

Основы теории информации

1. Виды информации.

Подходы к измерению информации

2. Свойства информации

3. Информационные технологии и информационные системы

Информация (от лат. *informatio*, разъяснение, изложение, осведомленность) — сведения о чем-либо не зависимо от формы их представления.

Информация — сведения об окружающем мире и протекающих в нем процессах, воспринимаемые человеком или специальным устройством.

- **Данные** - сведения, представленные в определенной знаковой системе и на определенном материальном носителе для обеспечения возможностей хранения, передачи, приема и обработки.
- **Информация** - это данные, сопровождающиеся смысловой нагрузкой, помещенные в некоторый контекст; данные, как-либо оцениваемые приёмником информации (в частном случае, это может быть человек). Как правило, получение информации связывают с уменьшением неопределенности существующего выбора; ответ на какой-либо заданный либо подразумеваемый вопрос. (При этом то, что для одних личностей может быть **данными**, для других вполне может быть **информацией**).
- **Знание** – зафиксированная и проверенная практикой информация, которая может многократно использоваться людьми для решения тех или иных задач.

Информационная революция — преобразование общественных отношений из-за кардинальных изменений в сфере обработки информации

Первая революция — изобретение письменности

Вторая революция (середина XVI в.) — изобретение книгопечатания

Третья революция (конец XIX в.) — изобретение электричества

Четвертая революция (70-е гг. XX в.) — изобретение микроспроцессорной технологии и появление персонального компьютера.

Этот период характеризуют три фундаментальные **инновации**:

- переход от механических и электрических средств преобразования информации к электронным;
- миниатюризация всех узлов, устройств, приборов, машин;
- создание программно-управляемых устройств и процессов.

Термин "информатика", - наука о получении, передаче, хранении и обработке информации. В настоящее время термин "информатика" все чаще заменяется более содержательным термином "информационные технологии" (ИТ).

Теоретическая информатика включала математическое моделирование информационных процессов.

Прикладная охватывала вопросы построения и проектирования ЭВМ, сетей, мультимедиа, компьютерные технологии информационных процессов и др.

Главной научной базой прикладной информатики были электроника (микроэлектроника) и теория искусственного интеллекта.

Информационное общество — общество, в котором большинство работающих занято производством, хранением, переработкой и реализацией информации, особенно высшей ее формы — знаний.

Информационная культура — умение целенаправленно работать с информацией и использовать для ее получения, обработки и передачи компьютерную информационную технологию, современные технические средства и методы.

Информационные ресурсы — отдельные документы и отдельные массивы документов в информационных системах (библиотеках, архивах, фондах, банках данных, других информационных системах).

Информационный продукт — совокупность данных, сформированная производителем для распространения в вещественной или невещественной форме.

Классификация информации

по способу восприятия:

- **Визуальная** — воспринимаемая органами зрения.
- **Аудиальная** — воспринимаемая органами слуха.
- **Тактильная** — воспринимаемая тактильными рецепторами.
- **Обонятельная** — воспринимаемая обонятельными рецепторами.
- **Вкусовая** — воспринимаемая вкусовыми рецепторами.

Классификация информации

по форме представления:

- **Текстовая** — передаваемая в виде символов, предназначенных обозначать лексемы языка.
- **Числовая** — в виде цифр и знаков, обозначающих математические действия.
- **Графическая** — в виде изображений, предметов, графиков.
- **Звуковая** — устная или в виде записи передача лексем языка аудиальным путём.

Классификация информации

по предназначению:

- **Массовая** — содержит тривиальные сведения и оперирует набором понятий, понятным большей части социума.
- **Специальная** — содержит специфический набор понятий, при использовании происходит передача сведений, которые могут быть не понятны основной массе социума, но необходимы и понятны в рамках узкой социальной группы, где используется данная информация.
- **Личная** — набор сведений о какой-либо личности, определяющий социальное положение и типы социальных взаимодействий внутри популяции.

Различные подходы к определению и измерению информации

	Подход к определению	Подход к измерению
в быту	разнообразные сведения, сообщения, их новизна	новизна не измеряется
в технике	сообщения, передаваемые в форме знаков и сигналов	информационная емкость = количеству символов
теория информации	снятая неопределенность	количество информации зависит от вероятности получения сообщения
теория принятия решений	увеличение вероятности достижения цели	попробуйте
кибернетика(теория управления)	часть знания, используемая в управлении	придумать
теория отражения	результат отражения, отраженное разнообразие	сами

В информатике используются следующие *подходы* к измерению информации:

- **Содержательный подход.** Информация - знания человека, поэтому сообщение должно быть информативно. Если сообщение не информативно, то количество информации с точки зрения человека = 0.
- **Алфавитный подход** - объективный подход к измерению информации, не зависит от содержания сообщения. Кол-во информации зависит от объема текста и мощности алфавита.
- **Вероятностный подход.** Все события происходят с различной вероятностью, но зависимость между вероятностью событий и количеством информации (энтропии), полученной при совершении того или иного события можно выразить формулой Шеннона.

Так как наступление каждого из N возможных событий имеет одинаковую вероятность

$p = 1 / N$, то $N = 1 / p$ и формула имеет вид

$$I = \log_2 (1/p) = - \log_2 p$$

Количество собственной информации,
переносимой одной буквой равно

$$h_i = \log_2 1/p_i = -\log_2 p_i,$$

где p_i - вероятность появления в сообщении i -го символа алфавита.

Средним значением количества информации,
приходящейся на один символ алфавита

$$H = \sum p_i h_i = -\sum p_i \log_2 p_i$$

Значение H достигает максимума при
равновероятных событиях, то есть при равенстве
всех p_i

$$p_i = 1 / N.$$

Формула Шеннона превращается в формулу
Хартли

$$I = \log_2 N \quad \text{или} \quad 2^I = N$$



Больцман ввел в 1872 Н-функцию, характеризующую состояние

замкнутой макроскопической системы, и показал, что с течением времени Н-функция не может возрастать (Н-теорема). Отождествив Н-функцию с энтропией S (с обратным знаком), Больцман связал энтропию с термодинамической вероятностью W .

Согласно Больцману, $S = k \log W$.

Эта формула украшает памятник Больцману в Вене. Еще лучшим памятником служит то, что его имя носит универсальная постоянная k .

А мы отметим пока, что именно в его работах сошлись энтропия и вероятность.

