

Тема 6. Информационно-аналитическая поддержка процессов разработки и принятия управленческих решений

6.1. СППР и современные информационные системы менеджмента

Основным объектом, основной категорией процесса управления является *информация*. Хотя информация является общенаучной (философской) категорией, строгого понятия этого феномена нет. Понятие «информация», условно распределяется на следующие 4-е группы, каждая из которых опирается на:

- содержательное понимание;
- формализованные модели: мера разнообразия;
- теорию познания и отражения;
- учет связи информации со свойствами материи.

Сущность информации описывается в рамках двух концепций: атрибутивной и функциональной и содержится в ответе на вопрос о наличии информации в неживой природе.

Первая концепция рассматривает информацию как атрибут, присущий всем уровням материи, в то время как вторая концепция связывает информацию только с самоуправляемыми и самоорганизующимися системами.

Информационный процесс - это реализация определенного информационного взаимодействия.

Взаимодействие - фундаментальная категория, отражающая процессы воздействия различных объектов друг на друга.

Основными особенностями информационного взаимодействия являются следующие:

- информация выступает в качестве объекта;
- взаимодействие осуществляется с помощью очень слабых сигналов;
- операции осуществляются со смыслами, образами, эмоциями;
- наличие памяти (интеллекта) у взаимодействующих объектов.

Информационные технологии - это последовательности действий по преобразованию информационных ресурсов в информационные продукты. Различают инфраструктурные технологии, технологии работы с данными, технологии совместной работы, технологии бизнес-приложений, интернет-технологии и др. Информационные технологии реализуются в *информационных системах*, которые можно описать с помощью функциональной структуры и информационного, технического (hardware), алгоритмического (brainware), программного (software), организационного (orgware) и кадрового обеспечений.

Первыми стали *информационные системы, предназначенные для обработки электронных данных – СОД (Electronic Data Processing – EDP)*. К таким системам относятся системы, реализующие контрольные (учетные) функции, например, бухгалтерские и банковские системы. СОД предназначены, как правило,

для решения задач *учета* и автоматизируют такие операции, как сбор, хранение, обработка, отображение и регистрация данных. Результаты работы такой системы представляются, как правило, в виде видеogramм и различных по форме отчетов. Классическим примером такой системы является электронная бухгалтерия. Основной функцией СОД является *сбор и обработка данных*, которые описывают работу фирмы в разные периоды времени. К обработке обычно относятся следующие операции: классификация; сортировка; вычисления; агрегирование; хранение; отображение и регистрация данных. К отличительным особенностям СОД можно отнести: выполнение задач, регламентированных законодательными актами (бухгалтерские системы); решение только хорошо структурированных задач, с заранее известным алгоритмом их решения; автоматизация рутинных расчетов с минимальным участием человека; использование подробных (хорошо детализированных) данных с акцентом на хронологию событий. Таким образом, СОД предназначены для решения в хронологическом порядке хорошо структурированных задач путем обработки детализированных данных с использованием стандартных процедур. Основной целью СОД является автоматизация рутинных расчетов.

Дальнейшее развитие информационных технологий позволяет переместить акцент с автоматизации функций контроля (учета) на автоматизацию других функций управления. Сначала появляются, так называемые, *информационные системы (оперативного) управления – ИСУ*, которые автоматизируют операции по составлению управленческих отчетов. Затем появляются *системы поддержки принятия решений – СППР*, которые облегчают решение слабо структурированных задач.

Информационные системы (оперативного) управления - ИСУ (Management Information System - MIS) предназначены для автоматизации таких функций, как: учет, регулирование и частично функции анализа. ИСУ используется при решении менее структурированных задач по сравнению с СОД. ИСУ включает СОД и осуществляет поиск и обработку информации, поступающей из СОД и внешнего окружения. ИСУ обеспечивает возможность работы в запросно-ответном режиме с помощью системы управления базами данных (СУБД). Результатом работы ИСУ является информация, которая представлена в виде, удобном для принятия решений менеджером. ИСУ не является полностью автоматической системой. Она предназначена для использования на любом уровне управления, где все решения принимает человек.

Система поддержки принятия решений - СППР (Decision Support System – DSS) представляет собой вид информационной системы, предназначенной для помощи менеджеру при решении плохо структурированных задач, возникающих в процессе принятия решений. Она предназначена для автоматизации таких функций, как: учет, регулирование, анализ и частично функции планирования. Отличительными особенностями такой системы является: возможность сочетания традиционных методов решения с методами математического моделирования; высокая адаптивность используемых моделей и методов; наличие человеко-машинного интерфейса, ориентированного на непрофессионального пользователя. Место СППР среди информационных систем можно определить,

построив информационную модель предприятия, которая будет включать следующие три уровня: уровень обработки данных (СОД); уровень обработки оперативной информации (ИСУ); уровень принятия решений (СППР). Таким образом, СППР включает в себя СОД и ИСУ, предназначена для решения плохо структурированных задач, имеет развитый человеко-машинный интерфейс и строится на основе экспертных систем.

