

Кодирование текста. Программы для обработки текста



Тексты на внешних носителях сохраняются в файлах.

Какие есть преимущества сохранения текстов в файлах на компьютерных носителях по сравнению с бумагой?

When I see the seven it reminds Me that
on the seventh day God rested from the
great work he had made and hallowed it
When I see the eight it reminds Me of the
Eight righteous persons that were saved
when God destroyed the world by Noah
and his wife his three sons and their wives
Then I see the nine it reminds Me of the
Nine Lepers that were cleansed by our Saviour
there were nine out of the ten who returned
thanks. When I see the ten it reminds Me
of the Ten Commandments which God
handed down to Moses on the table of stone
When I see the thing it reminds Me of the



Какие есть преимущества сохранения текстов в файлах на компьютерных носителях по сравнению с бумагой?

- ✓ **Компактное размещение**
- ✓ **Если текст становится ненужным, то с помощью компьютера его легко удалить с носителя**
- ✓ **Легко скопировать файлы в любом количестве**
- ✓ **Файл с текстом можно быстро переслать другому человеку по электронной почте**

Главное неудобство хранения текстов в файлах состоит в том, что прочитать их можно только с помощью компьютера.

Человек может просмотреть текст на экране монитора или напечатать на бумаге, используя принтер.



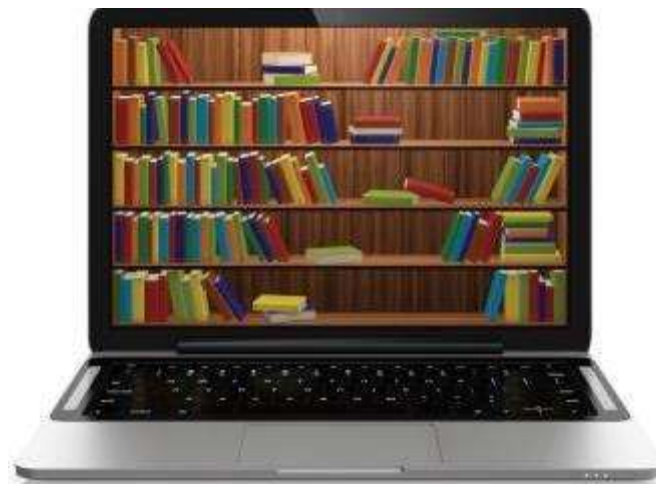
Уже сейчас существует издания, которые не печатаются на бумаге, а хранятся и распространяются в форме файлов.

С распространением компьютеров число таких безбумажных изданий с каждым годом увеличивается.



Представьте себе, что вся ваша личная библиотека разместится либо в коробке с дисками, либо на флэшке.

Причем по объему она не будет не меньше, чем сотни книг.



А экономя бумагу, мы сохраняем леса на нашей планете.



Как представляются тексты в памяти компьютера:

Текстовая информация состоит из символов: букв, цифр, знаков препинания, скобок и др.

Множество всех символов, с помощью которых записывается текст, называется алфавитом, а число символов в алфавите – его мощностью.

В компьютере используется алфавит мощностью

$$N = 256 \text{ символов}$$

Один символ такого алфавита несет 8 битов информации.

$$N = 2^b - \text{мощность алфавита}$$

$$V = k \cdot b - \text{информационный объем}$$

**Все символы компьютерного алфавита
пронумерованы от 0 до 255.**

**Каждому номеру соответствует восьмиразрядный
двоичный код:**

от 00000000 до 11111111

**Этот код – порядковый номер символа в двоичной
системе счисления.**

**Таблица, в которой всем символам компьютерного
алфавита поставлены в соответствие порядковые
номера, называется таблицей кодировки.**

Таблица стандартной части кода ASCII

32		00100000	56	8	00111000	80	P	01010000	104	h	01101000
33	!	00100001	57	9	00111001	81	Q	01010001	105	i	01101001
34	"	00100010	58	:	00111010	82	R	01010010	106	j	01101010
35	#	00100011	59	;	00111011	83	S	01010011	107	k	01101011
36	\$	00100100	60	<	00111100	84	T	01010100	108	l	01101100
37	%	00100101	61	=	00111101	85	U	01010101	109	m	01101101
38	&	00100110	62	>	00111110	86	V	01010110	110	n	01101110
39	'	00100111	63	?	00111111	87	W	01010111	111	o	01101111
40	(00101000	64	@	01000000	88	X	01011000	112	p	01110000
41)	00101001	65	A	01000001	89	Y	01011001	113	q	01110001
42	*	00101010	66	B	01000010	90	Z	01011010	114	r	01110010
43	+	00101011	67	C	01000011	91	[01011011	115	s	01110011
44	,	00101100	68	D	01000100	92	\	01011100	116	t	01110100
45	-	00101101	69	E	01000101	93]	01011101	117	u	01110101
46	.	00101110	70	F	01000110	94	^	01011110	118	v	01110110
47	/	00101111	71	G	01000111	95	_	01011111	119	w	01110111
48	0	00110000	72	H	01001000	96	`	01100000	120	x	01111000
49	1	00110001	73	I	01001001	97	a	01100001	121	y	01111001
50	2	00110010	74	J	01001010	98	b	01100010	122	z	01111010
51	3	00110011	75	K	01001011	99	c	01100011	123	{	01111011
52	4	00110100	76	L	01001100	100	d	01100100	124		01111100
53	5	00110101	77	M	01001101	101	e	01100101	125	}	01111101
54	6	00110110	78	N	01001110	102	f	01100110	126	~	01111110
55	7	00110111	79	O	01001111	103	g	01100111	127	.	01111111

Международным стандартом стала таблица кодировки под названием ASCII (ANSI) – американский стандартный код для обмена информацией.

Кодовая таблица 1251 (CP 1251) – содержит русский алфавит и используется в ОС Windows и ее приложениях.

Помимо восьмиразрядной кодировки символов все большее распространение получает шестнадцатиразрядная – двухбайтовая кодировка.

Международный стандарт такой кодировки носит название UNICODE.

