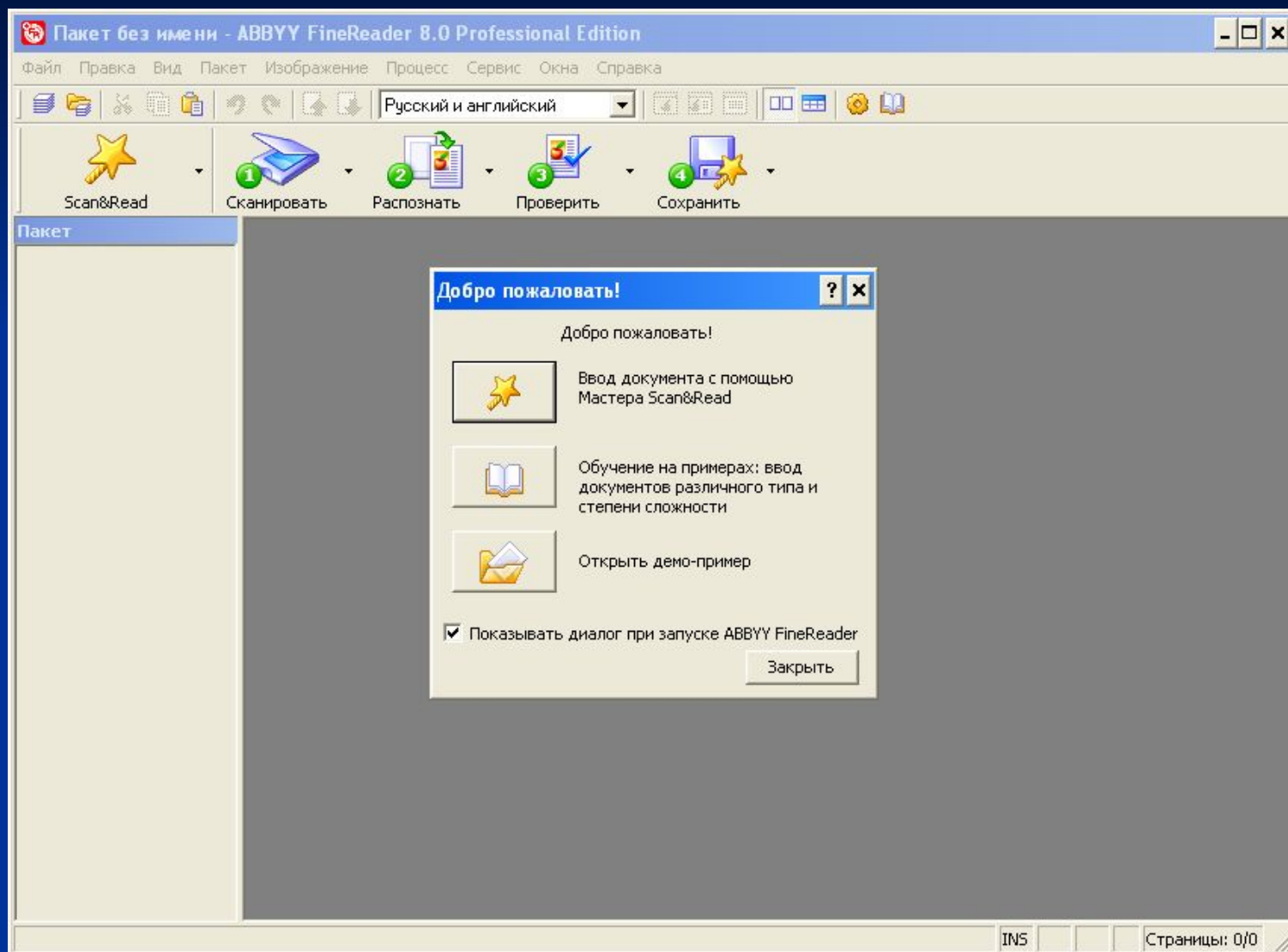
МЕНЮ:

*Пуск*

*Программы*

*ABBYY FineReader*

# Главное окно FineReader





# Интерфейс FineReader

Пакет без имени - ABBYY FineReader 8.0 Professional Edition - [1 - Изображение]

Файл Правка Вид Пакет Изображение Процесс Сервис Окна Справка

Русский и английский

Scan&Read Сканировать Распознать Проверить Сохранить

Пакет Изображение Текст

**ЗАГАДОЧНАЯ ВОДА И ЖИЗНЬ**

Многие люди считают воду источником жизни, а вода и чистая вода имеют свойства, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Вода — один из самых важных и загадочных веществ на Земле. Прежде всего известно, что вода имеет на своей поверхности уникальные свойства, а вода и чистая вода имеют свойства, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Дальнейшие исследования показали, что на молекулярном уровне вода имеет уникальные свойства, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Вода обладает уникальными свойствами, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Вода имеет уникальные свойства, которые делают ее уникальной. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Уникальные свойства и особая роль воды в биосфере Земли не могут быть описаны без знания ее структуры и свойств. Вода и чистая вода имеют свойства, которые делают ее уникальной.

