

# ИНФОРМАТИКА

ЮУрГУ, каф. информатики  
Конова Елена Александровна  
[E\\_Konova@mail.ru](mailto:E_Konova@mail.ru)

# Введение

---

- Предмет и задачи.
- Балльно-рейтинговая система.
- Методические материалы.

# 1. Предмет и задачи дисциплины.

## Основные понятия

---

**Информатика** (informatique) =  
information (информация) + automatique (автоматика).  
Computer science – компьютерная наука.

**Информатика** – дисциплина, изучающая структуру и общие свойства информации, а также закономерности и методы её создания, хранения, поиска, преобразования, передачи и применения в различных сферах человеческой деятельности.

## 1.1. Предмет дисциплины

---

Информатика изучает:

- 1) области, связанные с разработкой, созданием, использованием и материально-техническим обслуживанием систем обработки информации, включая компьютеры и их программное обеспечение;
- 2) организационные, коммерческие, административные и социально-политические аспекты компьютеризации.

## 1.1. Предмет дисциплины

---

Информатика как наука, это:

- это фундаментальная наука, которая изучает теорию информационных систем и технологий, методологию создания информационного обеспечения;
- это прикладная наука, которая изучает закономерности в информационных процессах, создает информационные модели и разрабатывает информационные системы и технологии для предметных областей;
- это область бизнеса, представляющая коммерческий интерес.

## 1.2. Теоретическая информатика

---

Теоретическая информатика занимается созданием системы понятий, выявлением общих закономерностей, позволяющих описывать информацию и информационные процессы, протекающие в различных сферах (в природе, обществе, человеческом организме, технических системах).

Теоретическая информатика используется для создания новых средств работы с информацией.

## Некоторые направления

---

1. **Теория информации** изучает процессы хранения, передачи, преобразования информации.
2. **Математическое моделирование**, методы вычислительной и прикладной математики.
3. **Искусственный интеллект** – моделирование логического и аналитического мышления в интеллектуальной деятельности человека, в том числе, в прикладной лингвистике.
4. **Системный анализ** изучает методологию принятия решений по сложным проблемам.

# Прикладная информатика

---

Прикладная информатика, это совокупность наук, изучающих: теорию алгоритмов, формальных языков, автоматов, вычислимости и вычислительной сложности, теорию графов, криптологию, логику, формальную семантику. Закладывает теоретические основы для разработки информационных систем.

- 1. Социальная информатика** изучает процессы информатизации общества.
- 2. Телекоммуникационные системы и сети**, в том числе, глобальные компьютерные сети.
- 3. Разработка** вычислительных систем и программного обеспечения.

# Практическая информатика

---

Разрабатывает методики и алгоритмы решения прикладных задач для различных предметных областей: наука, производство, образование, медицина, торговля, быт и другие.

Инструмент – компьютерная техника, программы, алгоритмы.

# Место информатики в системе наук

---



## 1.3. Инструментальные средства

---

Информатика базируется на компьютерной технике

- Технические средства, или аппаратура компьютеров – Hardware.
- Программные средства (совокупность всех программ, используемых компьютерами, и область деятельности по их созданию и применению) – Software.
- Алгоритмы – часть информатики, связанная с разработкой и изучением методов и приемов их построения – Brain ware.

## 1.4. Информация

---

**Информация** (лат. Informatio):

сведения, разъяснения, изложение.

**Обыденное определение** – данные или сведения, которые кого-либо интересуют.

**Техническое определение** – сообщения, передаваемые в форме знаков или сигналов.

**Определение К. Шеннона** – снятая неопределенность наших знаний о чем-то.

**Определение Н. Винера** – обозначение содержания, полученного из внешнего мира в процессе нашего приспособления к нему и приспособления к нему наших чувств.

## Прагматические определения

---

**Информация в лингвистике** – ресурс, используемый в профессиональной деятельности: письменная или звучащая речь. Структурированная информация: словари, корпуса, словари перевода, и т.д.

**Информация в информатике** – совокупность знаний о фактических данных и зависимостях между ними.

**Информация в информационных системах** – данные, подлежащие вводу в систему, хранимые, обрабатываемые, выдаваемые пользователю.

**Информация юридическая** – сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления.

## Информация в журналистике

---

Результат отражения многообразия действительности. Такое знание, которое потребно и у которого есть потребитель.

