

Курс «СОВРЕМЕННЫЕ ПРОБЛЕМЫ ИНФОРМАТИКИ И
ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ»

Основные направления развития информационных систем и вычислительной техники



Воробьев Евгений
Сергеевич



Принципы развития ИТ:

- интерактивный (диалоговый) режим работы с компьютером
- интегрированность (стыковка, взаимосвязь) с другими программными продуктами.
- гибкость процесса изменения как данных, так и постановок задач.

Области развития ИТ:

- Обработка и хранение информации;
- Способы передачи данных;
- Программное обеспечение;
- Аппаратурное оформление.

Курс «СОВРЕМЕННЫЕ ПРОБЛЕМЫ ИНФОРМАТИКИ И
ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ»

Знания. Способы представления знаний





Информация

Объем информации во всем мире увеличивается более чем в два раза каждые два года, так в 2002 году было создано и реплицировано 1,8 экзабайт данных (1,8 трлн. гигабайт). Следующие 5 лет увеличили информацию на больший объем, чем имелось за всю жизнь ранее. Каждый год информация увеличивается на 30%

Что такое 1,8 экзабайт данных:

- каждый человек в России должен грузить в "Твиттер" по 3 записи в минуту непрерывно в течение 58 716 лет;
- это 200 млрд. фильмов в HD-качестве продолжительностью 2 часа каждый.

Для хранения этих данных потребуется 57,5 млрд. Apple iPad со встроенной памятью 32 Гбайт, что позволит:

- построить из iPad'ов копию Великой Китайской стены, которая будет вдвое выше оригинала;
- построить вокруг всей Южной Америки стену высотой 6 м;

Современная революция в информационных технологиях характеризуется тем, что

на 7 млрд человек приходится 6 млрд телефонов.



В следующие десять лет вплоть до 2020 года ИТ-департаменты по всему миру могут столкнуться со следующими проблемами:

- в 10 раз станет больше серверов (как физических, так и виртуальных);
- объем информации, которой нужно управлять, вырастет в 50 раз;
- в 75 раз вырастет число файлов или контейнеров, в которых информация инкапсулируется;
- численность ИТ-персонала, который управляет всей этой информацией и серверами, вырастет всего лишь в 1,5 раза!

Число этих объектов будет расти даже быстрее, чем объем хранимой в них информации, поскольку будет использоваться все больше встроенных систем,



Информационные цивилизации

Этот тип цивилизации основан на информационных технологиях.

У ее истоков лежат процессы, сущность которых принято обозначать терминами "информатизация" и "глобализация".

Виды цивилизаций:

1. Аграрную (доиндустриальную),
2. Промышленную (индустриальную)



Причины возникновения Информационной цивилизации:

1. появление транснациональных корпораций, что способствует возникновению нового международного разделения труда (то есть появляется экономическая взаимозависимость различных обществ, глубоко проникшая в экономику всех стран мира);
2. появление новых глобальных торговых сетей, каналов экономического обмена, связывающих страны и компании;
3. процесс, сопровождаемый зарождением информационных сетей, основанных на новейших технологиях; их бурное развитие;
4. интенсификация миграционных потоков различных уровней и развитие мирового туризма;
5. распространение культуры постмодерна, новых интегральных форм репигиозных культов, усиление влияния



Знание, способность его хранить,
концентрировать, приобретать,
обмениваться, добывать качественно
новое знание, играет все большую, если
не доминирующую роль в развитии
Постиндустриальной (информационной)
цивилизации.



