

Глава 1. Теоретические основы информатики

Лекция 1. Информация и общество

Лекция 2. Информация

Лекция 3. Представление информации

Лекция 4. Информационные процессы

Лекция 5. ЭВМ — техническое средство информатики

Глава 1. Теоретические основы информатики

Лекция 2. Информация

- 1.2.1. Понятие информации
- 1.2.2. Информационные сигналы
- 1.2.3. Преобразование сообщений
- 1.2.4. Понятие данных
- 1.2.5. Свойства информации
- 1.2.6. Меры информации
- 1.2.7. Экономическая информация

1.2.1. Понятие информации

Information — разъяснение, осведомление, изложение.

Схема понятия информации



1.2.1. Понятие информации

Информация — это сведения об объектах, процессах или явлениях окружающей среды, полученные человеком непосредственно или с помощью технических средств, уменьшающие степень неопределенности, неполноты знаний о них.

Следовательно:

1. Существует источник информации (объект, процесс или явление окружающей среды).
2. Существует получатель информации (человек или техническое устройство, преобразующее получаемую информацию для человека).
3. Происходит взаимодействие двух материальных объектов (источника и получателя).
4. В результате взаимодействия у получателя образуются сведения об источнике (данные об объекте).
5. Вновь полученные данные сравниваются с уже имеющимися сведениями (храняемыми в памяти).
6. Информацией становится та часть вновь полученных данных, которая дополняет уже имеющиеся сведения.

1.2.1. Понятие информации

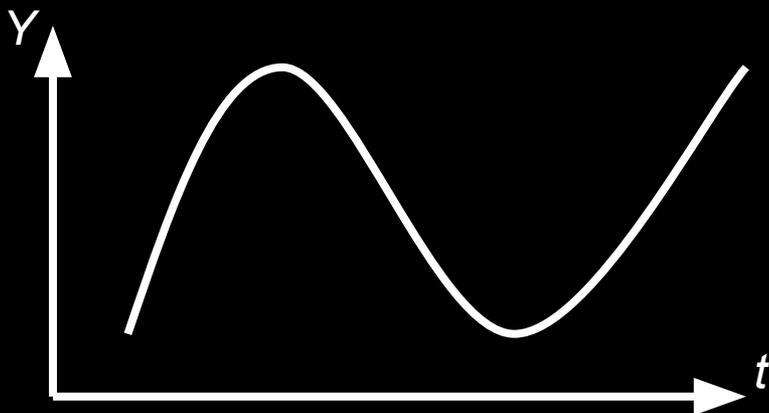
Любое взаимодействие материальных объектов имеет материальную или энергетическую природу и осуществляется посредством материального носителя — **сигналов**, распространяющихся в пространстве и времени.

Схема получения информации

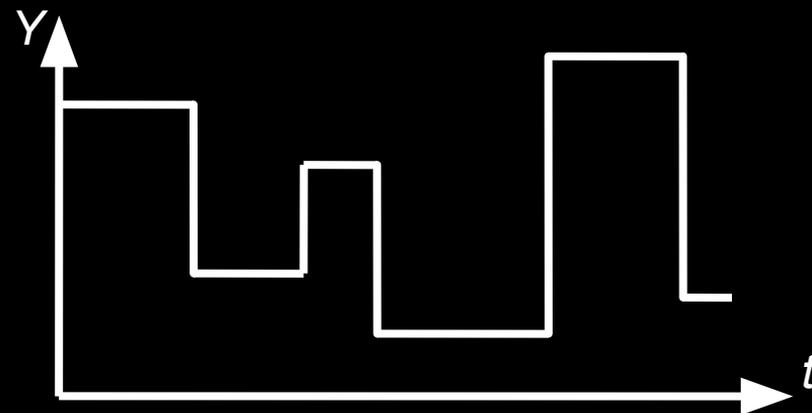


1.2.2. Информационные сигналы

Виды сигналов



а) непрерывный сигнал



б) дискретный сигнал

Сигнал называется аналоговым (или непрерывным), если его параметр может принимать любое значение в пределах некоторого интервала (параметр является непрерывной функцией времени).

