

[https://yadi.sk/d/n9JcoQt\\_pQDi7](https://yadi.sk/d/n9JcoQt_pQDi7)

*Здесь находятся материалы лекций*

---

## *Лекция \_3*

# *Интернет технологии*

---

*Кафедра прикладной математики  
и кибернетики СибГУТИ*

*Молородов Юрий Иванович*

*yuto@ict.sbras.ru*

*Новосибирск*

---

# *Содержание*

Классификация информационных технологий (лек. 2)

Информационный ресурс и его составляющие

Итология. Предмет и методы итологии.

Организационная структура в области стандартизации ИТ.

Промышленные, профессиональные, административные организации, консорциумы.

Методы работы с информацией. Извлечение информации.

Основные понятия объектно-ориентированного подхода.

Методы обогащения информации.

Обработка информации. Архитектуры ЭВМ.

Системы поддержки принятия решений.

## ■ *Классификация информационных технологий*

В *нулевом* поколении (4000 г. до н. э. – 1900 г.) в течение шести тысяч лет наблюдалась эволюция от глиняных таблиц к папирусу, затем к пергаменту и, наконец, к бумаге. *Первое* поколение (1900 – 1955 гг.) связано с технологией перфокарт, когда запись данных представлялась на них в виде двоичных структур.

*Второе* поколение (программируемое оборудование обработки записей, 1955 – 1980 гг.) связано с появлением технологии магнитных лент, каждая из которых могла хранить информацию десяти тысяч перфокарт.

# ■ *Классификация информационных технологий*

*Третье поколение* (оперативные базы данных, 1965 – 1980 гг.) связано с внедрением оперативного доступа к данным в интерактивном режиме, основанном на использовании систем баз данных с оперативными транзакциями.

*Четвертое поколение* (реляционные базы данных: архитектура «клиент – сервер», 1980 – 1995 гг.) явилось альтернативой низкоуровневому интерфейсу.

*Пятое поколение* (мультимедийные базы данных, с 1995 г.) связано с переходом от традиционных хранящих числа и символы, к объектно-реляционным, содержащим данные со сложным поведением.

# Информационный ресурс и его составляющие

**Информационный ресурс** – концентрация имеющихся фактов, документов, данных и знаний, отражающих реальное изменяющееся во времени состояние общества, и используемых в научных исследованиях и материальном производстве.

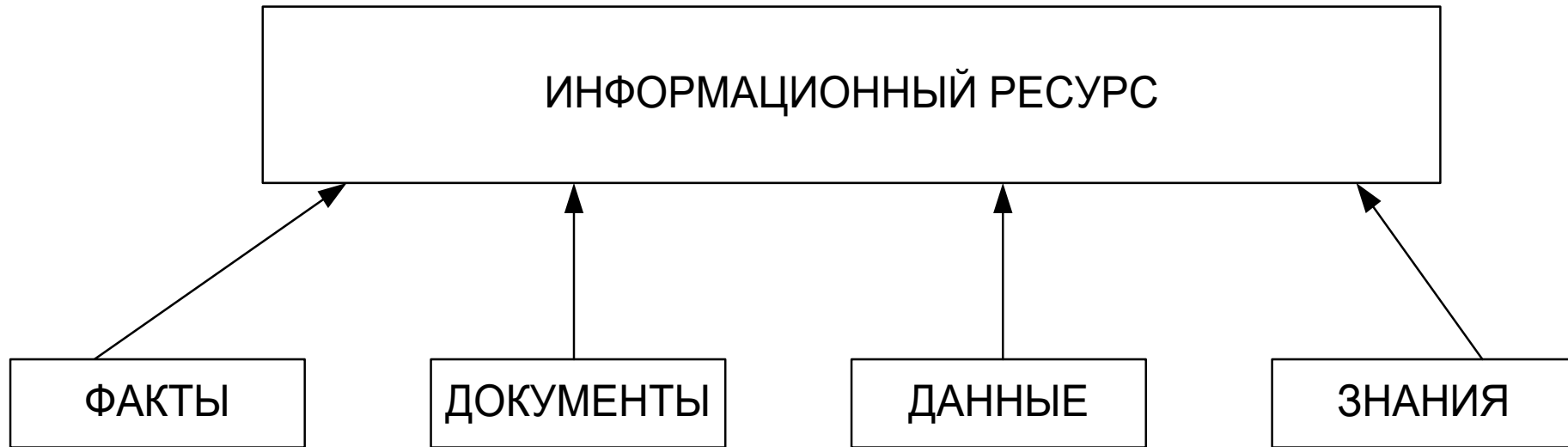
**Факты** – результат наблюдения за состоянием предметной области.