Таблица 2. Кодовая страница CP1251

128	Ъ	10000000	160	"	10100000	192	А	11000000	224	а	11100000
129	Ѓ	10000001	161	Ў	10100001	193	Б	11000001	225	б	11100001
130	,	10000010	162	ў	10100010	194	В	11000010	226	в	11100010
131	ѓ	10000011	163	Ј	10100011	195	Г	11000011	227	г	11100011
132	"	10000100	164	џ	10100100	196	Д	11000100	228	д	11100100
133	...	10000101	165	ѓ	10100101	197	Е	11000101	229	е	11100101
134	†	10000110	166		10100110	198	Ж	11000110	230	ж	11100110
135	‡	10000111	167	§	10100111	199	З	11000111	231	з	11100111
136	€	10001000	168	Ё	10101000	200	И	11001000	232	и	11101000
137	‰	10001001	169	©	10101001	201	Й	11001001	233	й	11101001
138	Љ	10001010	170	Є	10101010	202	К	11001010	234	к	11101010
139	«	10001011	171	«	10101011	203	Л	11001011	235	л	11101011
140	Њ	10001100	172	¬	10101100	204	М	11001100	236	м	11101100
141	Ќ	10001101	173		10101101	205	Н	11001101	237	н	11101101
142	Ь	10001110	174	®	10101110	206	О	11001110	238	о	11101110
143	Њ	10001111	175	Ы	10101111	207	П	11001111	239	п	11101111
144	ђ	10010000	176	•	10110000	208	Р	11010000	240	р	11110000
145	‘	10010001	177	±	10110001	209	С	11010001	241	с	11110001
146	’	10010010	178	І	10110010	210	Т	11010010	242	т	11110010
147	“	10010011	179	Ѕ	10110011	211	У	11010011	243	у	11110011
148	”	10010100	180	Г	10110100	212	Ф	11010100	244	ф	11110100
149	••	10010101	181	μ	10110101	213	Х	11010101	245	х	11110101
150	—	10010110	182	¶	10110110	214	Ц	11010110	246	ц	11110110
151	—	10010111	183	•	10110111	215	Ч	11010111	247	ч	11110111
152	"	10011000	184	ё	10111000	216	Ш	11011000	248	ш	11111000
153	™	10011001	185	№	10111001	217	Щ	11011001	249	щ	11111001
154	Љ	10011010	186	є	10111010	218	Ъ	11011010	250	ъ	11111010
155	›	10011011	187	»	10111011	219	Ы	11011011	251	ы	11111011
156	Њ	10011100	188	Ј	10111100	220	Ь	11011100	252	ь	11111100
157	Ќ	10011101	189	Ѕ	10111101	221	Э	11011101	253	э	11111101
158	ђ	10011110	190	ѕ	10111110	222	Ю	11011110	254	ю	11111110

Тексты вводятся в память компьютера с помощью клавиатуры. На клавишах написаны привычные нам буквы, цифры, знаки препинания и другие символы.

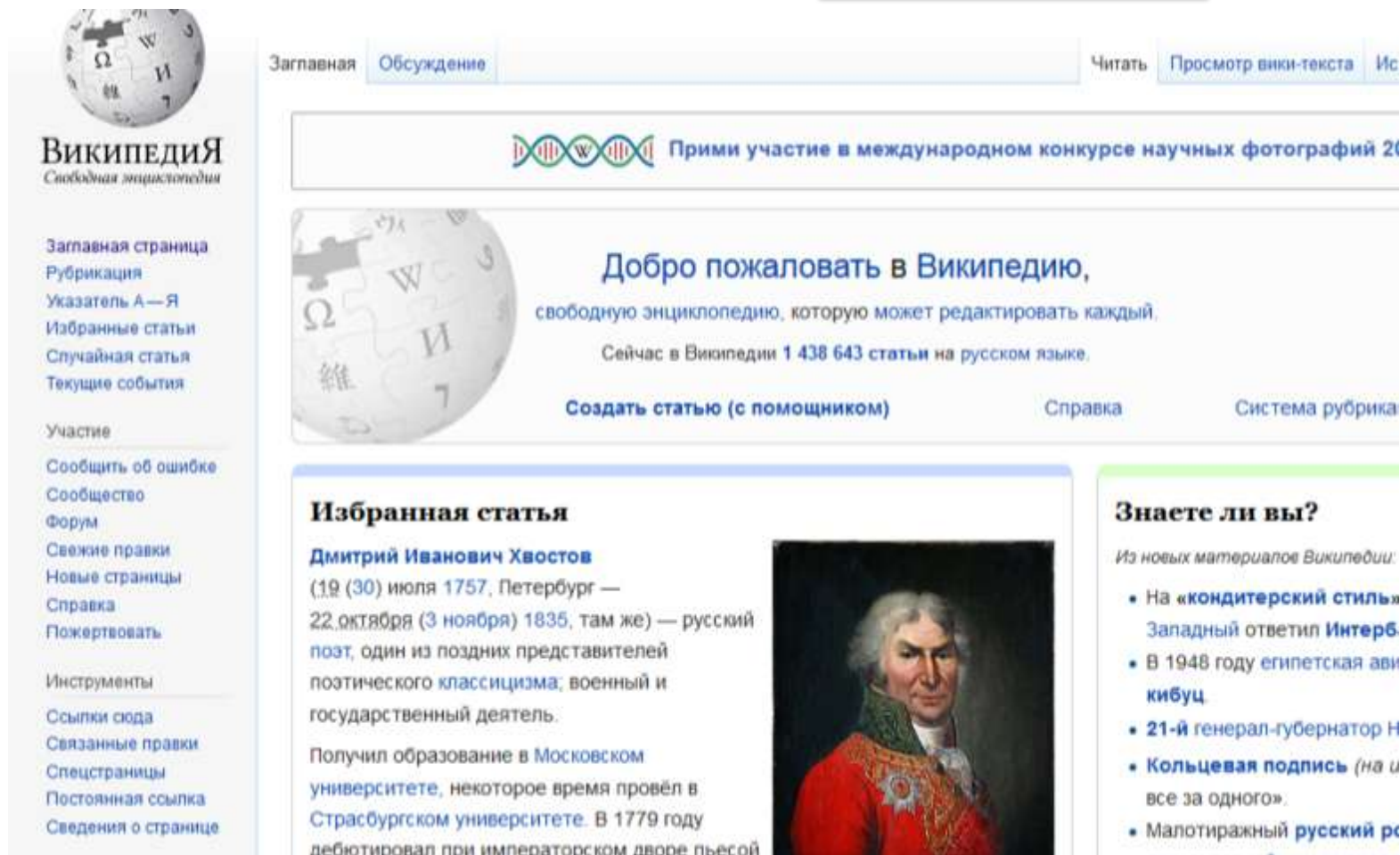
В оперативную память они попадают в форме двоичного кода.



