

Лекция 5,6

Этапы разработки ИС



Жизненный цикл информационных систем

Жизненный цикл - это период создания и использования ИС, начиная с момента возникновения потребности в ИС и заканчивая моментом полного ее выхода из эксплуатации.

Стадии жизненного цикла информационной системы:

1. Предпроектное обследование:

сбор материалов для проектирования, при этом выделяют формулирование требований, с изучения объекта автоматизации, даются предварительные выводы предпроектного варианта ИС;

анализ материалов и разработка документации, обязательно дается технико экономическое обоснование с техническим заданием на проектирование ИС.

2. Проектирование:

2.1 предварительное проектирование;

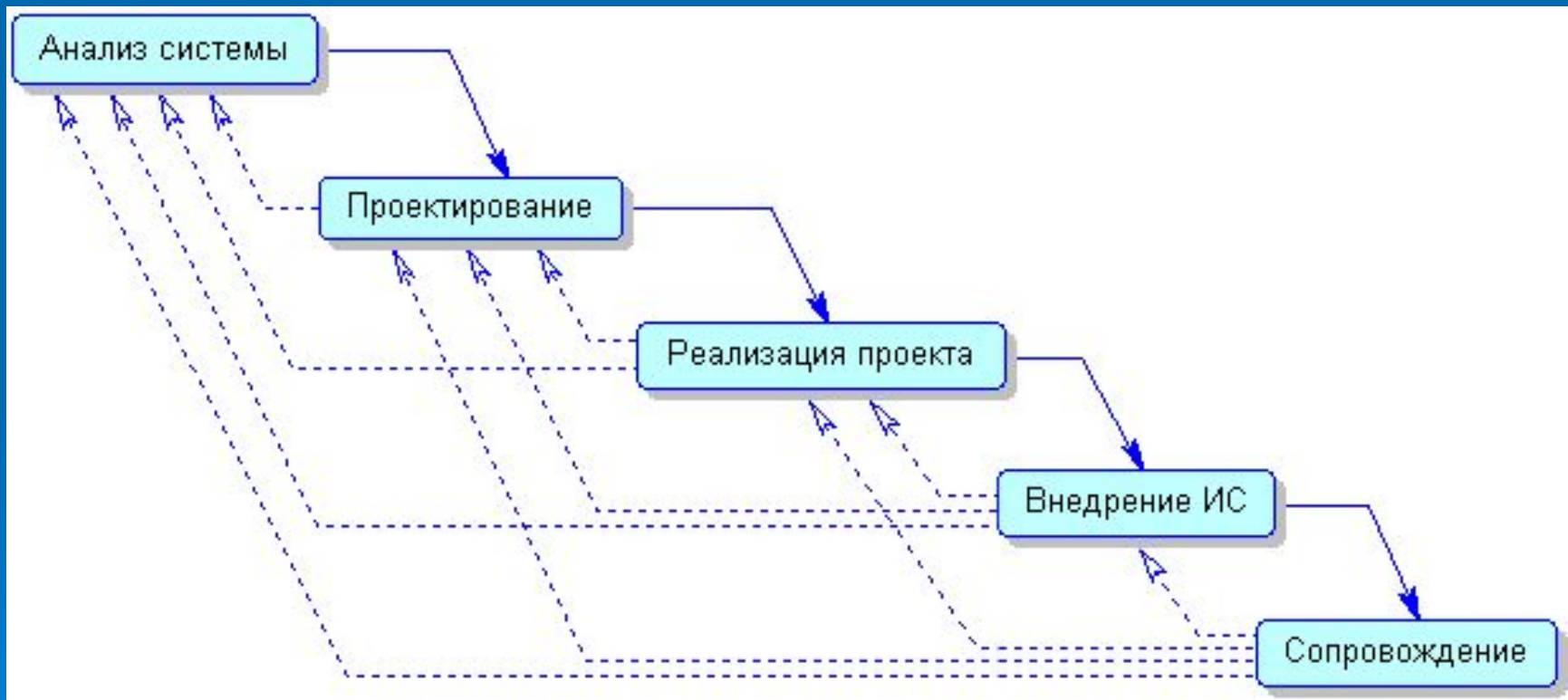
выбор проектных решений по аспектам разработки ИС;

описание реальных компонент ИС;

оформление и утверждение технического проекта (ТП).

