

Лекция 1. Введение в дисциплину «Поиск и обработка экономической информации средствами Интернет и офисных приложений»

Изучаемые вопросы:

1. Цели и задачи дисциплины, распределение учебного времени.
2. Информация и ее свойства.
3. Технологии поиска и обработки информации.
4. Операции с данными

1. Цели и задачи дисциплины, распределение учебного времени.

«Программа развития НИУ ВШЭ на период с 2009 по 2015 гг. и до 2020 г.» **основной задачей университета** определяет осуществление информационно-аналитического и экспертного обеспечения деятельности Правительства по вопросам экономической и социальной политики, а также государственного и муниципального управления. Следовательно, выпускники университета должны обладать современными компетенциями аналитика и эксперта в области экономической и социальной политики государства

Дисциплина «Поиск и обработка экономической информации средствами Интернет и офисных приложений» **имеет целью** формирование информационной культуры специалиста и предполагает изучение теоретических основ, принципов построения и организации функционирования современных средств обработки данных и способов их эффективного применения.

Задачей данного курса является приобретение студентами знаний, умений и навыков, позволяющих им найти необходимую экономическую информацию в локальной или глобальной вычислительной сети подготовить информацию к анализу, провести такой анализ и оформить результаты анализа в виде электронного документа, содержащего различные типы данных: текстовые, графические, рассчитываемые и т. д.

Тема 1. Экономическая информация и технологии поиска и обработки информации

Информация и ее свойства. Технологии поиска и обработки информации. Кодирование и структурирование данных. Аппаратное и программное обеспечение вычислительных систем как средство реализации информационных технологий. Вычислительные сети и технологии поиска информации.

Тема 2. Технологии обработки текстовой и графической информации

Информационные технологии, реализованные в современных текстовых процессорах. Общая характеристика Microsoft Word 2007. Интерфейс Word 2007 и его особенности. Разработка текстовых и графических документов с помощью Microsoft Word 2007. Поиск и просмотр документов. Разработка печатных и электронных документов.

Тема 3. Технологии создания документов, содержащих расчеты

Система математических расчетов MathCAD и особенности ее применения. Общая характеристика системы MathCAD. Информационная среда, предоставляемая в распоряжение пользователя. Входной язык, встроенные функции и модули системы MathCAD. Решение математических и экономических задач с помощью MathCAD .

Таблицы и табличные процессоры. Основные понятия электронных таблиц. Применение электронных таблиц для решения задач экономического содержания. Использование расчетных технологий в экономических документах

Тема 4. Информационные системы экономиста

Информационные системы и базы данных. Данные и программы, представление данных в компьютере. Базы данных и системы управления базами данных. . Назначение и классификация СУБД. Реляционная система управления базами данных и особенности их построения. Разработка структуры базы данных. СУБД Microsoft Access и ее использования в информационных системах

Распределение учебного времени

Модуль	Всего часов	Лекции	Лабораторные занятия	Контрольные работы	Домашние задания	Сам. работа	Отчетность
1	24	16	8			30	
2	24		24	2		30	
3	24		24		1	30	
Всего	72	16	56	2	1	90	Зачет

Рейтинговая система оценки!!!!

1. Оценка студенческой работы

Таблица соответствия оценок по десятибалльной и пятибалльной системам

По десятибалльной шкале

1 – неудовлетворительно

2 – очень плохо

3 – плохо

4 – удовлетворительно

5 – весьма удовлетворительно

6 – хорошо

7 – очень хорошо

8 – почти отлично

9 – отлично

10 – блестяще

По пятибалльной шкале

неудовлетворительно – 2

удовлетворительно – 3

хорошо – 4

отлично - 5

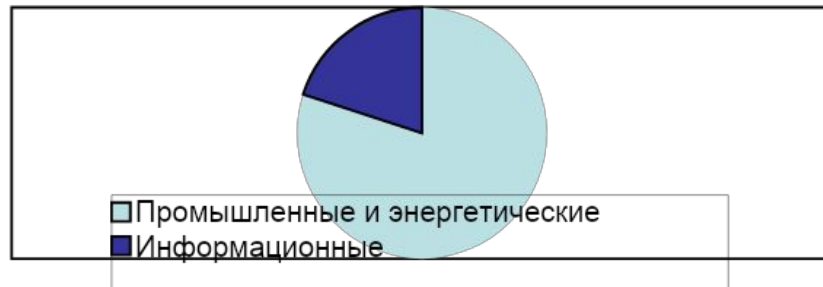
2. Итоговая оценка за учебную дисциплину