Знания

Проблема представления знаний в компьютерных системах – одна из основных в области искусственного интеллекта. Её решение позволит специалистам, не программирующим, непосредственно на языке «деловой прозы» в диалоговом режиме работать на ЭВМ и с ее помощью формировать необходимые решения. Решение проблемы представления

Знания – это богатство, которое добывается, обрабатывается и

Данные

это не долго живущие новости, временные записи и т.п., не предназначенные для длительного использования

Контекст

Информация

это полу структурированные (или агрегированные) данные, служащие опорой для периодического принятия каких-либо решений

Осмысление

Знания

являются результатом переработки информации, имеют длительный цикл жизни, несут определенную идею и снабжены контекстом, определяющим область ее эффективного применения в данном месте в данное время



Виды знаний

- **Обыденное** знание служит основой ориентации человека в окружающем мире, основой его повседневного поведения и предвидения, но обычно содержит ошибки, противоречия;
- **Научному** знанию присущи логическая обоснованность, доказательность, воспроизводимость результатов, проверяемость, стремление к устранению ошибок и преодолению противоречий;
- **Интуитивное**;



Знания можно подразделять на явные (формализованные) – это то, что можно потрогать, сохранить на диске или найти в Интернете и потом просмотреть в браузере, и неявные (личностные)- это, что находится в головах сотрудников, рекомендации, которые пересказываются из уст в уста, навыки, выработанные в результате тренингов.

Однако в определенных случаях

Представления знаний внутри ИС:

- **Внутренняя интерпретируемость** – наличие уникального имени не только внутри программы, но для системы в целом;
- **Структурированность** – наличие внутренней структуры, которая определяет вложенность единиц информации друг в друга, облегчает построение связей и поиски;
- **Связанность** – возможность построения различных связей с информационными единицами;
- **Сематическая метрика** - отношение близости понятий, силы ассоциативной связи между ними;
- **Активность** - появление новых фактов в системе приводит к инициации процедур, обрабатывающих уже известные системе факты, т. е. данные управляют программой;;
- **Конвертируемость** - свойство изменять форму представления, уровень детализации и степень неопределенности информации о предметной области в



Корпоративные знания делятся на

Внешние

- знания клиента (наиболее важное знание для большинства организаций);
- независимая аналитическая информация (маркетинговые отчеты и рейтинги, цены на международных фондовых биржах) и др.



Внутренние

- знания о ключевых для данной отрасли процессах – накопление лучшего опыта (ноу-хау) при выполнении основных задач;
- знания об изделиях (и услугах);
- лучшие решения, наиболее соответствующие текущим потребностям пользователей;
- знания сотрудников – выявление, накопление и использование интеллектуального капитала (наиболее ценный актив организации);
- «память» организации (прошлый опыт);
- знания о построении отношений – глубокие персональные знания, которые обеспечивают успешное сотрудничество;



Представление знаний

Когнитология (наука о мышлении) - как люди хранят и обрабатывают информацию.

Информатика – подбор представлений конкретных и обобщённых знаний, сведений и фактов для накопления и обработки информации в ЭВМ.

Искусственный интеллект – научиться хранить знания таким образом, чтобы программы могли осмысленно обрабатывать их и достигать тем подобия человеческого интеллекта.

