

Теория информации и кодирования

Введение в теорию информации и кодирования

*Профессор: Головань Вячеслав
Григорьевич*

Учебная карта курса

Виды учебной работы	Семестр
	4
Общая трудоемкость дисциплины	60
Аудиторные занятия	36
Самостоятельная работа	24
Реферат	+
Вид итогового контроля	Зачет

Рекомендуемая литература (1)



Рекомендуемая литература

(2)

- Цымбал В.П. Теория информации и кодирование. – М.: Высш. шк., 1986.
- Кловский Д.Д. Теория передачи сигналов. – М.: Связь, 1984.
- Стратонович Р.Л. Теория информации, М.: Советское радио, 1975, 424 с.
- Кудряшов Б.Д. Теория информации. Учебник для вузов Изд-во ПИТЕР, 2008. – 320 с.
- Цымбал В.П. Теория информации и кодирование. – М.: Высш. шк., 1986.

Вопросы для аудитории

Что вы знаете о теории информации и кодирования?

Считаете ли вы, что данная дисциплина является важной для вас? Почему?

Могут ли знания, полученные в ходе изучения дисциплины, помочь вам в достижении ваших целей? Если да, то каких именно?

Высказывания великих о теории информации

- Значение теории информации было, возможно, преувеличено и раздуто до пределов, превышающих ее реальные достижения ... Сейчас теория информации, как модный опьяняющий напиток, кружит голову всем вокруг.
- Сознавая, что теория информации является сильным средством решения проблем теории связи, нельзя забывать, что она не является панацеей для инженера-связиста и тем более для представителей всех других специальностей.
- Представителям различных наук следует ясно понимать, что основные положения теории информации касаются очень специфического направления исследования, направления, которое совершенно не обязательно должно оказаться плодотворным в психологии, экономике и в других социальных науках.

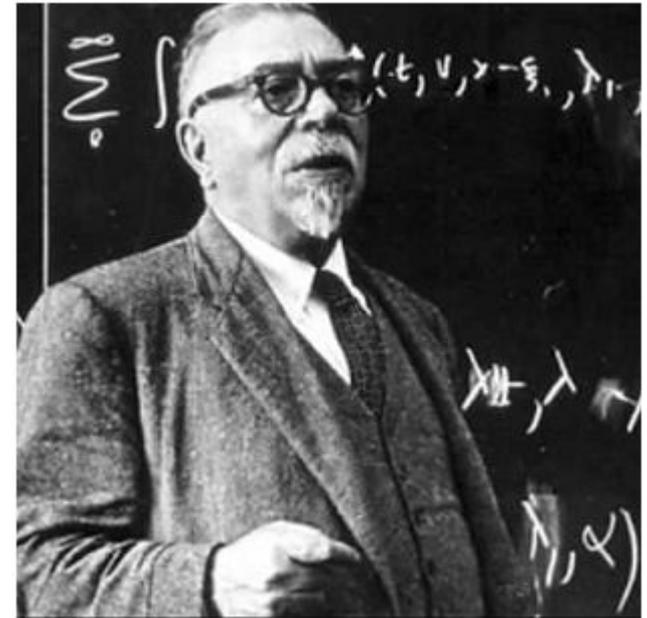
Высказывания К. Шеннона взяты из статьи "Бандвагон", 1963

Теория информации как часть кибернетики

Теория информации рассматривается как существенная часть кибернетики.

Кибернетика - это наука об общих законах получения, хранения, передачи и переработки информации.

Основными разделами (они фактически абсолютно самостоятельны и независимы) современной кибернетики считаются: *теория информации*, теория алгоритмов, теория автоматов, исследование операций, теория оптимального управления и теория распознавания образов.



Норберт Винер (*Norbert Wiener*) американский учёный, выдающийся математик и философ, основоположник кибернетики и теории искусственного интеллекта. 1894-1964

Немного истории. Становление теории информации

В настоящее время теория информации (в широком смысле) включает:

- теорию распознавания статистических свойств источников сообщений
- теорию кодирования источников сообщений (сжатия данных),
- теорию передачи сообщений по каналам связи,
- теорию помехоустойчивого кодирования,
- **теорию поиска информации и криптографию.**

Работой, заложившей основу математической теории информации, является статья К. Шеннона **«Математическая теория связи»**, опубликованная в 1948 году.

В этой статье впервые было введено понятие энтропии (количества информации) и были указаны основные задачи и направления развития теории.

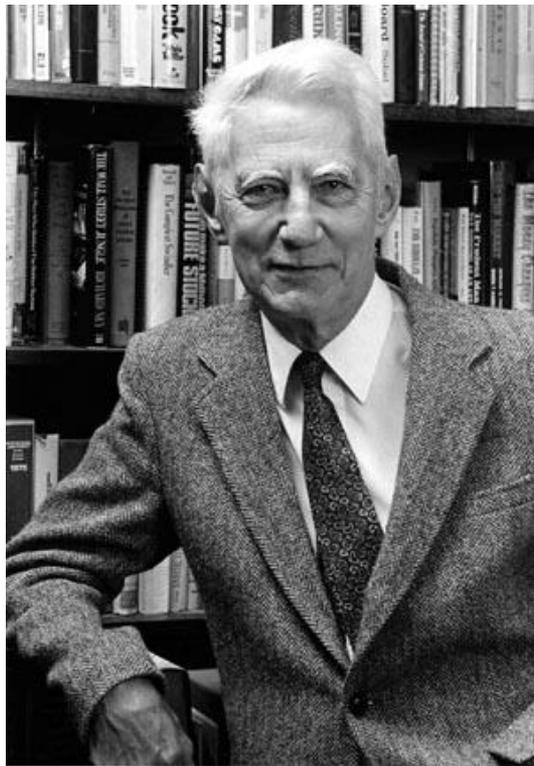
Теория информации

Возникновение теории информации связывают обычно с появлением фундаментальной работы американского ученого К. Шеннона «Математическая теория связи» (1948).

Однако в теорию информации органически вошли и результаты, полученные другими учеными.

Например, Р. Хартли, впервые предложил количественную меру информации (1928), акад. В. А. Котельников, сформулировал важнейшую теорему о возможности представления непрерывной функции совокупностью ее значений в отдельных точках отсчета (1933) и разработал оптимальные методы приема сигналов на фоне помех (1946). Акад. А. Н. Колмогоров, внес огромный вклад в статистическую теорию колебаний, являющуюся математической основой теории информации (1941).

Клод Шеннон. Вклад в развитие теории информации



- Клод Элвуд Шеннон – первый ученый взялся за решение проблем передачи информации по техническим каналам связи и создал новую для того времени науку – теорию информации
- Определил способ измерения количества информации, передаваемой по каналам связи.
- Им введено понятие пропускной способности канала как максимально возможной скорости передачи информации.
- Скорость передачи информации измеряется

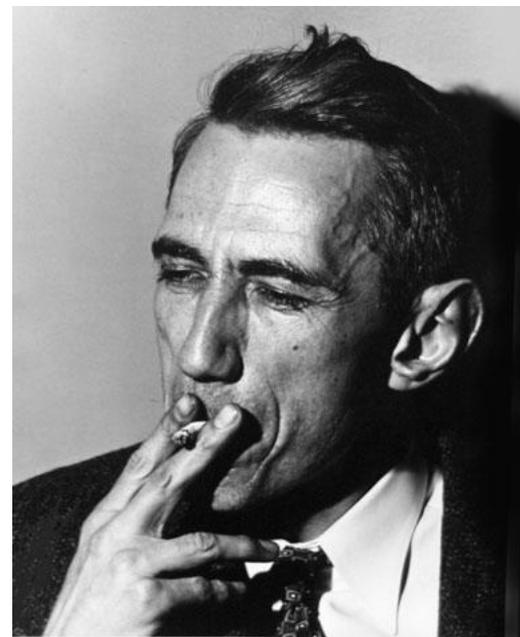
Клод Элвуд Шеннон (Shannon) (1916 – 2001) – американский инженер и математик. Основатель современных теорий информации (в битах в секунду, в килобитах в секунду, в мегабитах в секунду)

Клод Шеннон. Вклад в развитие криптографии

Секретный доклад "математическая теория криптографии", датированного 1 сентября 1945 года (1949 году опубликован в техническом журнале корпорации Bell System).

В статье были впервые сформулированы теоретические основы криптографии, понятия абсолютной секретности систем, доказано существование абсолютно стойких, невскрываемых шифров, условия, необходимые для этого.