Сущность журналистики в поиске, сборе, обработке, интерпретации и распространении информации. Информация – это «инструмент», посредник между журналистом и аудиторией.

# Информация в журналистике (Е. Прохоров)

---

1. В узком понимании - событийная информация: заметка, хроника.
2. В широком понимании - совокупность сведений, которые несет журналистика аудитории: опубликованные в СМИ произведения, характер заголовка газетной публикации в размере шрифта, интонации диктора или комментатора.
3. Специальное понимание. Информация - новое знание. Пример - хроникальная заметка, содержащая важную новость.

# Категории информации

---

Информация существует в трех видах:

- Данные (**Data**),
- Информация (**Information**),
- Знания (**Knowledge**)

# Данные

---

**Данные** – зафиксированные наблюдения, которые в данный момент времени не оказывают воздействия на управление или принятие решения.

Данные обычно представлены в форме, которая позволяет ее использовать для компьютерной обработки и передачи, то есть, закодированы, могут храниться.

Примеры:

Словарь – упорядоченный набор текстовых данных;

Произвольный текст (статья, конспект).

Из **данных** можно извлечь **информацию**

# Информация

---

Информация – обработанные данные, которые представлены в виде, пригодном для принятия получателем решения.

Примеры:

Извлеченное из словаря толкование слова;

Извлеченное из энциклопедии значение термина.

Информацией является содержание, значение данных, или факты, которые используются для принятия решения.

# Знания

---

**Знания** – процедуры и правила манипулирования фактами, и информация о том, когда и как следует применять эти процедуры и правила.

**Знания** – проверенный практикой результат познания действительности, вид информации, которая отображает знания человека, специалиста в предметной области.

**Знания:**

- Декларативные (факты) – знания об определенных событиях, явлениях, свойствах объектов. «Я знаю что...»
- Процедурные (правила) – знания о действиях, которые нежно предпринять для достижения какой-либо цели. «Я знаю как...»

# Отличия знаний от данных

---

1. Интерпретация. Хранимые данные могут быть интерпретированы человеком или программой. Данные не несут информации. Знания содержат как данные, так и их описание (правила интерпретации).
2. Наличие связей классификации. Данные не имеют эффективного описания связей между различными типами данных. Знания структурированы, можно установить соответствие между единицами знаний.
3. Наличие ситуационных связей. Связи описывают множество текущих состояний объекта. Данные трудно поддаются анализу.

## 1.13. Свойства информации

---

Свойства информации зависят от ?, и могут быть:

- достоверность;
- полнота;
- ценность;
- своевременность;
- понятность;
- доступность;
- краткость;
- и другие.

## 1.5. Представление информации

---

- Тексты, рисунки, чертежи, фотографии.
- Световые, звуковые сигналы.
- Радиоволны.
- Электрические и нервные импульсы.
- Жесты, мимика.
- Запахи, вкусовые ощущения.
- Хромосомы, посредством которых передаются по наследству признаки и свойства организмов.
- И другие. Примеры?

## Формализация как этап подготовки

---

Для хранения и обработки данные должны быть формализованы, возможно, структурированы, для компьютерной обработки иметь механизм хранения.

В информатике и информационных технологиях:

Данные – поддающееся многократной интерпретации представление информации в формализованном виде, пригодном для передачи, связи, или обработки (ISO/IEC 2382-1:1993).

Данные – формы представления информации для информационных систем и их пользователей (ISO/IEC 10746-2:1996).

## 1.7. Кодирование информации

---

**Кодирование** – представление данных в виде последовательности символических обозначений (букв, цифр, графических образов, звуков и т.п.).

Кодирование, сохраняя смысловую нагрузку данных, представляет их в формате, возможном для технического устройства.

Семиотика и семиотические системы.

## 1.8. Передача информации

---

Информация передаётся в форме сообщений.



Процесс кодирования и декодирования.

Помехи в канале связи приводят к искажению или потере информации.

Термин «Сообщение» из теории информации.

## 1.6. Сообщение как носитель информации

---

**Сообщение** – совокупность информации, достаточная для суждения о конкретном явлении, факте, процессе.

**Сообщение** – форма представления информации в виде речи, текстов, жестов, взглядов, изображений, цифровых данных, графиков, таблиц и т.п.

