



Информационные системы

Тема:
«Фактографические системы»

Фактографические информационные СИСТЕМЫ

накапливают и хранят данные в виде множества экземпляров одного или нескольких типов структурных элементов
(информационных объектов).

Каждый из таких экземпляров структурных элементов или некоторая их совокупность отражают сведения по какому-либо факту, событию и т. д., отделенному (вычлененному) от всех прочих сведений и фактов.

Структура каждого типа информационного объекта состоит из конечного набора реквизитов, отражающих основные аспекты и характеристики сведений для объектов данной **предметной** области.

Предметная область информационной системы

набор объектов, представляющих интерес для актуальных или предполагаемых пользователей, когда реальный мир отображается совокупностью конкретных и абстрактных понятий, между которыми фиксируются определенные связи.

Комплектование информационной базы в фактографических информационных системах

включает, как правило, обязательный процесс структуризации входной информации из документального источника.

Структуризация при этом осуществляется через определение (выделение, вычленение) экземпляров информационных объектов определенного типа, информация о которых имеется в документе, и заполнение их реквизитов

Основные признаки фактографической информационной системы

- *простая структура данных*
- *сложная система взаимосвязей между агрегатами данных*

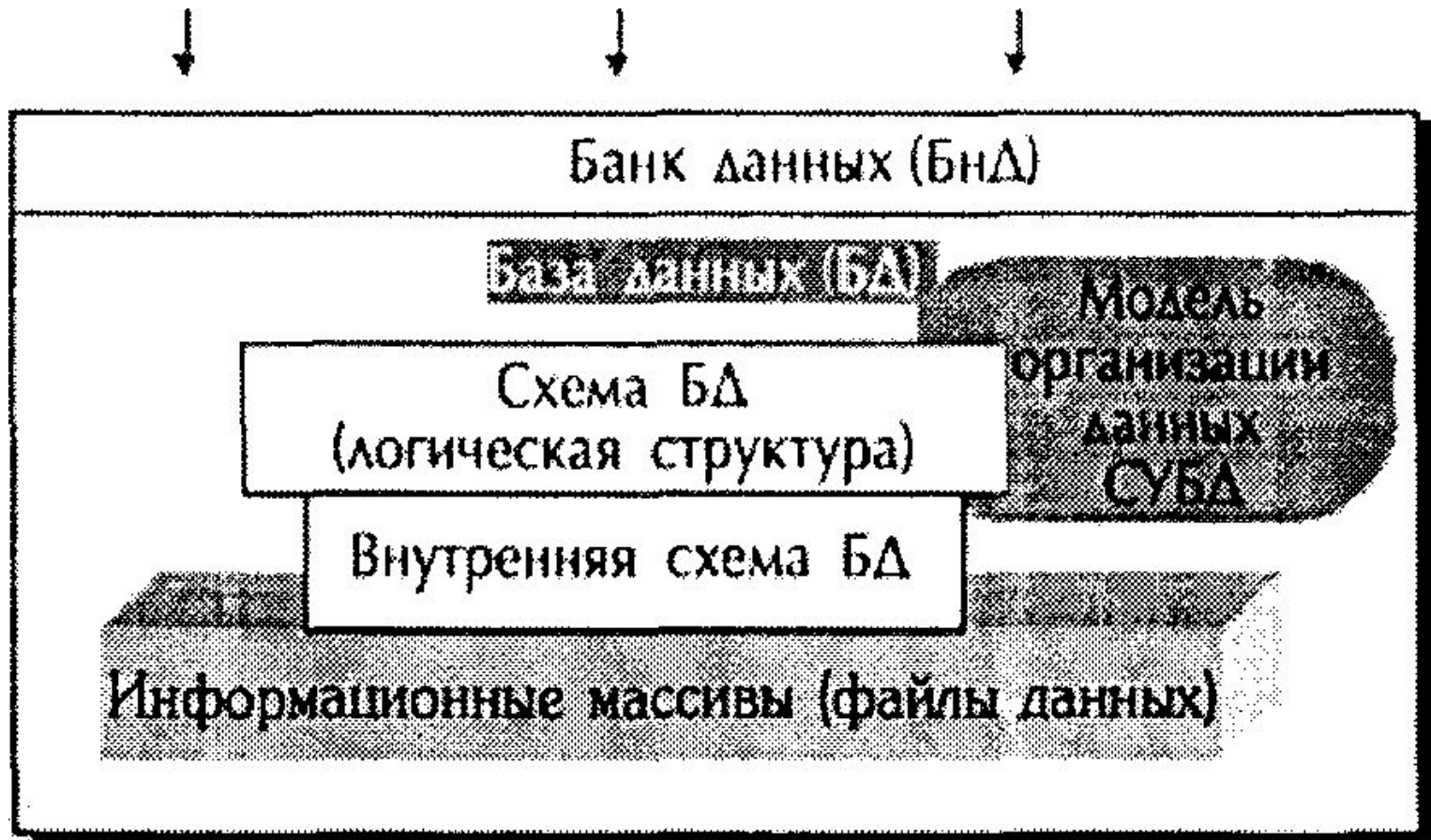
Уровни представления информации фактографической ИС