Шеннон определил основные принципы, которым должны соответствовать надежные шифры. Ввел в рассмотрение понятия перемешивания и рассеивания, предложил строить стойкие криптографические системы из относительно несложных преобразований



Клод Шеннон. Вклад в развитие теории связи

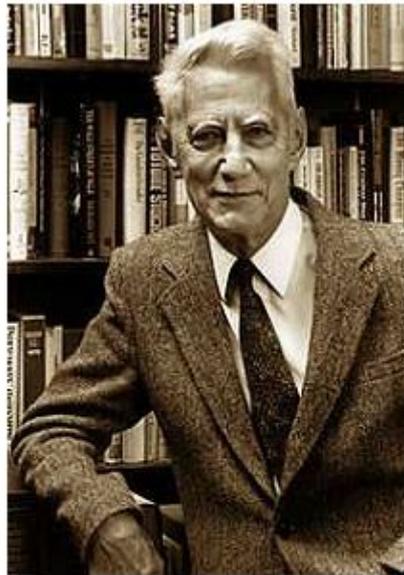
Разработанные Шенноном К. методы обработки и оценки сообщений, получили широкое применение при передаче информации по техническим каналам связи.

Разработанная теория информации позволила получить количественную оценку информационных сообщений, дала возможность определять и рассчитывать необходимые параметры передающих средств в зависимости от объема информации, и положила начало новым методам обработки сообщений, кодированию и сжатию информации.

Основоположники теории информации



Ральф Винтон Лайон Хартли (*Ralph Vinton Lyon Hartley*) американский учёный-электронщик, 1888-1970.



Клод Элвуд Шеннон (*Claude Elwood Shannon*) американский инженер и математик 1916-2001



Андрей Николаевич Колмогоров советский математик, один из основоположников современной теории вероятностей 1903-1987



Виктор Михайлович Глушков, выдающийся советский математик и кибернетик 1923-1982

Качественный VS количественный подход к информации

Существуют и альтернативные теории информации, например "Качественная теория информации", предложенная польским учёным М. Мазуром (в противовес «Количественной теории информации» которой по сути является теория, разработанная К. Шенноном)



Мариан Мазур, ученый, польский специалист по теории информации, участвующих в *elektrotermia* и кибернетики, создатель польской школы кибер-преступников. 1909-1983.

А как применить понятие "количество информации", например к географической карте? Ведь карта содержит самую различную информацию. ... О каких вероятностях здесь может идти речь? Ведь каждый элемент карты, как и каждый элемент территории, существует, а не "происходит" с какой-то вероятностью. На заданные вопросы можно ответить, что теория информации создана не для этих потребностей. Однако такой ответ означает признание того факта, что созданная теория до сих пор дает меньше, чем обещает ее название.

Основная задача теории информации и кодирования

Основная задача ТИК: обеспечить **надежную и эффективную** передачу информации от источника к приемнику при условии, что передаче препятствуют помехи.

"**надежную**" — означает, что в процессе передачи не должно происходить потери информации;

«эффективная» означает, что передача должна осуществляться наиболее быстрым способом, поскольку время эксплуатации линии связи – экономический фактор, который требуется минимизировать; помехи присутствуют в любой реальной линии связи.

Частные задачи теории информации:

- измерение количества информации;
- изучение свойств информации;
- разработка оптимальных методов кодирования;
- разработка методов передачи информации, обеспечивающих максимальную пропускную способность каналов связи при наличии помех, пр.

Связь теории информации с другими дисциплинами (1)

Теория информации тесно связана с такими разделами математики как теория вероятностей и математическая статистика, а также прикладной алгеброй, которые предоставляют для нее математический фундамент. С другой стороны теория информации исторически и практически представляет собой математический фундамент теории связи.

Теория информации и информационная безопасность

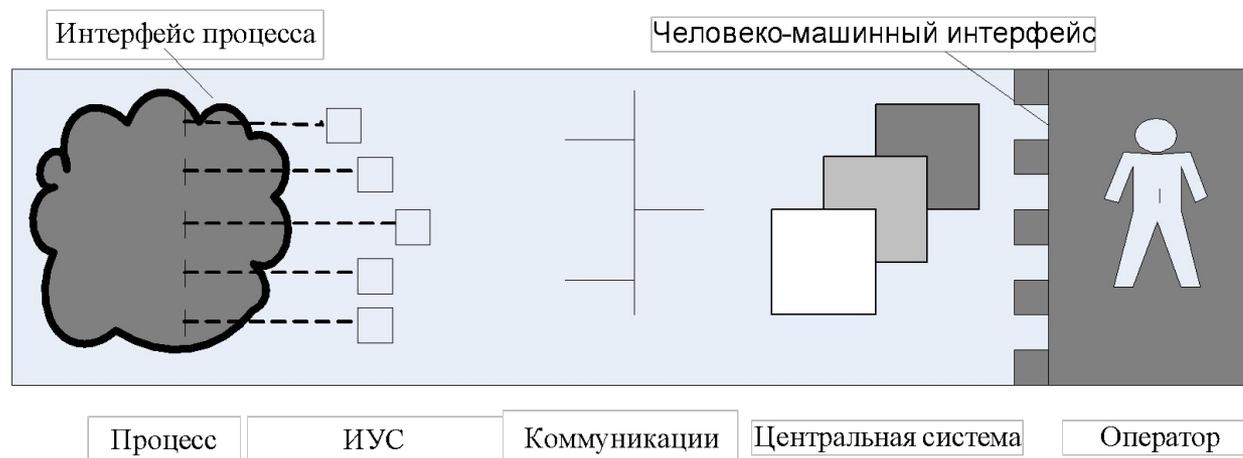
Теория информации изучает **теорию поиска информации и криптографию**. Криптография как метод защиты информации от умышленных деструктивных воздействий: криптографическое преобразование информации и реализованные на его основе криптографические протоколы.

Связь теории информации с другими дисциплинами (2)

Теория информации и информационные системы.

Информационная система (ИС) - автоматизированная система, предназначенная для организации, хранения, пополнения, поддержки и представления пользователям информации в соответствии с их запросами.

Пример информационно-управляющих систем



Понятие информации

Что такое информация?



Этимология понятия «информация»

в XIX веке: «информация» - **учение, наставление**,
«информатор» - домашний учитель

Информация – от слов «ин»-в, и «форма»-образ, вид, т. е. в смысле того, что вносит форму.

**Активн
ость**

до конца 40-х гг XX в.: под информацией понимались **сведения, передаваемые людьми** устным, письменным или другим способом.

В наст. время: информация – от латинского informatio - ознакомление, разъяснение, изложение.

**Содер
жание**

Перевод «message» - «**сообщение**»

Понятие информации

ИНФОРМАЦИЯ от латинского «informatio» «сведения, разъяснения, изложение»

ИНФОРМАЦИЯ есть отражение реального мира _____

ИНФОРМАЦИЯ есть сведения, являющиеся объектом переработки

ИНФОРМАЦИЯ — сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределенности, неполноты знаний.

Представление информации — способ, с помощью которого происходит отображение информации, ее внешняя форма.

Переход от представления к абстрактной информации называется *интерпретацией*. Этот переход может быть только воображаемый, мыслимый.