**Объем сообщения** зависит от предшествующих знаний потребителя, от уровня понимания этого сообщения и интереса к нему.

Информация есть характеристика соотношения между сообщением и его приемником.

## Примеры:

---

1. Информация о прогнозе погоды: источник – специалист-метеоролог, канал связи – телевизионная передающая аппаратура и телевизор, приемник – телезритель.
2. Живое существо своими органами чувств (осязание, обоняние и пр.) воспринимает информацию из внешнего мира, перерабатывает её в последовательность нервных импульсов, передает импульсы по нервным волокнам, хранит в памяти, воспроизводит, использует в процессе жизнедеятельности.

# Передача информации в журналистике

---

$Ж \rightarrow Т (И) \rightarrow А$

Журналистика (Ж) несет определенную информацию (И), заключенную в разнообразных текстах (Т) для аудитории (А).

В идеале полученная аудиторией информация становится источником новой информации, которая посредством обратной связи возвращается к журналистам.

## 1.9. Уровни адекватности информации

---

- Синтаксический.
- Семантический.
- Прагматический.

1. *Синтаксический* (формальный) уровень отражает внешние формально-структурные характеристики, независимо от ее смыслового содержания.

2. *Семантический* (смысловой) уровень рассматривает содержание информации, отражающее состояние объекта, безотносительно ее полезности для получателя.

3. *Прагматический* (практический) уровень рассматривает практическую ценность информации для принятия решений.

## 1.10. Измерение количества информации

---

Методы измерения количества информации приняты в теории информации с допущениями:

в определенных условиях можно пренебречь качественными особенностями информации, при этом выразить её количество числом, а также сравнить количество информации, содержащейся в различных группах данных.

Подходы к определению понятия «количество информации», основаны на том, что информацию, содержащуюся в сообщении, можно нестрого трактовать в смысле её новизны.

## 1.11. Единицы измерения информации

---

**Бит** – минимальная единица измерения информации  
(англ. bit – binary digit, двоичная цифра).

**В теории информации (К.Шеннон)** бит – количество информации, необходимое для различения двух равновероятных сообщений (типа «чет»-«нечет» и т.п.).

## 1.11. Единицы хранения информации

---

**В** вычислительной технике бит – наименьшая единица памяти, необходимая для хранения одного из двух знаков "0" и "1", используемых для машинного представления данных и команд.

**Байт** равен восьми битам: столько требуется, чтобы закодировать любой из 256 символов алфавита клавиатуры компьютера.

## Производные единицы информации

---

1 Килобайт (Кб) = 1024 байт =  $2^{10}$  байт,

1 Мегабайт (Мб) = 1024 Кбайт =  $2^{20}$  байт,

1 Гигабайт (Гб) = 1024 Мбайт =  $2^{30}$  байт.

1 Терабайт (Тбайт) = 1024 Гбайт =  $2^{40}$  байт,

1 Петабайт (Пбайт) = 1024 Тбайт =  $2^{50}$  байт.

# Три способа измерения информации

---

- Содержательный подход
- Алфавитный подход
- Вероятностный подход

## Содержательный подход

---

Количество информации, заключенное в сообщении, определяется объемом знаний, которое оно несет приемнику информации (получающему человеку).

Американский инженер Р. Хартли в 1928 г. процесс получения информации рассматривал как выбор одного сообщения из конечного, наперёд заданного, множества из  $n$  равновероятных сообщений, а количество информации  $i$ , содержащееся в выбранном сообщении, определял как двоичный логарифм  $n$ .

Формула Хартли:  $i = \log_2 n$ .

## Содержательный подход (примеры)

---

1. При бросании монеты количество возможных вариантов результата равно 2. Оба эти варианта равновероятны. Ответ  $2^i = 2$ .  $i = 1$  бит.

2. В барабане для розыгрыша лотереи находится 32 шара. Сколько информации содержит сообщение о первом выпавшем номере (например, выпал номер 15)? Вытаскивание любого из 32 шаров равновероятно.  $2^i = 32$ .  $i = 5$  бит.

## Алфавитный подход

---

Для определения количества информации, заключенной в тексте используется понятие мощности алфавита, на котором записан текст. Мощность алфавита - полное количество символов в алфавите. Если допустить, что все символы встречаются с одинаковой частотой, то количество информации, которое несет каждый символ, вычисляется по формуле Хартли  $i = \log_2 N$ , где  $N$  – мощность алфавита.