Что такое гипертекст?

Гипертекст – это текст, организованный так, что его можно просматривать в последовательности смысловых связей между его отдельными фрагментами.

Такие связи называются гиперссылками.



The image is a screenshot of the Russian Wikipedia homepage. At the top left is the Wikipedia logo, a globe with letters in various scripts. Below it, the text 'Википедия' and 'Свободная энциклопедия' is displayed. A sidebar on the left contains a list of links: 'Заглавная страница', 'Рубрикация', 'Указатель А—Я', 'Избранные статьи', 'Случайная статья', 'Текущие события', 'Участие', 'Сообщить об ошибке', 'Сообщество', 'Форум', 'Свежие правки', 'Новые страницы', 'Справка', 'Пожертвовать', 'Инструменты', 'Ссылки сюда', 'Связанные правки', 'Спецстраницы', 'Постоянная ссылка', and 'Сведения о странице'. The main content area has a navigation bar with 'Заглавная', 'Обсуждение', 'Читая', 'Просмотр вики-текста', and 'Ис'. Below this is a banner for a scientific photography competition. The central part of the page features a large globe icon and the text 'Добро пожаловать в Википедию, свободную энциклопедию, которую может редактировать каждый. Сейчас в Википедии 1 438 643 статьи на русском языке.' Below this are links for 'Создать статью (с помощником)', 'Справка', and 'Система рубрика'. The 'Избранная статья' section features a portrait of Dmitry Ivanovich Khvostov and a brief biography. The 'Знаете ли вы?' section lists several interesting facts, including the 'confectionery style' and the 'Egyptian aviatrix'.

Википедия
Свободная энциклопедия

Заглавная страница
Рубрикация
Указатель А—Я
Избранные статьи
Случайная статья
Текущие события

Участие
Сообщить об ошибке
Сообщество
Форум
Свежие правки
Новые страницы
Справка
Пожертвовать

Инструменты
Ссылки сюда
Связанные правки
Спецстраницы
Постоянная ссылка
Сведения о странице

Заглавная Обсуждение Читая Просмотр вики-текста Ис

Прими участие в международном конкурсе научных фотографий 20

Добро пожаловать в Википедию,
свободную энциклопедию, которую может редактировать каждый.
Сейчас в Википедии 1 438 643 статьи на русском языке.

Создать статью (с помощником) Справка Система рубрика

Избранная статья

Дмитрий Иванович Хвостов
(19 (30) июля 1757, Петербург —
22 октября (3 ноября) 1835, там же) — русский
поэт, один из поздних представителей
поэтического классицизма; военный и
государственный деятель.

Получил образование в Московском
университете, некоторое время провёл в
Страсбургском университете. В 1779 году
дебютировал при императорском дворе пьесой

Знаете ли вы?

Из новых материалов Википедии:

- На «кондитерский стиль»
Западный ответил **Интерб**
- В 1948 году египетская ави
кибуц
- 21-й генерал-губернатор Н
- Кольцевая подпись (на и
все за одного).
- Малотиражный русский рс

Чаще всего по принципу гипертекста организованы компьютерные справочники, энциклопедии, учебники.

БИОЛОГИЯ Учебник Практикум Определитель Контроль Поиск Личная папка Помощь Выход

24 8 25

СИСТЕМАТИКА РАСТЕНИЙ И ЖИВОТНЫХ

В результате эволюции возникло огромное разнообразие видов. До сих пор точно неизвестно, сколько видов живых существ обитает на Земле. По приблизительным подсчетам, их более 2,5 млн.

Систематика — наука, приводящая в систему знания о разнообразии живых организмов.

ВЫ УЗНАЕТЕ:

- Для чего необходима систематика как наука
- Как называют виды

ВСПОМНИТЕ:

- Что такое вид?
- Какие царства живых организмов вам известны?

Своим первым учеником, пытаясь систематизировать сведения о растениях и животных, был Аристотель. Животные были классифицированы им по среде обитания: водные, сухопутные и обитающие в воздухе. Ученый Аристотеля Теодорас всю жизнь посвятил изучению растений и описал множество их видов. Его труды служили для определения растений не одно столетие.

Анимация

Файл

▶ Систематические группы растений

Царство	Растения		
Подцарство	Багрянки	Водоросли	Высшие растения
Отдел	Зеленые водоросли	Бурые водоросли	Золотистые водоросли
	Диатомовые водоросли	Зеленовые водоросли	

Подцарство Настоящие водоросли включает по разным оценкам 8–10 отделов, среди которых Золотистые, Диатомовые, Бурые, Зеленые, Зеленовые и другие. Настоящие водоросли обитают в морях, пресных водоемах, наземные представители встречаются редко. Растения с хлорофиллом а, в качестве дополнительного хлорофилла присутствуют хлорофиллы b (у зеленых и зеленовато-бурых водорослей) или c (у бурых и диатомовых водорослей). Одноклеточные, колониальные или многоклеточные формы. Тело многоклеточных водорослей (слоевые) не дифференцировано на ткани и органы (в частности отсутствуют покровные, проводящие и механические ткани).



2.5 Систематические группы

7-778) — натуралист, ученой класс-организмов.

ИЗУЧАЮЩИЕ ВОДОРАСЛИ

НАТУРАЛИСТОВ ДОМАШНИЙ

ВНЕШНЕЕ ЛАБОРАТОРИЯ-ТО ДВОЙНОЕ И ЭТОМ ВИ-

КОМУ РОДУ, КЛАССУ И ВИДУ

КАКИ РОДА

КАКИ:

КАКИ СИСТЕ-

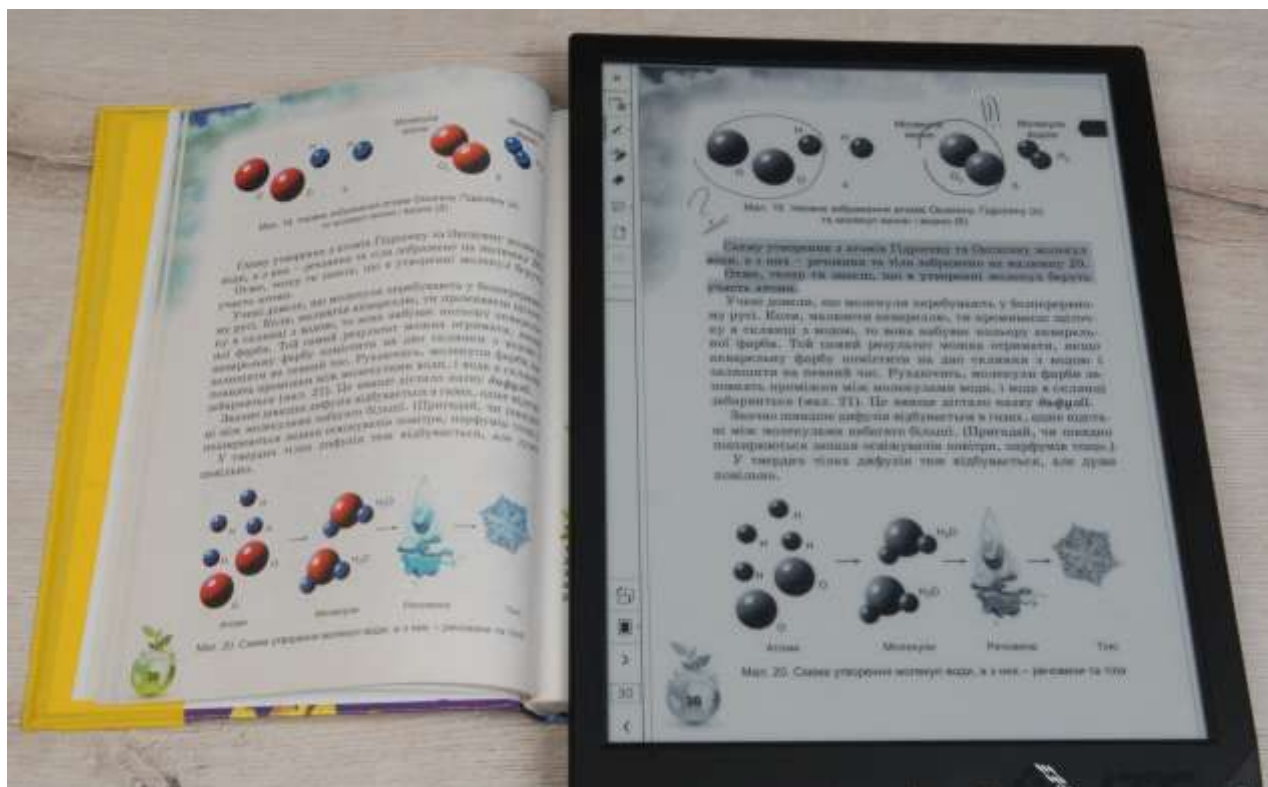
КАКИ ГРУППЫ

КАКИ СИСТЕМАТИКА ЖИВОТНЫХ?

ПО ОБЪЕДИНЯЕТ ДВОЙНОЕ НАЗВАНИЕ?

Такую «книгу» можно читать не только в обычном порядке, «листая страницы» на экране, но и перемещаться по смысловым связям в произвольном порядке.

После такой экскурсии вглубь материала ученик может вернуться в исходную точку, щелкнув мышью по кнопке «Назад», так как система запоминает весь маршрут продвижения по гиперссылкам.

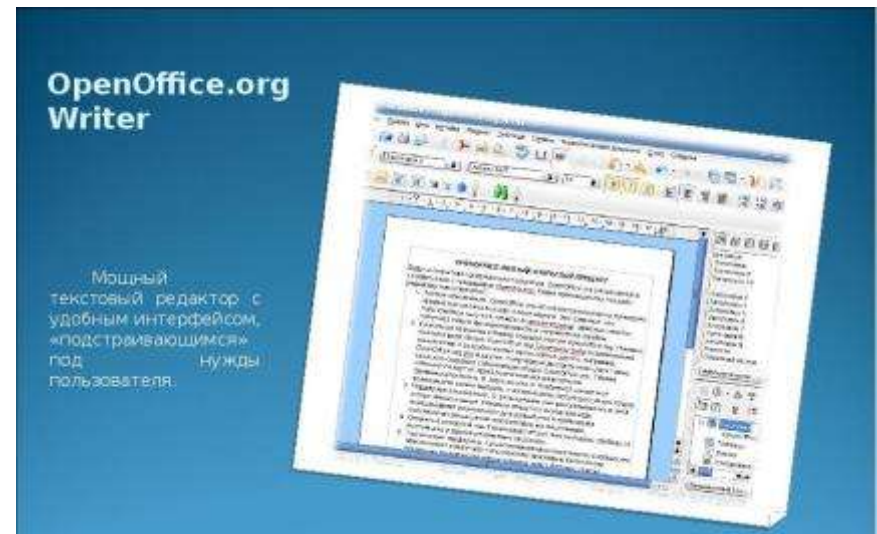
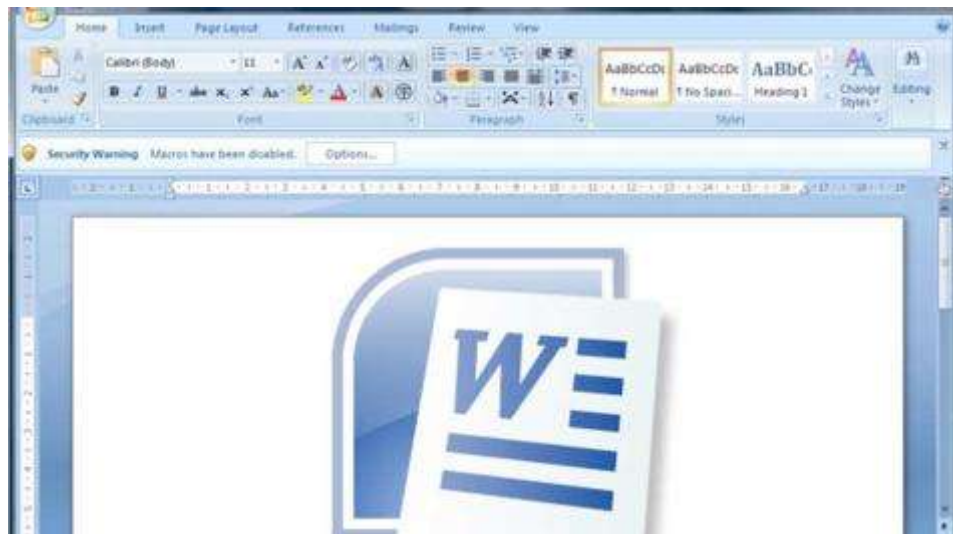


Текстовые редакторы и текстовые процессоры



Текстовый редактор – это простейшая программа для работы с текстом без оформления. (Блокнот)

Текстовый процессор - прикладная программа, которая позволяет форматировать части текста, добавлять рисунки, таблицы, списки и диаграммы.