В материальном мире все физические объекты (тела или поля), взаимодействуя друг с другом, порождают **сигналы (сообщения)** различных типов



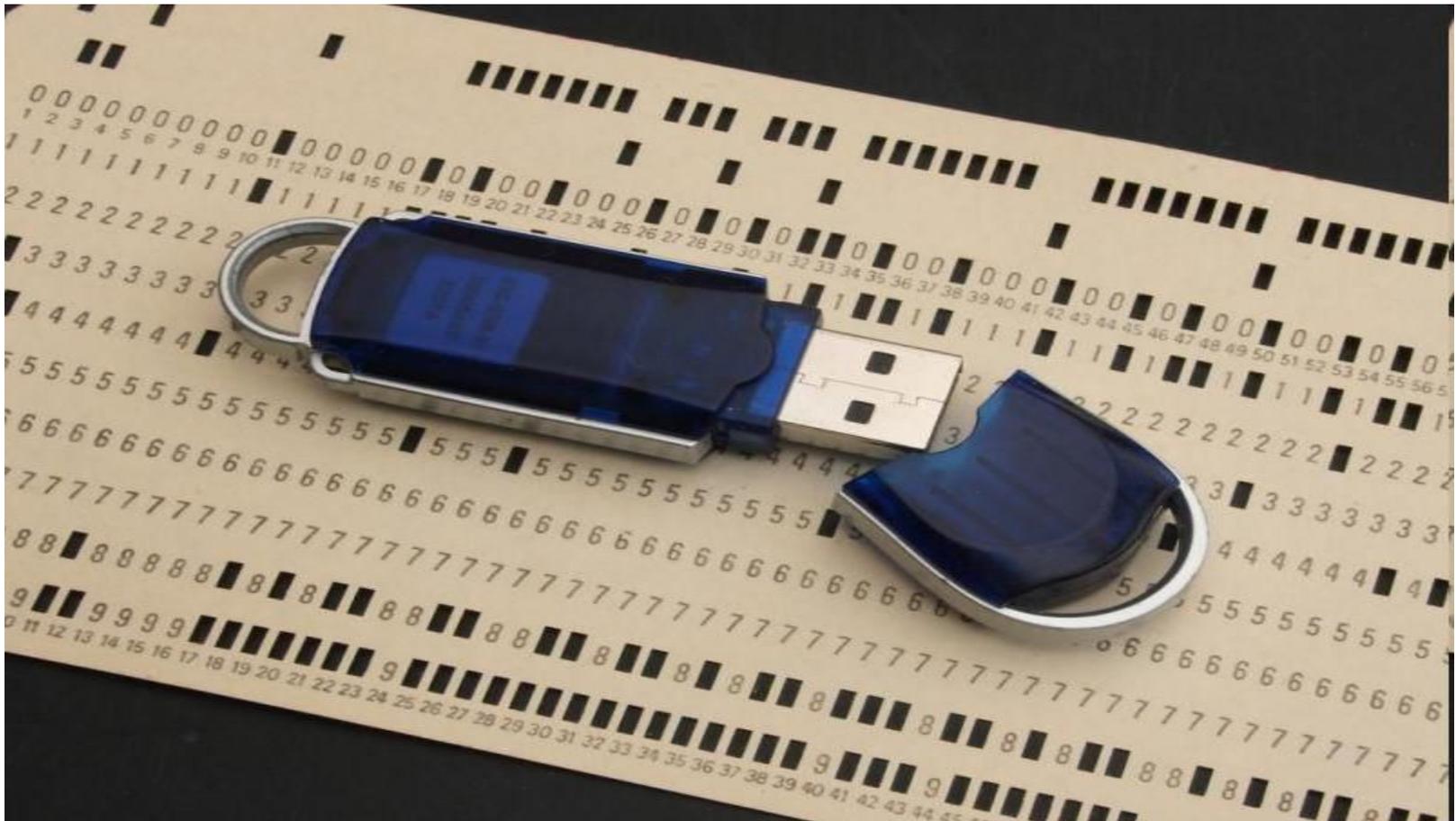
Понятие информации

Регистрация сигнала – это порождение в физических телах изменения свойств



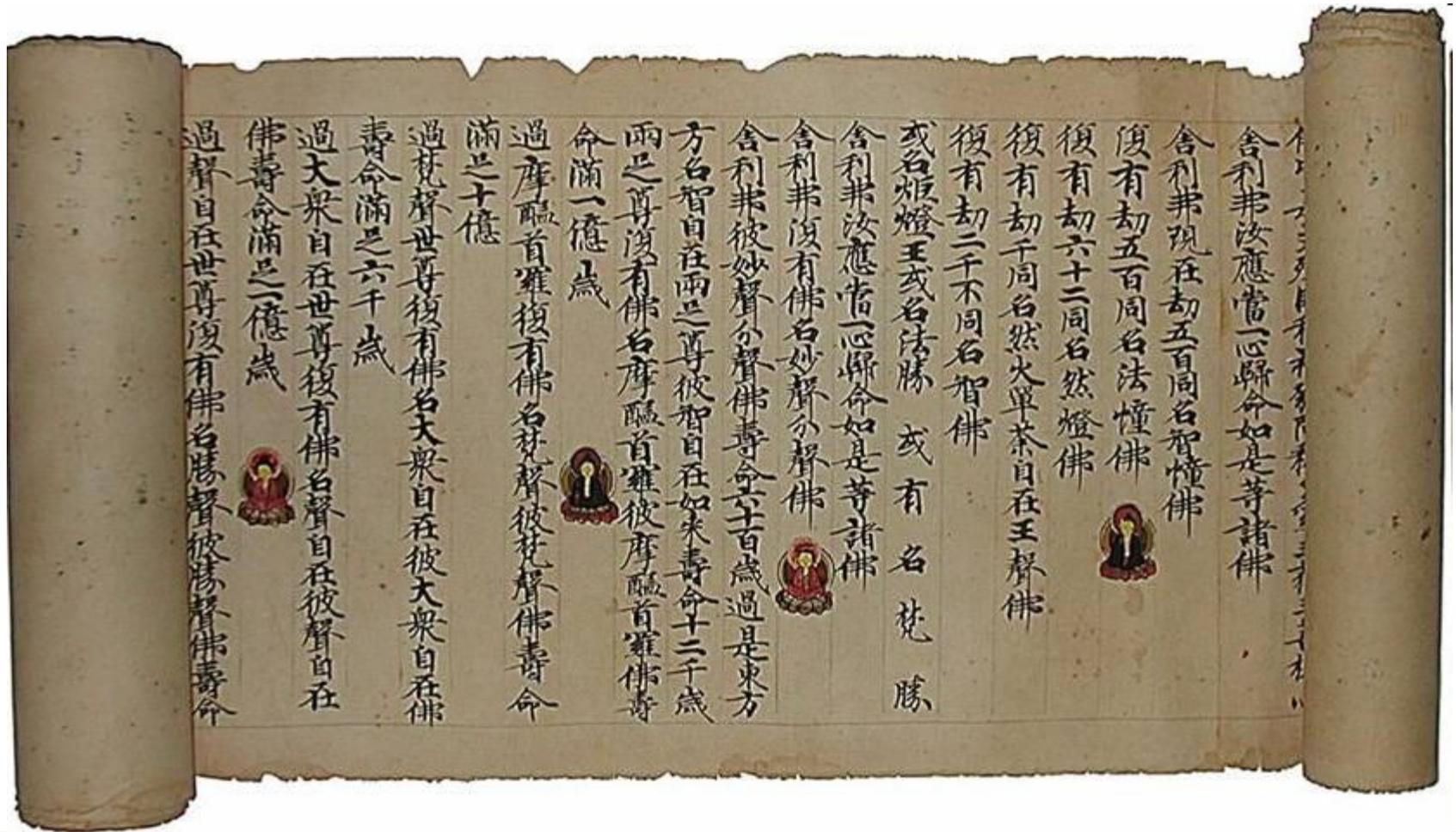
Понятие информации

Данные - это запись в соответствующем коде наблюдения, факта, объекта, пригодная для коммуникации, интерпретации, обработки и получения новой информации



Понятие информации

Физический метод регистрации сигналов



Информация в жизни человека

Память

- I Письменность (6 тыс. лет назад)
- II Печатный станок (XVI век)
- III Радио, телеграф (конец XIX века)
- IV ЭВМ, компьютеры (середина XX века)

Информация в жизни человека



Материальные

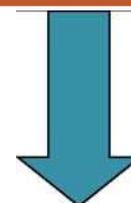
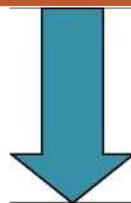
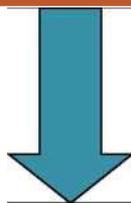
Энергетические

Финансовые

ресурсы

ресурсы

ресурсы



**Информационные
ресурсы**

Информация в жизни человека

Информационные ресурсы — это отдельные документы и массивы документов в библиотеках, архива) фондах, банках данных, информационных системах и других хранилищах.

Информационные ресурсы — это знания, подготовленные людьми для социального использования обществе и зафиксированные на материальных носителях.

Информационные ресурсы страны, региона, организации все чаще рассматриваются как стратегические ресурсы, аналогичные по значимости запасам сырья, энергии, ископаемых прочим ресурсам.



Свойства информации

Свойства:

объективные (не зависящие от потребителя);

субъективные (зависящие от потребителя).

Информация **объективна** в силу объективности данных, и **субъективна**, в силу субъективности применяемых потребителем методов.