- Локальные представления пользователей о предметной области (ПрО)
- Информационно-логическая модель ПрО
- Концептуальная модель использования ИС
- Информационные потребности абонентов
- Формализованное представление об объектах и отношениях ПрО
- Описание структуры базы данных ИС

Банк Данных информационной системы:

- *База данных*
 - *Модель организации данных СУБД*
 - *Информационные массивы*

Банк Данных фактографической информационной системы



Формализация представлений о предметной области

осуществляется в рамках модели «**объекты-связи**» (так называемая *ER-модель* — от англ. *Entity Relationship*).

При этом под **информационным объектом** в общем плане понимается некоторая сущность фрагмента действительности.

В предметной области выделяются различные **типы объектов**, представляемые в информационной системе в каждый момент времени конечным набором **экземпляров** данного типа.

Каждый тип объекта включает (идентифицируется) присущий ему набор **атрибутов** (свойств, характерных признаков, параметров).

Формализация представлений о предметной области

Атрибут представляет логически неделимый элемент структуры информации, характеризующийся множеством атомарных значений.

Экземпляр объекта образуется совокупностью конкретных значений атрибутов данного типа объекта.

Один или некоторая группа атрибутов объекта данного типа могут исполнять роль **ключевого атрибута**, по которому идентифицируются (различаются) конкретные экземпляры объектов.

Различные типы объектов и различные экземпляры одного типа объекта могут быть охвачены определенными **отношениями**, которые в рамках ER-модели выражаются **связями**.

Структурные и одноуровневые связи по признаку множественности могут быть трех типов:

- *«один-к-одному»*
 - *«один-ко-многим»*
 - *«многие-ко-многим»*

Пример диаграммы «сущность—связь»



Логическая структура данных

представляется *схемой базы данных*, представляющей описание средствами конкретной СУБД *мифологической схемы предметной области* (информационные объекты, реквизиты, связи).

Совокупность средств и способов реализации схемы базы данных в конкретной СУБД составляет ***модель организации данных***.

Схема базы данных содержит также ***ограничения целостности данных***.

Ограничения целостности представляют собой набор установок и правил по типам, диапазонам, соотношениям значений атрибутов объектов, характеристик и особенностей связей между объектами.

