

Д.Г. ПОДОБЕД, О.В. ПОДОБЕД

**ОСНОВЫ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**
(краткий словарь-справочник)

Учебное пособие

**Санкт-Петербург
2010**

**Министерство образования и науки
Российской Федерации**

**Государственное образовательное учреждение
высшего профессионального образования**

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ
РАСТИТЕЛЬНЫХ ПОЛИМЕРОВ**

Д.Г. ПОДОБЕД, О.В. ПОДОБЕД

**ОСНОВЫ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

(краткий словарь-справочник)

Учебное пособие

**Санкт-Петербург
2010**

ББК 22.18я7
П 444
УДК 681.3:65.5

Подобед Д.Г., Подобед О.В. Основы информационных технологий (краткий словарь-справочник): учебное пособие / ГОУВПО СПбГТУРП. - СПб., 2010.- 85 с.

В учебном пособии изложены основные определения базовых понятий по тематике «Основы информационных технологий», которые необходимы при изучении материалов и тестировании студентов по курсу «Информатика».

Предназначается для организации работы студентов, обучающихся по специальностям: «Экономика и управление на предприятии ЦБП», «Менеджмент организации».

Рецензенты:

доцент кафедры прикладной математики и информатики ГОУВПО СПбГТУРП Антонюк П.Е.;

доцент кафедры эконометрии ГОУВПО «Санкт-Петербургский государственный университет водных коммуникаций», канд. техн. наук Полянская Т.И.

Рекомендовано к изданию Редакционно-издательским советом университета в качестве учебного пособия.

© Подобед Д.Г., Подобед О.В., 2010

© ГОУВПО Санкт-Петербургский
государственный технологический
университет растительных
полимеров, 2010

О ПОЛЬЗОВАНИИ СЛОВАРЕМ

В тексте статей часто встречаются слова объект, тип объекта, понятие, термин и дефиниция. Определим их значение:

- объект [object] - любой материальный предмет или явление, с которым связана познавательная, информационная или любая другая практическая деятельность человека;
- тип объекта [object type] - обобщенное имя какого-либо множества объектов, характеризующееся определенным набором признаков (в том числе свойств и характеристик), например, "автомобиль", "гроза", "компьютерные игры" и т.д.
- понятие [concept, notion] - форма мысли, в которой отражаются общие и существенные признаки предмета или явления; целостная совокупность суждений о каком-либо объекте, отображающая его сущность и являющаяся результатом познания объекта;
- термин [term] - слово или словосочетание, обозначающее строго определенное понятие;
- дефиниция [definition, concept definition] - краткое научное определение какого-либо понятия; выделяет существенные черты определенного объекта таким образом, что он отличается от любых других объектов, и выражается законченным предложением, построенным по правилам логики.

Принципы построения словаря:

1. Словарные статьи вместо общепринятого алфавитного расположения представлены в тематических разделах в логической последовательности: от простого понятия к более сложному или от общего к частному.
2. Пояснительная часть словарных статей жестко не формализована, при необходимости расширена и содержит связанные с основным понятием термины, а также варианты толкования отдельных терминов, существующие в различных нормативных и других документальных источниках.
3. Варианты терминов - синонимы, обозначающие одно и то же понятие, отделены друг от друга запятой, точкой с запятой или косой чертой.
4. Составные части сложных терминов, которые используются как синонимы или носят факультативный характер (т.е. могут опускаться), заключены в круглые скобки.
5. Русскоязычные термины сопровождаются англоязычными эквивалентами, заключенными в квадратные скобки. В случае отсутствия устоявшегося англоязычного эквивалента термина, он не указывается.
6. Переводы англоязычных терминов, не имеющих устоявшихся русскоязычных эквивалентов, выделены кавычками и представлены с прописной буквы.
7. Англоязычные термины, для которых русскоязычный эквивалент отсутствует и вводить его признано нецелесообразным, помечаются знаком *

(звездочка).

8. Сокращенные термины представлены, как правило, через запятую вместе с их развернутой формой.

9. Сокращенные англоязычные термины и их развернутая форма выделяются круглыми скобками.

10. Составные термины могут содержать англо- и русскоязычные части, если каждая из них в отдельности и их комбинации являются широкоупотребительными, например Web-сервер, индекс iCOMP, USB-шина, накопитель CD-RW и т. п. В настоящее время в различных публикациях все чаще встречается русскоязычное написание - Веб. Однако, учитывая профессиональную специфику словаря-справочника, в нем преимущественно используется оригинальное международное англоязычное написание - Web.

11. В словарных статьях используются два близких по своему значению сокращения: ПК (рус.) - по отношению ко всем персональным компьютерам и РС (англ.) – преимущественно по отношению к IBM-совместимым ПК.

1. ДАННЫЕ, ИНФОРМАЦИЯ, ЗНАНИЯ, ЛОГИКА

ДАННЫЕ [data, information]

Сведения, факты, показатели, выраженные как в числовой, так и любой

другой форме.

Приведенный в скобках англоязычный эквивалент термина (**information**) показывает, что он весьма часто рассматривается как синоним термина **информация**. Так, словосочетание **data system** и производные от него могут соответствовать одному из совсем не однозначных, хотя и связанных между собой понятий - система данных и информационная система. Аналогичное явление прослеживается при русскоязычном использовании этого термина. Например, в ГОСТ 15971-84 - как "Информация, представленная на материальных носителях", а в ГОСТ 7.0-99 - "Информация, обработанная и представленная в формализованном виде для дальнейшей обработки".

С целью смыслового разделения понятий "информация" и "данные" Ассоциация стандартов Франции (АФНОР) дает следующее определение: "Данные - факты, понятия или инструкции, представленные в условной форме, удобной для пересылки, интерпретации и обработки человеком или автоматизированными средствами".

Согласно другому важному для понимания этого термина определению: "Данные - некоторый факт, то на чем основан вывод или любая **интеллектуальная система**". Компонентами данных являются цифры и символы естественного языка или их кодированное представление в виде двоичных **битов**.

ИНФОРМАЦИЯ [**information** - от лат. *information* - разъяснение, осведомление]

Данный термин и отражаемое им понятие являются сегодня одними из самых распространенных. Сказанное относится к их использованию как на бытовом, так и на профессиональном уровнях.

Существует множество различных определений этого понятия, например:

- "Информация - содержание какого-либо сообщения, сведения о чем-либо, рассматриваемые в аспекте их передачи в пространстве и времени ...";
- "Информация - сведения, подлежащие передаче";
- "Информация - это значение, вкладываемое человеком в данные на основании известных соглашений, используемых для их представления";
- "Информация - сведения, воспринимаемые человеком и/или специальными устройствами как отражение фактов материального или духовного мира в процессе коммуникации";
- "Информация - содержание, значение данных, которое видят в них люди. Обычно данные состоят из фактов, которые ставятся информацией в определенном контексте и понятны людям".

Следует упомянуть также классическое определение К. Шеннона, в соответствии с которым информация это то, что сокращает степень неопределенности (у Шеннона - **энтропии**) знаний у ее адресата о каком-либо объекте (в том числе явлении, передаваемом сигнале и т.п.).

Энтропия - мера неопределенности. Измеряется вероятностью наступления одного из N возможных событий. Если вероятность одного из

них становится равной 1, то неопределенность отсутствует (вероятность остальных равна 0). Максимальная неопределенность имеет место, если все события равновероятны. Другими словами, по Шеннону информация это то, что увеличивает степень знания ее адресатом интересующих его объектов и явлений окружающего мира. В указанном контексте количество информации можно даже рассчитать, в частности по увеличению вероятности успешного решения поставленной задачи.

Мы остановимся только на тех признаках понятия "информация", которые необходимы большинству пользователей, тем или иным образом преимущественно связанных с информационными сферами деятельности. Если на бытовом уровне смешение понятий "данные" и "информация" вполне допустимо, то для профессионалов это может привести и приводит к серьезным последствиям.

Чтобы стать информацией, данные должны:

1. Правильно отражать объекты описания, в противном случае мы будем иметь дело с **дезинформацией** (ее англоязычные эквиваленты: **false information, misleading information**). Сама по себе "правильность отражения действительности" в соответствии с теорией познания всегда носит условный характер, поскольку связана с уровнем развития знаний на данном этапе развития общества или отдельных социальных групп и индивидуумов. Так, состав и точность данных, которыми владеют или которые необходимы различным организациям и лицам об одном и том же объекте, будут существенно различаться в зависимости от образовательного, возрастного, социального состояний субъекта информирования, а также целей их использования (например, для "повышения общей эрудиции" или для решения конкретных научных, технических, производственных, коммерческих или других задач). В указанном плане данные, которые для одного субъекта будут представляться вполне точными, для другого — могут оказаться грубой дезинформацией;

2. Быть необходимы человеку для удовлетворения его **информационной потребности**;

3. Быть получены пользователем своевременно (не раньше и не позже) наступления в них **информационной потребности**. Всякое несвоевременное, а также повторное предоставление сведений, которыми адресат уже владеет, является информационным шумом;

4. Быть представлены в форме, удобной для восприятия тем, кому они предназначены.

Перечисленные признаки информации можно кратко обобщить следующим образом:

"Информация - это данные, удовлетворяющие информационную потребность того, кому они передаются, соответствующие действительности и материализованные в форме, удобной для использования, передачи, хранения и/или обработки (преобразования) человеком или автоматизированными средствами".

Информационный шум [information noise] - данные, не

соответствующие информационной потребности или не представляющие для субъекта информирования предмета новизны (другими словами, этими данными он уже владеет).

Понятие "информационный шум" может быть также распространено на данные, "не удобные для использования, передачи, хранения и/или обработки", поскольку и в этом случае они приводят к бесполезным, а, возможно, и вредным затратам материальных, временных и других ресурсов.

Дезинформация [misinformation] - передаваемые кому-либо и в любой форме данные, сведения, сообщения и т.п., неверно отражающие объекты описания реального мира или мыслительной деятельности человека.

Термин "информация" в сочетании с различными прилагательными широко употребляется в следующем случаях:

1. **Сведения или данные, полученные в процессе какого-либо вида деятельности**, отражают результаты этого вида деятельности или имеют отношение к ней и предназначены для справочно-информационного обслуживания и/или **информационного обеспечения** (см. далее) заинтересованных пользователей, например: научно-техническая информация [scientific-technical information], юридическая информация [juridical information], патентная информация [patent information] и т. п.

2. **Сведения или данные имеют определенное назначение**, например:

- **справочная информация [reference information]** - сведения или данные для выдачи справок о чем-либо;

- **сигнальная информация [alert information, current awareness information]** - информация, предназначенная для быстрого предварительного оповещения.

3. **Словарная статья определяет характер, принадлежность, форму или вид данных, используемых в информационном процессе**, например:

- **априорная информация [aprior information]** и **апостериорная информация [aposterior information]** - соответственно - данные, имевшиеся до проведения какого-либо опыта или другого действия, и сведения, полученные после его выполнения;

- **коммерческая информация [commercial information]** - данные, сведения и содержащие их документы, являющиеся объектом продажи их собственником;

- **личная информация [private information]** - сведения (данные) о гражданах и организациях, затрагивающие их интересы и запрещенные для распространения без их согласия;

- **библиографическая информация [bibliographic information]** - библиографические данные, описания и их перечни;

- **графическая информация [graphical (pictorial, image, pattern) information]** - сведения или данные, представленные в виде схем, эскизов, изображений, графиков, диаграмм, символов;

- **ретроспективная информация [retrospective information]** - сведения, содержащиеся в накопленных за два и более лет массивах данных или полученные в результате поиска в этих массивах (так называемого

ретроспективного поиска).

4. *Словарная статья характеризует средства закрепления, отображения и/или передачи данных*, например:

- **документальная информация [documentary information]** - сведения, закрепленные на каком-либо материальном носителе; содержание документа или текста;

- **устная информация [oral information]** - содержание устного сообщения и т.п.

ЗНАНИЯ [knowledge]

Совокупность хранимых в памяти человека или базах знаний фактов (данных, сведений) о некоторой предметной области, их взаимосвязей и правил, которые могут быть использованы для получения новых фактов и/или решения каких-либо задач, связанных с различными видами интеллектуальной деятельности людей или их сообществ.