Качество информации — совокупность *свойств* информации, характеризующих степень ее соответствия потребностям (целям, ценностям) пользователей (средств автоматизации, персонала...)

Свойства информации, отражающие ее качество делят:

внутренние — **содержательность** (присущее собственно информации и сохраняющееся при ее переносе в другую систему);

внешние — **защищенность** (присущее информации, находящейся или используемой только в определенной системе).



Свойства информации

Содержательность (внутреннее)

Значимость

- полнота
- адекватность
- идентичность
- актуальность

Кумулятивность

- избирательность
- гомоморфизм

Защищенность (внешнее)

Достоверность

Сохранность

- целостность
- готовность

Конфиденциальность

- доступность
- скрытность
- имитостойкость

Информационные процессы

Информационный процесс – это процесс, в ходе которого изменяется содержание информации или форма ее представления.

Процессы, связанные с поиском, хранением, передачей, обработкой и использованием информации, называются информационными процессами.

К основным информационным процессам относятся сбор, хранение, обработка и передача информации.



Информационные процессы:

- ▶ Поиск информации

- ▶ Хранение информации
- ▶ Передача информации
- ▶ Получение информации
- ▶ Обработка информации
- ▶ Защита информации
- ▶ Использование информации



Поиск информации

Методы поиска:

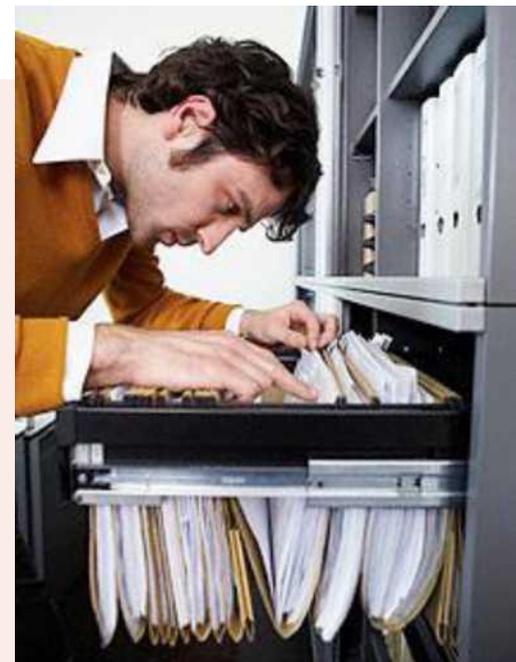
- ❖ непосредственное наблюдение;
- ❖ общение со специалистами по интересующему вас вопросу;
- ❖ чтение соответствующей литературы;
- ❖ просмотр видео, телепрограмм;
- ❖ прослушивание радиопередач;
- ❖ работа в библиотеках, архивах;
- ❖ другие методы.



Информационные процессы

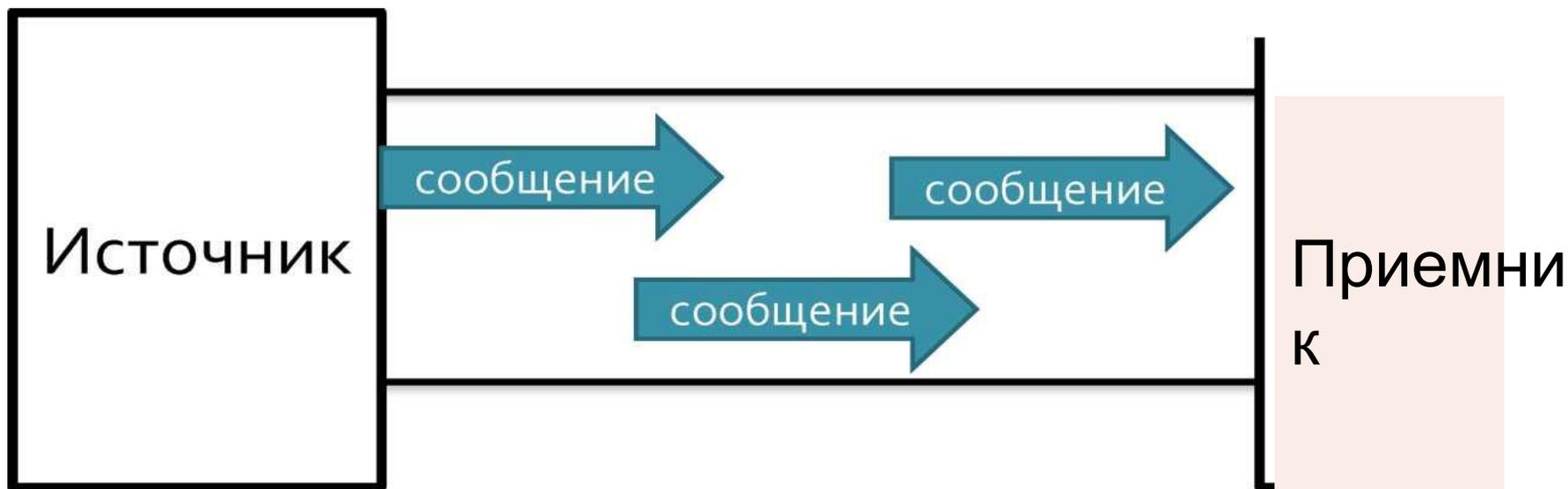
Сбор данных - это деятельность субъекта по накоплению данных с целью обеспечения достаточной полноты.

Информация = данные + адекватный метод



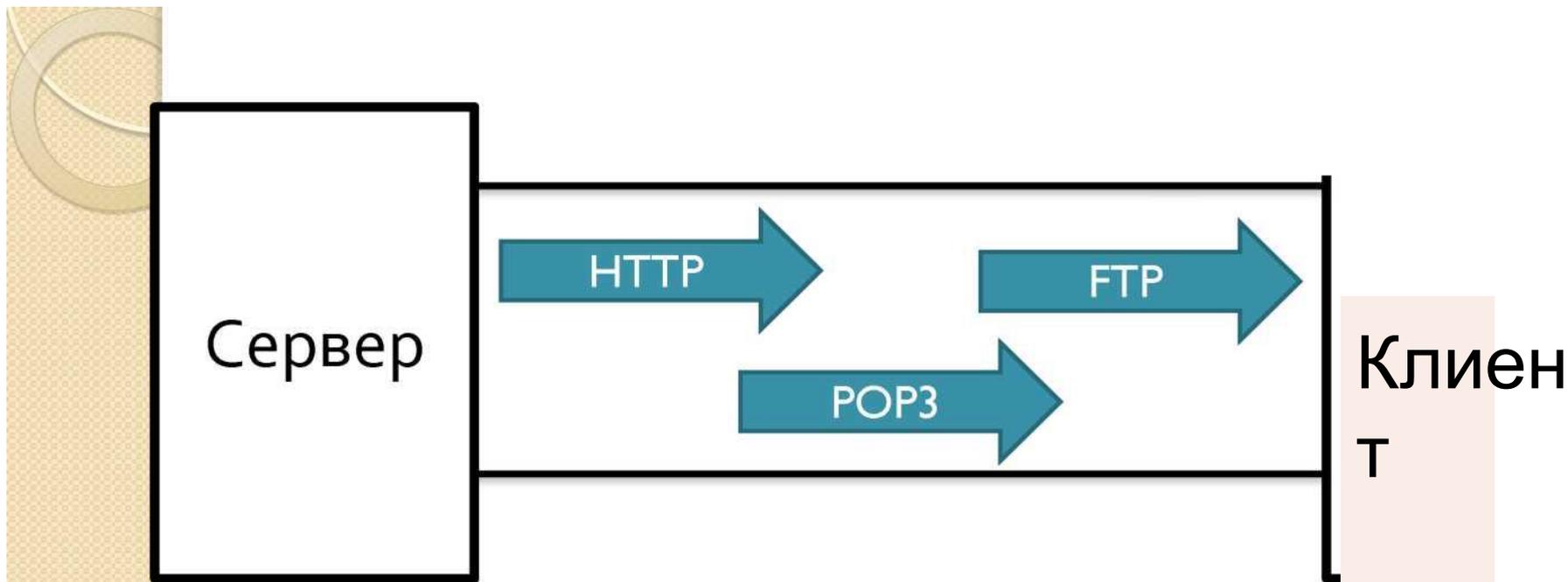
Информационные процессы

Передача данных - это процесс обмена данными.