Типичная структура простейшей табличной БД



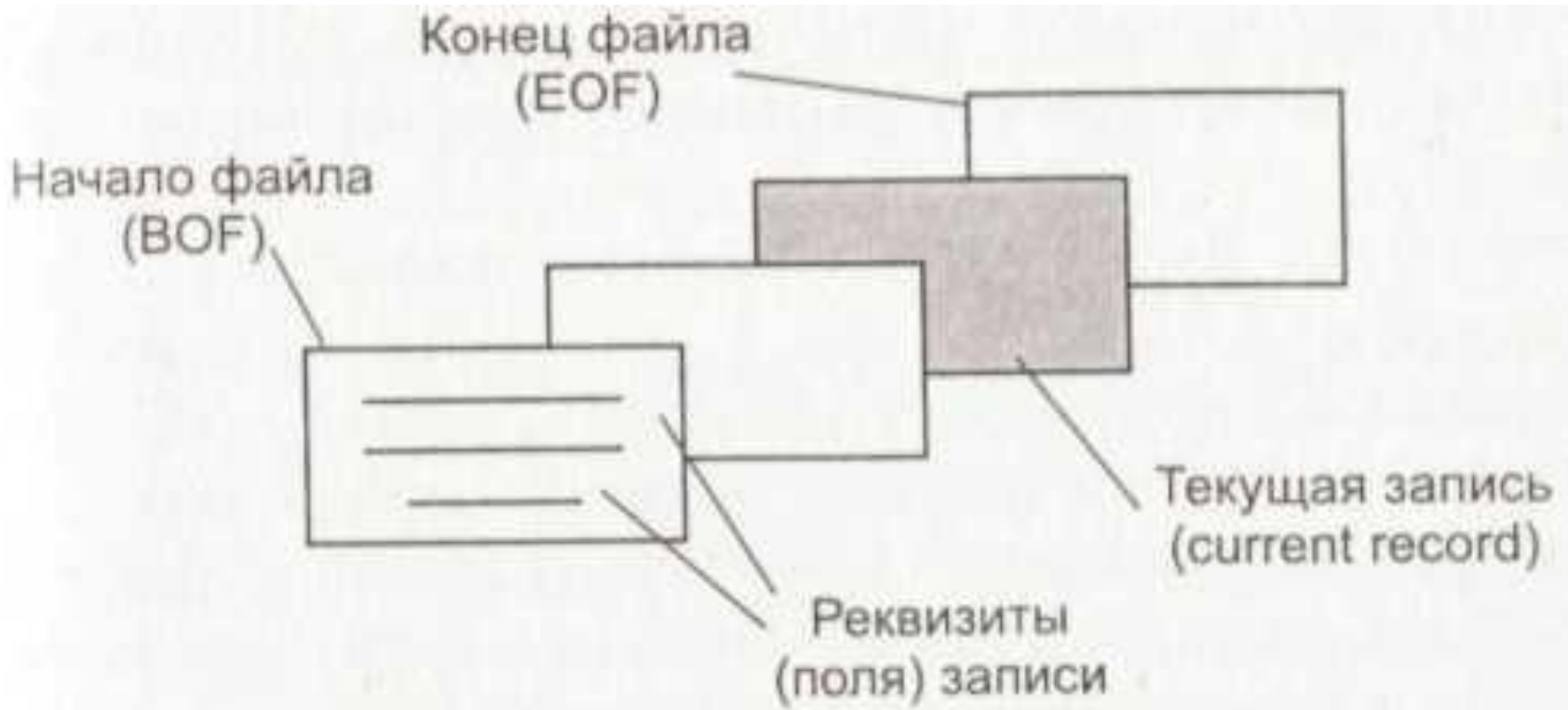
а) - общая структура БД; б) - структура PERSON.DBF;
 в) - структура AUTO.DBF; г) - структура FINANCES.DBF

Внутренняя схема базы данных

определяет структуру организации и особенности хранения информационных массивов, в которых и находятся собственно сами данные.

Более конкретные особенности представления и организации данных определяются конкретным типом и особенностями СУБД, используемой для создания фактографической информационной системы.

Основные понятия, связанные с отдельным файлом табличной БД



В общем плане можно выделить следующие функции, реализуемые СУБД:

- **организация и поддержание логической структуры данных (схемы базы данных);**
- **организация и поддержание физической структуры данных во внешней памяти;**
- **организация доступа к данным и их обработка в оперативной и внешней памяти.**

Организация и поддержание логической структуры данных

обеспечивается средствами *модели организации данных*.