Знания отражают множество возможных ситуаций, характеризующихся состоянием и конкретной реализацией объектов определенного типа, способы перехода от одного описания объекта к другому. Для знаний характерна внутренняя интерпретируемость, структурированность, связанность и активность. Условно можно записать: "знания = факты + убеждения + правила".

Часто понятие "знание" неверно отождествляют с "информацией". В указанном контексте "информация" может рассматриваться только как дополнение к "знанию", становящееся его частью после получения субъектом, которому она предназначена.

Различают следующие виды знаний:

- **базовые (фундаментальные) знания [deep knowledge]** - систематизированные знания, основанные на модели, описывающей все значимые аспекты некоторой предметной области, которые описывают ее сущности, их свойства и различные связи между ними;

- **компилятивные знания [compiled knowledge]** - знания, полученные на основе уже известных знаний путем их структурирования и/или систематизации в форме, необходимой для их использования в новых целях;

- **неполные знания [incomplete knowledge]** - знания, из которых не могут быть получены значимые факты в данной предметной области;

- **нечеткие знания [fuzzy knowledge]** - знания, основанные на нечеткой логике (см. далее);

- **поверхностные знания [surface knowledge]** - неполные, фрагментарные знания, часто связанные с недостаточной изученностью предметной области. Тем не менее, такие знания могут быть успешно использованы в случаях, не затрагивающих ситуаций, которые ими не охвачены;

- **предметные знания [domain knowledge]** - знания, относящиеся к конкретной узкой предметной области;

- **процедурные знания [procedural knowledge]** - знания, воплощенные в

компьютерных программах для решения тех или иных задач.

ЛОГИКА [logic]

Наука о законах и формах мышления, методах познания и условиях определения истинности знаний и суждений.

Нечеткая логика [fuzzy logic] — в математике и вычислительных системах: форма представления знаний или данных, связанных с описанием различных объектов понятиями вида тяжелый, громкий, горячий и т. п., имеющими неточные или неопределенные значения. Более строгое определение значений указанных понятий возможно только с привлечением ряда дополнительных сведений или данных, входящих в **нечеткие множества** и составляющих перечни дополнительных данных или условий. Центральным понятием нечеткой логики является понятие **вероятность члена множества**, определяющее степень правомерности отнесения данного члена к указанному множеству. Например, для члена нечеткого множества, относящего предметы к тяжелым, вес 20 кг может иметь такое значение с вероятностью 90% - для миниатюрных женщин и, скажем, только 20% - для мужчин.

Нечеткая логика широко используется в различного рода **экспертных системах** для автоматизированного принятия решений, близких к человеческим, на основе адекватного реагирования на сигналы, поступающие от связанных с ними датчиков, а также команды с пульта управления.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ [artificial intelligence]

1. Искусственная система, имитирующая решение человеком сложных задач, связанных с различными видами его деятельности.

2. Научное направление, связанное с созданием на базе средств вычислительной техники средств обработки больших объемов данных и выработки на основе **моделирования** органов человека и/или заданных им **алгоритмов** решений определенных практических задач. Примерами использования искусственного интеллекта являются "экспертные системы", "интеллектуальные системы" и "компьютерное зрение".

2. ИНФОРМАЦИОННЫЕ РЕСУРСЫ, ТЕОРИЯ ИНФОРМАЦИИ, ИНФОРМАТИКА

ИНФОРМАЦИОННЫЕ РЕСУРСЫ, ИР [information resources]

1. Федеральный закон РФ "Об информации, информатизации и защите информации" трактует данный термин как "... отдельные документы и отдельные массивы документов, документы и массивы документов в информационных системах: библиотеках, архивах, фондах, банках данных, других видах информационных систем".

2. В общем случае под ИР понимается вся совокупность сведений, получаемых и накапливаемых в процессе развития науки и практической деятельности людей для их многоцелевого использования в общественном

производстве и управлении. ИР отображают естественные процессы и явления, зафиксированные в результате научных исследований и разработок или других видов целенаправленной деятельности в различного рода документах (например, в отчетах о НИР, патентах, проектно-конструкторской документации, массивах данных и т. п.), понятиях и суждениях, а также более сложных моделях действительности.

Этот термин начал широко использоваться в конце 1970-х - начале 1980-х гг. в результате осознания растущей зависимости промышленно развитых стран, отдельных организаций и фирм от источников информации (технической, политической, военной и т.д.), а также от уровня развития и использования средств передачи и переработки информации. С ним связаны термины: **национальные информационные ресурсы** (в том числе **государственные и негосударственные информационные ресурсы**), информационные ресурсы территориально- административных образований, фирм (организаций), их подразделений и т. п.

В современном обществе ИР относятся к материальным и наиболее важным видам ресурсов, определяющим экономическую, политическую и/или военную мощь их владельца. В подтверждение этого тезиса можно привести ставший классическим пример: Япония, страна, практически лишенная природных ресурсов и обладающая весьма скромными людскими ресурсами, является крупнейшим в мире производителем и экспортером не только изделий микроэлектроники, но и такой материалоемкой продукции, как автомобили и супертанкеры.

Отличием ИР от других материальных видов ресурсов (например, полезных ископаемых) является их воспроизводимость. Как и другие виды ресурсов, ИР являются объектами импорта-экспорта, а также конкуренции, политической и экономической экспансии. Следует отметить, что границы понятия ИР в настоящее время четко не установлены.

Виртуальные (информационные) ресурсы [virtual (information) resources] - информационные ресурсы других организаций, предприятий, фирм и т.п., доступные пользователям в режиме теледоступа по каналам глобальной связи, например, Интернета.

ТЕОРИЯ ИНФОРМАЦИИ [information theory]

Раздел **Кибернетики**, изучающий общие стороны процессов передачи, хранения, извлечения и классификации информации различной природы (в том числе биологической, технической, социальной и др.) в независимости от ее семантического (смыслового) содержания. Общим средством анализа, описания и количественной оценки исследуемых процессов теории информации является ее математический аппарат, представляющий собой основу разрабатываемых и используемых теорий и методов.

Важнейшей частью теории информации является **Теория передачи информации [Theory of communication]**. Ее основоположник - американский математик К. Шеннон. Основные понятия этого раздела теории: **энтропия** (количественная мера неопределенности ситуации, от

греческого **entropia** - поворот, превращение) и **количество информации**, измеряемое величиной изменения энтропии в условиях, связанных с получением информации. С использованием этих понятий выражается пропускная способность канала связи между источником информации и ее адресатом, равная максимально допустимой скорости передачи информации со сколь угодно малой вероятностью ошибки.

Составной частью теории информации является также **Теория кодирования** (или **Теория оптимального кодирования**), рассматривающая вероятностные аспекты проблем кодирования и декодирования информации. Большой вклад в разработку теории информации внесли отечественные ученые А. Н. Колмогоров, А.Я. Хинчин, Р. Л. Добрушин, В.А. Котельников, А.А. Харкевич и др. Возросшая необходимость не только количественного, но и содержательного анализа информационных процессов породила появление новой науки - Информатики.

ИНФОРМАТИКА [informatics, Information science, Computer science]

Словарь по кибернетике содержит следующее определение: "**Информатика**: наука, изучающая информационные процессы и системы в социальной среде, их роль, методы построения, механизм воздействия на человеческую практику, усиление этого воздействия с помощью вычислительной техники. Информатика возникла как дополнение и конкретизация Теории информации из потребностей автоматизации социально-коммуникативных процессов и начала формироваться в 1970-е гг. как научная база использования электронных вычислительных машин в управлении, науке, проектировании, образовании, сфере услуг и т.д."

Как всякая относительно новая и быстро развивающаяся отрасль знания, не только связанная с социальной сферой, но и широко используемая в ней, Информатика получила в последние годы множество толкований, и не все они однозначны.

Наибольшие противоречия связаны с той частью этого понятия, которая определяет его семантические границы распространения. В качестве примера приведем другое определение: "**Информатика**: отрасль знания, изучающая закономерности сбора, преобразования, хранения, поиска и распространения документальной информации и определяющая оптимальную организацию информационной работы на базе современных технических средств".

Видимые отличия цитируемых определений заключаются, в частности, в том, что второе ограничивает понятие "Информатика" технологическими процессами, входящими в функции информационных органов, а также документальной информацией. Следует отметить, что, несмотря на давность этого определения (1971 г.), оно используется и в настоящее время в среде работников информационных органов и служб, в недрах которых изначально и было порождено.

Еще один подход связан с организациями, которые ранее были подведомственными Комитету по информатизации России. Основное внимание этот подход акцентирует на инструментальных (программных и

технических) средствах "**Информатики**" и "**информатизации**":
"Информатика: группа дисциплин, занимающихся различными аспектами применения и разработки ЭВМ: Прикладная математика, Программирование, Программное обеспечение, Искусственный интеллект, Архитектура ЭВМ, Вычислительные сети".

Анализируя сказанное, мы склонны предпочесть вариант "кибернетиков" как более объективный и полный.

С целью более глубокого понимания указанного термина продолжим выборочное цитирование соответствующей статьи словаря по кибернетике:

"Важнейшими категориями информатики являются понятия информационных сред (социальных подсистем, в которых осуществляются информационные процессы и куда внедряются ЭВМ как усилители человеческого интеллекта), полного информационного цикла (включающего зарождение информации, ее переработку, передачу, использование для снижения энтропии рассматриваемой социальной системы), полезной работы (отдачи) ЭВМ. Отдача ЭВМ, коэффициент полезного действия зависят от уровня функционирования социальной среды, в которой они задействованы, ее упорядоченности, системности, условий для творческой деятельности людей, сложности и важности задач, решаемых с помощью машин.

Информатика не заменяет собой кибернетику, теорию информации, электронику, системотехнику, а взаимодействует с ними, имея ряд общих проблем. Интегральный характер информатики заключается также в ее взаимодействии с такими дисциплинами, как теория познания, семиотика, лингвистика, документалистика, библиотековедение".

3. НОСИТЕЛИ ИНФОРМАЦИИ, ДОКУМЕНТЫ, ДОКУМЕНТАЦИЯ, ИЗДАНИЯ

НОСИТЕЛЬ [media]

Обобщающее наименование материала, на который можно записывать данные. Носители подразделяются на человекочитаемые (твердые) и машиночитаемые.

- **Машиночитаемый носитель [machine-readable media]** — носитель, пригодный для непосредственной записи и считывания данных программно-техническими средствами (ЭВМ). Термин обычно применяется к устройствам внешней памяти ЭВМ, например, магнитным и оптическим дискам, дискетам и т.п. Однако он может использоваться и по отношению к определенной части твердых носителей, если они допускают использование специальных считывающих устройств, например, сканеров.

- **Человекочитаемый (твердый) носитель [human readable media, hard media]** - носитель, пригодный или используемый для записи данных, непосредственно считываемых человеком, например — бумага.

- **Микроформа, микроноситель [microform]** - общее наименование носителей, на которых тексты или графические изображения представлены в уменьшенном фотографическом способе виде. Микроформы широко

используются в библиотеках для хранения в человекочитаемом виде (чтение их производится с использованием специальных читальных аппаратов) больших объемов копий документов и данных. Создание микроформ производится при помощи специальной (микрофильмирующей) фотоаппаратуры, а также компьютерной записи. Основными типами микроносителей являются:

1. **Микрофильм [microfilm]** - пленка обычного фотоаппарата содержащая уменьшенные копии текстов и графических изображений;

2. **Микрофиша [microfiche]** - стандартный (105 x 148 мм) прямоугольный лист фотопленки, на котором располагается 420 кадров уменьшенных изображений страниц текста или графики.

ДОКУМЕНТ [document — от лат. documentum — свидетельство, доказательство]

Материальный носитель информации, зафиксированной вне памяти человека. В соответствии с ГОСТ 16487-83, "**документ**" является средством "закрепления различным способом на специальном материале сведений о фактах, событиях явлениях объективной действительности и мыслительной деятельности человека". Документы могут содержать текстовую, цифровую, графическую и аудиоинформацию. Они могут подвергаться процессам записи (преобразования), хранения, поиска, передачи, получения, сбора и чтения. В свою очередь, ГОСТ 7.60-2003 понимает под документом "зафиксированную на материальном носителе информацию с реквизитами, позволяющими ее идентифицировать". В зависимости от характера обработки содержащихся в них данных документы принято разделять на **первичные** и **вторичные**:

• **Первичный документ [primary document, source]** - документ, содержащий исходную запись сведений, полученных в процессе исследований, разработок, наблюдений, анализа или других видов человеческой деятельности независимо от ее характера или тематики. Деление документов на первичные и вторичные в значительной степени условно, поскольку один и тот же документ может содержать сведения, относящиеся и к первой, и ко второй группе. Так, считающиеся безусловно первичными видами документов отчеты по НИР, проектные документы, монографии и т.д. могут содержать наряду с оригинальными сведения или данные ранее включенные в другие первичные или даже вторичные документы.