Море и океаны	1,4 млрд. км <sup>3</sup>	Уникальные свойства и особая роль воды в биосфере Земли не могут быть описаны без знания ее структуры и свойств. Вода и чистая вода имеют свойства, которые делают ее уникальной.
Площадь	36 млн. км <sup>2</sup>	
Вода в озерах	2 млн. км <sup>3</sup>	
Атмосфера	14 тыс. км <sup>3</sup>	
Живые организмы	65%	

37%

100%

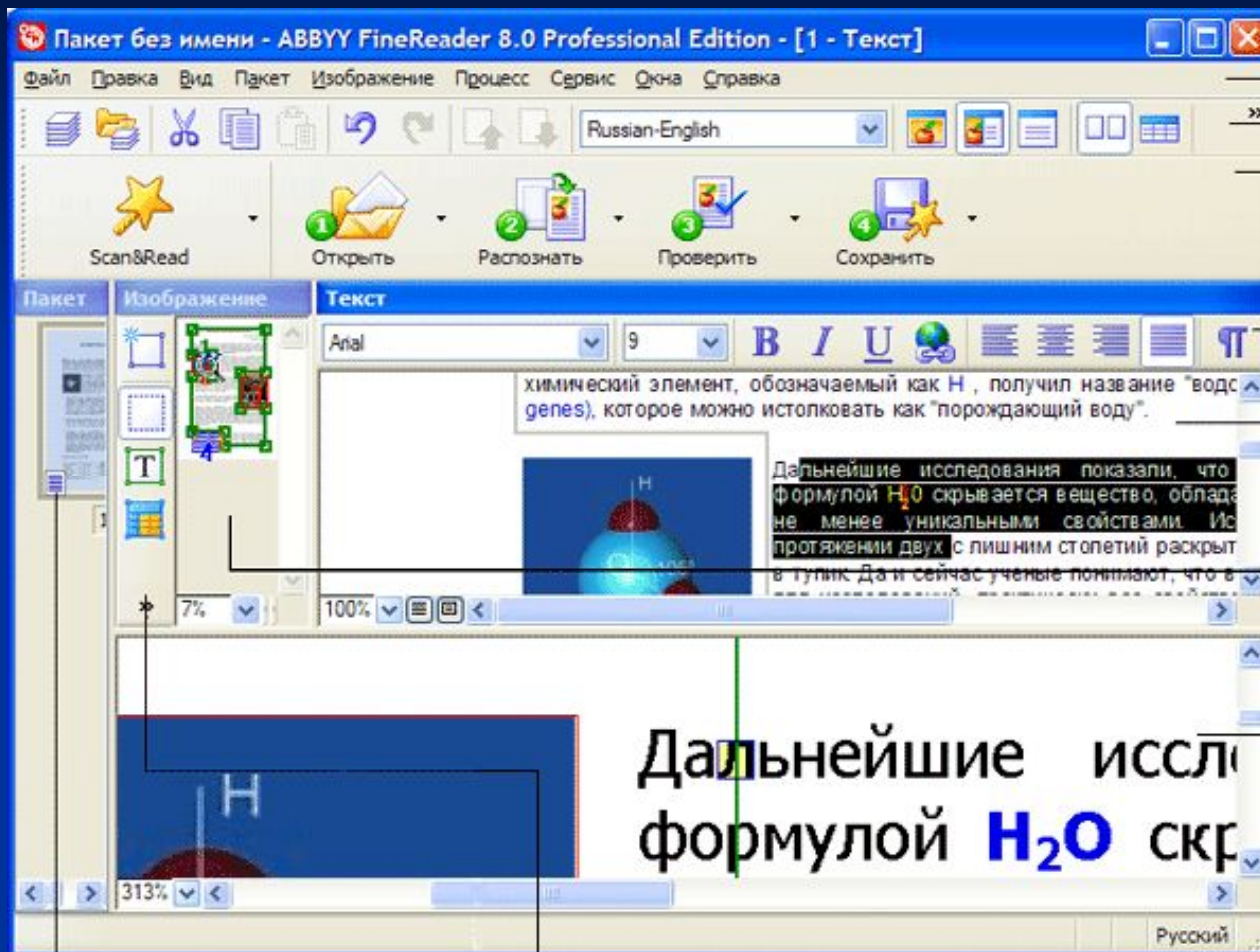
Страница не распознана

**ЗАГАДОЧНАЯ ВС**

313%

INS

# Основные панели FineReader



— Главное меню

»» Панель Стандартная

— Панель Scan&Read

позволяет произвести полную обработку текста: отсканировать, распознать, проверить и сохранить распознанный текст

— Панель Форматирование

— Окно Текст

в этом окне показывается распознанный текст для его проверки и редактирования

— Окно Изображение

в этом окне показывается отсканированное изображение, а также производится выделение блоков на изображении

— Окно Крупный план

в этом окне показывается увеличенное изображение редактируемой строки или обрабатываемого участка изображения

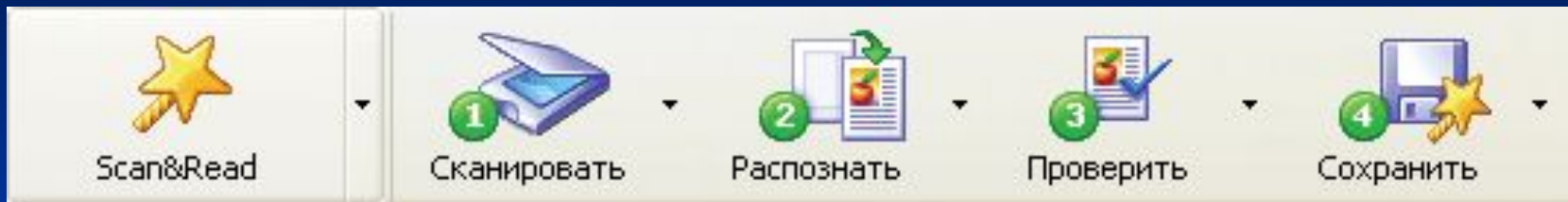
Окно Пакет

отображает страницы, входящие в открытый пакет; возможны 2 режима отображения страниц: Пиктограммы (см.рис.) или Вид со свойствами

Панель Изображение

позволяет произвести анализ макета страницы: выделить и отредактировать блоки

# Главная панель Scan&Read



Scan&Read

**Мастер Scan&Read...**  
Сканировать и распознать  
Открыть и распознать

PDF/изображения в Word  
Сканировать в PDF  
Сканировать в Word

Менеджер сценариев...

Сканировать

Открыть изображение...  
**Сканировать изображение...**  
Опции...