Содержательный подход

Примеры различных систем кодирования

C O M P U T E R

43 4F 4D 50 55 54 45 52

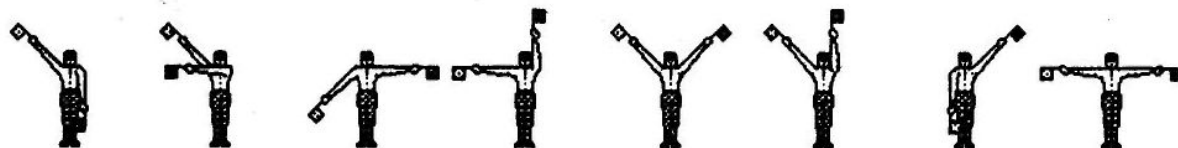
Код ASCII

-. .- .- .- .- .- .- .-

Код Морзе

●● ●●● ●●● ●●●● ●●● ●●●● ●●● ●●●●

Код Брайля

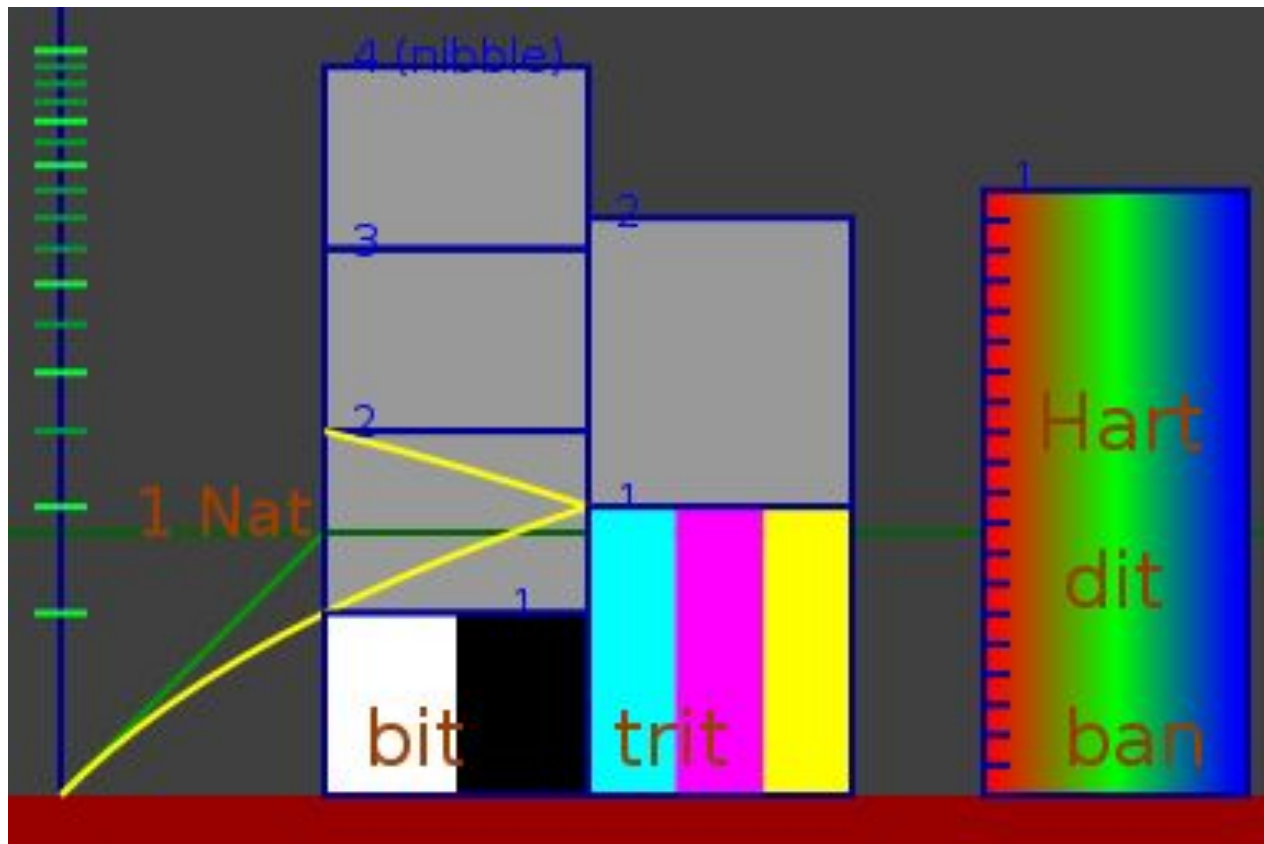
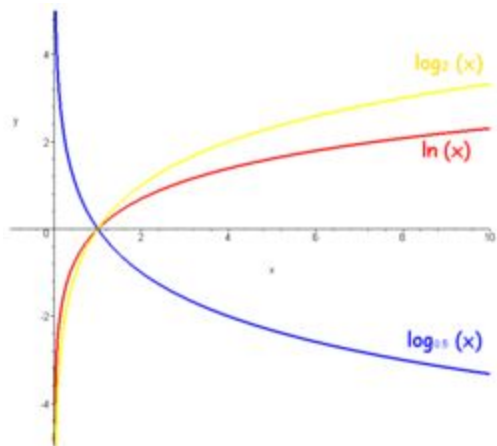


Код морской сигнальный

Единицы измерения количества информации

Единицы измерения информации служат для измерения объёма информации — величины, исчисляемой логарифмически.

Первичные единицы



Единицы, производные от бита

Измерения в байтах

Десятичная приставка			<u>Двоичная приставка</u>			
Название	Символ	Степень	Название	Символ	Степень	
ГОСТ			<u>МЭК</u>			
<u>байт</u>	В	10^0	<u>байт</u>	В	байт	2^0
<u>килобайт</u>	кВ	10^3	<u>кибибайт</u>	KiB	Кбайт	2^{10}
<u>мегабайт</u>	МВ	10^6	<u>мебибайт</u>	MiB	Мбайт	2^{20}
<u>гигабайт</u>	ГВ	10^9	<u>гибибайт</u>	GiB	Гбайт	2^{30}
<u>терабайт</u>	ТВ	10^{12}	<u>тебибайт</u>	TiB	Тбайт	2^{40}
<u>петабайт</u>	РВ	10^{15}	<u>пебибайт</u>	PiB	Пбайт	2^{50}
<u>эксабайт</u>	ЕВ	10^{18}	<u>эксбибайт</u>	EiB	Эбайт	2^{60}
<u>зеттабайт</u>	ЗВ	10^{21}	<u>зебибайт</u>	ZiB	Збайт	2^{70}
<u>йоттабайт</u>	УВ	10^{24}	<u>йобибайт</u>	YiB	Йбайт	2^{80}