**Документы** – часть информации, определенным образом структурированная и занесенная на бумажный носитель.

**Данные** – вид информации, отличающийся высокой степенью форматированности.

**Знания** – итог теоретической и практической деятельности человека.

# *Информационный ресурс и его составляющие*



# *Итология. Предмет итологии.*

*Итология* -наука об информационных технологиях (ИТ-науки).

*Предметом* итологии являются:

- информационные технологии (ИТ);
- процессы, связанные с созданием ИТ;
- процессы, связанные с применением ИТ.

Становление итологии как науки началось в 1990 г.

*Основная* роль итологии:

- методологического базиса формализации, анализа и синтеза знаний;
- инструмента, продвигающего интеллектуальные способности и конструктивные возможности человека.

## *Итология. Методы итологии.*

Основными методами итологии являются:

- *архитектурная спецификация* – создание эталонных моделей важнейших разделов ИТ;
- *фундаментальная спецификация* – представление ИТ-систем, которое может наблюдаться на интерфейсах (границах) этих систем;
- *таксономия* – классификация профилей ИТ, обеспечивающая уникальность идентификации в пространстве ИТ;
- *разнообразные методы формализации и алгоритмизации знаний*;
- *методы конструирования* прикладных информационных технологий (парадигмы, языки программирования, базовые открытые технологии, функциональное профилирование ИТ и пр.).



## Организационная структура в области стандартизации ИТ.

Особенностью информационных технологий является их строгая стандартизация во всем мире. Организационная структура, поддерживающая процесс стандартизации ИТ, включает три основные группы организаций.

*Международные организации, входящие в структуру ООН:*

- **ISO** (International Organization for Standardization – Международная организация по стандартизации);
- **IEC** (International Electrotechnical Commission – Международная электротехническая комиссия);

## Организационная структура в области стандартизации ИТ.

*ITU-T* (International Telecommunication Union-Telecommunications – Международный союз по телекоммуникации – телекоммуникация). Она несет ответственность за разработку и согласование рекомендаций, которые обеспечивают *интероперабельность* (возможность совместного использования информации и ресурсов компонентами распределенной системы) телекоммуникационного сервиса в глобальном масштабе.

До 1993 г. эта организация имела другое название – *CC/TT* (International Telegraph and Telephone Consultative Committee – Международный консультативный комитет по телефонии и телеграфии, или МККТТ).

## *Промышленные, профессиональные, административные организации.*

- *IEEE* (Institute of Electrical and Electronic Engineers – Институт инженеров по электротехнике и электронике);
- *Internet u IAB* (Internet Activities Board – Совет управления деятельностью Internet);
- *Regional WOS* (Workshops on Open Systems – Рабочие группы по открытым системам).

## *Промышленные, консорциумы.*

- *ЕСМА* (European Computer Manufactureres Association – Европейская ассоциация производителей вычислительных машин);
- *OMG* (Object Management Group – Группа управления объектами);
- *NMF* (Network Management Forum – Форум управления сетями);
- *OSF* (Open Software Foundation – Основание открытого программного обеспечения).

## Промышленные, консорциумы

В 1987 г. *ISO* и *IEC* объединили свою деятельность в области стандартизации ИТ, создав единый орган *JTC1* (Joint Technical Committee 1 – Объединенный технический комитет 1), предназначенный для формирования всеобъемлющей системы базовых стандартов в области ИТ и их расширений для конкретных сфер деятельности. Работа над стандартами ИТ в JTC1 тематически распределена по подкомитетам (Subcommittees – *SC*).