Распознать

**Распознать**  
Распознать все  
Опции...

Проверить

**Проверка...**  
Опции...

Сохранить

**Мастер сохранения результатов...**  
Сохранить страницы... Ctrl+S  
Передать страницы в  
Передать все страницы в  
Опции...

# Панель Изображение

Группа "Создание новых блоков"

- Анализ макета страницы
- Выделить зону распознавания
- Выделить текстовый блок
- Выделить табличный блок
- Выделить картинку
- Выбор объектов

Группа "Редактирование формы и положения блоков"

- Добавить часть к блоку
- Удалить часть блока
- Удалить блоки

Группа "Редактирование табличных блоков"

- Добавить вертикаль
- Добавить горизонталь
- Удалить линии

Группа "Работа с изображением"

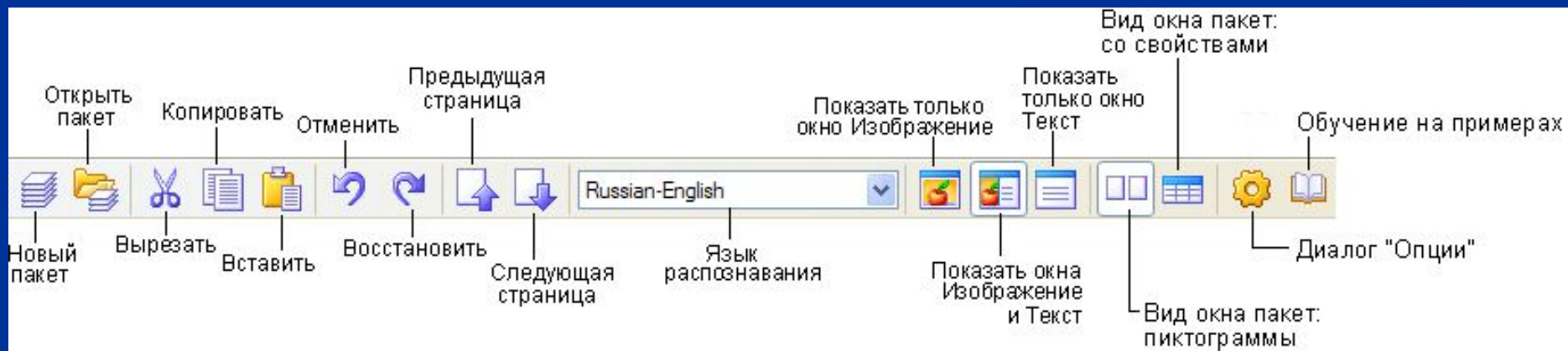
- Обрезать изображение
- Ластик
- Устранить искажение строк
- Повернуть по часовой
- Повернуть против часовой

**Содержит:**

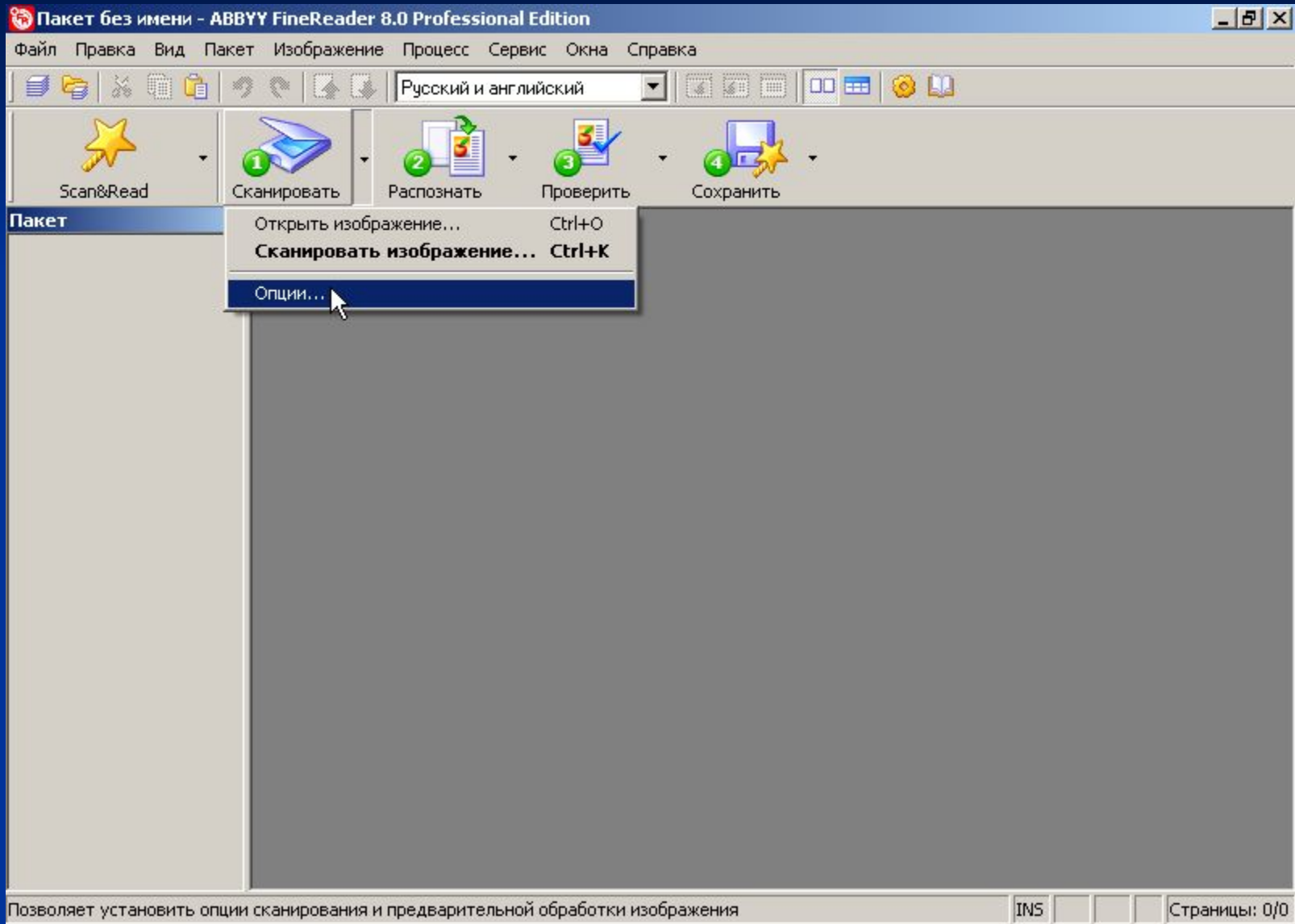
- кнопки для выделения и редактирования блоков,
- кнопки для редактирования изображения.