Классификация способов

Способы представления знаний

Теоретические способы

Логические способы

Формальные
грамматики

Комбинаторные
способы

Алгебраические
способы

Нейронные сети, генетические алгоритмы

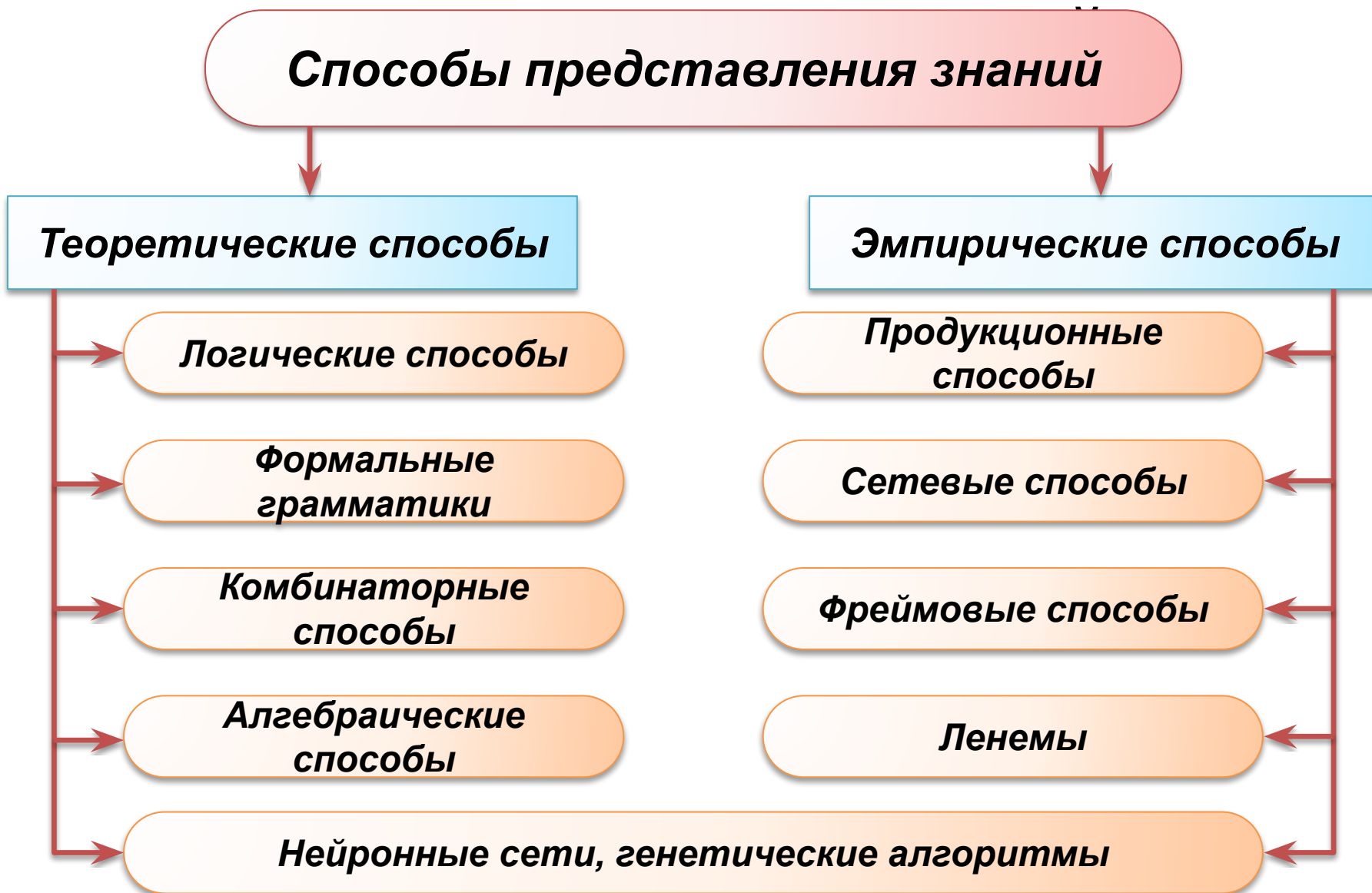
Эмпирические способы

Продукционные
способы

Сетевые способы

Фреймовые способы

Ленемы





Логическая модель

Основная идея при построении логических моделей знаний заключается в следующем – вся информация, необходимая для решения прикладных задач, рассматривается как совокупность фактов и утверждений, которые представляются как формулы в некоторой логике. Знания отображаются совокупностью таких формул, а получение новых знаний сводится к реализации процедур логического вывода.

Основные достоинства логических моделей знаний:

- в качестве «фундамента» здесь используется классический аппарат математической логики, методы которой достаточно хорошо изучены и формально обоснованы;
- существуют достаточно эффективные процедуры вывода, в том числе реализованные в языке логического программирования «Пролог»;
- в базах знаний можно хранить лишь множество аксиом, а



Продукционная модель

Её можно считать наиболее распространенной моделью представления знаний. Продукционная модель основанная на правилах, позволяющая представить знание в виде предложений типа:

«ЕСЛИ условие, ТО действие»

Продукционная модель обладает тем недостатком, что при накоплении достаточно большого числа (порядка нескольких сотен) продукций они начинают противоречить друг другу.

Данная модель используется в экспертных системах с которыми мы сталкиваемся в повседневной жизни.