# *Методы работы с информацией.*

## *Извлечение информации.*

Объекты и их свойства, процессы и функции, выполняемые этими объектами или для них являются источниками данных в любой предметной области. Любая предметная область рассматривается в виде трех представлений:

1. реальное представление предметной области;
2. формальное представление предметной области;
3. информационное представление предметной области

При *извлечении* информации важное место занимают различные формы и методы исследования данных:

- *поиск* ассоциаций, связанных с привязкой к какому-либо событию;
- *обнаружение* последовательностей событий во времени;

# Методы работы с информацией.

## Извлечение информации.

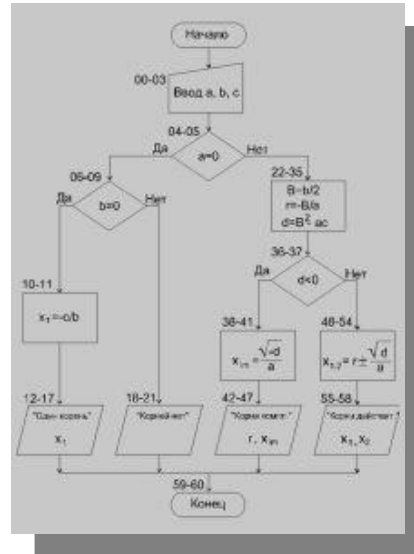
- **выявление** скрытых закономерностей по наборам данных, путем определения причинно-следственных связей между значениями определенных косвенных параметров исследуемого объекта (ситуации, процесса);
- **оценка** важности (влияния) параметров на развитие ситуации;
- **классифицирование** (распознавание), осуществляемое путем поиска критериев, по которым можно было бы относить объект (события, ситуации, процессы) к той или иной категории;
- **кластеризация**, основанная на группировании объектов по каким-либо признакам;
- **прогнозирование** событий и ситуаций.

# Извлечение информации

## Реальное представление предметной области



## Формальное представление предметной области



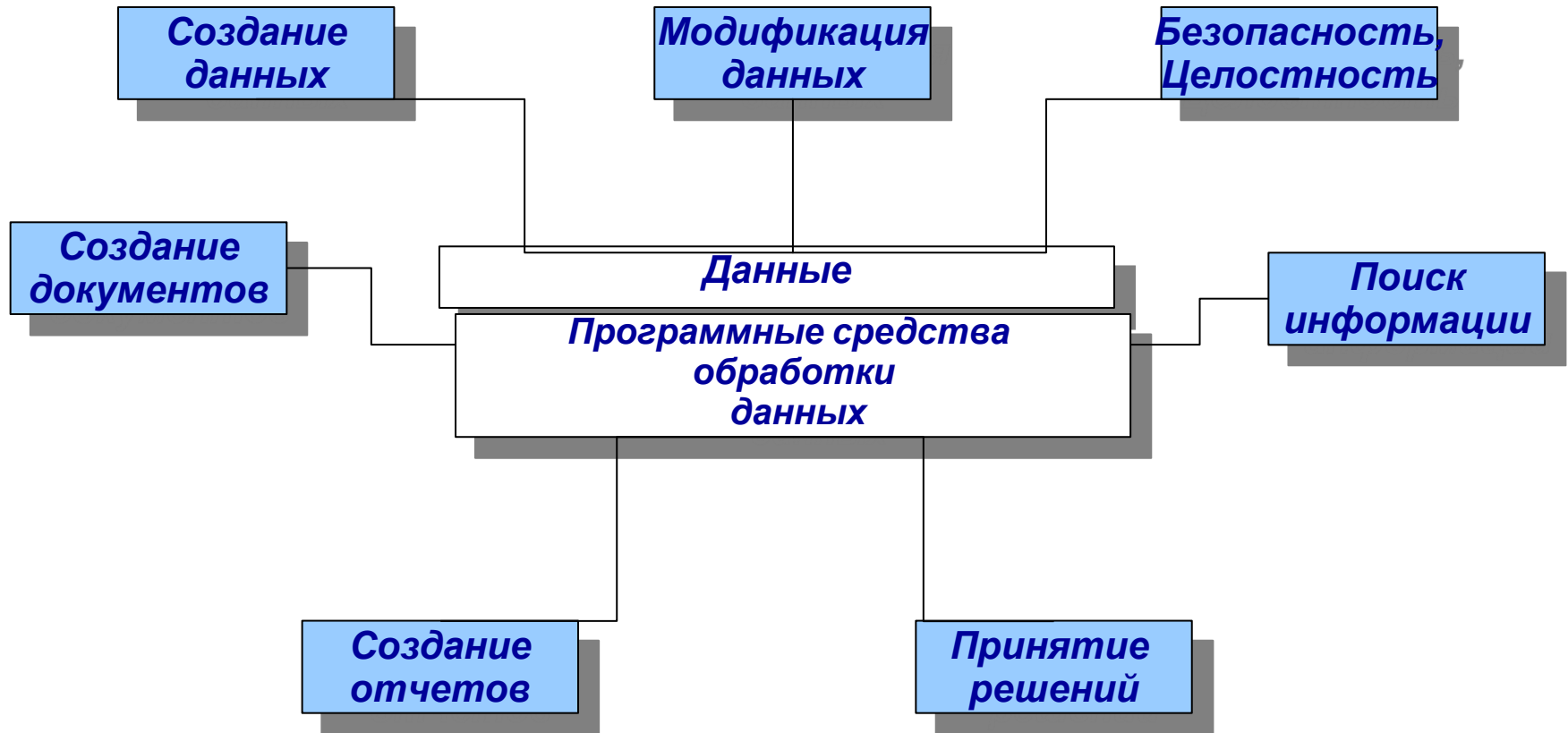
## Информационное представление предметной области