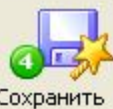
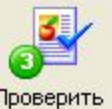
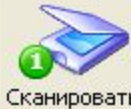
# На панели СТАНДАРТНАЯ находятся кнопки

- для управления работой с файлами и изображением:
- отмена и повтор действия
- перемещение по страницам пакета
- очистка и поворот изображения
- а также список языков распознавания.



# 0.5 Выбор параметров сканирования





Пакет

Опции

Общие Вид 1. Сканировать/Открыть 2. Распознать 3. Проверить 4. Сохранить

Сканер

TWAIN-драйвер сканера:

HP PrecisionScan LT 3.0

Выбрать драйвер...

Использовать интерфейс ABBYY FineReader

Настройки сканера...

Запрашивать опции перед началом сканирования

Использовать интерфейс TWAIN-драйвера сканера

Многостраничное сканирование

Сканировать несколько страниц

Обработка изображений

Устранить искажение строк

Делить разворот книги

Определять ориентацию страницы (при распознавании)

Запрашивать номер страницы перед добавлением в пакет

Закрыть

Отмена

Справка

# Диалог «Настройки сканера»

**Настройки сканера**

Ориентация изображения

Книжная  
 Альбомная

Подбор яркости

Автоматический  
 Средствами сканера  
 Ручной 50

Единицы измерения: Миллиметры

Размер бумаги: A4

Отступы

Сверху: 0  
Слева: 0

Размер

Высота: 297  
Ширина: 210

Режим сканирования картинок: Черно-белое изображение

Разрешение: 600 dpi

Загружать страницы из автоподатчика бумаги (ADF)

Пауза между страницами 5 секунд

Останавливаться между страницами

Показывать этот диалог перед сканированием

OK Отмена

**В диалоге Настройки сканера установите нужные опции**



# Тип изображения

**ЧЕРНО-БЕЛЫЙ (2 цвета)** – для сканирования графиков или текста.



**СЕРЫЙ (256 цветов)** – для сканирования нецветных или предназначенных для печати на черно-белом принтере рисунков. В этом случае осуществляется автоматический подбор яркости.



**ЦВЕТНОЙ (24 млн. цветов)** – для сканирования фотографий, цветных рисунков и т.п.



# Разрешение

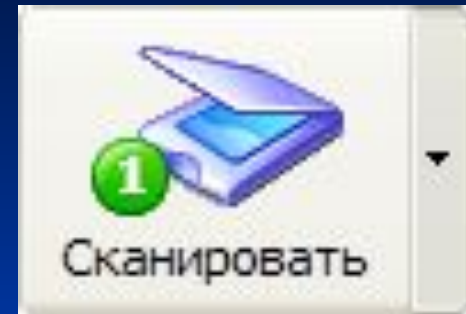
Используйте **300 dpi** для обычных текстов  
(размер шрифта 10 и более пунктов)

**400-600 dpi** для текстов, набранных мелким  
шрифтом (9 и менее пунктов)

Назначение	Разрешение
Для изображений, предназначенных к использованию только на экране компьютера	72 (экранное)
Распознавание текста	75 – 300
Если картинка подлежит увеличению	600 - 1200
Сканирование слайдов и	1200 и выше