Модель данных определяется способом организации данных, ограничениями целостности и множеством операций, допустимых над объектами организации данных.

Модель данных разделяют на три составляющие:

- *структурную,*
 - *целостную,*
 - *манипуляционную*

Организация и поддержание физической структуры данных во внешней памяти

Эта функция включает организацию и поддержание внутренней структуры файлов базы данных, иногда называемой **форматом файлов базы данных**, а также создание и поддержание специальных структур (индексы, страницы) для эффективного и упорядоченного доступа к данным.

Организация доступа к данным и их обработка в оперативной и внешней памяти

осуществляется через реализацию процессов, получивших название транзакций.

Транзакцией называют последовательную совокупность операций, имеющую отдельное смысловое значение по отношению к текущему состоянию базы данных.

Транзакции принято разделять на две разновидности:

- изменяющие состояние базы данных после завершения транзакции
- изменяющие состояние БД лишь временно, с восстановлением исходного состояния данных после завершения транзакции.

Совокупность функций СУБД по организации и управлению транзакциями называют **монитором транзакции**.

Ввод, обработка и вывод данных в фактографических ИС

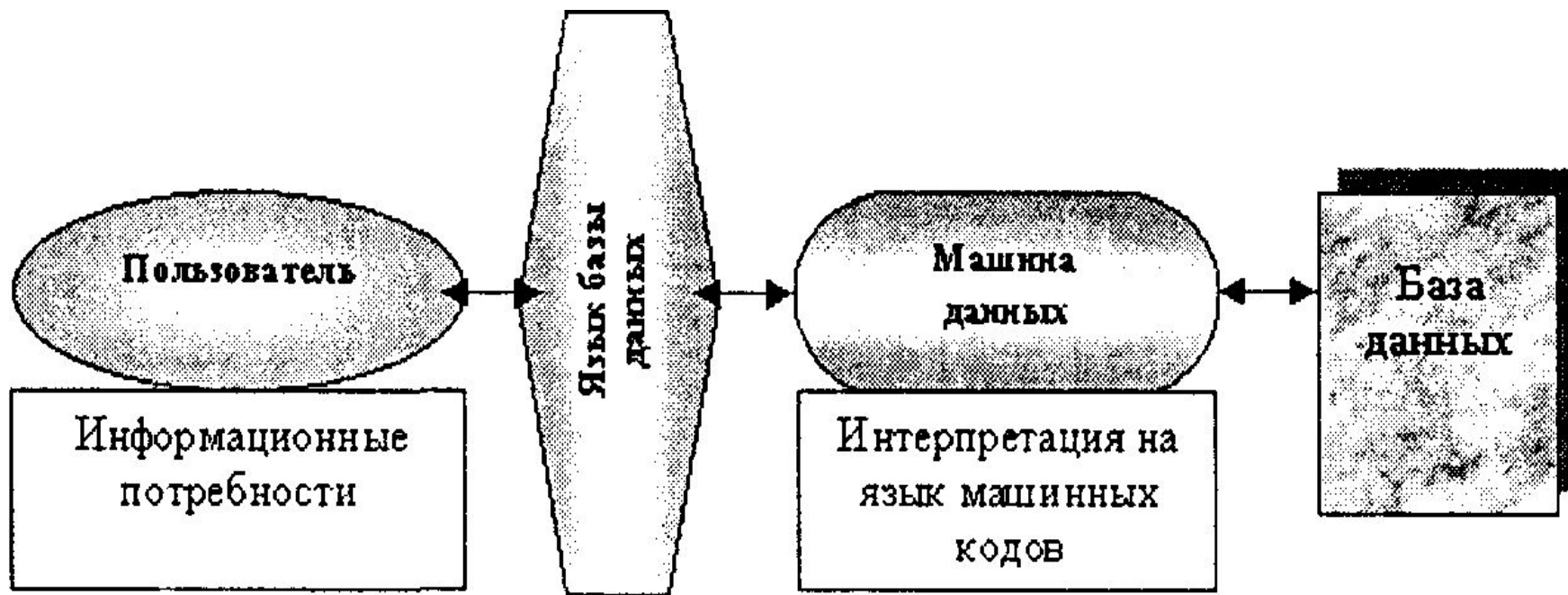
Если исключить организационно-технологические аспекты сбора, комплектования и выдачи информации, технология работы пользователей с базами данных ИС включает ввод (загрузку), обработку и вывод данных.