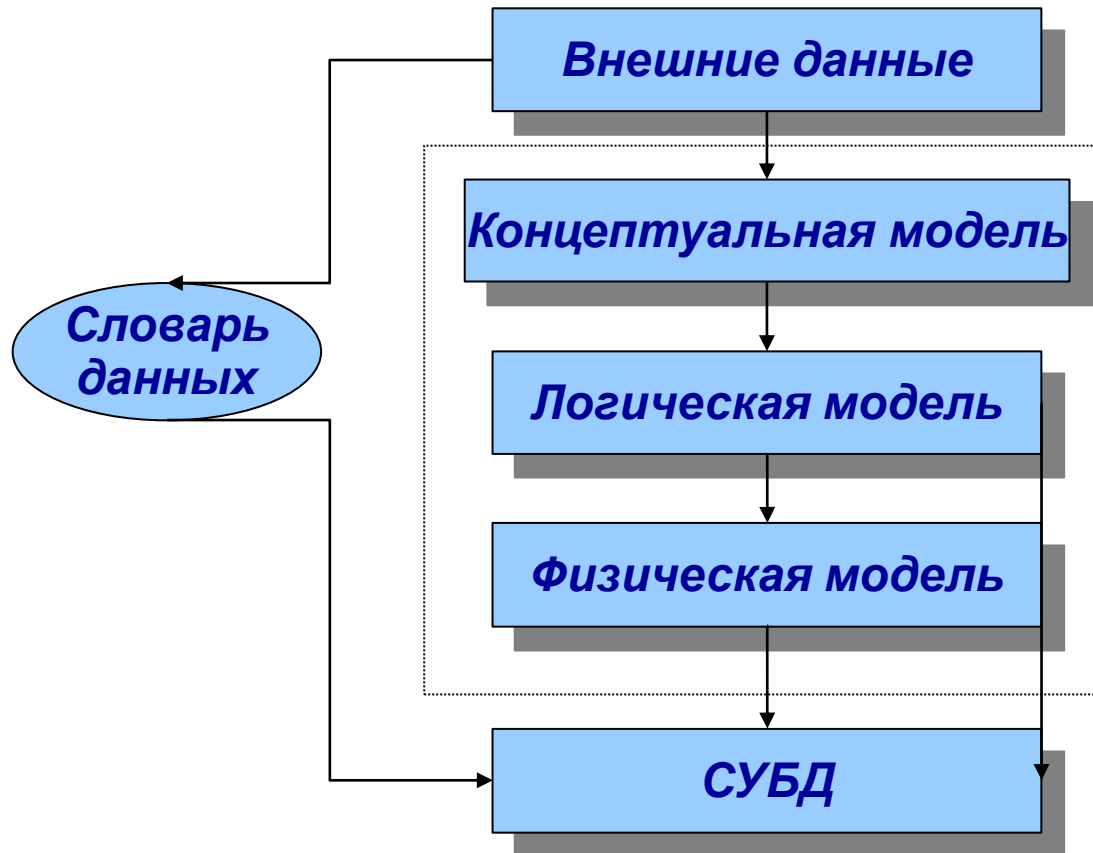
Т.о., *процесс извлечения* информации связан с переходом от реального представления предметной области к его описанию в формальном виде и в виде данных, которые отражают это представление



# Обработка информации



# Хранение информации



# *Представление и использование информации*



## Методы работы с информацией.

**Декомпозиция** – научный метод, использующий структуру задачи и позволяющий заменить решение одной большой задачи решением серии меньших задач. На основе объектно-ориентированного подхода она основана на выделении следующих основных понятий: объект, класс, экземпляр.