# 1а. Сканирование

ДЛЯ ЭТОГО НАЖМИТЕ КНОПКУ



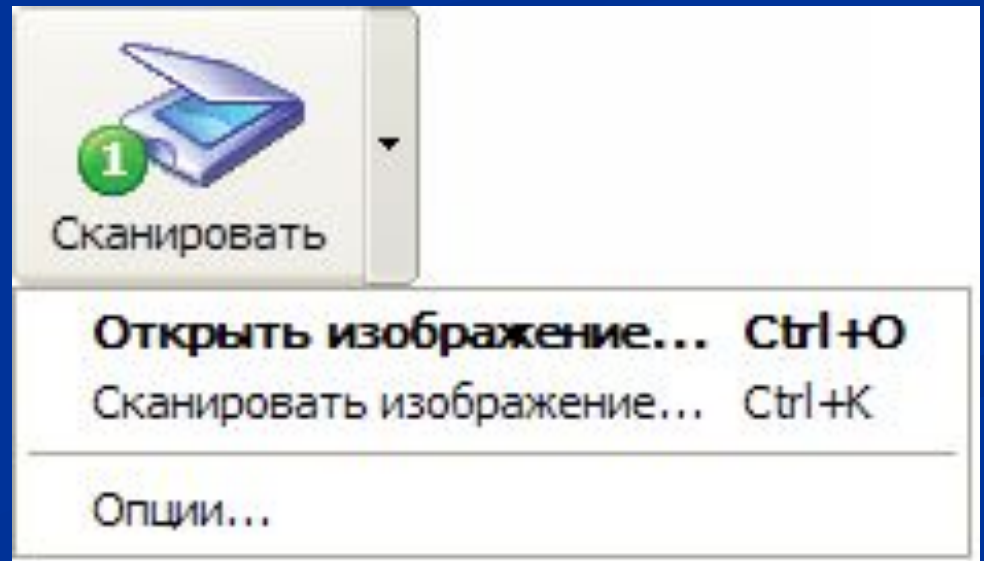
ХОД ПРОЦЕССА СКАНИРОВАНИЯ



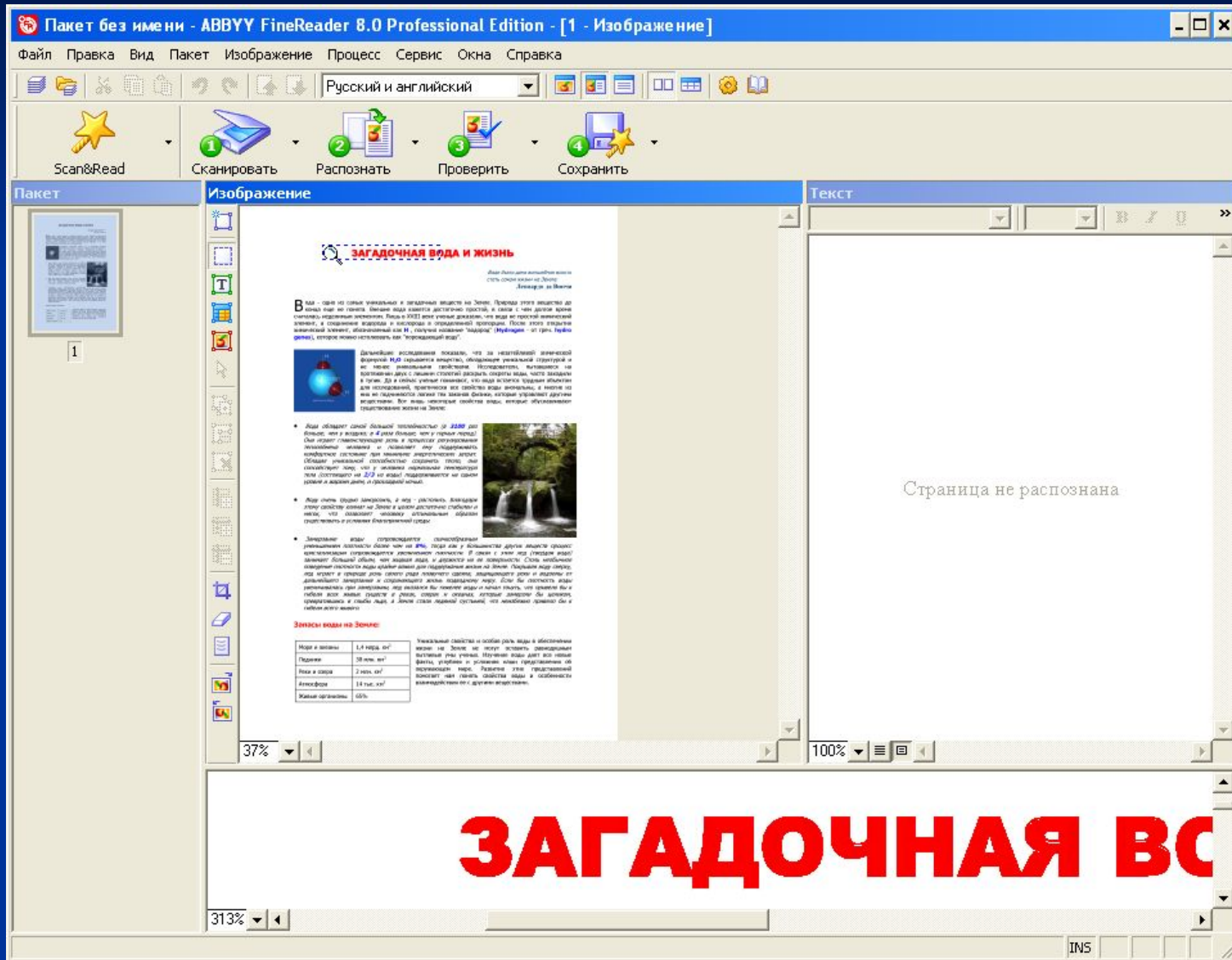
# 16. Открытие файлов с изображениями

Если у Вас нет сканера,  
Вы можете распознавать графические файлы  
ЧТОБЫ ОТКРЫТЬ ИЗОБРАЖЕНИЕ

Нажмите стрелку  
справа от кнопки  
**1-Сканировать**  
и выберите пункт  
**Открыть**  
**изображение**



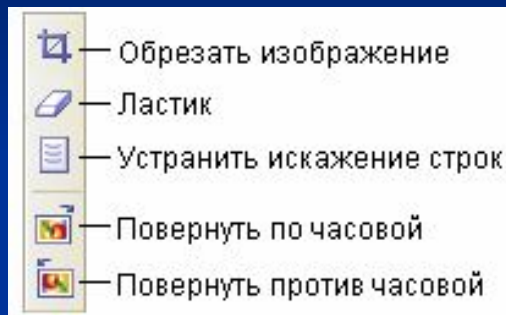
# Спустя некоторое время в Главном окне программы FineReader появится окно изображение с "фотографией" сканируемого листа



