# Интегрированные ИС

Интегрированный подход подразумевает, что данные рассматриваются как информационные ресурсы для разноаспектного и многократного использования.

## Банк данных

**Банк данных (БнД)** – это система специальным образом организованных данных (баз данных), программных, технических, языковых, организационно-методических средств, предназначенных для обеспечения централизованного накопления и коллективного многоцелевого использования данных.

Это вид ИС. Главная отличительная черта БнД: централизованное управление данными.

1. БнД пользуется большое число различных пользователей.
2. В БнД хранится большой объем многоаспектной информации.

**БнД = n\*БД + СУБД [+ Словарь данных]
[+ ОС] [+ Аппаратное обеспечение] [+ Персонал]**

**Словарь данных** – специальная система для хранения *метаданных*:

* об объектах (таблицах), их свойствах и отношениях;
* о данных (полях, атрибутах): наименование, структура, связи с другими данными, возможные значения и формат представления, об источниках их возникновения;
* о кодах защиты, разграничениях доступа к данным со стороны пользователей.

## Информационное хранилище

**Информационное хранилище (склад данных)** – это автоматизированная информационно-технологическая система, которая собирает данные из внутренних баз и внешних источников, формирует, хранит и эксплуатирует информацию как единую.

Основной целью является бизнес-анализ или информационная поддержка принятия управленческих решений.

Это не просто объединение разных БД, это отдельная ИС, со своими механизмами управления данными. В хранилище располагаются:

* **метаданные**;
* **фактические данные** (архивы) за различные моменты времени;
* **суммарные (аналитические) данные** - итоги расчетов по фактическим данным.

**Предметная ориентированность** – ориентированность на предметы (определенные группы данных), а не на их конкретные приложения.

**Интегрированность** – возможность объединения данных из разных источников, имеющих разный формат и несогласованных.

**Привязка ко времени** – для всех данных указан момент или промежуток времени, в который они корректны.

Данные в хранилище **не изменяются**, они лишь регулярно пополняются из оперативных баз данных.

Информационное хранилище может быть как *централизованным* (БнД), так и *распределенным*. Зачастую используют **вторичные устройства** хранения на сервере, например, оптические диски.

Пример. JukeBox – библиотека-автомат с автоматической сменой оптических дисков. Ненужная и редко используемая информация автоматически выгружается на вторичный носитель (**миграция данных**).

**Витрины данных** – небольшие хранилища с упрощенной архитектурой, предназначенные для хранения части данных с целью снятия нагрузки с основного информационного хранилища.

Примеры систем для хранилищ данных: IBM Red Brick Warehouse,SAP NetWeaver Business Warehouse.

## Базы знаний и экспертные системы

**База знаний (БЗ)** предназначена для хранения информации более высокого уровня – знаний.

Данные

Информация

Знания

БЗ, в отличие от БД, содержит не только данные (факты), но и **правила вывода** новых данных.

БЗ – важный компонент интеллектуальной системы, из которых наиболее распространены экспертные системы. В них хранятся знания человека-эксперта в определенной предметной области.

**Экспертная система** (ЭС)– система искусственного интеллекта, включающая БЗ с набором правил и механизмом вывода, позволяющая на основании предоставляемых пользователем фактов распознать ситуацию, сделать выводы, и сформулировать решение или дать рекомендацию для выбора действия.

Примеры организации ЭС:

**Семантические сети** – множество понятий и связей между ними.

**Продукционные системы** – делают выводы по заранее заданным логическим правилам (ЕСЛИ-ТО).

**Нейронные сети** – обучающиеся системы, имитирующие работу нервной системы человека. Чаще всего обучаются на примерах.

**Фреймы** – минимально возможное описание представления человека о реальных объектах. Выделяют основные свойства объекта, которые отличают его от остальных.

**Prolog** – язык программирования, основанный на правилах логического вывода (предикатах). Содержит факты, правила вывода и запросы (к пользователю).

Вывод :- Условие.

%**Факты**

Отец(Анастасия, Александр).

Мать(Анастасия, Ирина).

Отец(Ксения, Александр).

Мать(Ксения, Ирина).

Отец(Александр, Евгений).

Мать(Александр, Раиса).

Отец(Ирина, Борис).

Мать(Ирина, Римма).

%**Правила**

Родитель(A,B) :- Отец(A,B).

Родитель(A,B) :- Мать(A,B).

Дед(A,B) :- Родитель(А,X), Отец(X,B).

Бабушка(A,B) :- Родитель(А,X), Мать(X,B).

Система может сделать выводы о том, что Евгений - дед Анастасии, Римма - бабушка Ксении и т.д.

## Корпоративные ИС

**Корпоративная информационная система** (КИС) – АИС управления крупными, территориально рассредоточенными предприятиями, имеющими несколько уровней управления, построенная посредством интегрированных информационных технологий и систем.

Назначение КИС – обеспечить решение внутренних задач управления:

* бухгалтерский учет;
* финансовое планирование и финансовый анализ;
* управление договорными отношениями;
* расчеты с поставщиками и покупателями;
* анализ рынка;
* управление себестоимостью;
* управление кадрами
* и др.

Обычно КИС полностью или частично закрыты для внешних пользователей и обслуживаются корпоративной сетью.

КИС предполагает автоматизацию основных **бизнес-процессов** предприятия, поддержку принятия управленческих решений, т.е. создание **бизнес-модели** предприятия.

КИС могут включать:

* СУБД;
* Workflow – управление деловыми процессами;
* GroupWare – система групповой работы в пределах каждой рабочей группы/отдела;
* EDMS – система управления электронными документами и ведения электронного архива;
* OCR – система массового ввода печатной информации в компьютер;
* системы информационной безопасности;
* прикладные программные средства.

## Системы интеллектуального анализа данных

Технологии интеллектуального анализа данных обеспечивают формирование аналитических данных посредством очищения данных из локальных баз организации, применения статистических методов и других сложных алгоритмов.

* оперативный анализ данных посредством **OLAP-систем**;
* поиск и интеллектуальный выбор данных **Data Mining**;
* деловые интеллектуальные технологии BIS;
* интеллектуальный анализ текстовой информации.

**OLAP-куб** – многомерная агрегированная выборка данных из плоских (двумерных) таблиц. Ускоряет получение запросов по разным (большим) таблицам.

Двумерный запрос с итогами (агрегированием):

|  |  |  |
| --- | --- | --- |
| Страна | Поставщик | Суммарная стоимость заказов |
|  |  |  |

Добавив годы, получим трехмерный куб:



**Data Mining** - поиск новых связей и зависимостей в больших объемах «сырых» данных. Например:

* классификация покупателей по группам,
* прогнозирование на основе тренда продаж,
* ранжирование товаров по популярности,
* поиск аномальных значений (выявление ошибок, мошенничества),
* и др.