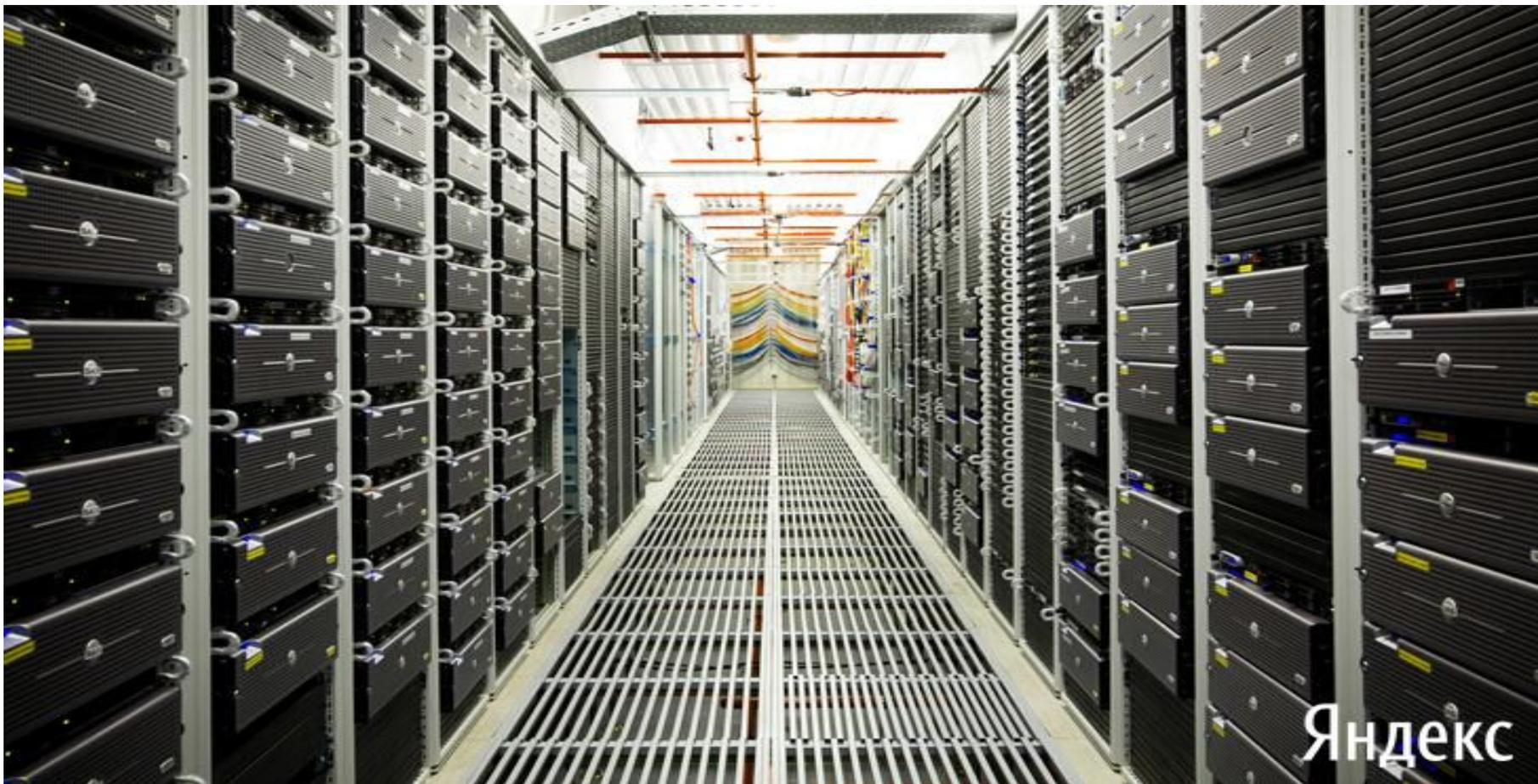
Сообщение — форма представления информации в виде совокупности знаков (символов), используемая для передачи.

Информационные процессы



Канал связи — это естественный или искусственный материальный объект, обеспечивающий передачу сигнала от передатчика к приемнику.

Информационные процессы



Хранение данных — это поддержание данных в форме, постоянно готовой к выдаче их потребителю.

Хранение информации

Сбор информации не является самоцелью. Чтобы полученная информация могла использоваться, причем многократно, необходимо ее хранить.

Хранение информации – это

способ распространения информации в пространстве и времени.

Способ хранения информации зависит от ее носителя.



Основные хранилища информации

Для человека

память

Для общества

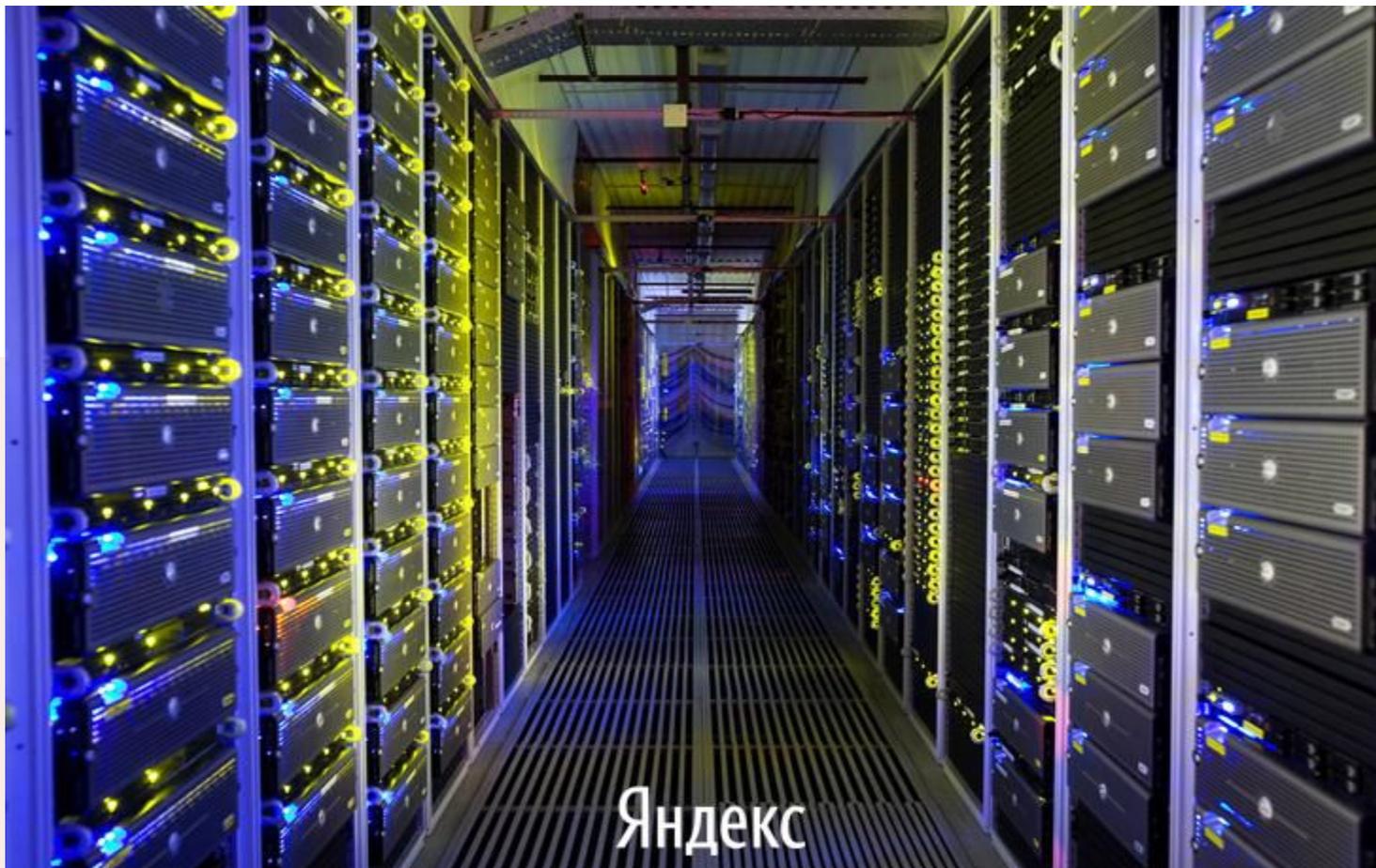
Библиотеки,
видеотеки,
фонотеки, архивы,
патентные бюро,
музеи,
картинные
галереи.