Если весь текст состоит из  $K$  символов, то размер содержащейся в нем информации равен  $I = K \times i$ , где  $i$  – информационный вес одного символа в алфавите.

## Алфавитный подход (пример)

---

Книга, набранная с помощью компьютера, содержит 150 страниц; на каждой странице – 40 строк, в каждой строке – 60 символов. Каков объем информации в книге?

Мощность компьютерного алфавита равна 256. Один символ несет  $i = \log_2 N = \log_2 256 = 8$  бит информации. Объем всей информации в книге  $8 \times 60 \times 40 \times 150 = 2880000$  бит = 360 000 байт.

## Вероятностный подход

---

Определяет количественную связь между вероятностью появления некоторого события ( $p$ ) и количеством информации в сообщении о наступлении этого события, учитывающую возможную неодинаковую вероятность сообщений в наборе:  $i = \log_2(1/p)$ .

Американский учёный Клод Шеннон предложил в 1948 г. формулу определения количества информации:

$$I = -\sum_{i=1}^n p_i \log_2 p_i,$$

где  $p_i$  вероятность того, что система находится в  $i$ -м состоянии.

## Вероятностный подход (пример)

---

В коробке имеется 50 шаров. Из них 40 белых и 10 черных. Определить количество информации в сообщении о том, что из корзины был вытащен черный шар.

Вероятность вытаскивания черного шара  $p = 10/50 = 0,2$ .  
Количество информации в сообщении о том, что вытащен черный шар  $i = \log_2(1/0,2) = \log_2 5 = 2,32$  бита. .

## 1.12. Информационные процессы

---

- искать;
- собирать;
- формализовать;
- хранить;
- воспринимать;
- использовать;
- передавать;
- принимать;
- копировать;
- распространять;
- преобразовывать;
- комбинировать;
- обрабатывать;
- делить на части;
- упрощать;
- измерять;
- разрушать;
- и другие.

## 1.14. Обработка информации

---

Информация нематериальна, поэтому обработка, это различные преобразования. Информация, предназначенная для обработки, называется данными.

**Обработка информации**, это получение одних информационных объектов из других путем выполнения специальных алгоритмов, ориентированных на прикладные задачи.

**Назначение** – увеличение объёма и разнообразия информации.

**Средства обработки информации** – устройства, в первую очередь, компьютер, и программные системы.

Основа компьютерной обработки информации – Brain Ware, алгоритмы.

## 1.16. Информационные технологии

---

Согласно определению ЮНЕСКО, информационная технология, это комплекс взаимосвязанных, научных, технологических, инженерных дисциплин, изучающих: методы эффективной организации труда людей, занятых обработкой и хранением информации; вычислительную технику и методы организации и взаимодействия с людьми и производственным оборудованием, их практические приложения, а также связанные со всем этим социальные, экономические и культурные проблемы.

## 1.16. Информационные технологии

---

**Информационные технологии** – совокупность методов и устройств для обработки информации.

**Информационные технологии** – последовательность прохождения информационных потоков в структуре системы управления.

Этапы:

- получение данных;
- хранение;
- сортировка;
- обработка;
- внешняя и внутренняя передача.

# Развитие ИТ по инструментарию

---

- 1-й этап: до второй половины XIX в., это «ручная» информационная технология, инструменты которой – письменный текст, бумага, перо.

Коммуникации: ручной способ – почта, курьерский способ.

Основная цель: формализованное представление информации.

- 2-й этап: с конца XIX в. , это «механическая» технология, инструменты которой – пишущая машинка, телефон, диктофон.

Коммуникации: почта, курьерский способ, телефон, телеграф.

Основная цель: формализованное представление информации по способу обработки, передачи.

# Развитие ИТ по инструментарию

---

- 3-й этап: 40-60-е гг. XX в., это «электрическая» технология, инструменты которой – большие ЭВМ и программное обеспечение, электрические пишущие машинки, ксероксы, диктофоны.

Коммуникации: телефон, телеграф, линии передачи.

Основная цель: формирование содержания информации.

- 4-й этап: с начала 70-х гг., это «электронная» технология, инструменты которой – большие и персональные ЭВМ, автоматизированные системы управления (АСУ), информационно-поисковые системы (ИПС).

Основная цель: формирование содержательной стороны информации для управленческой среды различных прикладных областей.