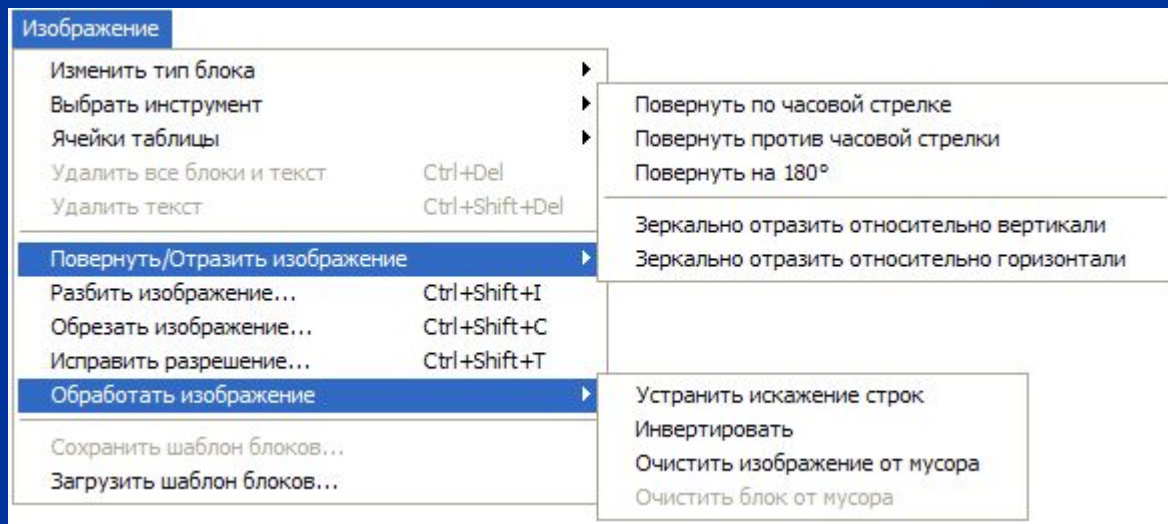
# ЗАГАДОЧНАЯ ВС

# Обработка изображения

1. С помощью кнопок на панели «Изображение».



2. С помощью меню Изображение



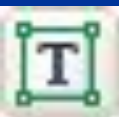
# 3. Выделение блоков

Для корректного распознавания требуется указать программе участки изображения, требующие распознавания.

Для этого проводится **анализ макета страницы**, во время которого выделяются **блоки** с *текстом, картинками и таблицы*.

**БЛОКИ** — участки изображения, позволяющие указать системе, какие части отсканированной страницы надо распознавать, и в каком порядке.

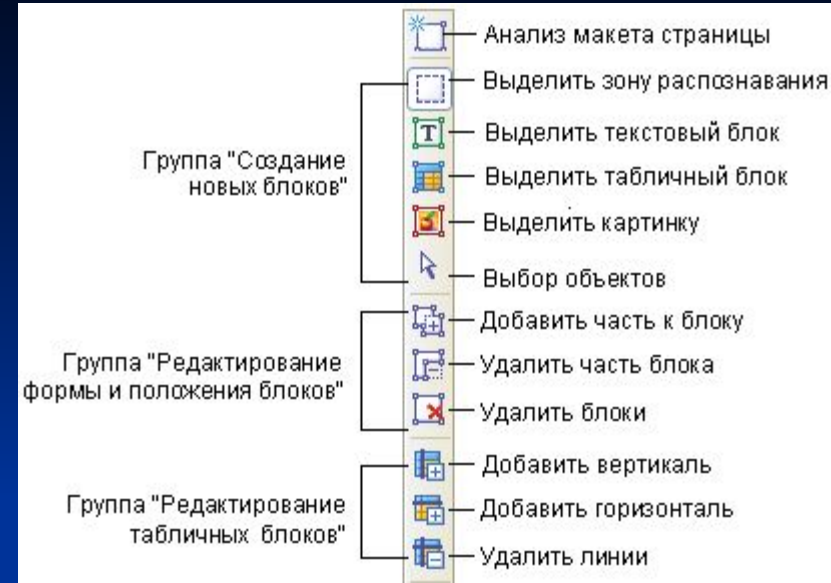
## Виды блоков

 **ТЕКСТ** — для обозначения текста

 **ТАБЛИЦА** — для обозначения таблиц или текста, имеющего табличную структуру

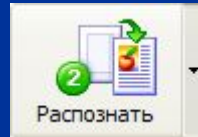
 **КАРТИНКА** — передаётся в исходном виде, в качестве картинки