Оценка за лабораторные работы	Оценка за контрольные работы	Оценка за домашние задания	Оценка на зачете
Накопленная оценка			
Итоговая оценка			

Информационное общество

История цивилизации - непрерывный процесс развития **отношений**, связанных с созданием, распределением и потреблением **ресурсов**.

XX век - Индустриальный



Усиление средств массовой информации (политические технологии)

Электронная коммерция
Автоматизированные системы проектирования и производства
Усиление роли индивидуальной деятельности

XXI век - Информационный



Нормативная база электронного документооборота
Налогообложение электронных сделок
Защита информации

- Выводы:
1. Усиление роли информационных работников
 2. Экономист – важнейший информационный работник

2. Информация и ее свойства.

В настоящее время не существует единого определения информации как научного термина. С точки зрения различных областей знания, данное понятие описывается своим специфическим набором признаков.

В информатике

Предметом изучения науки **информатика** являются именно данные: методы их создания, хранения, обработки и передачи^[1]. А сама информация, зафиксированная в данных, её содержательный смысл интересны пользователям информационных систем, являющимся специалистами различных наук и областей деятельности: медика интересует медицинская информация, геолога — геологическая, предпринимателя — коммерческая и т. п. (в том числе специалиста по информатике интересует информация по вопросам работы с данными).

Системология

Работа с информацией связана с преобразованиями и всегда подтверждает её, работы, материальную природу:

- запись — формирование структуры материи и модуляции потоков путём взаимодействия инструмента с носителем;
- хранение — стабильность структуры (квазистатика) и модуляции (квазидинамика);
- чтение (изучение) — взаимодействие зонда (инструмента, преобразователя, детектора) с субстратом или потоком материи.

Системология рассматривает информацию через связь с другими основаниями: $I=S/F[MvRvT]$, где: I — информация; S — системность мироздания; F — функциональная связь; M — материя; v — (v подчёркнутое) знак великого объединения (системности, единства оснований); R — пространство; T — Время.

В физике

Объекты материального мира находятся в состоянии непрерывного изменения, которое характеризуется обменом **энергией** объекта с окружающей средой. Изменение состояния одного объекта, всегда приводит к изменению состояния, некоторого другого объекта окружающей среды. Это явление, вне зависимости от того, как, какие именно состояния и каких именно объектов изменились, может рассматриваться, как передача **сигнала** от одного объекта, другому. Изменение состояния объекта при передаче ему сигнала, называется регистрацией сигнала.

Сигнал или последовательность сигналов образуют **сообщение**, которое может быть воспринято получателем в том или ином виде, а также в том или ином объёме. Информация в физике есть термин, качественно обобщающий понятия «сигнал» и «сообщение». Если сигналы и сообщения можно исчислять количественно, то можно сказать, что сигналы и сообщения являются единицами измерения объёма информации.

Одно и тоже сообщение (сигнал) разными **системами** интерпретируется **по-своему**. Например, последовательно длинный и два коротких **звуковых** (а тем более в **символьном кодировании** ...) сигнала в терминологии азбуки Морзе — это буква Д (или D), в терминологии БИОС от фирмы AWARD — неисправность видеокарты.

В математике

Основные статьи: Собственная информация, Взаимная информация

В математике **теория информации** (*математическая теория связи*) — раздел прикладной математики, определяющий понятие информации, её свойства и устанавливающий предельные соотношения для систем **передачи данных**. Основные разделы теории информации — кодирование источника (*сжимающее кодирование*) и канальное (*помехоустойчивое*) кодирование

(Осталось: 1) Загрузка изображения http://upload.wikimedia.org/wikipedia/commons/thumb/d/dc/Nuvola_apps_important_yel

Интернет | Защищенный режим: вкл.