Сигнал называется дискретным, если его параметр может принимать конечное число значений в пределах некоторого интервала.

1.2.2. Информационные сигналы

Дискретные сигналы

Принципиальным и важнейшим различием непрерывных и дискретных сигналов является то, что дискретные сигналы можно обозначить, т. е. приписать каждому из конечного числа возможных значений сигнала знак, который будет отличать данный сигнал от другого.

Знак — это элемент некоторого конечного множества отличных друг от друга сущностей.

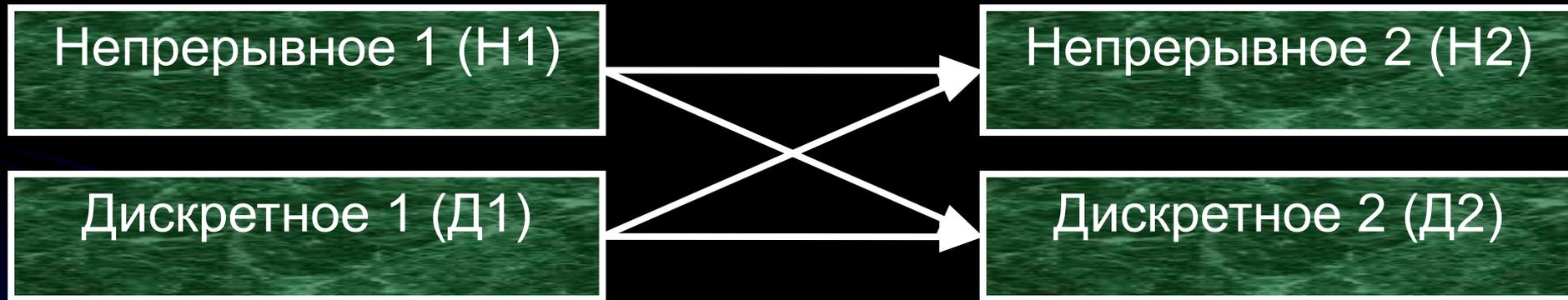
Природа знака может быть любой — жест, рисунок, буква, сигнал светофора, определенный звук и т. д.

Вся совокупность знаков, используемых для представления дискретной информации, называется **набором знаков**.

Набор знаков, в котором установлен порядок их следования, называется **алфавитом**.

1.2.3. Преобразование сообщений

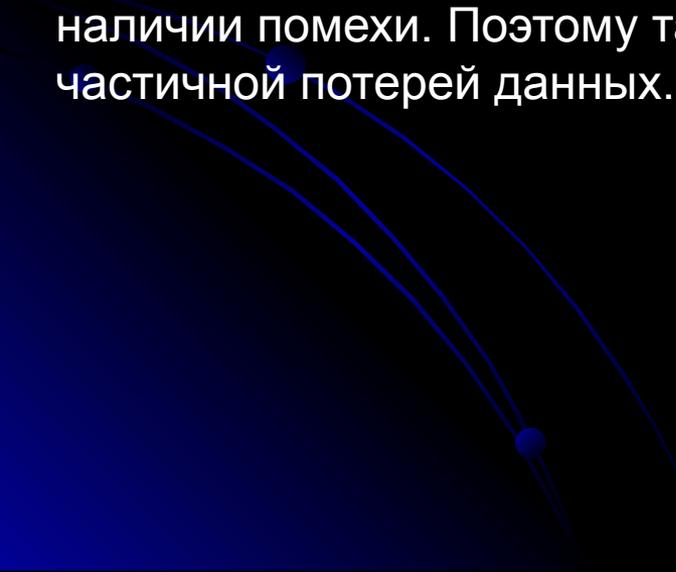
Варианты преобразования сообщений



1.2.3. Преобразование сообщений

Преобразование H_1 в H_2

Примеры: микрофон — звук преобразуется в электрические сигналы; магнитофон; телекамера; радио и телеприемник. Особенностью данного преобразования является то, что оно всегда сопровождается шумами (помехами), формируемыми самими информационными техническими устройствами, искажающими реальный информационный сигнал. Определить искажение не представляется возможным, т. к. реальный сигнал может принимать любые значения и по его величине нельзя судить о наличии помехи. Поэтому такое преобразование всегда сопровождается частичной потерей данных.