Дальнейшее развитие информационных систем, используемых для автоматизации процессов разработки управленческих решений, идет по пути моделирования процессов человеческого мышления, который получил название искусственный интеллект. Важнейшим приложением искусственного интеллекта стали *экспертные системы*, представляющие собой новый класс компьютерных информационных систем, основанных на обработке знаний.

Экспертные системы (ЭС) представляют собой раздел искусственного интеллекта и используются в СППР для повышения производительности и качества принимаемых решений. Экспертная система в СППР реализует следующие функции: распознавание сложившейся ситуации, ее анализ, постановку диагноза и формулирование ближайших целей для возвращения на желаемую траекторию развития предприятия; построение путей достижения сформулированных целей с учетом резервов предприятия; пополнение Базы Знаний; обеспечение дружественного интерфейса.

Экспертные системы (ЭС), используемые в менеджменте для поддержки процессов принятия решений, состоят из блоков, представленных на рис.6.1-1, и включают: базы знаний и данных; блок логических выводов; блок объяснений; блок ввода и корректировки данных; блок приобретения знаний; блок проведения расчетов; пользовательский интерфейс.

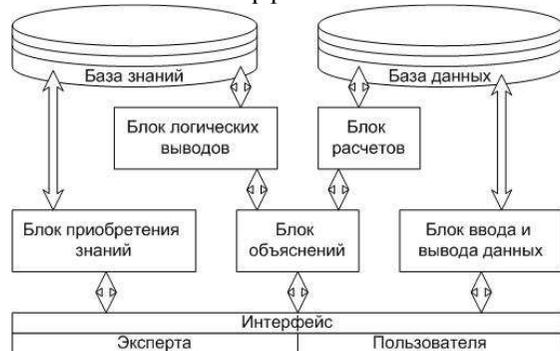


Рис.6.1-1. Структура экспертной системы поддержки принятия решения

Базы знаний строятся на основе моделей, с помощью которых отражают знания эксперта о предметной области, способы анализа поступающих фактов и методы вывода, т.е. порождения новых знаний на основе имеющихся и вновь поступивших. Наиболее распространенными являются следующие виды моделей: логические, продукционные, фреймовые и семантические сети.

Логические модели базируются на представлении знаний и системе логики предикатов первого порядка. Вывод новых знаний осуществляется на основании силлогизмов. Правила формальной логики постепенно расширяются, приближаясь к "человеческой" логике. Последняя характеризуется нечеткостью. В связи с этим появляются модальная, многозначная, немонотонная, псевдофизическая и другие виды логики.

Продукционные модели представляют знания в форме предикатов первого порядка, а *правила* манипулирования ими - с помощью конструкций "ЕСЛИ - ТО". *Фреймовое представление* знаний отражает систематизированную в виде единой теории психологическую модель памяти человека. Основным элементом *модели* - фрейм отражает структуру данных для описания концептуальных (понятийных) объектов. Информация, относящаяся к одному фрейму, содержится в слоте. Все фреймы взаимосвязаны и образуют единую систему, в которой объединены факты (описательные знания) и правила манипулирования ими.

Семантическая сеть - граф, узлы которого соответствуют понятиям или объектам. *Логические* выводы базируются на прямом или обратном рассуждениях. Прямые рассуждения, ведутся от данных к цели рассуждения, а обратные - от цели к данным. Обратные рассуждения базируются на графе И/ИЛИ, связывающем в единое целое факты и заключения. Оценка этого графа и есть логический вывод. При этом оцениваются лишь те части графа, которые имеют отношение к заключению.

Блок логических выводов предназначен для извлечения знаний из введенной в систему информации и должен быть рассчитан для работы с ненадежными данными. Он использует нечеткую логику, коэффициенты уверенности, байесовскую логику, меры доверия и т.д. Этот блок является главным, так как с его помощью пользователь генерирует альтернативы и вырабатывает решение. Здесь реализуются алгоритмы факторного анализа показателей, результаты которого используются для постановки диагноза. Очевидно, что количество диагнозов может быть велико, поэтому используют иерархические способы их описания, что позволяет снизить их сложность на несколько порядков и облегчить процесс выбора решения.

Блок объяснений показывает весь путь, которым система пришла к тому или иному выводу. В ЭС, основанных на правилах, объяснения получают путем прослеживания шагов рассуждения, которые привели к данному выводу.

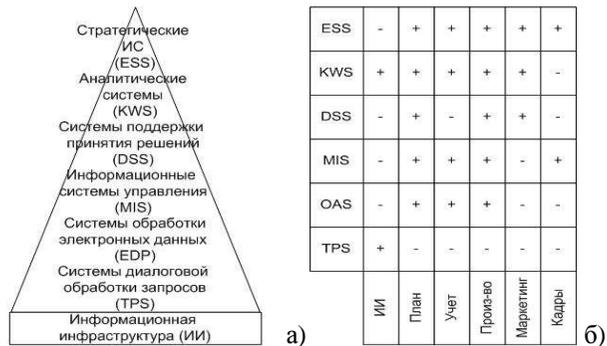
База данных, *Блок проведения расчетов* и *Блок ввода/вывода и корректировки данных* являются отличительной чертой, присущей только ЭС поддержки процессов принятия решений. Их наличие продиктовано большим объемом и жесткостью требований к точности расчетов, необходимых для принятия решений по управлению именно экономическими системами. В базе данных находятся плановые, фактические, расчетные, отчетные и другие показатели. Если эти блоки входят в соответствующие ИСУ и СОД как составляющие их части, то в ЭС они не рассматриваются.