**Компьютерны
е
хранилища**

Базы и банки
данных,
информационно-
поисковые
системы,
электронные
энциклопедии,
медиатеки.



Информационные процессы



Обработка данных — это процесс преобразования информации от исходной ее формы до определенного результата.

Информационные процессы

Сообщение



Получателем сообщение рассматривается на трех уровнях:

- синтаксическом** — рассматриваются внутренние свойства сообщений, т. е. отношения между знаками, отражающие структуру данной знаковой системы;
- семантическом** — анализируются отношения между знаками и обозначаемыми ими предметами, действиями, качествами, т. е. смысловое содержание сообщения;
- прагматическом** — рассматриваются отношения между сообщением и получателем, т. е. потребительское содержание сообщения, его значимость получателя.

Основные определения теории информации и кодирования

Информация — нематериальная сущность, при помощи которой с любой точностью можно описывать реальные (материальные), виртуальные (возможные) и понятийные сущности.

Информация — противоположность неопределенности.

Информация (по Шеннону) — снятая неопределенность.

Информация имеет количественную оценку.

Сообщение — форма представления информации. Понятие сообщение = понятие случайной величины

Понятие последовательность сообщений = случайный процесс.



Основные определения теории информации и кодирования

Классификация информации

1. По форме представления (2 вида)

Различают две формы представления информации — непрерывную и дискретную.

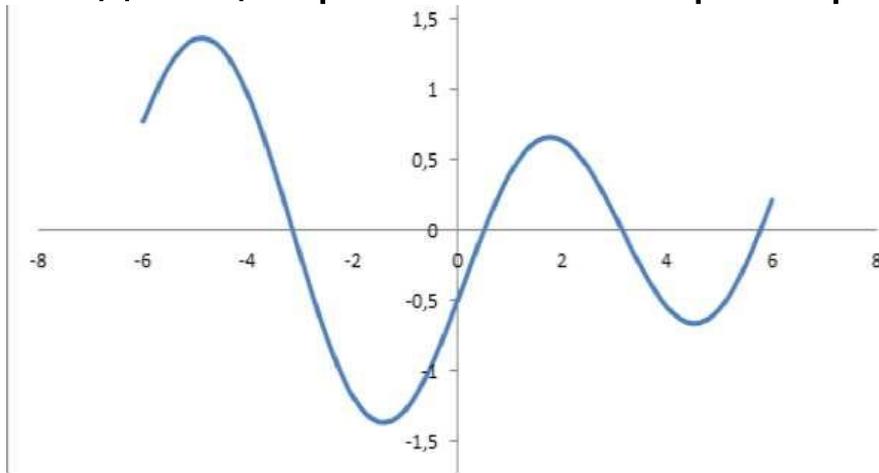
Дискретные сообщения формируются в результате последовательной выдачи источником сообщений отдельных элементов - знаков.

Множество различных знаков называют алфавитом источника сообщения, а число знаков - объемом алфавита.

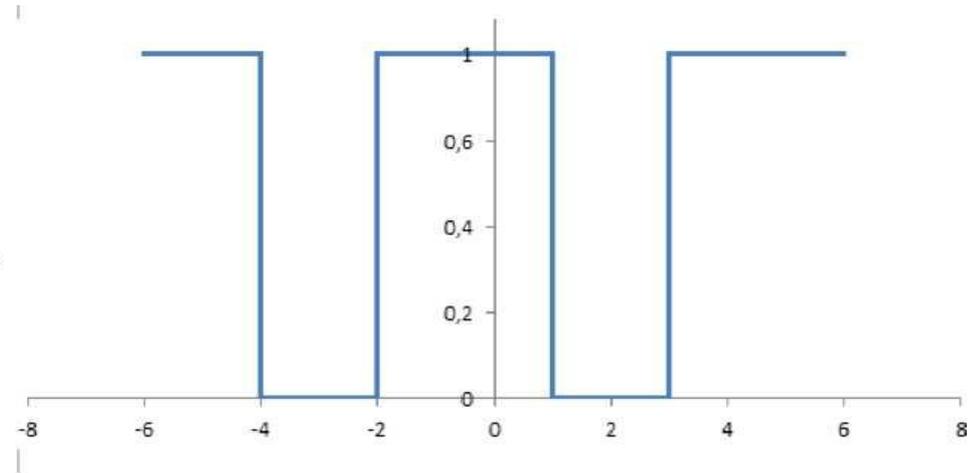
Непрерывные сообщения не разделены на элементы. Они описываются непрерывными функциями времени, принимающими непрерывное множество значений (речь, телевизионное изображение).

Основные определения теории информации и кодирования

Сигнал — это изменяющийся во времени физический процесс, обладающий различными характеристиками



а — непрерывный сигнал



б — дискретный сигнал

которого в заданном интервале времени можно отсчитать бесконечно большое число значений.

Дискретный сигнал в том же интервале времени имеет конечное число значений. Примером дискретного сигнала является импульсный, т.е. такой, длительность которого соизмерима с длительностью установления переходного процесса в системе, на вход которой он действует.

Основные определения теории информации и кодирования

Классификация информации

2. По области возникновения выделяют информацию:

- **механическую**, которая отражает процессы и явления неодушевленной природы;
- **биологическую**, которая отражает процессы животного и растительного мира;
- **социальную**, которая отражает процессы человеческого общества.

3. По способу передачи и восприятия различают следующие виды информации:

- **визуальную**, передаваемую видимыми образами и символами;
- **аудиальную**, передаваемую звуками;
- **тактильную**, передаваемую ощущениями прикосновений;
- **органолептическую**, передаваемую запахами и вкусами;
- **машинную**, выдаваемую и воспринимаемую средствами вычислительной техники.

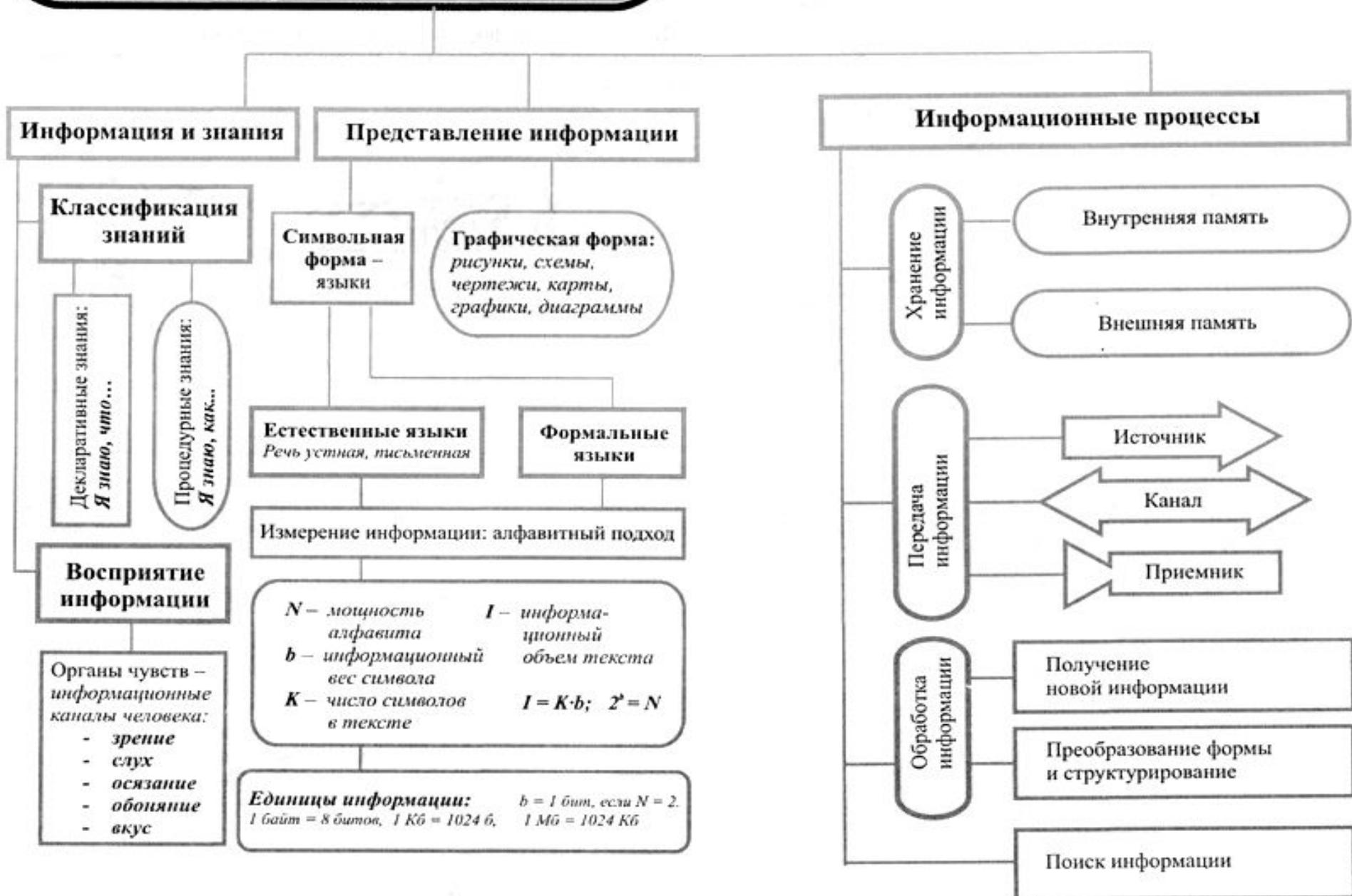
Основные определения теории информации и кодирования