Жизненный цикл информационных систем

(поэтапная модель с промежуточным контролем)



2.2 детальное проектирование:

выбор или разработка математических методов или алгоритмов программ;

корректировка структур БД;

создание документации на доставку и установку программных продуктов;

выбор комплекса технических средств с документацией на ее установку.

2.3 разработка техно-рабочего проекта ИС (ТРП).

2.4 разработка методологии реализации функций управления с помощью ИС и описанием регламента действий аппарата управления.

3. Разработка ИС:

получение и установка технических и программных средств;

тестирование и доводка программного комплекса;

разработка инструкций по эксплуатации программно-технических средств.

4. Ввод ИС в эксплуатацию:

ввод технических средств;

ввод программных средств;

обучение и сертификация персонала;

опытная эксплуатация;

сдача и подписание актов приемки-сдачи работ.

5. Эксплуатация ИС:

повседневная эксплуатация;

общее сопровождение всего проекта.

Основные способы построения ИС:

разработка системы "под себя";

использование прототипов - вместо полной системы создается прототип, отвечающий основным потребностям пользователей: - определение основных запросов;

- создание рабочего прототипа;

- использование рабочего прототипа;

- пересмотр и улучшение прототипа;

- работа с окончательной версией прототипа;

использование готовых решений - рекомендуется в максимальной степени использовать стандартные технологии и автоматизации бизнеса;

использование услуг сторонней организации для передачи функций управления ИС - организация использует специализированную фирму, которая выполняет управляющие функции по функционированию и развитию ИС компании.

Плюсы:

гарантийное качество обслуживания;

экономия денежных средств;

человеческие ресурсы.

Минусы:

не дешево;

утечка информации;

зависимость;

потеря контроля за ИТ.

Процессы в информационной системе:

ввод информации из внешних и внутренних источников;

обработка входящей информации;

хранение информации для последующего ее использования;

вывод информации в удобном для пользователя виде;

обратная связь, т.е. представление информации, переработанной в данной организации, для корректировки входящей информации.

С учетом сферы применения выделяют: технические ИС, экономические ИС, ИС в гуманитарных областях и т.д.



Система поддержки принятия решений или **СППР** (Decision Support Systems, DSS) — это компьютерная система, которая путем сбора и анализа большого количества информации может влиять на процесс принятия решений организационного плана в управлении и производстве.

Интерактивные системы позволяют руководителям получить полезную информацию из первоисточников, проанализировать ее, а также выявить существующие информационные модели для решения определенных задач.

Классификация СППР

По взаимодействию с *пользователем* :

пассивные помогают в процессе принятия решений, но не могут выдвинуть конкретного предложения;

активные непосредственно участвуют в разработке правильного решения;

кооперативные предполагают взаимодействие СППР с пользователем. Выдвинутое системой предложение пользователь может доработать, усовершенствовать, а затем отправить обратно в систему для проверки. После этого предложение вновь представляется пользователю, и так до тех пор, пока он не одобрит решение.

Классификация СППР

По *способу поддержки* различают:

модельно-ориентированные, используют в работе доступ к статистическим, финансовым или иным моделям;

основанные на коммуникациях, поддерживают работу двух и более пользователей, занимающихся общей задачей;

ориентированные на данные, имеют доступ к временным рядам организации. Они используют в работе не только внутренние, но и внешние данные;

ориентированные на документы, манипулируют неструктурированной информацией, заключенной в различных электронных форматах;

ориентированные на знания, предоставляют специализированные решения проблем, основанные на фактах.

Структура СППР

Выделяют четыре основных компонента:

- 1) Информационные хранилища данных;
- 2) средства и методы извлечения, обработки и загрузки данных (ETL);
- 3) многомерная база данных и средства анализа OLAP;
- 4) средства Data Mining.

Информационное хранилище данных

Хранилище данных — предметно-ориентированная информационная БД, предназначенная для подготовки отчётов и анализа с целью поддержки принятия решений в организации.

Строится на базе СУБД и СППР.

Данные доступны только для чтения.

Обновление данных в хранилище:

Полное обновление данных в хранилище.


~~Старые данные – новые данные~~

Инкрементальное обновление — обновляются только те данные, которые изменились.

Средства и методы извлечения, обработки и загрузки данных (ETL)

ETL (от англ. *Extract, Transform, Load* — извлечение, преобразование, загрузка) — процесс управления хранилищами данных, а также наименование класса утилит автоматизации этого процесса.