1 байт = 8 бит

1 Кб (килобайт) = 2^{10} байт = 1024 байт

1 Мб (мегабайт) = 2^{10} Кб = 1024 Кб

1 Гб (гигабайт) = 2^{10} Мб = 1024 Мб

Употребление десятичных приставок

Приставка	Обозначение	Ошибочное применение	Корректное применение	Относит. ошибка, %
кило	к, k	$2^{10} = 1\ 024$	$10^3 = 1\ 000$	2,40
мега	М, M	$2^{20} = 1\ 048\ 576$	$10^6 = 1\ 000\ 000$	4,86
гига	Г, G	$2^{30} = 1\ 073\ 741\ 824$	$10^9 = 1\ 000\ 000\ 000$	7,37
тера	Т, T	$2^{40} = 1\ 099\ 511\ 627\ 776$	$10^{12} = 1\ 000\ 000\ 000\ 000$	9,95
пета	П, P	$2^{50} = 1\ 125\ 899\ 906\ 842\ 624$	$10^{15} = 1\ 000\ 000\ 000\ 000\ 000$	12,59
экса	Э, E	$2^{60} = 1\ 152\ 921\ 504\ 606\ 846\ 976$	$10^{18} = 1\ 000\ 000\ 000\ 000\ 000\ 000$	15,29
зетта	З, Z	$2^{70} = 1\ 180\ 591\ 620\ 717\ 411\ 303\ 424$	$10^{21} = 1\ 000\ 000\ 000\ 000\ 000\ 000\ 000$	18,06
йотта	Й, Y	$2^{80} = 1\ 208\ 925\ 819\ 614\ 629\ 174\ 706\ 176$	$10^{24} = 1\ 000\ 000\ 000\ 000\ 000\ 000\ 000\ 000$	20,89

ПРИМЕР

В 100 Мб можно “уместить”:

страниц текста 000 или	50	150 романов
цветных слайдов высочайшего качества	150	
аудиозапись речи видного политического деятеля	1.5 часа	
музыкальный фрагмент качества CD -стерео	10 минут	
фильм высокого качества записи	15 секунд	

Кодирование числовых данных

Целые числа

$$19:2 = 9 + 1$$

$$9:2=4+1$$

$$4:2=2+0$$

$$2:2 = 1$$

Таким образом,

$$19_{10} = 1011_2.$$

Действительные числа

$$3,1415=0,31415 \cdot 10^1$$

$$700\ 000=0,7 \cdot 10^6$$

КОДИРОВАНИЕ ТЕКСТОВЫХ

ДАННЫХ

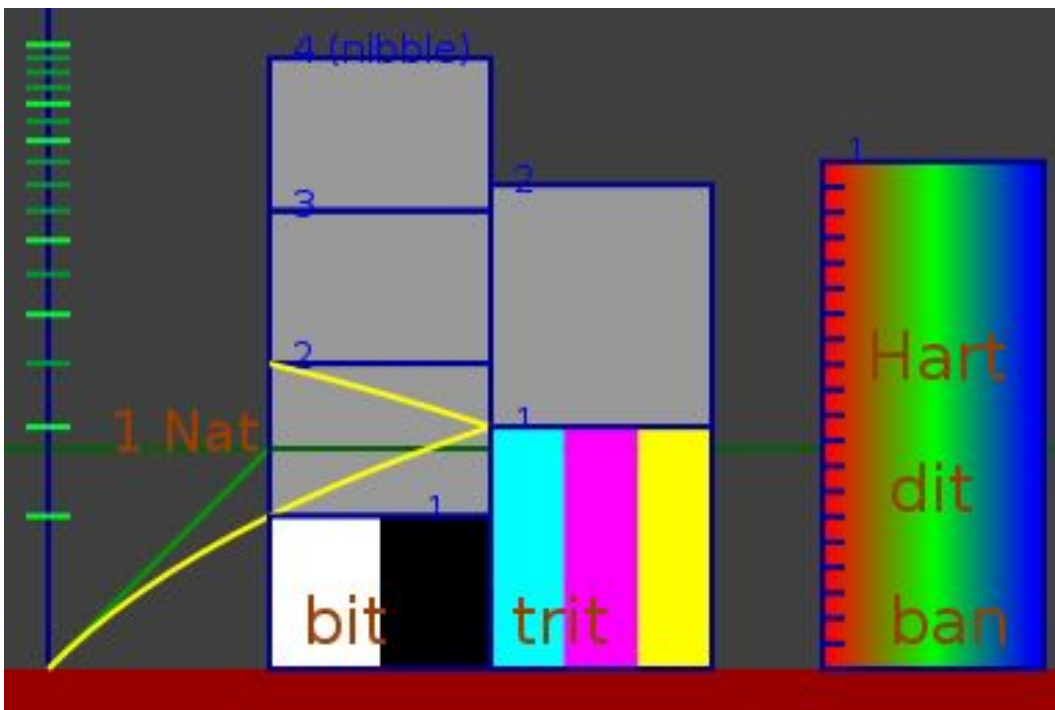
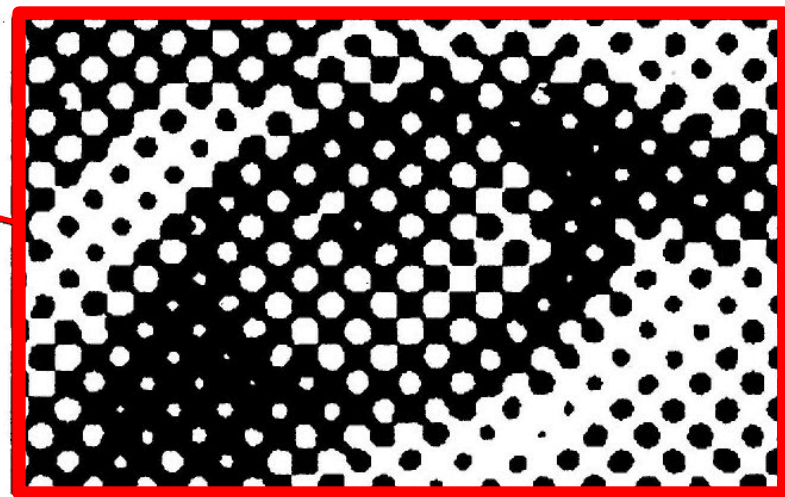
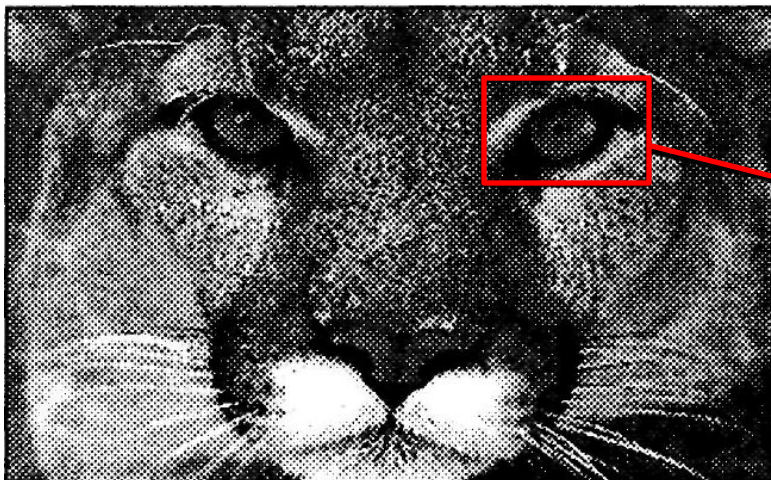
Базовая таблица кодировки ASCII (American Standard Code for Information Interchange)