1.2.3. Преобразование сообщений

Преобразование Н1 в Д2

С математической точки зрения перевод сигнала из аналоговой формы в дискретную означает замену описывающей его непрерывной функции времени $Y(t)$ на некотором интервале $[t_1; t_2]$ конечным множеством (массивом) $\{Y_i, t_i\}$ ($i = \overline{0, n}$, где n — количество точек разбиения временного интервала). Подобное преобразование называется дискретизацией непрерывного сигнала и осуществляется посредством двух операций: развертки по времени и квантования по величине сигнала.

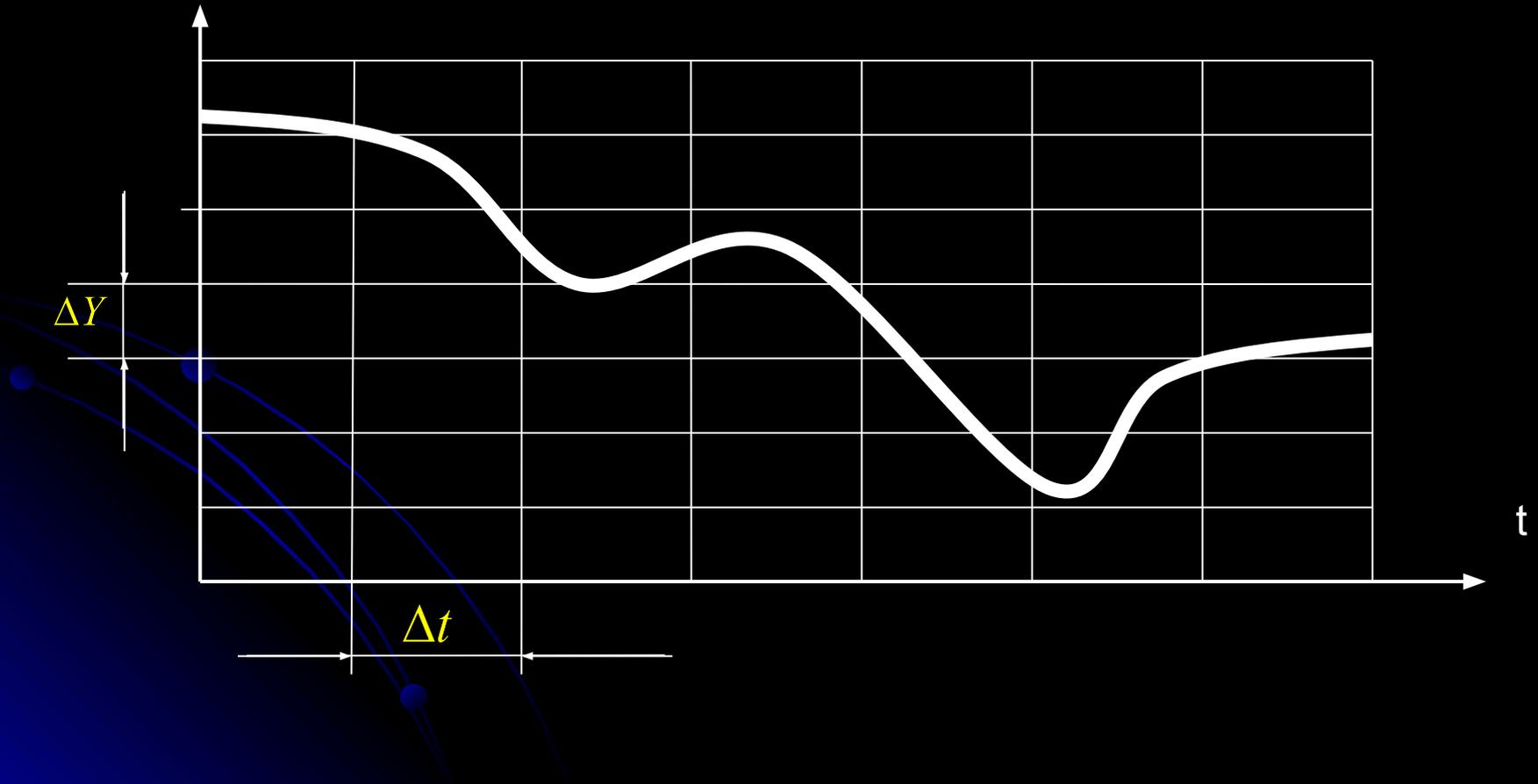


1.2.3. Преобразование сообщений

Преобразование Н1 в Д2

Развертка во времени и квантование по величине сигнала

Схема



1.2.3. Преобразование сообщений

Преобразование Н1 в Д2

Развертка во времени

Развертка по времени состоит в том, что наблюдение за значением величины Y производится не непрерывно, а лишь в определенные моменты времени с интервалом

$$\Delta t = \frac{t_2 - t_1}{n}$$

Теорема Котельникова В. А.:

Непрерывный сигнал можно полностью отобразить и точно воссоздать по последовательности измерений или отсчетов величины этого сигнала через одинаковые интервалы времени, меньшие или равные половине периода максимальной частоты, имеющейся в сигнале.

$$\Delta t \leq \frac{1}{2f_{\max}}$$

1.2.3. Преобразование сообщений

Преобразование Н1 в Д2

Квантование по величине

Квантование по величине — это отображение вещественных значений параметра сигнала в конечное множество чисел, кратных некоторой постоянной величине — шагу квантования (ΔY).

Выбор величины квантования:

Любой получатель данных — человек или устройство — всегда имеют конечную предельную точность распознавания градаций сигнала по уровню. Например, человеческий глаз в состоянии различить около 16 млн. цветовых оттенков, т. е. при квантовании нет смысла делать большее число градаций. При передаче речи точность воспроизведения градации около 1%. Следовательно, величину градации звуковых колебаний выбирают из соотношения $\Delta Y = 0,001 \cdot Y_{\max}$, а алфавит для обозначения всех градаций громкости должен содержать 100 знаков.

1.2.3. Преобразование сообщений

Преобразование D1 в H2