Основные свойства информации:

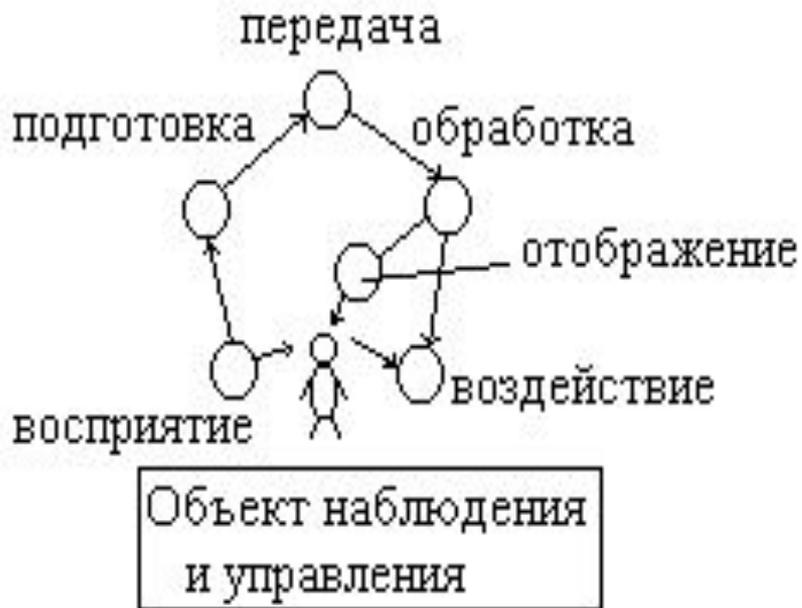
- **достоверность** (адекватность) – степень соответствия _____ реальному состоянию дел;
- **актуальность** – степень соответствия текущему моменту времени;
- **полнота** – достаточность для её понимания и принятия правильных решений или для создания новых данных на её основе;
- **доступность** – мера возможности получить нужные сведения;
- **понятность** – способ представления информации должен быть понятен её получателю
- **ценность** – степень важности для решения текущей задачи или дальнейшего применения в каких-либо видах деятельности человека.

ИНФОРМАЦИЯ

Система ОСНОВНЫХ ПОНЯТИЙ



Основные определения теории информации и кодирования



Этапы обращения информации

Можно выделить следующие этапы обращения информации:

- 1) восприятие информации;
- 2) подготовка информации;
- 3) передача и хранение информации;
- 4) обработка информации;
- 5) отображение информации;
- 6) воздействие информации.

Основные определения теории информации и кодирования

Упорядоченную последовательность взаимосвязанных действий, выполняемых в строго определенной последовательности с момента возникновения информации до получения заданных результатов, **называют технологией обработки информации**

Под информационной технологией следует понимать систему методов и способов сбора, накопления, хранения, поиска, обработки, анализа, выдачи данных, информации и знаний на основе применения аппаратных и программных средств в соответствии с требованиями, предъявляемыми пользователями.

Что такое неопределенность?



Неопределенность. Общее определение

НЕОПРЕДЕЛЕННОСТЬ в системе [systems uncertainty] — *ситуация*, когда полностью или частично отсутствует *информация* о возможных *состояниях системы* и *внешней среды*. Иначе говоря, когда в системе возможны те или иные непредсказуемые события (вероятностные характеристики которых не существуют или неизвестны).

Это неизбежный признак *больших (сложных) систем*; чем сложнее система, тем большее значение приобретает фактор неопределенности в ее *поведении* (развитии).

Можно ли измерить неопределенность?

Неопределенность по Шеннону

Неопределенность события – число возможных его исходов.



Где больше неопределенность? От чего она зависит?

От числа исходов и о вероятности каждого из них

Что такое данные?

Основные определения теории информации и кодирования

Данные могут рассматриваться как признаки или записанные наблюдения, которые по каким-то причинам не используются, а только хранятся. В том случае, если появляется возможность использовать эти данные для уменьшения неопределенности о чем-либо, данные превращаются в информацию.

Шум — это помехи в канале связи при передаче информации.

Форма представления информации, имеющая признаки начала и конца, предназначенная для передачи через среду связи называется **сообщением**.

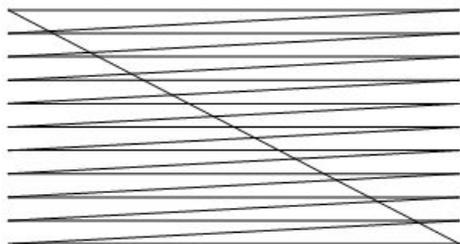
Сообщением также является форма предоставления информации как совокупность первичных сигналов или знаков, содержащих информацию.

Основные определения теории информации и кодирования

Примеры сообщений:

Сообщения

- В системе **черно-белого телевидения** сообщение - последовательность кадров, каждый кадр - последовательность значений яркости, упорядоченных согласно схеме телевизионной развертки
- В **телефонии** сообщение – непрерывная последовательность значений изменяющегося во времени звукового давления на мембрану микрофона:



Основные определения теории информации и кодирования

Глубина сообщения q - количество различных элементов (символов, знаков), принятых для представления сообщений. В каждый момент времени реализуется только один какой-либо элемент. Например, если сообщение представляется символами - буквами русского алфавита, то $q = 32$ (для простоты буква ё во внимание не принимается), английского - $q = 26$.

Длина сообщения n - количество позиций, необходимых и достаточных для представления сообщений заданной величины.

Если символы сообщения - цифры, то сами сообщения - числа и их последовательности.

При заданных глубине и длине сообщения количество всех возможных сообщений (N), которое можно представить