32	пробел	48	0	64	@	80	P	96	`	112	p
33	!	49	1	65	A	81	Q	97	a	113	q
34	"	50	2	66	B	82	R	98	b	114	r
35	#	51	3	67	C	83	S	99	c	115	s
36	\$	52	4	68	D	84	T	100	d	116	t
37	%	53	5	69	E	85	U	101	e	117	u
38	&	54	6	70	F	86	V	102	f	118	v
39	'	55	7	71	G	87	W	103	g	119	w
40	(56	8	72	H	88	X	104	h	120	x
41)	57	9	73	I	89	Y	105	i	121	y
42	*	58	:	74	J	90	Z	106	j	122	z
43	+	59	:	75	K	91	[107	k	123	{
44	,	60	<	76	L	92	\	108	l	124	
45	-	61	=	77	M	93]	109	m	125	}
46	.	62	>	78	N	94	^	110	n	126	~
47	/	63	?	79	O	95	_	111	o	127	

Кодировка Windows 1251

128	Ђ	144	ђ	160	Ў	176	•	192	А	208	Р	224	а	240	р
129	Ѓ	145	ѓ	161	ў	177	±	193	Б	209	С	225	б	241	с
130	„	146	ђ	162	џ	178	і	194	В	210	Т	226	в	242	т
131	ѓ	147	„	163	Ј	179	і	195	Г	211	У	227	г	243	у
132	„	148	”	164	Ѡ	180	г	196	Д	212	Ф	228	д	244	ф
133	…	149	•	165	Г	181	μ	197	Е	213	Х	229	е	245	х
134	†	150	—	166	Г	182	¶	198	Ж	214	Ц	230	ж	246	ц
135	‡	151	—	167	§	183	•	199	З	215	Ч	231	з	247	ч
136	•	152	’	168	£	184	ё	200	И	216	Ш	232	и	248	ш
137	%o	153	™	169	©	185	№	201	Й	217	Щ	233	й	249	щ
138	Љ	154	љ	170	€	186	е	202	К	218	Ъ	234	к	250	ъ
139	‹	155	›	171	«	187	»	203	Л	219	Ы	235	л	251	ы
140	Њ	156	њ	172	¬	188	ј	204	М	220	Ь	236	м	252	ь
141	Ќ	157	ќ	173	-	189	ѕ	205	Н	221	Э	237	н	253	э
142	Ћ	158	ћ	174	®	190	s	206	О	222	Ю	238	о	254	ю
143	Ќ	159	џ	175	Ї	191	ї	207	П	223	Я	239	п	255	я

Кодирование графических данных

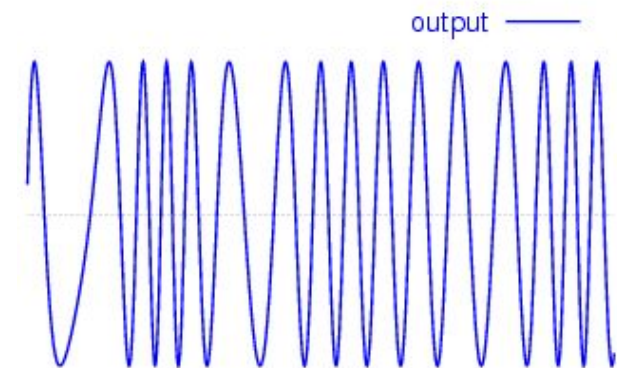
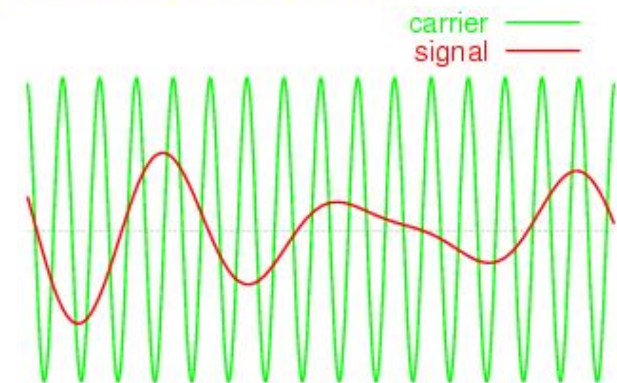
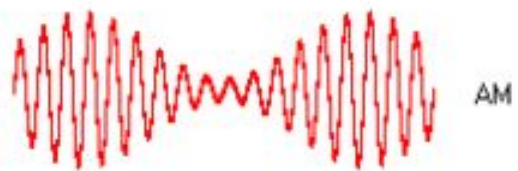


Системы (True Color):

24 разряда – **RGB** используют три основных цвета: красный (*Red, R*), зеленый (*Green, G*) и синий (*Blue, B*).

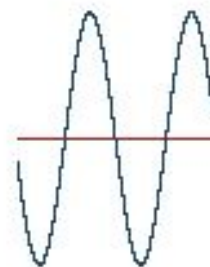
32 двоичных разряда – **СМУК** дополнительные цвета: голубой (*Cyan, C*), пурпурный (*Magenta, M*) и желтый (*Yellow, Y*)+(Black, B)