# Развитие ИТ по инструментарию

---

5-й этап: с середины 80-х гг., это «компьютерная» технология, инструменты которой – компьютер с широким спектром программных продуктов разного назначения.

Коммуникации: глобальные и локальные компьютерные сети.

Основная цель: создание систем поддержки принятия решений, имеющих встроенные элементы анализа и интеллекта для разных уровней управления, а также развитие технических средств бытового и культурного назначения.

# Инструментарий информационной технологии

---

Одно или несколько приложений для определенного типа компьютера, технология работы в котором позволяет достичь поставленную пользователем цель.

Инструменты лингвиста: текстовый процессор, настольные издательские системы, электронные таблицы, системы управления базами данных, электронные записные книжки, электронные словари, различные информационные системы функционального назначения.

# Информационные технологии в лингвистике

---

Информационные технологии в лингвистике – это совокупность законов, методов и средств получения, хранения, передачи, распространения, преобразования информации о языке и законах его функционирования с помощью компьютеров.

Информатика в лингвистике: [мнения студентов](#).

# Классификация

---

- Фундаментальная лингвистика.
- Прикладная лингвистика.

## 1.18. Компьютерная лингвистика

---

Компьютерная (математическая) лингвистика (computational linguistics) – научное направление в области математического и компьютерного моделирования интеллектуальных процессов у человека и животных при создании систем искусственного интеллекта.

Цель: использование математических моделей для описания естественных языков.

## Направления компьютерной лингвистики

---

- Обработка естественного языка (natural language processing): синтаксический, морфологический, семантический анализы текста.
- Оптическое распознавание символов (OCR).
- Автоматическое распознавание речи (ASR).
- Автоматический синтез речи.

# Обработка естественных языков

---

Компьютерная лингвистика:

разработка абстрактных моделей языковых систем.

**Обработка естественных языков:** реализация прикладных методов описания и обработки языка для компьютерных систем.

Цель: разработка алгоритмов и прикладных программ для обработки языковой информации.

## Обработка естественных языков

---

Создание и использование электронных корпусов текстов (корпусная лингвистика).

Создание электронных словарей, тезаурусов, онтологий. Например, словари используют для автоматического перевода, проверки орфографии.

Автоматический перевод текстов.

Извлечение информации (fact extraction, text mining): автоматическое извлечение фактов из текста.

Автореферирование (automatic text summarization). Эта функция включена в Microsoft Word.

## Обработка естественных языков

---

Построение систем управления знаниями (экспертные системы).

Создание вопросно-ответных систем (question answering systems).

## 1.15. Информационные ресурсы

---

**Информационные ресурсы** – это знания и указания по их реализации, накопленные в форме, позволяющей их воспроизводство:

книги, статьи, словари, патенты, диссертации, научно-исследовательская документация, технические переводы, данные о производственном опыте и другое.

## Русская виртуальная библиотека <http://rvb.ru/>

---

Каталог лингвистических программ и ресурсов в Сети, составленный С.В. Логичевым.

<http://rvb.ru/soft/catalogue/index.html>

Каталог включает в себя описание:

- программ, связанных с анализом текстов и вычислительной лингвистикой;
- лингвистических ресурсов, доступных в Интернет.

В основном, это бесплатные программы, доступные для загрузки или использования on-line.

## 1.17. Информатизация общества

---

Информатизация общества – социально-экономический и научно-технический процесс создания оптимальных условий для удовлетворения информационных потребностей и реализации прав граждан, органов государственной власти, органов местного самоуправления организаций, общественных объединений на основе формирования и использования информационных ресурсов.

# Информатизация

---

Цель информатизации – улучшение качества жизни людей за счет увеличения производительности и облегчения условий их труда.

Информатизация – социальный процесс, связанный с изменениями в образе жизни населения путем внедрения компьютерной грамотности, формирования культуры использования новых информационных технологий и пр.

## Задачи журналистики

---

Основными задачами журналистики являются подготовка информации и доведение ее до аудитории.. Информация должна быть социально-значимой, современной, новой. Такая информация воздействует на мышление, духовный мир человека и вызывает определенные реакции. Так формируется воспитательная функция журналистики. Каждый раз журналист своим произведением влияет определенным образом на сознание массово. Это влияние может быть, как позитивным, так и негативным.