Блок приобретения знаний, в конечном счете, отвечает за самообучение экспертной системы и, следовательно, за надежность и точность выдаваемых рекомендаций. Извлечение и формализация знаний, которыми обладает эксперт, -

достаточно сложный процесс, который изучает такой раздел науки, как инженерия знаний. Здесь эти вопросы мы рассматривать не будем. Лишь отметим, что в настоящее время приобретение знаний экспертной системой сводится к вводу вербализованных знаний эксперта с помощью соответствующего пользовательского интерфейса. Рассмотрим место СППР в корпоративной информационной системе управления предприятия.

Корпоративные (интегрированные) информационные системы. В каждой организации имеются различные уровни управления, на которых циркулируют специфичные информационные потоки. Для обработки информации используются различные информационные технологии, которые реализуются с помощью соответствующих информационных систем, имеющих собственные названия.

Корпоративные (интегрированные) информационные системы (КИС) управления в каждой организации можно описывать: по уровням управления (рис.6.1-2 а), по базовым функциям (рис.6.1-2 б) и по операциям обработки информации (рис.6.1-2 в).



ESS	-	+	+	+	+	+
KWS	+	+	+	+	+	-
DSS	-	+	-	+	+	-
MIS	-	+	+	+	-	+
OAS	-	+	+	+	-	-
TPS	+	-	-	-	-	-
	ИИ	План	Учет	Произво	Маркетинг	Кадры

б)

Уровень	Вход	Обработка	Выход	Пользователи
ESS	С овокупные данные	Анализ и принятие решений, моделирование	Решения, стратегии, планы	Высшее руководство
KWS	Технологические данные, баз знаний	Моделирование, анализ, прогнозирование	Модели, результаты анализа, графика, таблицы, отчеты	Аналитики, ИТ профессионалы
DSS	Слабоформализованные данные, аналитические модели	Моделирование, разработка альтернатив	Альтернативы и результаты их анализа	Средний персонал управления
MIS	Итоговые оперативные данные, данные большого объема, простые модели	Обычные отчеты, простые модели, простейший анализ	Предложения, возражения, указания	Управляющие, линейные менеджеры, операторы
OAS	Документы, расписания	Контроль выполнения, распоряжения, связь	Документы, графика, почта, сводки	Служащие, персонал
TPS	Запросы, документы	Организация, списание, модификация	Отчеты, доклады, списки	Оперативный и технический персонал

в)

Рис.6.1-2. Стратификация КИС по уровням управления (а), базовым функциям (б) и операциям обработки информации (в).

Стратегические информационные системы – СИС корпоративного типа (*Enterprise Strategic System - ESS*) предназначены для оказания помощи высшему руководству компании (Top Managers) в процессе поддержки принятия стратегических решений. ESS учитывают долгосрочные изменения, происходящие в окружающей среде и деловом окружении предприятия, интегрируют в себе знания и данные всех информационных систем предприятия и строятся, как правило, на базе систем искусственного интеллекта (экспертных систем – ЭС). Их назначение – приводить в соответствие изменения в условиях эксплуатации с существующей организационной возможностью.

Аналитические системы или чаще *Система управления знаниями* – СУЗ (*Knowledge Management System – KWS*) - это набор повторяемых на регулярной основе управленческих процедур, призванных повысить эффективность сбора, хранения, распространения и использования ценной информации с точки зрения компании.

Классическое управление знаниями включает в себя следующие пять фундаментальных процессов: *создание* (результатом является новое знание); *поиск* (поиск и представление неявных знаний в явной форме, что делает возможным сбор индивидуальных знаний для коллективного использования); *систематизация* или организация (классификация и категоризация знаний с целью их последующего целенаправленного извлечения); *поддержание* целостности данных за счет реализации соответствующих процессов); *доступ* (действие, с помощью которого знания посылаются или запрашиваются конкретным пользователем); *использование* (применение знаний в работе, принятии решений и реализации возможностей).

Наиболее популярными направлениями при внедрении системы управления знаниями являются: формирование лучших практик компании; повышение эффективности обучения; сбор идей и информации для создания новых товаров; формирование библиотеки выполненных работ, для минимизации повторов и нерационального расходования времени; оптимизация информационных потоков предприятия; сокращение уровней управления и повышение эффективности деятельности управляющей структуры; создание новых знаний и подходов; формирование базы знаний для регулярного пересмотра стратегии компании и т.д.