Кодирование звуковой информации

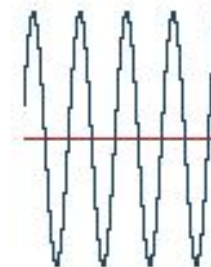


Carrier

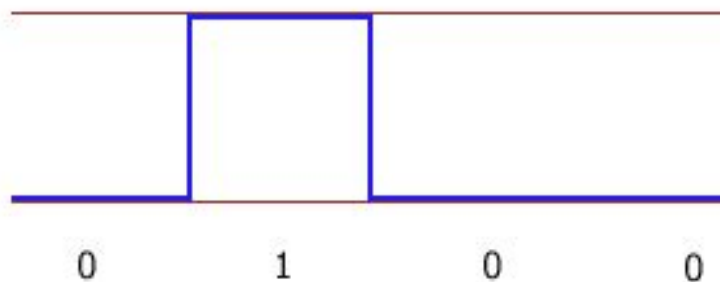
0



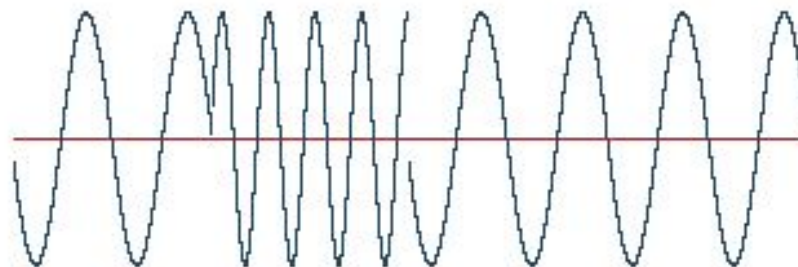
1



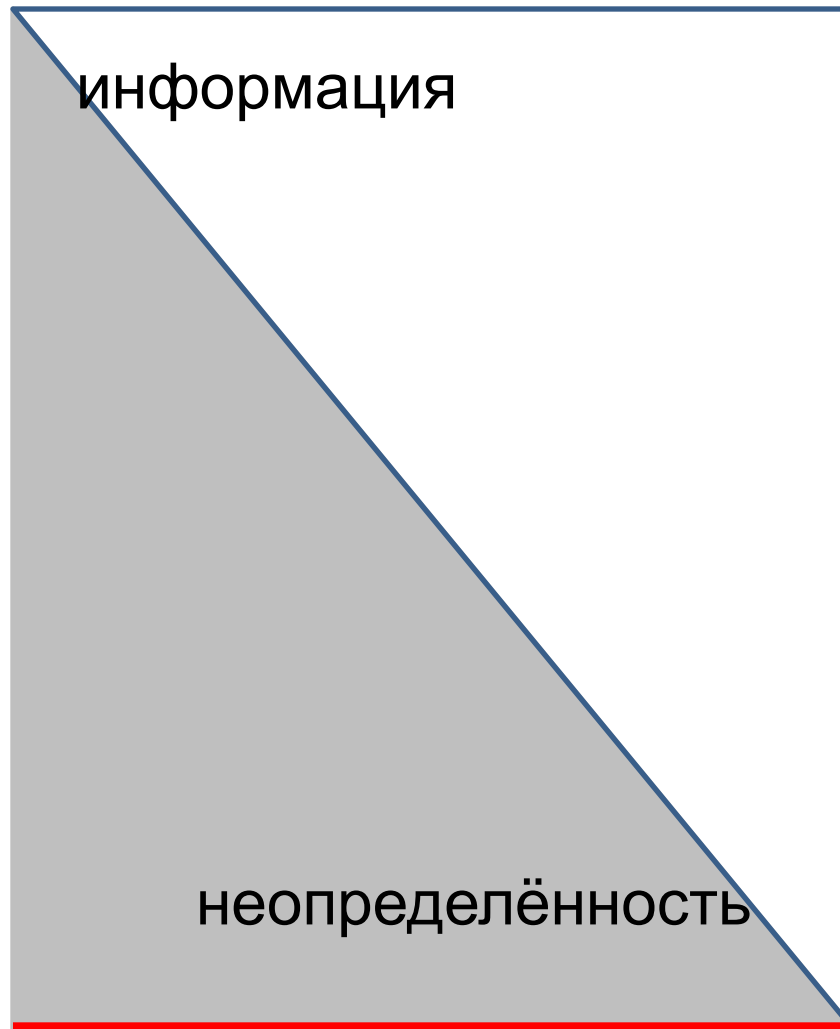
Modulating Wave (digital)



Modulated Result



- **Вероятностный подход.** Получение информации (ее *увеличение*) одновременно означает увеличение знания, что, в свою очередь, означает уменьшение незнания или информационной неопределенности.



Расчетная формула Хартли имеет вид:

$$I = \log_2 N \quad \text{или} \quad 2^I = N,$$

где N - количество *равновероятных* событий
(число возможных выборов),

I - количество информации.

Если $N = 2$ (выбор из двух возможностей), то $I = 1$
бит.

Таблица вычисления двоичных логарифмов чисел от 1 до 64 по формуле $I = \log_2 N$ (пример, $\log_2 8 = 3$, потому что $2^3 = 8$)

N	x	N	x	N	x	N	x
1	0,00000	17	4,08746	33	5,04439	49	5,61471
2	1,00000	18	4,16993	34	5,08746	50	5,64386
3	1,58496	19	4,24793	35	5,12928	51	5,67243
4	2,00000	20	4,32193	36	5,16993	52	5,70044
5	2,32193	21	4,39232	37	5,20945	53	5,72792
6	2,58496	22	4,45943	38	5,24793	54	5,75489
7	2,80735	23	4,52356	39	5,28540	55	5,78136
8	3,00000	24	4,58496	40	5,32193	56	5,80735
9	3,16993	25	4,64386	41	5,35755	57	5,83289
10	3,32193	26	4,70044	42	5,39232	58	5,85798
11	3,45943	27	4,75489	43	5,42626	59	5,88264
12	3,58496	28	4,80735	44	5,45943	60	5,90689
13	3,70044	29	4,85798	45	5,49185	61	5,93074
14	3,80735	30	4,90689	46	5,52356	62	5,95420
15	3,90689	31	4,95420	47	5,55459	63	5,97728
16	4,00000	32	5,00000	48	5,58496	64	6,00000

Обстоятельства определения информации:

- 1. Динамический характер информации.** Информация не является статичным объектом — она динамически меняется и существует только в момент взаимодействия данных и методов. Информация существует только в момент протекания *информационного процесса*. Все остальное время она содержится в виде данных.
- 2. Требование адекватности методов.** Одни и те же данные могут в момент потребления поставлять разную информацию в зависимости от степени адекватности взаимодействующих с ними методов.
- 3. Диалектический характер взаимодействия данных и методов.** Информация возникает и существует в момент диалектического взаимодействия объективных данных и субъективных методов.

Свойства информации

● объективность

Более объективной принято считать ту информацию, в которую методы вносят меньший субъективный элемент.

● полнота

Полнота информации во многом характеризует *качество информации* и определяет *достаточность* данных для принятия решений или для создания новых данных на основе имеющихся.

● достоверность

Данные возникают в момент регистрации сигналов, но не все сигналы являются «полезными» — всегда присутствует какой-то уровень посторонних сигналов, в результате чего полезные данные сопровождаются определенным уровнем «информационного шума».