**Объект** – абстракция множества предметов реального мира, обладающих одинаковыми характеристиками и законами поведения. Он характеризует собой типичный неопределенный элемент такого множества. Основной характеристикой объекта является состав его атрибутов (свойств).

## Основные понятия объектно-ориентированного подхода.

**Полиморфизм** – способность объекта принадлежать более чем одному типу.

**Наследование** выражает возможность определения новых классов на основе существующих с возможностью добавления или переопределения данных и методов.

**Инкапсуляция** – это процесс отделения друг от друга элементов объекта, определяющих его устройство и поведение; инкапсуляция служит для того, чтобы изолировать контрактные обязательства абстракции от их реализации.

Понятия **полиморфизма** и **наследования** определяют эволюцию объектно-ориентированной системы, что подразумевает определение новых классов объектов на основе базовых.

## *Методы обогащения информации.*

*Структурное, статистическое, семантическое и прагматическое.*

*Структурное обогащение* предполагает изменение параметров сообщения, отображающего информацию в зависимости от частотного спектра исследуемого процесса, скорости обслуживания источников информации и требуемой точности.

При *статистическом обогащении* осуществляют накопление статистических данных и обработку выборок из генеральных совокупностей накопленных данных.

## *Методы обогащения информации.*

*Семантическое обогащение* означает минимизацию логической формы, исчислений и высказываний, выделение и классификацию понятий, содержания информации, переход от частных понятий к более общим. В итоге семантического обогащения удастся обобщенно представить обрабатываемую либо передаваемую информацию и устранить логическую противоречивость в ней.

*Прагматическое обогащение* является важной ступенью при использовании информации для принятия решения, при котором из полученной информации отбирается наиболее ценная, отвечающая целям и задачам

пользователя.

## Обработка информации.

*Обработка* состоит в получении одних «информационных объектов» из других «информационных объектов» путем выполнения некоторых алгоритмов и является одной из основных операций, осуществляемых над информацией, и главным средством увеличения ее объема и разнообразия.

На самом верхнем уровне можно выделить *числовую* и *нечисловую* обработку. При числовой обработке используются такие объекты, как переменные, векторы, матрицы, многомерные массивы, константы и пр. При нечисловой обработке объектами могут быть файлы, записи, поля, иерархии, сети, отношения и пр.



## Обработка информации.

С точки зрения реализации на основе современных достижений вычислительной техники выделяют такие виды обработки информации:

- *последовательная* обработка, применяемая в традиционной фонеймановской архитектуре ЭВМ, располагающей одним процессором;
- *параллельная* обработка, применяемая при наличии нескольких процессоров в ЭВМ;
- *конвейерная* обработка, связанная с использованием в архитектуре ЭВМ одних и тех же ресурсов для решения разных задач.

Если эти задачи тождественны, то это *последовательный* конвейер, если задачи *одинаковые* – векторный конвейер.

## *Обработка информации. Архитектуры ЭВМ.*

Существующие архитектуры ЭВМ с точки зрения обработки информации относятся к одному из следующих классов (*Классификация параллельных архитектур по Флинну*).

*Архитектуры с одиночным потоком команд и данных*

*Single Instruction, Single Data (**SISD**)*. Традиционная архитектура фон Неймана + КЭШ + память + конвейеризация

*Архитектуры с одиночными потоками команд и данных*

*Single Instruction, Multiple Data (**SIMD**)*. Особенностью данного класса является наличие одного (центрального) контроллера, управляющего рядом одинаковых процессоров.

# Обработка информации. Архитектуры ЭВМ.

Архитектуры с множественным потоком команд и одиночным потоком данных - *Multiple Instruction stream, Single Data stream (MISD)*.