# Ручное выделение блоков — с помощью кнопок на панели Изображение

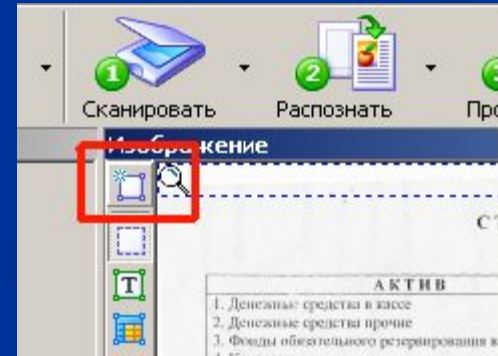


## Автоматическое:

1. При распознавании



2. «Анализ макета страницы»  
на панели Изображение



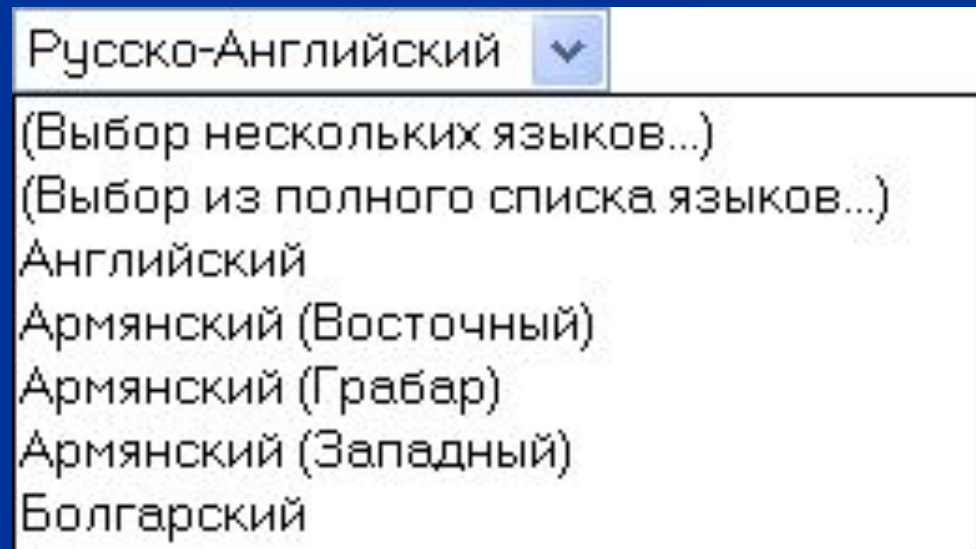


# 4. Распознавание

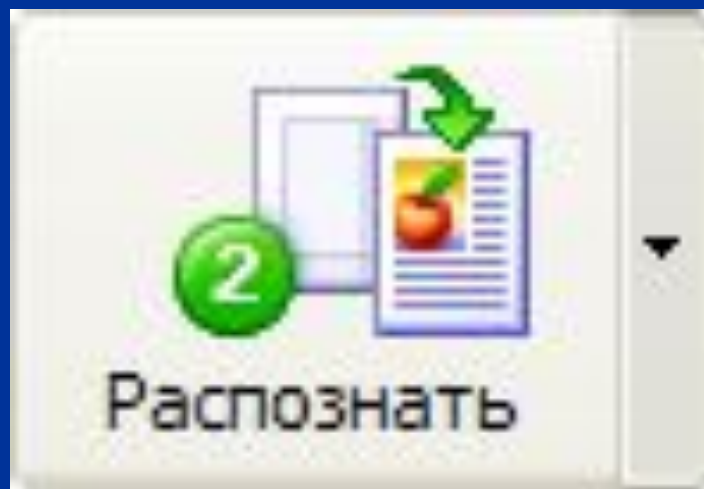
**РАСПОЗНАВАНИЕ** - это преобразование отсканированного изображения в текст

Перед распознаванием, необходимо установить  
**Язык распознавания**

Чтобы указать язык распознавания, выберите соответствующую строку в списке на панели **Распознавание**



**ЧТОБЫ ЗАПУСТИТЬ РАСПОЗНАВАНИЕ,  
нажмите на кнопку 2-Распознать,  
которая запускает распознавание  
открытого изображения**



# После завершения распознавания результат появляется в окне Текст, в котором можно проверить и отредактировать распознанный текст

Пакет без имени - ABBYY FineReader 8.0 Professional Edition - [1 - Текст]

Файл Правка Вид Пакет Изображение Процесс Сервис Окна Справка

Русский и английский

Scan&Read Сканировать Распознать Проверить Сохранить

Пакет Изображение Текст

Аrial 16 B I U x<sup>2</sup> x<sub>2</sub>

### ЗАГАДОЧНАЯ ВОДА И ЖИЗНЬ

Воды была одна из загадочных веществ, обладающих уникальной структурой и не менее уникальными свойствами. Последователи, пытавшиеся на протяжении двух с лишним столетий раскрыть тайны воды, часто заканчивали в тупик. Да и сейчас ученые спорят, что вода является основой жизни на Земле, а следовательно и основой человеческого существования. После этого открытия ученые обнаружили, что вода имеет необычные свойства, которые позволяют другим веществам. Вот лишь некоторые из них.

Дальнейшие исследования показали, что за необычной химической формулой H<sub>2</sub>O скрывается вещество, обладающее уникальной структурой и не менее уникальными свойствами. Последователи, пытавшиеся на протяжении двух с лишним столетий раскрыть тайны воды, часто заканчивали в тупик. Да и сейчас ученые спорят, что вода является основой жизни на Земле, а следовательно и основой человеческого существования. После этого открытия ученые обнаружили, что вода имеет необычные свойства, которые позволяют другим веществам. Вот лишь некоторые из них.

Воду очень трудно заморозить и растопить. Всплывание льдины в воде позволяет человеку оптимальным образом существовать в условиях безледяной среды.

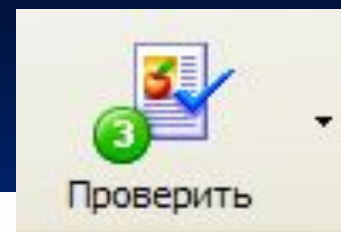
История воды на Земле:	
Миллионы лет	1,4 млрд. лет
Возникновение жизни	380 млн. лет
Появление человека	2 млн. лет
Изобретение	14 тыс. лет
Индустриальная революция	1850