Получение непрерывного сигнала из дискретного широко применяется в вычислительной технике, например, аудио система компьютера. Когда цифровой звук передается на колонки необходимо произвести обратное преобразование, чтобы человек что-то услышал. Данный процесс осуществляется при помощи специальных микросхем ЦАП (цифроаналоговых преобразователей), принцип работы которых состоит в периодическом намагничивании и размагничивании конденсаторов в зависимости от значения дискретного сигнала. Наличие в ЦАП катушек индуктивности позволяет добиться непрерывного (плавного) изменения сигнала на выходе такого устройства. Если оцифровка сигнала была осуществлена правильно, то при помощи обратного преобразования будет получен сигнал практически идентичный исходному (изменения сигнала не будут восприниматься человеком).

1.2.3. Преобразование сообщений

Преобразование Д1 в Д2

Преобразование состоит в переходе при представлении сигналов от одного алфавита к другому — такая операция носит название перекодировка и может осуществляться без потерь.

Достоинства:

- высокая помехоустойчивость;
- простота и, как следствие, надежность и относительная дешевизна устройств по обработке данных;
- точность обработки, которая определяется количеством обрабатываемых элементов и не зависит от точности их изготовления;
- универсальность устройств.

1.2.4. Понятие данных

Хранимые на собственном материальном носителе получателя результаты регистрации сигналов называют **данными**.

Сообщение — это информационный поток, который в процессе передачи данных поступает к принимающему его субъекту.

Информативным называется **сообщение**, которое пополняет знания человека, т. е. несет для него информацию.

Схема процесса обработки данных



1.2.4. Понятие данных

При регистрации, воспроизведении и обработке данных получатель использует естественные, аппаратные и программные методы.

Естественные методы основаны на использовании естественных возможностей человека (органов чувств и логического мышления). Они предполагают использование данных в привычном для человека виде — аналоговом.

Аппаратные методы и устройства — регистрируют сигналы и преобразуют данные к форме, доступной для естественных методов. В основном технические средства работают с дискретными сигналами. Поэтому и возникает необходимость в обеспечении взаимодействия человека и машины.