RU 10:25 13.09.2012

2. Информация и ее свойства

В **философском смысле информация** есть отражение реального мира, это сведения, которые один реальный объект содержит о другом реальном объекте.

Информация, как показано в кибернетике, имеет непосредственное отношение к процессам управления и развития, обеспечивающим устойчивость и выживаемость различных систем.

Информация имеет двойственный характер

Информация может быть отнесена к **абстрактным понятиям**, наряду с математическими понятиями

- в передающей системе количество информации не уменьшается, а в принимающей системе – увеличивается.
- информация обладает определенной независимостью от носителя
- информация может быть преобразована и передана с помощью различных сигналов, безотносительно к ее содержанию.

Информация обладает свойствами приближающими ее к **материальному миру**

- информация не может возникать из ничего;
- информацию можно получать, записывать, пересылать, стирать и т.д.

Понятие информации

Информация отождествляется со сведениями или фактами, которые теоретически могут быть получены и усвоены, то есть, преобразованы в знания.

"Под информацией понимают сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления" (Федеральный Закон № 24-ФЗ "Об информации, информатизации и защите информации" от 25.01.95 г.).

Информацию отождествляют с данными. Здесь налицо подмена понятий. Станут ли данные информацией и если да, то какой, зависит не только от данных, а и от многочисленных аппаратных, программных и естественных методов.

Он состоит в отказе от определения информации на том основании, что оно является фундаментальным, как, например, материя и энергия.

Информатика. Базовый курс. 2-е издание/Под редакцией С.В. Симоновича.- СПб.:Питер, 2008. -640 с

Информация - это действительно фундаментальное научное понятие. Информация - это динамический объект, не существующий в природе сам по себе, а образующийся в ходе взаимодействия данных и методов. Он существует ровно столько, сколько длится взаимодействие, а все остальное время пребывает в виде данных.

Информация - это продукт взаимодействия данных и методов, рассмотренный в контексте этого взаимодействия [2]

Любые изменения во внутренней структуре вещества или в энергетических полях сопровождаются образованием **сигналов**

Сигналы распространяются в пространстве и времени

Распространение сигналов завершается взаимодействием с веществом физического тела – **регистрацией сигналов**

Результат регистрации сигналов – **данные**. Данные всегда объективны

Если у вас есть **средства для регистрации данных**, значит у вас есть и **средства для их воспроизведения**. Эти средства предоставляют в ваше распоряжение **метод доступа к данным**