• **Вторичный документ [secondary document]** - документ, полученный в результате аналитико-синтетической и логической переработки сведений или данных, содержащихся в первичных документах. Примерами вторичных документов являются справочные и энциклопедические издания, рефераты и реферативные издания, библиографические издания, указатели и списки, обзоры (за исключением аналитических обзоров) и т. п.

Юридический документ [juridical (legal) document] - документ, оформленный в соответствии с действующим юридическим законодательством и имеющий правовое значение.

В зависимости от характера связи документов с технологическими процессами в автоматизированных системах и вида физического носителя информации различают:

- **машинно-ориентированный документ [machine-oriented document]** - документ, предназначенный для обработки части содержащейся в нем информации средствами вычислительной техники. Примерами машинно-ориентированных документов могут служить заполненные специальные формы бланков библиографической записи, различного рода анкет и т.п., предназначенные для последующего считывания в ЭВМ записанных в них данных с использованием клавиатурных операций;

- **машиночитаемый документ [machine-readable document]** - документ, пригодный для автоматического считывания содержащейся в нем информации. Средства автоматического считывания - **сканеры** предъявляют определенные требования к характеру оформления соответствующих текстовых, графических и других видов записей, включая их формат, виды шрифтов, наличие специальных служебных знаков, контраст и т. п.;

- **документ на машиночитаемом носителе, электронный документ [electronic document]** - документ, созданный средствами вычислительной техники, записанный на машиночитаемый носитель: магнитную ленту (МЛ), магнитный диск (МД), дискету, оптический диск и т. п. и оформленный в установленном порядке. В зависимости от того, на каком носителе записан машиночитаемый (электронный) документ или документы, принято указывать его вид, например, "**документ(ы) на магнитном (оптическом) диске**", "**документы на магнитной ленте**" и т. п.;

- **документ-машинограмма, распечатка [hard copy document]** - документ на бумажном носителе, созданный средствами вычислительной техники и оформленный в установленном порядке;

- **документ на экране дисплея** — документ, созданный средствами вычислительной техники, отображенный на экране дисплея (монитора) и оформленный в установленном порядке;

- **title** — **электронный документ**, который идентифицируется и передается как единое целое;

- **служебный документ [internal document, in-house document]** — документ, содержание которого отражает способ или результаты решения какой-либо функциональной задачи автоматизированной системы;

- **входной документ [input document]** - документ, составленный по определенной форме и содержащий данные, предназначенные для ввода в ЭВМ. Входные документы могут быть условно разделены на две категории: **информационные документы и запросы**;

- **информационный документ [information document]** - документ, основное назначение которого - пополнение массивов или баз данных (БД) ЭВМ;

- **выходной документ [output document]** - документ, являющийся носителем результатов обработки данных ЭВМ и/или формируемый автоматизированной системой и выданный системными средствами вывода.

ДОКУМЕНТАЦИЯ [documentation, collection of documents, file of documents]

Совокупность документов, объединенных по определенным признакам (например, по назначению, содержанию и т. п.) и оформленных по единым правилам.

Применительно к программным продуктам и средствам вычислительной техники термином "*документация*" обозначают:

1. Руководства по использованию;
2. Совокупность текстов, описывающих строение и применение соответствующих средств или изделий.

Документация предназначена для облегчения использования программных и технических средств и включает руководства, справочники, учебники, краткие справочники, обучающие программы, а также средства диалоговой документации и подсказки.

СИСТЕМА ДОКУМЕНТАЦИИ [documentation system]

Совокупность документов, состав, содержание, структура и правила оформления которых определены государственными стандартами. Часто понятие "*система документации*" подменяется другим - **вид документации**, которое, строго говоря, не является его синонимом. Так, различают следующие виды документации: научно-техническая (научная и техническая), директивная, юридическая, нормативная (нормативно-техническая), технологическая, проектно-конструкторская (проектная и конструкторская), эксплуатационная, технико-экономическая и др. Данная классификация достаточно условна в том смысле, что состав и правила оформления большинства используемых в реальной практике видов документации нормативно не определены (ГОСТ 6.10.1-88).

ДОКУМЕНТАЛИСТИКА [Documentation]

Направление в Кибернетике, занимающееся изучением и оптимизацией документальных систем независимо от их назначения (в первую очередь - документов).

ИЗДАНИЕ [publication, edition]

ГОСТ 7.60-2003 трактует издание, как "документ, предназначенный для распространения содержащейся в нем информации, прошедший редакционно-издательскую обработку, самостоятельно оформленный, имеющий выходные сведения". В зависимости от способа исполнения издания делятся на печатные и электронные.

Печатное издание [publication, edition] — издание, полученное печатанием или тиснением и имеющее самостоятельное полиграфическое оформленное.

Печатные издания различаются по многим признакам, включая периодичность выпуска, содержание, знаковую природу информации, ее

целевое назначение, вид аналитико-синтетической переработки содержащегося в них материала и т.д.

В зависимости от периодичности выхода различают:

- **непериодическое издание [non-periodical edition]** — издание, выходящее однократно, не имеющее продолжения;

- **сериальное издание [serial edition]** - издание, выходящее в течение времени, продолжительность которого заранее не установлена, как правило, нумерованными и/или датированными выпусками (томами), имеющими одинаковое заглавие;

- **периодическое издание [periodical edition]** - сериальное издание, выходящее через определенные промежутки времени, как правило, с постоянным для каждого года числом номеров (выпусков), неповторяющимися по содержанию, однотипно оформленными, нумерованными и/или датированными выпусками, имеющими одинаковое заглавие. Периодические издания могут быть ежедневными, еженедельными, ежемесячными, ежеквартальными или ежегодными;

- **продолжающееся издание [continued edition]** - сериальное издание, выходящее через неопределенные промежутки времени по мере накопления материала, не повторяющимися по содержанию, однотипно оформленными, нумерованными и/или датированными выпусками, имеющими общее заглавие.

Виды научных изданий:

- **монография [monograph]** — научное или научно-популярное издание, содержащее полное и всестороннее рассмотрение одной проблемы или темы и принадлежащее одному или нескольким авторам;

- **сборник научных трудов [collection of scientific papers]** - сборник (т.е. издание, содержащее ряд произведений одного или разных авторов), содержащий исследовательские материалы научных учреждений, учебных заведений или обществ;

- **материалы конференции (съезда, симпозиума) [proceedings]** - непериодический сборник, содержащий итоги конференции в виде докладов, рекомендаций, решений;

- **препринт [preprint]** - научное издание, содержащее материалы предварительного характера, опубликованные до выхода в свет издания, в котором они могут быть помещены;

- **пролегомены, введение [prolegomena]** - научное или учебное издание, содержащее первичные сведения и основные принципы какой-либо науки;

- **тезисы докладов (научной конференции, съезда, симпозиума) [scientific conference abstracts]** - научный непериодический сборник, содержащий опубликованные до начала конференции материалы предварительного характера (аннотации, рефераты докладов и/или сообщений);

- **автореферат диссертации [authors abstract, synopsis of thesis]** - научное издание в виде брошюры, содержащее составленный автором реферат проведенного им исследования, представляемого на соискание ученой сте-

пени;

- **стандарт [standard]** — нормативное производственно-практическое издание, содержащее комплекс норм, правил, требований к объекту стандартизации, которые устанавливаются на основе достижений науки, техники и передового опыта и утверждаются в соответствии с действующим законодательством. В зависимости от уровня организации, утвердившей и выпустившей стандарт, стандарты подразделяются на международные, государственные, отраслевые и стандарты предприятий. Требования стандартов вышестоящих уровней обязательно распространяются на все стандарты нижестоящих уровней;

- **электронное издание, ЭИ [electronic publication]** - существуют два определения ЭИ, которые хорошо дополняют друг друга и поэтому могут рассматриваться в непосредственной взаимосвязи:

1. ЭИ - электронный документ (документ или группа документов на машиночитаемом носителе), прошедший редакционно-издательскую обработку, предназначенный для распространения в неизменном виде, имеющий выходные сведения. По наличию печатного эквивалента ЭИ подразделяются на электронные аналоги печатного издания (полностью их воспроизводящие) и самостоятельные электронные издания;

2. ЭИ - самостоятельный законченный продукт, содержащий информацию, представленную в машиночитаемой (электронной) форме, все копии (экземпляры) которого соответствуют оригиналу, и предназначенный для длительного хранения, широкого распространения и многократного использования неопределенным кругом пользователей.

Электронные издания представляют собой основную часть **контента электронных библиотек**.

По характеру содержащейся в ЭИ информации различают:

- **текстовое (символьное) ЭИ [textual (symbol) electronic publication]** - ЭИ, которое содержит преимущественно текстовую информацию, представленную в форме, допускающей посимвольную обработку;

- **изобразительное ЭИ [graphic(al) electronic publication]** - ЭИ, содержащее преимущественно электронные образы объектов (в том числе текстов), рассматриваемых как целостные графические сущности, которые представлены в форме, допускающей просмотр и печатное воспроизведение, но не допускающей посимвольной обработки;

- **звуковое ЭИ [audio (sound) electronic publication]** - ЭИ, содержащее цифровое представление звуковой (аудио-) информации в форме, допускающей ее прослушивание, но не предназначенной для печатного воспроизведения;

- **программный продукт [software product]** - самостоятельное отчуждаемое произведение, представляющее собой публикацию текста программы (или программ) на языке программирования или в виде исполняемого кода;

- **мультимедийное ЭИ [multimedia electronic publication]** - ЭИ, в котором информация различной природы присутствует равноправно и

взаимосвязанно для решения определенных разработчиком задач, причем эта взаимосвязь обеспечена программными средствами.

По технологии распространения различают:

- **локальное ЭИ [local electronic publication]** - ЭИ, предназначенное для локального использования и выпускающееся в виде определенного количества идентичных экземпляров (тиража) на переносных машиночитаемых носителях;
- **сетевое ЭИ [network electronic publication]** - ЭИ, доступное потенциально неограниченному кругу пользователей через телекоммуникационные сети;
- **ЭИ комбинированного распространения [local and wide-spread propagation electronic publication]** - ЭИ, которое может использоваться как в качестве локального, так и сетевого.

4. ПРИНЦИПЫ СТРУКТУРИРОВАННОГО ПРЕДСТАВЛЕНИЯ ДОКУМЕНТОВ И ДАННЫХ

4.1. ИНФОРМАЦИОННЫЕ ЭЛЕМЕНТЫ И ИХ ВИДЫ

ИНФОРМАЦИОННЫЙ ЭЛЕМЕНТ, ИЭ [information element, item]

Единица информации, подлежащая обработке, хранению и передаче пользователям системы или предназначенная для обеспечения ее работы. ИЭ является обобщенным наименованием структурной единицы информации, не зависящей от ее назначения, состава данных, характера или материала носителя и т.п. Например, ИЭ может быть названо содержание книги, статьи, библиографического описания или его части. С учетом сказанного, ИЭ могут быть делимыми на части или не делимыми.

Виды информационных элементов:

- **Составной ИЭ [composite data item]** - информационный элемент, который может быть разделен на части средствами системы (без участия человека в выполнении логических или интеллектуальных операций, связанных с анализом и разделением содержащихся в этом информационном элементе данных), причем каждая из его частей также является информационным элементом системы.

Исходя из этого определения, составной ИЭ характеризуется видом и организацией содержащихся в нем данных. Так, неструктурированные вторичные документы — библиографическое описание, реферат и аннотация - не могут считаться составными ИЭ, если **автоматизированная система** не позволяет на физическом уровне своими средствами выделить содержащиеся в них различные виды данных (например, в библиографическом описании: сведения об ответственности, фамилии авторов, продолжение названия и т.п.) или аспекты описания.

В общем случае ИЭ, которые не могут быть разделены на части автоматизированными средствами, принято называть **данными**.