Предоставление пользователю средств реализации (функций ввода, обработки и выдачи данных является одной из основных функций интерфейса информационных систем.

Языки баз данных

- **Язык описания данных (ЯОД)** – средство внутрисистемного определения данных, представляющего обобщение внешних взглядов;
- **Язык манипулирования данными (ЯМД)** – включает в себя средства запросов и поддержания БД;
- **Язык структурированных запросов (SQL)** – основной интерфейс пользователя и автоматизированной базы данных для запоминания и поиска информации в БД для ряда СУБД

Схема взаимодействия пользователя с базой данных через язык баз данных

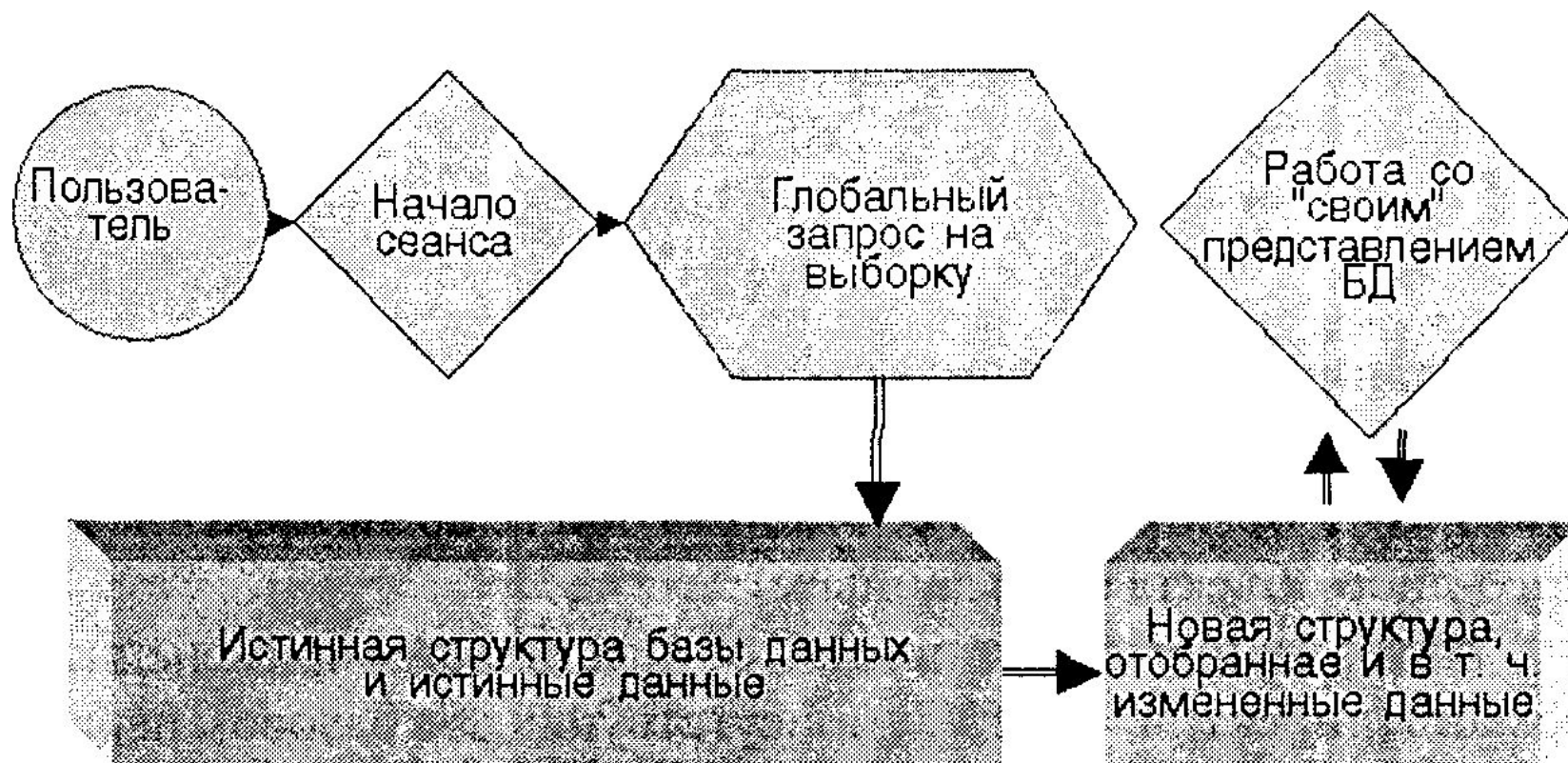


Распределенных информационных системы

В основе *распределенных ИС* лежат две основные идеи:

- много организационно и физически распределенных пользователей, одновременно работающих с общими данными — общей базой данных (пользователи с разными именами, в том числе располагающимися на различных вычислительных установках, с различными полномочиями и задачами);
- логически и физически распределенные данные, составляющие и образующие тем не менее единое взаимосогласованное целое — общую базу данных (отдельные таблицы, записи и даже поля могут располагаться на различных вычислительных установках или входить в различные локальные базы данных).

Данные должны быть такими и логически организованы таким образом, чтобы можно было решать требуемые информационные задачи и выполнять установленные функции.



Фактографические ИС можно классифицировать по признаку представления предметной области:

- универсальные

- специализированные

Универсальные фактографические информационные системы

могут работать с любой предметной областью, с любой “начинкой”, если данные организованы в соответствии с рассмотренными нами принципами и формализованы согласно дополнительным требованиям конкретной ИС.

Кроме того, они могут сортировать тексты, составлять словари, выполнять ряд других операций по формальной обработке данных.