Текстовые процессоры позволяют:

- **использовать различные шрифты, выделять фрагменты текста;**
- **создавать составные документы, включающие списки, рисунки, таблицы, диаграммы;**
- **использовать стили оформления (например, заголовки разного уровня);**
- **использовать шаблоны документов (заранее оформленные заготовки);**
- **выполнять несложные вычисления в таблицах.**



Форматы файлов

Формат файла – это порядок расположения данных в файле.

doc, docx – документы текстового процессора Word

odt – открытый формат текстовых документов, с ним работает OpenOffice Writer

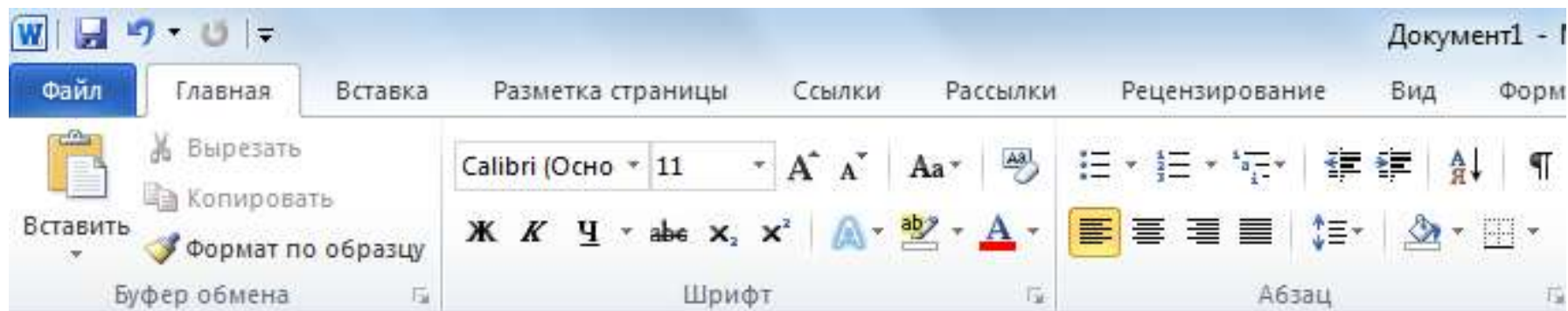
rtf – формат текстового документа с дополнительными возможностями (вставка таблиц, списков, рисунков)

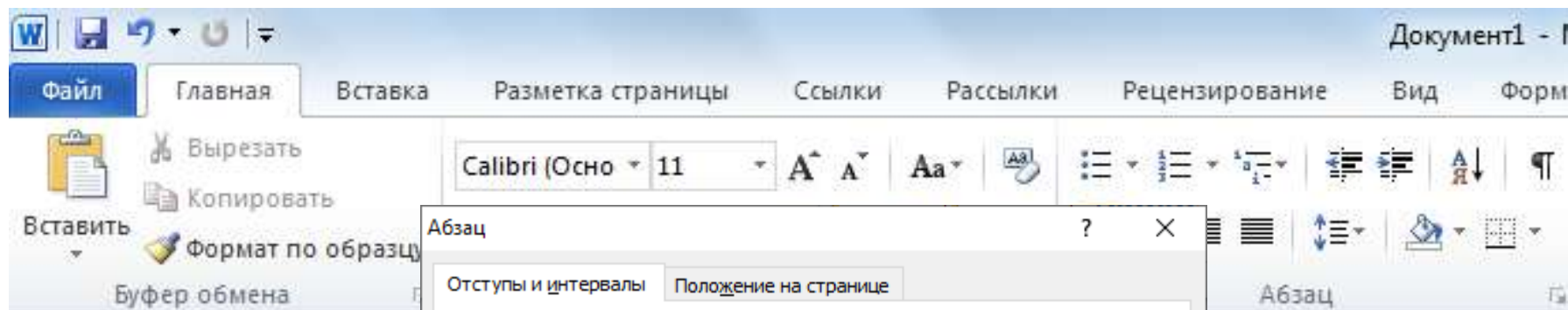
html, htm – форматы документов, которые можно размещать в Интернете

pdf - переносимый формат документов, этот формат был разработан для распространения документов, которые должны везде выглядеть одинаково

Интерфейс текстовых редакторов и процессоров

Интерфейс – это средства взаимодействия программы с пользователем.





Абзац

Отступы и интервалы Положение на странице

Общие

Выравнивание: По левому краю

Уровень: Основной текст

Отступ

Слева: 0 см

Справа: 0 см

первая строка: на:

(нет)

☐ Зеркальные отступы

Интервал

Перед: 0 пт

После: 10 пт

междустрочный: значение:

Множитель 1,15

☐ Не добавлять интервал между абзацами одного стиля

Образец

Примерный абзац Примерный абзац Примерный абзац Примерный абзац Примерный абзац
Примерный абзац Примерный абзац Примерный абзац Примерный абзац Примерный абзац
Примерный абзац

Образец текста Образец текста Образец текста Образец текста Образец текста Образец текста Образец текста
Образец текста Образец текста Образец текста Образец текста Образец текста Образец текста Образец текста
Образец текста Образец текста

Табуляция... По умолчанию ОК Отмена

Диалоговое окно – это окно, созданное программой, в котором пользователь может ввести данные или выбрать нужный вариант действий.