В указанном контексте уместно упомянуть еще одно определение

понятия "**данные**": это "информация, представленная в виде пригодном для обработки автоматическими средствами при возможном участии человека".

- **Элемент данных, ЭД [data element, data item]** - неделимый информационный элемент. ЭД является минимальной структурной единицей информации, поскольку части ЭД теряют признаки информационного элемента. Примерами ЭД в библиографическом описании могут служить: фамилия автора, год издания, название издания и т.п. Деление на части лексических единиц, образующих соответствующие понятия, не приводит к образованию новых ИЭ.

Из вышеизложенного следует, что вид ЭД определяется как характером содержащихся в нем сведений, так и особенностью его организации или записи.

В зависимости от характера содержащихся в ИЭ данных они могут подразделяться на следующие виды:

- **документографический (документальный) ИЭ [documentary data item]** - информационный элемент, содержащий сведения о документах (например: библиографическое описание, реферат, аннотация, поисковый образ документа);

- **полнотекстовый ИЭ [full text(ual) data item]** - информационный элемент, содержащий полные тексты документов или их частей;

- **фактографический ИЭ [fact(ual) data item]** - информационный элемент, содержащий описание отдельных фактов или некоторой ограниченной совокупности фактов, объединенных по каким-либо признакам (например: сведения о курсе валют, расписания движения транспорта и т. п.);

- **объектографический ИЭ [object(ual) data item]** - информационный элемент, содержащий структурированное и логически связанное описание различных объектов науки, техники или управления. Объектографические информационные элементы являются более сложным вариантом фактографических ИЭ, поэтому многие специалисты и организации, работающие в области информатики, используют в рассматриваемых случаях термин фактографический ИЭ или его синонимы.

ТИП ДАННЫХ [data type]

Множество допустимых значений данных, объединенных общим содержанием и именем (например: "Библиографическое описание", "Автор", "Год издания", "Стоимость" и т. п.), а также совокупностью допустимых операций, которые можно выполнять над этими данными, включая способ их хранения в памяти ЭВМ. Понятие "тип данных" делает манипулирование данными с использованием средств вычислительной техники абстрактным процессом и скрывает лежащее в основе обращения с ними представление их в виде двоичного кода.

Виды типов данных:

- **аналоговые данные [analog data]** - данные, принимающие произвольные значения из заданного диапазона и представляемые в виде

непрерывно изменяющихся физических величин, например, напряжения, длительности сигнала и т. п.;

- **дискретные (цифровые) данные [digital data]** - данные, представленные в дискретном коде в определенной, например, двоичной системе счисления;

- **аналого-цифровые данные [analog-digital data]** - аналоговые данные, преобразуемые для обработки в цифровой код;

- **двоичные данные [binary data]** - данные, представленные в двоичном коде;

- **десятичные данные [decimal data]** - данные, представленные в десятичном коде;

- **алфавитно-цифровые (текстовые) данные [alphanumeric data]** - данные, значения которых составлены из любых знаков алфавита;

- **числовые (арифметические) данные [arithmetic data]** - данные, над которыми можно выполнять арифметические операции.

МНОЖЕСТВО [set]

Совокупность каких-либо объектов, представляемых как единое целое. Множество может включать в себя как однородные объекты, объединенные каким-либо общим признаком, так и неоднородные. Объединение последних в конкретное "множество" определяется на основе ассоциативных и других связей между его элементами.

Виды множеств:

- **конечное множество [finite set]** - множество, содержащее определенное (конечное) число элементов;

- **нечеткое множество [fuzzy set]** - множество, принадлежность объекта к которому определяется функцией, принимающей значения в области $[0,1]$;

- **ограниченное множество** - упорядоченное множество, имеющее нижнюю и верхнюю границы;

- **пустое множество [empty set]** - множество, не имеющее ни одного элемента;

- **подмножество [subset]** - множество, являющееся частью другого множества;

- **эквивалентное множество [equivalent set]** - множество, имеющее взаимно однозначное соответствие с другим множеством, при котором каждому элементу одного из них соответствует один единственный элемент другого и наоборот.

ТЕОРИЯ МНОЖЕСТВ [set theory]

Раздел математики, исследующий общие свойства множеств и операций над ними. Понятия и методы теории множеств широко используются в информатике, в частности, в теории и практике построения информационно-поисковых систем.

4.2. ЗАПИСЬ, ФАЙЛ, МАССИВ, КЛЮЧ

ЗАПИСЬ [record, writing]

1. Единица обмена данными между программой и внешней памятью ЭВМ (record).
2. Процесс или результат закрепления (фиксирования) данных на носителе информации (writing).

Запись может содержать один информационный элемент, включая и составной.

Логическая запись [logical record] - совокупность записей взаимосвязанных ИЭ (в том числе элементов данных, данных и составных ИЭ), рассматриваемая в логическом плане как единое целое. Одна логическая запись может состоять из нескольких физических или быть частью одной физической записи.

Физическая запись [physical record] - порция данных, пересылаемых как единое целое между основной и внешней памятью ЭВМ.

На машиночитаемом носителе (например, магнитном или оптическом диске, дискете и т.п.) физическая запись реализуется в форме поля данных, файла (см. далее) и др. Физическая запись может содержать одну логическую запись, ее часть или несколько логических записей.

ТИП ЗАПИСИ [record type]

1. Обобщенное имя записи типа данных.
2. В базах данных: тип, к которому относится данная запись.
3. Характеристика, определяющая возможность записи менять свою длину (по этому признаку различают записи: **фиксированной, переменной и неопределенной длины**).

БЛОК ДАННЫХ, БЛОК [data block, block]

1. Несколько последовательных логических записей, объединенных в одну физическую.
2. Единица доступа к внешней памяти ЭВМ.
3. Выделенный фрагмент текстового материала, который можно удалить, переместить или выполнить над ним некоторые преобразования.
4. В вычислительных сетях: последовательность передаваемых данных, рассматриваемая как единое целое. Обычно имеет **адрес и контрольную сумму**, позволяющую обнаруживать ошибки, и другие служебные элементы, требуемые соответствующим **протоколом** канала связи.

ПАКЕТ [packet]

Единица информации в сети передачи данных. Пакет имеет строго определенную структуру. Она предусматривает наличие заголовка, который содержит адреса получателя и отправителя, данные для контроля ошибок, а также самого сообщения или его части, поскольку передаваемое сообщение (например текстовый файл) может быть разделено и пересылаться в виде

последовательности пакетов.

ФАЙЛ [file]

1. Идентифицированная последовательность или множество записей однотипных **информационных элементов**.

2. Поименованная целостная совокупность данных на машиночитаемом носителе.

3. Поименованная область внешней памяти ЭВМ.

4. В английском языке: картотека, архив, комплект, подшивка и т. п.

Файл - основной структурный элемент хранения данных в ЭВМ, обеспечивающий возможность машине и человеку отличать один набор данных от другого при их поиске, изменении, удалении или выполнении с ними других операций. Содержанием файла может стать одна поименованная логическая запись или несколько логических записей, объединенных общим именем. Например, файл может содержать записи некоторого множества библиографических описаний, каждое из которых по своей сути является логической записью, объединенной с другими каким-либо общим признаком (вид документа, тематика и т.п.). В других случаях содержанием записи в файле могут быть отдельные документы, управляющие работой ПЭВМ, программы или их части. "Идентификация" или "поименование" указанных записей или групп записей выражается присвоением каждому из них "уникального" или отличного от других имени, позволяющего автоматизированным средствам находить их в массивах других записей, в том числе файлов.

Наименования файлов состоят из двух частей:

1) **основного имени файла [basic file name (filename)]**, отражающего в краткой форме содержание и/или назначение находящихся в нем данных;

2) **расширения имени файла [file name broadening]**, которое характеризует вид записанных данных и их организацию. Расширение отделяется от основного имени файла точкой и записывается в форме кода, принятого для каждого вида формата файлов. Код расширения, как правило, содержит от двух до четырех буквенных или буквенно-цифровых символов, например: "*.aiff", "*.au", "*.exe", "*.doc", "*.txt", "*.tif", "*.jpg", "*.wav", "*.3ds" и т. п.

На физическом уровне содержимое файла может быть не структурировано и представлять собой единственную физическую запись, или структурировано, например, полями данных. В последнем случае оно будет включать соответствующее множество физических записей.

Некоторые типы файлов:

• **Файл данных [data file]** - файл, содержащий данные, в отличие от программных файлов, содержащих записи программ или их частей.

• **Файл регистрации [log file]** - файл, в котором хранятся записи о других файлах, например, если какие-либо файлы были архивированы, список имен этих файлов может храниться в файле регистрации.

• **Файл с произвольным доступом [random-access file]** - файл, в

котором любая запись может быть считана, записана или изменена без необходимости считывания других записей. Произвольный доступ должен быть обеспечен операционной системой и реализован специальной машинной программой.

- **Основной файл [master file]** - файл, являющийся основным источником данных для решения определенного класса задач, для какой-либо цели или назначения. Поддержка этого файла в актуальном состоянии, при котором содержащиеся в нем данные отвечают требованиям новизны и точности, производится с использованием операций обновления файла.

- **Упорядоченный файл [sequential file]** - файл, в котором записи упорядочены по ключевому полю для ускорения доступа к определенной записи.

- **Последовательный файл [serial file]** - файл, в котором записи не упорядочены; поэтому чтобы прочитать нужную запись, требуется прочитать все предыдущие.

- **Программный файл [program file]** - файл, содержащий программу или ее часть, написанную на каком-либо из языков программирования.

- **Исполняемый файл [executable file]** - программный файл, предназначенный для запуска операционной системой. К исполняемым файлам относятся: командный файл, имеющий расширение "*.bat", и машинный файл, имеющий расширение имени (см. ранее) "*.com" или "*.exe".

- **Скрытый файл [hidden file]** - файл, который не отображается на экране монитора при просмотре каталога файлов (директории). Скрытым можно сделать любой файл с целью затруднения его нахождения, чтения и/или порчи другими лицами, которые могут иметь доступ к ЭВМ и др.

МАССИВ [array]

1. Упорядоченная структура множества документов или данных одного типа.

2. Поименованная совокупность однотипных (логически однородных), упорядоченных по индексам записей ИЭ.

3. Упорядоченное множество элементов одного типа. Каждый элемент массива должен иметь имя (идентификатор, индекс), обеспечивающее возможность его нахождения. Элементы массива могут быть одномерными и многомерными.

Применительно ко многим задачам автоматизированной обработки данных термины массив и файл могут использоваться как синонимы. Однако соответствующие понятия имеют и отличия: так, понятие массив не обязательно связано с записями информационных элементов (включая полнотекстовые документы) на машиночитаемых носителях. Понятие же файл в русскоязычной практике, как правило, предполагает указанное условие.

Информационно-поисковый массив, поисковый массив [information collection, file] - массив документов или данных (соответствующие англ.

эквиваленты - **document collection, data collection**), в котором производится информационный поиск.

АРХИВ [archiwes]

Организованная совокупность массивов данных или программ, длительно хранимых на внешних машиночитаемых носителях (например, на гибких магнитных дисках, магнитных лентах и/или CD™ROM) с целью обеспечения возможности их дальнейшего использования.

Целями создания архивов являются:

1. Создание страховочных копий информационных и программных продуктов на случай их утраты или порчи в ходе эксплуатации вычислительных средств;
2. Освобождение внешней памяти ЭВМ (например, накопителя на жестком магнитном диске) от программ и данных, потребность в оперативном использовании которых частично, полностью или временно отпала.

КЛЮЧ [key]

1. Информационный элемент, однозначно идентифицирующий запись или указывающий ее местоположение. Ключ может также служить средством для идентификации некоторого множества (в том числе массива) записей и располагаться в одном или нескольких из его полей.

В качестве ключа может служить элемент данных (в том числе индекс, адрес, код и др.) или группа элементов данных. Так, в индексно-последовательном файле ключ является обязательным элементом записи в каждом поле данных. Набор значений некоторой совокупности атрибутов в реляционной модели данных служит ключом, который однозначно идентифицирует кортеж или группу кортежей конкретного отношения.

2. Параметр шифрования, представляющий один из возможных вариантов шифра. Для дешифрирования помимо ключа необходимо знать последовательность операций или правило его использования.