$$N = q^n$$

Основные определения теории информации и кодирования

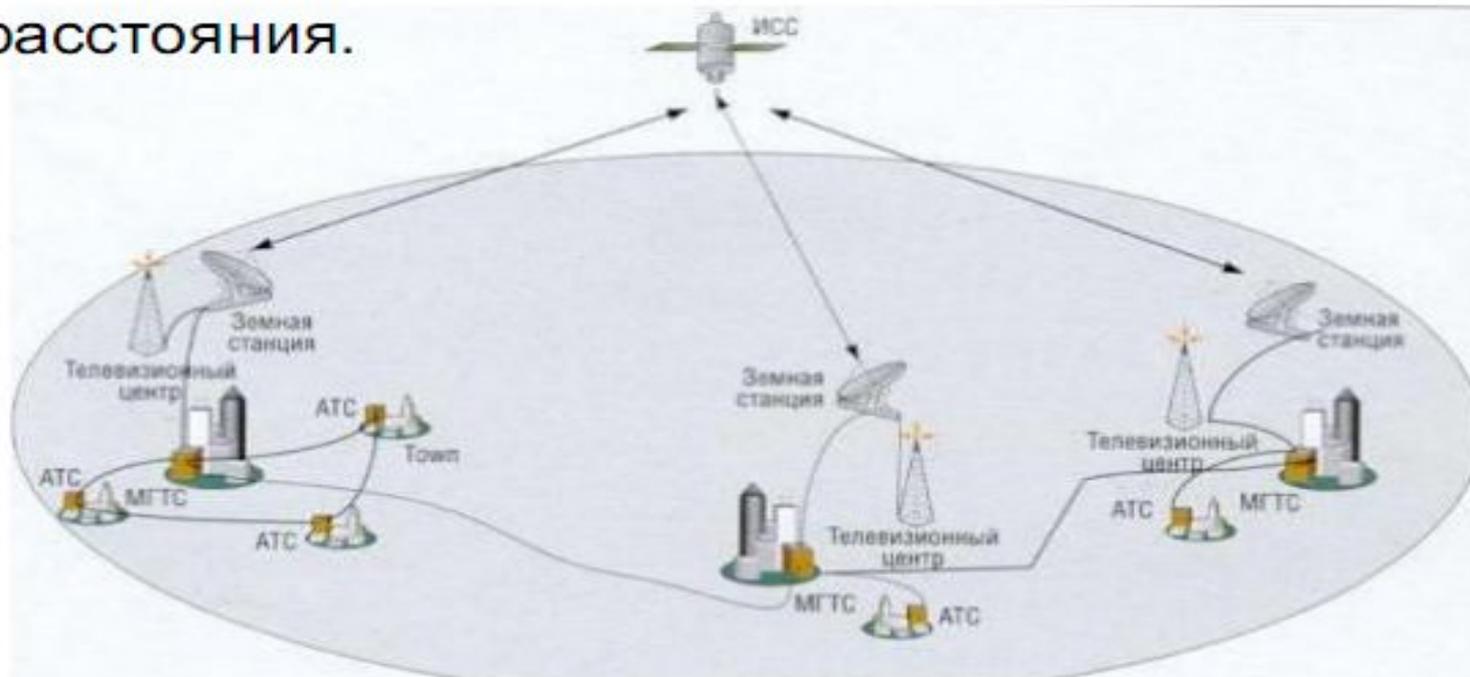
Материальным носителем информации, используемым для передачи сообщений, является **сигнал**.



А	Б	В	Г	Д	Е	Е
Ж	З	И	Й	К	Л	М
Н	О	П	Р	С	Т	У
Ф	Х	Ц	Ч	Ш	Щ	Ъ
Ы	Ь	Э	Ю	Я		

Основные определения теории информации и кодирования

В радиотехнике и электрической связи используются **электрические сигналы**, которые наилучшим образом приспособлены для передачи больших объемов данных на большие расстояния.



Основные определения теории информации и кодирования

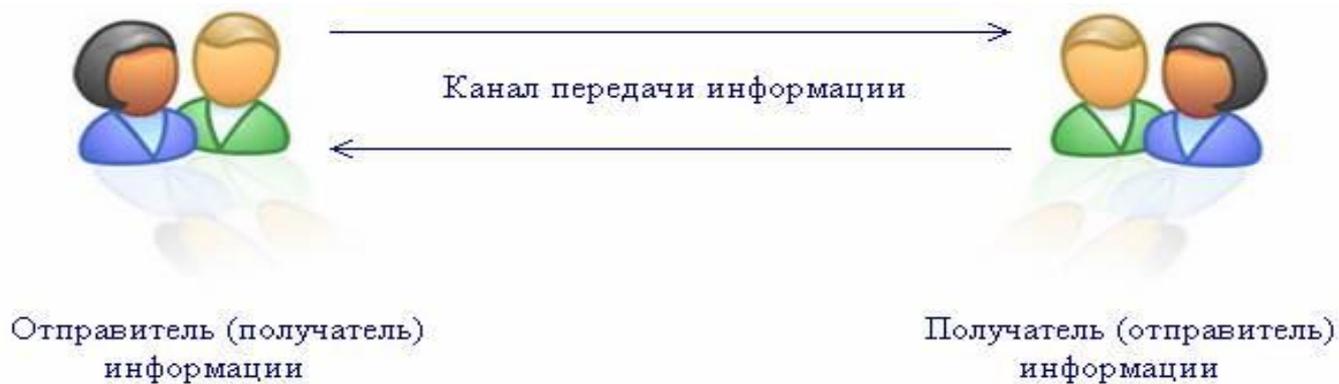
Кодирование — преобразование дискретной информации одним из следующих способов: шифрование, сжатие, защита от шума.

Входное кодовое слово – последовательность букв, символов, качественных признаков на входе канала связи называется входным кодовым словом, а на выходе – **выходным кодовым словом**.

Сообщения, составленные из входных кодовых слов, называют **первичными**, а сообщения, составленные из выходных кодовых слов – **вторичными**.

Кодирование может быть представлено как переход от первичных сообщений ко вторичным, а код как алгоритм, по которому осуществляется этот переход.

Обобщенная модель передачи информации. Как это было в теории информатики



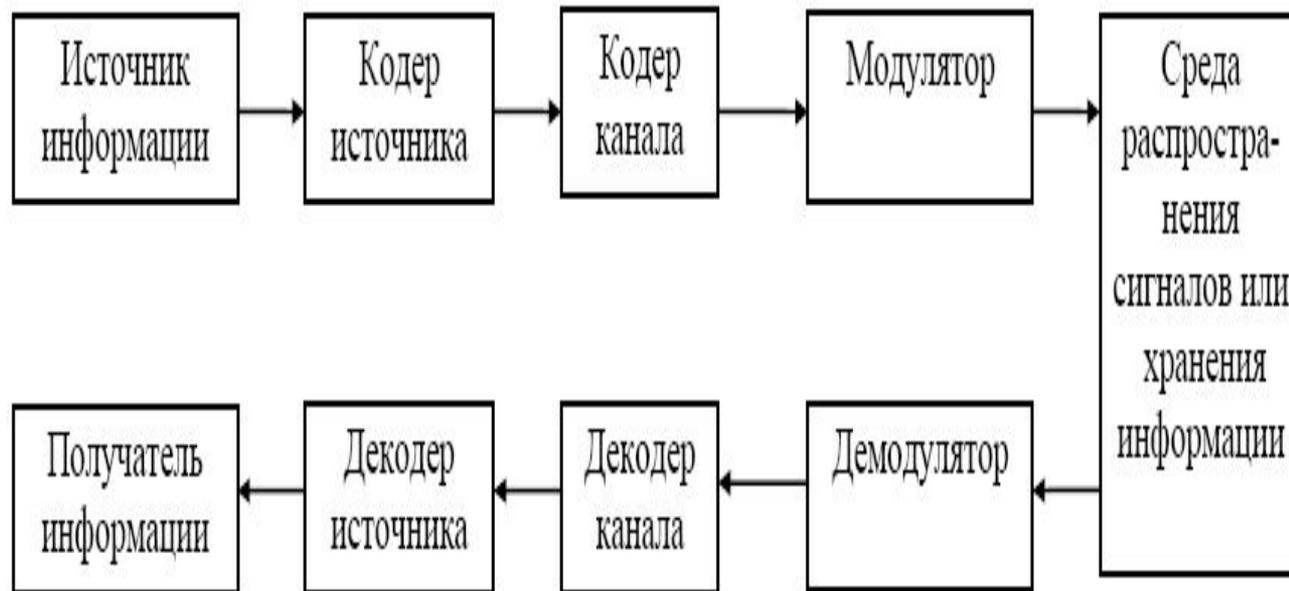
Источником информации является любое устройство или объект живой природы, порождающие сообщения, которые должны быть перемещены в пространстве или во времени (клавиатура ЭВМ, человек, аналоговый выход видеокамеры и т.п.) Далее считаем, что независимо от изначальной физической природы, все подлежащие передаче сообщения по умолчанию преобразуется в форму электрических сигналов

Приемник информации (получатель) – это субъект или объект, принимающий сообщение и способный правильно его интерпретировать.

Канал связи — это среда передачи информации, которая характеризуется в первую очередь максимально возможной для нее скоростью передачи данных (емкостью канала связи).

Обобщенная модель передачи информации. Как это должно быть (1)

Общая схема передачи информации.



Обобщенная модель передачи информации. Как это должно быть (2)

Кодер источника представляет информацию в наиболее компактной форме. Это нужно для того, чтобы эффективно использовать ресурсы канала связи.

Кодер канала обрабатывает информации с целью защиты сообщений от помех при передаче по каналу связи либо возможных искажений при их хранении.

Модулятор служит для преобразования сообщений, формируемых кодером канала, в сигналы, согласованные с физической природой канала связи или средой накопителя информации.

Остальные блоки, расположенные на приемной стороне, выполняют обратные операции и предоставляют получателю информацию в удобном для использования виде.

Классификация каналов связи