Формальная грамматика

- Формальная грамматика или просто грамматика в теории формальных языков — способ описания формального языка, то есть выделения некоторого подмножества из множества всех слов некоторого конечного алфавита. Различают порождающие и распознающие (или аналитические) грамматики — первые задают правила, с помощью которых можно построить любое слово языка, а вторые позволяют по данному слову определить, входит оно в язык или нет.



Семантические сети или сетевые модели

Знаний
Знания представляются в виде графов, где узлы соответствуют понятиям и объектам, а дуги – отношениям между ними.

Семантическая сеть как модель наиболее часто используется для представления декларативных знаний. С помощью неё реализуются такие свойства системы, как интерпретируемость и связность, в том числе по отношениям **IS-A** и **PART-OF**. За счет этих свойств семантическая сеть позволяет снизить объем хранимых данных, обеспечивает вывод умозаключений по ассоциативным связям.

Сейчас семантические сети используются в системах распознавания образов, экспертных системах и



Фреймовая модель знаний

Фреймовая модель представляет собой систематизированную психологическую модель памяти человека и его сознания.

Фрейм (frame – рамка, каркас) – структура данных для представления некоторого концептуального объекта. Информация, относящаяся к фрейму, содержится в составляющих его слотах.

Слот (slot – щель, прорезь) может быть терминальным (листом иерархии) или представлять собой фрейм нижнего уровня.

Каждый фрейм состоит из произвольного числа слотов, причем несколько из них обычно