3. Значение, используемое для подтверждения полномочий на доступ к некоторой информации (в том числе к базам данных, отдельным файлам и т.п.), программным и/или техническим средствам.

Единицы измерения количественных показателей записей данных в ЭВМ:

БИТ [bit - от англ. binary digit] - простое двоичное число (цифра или символ), принимающее значения 1 или 0 и служащее для записи и хранения данных в ЭВМ. Бит является минимальной двоичной единицей измерения энтропии и количества информации в ЭВМ, соответствующей одному двоичному разряду. Энтропия сообщения, выраженная в битах, определяется средним числом символов, необходимых для записи этого сообщения. Определенное количество бит составляет размер других единиц - двоичных слов, в том числе – байта, килобайта, мегабайта и т.д.

Байт [byte] - двоичное слово, способное записывать и хранить в памяти ЭВМ один буквенно-цифровой или другой символ данных. Каждый символ записывается в виде набора двоичных цифр (битов) при помощи определенного кода, например, ASCII. Количество бит в байте определяет его разрядность, которая может составлять 8, 16, 32 и т.д. Соответственно байт называют 8-разрядным, 16-разрядным и т.д. Один 8-разрядный байт может определять 256 разных значений, например, десятичных чисел от 0 до 256. Увеличение разрядности ведет к соответствующему увеличению числа возможных вариантов комбинаций, кодируемых одним байтом. Например, 16-разрядным — до 65536 или 2^{16} , 32-разрядным - до 2^{32} и т. д.

Килобайт, Кбайт [kilobyte] - единица измерения емкости памяти или длины записи, равная 1024 байтам. Часто под килобайтом понимается также величина, равная 10^3 байт.

Мегабайт, Мбайт [megabyte] - единица измерения емкости памяти или длины записи, равная 1024 Кбайт. Часто под мегабайтом понимается также величина, равная 10^3 килобайт или 10^6 байт.

Гигабайт, Гбайт [gigabyte] - единица измерения емкости памяти или длины записи, равная 1024 Мбайт. Часто под гигабайтом понимается также величина, равная 10^3 мегабайт, 10^6 килобайт или 10^9 байт.

Терабайт, Тбайт [terabyte] - единица измерения емкости памяти или длины записи, равная 1024 Гбайт. Часто под терабайтом понимается также величина, равная 10^3 гигабайт, 10^6 мегабайт, 10^9 килобайт или 10^{12} байт.

Кубит [quantum bit, qubit] - "**Квантовый бит**": мера и измерения объема памяти в теоретически возможном виде компьютера, использующем квантовые носители, например, спины электронов. Кубит может принимать не два различных значения (0 и 1), а несколько, соответствующих нормированным комбинациям двух основных состояний спина, что дает большое число возможных сочетаний. Так, 32 кубита могут образовать около 4 млрд состояний. **Спин** (от англ. spin - вращаться, вертеться) - собственный момент количества движения элементарных частиц (например, электрона, протона, нейтрона, нейтрино и т.д.), имеющий квантовую природу и не связанный с перемещением частицы как целого.

ЗНАК [character]

Один символ, который может быть представлен и воспринят ЭВМ. К знакам относятся буквы, цифры, пробелы, знаки препинания, специальные символы (например, математические, кодовые и т. п.).

4.3. СТРУКТУРЫ, МОДЕЛИ ДАННЫХ

СТРУКТУРА [structure]

Фиксированное упорядоченное множество объектов и связей между ними.

С понятием "*структура*" связаны следующие термины:

- **структура данных [data structure]** - множество элементов данных,

объединенных и упорядоченных определенным образом;

- **структура информационной базы [information support structure]** - упорядоченная по определенным правилам совокупность подмножеств записей информационных элементов, образующих информационную базу и необходимых для реализации функций автоматизированной системы;

- **структура базы данных [DB structure]** - принцип или порядок организации записей в базе данных и связей между ними. Структуру БД принято рассматривать на разных уровнях абстракции (представления) и, в частности, концептуальном (с позиции администратора предприятия), реализации или внешнем (с позиций конечного пользователя и прикладного программиста), физическом или внутреннем (с позиций системного аналитика и системного программиста). Соответственно этим уровням различают концептуальную, внешнюю и физическую модели и/или схемы организации данных;

- **абстракция [abstraction]** - использование для описания или представления общих свойств объекта без конкретной его реализации (например, типов объектов — "читатель", "фирма", "автомобиль", но не конкретно названных читателей, фирм, марок автомобилей и т. п.);

- **абстрактная структура данных [abstract logic design]** - структура данных, определенная функционально посредством выполняемых на ней операций. Такая структура не связана с поименованными типами объектов.

СПИСОК [list]

1. Структура данных, представляющая собой логически связанную последовательность записей - элементов списка;

2. Перечень различных объектов.

Связанные со "*структурой*" и "*списком*" понятия:

- **цепной список, связанный список [chained (linked) list]** - список данных, в котором порядок элементов списка задан посредством указателей, включенных в их запись;

- **подсписок [sublist]** - ветвь списковой структуры, представляющая собой отдельный цепной список, на который имеется указатель от одного из элементов цепного списка вышестоящего уровня иерархии в данной структуре;

- **цепная структура [chain structure]** - структура цепного списка;

- **ассоциативная (ассоциативно-адресная) структура [associative structure]** - совокупность цепного списка и всех связанных с ним подсписков. Различают объектные и признаковые списковые структуры;

- **однородная структура [structure]** - структура, состоящая из однотипных элементов;

- **комбинированная структура [combined structure]** - структура данных, полученная путем объединения (композиции) нескольких исходных структур.

МОДЕЛЬ [model]

1. Результат корректного воспроизведения каким-либо способом или средствами различных объектов (в том числе процессов и явлений реального мира или мыслительной деятельности человека). Модели являются, с одной стороны, продуктом изучения свойств соответствующих объектов, процессов и явлений предметной области, с другой — служат инструментом для углубления знаний о них, а также решения разнородных прикладных задач (см. далее также "Моделирование"). В зависимости от характера средств, используемых для построения (создания) "моделей", последние подразделяются на описательные, математические, физические и комбинированные (например, физико-математические модели). Различают также статические и динамические модели (в том числе кибернетические модели) и др.

2. Тип, марка, образец конструкции (например, модель автомобиля - "ВАЗ 21099").

3. Образец для подражания, образцовый экземпляр какого-либо изделия.

4. Оригинал, который служит для снятия копии, изображения или создания другого произведения, имеющего признаки сходства с ним (примерами могут служить неодушевленные предметы и люди, в том числе - "фотомодели", "натурщики" в живописи и т. п.).

МОДЕЛЬ ДАННЫХ [data model]

1. Представление данных и их взаимосвязей (отношений), описывающих понятия проблемной среды. Модели данных используются для представлений структур данных на концептуальном и внешнем уровнях, но не физическом.

2. Совокупность правил порождения структур данных в БД и выполнения операций над ними.

3. Формализованное описание структур данных и операций над ними.

МОДЕЛИРОВАНИЕ [modeling]

Методология выполнения экспериментальных работ путем исследования свойств различных объектов на их моделях. Виды моделирования различаются в зависимости от целей его выполнения, характера исследуемых объектов и выбранных для исследования средств. В связи с развитием вычислительной техники наиболее широкое применение в различных областях человеческой деятельности получило математическое моделирование.

Математическое моделирование [mathematical modeling] - процесс построения и исследования в динамике поведения математических моделей различных процессов, явлений и физических объектов с использованием средств вычислительной техники. В основе математического моделирования лежит использование фундаментальных законов естествознания и конкретных наук, связанных с целями и предметами моделирования. На основании этих законов разрабатывается математический аппарат, описывающий исследуемые явления и объекты, который преобразуется в соответствующий алгоритм и программу для реализации его на ЭВМ. В

зависимости от общих целей математического моделирования, полученные в ходе "испытания" математической модели данные могут использоваться для принятия определенных решений, в частности, для выбора альтернативных вариантов поведения специалистов или для уточнения исходной математической модели (например, в виде поправок, вводимых в математический аппарат, и таким образом использоваться в качестве средства ее совершенствования). Методы математического моделирования нашли широчайшее применение при построении так называемых интеллектуальных (экспертных) систем, систем автоматизированного проектирования и др.

Некоторые общие термины, связанные с моделированием и моделями данных:

- **логическая структура [logical structure]** - представление логической организации данных в виде множества типов записей данных и связей между ними;

- **структурная модель данных [structured data model]** - модель данных, представленная в виде структуры - множества типов данных и связей между ними. Различают три основных вида структурной модели (логической структуры) организации данных: иерархической, сетевой и реляционной;

- **связь (между данными) [binding, link, relationship]** - установленный характер взаимозависимости данных в различных информационных моделях и структурах данных. Связи между данными идентифицируются видом связи и направлением. Характер вида и направления связи могут отражаться ее именем - указателем связи (поименованная связь, характеристика отношения). Различают иерархическую или вертикальную, горизонтальную, ассоциативную, логическую, двунаправленную связи и др.;

- **иерархическая (вертикальная) связь [hierarchical binding, vertical binding]** - вид связи, устанавливаемый между данными, находящимися на разных уровнях иерархической структуры (например, связи: "род-вид", "вид-подвид", "отец-сын" и т. п.);

- **горизонтальная связь [horizontal binding]** - вид связи, устанавливаемый между данными, находящимися на одном уровне иерархической структуры;

- **логическая связь [logical relationship]** - вид связи, устанавливаемый между типами данных в иерархической и сетевой моделях данных, в отличие от связи между конкретными экземплярами данных в базах данных;

- **ассоциативная связь [associative link]** - вид связи, устанавливаемый исходя из заданного сочетания признаков данных или информационных элементов, которые образуют таким путем упорядоченные последовательные цепочки. Указатели на связанные данные (адреса связи) могут размещаться в самих данных или программных средствах управления базой данных ;

- **двунаправленная связь [bidirectional binding]** - совокупность связей в прямом и обратном направлениях.

ИЕРАРХИЧЕСКАЯ МОДЕЛЬ (СТРУКТУРА) ДАННЫХ [hierarchical data model]

Модель организации данных, представляющая собой древовидный граф, состоящий из ряда типов записей (типов данных) и связей между ними (отношений или характеристик отношений), причем один из типов записей определяется как корневой, или входной, а остальные связаны с ним или друг с другом отношениями "один-ко-многим" или (реже) "один-к-одному". При этом запись, идентифицируемая элементом "один", рассматривается как исходная, а соответствующая элементу "многие", как порожденная. Каждая запись может быть порожденной только в одной связи, следовательно ей соответствует только одна исходная запись. Однако каждая запись может быть исходной во многих связях. Корневая запись может быть только исходной. Пример иерархической модели приведен на рис. 1.