Область применения — от личных (“персональных”) баз данных всевозможного назначения до профессиональных систем (гуманитарные БД, медицина, небольшие офисы).

Специализированные фактографические информационные системы

Отличительные черты:

- специальные алгоритмы ввода, поиска, обработки и выдачи информации, которые учитывают и специфику данных, и характер запросов;
- учет динамической природы профессиональных систем, которая выражается в том, что база данных непрерывно обновляется, и не одним, а десятками и даже тысячами пользователей одновременно.

Специализированные фактографические информационные системы широко применяются в материальном производстве, банковском деле, криминалистике, научных исследованиях.

Создание и эксплуатация фактографических информационных систем

требуется либо изначально структурированных данных, либо предварительной структуризации данных. При этом зачастую структуризация данных требует больших накладных, в том числе и организационных расходов, что, в конечном счете, приводит к материальным издержкам информатизации.

Входные информационные потоки могут быть в представлении неструктурированными данными в виде служебных документов и иных текстовых источников. Извлечение из текстов данных по формализованным позициям для ввода в фактографические системы может приводить к ошибкам и потере части информации, которая в исходных источниках имеется, но в силу отсутствия в схеме базы данных адекватных элементов не может быть отражена в банке данных фактографических ИС.

Контрольные вопросы:

- Что подразумевается под предметной областью информационной системы?
- Дайте определение информационного объекта.
- Охарактеризуйте основные признаки фактографической информационной системы.
- Перечислите уровни представления информации в фактографической информационной системе.
- Что подразумевается под логической структурой данных?
- Дайте определение ключевого атрибута.
- Опишите особенности специализированных фактографических информационных систем.