Информационные системы управления - ИСУ (Management Information System - MIS) ориентированы на тактический уровень управления: среднесрочное планирование, анализ и организацию работ в течение нескольких недель (месяцев), например анализ и планирование поставок, сбыта, составление производственных программ. Для данного класса задач характерны регламентированность (периодическая повторяемость) формирования результирующих документов и четко определенный алгоритм решения задач, например свод заказов для формирования производственной программы и определение потребности в комплектующих деталях и материалах на основе спецификации изделий. Решение подобных задач предназначено для руководителей различных служб предприятий (отделов материально-технического снабжения и сбыта, цехов и т.д.). Задачи решаются на основе накопленной базы оперативных данных.

Системы обработки (электронных) данных - СОД (Electronic Data Processing - EDP) предназначены для учета и оперативного регулирования хозяйственных операций, подготовки стандартных документов для внешней среды (счетов, накладных, платежных поручений). Горизонт оперативного управления хозяйственными процессами составляет от одного до несколько дней и реализует регистрацию и обработку событий, например оформление и мониторинг выполнения заказов, приход и расход материальных ценностей на складе, ведение табеля учета рабочего времени и т.д. Эти задачи имеют итеративный, регулярный характер, выполняются непосредственными исполнителями хозяйственных процессов (рабочими, кладовщиками, администраторами и т.д.) и связаны с оформлением и пересылкой документов в соответствии с четко определенными алгоритмами. Результаты выполнения хозяйственных операций через экранные формы вводятся в базу данных.

Информационная система, которая включает первые три типа перечисленных выше информационных систем, называется стратегической информационной системой - СИС. В зависимости от охвата функций и уровней управления различают корпоративные (интегрированные) и локальные ИС.

Корпоративная (интегрированная) информационная система (КИС) автоматизирует все функции управления на всех уровнях управления, является многопользовательской и функционирует в распределенной вычислительной сети. *Локальная* информационная система (ЛИС) автоматизирует отдельные функции управления на отдельных уровнях управления и может быть однопользовательской, функционирующей в отдельных подразделениях системы управления.

Системы поддержки принятия решений.

Системы поддержки принятия решений - СППР (Decision Support System - DSS) используются в основном на верхнем уровне управления (руководства фирм, предприятий, организаций), имеющего стратегическое долгосрочное значение в течение года или нескольких лет. К таким задачам относятся формирование стратегических целей, планирование привлечения ресурсов, источников финансирования, выбор места размещения предприятий и т.д. Реже задачи класса СППР решаются на тактическом уровне, например при выборе поставщиков или заключении контрактов с клиентами. Задачи СППР имеют, как правило, нерегулярный характер. Для задач СППР свойственны недостаточность имеющейся информации, ее противоречивость и нечеткость, преобладание качественных оценок целей и ограничений, слабая формализуемость алгоритмов решения. В качестве инструментов обобщения чаще всего используются средства составления аналитических отчетов произвольной формы, методы статистического анализа, экспертных оценок и систем, математического и имитационного моделирования. При этом используются базы обобщенной информации, информационные хранилища, базы знаний о правилах и моделях принятия решений.

Характеристики СППР.

Согласно Е. Turban², СППР обладает следующими четырьмя основными характеристиками:

1. СППР использует и данные, и модели;
2. СППР предназначены для помощи менеджерам в принятии решений для слабоструктурированных и неструктурированных задач;
3. Они поддерживают, а не заменяют, выработку решений менеджерами;
4. Цель СППР — улучшение эффективности решений.

Е. Turban предложил список характеристик идеальной СППР (которая имеет мало общих элементов с определенным, приведенным выше). Идеальная СППР:

1. оперирует со слабоструктурированными решениями;
2. предназначена для ЛПР различного уровня;
3. может быть адаптирована для группового и индивидуального использования;
4. поддерживает как взаимозависимые, так и последовательные решения;
5. поддерживает 3 фазы процесса решения: интеллектуальную часть, проектирование и выбор;
6. поддерживает разнообразные стили и методы решения, что может быть полезно при решении задачи группой ЛПР;
7. является гибкой и адаптируется к изменениям как организации, так и ее окружения;
8. проста в использовании и модификации;
9. улучшает эффективность процесса принятия решений;
10. позволяет человеку управлять процессом принятия решений с помощью компьютера, а не наоборот;
11. поддерживает эволюционное использование и легко адаптируется к изменяющимся требованиям;
12. может быть легко построена, если может быть сформулирована логика конструкции СППР;
13. поддерживает моделирование;
14. позволяет использовать знания.

Классификации СППР

Для СППР отсутствует не только единое общепринятое определение, но и исчерпывающая классификация. Разные авторы предлагают разные классификации. На уровне пользователя Р. Haettenschwiler³ (1999) делит СППР на пассивные, активные и кооперативные СППР. *Пассивной* СППР называется система, которая помогает процессу принятия решения, но не может вынести предложение, какое решение принять. *Активная* СППР может сделать предложение, какое решение следует выбрать.