Один из немногих – систолический массив процессоров, в котором процессоры находятся в узлах регулярной решетки, роль ребер которой играют межпроцессорные соединения. К классу MISD ряд исследователей относит конвейерные ЭВМ, однако это не нашло окончательного признания, поэтому можно считать, что реальных систем – представителей данного класса не существует.

# Обработка информации. Архитектуры ЭВМ.

*Архитектуры с множественным потоком команд и множественным потоком данных - Multiple Instruction stream, Multiple Data stream (**MIMD**).*

К этому классу могут быть отнесены следующие конфигурации:

- мультипроцессорные системы,
- системы с мультобработкой,
- вычислительные системы из многих машин, вычислительные сети.

## *Обработка информации.*

*Создание* данных, как процесс обработки, предусматривает их образование в результате выполнения некоторого алгоритма и дальнейшее использование для преобразований на более высоком уровне.

*Их модификация* связана с отображением изменений в реальной предметной области, осуществляемых путем включения новых данных и удаления ненужных.

*Контроль, безопасность и целостность* направлены на адекватное отображение реального состояния предметной области в информационной модели и обеспечивают защиту информации от несанкционированного доступа (безопасность) и от сбоев и повреждений технических и программных средств.

## Обработка информации.

*Поддержка принятия решения* является наиболее важным действием, выполняемым при обработке информации.

*Создание* документов, сводок, отчетов заключается в преобразовании информации в формы, пригодные для чтения как человеком, так и компьютером. С этим действием связаны и такие операции, как *обработка*, считывание, сканирование и сортировка документов.

При *преобразовании* информации осуществляется ее перевод из одной формы представления или существования в другую, что определяется потребностями, возникающими в процессе реализации информационных технологий.

## Обработка информации.

В *зависимости* от степени информированности о состоянии управляемого процесса, полноты и точности моделей объекта и системы управления, взаимодействия с окружающей средой, процесс принятия решения протекает в различных условиях:

### 1. Принятие решений в условиях определенности.

В этой задаче модели объекта и системы управления считаются заданными, а влияние внешней среды – несущественным. Поэтому между выбранной стратегией использования ресурсов и конечным результатом существует однозначная связь. Отсюда следует, что в условиях определенности достаточно использовать правило для оценки полезности вариантов решений, которое приводит к наилучшему эффекту.

## 2. Принятие решений в условиях риска.

В отличие от предыдущего случая для принятия решений в условиях риска необходимо учитывать влияние внешней среды, которое не поддается точному прогнозу, а известно только вероятностное распределение ее состояний.

В этих условиях использование одной и той же стратегии может привести к различным исходам, вероятности появления которых считаются заданными или могут быть определены.



## *Обработка информации.*

### 3. *Принятие решений в условиях неопределенности.*

Как и в предыдущем случае, между выбором стратегии и конечным результатом отсутствует однозначная связь. Кроме того, неизвестны также значения вероятностей появления конечных результатов, которые либо не могут быть определены, либо не имеют в контексте содержательного смысла.

## *Обработка информации.*

### 4. Принятие решений в условиях многокритериальности.

В любой из перечисленных выше задач многокритериальность возникает в случае наличия нескольких самостоятельных, не сводимых одна к другой целей.

Наличие большого числа решений усложняет оценку и выбор оптимальной стратегии.

Одним из возможных путей решения является использование методов моделирования.

## Обработка информации.

*Экспертные системы* пользуются знаниями, которыми они обладают в своей узкой области. Это ограничивает поиск на пути к решению задачи путем постепенного сужения круга вариантов.

Для решения задач в экспертных системах используют:

- метод логического вывода, основанный на технике доказательств, называемой резолюцией и использующей опровержение отрицания (доказательство «от противного»);
- метод структурной индукции, основанный на построении дерева принятия решений для определения объектов из большого числа данных на входе;

## *Обработка информации.*

- метод эвристических правил, основанных на использовании опыта экспертов, а не на абстрактных правилах формальной логики;
- метод машинной аналогии, основанный на представлении информации о сравниваемых объектах в удобном виде, например, в виде структур данных, называемых фреймами.