определяются самой системой для выполнения

Классификация способов

Способы представления знаний

Теоретические способы

Логические способы

Формальные
грамматики

Комбинаторные
способы

Алгебраические
способы

Нейронные сети, генетические алгоритмы

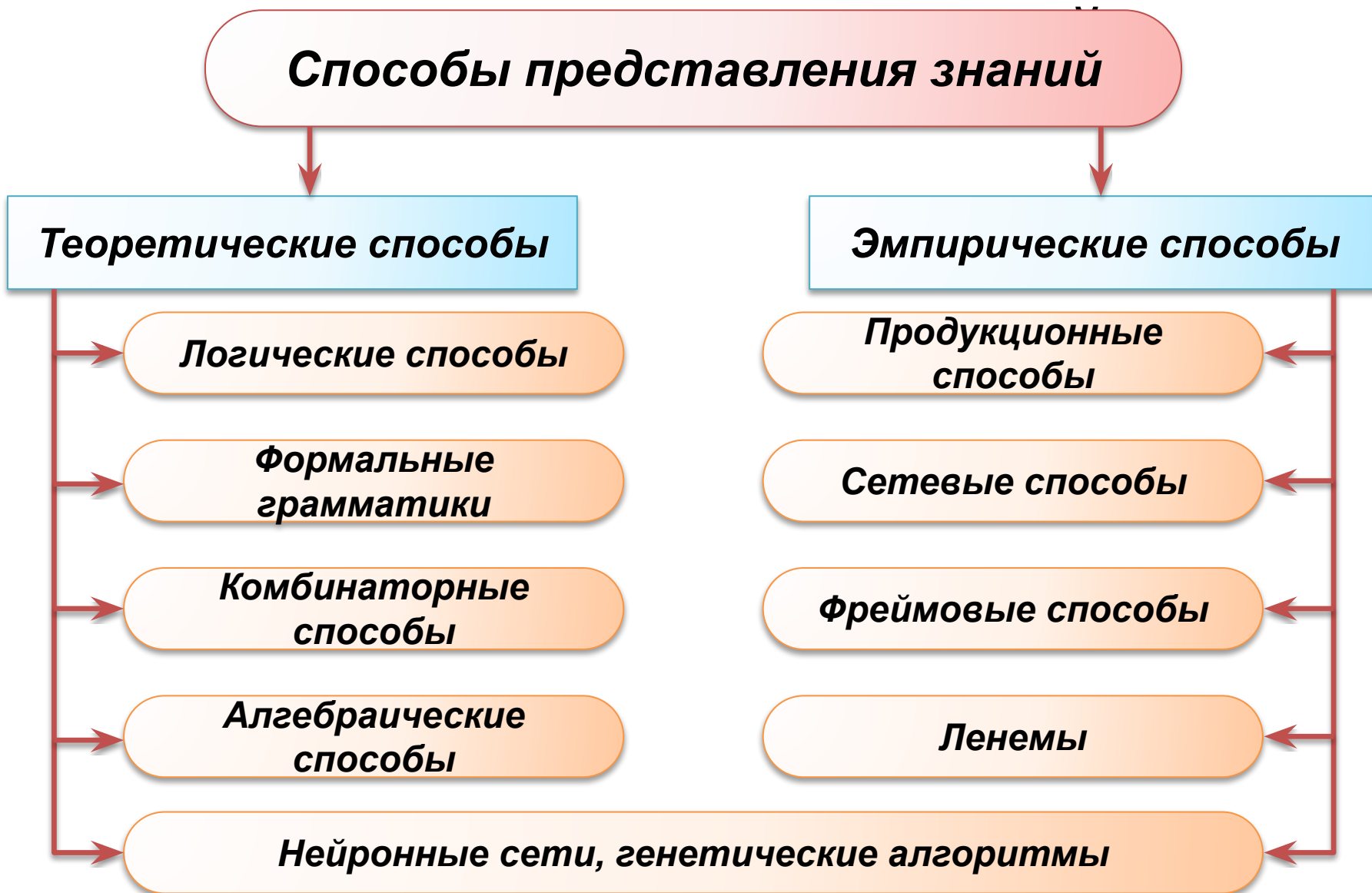
Эмпирические способы

Продукционные
способы

Сетевые способы

Фреймовые способы

Ленемы





Ленемы

Ленема очень напоминает фрейм, но сходство между ними чисто внешнее. **Ленема** предназначена для структурного комплексного описания понятий предметной области. По изобразительным возможностям **ленемы** более совершенны, чем такие традиционные модели представления знаний, как семантическая сеть, фрейм, система продукций.

Ленемы представляют собой смешанный тип модели, являющийся как бы «развитием» других моделей. Они предназначена для структурного комплексного описания понятий предметной области. По изобразительным возможностям **ленемы** более совершенны, чем такие традиционные модели



Формализованный способ представления знаний «Узел-Функция-Объект» - это современный способ представления знаний. В его основе лежит графоаналитическая технология моделирования организационных систем «Узел-Функция-Объект». Данный способ представления знаний совмещает в себе все традиционные подходы и позволяет хранить знания как понятийные, так и процедурные. В рамках данного способа представления знаний изучаемые объекты представляются в виде взаимосвязанных элементов: **узел** – структурная характеристика изучаемого объекта, позволяет описать взаимосвязь текущего объекта с другими; **функция** – функциональная характеристика изучаемого объекта



Нейронные сети и генетические

алгоритмы

Искусственная нейронная сеть (ИНС) — математическая модель и её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Используются в практических задачах: в прогнозировании, распознавании образов, управления и др.

Генетический алгоритм — это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путём случайного подбора, комбинирования и вариации искомых

Управление знаниями





Управление знаниями (knowledge management) – это формальный процесс, который состоит в оценке организационных процедур, людей и технологий, в создании системы, использующей взаимосвязи между этими компонентами с целью предоставления нужной информации нужным людям в нужное время, что приводит к повышению продуктивности.



Управление знаниями включает в себя комплекс формализованных методов

- поиск и извлечение знаний из живых и неживых объектов (носителей знаний);
- структурирование и систематизацию знаний (для обеспечения их удобного хранения и поиска);
- анализ знаний (выявление зависимостей и аналогий);
- обновление (актуализацию) знаний;
- распространение знаний;
- генерацию новых знаний;



Системы управления знаниями (СУЗ)

Для управления знаниями в организациях используют СУЗы (электронные системы хранения и управления данными – сервера, облака и др.). Предприятие не способно управлять своим интеллектуальным капиталом в отсутствии СУЗ.