Чтобы данные стали **информацией** надо применить систему методов доступа

Методы доступа к текстовым данным

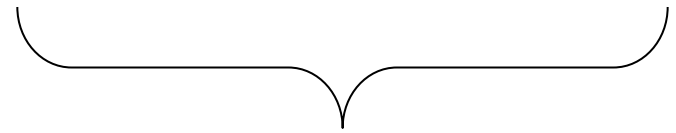
Метод зрения

Метод освещения

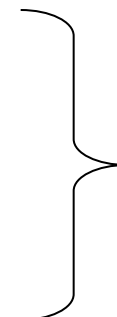
Знание азбуки

Знание языка

Знание терминов



Предметная область Информатики



Текстовая и
графическая
информация

Текстовая, графическая и
звуковая информация

Видеоинформация

Звуковая информация

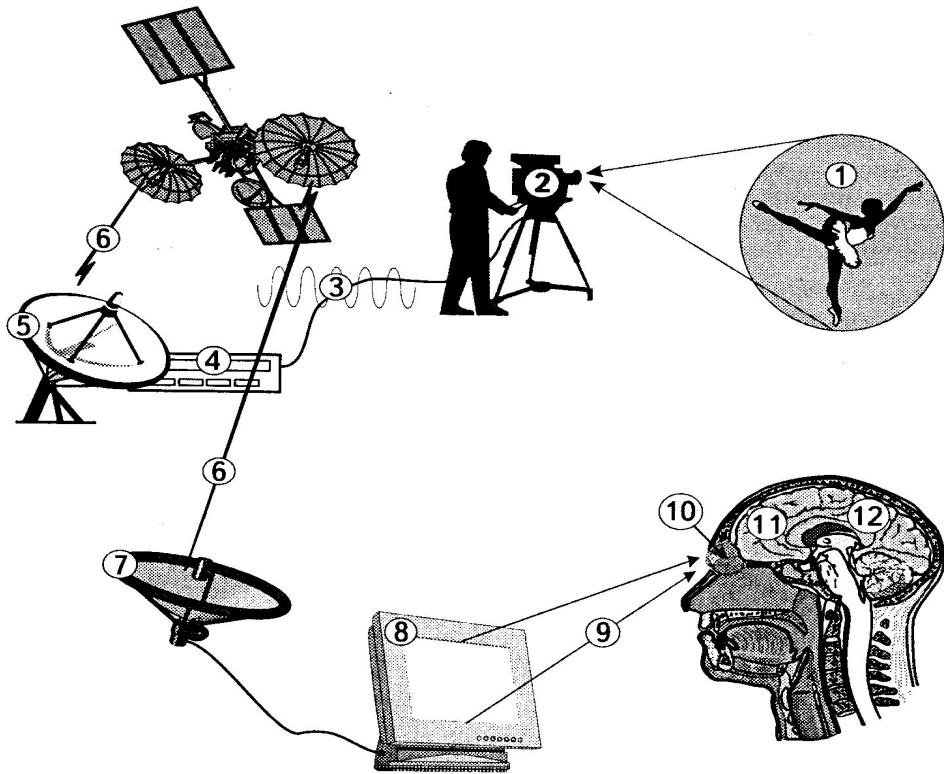
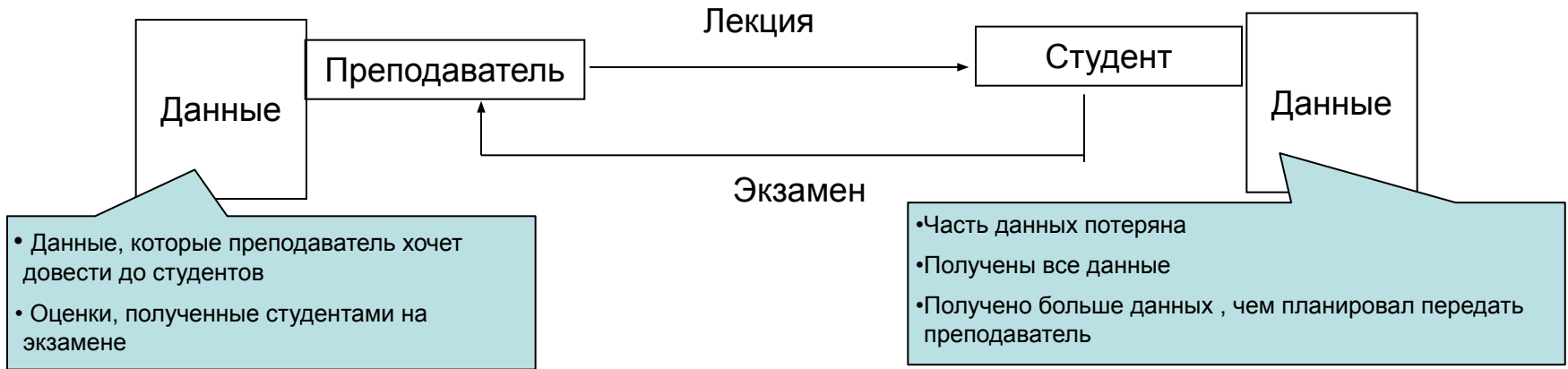
Видеоинформация