Кооперативная позволяет ЛПР изменять, пополнять или улучшать решения, предлагаемые системой, посылая затем эти изменения в систему для проверки.

² Turban, E. Decision support and expert systems: management support systems. -Englewood Cliffs, N.J.: Prentice Hall, 1995. — 887 p.

³ Haettenschwiler P. Neues anwenderfreundliches Konzept der Entscheidungs-unterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft. Zurich: Hochschulverlag AG, 1999. — S. 189—208.

Система изменяет, пополняет или улучшает эти решения и посылает их опять пользователю. Процесс продолжается до получения согласованного решения. На концептуальном уровне D.J. Power⁴ (2003) отличает:

1. СППР, управляемые сообщениями (Communication-Driven DSS, ранее групповая СППР — GDSS), поддерживает группу пользователей, работающих над выполнением общей задачи.
2. СППР, управляемые данными (Data-Driven DSS) или СППР, ориентированные на работу с данными (Data-oriented DSS) в основном ориентируются на доступ и манипуляции с данными.
3. СППР, управляемые документами (Document-Driven DSS), управляют, осуществляют поиск и манипулируют неструктурированной информацией, заданной в различных форматах.
4. СППР, управляемые знаниями (Knowledge-Driven DSS), обеспечивают решение задач в виде фактов, правил, процедур.
5. СППР, управляемые моделями (Model-Driven DSS), характеризуются в основном доступом и манипуляциями с математическими моделями (статистическими, финансовыми, оптимизационными, имитационными).
6. Отметим, что некоторые OLAP-системы, позволяющие осуществлять сложный анализ данных, могут быть отнесены к гибридным СППР, которые обеспечивают моделирование, поиск и обработку данных.

На техническом уровне D. Power⁵ (1997) различает СППР всего предприятия и настольную СППР. СППР всего предприятия подключена к большим хранилищам информации и обслуживает многих менеджеров предприятия. Настольная СППР — это малая система, обслуживающая лишь один компьютер пользователя. Существуют и другие классификации (Alter⁶, Holsapple и Whinston⁷, Golden, Hevner и Power⁸). Отметим лишь, что превосходная для своего времени классификация Alter'a, которая разбивала все СППР на 7 классов, в настоящее время несколько устарела.

В зависимости от данных, с которыми эти системы работают, СППР условно можно разделить на оперативные и стратегические.

Оперативные СППР предназначены для немедленного реагирования на изменения текущей ситуации в управлении финансово-хозяйственными процессами компании.

Стратегические СППР ориентированы на анализ значительных объемов разнородной информации, собираемых из различных источников. Важнейшей целью этих СППР является поиск наиболее рациональных вариантов развития

⁴ Power D.J. A Brief History of Decision Support Systems. DSSResources.COM, World Wide Web, <http://DSSResources.COM/history/dsshistory.html>, version 2.8, May 31, 2003.

⁵ Power D. J. «What is a DSS?» // The On-Line Executive Journal for Data-Intensive Decision Support, 1997. — v. 1. — N3.

⁶ Alter S. L. Decision support systems: current practice and continuing challenges. Reading, Mass.: Addison-Wesley Pub., 1980.

⁷ Holsapple C.W., Whinston A.B. Decision Support Systems: A Knowledge-based Approach. — Minneapolis: West Publishing Co., 1996.

⁸ Golden B., Hevner A., Power D.J. Decision Insight Systems: A Critical Evaluation // Computers and Operations Research, 1986. — v. 13. — N2/3. — p. 287—300.

бизнеса компании с учетом влияния различных факторов, таких как конъюнктура целевых для компании рынков, изменения финансовых рынков и рынков капиталов, изменения в законодательстве и др.

СППР *первого* типа получили название Информационных Систем Руководства (Executive Information Systems - EIS). По сути, они представляют собой конечные наборы отчетов, построенных на основании данных из транзакционной информационной системы предприятия, в идеале адекватно отражающих в режиме реального времени основные аспекты производственной и финансовой деятельности. Для EIS характерны следующие основные черты:

- отчеты, как правило, базируются на стандартных для организации запросах; число последних относительно невелико;
- ИСР представляет отчеты в максимально удобном виде, включающем, наряду с таблицами, деловую графику, мультимедийные возможности и т. п.;
- как правило, ИСР ориентированы на конкретный вертикальный рынок, например финансы, маркетинг, управление ресурсами.

СППР *второго* типа предполагают достаточно глубокую проработку данных, специально преобразованных так, чтобы их было удобно использовать в ходе процесса принятия решений. Неотъемлемым компонентом СППР этого уровня являются правила принятия решений, которые на основе агрегированных данных дают возможность менеджерам компании обосновывать свои решения, использовать факторы устойчивого роста бизнеса компании и снижать риски. СППР второго типа в последнее время активно развиваются. Технологии этого типа строятся на принципах многомерного представления и анализа данных (OLAP).

При создании СППР можно использовать Web-технологии. В настоящее время СППР на основе Web-технологий для ряда компаний являются синонимами СППР предприятия.