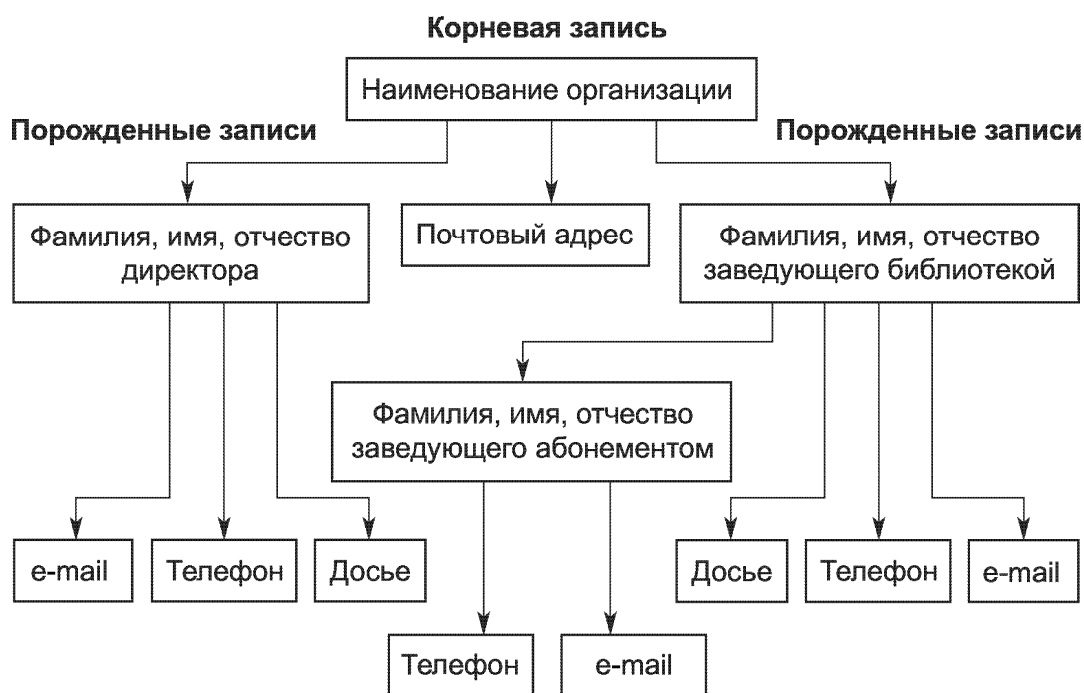


Рис. 1. Принцип построения иерархической модели организации БД

Граф [graph от греч. grapho - пишу, изображаю] - графическое представление математической модели системы связей между объектами любой природы. Объекты задаются в графе точками - **вершинами**, связи - линиями, соединяющими вершины, которые называются **ребрами** или **дугами** графа. Каждое ребро может быть ориентированным (т.е. иметь определенное направление от одной вершины к другой) либо неориентированным (двунаправленным). Ребро, соединяющее вершину с нею самой, называется **петлей**. Вершины, которым не соответствует ни одно ребро, являются **изолированными**. Число ребер, соединяющих две фиксированные вершины, может быть произвольным, поскольку оно определяется количеством и характером связей между соответствующими этим вершинам объектами. В то же время каждому ребру соответствует не более двух вершин. Раздел математики, изучающий свойства различных

геометрических схем, образованных множеством точек и соединяющих их линий, (графов), называется **теорией графов**.

СЕТЕВАЯ МОДЕЛЬ (СТРУКТУРА) ДАННЫХ [network data model]

Модель организации данных, подобная иерархической, но отличающаяся от нее тем, что каждая запись может вступать в любое количество поименованных связей с другими записями как исходная или порожденная, или как то и другое. Примеры 1 и 2 сетевых моделей приведены на рис. 2. и рис. 3.

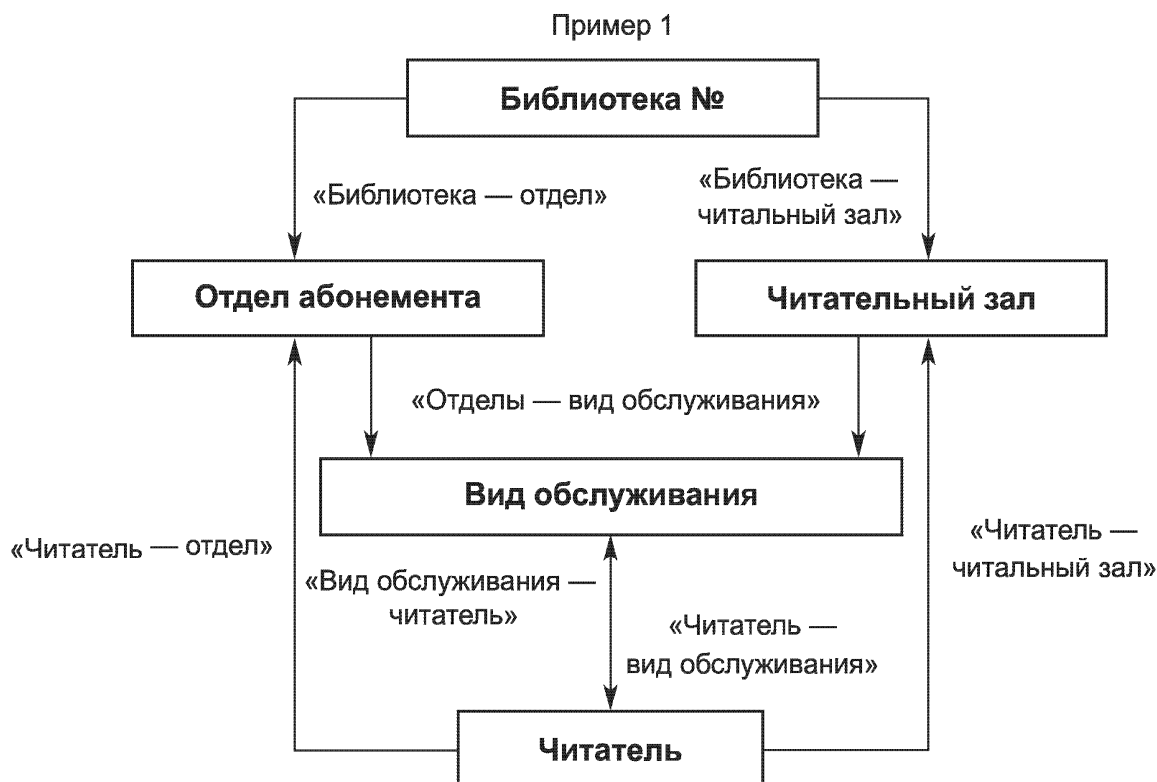


Рис. 2. Принцип построения сетевой модели организации БД

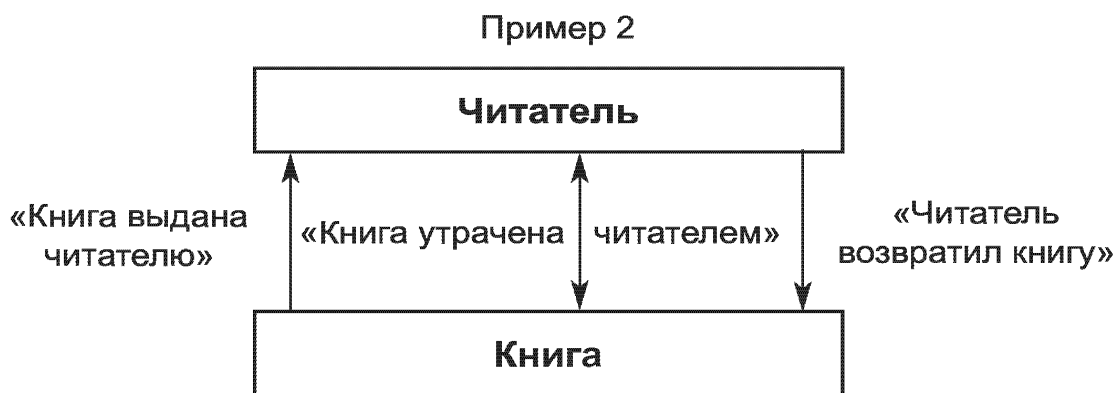


Рис. 3. Принцип построения сетевой модели организации БД

РЕЛЯЦИОННАЯ МОДЕЛЬ (СТРУКТУРА) ДАННЫХ

[relation data model]

Реляционная модель данных была предложена в 1969 г. сотрудником фирмы IBM **Е.Ф. Коддом**. Она представляет собой набор **плоских файлов** - таблиц, называемых отношениями, к которым применимы операции **реляционной алгебры** для реализации автоматизированного ответа на запросы пользователей системы. Весьма распространенной ошибкой является употребление термина реляционная БД по отношению к любым массивам данных, имеющих табличную форму организации, однако не обеспечивающим выполнение указанного условия.

Потенциально в реляционной модели может быть организовано очень большое количество связей между данными, значительная часть которых является избыточными (т.е. неиспользуемыми). Поэтому разработаны формы (варианты) нормализации отношений: первая (**1НФ**), вторая (**2НФ**), третья (**3НФ**) и четвертая (**4НФ**).

Примеры реализации реляционной модели приведены на рис. 4. и 5. В настоящее время существует достаточно большое число различных вариантов построения реляционных моделей. Одной из них является **постреляционная модель**.

Термины, связанные с реляционной моделью данных:

• Отношение [relation]

1. Форма связи между объектами (в рассматриваемом случае - разными типами данных или атрибутами), отражающая то общее, что их объединяет.
2. Два математических выражения, связанных знаком операции сравнения ("=", ">", "<" и т.п.).
3. Таблица реляционной модели данных.
4. Заданное подмножество n-ой декартовой степени некоторого множества.

Отношения в реляционной модели данных включают атрибуты и кортежи, составляющие соответственно столбцы и строки таблицы. Количество атрибутов в отношении соответствует количеству содержащихся в нем элементов данных, количество кортежей - числу экземпляров (реализаций) записей. Порядок следования атрибутов и кортежей может быть произвольным. Значения атрибутов определяются из доменов. Наличие доменов обеспечивает связи между разными отношениями.

• **Реляционная алгебра [relational algebra]** - язык для описания операций над отношениями. Основные операции реляционной алгебры: проекция, соединение, пересечение и объединение.

• **Сущность [entity]** - то же, что **тип объекта**: обобщенное наименование множества однотипных объектов, называемых **экземплярами**. Каждый экземпляр обладает набором свойств — **атрибутов**, отличающих его от всех остальных. Примеры сущностей: автомобиль, самолет, врач и т.д. Примеры экземпляров: "автомобиль ВАЗ-2106", "самолет ТУ-104", "врач Петров Иван Федорович" и т. д.

• Атрибут [attribute]

1. Признак, характеризующий **объект** или **сущность**, его свойства.

2. Описатель данных, содержащий одну из его характеристик (например, имя, тип, длину записи, форму представления, систему счисления и т. п.).

3. В реляционных моделях баз данных это столбец **отношения** (таблицы), содержащий реализации записей одного типа данных.

• **Кортеж [tuple]**

1. Запись группы взаимосвязанных элементов данных.

2. В реляционных моделях баз данных - строка отношения (таблицы), содержащая реализацию записей взаимосвязанных значений атрибутов.

3. Упорядоченный набор из N элементов.

• **Связь [глагол connection, interconnection]** - логическое отношение между сущностями (объектами). Связи обладают свойством, называемым **кардинальностью**. Например, "Заказчик может иметь 0, 1 или много заказов" (связь типа "0, 1 или много"), "Заказ содержит 1 или много товаров" (связь типа 1 или много"), "У автомобиля 4 колеса" (связь типа "равно 4"), "Билет резервируется для 0 или 1 пассажира" (связь типа "0 или 1"). Наиболее типичными являются связи типа "*один ко многим*" и "*многие к одному*".

• **Домен [domain]**

1. Область.

2. В реляционных моделях баз данных - область определения значений одного или нескольких **атрибутов**.

3. Крупная часть иерархической адресной структуры в службе **DNS** Интернета (например, правительственная или коммерческая организация, страна и т.п.);

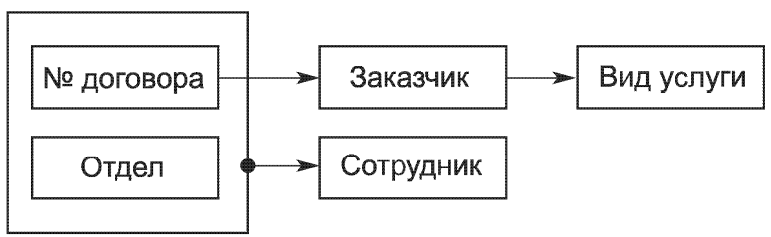
4. (мат.) Область определения функции.

• **Имя домена [domain name]** - полный адрес домена в Интернете, включает сам домен и наименование соответствующей ему организации.