Звуковая информация

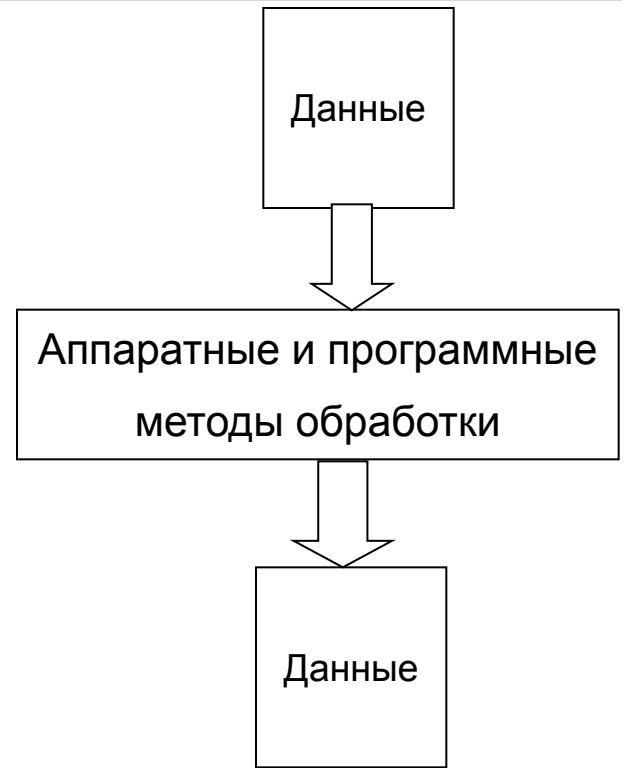
Текстовая информация

Текстовая и графическая информация

Информационные процессы



Информационные процессы в телевидении



Информационные процессы в компьютере

Информационный процесс – это цикл образования информации из данных и немедленного ее сохранения в виде новых данных.

Информация существует крайне непродолжительное время, но сам информационный процесс длится столько, сколько существуют носители данных.

Данные и методы находятся в постоянном диалектическом взаимодействии. Данные являются объективными, методы – субъективными.

Информация возникает и существует в момент диалектического взаимодействия объективных данных и субъективных методов.

Информатизация – совокупность процессов и явлений, связанных с целенаправленной обработкой информации, с применением средств вычислительной техники (СВТ) и связи, информационных технологий, а также соответствующего программного обеспечения (ПО).

Информационная технология (ИТ) – это система методов, производственных процессов и программных средств, интегрированных с целью обработки хранения, распространения, отображения и использования информации в интересах пользователей.

Проблема: Как оценить количество информации, полученной в ходе информационного процесса?

Для оценки и измерения **количества информации** могут быть применены различные подходы

Статистический подход

(К. Шеннон. 1948 г.)

Количество информации измеряется уменьшением (изменением) **неопределенности состояния системы** (энтропии). При получении информации энтропия уменьшается. Если энтропия равна нулю, то о системе имеется полная информация.

Структурный подход

Предполагает абстрагирование от смыслового содержания информации с целью организации таких логических и физических структур, которые позволили бы наиболее эффективное ее использование (быстрый поиск, извлечение, копирование, изменение и т.д.)

Подход предполагает преобразование информации в машинные коды и обратное преобразование.

Структурной единицей информации является байт.

Структурными элементами информации являются поля, записи, массивы, базы и банки данных, банки знаний.

Семантический подход

(Н. Винер, Ю. Шнайдер)

Для получения и использования информации ее получатель должен обладать определенным запасом знаний – **тезаурусом**. Изменение индивидуального тезауруса потребителя под воздействием сообщения свидетельствует о количестве смысловой информации. Если тезаурус потребителя близок к нулю, то и количество воспринятой им смысловой информации равно нулю.

Информация обладает свойством относительности и имеет субъективную ценность.

Для объективной оценки информации используется понятие общечеловеческого тезауруса.

Прагматический подход

(А. Харкевич)

Количество информации определяется как мера, способствующая достижению поставленной цели. Мера ценности информации – приращение вероятности достижения цели.

Качество информации определяется ее свойствами !!!

Адекватность - степень соответствия информации, полученной потребителем, тому, что автор вложил в ее содержание

Достоверность - соответствие информации реальности (как текущей, так и прошедшей) окружающего мира.

Полнота - достаточность информации для принятия решения.