Методы поддержки принятия управленческих решений на основе ИТ.

Для поддержки принятия решений с помощью информационных технологий, включая анализ и выработку альтернатив, в СППР используются следующие методы:

- 1) информационный поиск;
- 2) интеллектуальный анализ данных;
- 3) извлечение (поиск) знаний в базах данных;
- 4) рассуждение на основе прецедентов;
- 5) имитационное моделирование;
- 6) генетические алгоритмы;
- 7) искусственные нейронные сети;
- 8) методы искусственного интеллекта.

Информационный поиск (ИП) (англ. *Information retrieval*) — процесс поиска *неструктурированной* документальной информации и наука об этом поиске⁹.

⁹ <http://ru.wikipedia.org/>

Термин «информационный поиск» был впервые введен Кельвином Муром в 1948. Процесс поиска включает последовательность операций, направленных на сбор, обработку и предоставление необходимой информации заинтересованным лицам. В общем случае поиск информации состоит из четырех этапов:

1. определение (уточнение) информационной потребности и формулировка информационного запроса;
2. определение совокупности возможных держателей информационных массивов (источников);
3. извлечение информации из выявленных информационных массивов;
4. ознакомление с полученной информацией и оценка результатов поиска.

Различают: *полнотекстовый поиск* — поиск по всему содержимому документа; *поиск по метаданным* — это поиск по неким атрибутам документа, поддерживаемым системой (название документа, дата создания, размер, автор и т. д.); *поиск по изображению* — поиск по содержанию изображения.

К методами информационного относятся: *адресный поиск* - процесс поиска документов по чисто формальным признакам, указанным в запросе; *семантический поиск* - процесс поиска документов по их содержанию; *документальный поиск* - процесс поиска в хранилище информационно-поисковой системы первичных документов или в базе данных вторичных документов, соответствующих запросу пользователя; *фактографический поиск* - процесс поиска фактов, соответствующих информационному запросу.

Интеллектуальный анализ данных (англ. *Data Mining*) — выявление скрытых закономерностей или взаимосвязей между переменными в больших массивах необработанных данных. Подразделяется на задачи классификации, моделирования и прогнозирования и другие. Термин «Data Mining» введен Григорием Пятецким-Шапиро в 1989 году. Задачи, решаемые Data Mining: классификация; кластеризация; сокращение описания; ассоциация; прогнозирование; анализ отклонений; визуализация.

Извлечение (поиск) знаний в базах данных (Knowledge Discovery in Databases – KDD) - процесс обнаружения полезных знаний в базах данных. Эти знания могут быть представлены в виде закономерностей, правил, прогнозов, связей между элементами данных и др. Главным инструментом поиска знаний в процессе KDD являются аналитические технологии Data Mining, реализующие задачи классификации, кластеризации, регрессии, прогнозирования, предсказания и т.д. KDD включает последовательность операций, необходимых для поддержки аналитического процесса. К ним относятся: консолидация данных; подготовка анализируемых выборок данных (в том числе, обучающих); очистка данных от факторов, мешающих их корректному анализу; трансформация – оптимизация данных для решения определенной задачи; анализ данных; интерпретация и визуализация результатов анализа, их применение в бизнес-приложениях.

Рассуждение на основе прецедентов (Case-Based Reasoning). Прецедент - случай, имевший место ранее и служащий примером или оправданием для последующих случаев подобного рода. Вывод на основе прецедентов (*CBR – Case-Based Reasoning*) является подходом, позволяющим решить новую задачу, используя или адаптируя решение уже известной задачи. Как правило, такие ме-

тоды рассуждений включают в себя четыре основных этапа, образующие так называемый цикл рассуждения на основе прецедентов или CBR-цикл. Существует целый ряд методов извлечения прецедентов и их модификаций: метод ближайшего соседа (NN – Nearest Neighbor); метод извлечения прецедентов на основе деревьев решений; метод извлечения прецедентов на основе знаний; метод извлечения с учетом применимости прецедентов. Помимо рассмотренных методов для извлечения прецедентов могут успешно применяться и другие методы (например, аппарат искусственных нейронных сетей).

Имитационное моделирование — это метод, позволяющий строить модели, описывающие процессы так, как они проходили бы в действительности. Такую модель можно «проиграть» во времени как для одного испытания, так и заданного их множества. При этом результаты будут определяться случайным характером процессов. По этим данным можно получить достаточно устойчивую статистику.

Различают следующие виды имитационного моделирования: *агентное моделирование* – относительно новое (1990е-2000е гг.) направление в имитационном моделировании, которое используется для исследования децентрализованных систем; *дискретно-событийное моделирование* – подход к моделированию, предлагающий абстрагироваться от непрерывной природы событий и рассматривать только основные *события* моделируемой системы, такие как: «ожидание», «обработка заказа», «движение с грузом», «разгрузка» и другие; *системная динамика* – парадигма моделирования, где для исследуемой системы строятся графические диаграммы причинных связей и глобальных влияний одних параметров на другие во времени, а затем созданная на основе этих диаграмм модель имитируется на компьютере. Популярными системами имитационного моделирования: AnyLogic; Arena; eM-Plant; Powersim; GPSS.