Имя

Текстовое поле

Масштаб по ширине

Поле для ввода числа

Положение

- ☐ Верхний индекс
- ☒ Обычный
- ☐ Нижний индекс

Радиокнопки

☒ Контур

☒ Тень

☐ Мигание

☐ Скрытый

Флажки

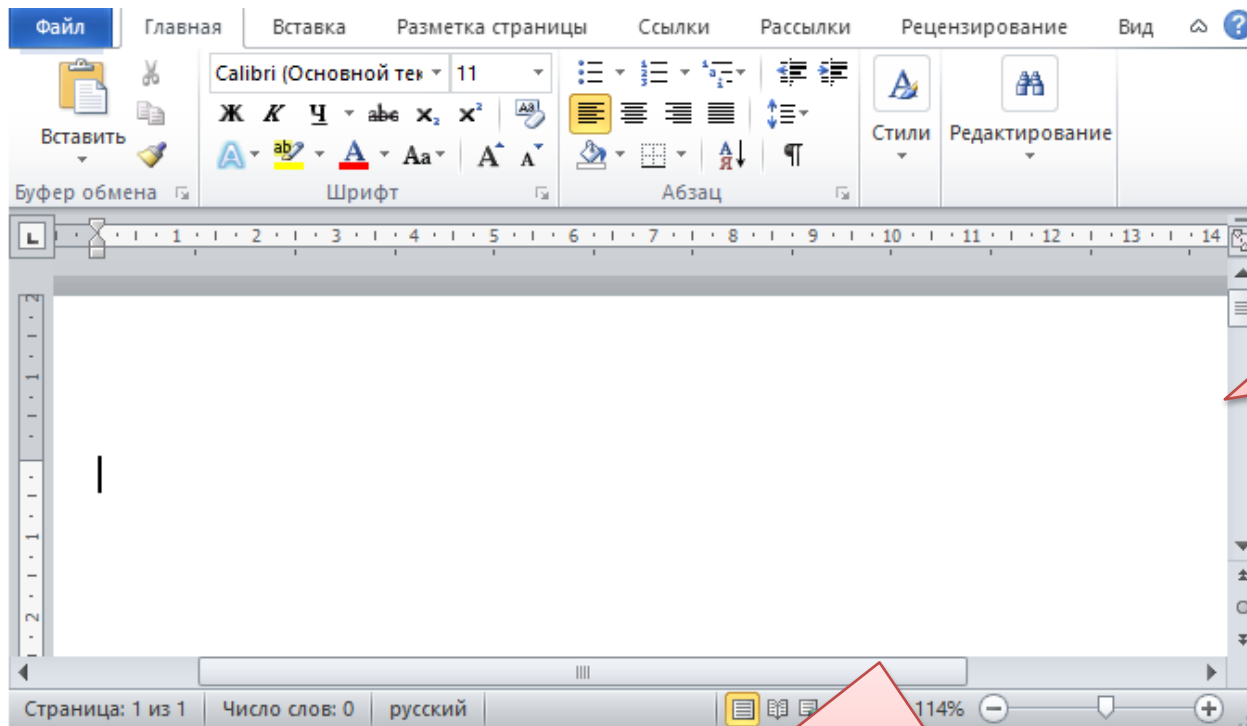
Интервал

Интервал

Выпадающий список

Полосы прокрутки

Если весь документ не помещается в отведенную для него часть окна, справа и снизу появляются полосы прокрутки.



**вертикальная
полоса
прокрутки**

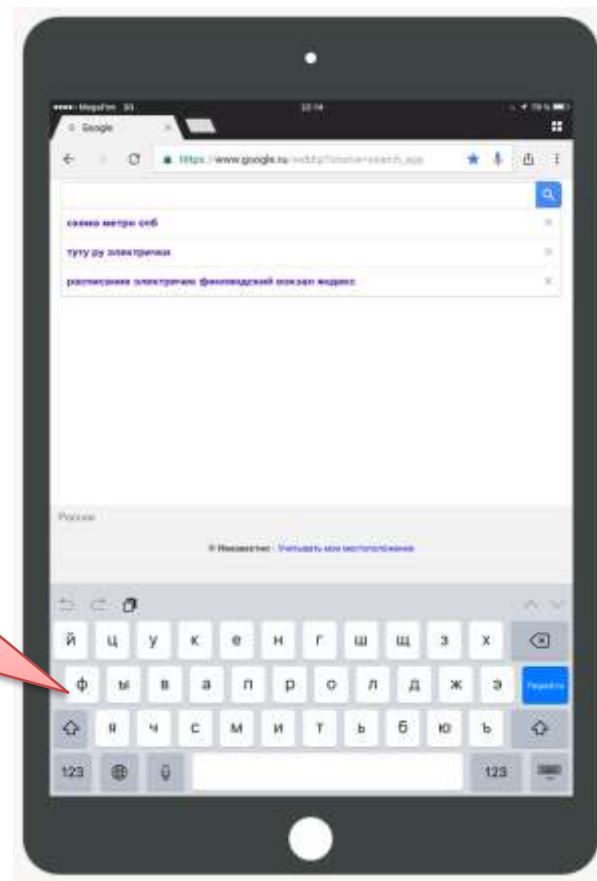
**горизонтальная полоса
прокрутки**

Редактирование текста

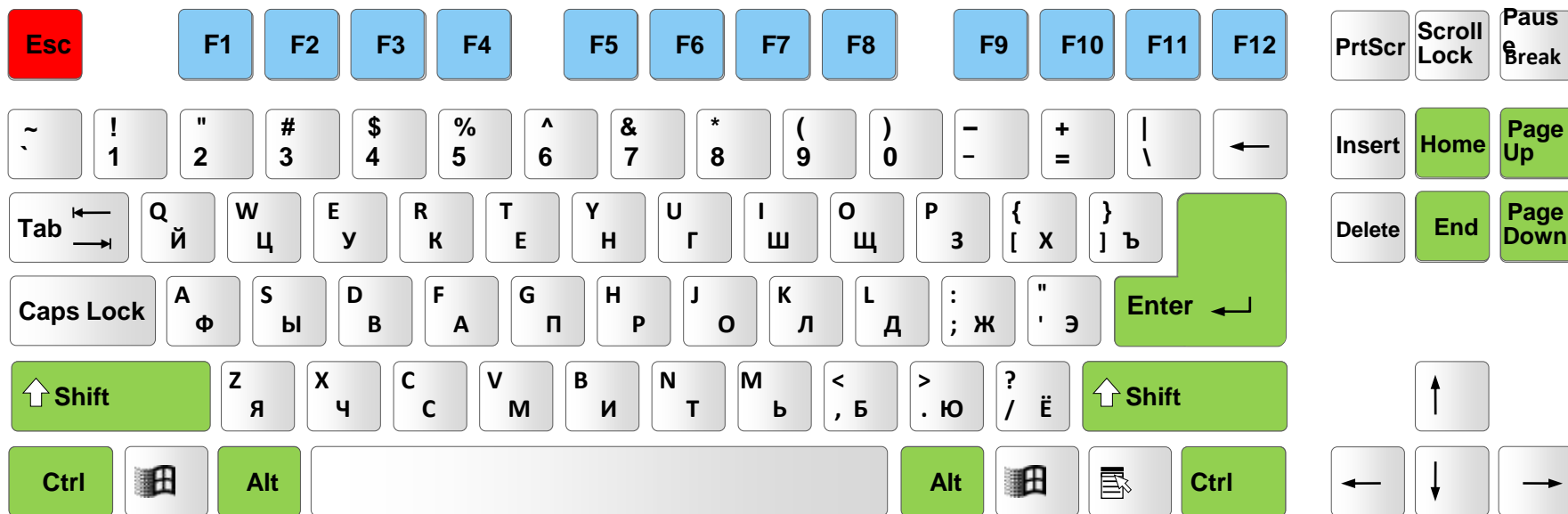
Редактирование – это изменение содержания текста,
т.е. вставка и удаление символов.