Интеллектуальные активы (авторские свидетельства, патенты, знания о товарах, покупателях, конкурентах, знания своих сотрудников) предприятия увеличивают его конкурентоспособность и рыночную стоимость.

В рамках СУЗ особое внимание надо уделить



В процессе управления знаниями выделяются следующие функции:

- создание** – функция, результатом которой являются новые знания или новые конфигурации существующих знаний;
- выявление** – функция, которая делает неявные знания явными, т.е. преобразует индивидуальные знания в знания предприятия (его сотрудников);
- организация знаний** – функция по классификации и категоризации знаний для навигации, запоминания, поиска и сопровождения знаний;
- доступ** – функция по передаче и распространению знаний между сотрудниками;
- использование** – функция по применению знаний для принятия решений и расширения возможностей;



Три основных компонента управления знаниями:

- **люди** получают, генерируют и передают знания;
- **процессы** используются для распространения знаний;
- **технологии** обеспечивают быструю и эффективную работу людей и процессов.

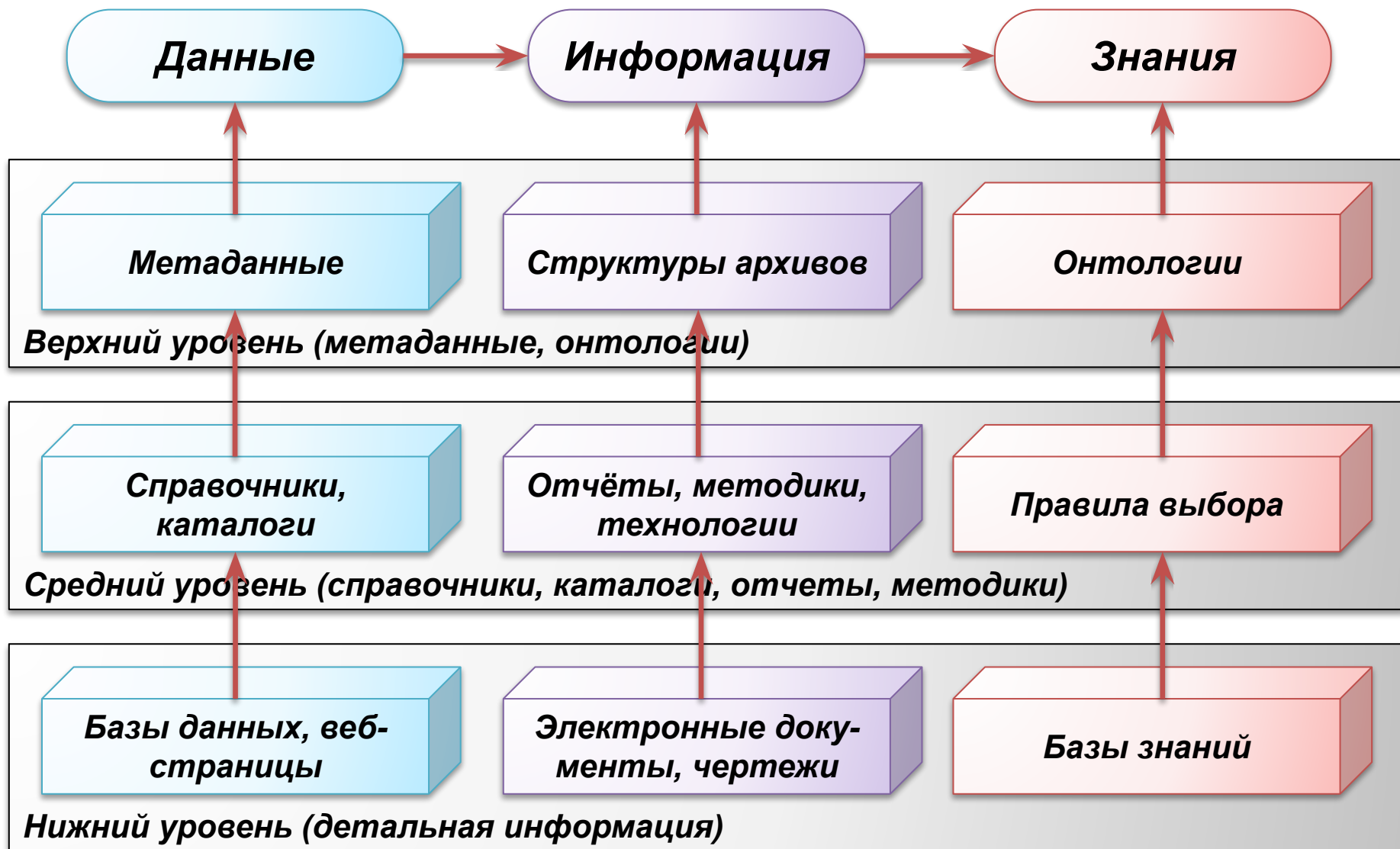


**В зависимости от архитектуры и
подхода к построению СУЗ, можно**

строить :

- **корпоративная память** (база знаний) – упор делается на программную реализацию задач получения, хранения и интегрирования знаний;
- **корпоративный портал знаний** – упор делается на программную реализацию задач уровня пользовательских приложений поиска и визуализации знаний;
- **семантический веб (СВ)** – веб-сайты с возможностью осмысленного манипулирования их данными с помощью