Генетический алгоритм (англ. *genetic algorithm*) — это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем последовательного подбора, комбинирования и вариации искомым параметров с использованием механизмов, напоминающих биологическую эволюцию. Является разновидностью эволюционных вычислений (англ. *evolutionary computation*). Отличительной особенностью генетического алгоритма является акцент на использование оператора «скрещивания», который производит операцию рекомбинации решений-кандидатов, роль которой аналогична роли скрещивания в живой природе. Генетические алгоритмы служат, главным образом, для поиска решений в очень больших, сложных пространствах поиска.

Искусственные нейронные сети (ИНС) — математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге при мышлении, и при попытке смоделировать эти процессы. Первой такой моделью мозга был перцептрон. Впоследствии эти модели стали использовать в практических целях, как правило, в задачах прогнозирования. С точки зрения *машинного обучения*, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов

кластеризации и т. п. С *математической* точки зрения обучение нейронных сетей, это многопараметрическая задача нелинейной оптимизации. С точки зрения *кибернетики*, нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники. С точки зрения *развития вычислительной техники и программирования*, нейронная сеть — способ решения проблемы эффективного параллелизма. А с точки зрения *искусственного интеллекта*, ИНС является основой философского течения коннективизма и основным направлением в структурном подходе по изучению возможности построения (моделирования) естественного интеллекта с помощью компьютерных алгоритмов.

Искусственный интеллект (ИИ) (англ. *Artificial intelligence, AI*) — это наука и разработка интеллектуальных машин и систем, особенно интеллектуальных компьютерных программ, направленных на то, чтобы понять человеческий интеллект. При этом используемые методы не обязаны быть биологически правдоподобны. Но проблема состоит в том, что неизвестно какие вычислительные процедуры мы хотим называть интеллектуальным. А так как мы понимаем только некоторые механизмы интеллекта, то под интеллектом в пределах этой науки мы понимаем только вычислительную часть способности достигнуть целей в мире. Если в основе работы СППР лежат методы искусственного интеллекта, то говорят об *интеллектуальной СППР*, или ИСППР.

Архитектура СППР

Архитектура СППР представляется разными авторами по-разному. Приведем пример. G.M. Marakas¹⁰ (1999) предложил обобщенную архитектуру, состоящую из 5 различных частей:

- (a) система управления данными (the data management system — DBMS),
- (b) система управления моделями (the model management system — MBMS),
- (c) машина знаний (the knowledge engine (KE)),
- (d) интерфейс пользователя (the user interface),
- (e) пользователи (the user).

На сегодняшний день можно выделить четыре наиболее популярных типа архитектур СППР:

1. Функциональная СППР.
2. Независимые витрины данных.
3. Двухуровневое хранилище данных.
4. Трехуровневое хранилище данных.

Функциональная СППР (Рис.6.2-3) является наиболее простой с архитектурной точки зрения. Такие системы часто встречаются на практике, в особенности в организациях с невысоким уровнем аналитической культуры и недостаточно развитой информационной инфраструктурой. Характерной чертой функциональной СППР является то, что анализ осуществляется с использованием данных из оперативных систем.

Преимущества: быстрое внедрение за счет отсутствия этапа перегрузки данных в специализированную систему; минимальные затраты за счет использования одной платформы.

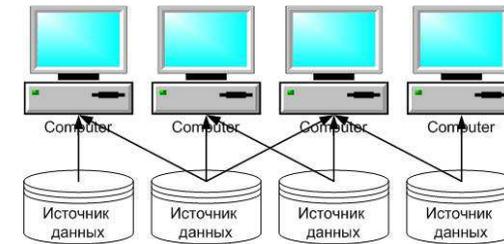


Рис. 6.2-3. Функциональная СППР

Недостатки: единственный источник данных, потенциально сужающий круг вопросов, на которые может ответить система; оперативные системы характеризуются очень низким качеством данных с точки зрения их роли в поддержке принятия стратегических решений. В силу отсутствия этапа очистки данных, данные функциональной СППР, как правило, обладают невысоким качеством; большая нагрузка на оперативную систему. Сложные запросы могут привести к остановке работы оперативной системы, что весьма нежелательно *СППР с использованием независимых витрин данных*. Независимые витрины данных (Рис.6.2-4) часто появляются в организации исторически и встречаются в крупных организациях с большим количеством независимых подразделений, зачастую имеющих свои собственные отделы информационных технологий.