37% 58% 313%

INS Русский

# 5. Проверка на ошибки

Осуществляется кнопкой 3-Проверить

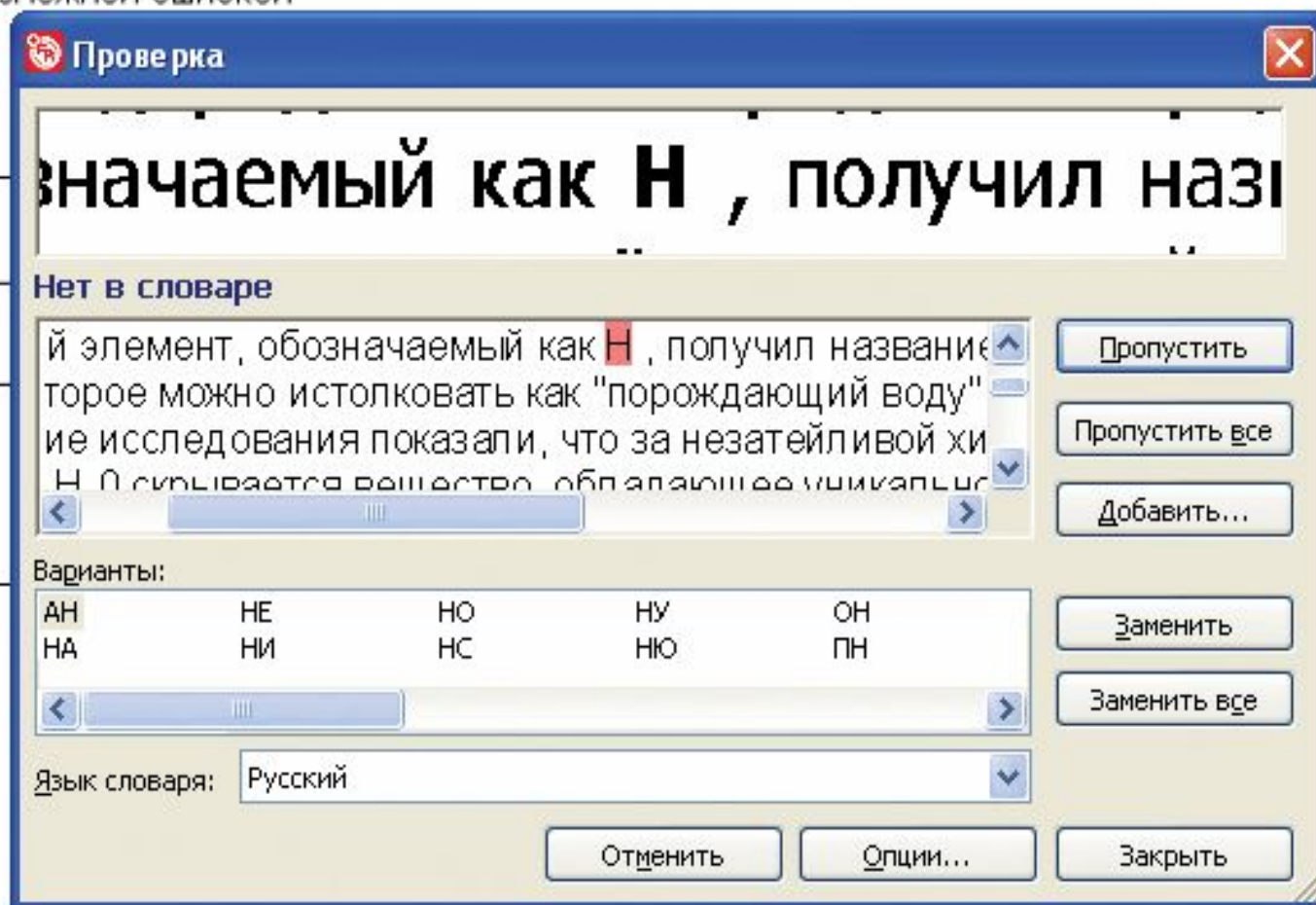


Аналог окна Крупный план, в нем показано изображение слова с возможной ошибкой

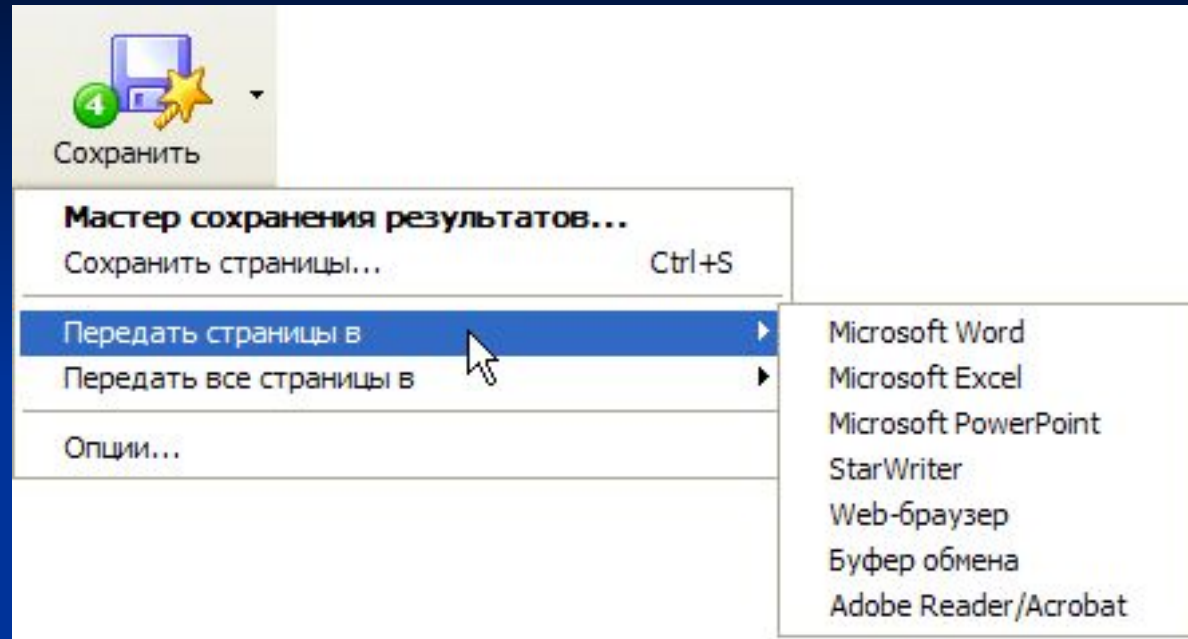
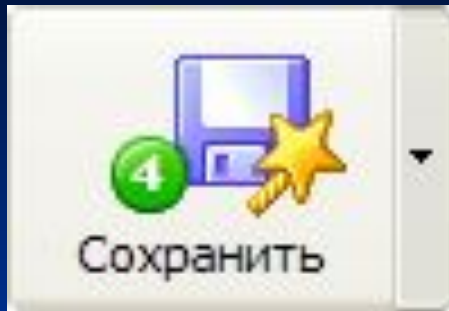
Название типа ошибки

Распознанный текст

Варианты замены выделенного слова



# 6. Сохранение



**Позволяет результаты распознавания:**

- 1. Сохранить в файл.**
- 2. Передать в выбранное приложение (Word, Excel, браузер).**
- 3. Скопировать в буфер обмена.**

# Настройка передачи в Word (или файл DOC/RTF)