Корпоративная память (КП)



Основные функции корпоративной

памяти:

- сбор и систематическая организация информации;
- интеграция с существующими автоматизированными системами;
- обеспечение нужной информации по запросу (пассивная форма) и при необходимости (активная форма).

Конечная цель КП состоит в обеспечении доступа к знанию всякий раз, когда это необходимо. КП реализует активный подход к распространению знаний, который не полагается на запросы пользователей, а автоматически предоставляет полезное для решения задачи знание.

Законченная система должна действовать как



Онтологія :

Онтологія — учение о сущем; учение о бытии как таковом; раздел философии, изучающий фундаментальные принципы бытия, его наиболее общие сущности и категории, структуру и закономерности.

Онтологія в информатике — это попытка всеобъемлющей и подробной формализации некоторой области знаний с помощью концептуальной схемы. Обычно такая схема состоит из структуры данных, содержащей все релевантные классы объектов, их связи и правила (теоремы, ограничения), принятые в этой области.

Современные онтологии строятся по следующей



Экземпляры — это основные компоненты онтологии. Экземпляры могут представлять собой как физические объекты (люди, дома, планеты), так и абстрактные (числа, слова). Строго говоря, онтология может обойтись и без конкретных объектов. Однако, одной из главных целей онтологии является классификация таких объектов, поэтому они также включаются.

Понятия — абстрактные группы, коллекции или наборы объектов. Они могут включать в себя экземпляры, другие классы, либо же сочетания и того, и другого.