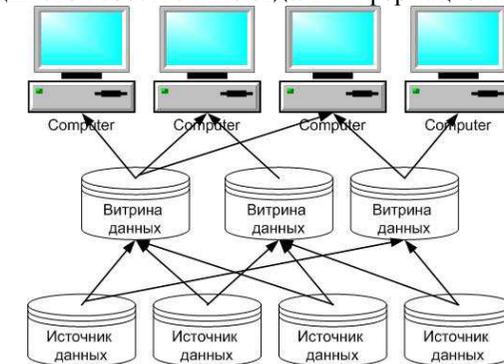


Рис.6.2-4. Независимые витрины данных

Преимущества: витрины данных можно внедрять достаточно быстро; витрины проектируются для ответов на конкретный ряд вопросов; данные в витрине оптимизированы для использования определенными группами пользователей, что облегчает процедуры их наполнения, а также способствует повышению производительности.

Недостатки: данные хранятся многократно в различных витринах данных (это приводит к дублированию данных и, как следствие, к увеличению расходов на

¹⁰ Marakas G. M. Decision support systems in the twenty-first century. Upper Saddle River, N.J.: Prentice Hall, 1999.

хранение и потенциальным проблемам, связанным с необходимостью поддержания непротиворечивости данных); потенциально очень сложный процесс наполнения витрин данных при большом количестве источников данных; данные не консолидируются на уровне предприятия, таким образом, отсутствует единая картина бизнеса.

СППР на основе двухуровневого хранилища данных. Двухуровневое хранилище данных (рис.6.2-5) строится централизованно для предоставления информации в рамках компании. Для поддержки такой архитектуры необходима выделенная команда профессионалов в области хранилищ данных. Это означает, что вся организация должна согласовать все определения и процессы преобразования данных.

Преимущества: данные хранятся в единственном экземпляре; минимальные затраты на хранение данных; отсутствуют проблемы, связанные с синхронизацией нескольких копий данных; данные консолидируются на уровне предприятия, что позволяет иметь единую картину бизнеса.

Недостатки: данные не структурируются для поддержки потребностей отдельных пользователей или групп пользователей; возможны проблемы с производительностью системы; возможны трудности с разграничением прав пользователей на доступ к данным.

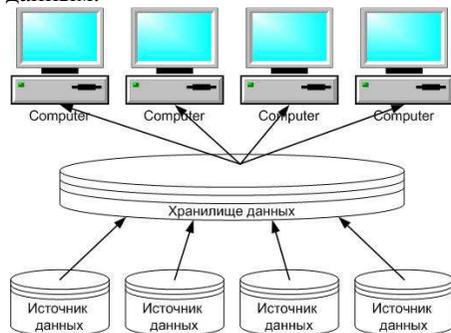


Рис.6.2-5. Двухуровневое хранилище данных

СППР на основе трёхуровневого хранилища данных. Хранилище данных представляет собой единый централизованный источник корпоративной информации (рис.6.2-6). Витрины данных представляют подмножества данных из хранилища, организованные для решения задач отдельных подразделений компании. Конечные пользователи имеют возможность доступа к детальным данным хранилища, в случае если данные в витрине недостаточно, а также для получения более полной картины состояния бизнеса.

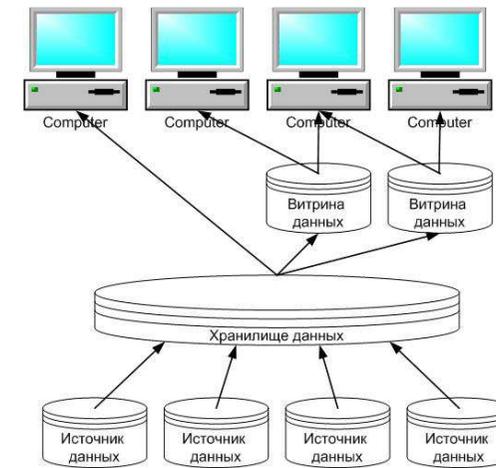


Рис. 6.2-6. Трёхуровневое хранилище данных

Преимущества: создание и наполнение витрин данных упрощено, поскольку наполнение происходит из единого стандартизованного надежного источника очищенных нормализованных данных; витрины данных синхронизированы и совместимы с корпоративным представлением (имеется корпоративная модель данных, существует возможность сравнительно лёгкого расширения хранилища и добавления новых витрин данных); гарантированная производительность.

Недостатки: существует избыточность данных, ведущая к росту требований на хранение данных; требуется согласованность с принятой архитектурой многих областей с потенциально различными требованиями (например, скорость внедрения иногда конкурирует с требованиями следовать архитектурному подходу).

Выше приведены основные варианты архитектур СППР. Выбор конкретного варианта зависит от условий, в которые поставлена проектная группа. На выбор архитектуры влияют ответы на такие вопросы как: нужен ли быстрый возврат от инвестиций; является ли проектная группа профессиональной; существует ли формализованная методология.

6.2. Информационные технологии создания и распределения знания

Данные и знания в ИС.

Данные - это отдельные факты, характеризующие объекты, процессы и явления предметной области, а также их свойства. При обработке на ЭВМ данные трансформируются, условно проходя следующие этапы:

↓ D1 – данные как результат измерений и наблюдений;

↓ D2 – данные на материальных носителях информации (таблицы, протоколы, справочники);