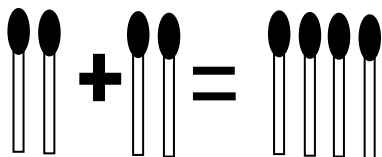
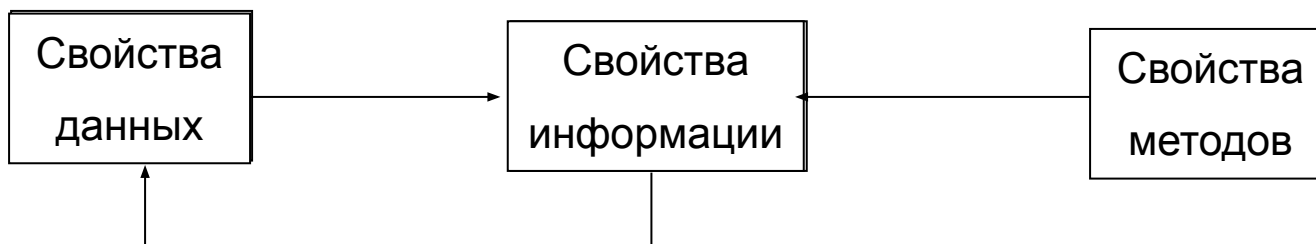
Избыточность – мера качества. Обычный текст имеет избыточность 20-25%, визуальная информация - более 90 %, видеоинформация - 98-99%.

Более **объективной** принято считать ту информацию, в которую методы вносят меньше субъективных элементов.

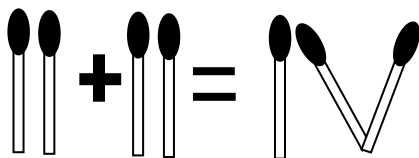
Доступность информации - это мера возможности получить ту или иную информацию

Актуальность - это степень соответствия информации текущему моменту времени.

Свойства данных, информации и методов доступа взаимосвязаны!!!



Естественный метод наблюдения и логический метод счета



Естественный метод наблюдения и метод записи римских цифр

Чем больше методов вам известны, тем больше информации вы получите!

3. Технологии поиска и обработки информации.

Информационные технологии в экономике – процесс , использующий совокупность методов и средств реализации операций сбора, регистрации, передачи, накопления и обработки информации на базе программно-аппаратного обеспечения для решения управленческих задач экономического объекта [1].

Основными техническими средствами реализации информационных технологий являются персональные компьютеры, локальные и глобальные вычислительные сети.

Технологии поиска и обработки информации реализуются в системном и прикладном программном обеспечении

Технологии поиска документов базируются на использовании имен файлов, технологии поиска данных – на ключевых словах.

Выбор технологии обработки информации зависит от типа данных: текстовые, Графические, рассчитываемые, звуковые, видео- и т.д.

4. Операции с данными

В ходе информационного процесса в компьютере данные преобразуются из одного вида в другой с помощью различных методов. Компьютер позволяет вести обработку данных с помощью множества различных операций.

Основные операции
конспектирования

накопление данных с целью обеспечения достаточной полноты информации для принятия решения

приведение данных, поступающих из различных источников, к одинаковой форме, чтобы сделать их сопоставимыми между собой, то есть повысить уровень их доступности (формальные языки)

отсеивание "лишних" данных, которые не нужны для принятия решения; при этом должен уменьшаться уровень "шума", а достоверность и адекватность данных должны возрастать

упорядочение данных по заданному признаку с целью удобства использования; повышает доступность информации

объединение данных по заданному признаку с целью повышения удобства использования; повышает доступность данных

организация хранения данных в удобной и легкодоступной форме; служит для снижения экономических затрат на хранение данных и повышает общую надежность информационного процесса в целом

комплекс мер, направленных на предотвращение утраты, воспроизведения и модификации данных

прием и передача данных между удаленными участниками информационного процесса; при этом источник данных в информатике принято называть *сервером*, а потребителя - *клиентом*

перевод данных из одной формы в другую или из одной структуры в другую

Основная цель обработки – повышения качества информации!!!