Программные методы осуществляют воспроизведение и обработку данных в соответствии с заранее предписанными условиями посредством компьютера, в котором сочетаются все три метода работы с сигналами и данными.

1.2.5. Свойства информации

Динамичность. Информация существует ровно столько, сколько продолжается взаимодействие данных об объекте с методами их обработки. Как только это взаимодействие завершается, опять получают данные, но представленные в другой форме.

Важность (ценность, полезность). Информация может быть важной для собеседников или представлять известные им сведения — в этом проявляется субъективность в оценке информации. Цена информации определяется связанными с ней материальными и духовными приобретениями или потерями.

Адекватность. Степень соответствия создаваемого посредством полученной информации образа содержанию оригинала (реального объекта, процесса или явления). Пример: информация об удешевлении иномарок в четыре раза на территории Московской области преподнесенная 1 апреля вполне соответствует оригиналу — первоапрельской шутке. Хотя в любой другой день она является ложной и не адекватной оригиналу.

Достоверность. Степень соответствия содержания информации реальному объекту. Это свойство связано с истинностью или ложностью информации.

Точность. Степень близости информации к описанию реального состояния объекта, процесса или явления.

Актуальность. Степень соответствия информации текущему моменту времени или степень сохранения ценности информации к моменту ее использования.

1.2.6. Меры информации

Информация → сообщение → информативность сообщения → единица измерения информации → информационный объем сообщения.

Количество информации I можно оценить величиной неопределенности, устраняемой в результате принятия сообщения.

Пусть существует объект L , о котором в настоящий момент времени получатель не имеет сведений. Тогда его неопределенность $H(L)$ равна полному незнанию об источнике данных. С получением сообщения часть неопределенности об источнике сообщения снята, но какое-то незнание осталось — $H_a(L)$. Тогда количество информации в сообщении можно измерить как разность между неопределенностью об источнике до сообщения и неопределенностью, оставшейся после получения сообщения: $I(L) = H(L) - H_a(L)$.

Мера неопределенности об источнике сообщений получила название энтропия.

1.2.6. Меры информации

Энтропия является мерой неопределенности опыта, в котором проявляются случайные события, и равна средней неопределенности всех возможных его исходов.

Для практики важно иметь возможность произвести численную оценку энтропии разных опытов. Пусть опыт имеет n равновероятных исходов. Исходы равновероятны, если ни одно из них не имеет преимущества перед другими. Очевидно, что энтропия данного опыта зависит от величины n , т. е. является функцией числа возможных равновероятных исходов $f(n)$, обладающая свойствами:

1. $f(1)=0$, поскольку возможный исход не является случайным.
2. $f(n)$ возрастает с ростом числа возможных исходов, поскольку чем больше число возможных исходов, тем трудней предугадать результат опыта.
3. $f(na \wedge nb) = f(na) + f(nb)$ (свойство аддитивности). Пусть проводятся два (или больше) независимых опыта a и b с количествами равновероятных исходов в каждом соответственно na и nb . Очевидно, мера суммарной неопределенности опытов $a \wedge b$ равна сумме неопределенностей.

1.2.6. Меры информации

Таким образом, энтропия должна описываться функцией, обладающей тремя необходимыми свойствами. Такому набору свойств соответствует лишь логарифмическая функция, причем ее основание не имеет значения. На практике выбирают основание 2, так как это оказывается удобным при принятии единицы информации бит — при $n=2$ равновероятных исходах $\log_2 n=1$ бит.

Формула Р. Хартли

Энтропия опыта с равновероятными исходами — $H(a) = \log_2 n$.

Энтропия одного исхода в опыте с возможными равновероятными исходами:

$$H(a) = \frac{1}{n} \log_2 n = -\frac{1}{n} \log_2 \frac{1}{n} = -p \log_2 p$$

где $p = \frac{1}{n}$ вероятность любого из отдельных исходов.