● адекватность

Степень соответствия реальному объективному состоянию дела.

● доступность

Мера возможности получить ту или иную информацию.

● актуальность

Степень соответствия информации текущему моменту времени.

Свойства информации

Качество информации — степень её соответствия потребностям потребителей

1. Объективность информации характеризует её независимость от чьего-либо мнения или сознания, а также от методов получения. Более объективна та информация, в которую методы получения и обработки вносят меньший элемент субъективности.

2. Полнота — информацию можно считать полной, когда она содержит минимальный, но достаточный для принятия правильного решения набор показателей. Как неполная, так и избыточная информация снижает эффективность принимаемых на основании информации решений.

Свойства информации

3. **Достоверность** — верность информации, не вызывающая сомнений. Объективная информация всегда достоверна, но достоверная информация может быть как объективной, так и субъективной. Причинами недостоверности могут быть:

- преднамеренное искажение (дезинформация);
- непреднамеренное искажение субъективного свойства;
- искажение в результате воздействия помех;
- ошибки фиксации информации.

В общем случае достоверность информации достигается:

- указанием времени свершения событий, сведения о которых передаются;
- сопоставлением данных, полученных из различных источников;
- своевременным вскрытием дезинформации;
- исключением искажённой информации и др.

Свойства информации

4. **Адекватность** — степень соответствия смысла реально полученной информации и её ожидаемого содержания.
5. **Доступность** — мера возможности получить ту или иную информацию.
6. **Актуальность** — это степень соответствия информации текущему моменту времени.
7. **Эмоциональность** — свойство информации вызывать различные эмоции у людей.

Информационная технология (Information technology) – система взаимосвязанных методов и способов сбора, хранения, накопления, поиска, обработки информации на основе применения вычислительной техники

Виды обрабатываемой информации

- Данные
 - СУБД, алгоритмические языки, табличный процессор
- Текст
 - Текстовый процессор
- Графика
 - Графический процессор
- Знания
 - Экспертные системы
- Объекты реального мира
 - Технология мультимедиа

Интегрированные пакеты: объединение различных технологий

Информационная технология (ИТ) — процесс, использующий совокупность средств и методов сбора, обработки и передачи данных (первичной информации) для получения информации нового качества о состоянии объекта, процесса или явления.

Этапы развития информационных технологий по видам инструментария:

1 этап (до второй половины XIX в.) — "**ручная**" (перо, чернильница, книга)

2 этап (с конца XIX в.) — "**механическая**" (пишущая машинка, телефон)

3 этап (40 - 60-е гг.) — "**электрическая**" (большие ЭВМ и создаваемые на их базе автоматизированные системы управления (АСУ) и информационно-поисковые системы (ИПС))

4 этап (с начала 70-х гг.) — "**электронная**" (большие ЭВМ, электрические пишущие машинки, копировальные аппараты, портативные диктофоны)

5 этап (с середины 80-х гг.) — "**компьютерная**" (персональный компьютер)

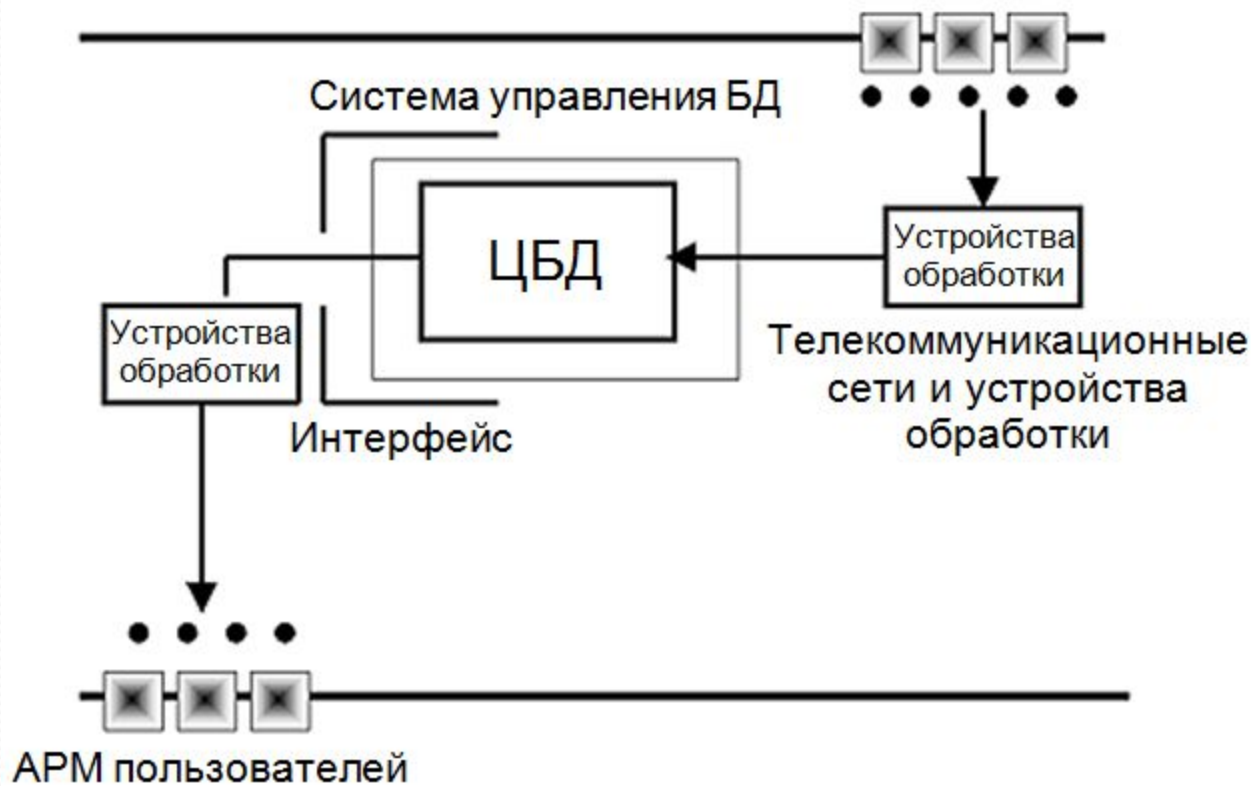
Информационная система (information system) это взаимосвязанная совокупность информационных, технических, программных, математических, организационных, правовых, эргономических, лингвистических, технологических средств, а также персонала, предназначенная для сбора, обработки, хранения и выдачи информации

- **Оргкомитет спортивного соревнования (мероприятия)**
 - Информация для **тренеров и спортсменов**
 - Информация для **зрителей**
 - Информация для **СМИ**
 - Информация для **спонсоров**
 - Информация для **рекламодателей**
 - Информация для **федераций**
 - Информация для **органов власти**