4.1. Кодирование данных

Один из методов
формализации

Выражение данных одного вида через
данные другого вида.

1 бит - 2 понятия {1 или 0}

2 бита - 4 понятия {00, 01, 10, 11}

3 бита - 8 понятий {000, 001, 010, 011, 100, 101, 110, 111}

.....

8 бита - 1 байт - 256 символов

2 байта - 65 536 символов

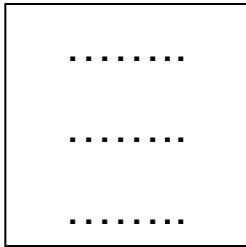
Кодирование действительных чисел:

$$3, 1415926 = 0,31415926 * 10^1$$

$$300000 = 0,3 * 10^6$$

$$123456789 = 0,123456789 * 10^9$$

Кодирование графических данных



Растр

Представляются системой целых чисел (двоичное кодирование)

Черно-белые иллюстрации представляются с помощью 256 градаций серого (1 байт на одну точку).

Для кодирования цветных графических изображений применяется принцип декомпозиции произвольного цвета на основные составляющие (красный, зеленый и синий). Такая система кодирования называется системой RGB.

На кодирование одной точки надо затратить 24 разряда. Такая система кодирования обеспечивает однозначное определение 16,5 млн. различных цветов, что близко к чувствительности человеческого глаза. Режим представления цветной графики с использованием 24 двоичных разрядов называется полноцветным.

Выводы:

1. Двоичная система кодирования позволяет представить числовые, символьные, графические и звуковые данные
2. Работа с данными чрезвычайно трудоемка и ее надо автоматизировать!

Основные структуры данных

Структуризация данных- обязательное условие автоматизации обработки данных

При создании любой структуры данных надо решить два вопроса: как разделять элементы данных между собой и как разыскивать нужный элемент

это простейшая структура данных, в которой адрес каждого элемента данных однозначно определяется уникальным номером

адрес каждого элемента определяется путем доступа (маршрутом), ведущим от вершины структуры к данному элементу

это упорядоченные структуры, в которых адрес элемента определяется номером строки и номером столбца, на пересечении которых находится ячейка, содержащая искомый элемент

1. Абрамов Александр Петрович
2. Бобров Борис Всеволодович
3. Воробьев Валентин Иванович