Формула К. Шеннона

Обобщая это выражение на опыт с не равновероятными исходами $A_1 \dots A_n$ получим:

$$H(a) = -\sum p(A_i) \log_2 p(A_i)$$

1.2.6. Меры информации

Свойства энтропии

1. $H(a)$ — вещественная и неотрицательная величина (знак минус учитывает, что значение величины под логарифмом меньше единицы).
2. Энтропия не учитывает содержательную сторону информации, т. е. ее ценность.
3. Энтропия нескольких независимых источников сообщений равна сумме энтропии каждого из них: $H(ab) = H(a) + H(b)$.
4. Энтропия — величина ограниченная. При прочих равных условиях наибольшую энтропию имеет опыт с равновероятными исходами.

1.2.6. Меры информации

Алфавитный подход

Алфавитный (кибернетический) подход — это единственный способ измерения информации, циркулирующей в информационной технике (компьютерах). Здесь речь идет об измерении информации в тексте (символьном сообщении), составленном из символов некоторого алфавита. содержанию текста (семантике сообщения) такая мера информации отношения не имеет. Поэтому такой подход можно назвать объективным, т. е. не зависящим от воспринимающего его субъекта.

Алфавит — это конечное множество символов, используемых для представления информации.

Число символов в алфавите называется **мощностью алфавита**.

1.2.6. Меры информации

Информационный вес символа

В предположении, что вероятности появления каждого символа алфавита одинаковы (что на практике далеко не так), количество информации, которое несет в тексте каждый символ (i) вычисляется из уравнения Хартли:

$$2^i = N$$

где N — мощность алфавита.

Информационный вес символа английского алфавита $i = \log_2 27 \approx 4,755$

Информационный вес символа русского алфавита $i = \log_2 34 \approx 5,087$

Количество информации в тексте

Количество информации во всем тексте, состоящим из K символов, равно произведению информационного веса символа на их количество:

$$I = iK$$

Величина I — информационный объем текста.

1.2.6. Меры информации

Единица измерения информации

Минимальная мощность алфавита, пригодного для передачи информации, равна 2. Такой алфавит называют двоичным. Информационный вес символа в двоичном алфавите $i = 1$, так как $2^i = 2$. Эта величина получила название бит. Один символ двоичного алфавита несет 1 бит информации.

$$1 \text{ байт} = 8 \text{ бит}$$

$$1 \text{ килобайт} = 2^{10} = 1024 \text{ байт}$$

$$1 \text{ мегабайт} = 2^{10} = 1024 \text{ килобайт}$$

$$1 \text{ гигабайт} = 2^{10} = 1024 \text{ мегабайт}$$

$$1 \text{ терабайт} = 2^{10} = 1024 \text{ гигабайт}$$

1.2.7. Экономическая информация

Экономическая информация — это та информация, которая возникает в процессе производственно-хозяйственной деятельности и используется при подготовке и принятии управленческих решений. Это совокупность сведений, отражающих социально-экономические процессы и служащих для управления этими процессами и коллективами людей в производственной и непроизводственной сферах.

Виды информации



1.2.7. Экономическая информация

Особенности экономической информации:

- она дискретна, т. е. характеризует состояние объекта, процесса или явления либо на определенный момент времени, либо за определенный интервал времени;
- представляется преимущественно в цифровой форме;
- характеризуется большим объемом, как переменных, так и условно постоянных данных;
- имеет строго определенные сроки своего представления;
- обладает способностью к преобразованию, агрегированию по определенным признакам;
- характеризуется возможностью использования одних и тех же данных, в том числе и у разных потребителей.

1.2.7. Экономическая информация

Классификация экономической информации:

1. **По месту возникновения** — входная и выходная, внутренняя и внешняя. Входная информация — поступает в организацию извне. Выходная — информация из организации. Внутренняя информация возникает внутри объекта управления, а внешняя — за его пределами.
2. **По стадии обработки** — первичная, вторичная, промежуточная, результатная. Первичная — возникает непосредственно в процессе деятельности объекта и регистрируется на начальной стадии. Вторичная — результат обработки первичной и может быть промежуточной или результатной. Промежуточная используется в качестве исходных данных для дальнейших расчетов. Результатная — используется для выработки управленческих решений.
3. **По способу отображения** — текстовая или графическая.
4. **По признаку стабильности** — переменная (текущая) или постоянная (условно).
5. **По функциям управления** — плановая, нормативно-справочная, учетная, оперативная.