Процесс ETL включает в себя:

- извлечение данных из внешних источников;
- их преобразование в соответствии с требованиями информационной модели;
-  загрузку преобразованных данных в целевую систему (например, хранилище данных).

Проблемы разработки (ETL)

- 1) В качестве внешних источников информации могут выступать различные информационные системы, форматы хранения данных которых и процедуры их извлечения могут существенно различаться.
- 2) ETL-процесс не сводится исключительно к техническому преобразованию форматов — данные из разнородных источников должны быть унифицированы и с точки зрения бизнес-правил, единства применяемых систем кодирования информации, классификаторов и справочников.
- 3) Процесс должен учитывать и особенности бизнес-процессов компании, в том числе, функционирования выступающих в качестве источников данных отдельных информационных систем, периодичности обновления данных в них и т. д.

Примеры готовых решений

[IBM InfoSphere DataStage](#)

[Talend Open Studio](#)

[Pentaho Data Integration](#)

[Informatica](#)



IBM InfoSphere DataStage

DataStage поддерживает:

Текстовые файлы

Сложно структурированные XML-данные

Бизнес-системы уровня предприятия от компаний

SAP Бизнес-системы уровня предприятия от компаний SAP, Siebel Бизнес-системы уровня предприятия от компаний SAP, Siebel, Oracle

Практически любая СУБД Практически любая СУБД, включая системы управления секционированными базами данных (содержимое которых физически разделено между различными компьютерами), такие, как Oracle Практически любая СУБД, включая системы управления секционированными базами данных (содержимое которых физически разделено между различными компьютерами), такие, как Oracle, IBM DB2 Практически любая СУБД, включая системы управления секционированными базами данных (содержимое которых физически разделено между

Многомерная база данных и средства анализа OLAP

Системы аналитической обработки в реальном времени (Online Analytical Processing) - технология обработки информации, включающая составление и динамическую публикацию отчётов и документов.

Используется аналитиками для быстрой обработки сложных запросов к базе данных. Служит для подготовки бизнес-отчётов по продажам, маркетингу, в целях управления.

Многомерная база данных и средства анализа OLAP

OLAP (от англ. OnLine Analytical Processing — оперативная аналитическая обработка данных, также: аналитическая обработка данных в реальном времени, интерактивная аналитическая обработка данных) — подход к аналитической обработке данных, базирующийся на их многомерном (от англ. OnLine Analytical Processing — оперативная аналитическая обработка данных, также: аналитическая обработка данных в реальном времени, интерактивная аналитическая обработка данных) — подход к аналитической обработке данных, базирующийся на их многомерном иерархическом представлении, являющийся частью более широкой области информационных технологий — бизнес-аналитики (BI).

OLAP-системы представляют средства гибкого

Многомерная база данных и средства анализа

OLAP

Основное требование к OLAP-системам — скорость, позволяющая использовать их в процессе интерактивной работы аналитика с информацией.

OLAP-системы противопоставляются,

во-первых, традиционным РСУБД во-первых, традиционным РСУБД, выборки из которых с типовыми для аналитиков запросами, использующими группировку и агрегирование данных, обычно затратные по времени ожидания и загрузке РСУБД, поэтому интерактивная работа с ними при сколь угодно значительных объемах данных сложна.