информационного обеспечения соревнований:

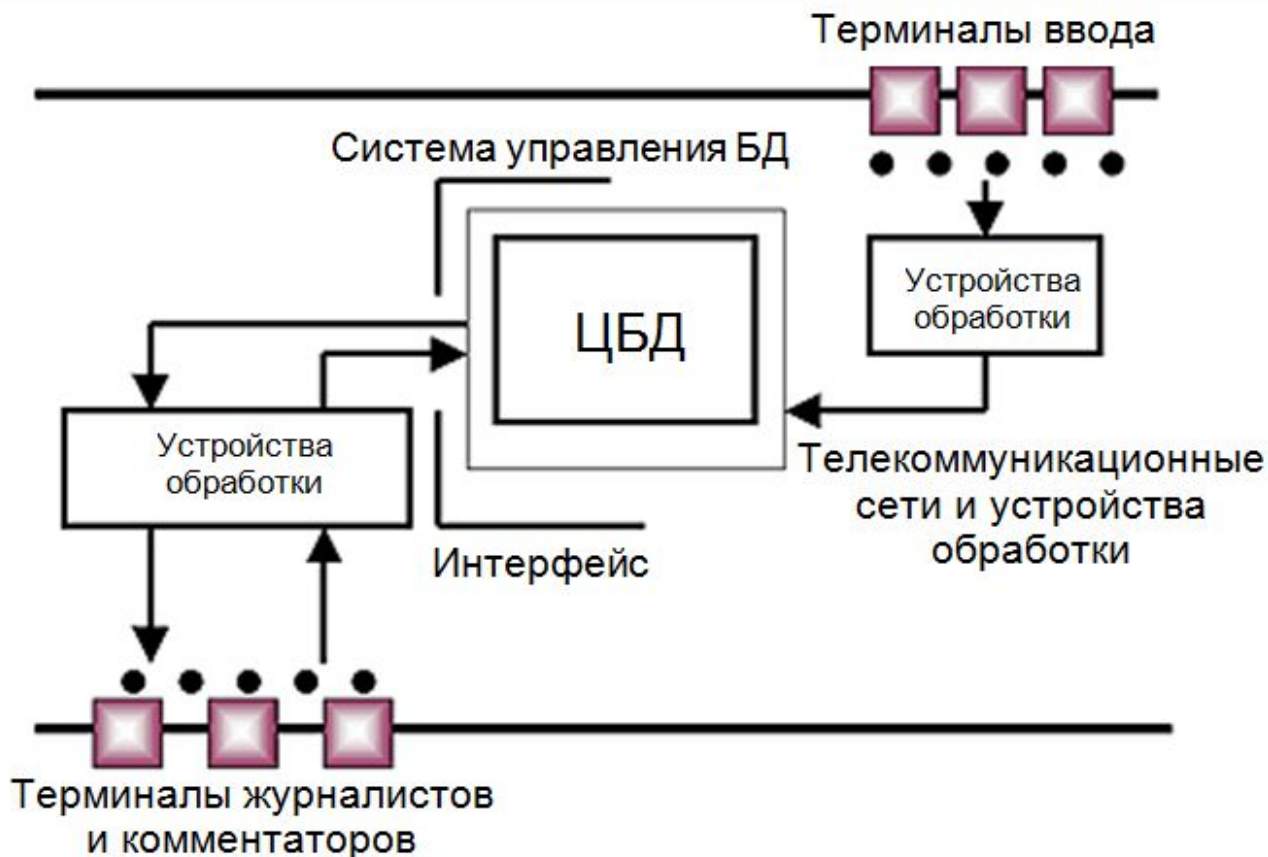
- **организация автоматизированного документооборота** – создание автоматизированной системы, обеспечивающей организационно-техническое обслуживание соревнований;
- **создание информационных терминалов журналиста и комментатора;**
- **создание информационного интерфейса TV-трансляции и табло соревнований** – визуальное оформление средствами компьютерной и TV-графики информационного потока для эфирного и кабельного телевидения (бегущее время, счет, текущий результат, сведения об участниках соревнований и т. п.) и внутренней телевизионной и звуковой сети (ВТЗС);
- **организация Web-сайта соревнования** – создание средств презентации соревнований в глобальной компьютерной сети, позволяющих образовывать интерактивный информационно-зрелищный комплекс на основе различных медиа-продуктов.

1. Схема работы системы автоматизированного документооборота



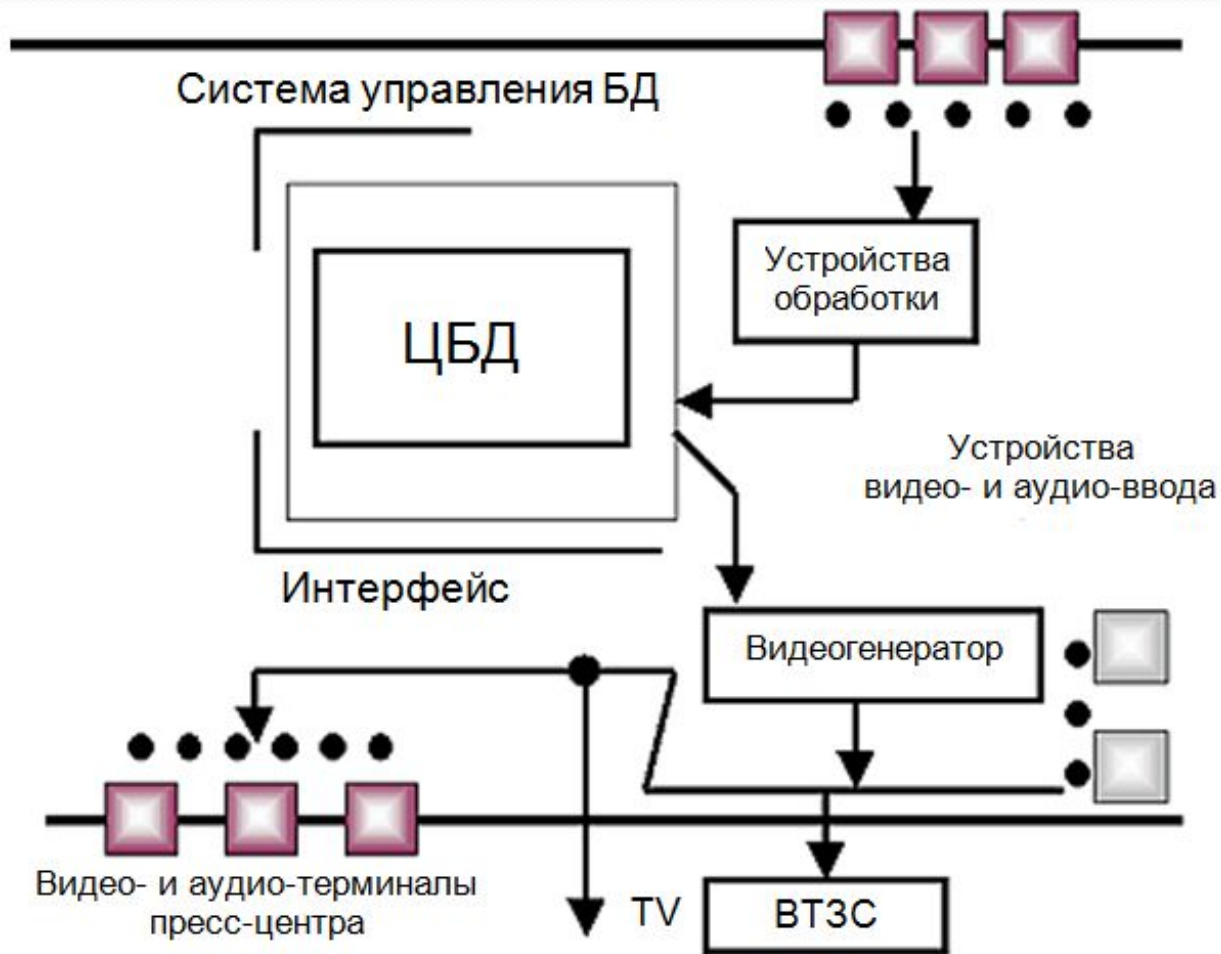
- аккредитация всех категорий участников и обслуживающего персонала; подготовка соответствующих удостоверений;
- жеребьевка;
- сбор и обработка результатов соревнований;
- распечатка судебных протоколов и других материалов;
- архивирование необходимых документов.