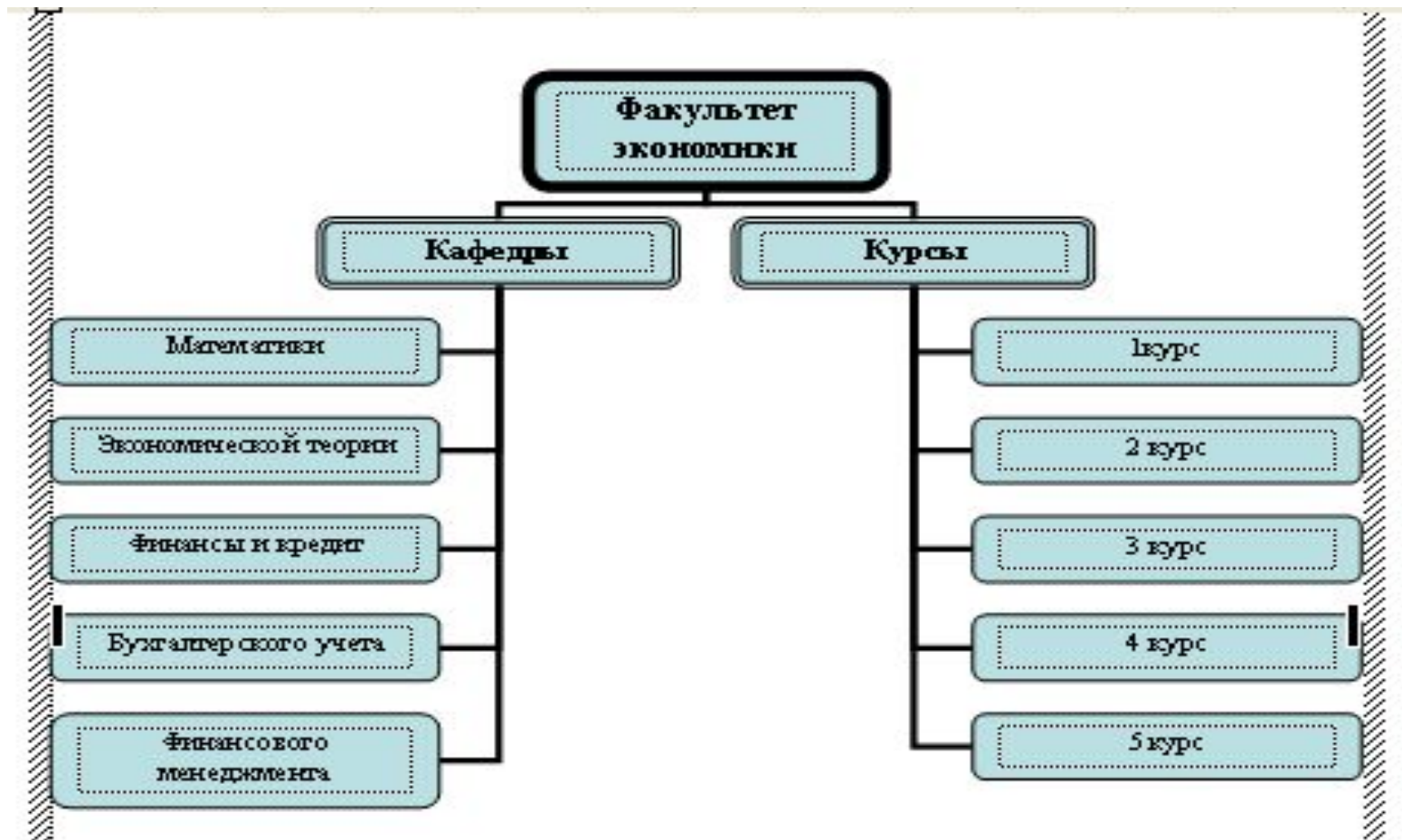
Достоинства: Простота адресации

Легко упорядочить

Недостаток: Трудность обновления

№ п/п	Фамилия	Имя	Отчество
1	Абрамов	Александр	Петрович
2	Бобров	Борис	Всеволодович
3	Воробьев	Валентин	Иванович
...
20	Сорокин	Сергей	Викторович

Пример иерархической структуры



Достоинства: Легко развивать и модифицировать

Недостатки: Трудности адресации (маршрут)

Трудно упорядочить, сортировать (индексация)

Структуризация данных приводит к увеличению объема данных за счет необходимости хранения и обработки адресов.

3. Файлы и файловая структура



Формат имени файла: имя носителя\имя каталога-1\...\имя каталога-n\,собственное имя файла

Пример: С:\электронная коммерция\ доклады\безопасность\цифровая подпись

Выводы: 1. Обязательным условием автоматизации обработки данных является структуризация и кодирование.

2. Структура данных определяется их типом.

Схема построения отчета о выполнении лабораторной работы:

1. Анализ решаемой информационной задачи на содержательном уровне:
 - Сформулировать решаемую задачу в терминах предметной области.
 - Какие входные данные и каком виде должны быть введены в компьютер?
 - Какие выходные данные и в каком виде предполагается получить от компьютера?
 - Какие операции нужно проделать над входными и промежуточными данными, чтобы получить конечные результаты?
2. Анализ информационной среды, созданной для эффективного решения поставленной задачи:
 - Обоснование выбора аппаратных и программных средств, позволяющих решить поставленную задачу.
 - Особенности настройки окна программы, используемой для решения задачи.
 - Какие информационные технологии, подпрограммы, инструменты и средства использовались в ходе решения?
3. Анализ результатов, полученных в ходе выполнения лабораторной работы:
 - Какой документ сформирован в ходе работы, где он хранится?
 - Какие компьютерные технологии использовались для анализа результатов решения?
 - Какое решение принято по результатам анализа?