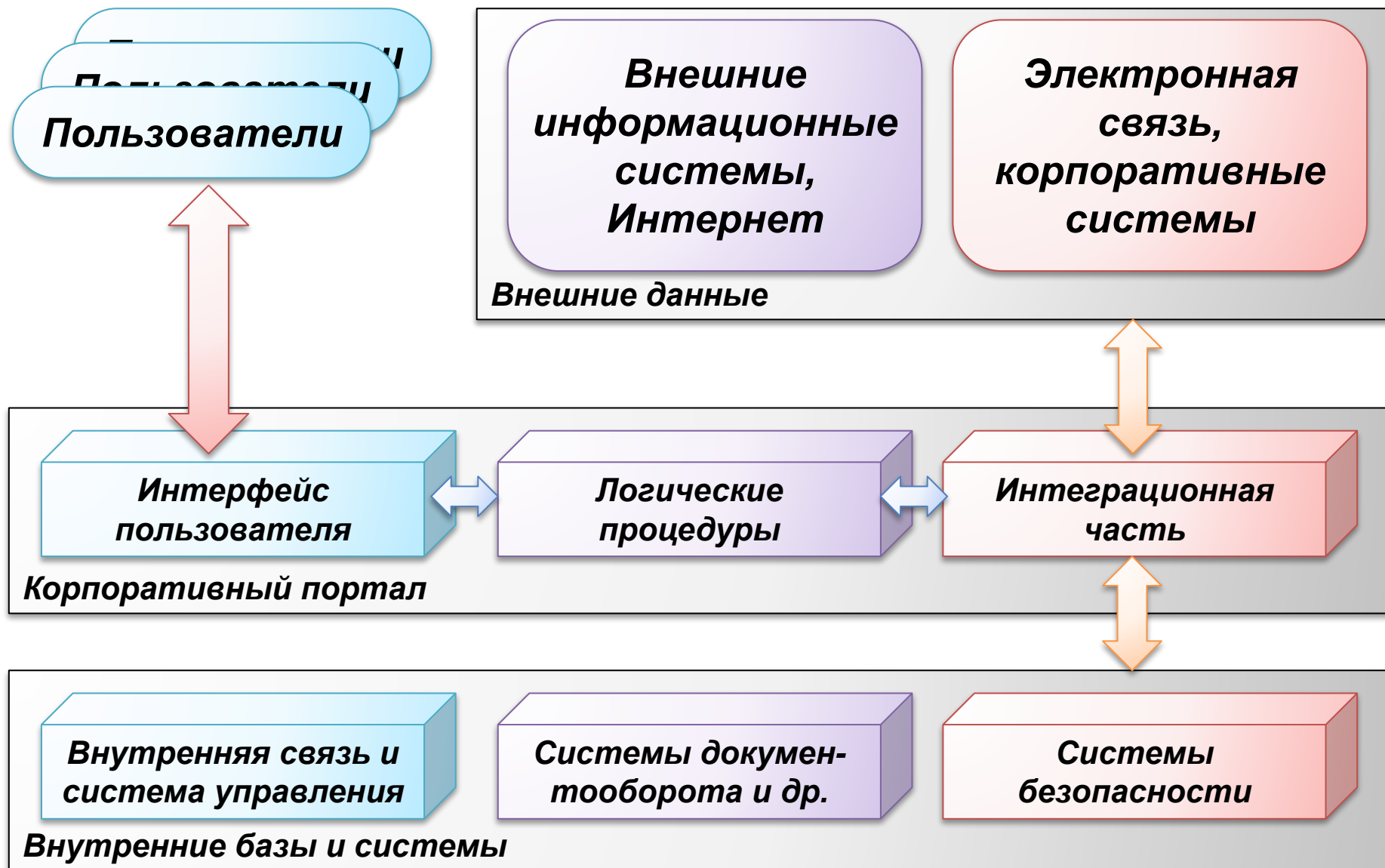
Пример: Понятие «люди», вложенное понятие «человек». Чем является «человек» — вложенным понятием, или экземпляром



Атрибуты — все объекты в онтологии могут иметь атрибуты. Каждый атрибут имеет по крайней мере имя и значение и используется для хранения информации, которая специфична для объекта и привязана к нему.

Отношения — важная роль атрибутов заключается в том, чтобы определять отношения (зависимости) между объектами онтологии. Обычно отношением является атрибут, значением которого является другой объект.

Корпоративный портал знаний





Основные свойства порталов :

- систематизация контента и предоставление удобных средств поиска и навигации;
- управление контентом для различных групп сотрудников (управление доступом);
- внутрикорпоративный обмен знаниями и совместная работа за счет наличия различных конференций, форумов и единого рабочего пространства;
- индивидуальный персонализированный интерфейс;
- возможность работы с онтологиями.



Семантический Веб :

Это направление развития Всемирной паутины, целью которого является представление информации в виде, пригодном для машинной обработки.

Основная идея этого проекта заключается в организации такого представления данных в сети, чтобы допускалась не только их визуализация, но и их эффективная автоматическая обработка программами разных производителей. По замыслу создателя (Тим Бернерс-Ли) Семантический Web должен обеспечить "понимание" информации компьютерами, выделение ими наиболее подходящих по тем или иным критериям данных, и уже после этого предоставление информации под запросам