## *Обработка информации.*

Источники «интеллекта», проявляющегося при решении задачи, могут оказаться бесполезными либо полезными или экономичными в зависимости от определенных свойств области, в которой поставлена задача. Исходя из этого, может быть осуществлен выбор метода построения экспертной системы или использования готового программного продукта.

Процесс выработки решения на основе первичных данных, можно разбить на два этапа: выработка допустимых вариантов решений путем математической формализации с использованием разнообразных моделей и выбор оптимального решения на основе субъективных факторов.

## *Обработка информации.*

Информационные потребности лиц, принимающих решение, во многих случаях ориентированы на интегральные технико-экономические показатели, которые могут быть получены в результате обработки первичных данных, отражающих текущую деятельность предприятия. Для поддержки принятия решений обязательным является наличие следующих компонент:

- обобщающего анализа;
- прогнозирования;
- ситуационного моделирования.

## *Системы поддержки принятия решений.*

Аналитические системы поддержки принятия решений (СППР) позволяют решать три основных задачи: ведение отчётности, анализ информации в реальном времени (OLAP) и интеллектуальный анализ данных.

***OLAP (On-Line Analytical Processing);***

***EIS (Execution Information System);***

***DSS (Decision Support System);***

***PS (Presentation Services);***

***PL (Presentation Logic);***

***BL (Business or Application Logic);***

***DL (Data Logic);***

***DS (Data Services);***

***FS (File Services).***

## *Системы поддержки принятия решений.*

Аналитические системы поддержки принятия решений (СППР) позволяют решать три основных задачи: ведение отчётности, анализ информации в реальном времени (OLAP) и интеллектуальный анализ данных.

***OLAP (On-Line Analytical Processing);***

***EIS (Execution Information System);***

***DSS (Decision Support System);***

***PS (Presentation Services);***

***PL (Presentation Logic);***

***BL (Business or Application Logic);***

***DL (Data Logic);***

***DS (Data Services);***

***FS (File Services).***



## *Системы поддержки принятия решений.*

**OLAP (On-Line Analytical Processing)** – сервис представляет собой инструмент для анализа больших объемов данных в режиме реального времени. Взаимодействуя с OLAP-системой, пользователь сможет осуществлять гибкий просмотр информации, получать произвольные срезы данных, и выполнять аналитические операции детализации, свертки, сквозного распределения, сравнения во времени. В зависимости от функционального наполнения интерфейса системы выделяют два основных типа систем поддержки принятия решений : EIS и DSS.

## *Системы поддержки принятия решений.*

***EIS (Execution Information System)*** – информационные системы руководства предприятия. Эти системы ориентированы на неподготовленных пользователей, имеют упрощенный интерфейс, базовый набор предлагаемых возможностей, фиксированные формы представления информации.

***DSS (Desicion Support System)*** – полнофункциональные системы анализа и исследования данных, рассчитанные на подготовленных пользователей, имеющих знания как в части предметной области исследования, так и в части компьютерной грамотности.

## *Системы поддержки принятия решений.*

***PS (Presentation Services)*** – средства представления. Обеспечиваются устройствами, принимающими ввод от пользователя и отображающими то, что сообщает ему компонент логики представления PL, плюс соответствующая программная поддержка.

***PL (Presentation Logic)*** – логика представления. Управляет взаимодействием между пользователем и ЭВМ. Обработывает действия пользователя по выбору альтернативы меню, по нажатию кнопки или выбору элемента из списка.

## *Системы поддержки принятия решений.*

***BL (Business or Application Logic)*** - прикладная логика.

Набор правил для принятия решений, вычислений и операций, которые должно выполнить приложение.

***DL (Data Logic)*** – логика управления данными. Операции с базой данных (SQL-операторы SELECT, UPDATE и INSERT), которые нужно выполнить для реализации прикладной логики управления данными.

## *Системы поддержки принятия решений.*

**DS (Data Services)** - операции с базой данных. Действия СУБД, вызываемые для выполнения логики управления данными, такие как манипулирование данными, определения данных, фиксация или откат транзакций и т.п. СУБД обычно компилирует SQL-приложения.

**FS (File Services)** – файловые операции.

Дисковые операции чтения и записи данных для СУБД и других компонент.

Обычно являются функциями ОС.

## *Хранение информации.*

Хранение и накопление являются одними из основных действий, осуществляемых над информацией и главным средством обеспечения ее доступности в течение некоторого промежутка времени.