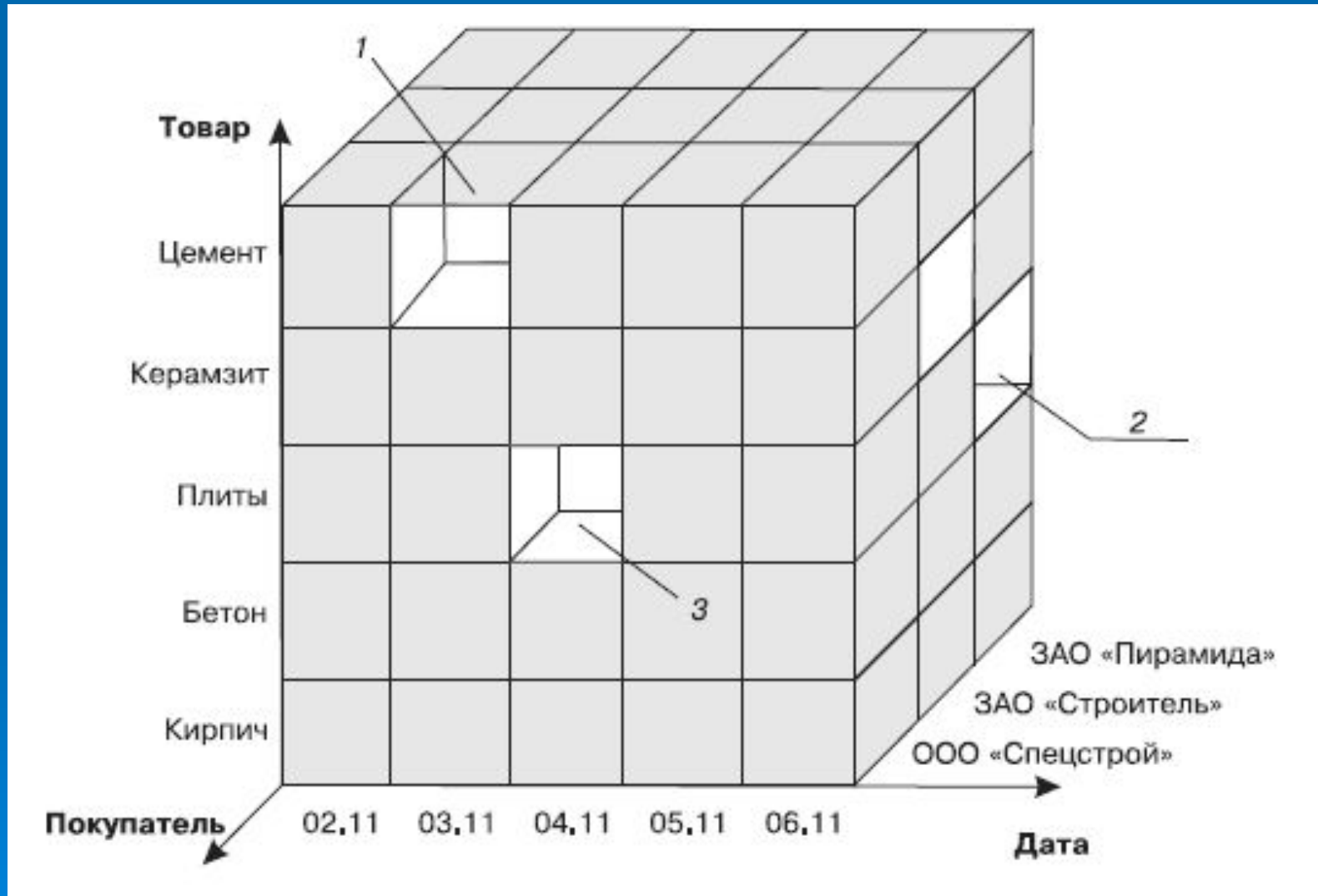
во-вторых, OLAP-системы противопоставляются и обычному плоскофайловому представлению данных, например, в виде часто используемых традиционных электронных таблиц, представление многомерных данных в которых сложно и не интуитивно, а операции по смене среза — точки зрения на данные — также требуют временных затрат и усложняют интерактивную работу с данными.

Многомерная база данных и средства анализа OLAP

<u>модель</u>	<u>месяц</u>	<u>объем</u>
BMW	Январь	12
BMW	Февраль	14
BMW	Март	16
Opel	Январь	2
Opel	Февраль	3
Lexus	январь	1

<u>модель</u>	<u>январь</u>	<u>февраль</u>	<u>март</u>
BMW	12	14	16
Opel	2	3	No
Lexus	1	No	No

Многомерная база данных и средства анализа OLAP



Средства Data Mining

Data mining (интеллектуальный анализ данных) — процесс выявления скрытых, полезных фактов и взаимосвязей в крупных массивах данных. Дословно переводится как «извлечение данных».

Data mining решает задачи:

Классификация - структуризация объектов по заданным классам

Ассоциация - выявление ассоциативных цепей.
впервые такой метод был применен для анализа рыночной корзины типичного потребителя.

Кластеризация - группировка событий и наблюдений в кластеры. В основу берутся свойства описывающие суть самих группируемых событий

Средства Data Mining

Data mining решает задачи:

Прогнозирование - предсказание на основе имеющихся данных возможного развития событий как прогрессивное, так и регрессивное.

Анализ изменений - выявление типичных ситуаций, шаблонов. Сюда входит и установление закономерностей между различными временными событиями, равно как и обнаружение зависимостей и причинно-следственных связей.

Готовые решения:

SAS

Solutions-Center

BaseGroup

Ridel

DataMiner

Средства Data Mining