2. Схема работы информационного терминала журналиста и комментатора



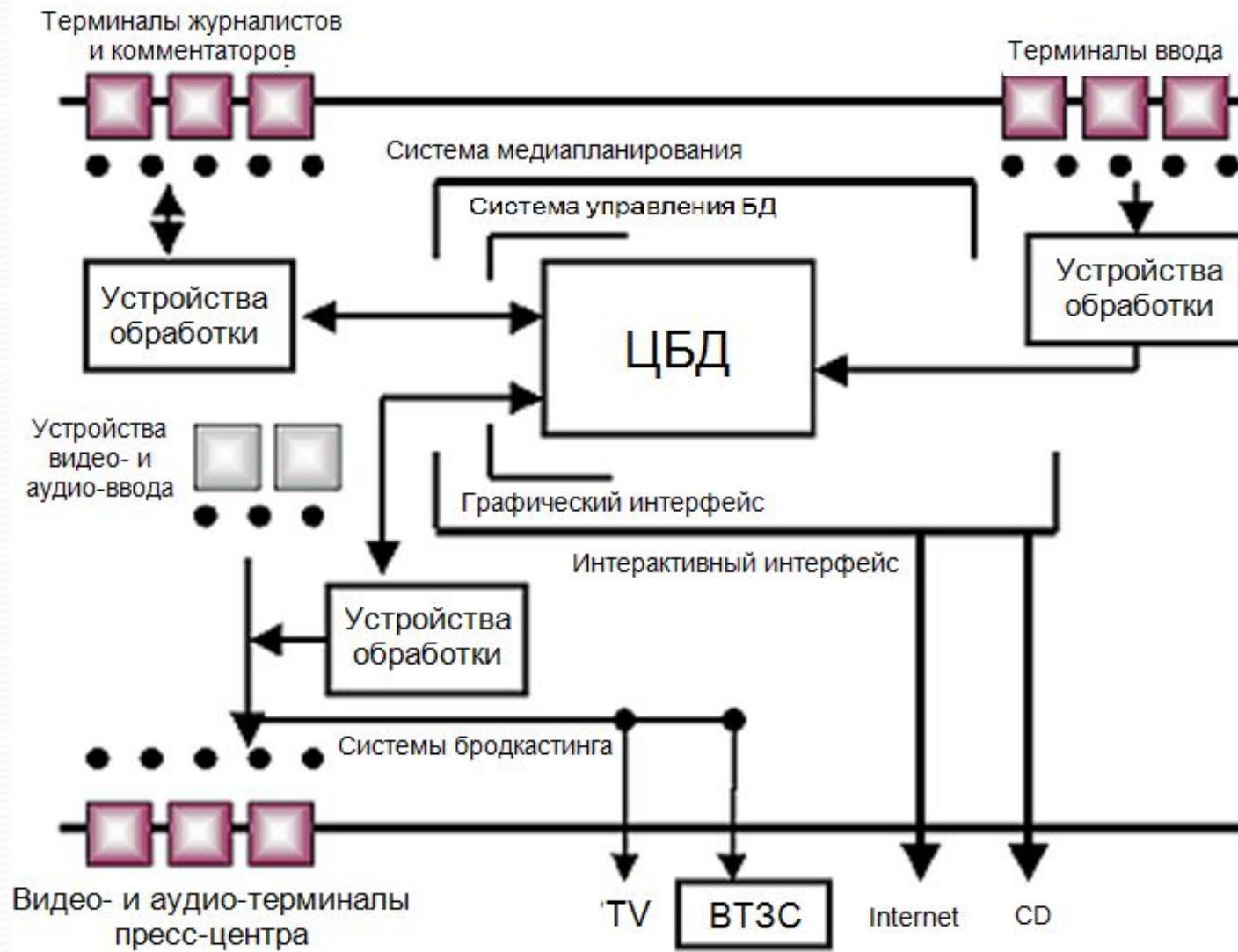
- сведения об участниках (биографические данные, спортивная карьера, индивидуальные особенности и т.п.),
- история события (места проведения, годы, чемпионы, медали и т.п.);
- сведения об инфраструктуре соревнования (место проведения, размещение, транспорт, погода, культурная программа и т.п.);
- информационные подборки по состоявшимся соревнованиям и т.д.

3. Схема работы информационного обеспечения TV-трансляции и внутренней телевизионной и звуковой системы (ВТЗС)



- вывод на экраны текущей информации: текущее время, счет, текущий результат, сведения об участниках соревнований и т.п.;
- и наборов статистической информации (рейтинги и другие сравнительные таблицы) и т.д.

4. Схема работы Web-презентации мероприятия



Данные – это краткая информация, пригодная для дальнейшей обработки. Как правило, данные предоставляются в структурированном виде, через заполнение какой-либо формы.