экранная
клавиатура



«Горячие клавиши»:



Ctrl + C – копирование

Ctrl + V – вставка

Ctrl + Z – отмена действия

Ctrl + Y – вернуть действие

Shift + (буква) – заглавная буква

Caps Lock – заглавные буквы

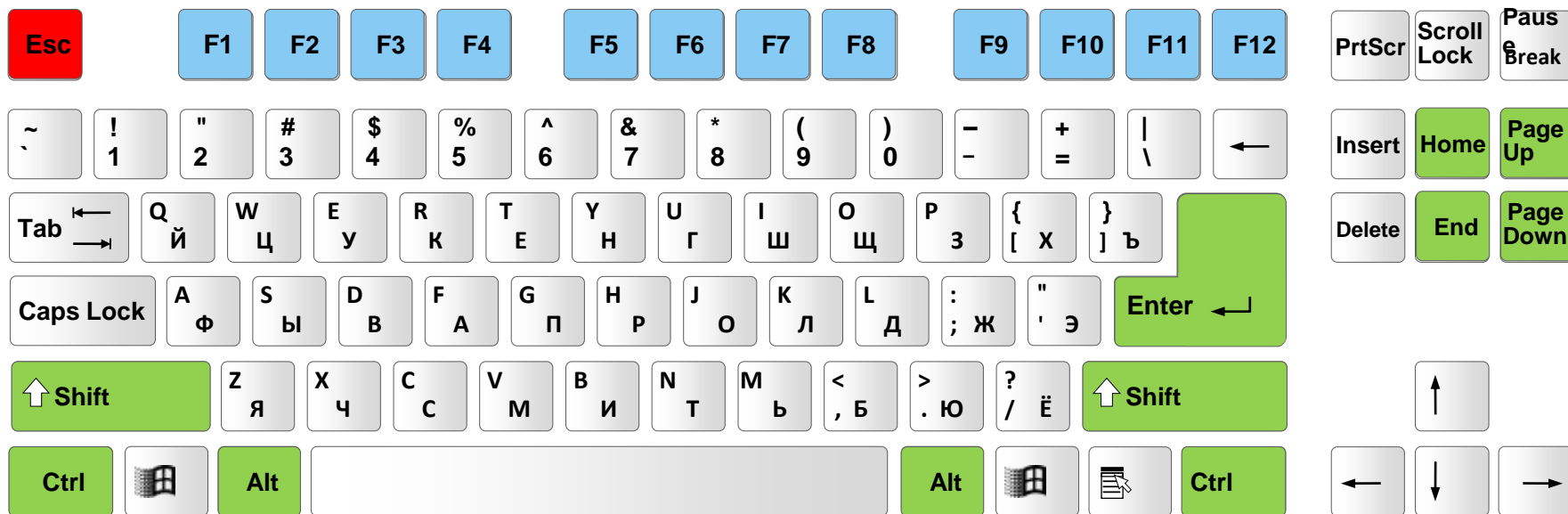
Enter – переход на новую строку

Backspace (←) – удалить

Del (Delete) – удалить

***Shift+Alt или Shift+Ctrl –
переход с английского на
русский и наоборот***

«Горячие клавиши»:



Home – в начало документа

End – в конец документа

Esc – отмена

Page up – на страницу вверх

Page down – на страницу вниз

F1-F12 – функциональные клавиши

Shift + 1 – !

Shift + 2 – « »

Shift + ?/ – ,

Shift + 3 – №

Shift + 4 – ;

Shift + 6 – :

Shift + 7 – ?

Shift + 9 – (

Shift + 0 –)

Пробелы и знаки препинания

- ✓ Между словами ставится только один пробел.
- ✓ В конце заголовка документа или раздела НЕ ставится точка:

Правильно:

Утро на озере

Однажды летом мы поехали
на рыбалку...

Неправильно:

Утро на озере✗

Однажды летом мы поехали
на рыбалку...

Пробелы и знаки препинания

- ✓ **Перед знаком препинания (точкой, запятой, точкой с запятой, двоеточием и др.) пробел не ставится, после знака ставится обязательно:**

Правильно:

Здравствуй, Семён!

Неправильно:

Здравствуй~~Х~~Семён!

нужен пробел

Пробелы и знаки препинания

- ✓ С внутренней стороны скобки и кавычек пробел НЕ ставятся, снаружи – ставятся (если после скобки или кавычки нет знака препинания):

нужен пробел

Правильно:

Он сказал «Поехали!» и
махнул рукой...

Правильно:

Петя (он был старше)
важно заметил, что...

Неправильно:

Он сказал «XПоехалиX!X»и
махнул рукой...

Неправильно:

Петя (Xон был старшеX)
важно заметил, что...

Пробелы и знаки препинания

- ✓ Тире выделяется пробелами с двух сторон, дефис записывается без пробелов:

Правильно:

Ты — мечтатель!
тёмно-вишнёвый

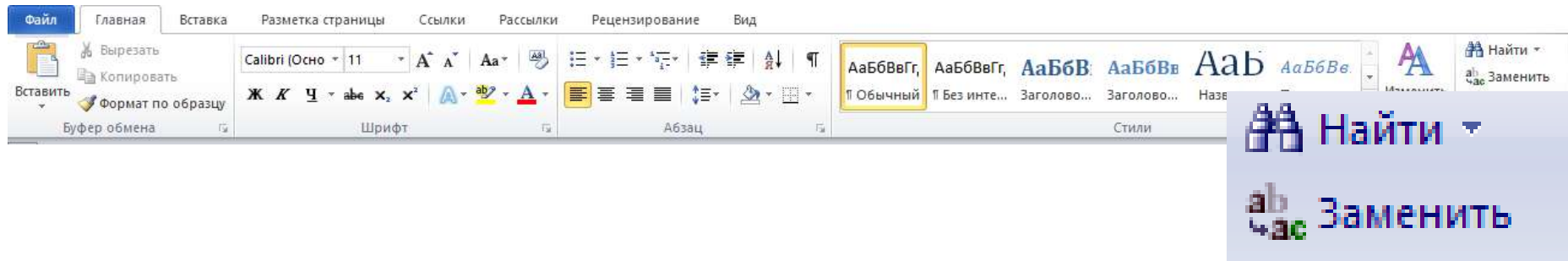
нужен
пробел

Неправильно:

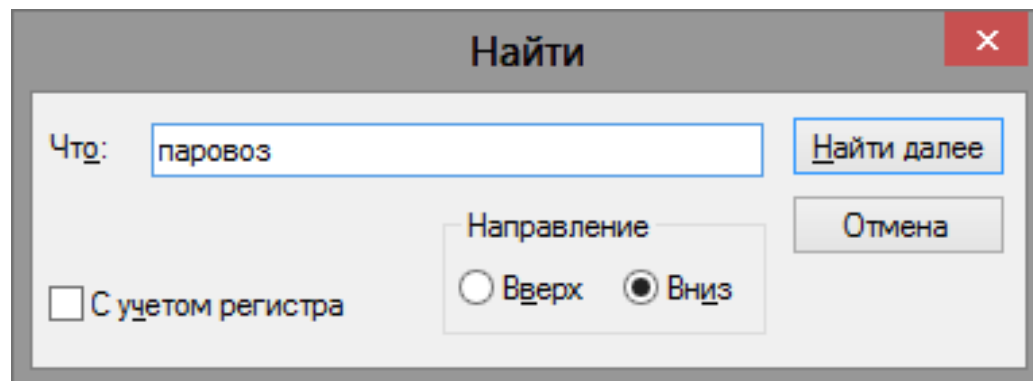
Ты—мечтатель!
тёмноX - Xвишнёвый

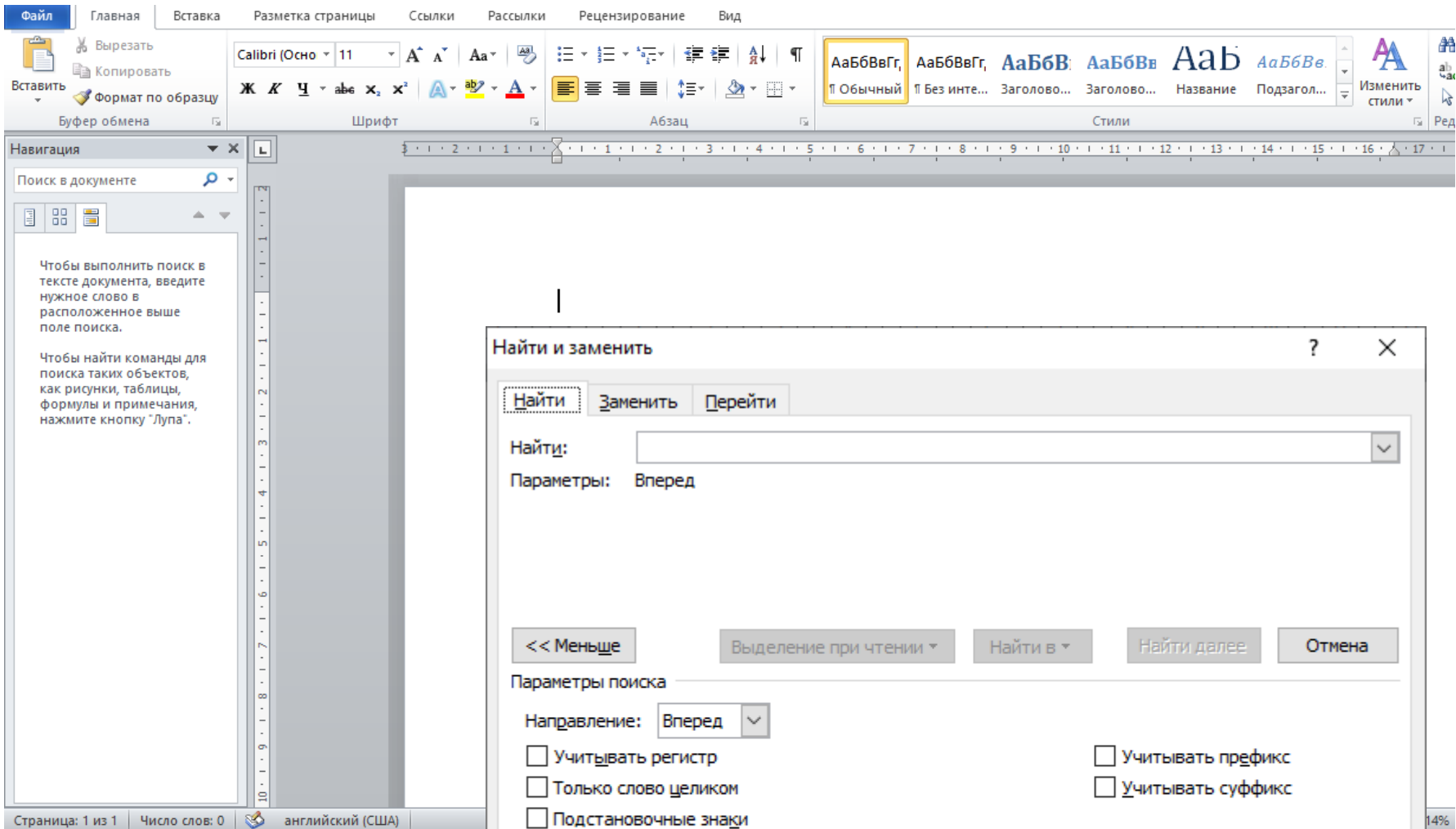
нужен
пробел

Поиск и замена



Ctrl+F:





Ctrl+F:

Ctrl+N:

Заменить [X]

Что:

Чем:

Найти и заменить [?] [X]

Найти:

Параметры:

Заменить на:

Параметры поиска

Направление:

<input type="checkbox"/> Учитывать регистр	<input type="checkbox"/> Учитывать префикс
<input type="checkbox"/> Только слово целиком	<input type="checkbox"/> Учитывать суффикс
<input type="checkbox"/> Подстановочные знаки	<input type="checkbox"/> Не учитывать знаки препинания
<input type="checkbox"/> Произносится как	<input type="checkbox"/> Не учитывать пробелы
<input type="checkbox"/> Все словоформы	

Заменить