*База данных* может быть определена как совокупность взаимосвязанных данных, используемых несколькими пользователями и хранящихся с регулируемой избыточностью.

*Банк данных* – система, представляющая определенные услуги по хранению и поиску данных определенной группе пользователей по определенной тематике.

## *Хранение информации.*

*Система баз данных* – совокупность управляющей системы, прикладного программного обеспечения, базы данных, операционной системы и технических средств, обеспечивающих информационное обслуживание пользователей.

*Хранилище данных* (ХД – используют также термины Data Warehouse, «склад данных», «информационное хранилище») – это база, хранящая данные, агрегированные по многим измерениям.

Альтернативой хранилищу данных является концепция витрин данных (Data Mart). **Витрины данных** – множество тематических БД, содержащих информацию, относящуюся к отдельным информационным аспектам предметной области.

## *Хранение информации.*

Важным направлением развития баз данных являются репозитории.

*Репозиторий*, в упрощенном виде, можно рассматривать просто как базу данных, предназначенную для хранения не пользовательских, а системных данных.

Каждый из участников действия (пользователь, группа пользователей, «физическая память») имеет свое представление об информации.

По отношению к пользователям применяют трехуровневое представление для описания предметной области:

концептуальное, логическое и внутреннее (физическое).



## *Хранение информации.*

*Концептуальный уровень* связан с частным представлением данных группы пользователей в виде внешней схемы, объединяемых общностью используемой информации. Каждый конкретный пользователь работает с частью БД и представляет ее в виде внешней модели. Этот уровень характеризуется разнообразием используемых моделей (модель «сущность –связь», ER-модель, модель Чена), бинарные и инфологические модели, семантические сети.

## *Хранение информации.*

*Логический уровень* является обобщенным представлением данных всех пользователей в абстрактной форме. Используются три вида моделей: иерархические, сетевые и реляционные.

*Сетевая модель* является моделью объектов-связей, допускающей только бинарные связи «многие к одному» и использует для описания модель ориентированных графов.

*Иерархическая модель* является разновидностью сетевой, являющейся совокупностью деревьев (лесом).

*Реляционная модель* использует представление данных в виде таблиц (реляций), в ее основе лежит математическое понятие теоретико-множественного отношения, она базируется реляционной алгебре и теории отношений.

## *Хранение информации.*

*Физический (внутренний) уровень* связан со способом фактического хранения данных в физической памяти ЭВМ. Основными компонентами физического уровня являются хранимые записи, объединяемые в блоки; указатели, необходимые для поиска данных; данные переполнения; промежутки между блоками; служебная информация.

## *Обработка информации.*

По наиболее характерным признакам БД можно классифицировать следующим образом по способу хранения информации:

- интегрированные;
- распределенные;
- по типу пользователя:
- монопользовательские;
- многопользовательские;
- по характеру использования данных:
- монопользовательские;
- предметные.

## *Обработка информации.*

Важным аспектом проектирования БД является проблема интеграции и распределения данных. Распределение данных по месту их использования может осуществляться различными способами:

- *Копируемые данные.*
- *Подмножество данных.*
- *Реорганизованные данные.*
- *Секционированные данные.*
- *Данные с отдельной подсхемой.*
- *Несовместимые данные.*

## Обработка информации.

1. *Копируемые данные.* Одинаковые копии данных хранятся в различных местах использования, так как это дешевле передачи данных. Модификация данных контролируется централизованно;
2. *Подмножество данных.* Группы данных, совместимые с исходной базой данных, хранятся отдельно для местной обработки;
3. *Реорганизованные данные.* Данные в системе интегрируются при передаче на более высокий уровень;
4. *Секционированные данные.* На различных объектах используются одинаковые структуры, но хранятся разные данные;

## Обработка информации.

### 5. *Данные с отдельной подсхемой.*

На различных объектах используются различные структуры данных, объединяемые в интегрированную систему;

### 6. *Несовместимые данные.*

Независимые базы данных, спроектированные без координации, требующие объединения.

---

*Лекция окончена!*

*Благодарю за внимание!*

*Ваши вопросы.*

*Молородов Юрий Иванович*  
*yuto@ict.sbras.ru*