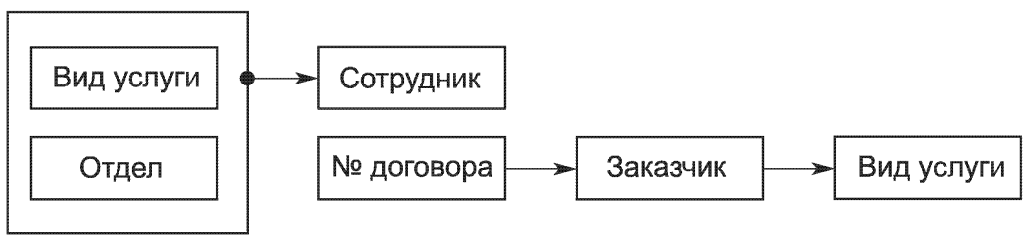
Отношение «Обслуживание заказчиков»

Отдел	Сотрудник	Заказчик	Вид услуги	№ договора

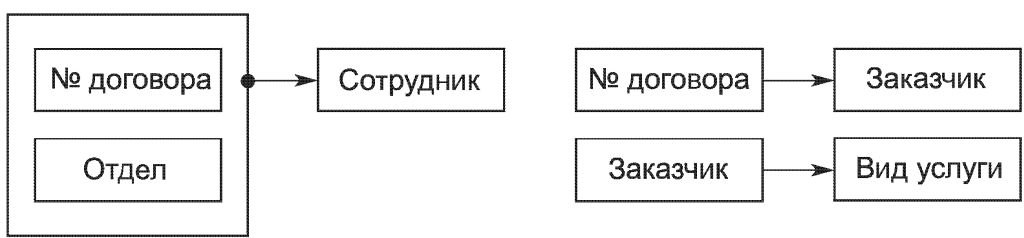
а



б



в



г

Рис. 4. Принцип построения реляционной модели (а) и связей внутри единственного отношения: б — первая нормальная форма (1НФ); в - вторая нормальная форма (2НФ); г - третья нормальная форма (3НФ)



Рис. 5. Принцип организации связей между разными отношениями в реляционной модели с использованием двух доменов: первый определяет область возможных значений атрибутов "Наименование фирмы" в отношении 1 и "Фирма изготовитель" в отношениях 2 и 3; второй определяет область возможных значений атрибутов "Наименование продукции и/или услуг" в отношении 1, "Марка оборудования" в отношении 2 и "Марка комплектующих" в отношении 3

ПОСТРЕЛЯЦИОННАЯ МОДЕЛЬ ДАННЫХ [postrelation data model]

Разновидность **реляционной модели**, позволяющая встраивать в **отношения** табличные записи реализаций отдельных атрибутов (так называемая многомерная табличная форма).

Записи в БД могут быть многомерными, т.е. включать несколько близких по содержанию **кортежей**, различающихся значением только одного или ограниченного числа полей (**атрибутов**). Например, в отношении "Владельцы информационно-программных продуктов (ИПП)" каждый конкретный "Владелец" может быть соотнесен с одним, двумя, тремя и т.д. разных ИПП. Поэтому использование одной из нормальных форм реляционной структуры БД может быть неэффективно из-за необходимости ввода и хранения значительного числа повторяющихся записей. Условиям более предпочтительной организации БД по данному признаку отвечает постреляционная модель данных **2НФ**, позволяющая создать компактную форму хранения путем использования "вложенных таблиц".

Указанным требованиям отвечает СУБД Universe фирмы Vmark Software (США). К концу 1993 г. этой фирмой продано около 14 тыс. лицензий на указанную программу для 430 тыс. конечных пользователей (среди клиентов фирмы Министерство безопасности Великобритании, Министерство

образования Австралии и др.). Популярность программы определяется возможностью строить на ее основе высокоэффективные информационные системы на недорогих платформах, обеспечивающих одновременный доступ к информации большого числа пользователей. Так, фирма Equifax (г. Атланта, шт. Джорджия, США) с помощью Universe смогла организовать работу более 775 пользователей на одном сервере Hewlett-Packard. В то же время, по выполненным подсчетам применение реляционной СУБД потребовало бы установки отдельной системы на каждые 150 рабочих мест. Другой пример реализации постреляционной модели - СУБД D3 Server 7.0 для операционных систем Windows NT фирмы Pick Systems.

В 1998 г. на российском рынке появилась постреляционная СУБД Cache фирмы Intersystems (<http://www.intersystems.ru>), отличающаяся более высокой производительностью по отношению к объектно-реляционным системам. В Cache используется специальный язык запросов - Cache SQL, позволяющий вместе со словарем данных создавать реляционные таблицы для сохранения данных. Cache Script позволяет сохранить скомпилированные запросы, а Cache ScriptObject обеспечивает возможность работы с данными как с кодом. Существует также возможность обращения к данным на языках Java и Си++.

ОБЪЕКТНАЯ МОДЕЛЬ [object model]

Модель, соответствующая структуре построения реальных **объектов** — их характеру, составу, состоянию и/или представлению о них, не связанная с какой-либо предопределенной схемой построения. Объекты можно сохранять и использовать непосредственно, не раскладывая их по таблицам, как это, например, имеет место в **реляционной модели**. Типы данных в объектной модели определяются разработчиком. Управление объектными БД осуществляется **объектными СУБД**, в которых данные об объектах и методах их описаний помещаются в хранилища как единое целое. Объектные модели используются в настоящее время в так называемых правовых информационных системах типа "Гарант", "Консультант" и др. Предполагается, что на их основе могут быть созданы консультационные экспертные системы.

ФРЕЙМОВАЯ МОДЕЛЬ [frame model]

Модель, основанная на использовании **фреймов** и их сетей. Является одним из способов представления **знаний**.

ФРЕЙМ [frame]

1. Единица представления знаний в искусственном интеллекте, описывающая понятие или объект. Фрейм состоит из ссылки на суперпонятие (т.е. понятие исходное или родовое) и описаний свойств, отличающих данный объект от суперпонятия.

2. Порция данных, передаваемая в сети передачи данных. Обычно кадр содержит набор служебных данных, включающих **адрес** устройства, которому он должен быть доставлен, адрес передающего устройства и др.

Кадры могут быть **unicast** - предназначенные одному устройству, **multicast** - предназначенные группе устройств или **broadcast** - предназначенные всем устройствам.

3. Кадр изображения.

4. Рамка - часть структурированного документа в интегрированных системах типа Framework, содержащая таблицу, график, текст или несколько вложенных рамок и отображаемая на экране дисплея в отдельном окне.

Framework - набор программ, разработанных для IBM PC автором dBase A. Тейтом. Он включает в себя: текстовый процессор, электронную таблицу, базу данных и генератор графических символов. Управление программами производится при помощи командного языка, в котором каждая команда начинается с символа @.

BizTalk Framework - архитектура, упрощающая обмен данными между приложениями масштаба предприятий и позволяющая вести переписку с торговыми партнерами и заказчиками через Интернет. BizTalk Framework определяет XML-схемы и набор XML-элементов, используемых в сообщениях, которыми обмениваются приложения. Фирма Microsoft выпустила набор средств BizTalk JumpStart Kit, позволяющих разработчикам быстро создавать BizTalk-совместимые решения с использованием схем, инструментальных средств, сервисов и образцов приложений XML.

Активность [activity] - состояние данных, записей, файлов, баз данных и др., когда к ним производится обращение из прикладных программ или запросов пользователей. При отсутствии обращений эти данные, записи, файлы или базы данных считаются неактивными.

Интерпретируемость (от англ. interpretation) - свойство знаний, отличающее их от данных возможностью получения с использованием ЭВМ четких ответов на любые вопросы относительно содержания объектов описания и отношений между ними, которые зафиксированы в базе знаний.

Другие термины, связанные с понятием "**модель данных**":

• **Бинарная модель [binary model]** - реляционная модель данных, в которой каждое отношение содержит по два атрибута.

• **Идеальная модель [ideal model]** - воображаемая модель, построенная на основе "идеальных" образов объектов **предметной области** с использованием чисто функциональной аналогии. Различают наглядно образные и знаковые идеальные модели.

• **Мифологическая (информационно-логическая, инфологическая) модель [information-logical model]** - модель предметной области, определяющая совокупность информационных объектов, их атрибутов и отношений между ними, динамику изменений **предметной области** и изменений информационных потребностей пользователей. Создается на основе **предпроектного обследования** предметной области с целью технико-экономического обоснования **банка данных** и составления технического задания на его проектирование.

• **Предметная (прикладная) область [application domain, knowledge domain]** - совокупность объектов, представляющих часть реального мира,

относящихся к ним понятий, а также связей между ними, сведения о которых обрабатываются и хранятся в базе данных автоматизированной системы.

- **Информационная модель [information model]**

1. То же, что модель данных.

2. Параметрическое представление процессов циркуляции информации (данных), подлежащей автоматизированной обработке в системе управления.

- **Имитационная модель [simulated model]** - алгоритм или программа, имитирующие функционирование системы.

- **Логико-лингвистическая модель [logical-linguistic model]** - модель знаний, в которой представление знаний основано на учете объектов предметной области, отношений между ними и использовании лингвистических средств. Примерами логико-лингвистических моделей являются семантические сети и сети фреймов.

- **Семантическая модель [semantic model]** - модель, представленная в виде графа, в вершинах которого расположены понятия, а дуги выражают отношения между понятиями. Примером семантической модели является семантическая сеть.

- **Семантическая сеть [semantic network]** - способ представления знаний в виде ориентированного графа, в котором вершины соответствуют семантическим (смысловым) единицам языка или речи (в том числе - понятиям, объектам, действиям, ситуациям и т.п.), а дуги - свойствам или отношениям между ними.

- **Семиотическая модель [semiotic model]**

1. Знаковая модель.

2. Модель, отражающая свойства знаковой системы.

- **Модель предметной области** - см. "Модель данных" и "Мифологическая модель".

- **Расширенная реляционная модель данных [extended relation model]** реляционная модель данных, в которую добавлен специальный **Е-домен** (домен сущностей), содержащий **суррогаты** всех доступных системе объектов предметной области. Каждый из суррогатов создается системой при первом упоминании объекта и находится под ее контролем.

- **Суррогат [surrogate]** - в реляционной модели данных: специальный домен, содержащий идентификаторы объектов в системе.

- **Структурная модель данных [structured data model]** - модель данных, представленная в виде структуры - множества типов данных и связей между ними. Различают **иерархические, сетевые и реляционные структуры**.

- **Модель Чена (сущностей и связей)[Chen (entity-relationship) model]** семантическая реляционная модель данных, в основе которой лежит деление реального мира на отдельные различимые сущности, находящиеся в определенных связях друг с другом, причем как сущность, так и связь полагаются первичными, неопределенными понятиями. Модель представляется в виде так называемой "**схемы E-R**" [**E-R-scheme**], на которой прямоугольниками отображаются имена типов сущностей, ромбами -

имена типов соотношений между сущностями, овалами — имена атрибутов.

• **Сеть фреймов [frame network]** - класс логико-лингвистических моделей, основанных на использовании **фреймов** для представления знаний о предметной области и позволяющих эффективно моделировать сложные и разнородные системы.

ГИПЕРТЕКСТ [hypertext]

1. Принцип организации информационно-поисковых массивов, при котором отдельные **информационные элементы** (в том числе **документографические, фактографические, полнотекстовые, графические** и др.) связаны между собой ассоциативными отношениями, обеспечивающими быстрый поиск необходимой информации и/или просмотр данных, взаимосвязанных указанными отношениями. Термин предложен в 1967 г. **Теодором Нельсоном**, представившим гипертекстовую структуру в виде **ориентированного графа**, в вершинах которого находятся записи разнородных видов данных (информационных элементов), а на ребрах - связи между ними. Средством организации связей в гипертексте и ориентации в нем является **навигатор**.

2. СУБД, реализующая гипертекст.

С гипертекстом связаны следующие понятия:

• **гиперсвязь [hyperlink]**

1. Вид связи между записями в документе или различными документами, реализующей **гипертекст**. Место связи выделяется каким-либо образом (например, цветом, шрифтом и т.д.); активизация гиперсвязи осуществляется при помощи мыши.

2. Программа, обеспечивающая гиперсвязь;

• **гиперссылка [hyperlink]** - слово или изображение в электронном документе, содержащие ссылку на другие файлы или части документа, на которые можно перейти щелчком "мыши" по гиперссылке. Гиперссылки в тексте, как правило, выделяются цветом;

• **навигация в базе данных [database navigation]** - процесс продвижения по логически связанным данным в иерархической и сетевой базах данных с целью поиска требуемых записей. Маршрут задается пользователем либо программными средствами СУБД;

• **автоматическая навигация [automatic navigation]** - выбор рационального пути (маршрута) поиска записей в базе данных, выполняемый программными средствами СУБД без участия человека;

• **навигатор [navigator]** - схема маршрутов, алгоритм или совокупность алгоритмов поиска записей в базе данных или в информационно-поисковых массивах.

4.4. ФОРМАТ, ПОЛЕ ДАННЫХ

ФОРМАТ [format, form, data format]

1. Совокупность правил записи и представления данных в памяти ЭВМ,

в базе данных, на экране монитора или на внешнем носителе (в том числе твердом). Основной структурной единицей формата является **элемент данных**, который записывается в поле данных. Формат определяет перечень **полей данных**, их характеристики, содержание вносимых данных и размещение. Некоторое отличие от представленного определения имеют специальные форматы, например, формат **SGIL**, в котором в качестве полей данных служат "абзацы", "главы" и "списки".

2. Элемент языка, в символическом виде описывающий представление информационных объектов в записи (в том числе в файле, базе данных и т.п.).

3. Способ кодирования записи двоичной информации (например, **текстовый формат [ASCII format]** - представление текстовой информации в коде ASCII).

4. Способ разбиения поверхности носителя, например, магнитного диска, - **формат диска [disk format]** на адресуемые элементы (дорожки и сектора).

5. Характеристика линейных размеров (ширины и высоты) печатного издания, листа, полосы набора, текстовой полосы и т.п. (например, **формат издания [format (size) of publication]**) и/или порядок размещения реквизитов (например, **формат документа**).

В зависимости от характера информационного объекта, его структуры или способа записи различают:

- **формат адреса [address format]** - структура адресной части команды, определяющая элементы и способ формирования действительного адреса;

- **формат данных [data format]** - способ представления данных в памяти ЭВМ и вне ее;

- **формат документа [document format]** - структура документа, в том числе порядок расположения реквизитов и его линейные размеры;

- **формат записи [record format]** - способность записи изменять свою длину ;

- **формат файла [file format]** - способ организации записи данных на машиночитаемом носителе в файле. Вид формата файла принято указывать в расширении его имени, например "*.doc", "*.txt", "*.tif", "*.jpg" и т. п.;

- **формат пакета [packet format]** - в вычислительных сетях: формат, определяющий размеры, размещение и содержание полей, входящих в состав пакета. Формат пакета регламентируется протоколом управления соответствующей сети передачи данных;

- **переменный формат [variable format]** - формат записи переменной длины, при котором длина записи указана в ее начале;

- **текстовый формат [ASCII format]** - представление текстовой информации в коде ASCII;

- **двоичный формат [binary format]** - формат, в котором байт представляет двоичную (по основанию 2), а не десятичную величину.

ФОРМАТИРОВАНИЕ, ИНИЦИАЛИЗАЦИЯ, РАЗМЕТКА (диска)
[formatting, format, initialization]

Процедура записи на магнитный диск меток, определяющих последующее расположение записей данных (блоков, секторов, дорожек), участков, не пригодных для записи, а также другой управляющей информации. Форматирование выполняется перед первым использованием диска.

Форматирование текста [text formatting] - преобразование текста в вид, в котором он должен выводиться на экран монитора или печать: установка границ страницы, формирование абзацев, центрирование заголовков, выравнивание полей, разбиение на страницы, установка колонтитулов и т. п.

Форматировать [formatting]

1. Размещать данные в соответствии с предписанным форматом.
2. Производить разметку или форматирование диска.

Переформатирование [reformatting]

1. Изменение формата данных.
2. Повторное форматирование магнитного диска.

RTF (Rich Text Format) - пакет соглашений, разработанных фирмой **Microsoft** для включения форматированных данных в текстовый документ независимо от используемой аппаратуры или типа системы обработки текстов.

ПОЛЕ ДАННЫХ, ПД, ПОЛЕ [data field, field]

1. Область на носителе информации (машино- или человекочитаемом), выделенная для записи определенного вида данных (в том числе информационных элементов, данных, элементов данных).

2. Часть записи или заполняемой формы, имеющая функционально самостоятельное значение и обрабатываемая как отдельный элемент данных.

Поле данных идентифицируется "*внешним*" и "*внутренним*" именами и характеризуется "*видом данные*", "*длиной*" и "*структурой*" записи. Поля данных могут быть "*элементарными*", "*множественными*" и "*групповыми*", а также "*повторяющимися*" (в одной записи) и "*неповторяющимися*".

Виды полей данных и связанные с ними термины:

• **Элементарное ПД [elementary (simple) data field]** - минимальная структурная единица физической записи, не содержащая в своем составе других полей или подполей данных.

• **Подполе данных [subfield]** - поле данных, которое является составной частью множественного или группового ПД.

• **Групповое ПД [group data field]** - поле данных, включающее в себя два и более элементарных, других групповых и/или множественных полей. Служит средством объединения нескольких полей, которые можно выделить (идентифицировать), используя имя группового поля данных соответствующего уровня иерархии. В библиографических и полнотекстовых БД примером группового поля является "Библиографическое описание".

• **Множественное ПД [multiple data field]** - поле данных, содержащее в

своем составе несколько элементарных полей и подполей с записями однотипных данных (элементов данных), каждая из которых может быть выделена средствами системы как самостоятельная. В фактографических базах данных групповые поля используются для записи взаимосвязанных значений, например, "длины", "высоты" и "ширины" объектов в соответствующих подполях поля — "Габариты объекта".

- **Повторяющееся ПД [recurring data field]** - поле данных, с кратностью повторения выше единицы. Повторяющимися могут быть элементарные, множественные и групповые поля данных. Примером повторяющегося ПД в международном библиографическом формате MARC21 является поле "Другие авторы".

- **Длина ПД [data field length]** - количество символов (байт), разрешенных для записи в поле данных. По указанному признаку поля данных могут быть фиксированной, переменной и неопределенной длины.

- **ПД фиксированной длины [fixed-length data field]** - поле данных, длина которого не изменяется.

- **ПД переменной длины [variable-length (data) field]** - поле данных, длина записи в котором может содержать произвольное число символов (байт).

- **ПД неопределенной длины [undefined-length (data) field]** - поле данных, для которого резервируется память, рассчитанная на максимальную длину, а реальная его длина определяется длиной конкретной реализации записи.

- **Структура (записи) поля данных [structure data field]** - принцип упорядочения записей в поле данных. По указанному признаку различают следующие виды полей данных: текстовые, нормированные, форматированные и числовые; говорят также: "Поле данных с текстовой (числовой и т. п.) структурой или формой записи".

- **Текстовое ПД [alphanumeric/character data field]** - поле данных, запись в котором ограничивается только длиной (другими словами, она не регламентируется порядком или правилами набора буквенно-цифровых символов, составляющих ее содержание).

- **Числовое ПД [numeric(al) data field]** - поле данных, содержащее записи только количественных значений измеряемых величин (например, денежных сумм, показателей веса, геометрических размеров - длина, высота или ширина и т.п., над которыми можно производить арифметические операции).

- **Нормированное ПД [normalized data field]** - поле данных, в которое производится запись только разрешенных терминов, контролируемых машинными словарями автоматизированной системы. Нормированными объявляются поля данных, содержащие сведения, наиболее важные для идентификации (при поиске или машинной обработке) документов и данных, например, наименований стран, фирм, организаций, "марок" изделий и т.п.

- **Форматированное ПД [formatted data field]** - поле данных, с однозначно установленной последовательностью записи буквенно-цифровых

символов. Примеры полей с форматированной записью: "дата" - мм-дд-гг (в конкретной реализации записи: 02-23-99 означает 23 февраля 1990 года): "время" - чч:мм:сс (в конкретной реализации записи: 12:34:45, что означает 12 час. 34 мин. 45 с.) и т. п.

Форматированные поля данных служат средством обеспечения сортировки и поиска данных по содержащимся в них сведениям. Для решения ряда задач некоторые форматированные ПД могут быть одновременно объявлены также нормированными. Так, для контроля технологической дисциплины обработки документальных потоков с установленными сроками прохождения разных видов документов по звеньям технологической цепи (например, "заказ книги", "поступление в отд. комплектования", "поступление в отд. обработки", "поступление в отд. абонемента" и т.п.) соответствующие технологические поля могут контролироваться обязательным вводом текущей (никакой иной) даты и времени обработки и/или передачи документов;

- **Внешнее имя ПД [data field external name]** - наименование типа данных или обобщенное наименование характера информационного элемента, записи реализаций которых производятся в полях с указанным именем. Внешнее имя поля данных является средством его идентификации для пользователей автоматизированной системы и реализуется в рабочих листах при вводе, обработке и/или поиске данных.

- **Внутреннее имя ПД, внутренняя метка, имя данных [data field internal name/label]** - короткий символьный код, открывающий запись в поле данных и хранимый вместе с данными, которые он идентифицирует. Внутреннее имя поля данных является средством его идентификации для программных средств автоматизированной системы.

- **Метка конца поля (данных) [field marc]** - ограничитель поля, символ конца записи в поле данных, например "@", "!".

5. ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ

5.1. ОБЩИЕ ПОНЯТИЯ И ТЕРМИНЫ

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ, ИТ [information technology, IT]

Комплекс методов, способов и средств, обеспечивающих создание, хранение, обработку, передачу, защиту и отображение информации, ориентированных на повышение эффективности и производительности труда.

Существуют и другие определения ИТ, например:

1. **ИТ** - термин, относящийся ко всем технологическим средствам, используемым для создания, хранения, обмена и использования информации в ее различных формах (деловые данные, телефонные переговоры, фотографии, видеозаписи, мультимедийные представления, а также какие-то иные, которые могут появиться в будущем);

2. **Информационные технологии или информационные и**

коммуникативные технологии, ИКТ [Information & communication technology, ICT] - это технологии, применяемые для обработки информации. В частности, они используют компьютеры и программное обеспечение для преобразования, хранения, защиты, передачи и извлечения информации в любом месте и в любое время.

Информационная технология является неременной составной частью большинства видов интеллектуальной, управленческой и производственной деятельности человека и общества. Развитие ИТ в современных условиях основано на применении вычислительной техники и связанных с ней методов и средств автоматизации информационных процессов. В зависимости от степени использования этих средств, ИТ иногда условно разделяют на традиционную ИТ и современную ИТ.

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ, ИО **[information provision (ensuring)]**

1. ГОСТ 7.27-80 трактует "ИО" как "Совокупность процессов по подготовке и предоставлению специально подготовленной научно-технической информации (НТИ) для решения управленческих и научно-технических задач в соответствии с этапами их решения".

Учитывая современные тенденции, понятие "ИО" можно определить более широко, как подсовокупность процессов по подготовке и предоставлению специально подготовленной информации для решения управленческих, научных, технических, производственных, коммерческих и других задач в соответствии с этапами их решения.

Важными отличительными признаками ИО являются:

- избирательный характер содержания информации, а также видов работ, связанных с ее подготовкой, формой и сроками представления, ориентированный не только на конкретного потребителя, но и на порядок выполнения им определенных этапов задач (в том числе связанных или смежных проблем, работ, заданий и т.п.);
- активный и упреждающий характер выполняемых информационных работ вне зависимости от того, были ранее точно сформулированы возможные частные запросы пользователя на информацию или нет.

Другими словами, ИО рассматривается как индивидуальное непрерывное сопровождение и поддержка "решения задач пользователя" комплексом информационных услуг и продукции, обеспечивающим успешное решение этих задач.

ИНФОРМАЦИОННАЯ ПОТРЕБНОСТЬ **[information demand, information need, information requirement]**

Состояние отдельного лица, коллектива или системы, характеризующееся необходимостью получения информации для успешного достижения каких-либо целей или выполнения работы. Специалисты в области информатики различают: истинную, или объективную

информационную потребность, определяющую действительно необходимый в каждом конкретном случае состав и объем сведений или данных; и субъективную, характеризующуюся представлением субъектов об информации, которая им необходима. Различия между указанными видами информационных потребностей определяются профессиональными и личностными характеристиками конкретных пользователей информации. В практике работы информационных органов и библиотек учет этих разновидностей информационных потребностей имеет чрезвычайно важное значение.

ИНФОРМАЦИОННЫЙ ЗАПРОС [information request (inquiry), question]

Выраженная на естественном языке информационная потребность.

Запрос [query, request, interrogation] — в автоматизированной системе: входное сообщение в автоматизированную систему, содержащее требование на выдачу информации или (в более общем случае) на выделение ресурсов. Следует отметить, что запрос может иметь и побочную информационную роль, поскольку содержит сведения об информационных потребностях пользователей, о самих пользователях, режимах их обращения к системе и др. данные, которые могут быть объектами обработки и накопления в базах данных ЭВМ.

5.2. ОБРАБОТКА И ПЕРЕРАБОТКА ДОКУМЕНТОВ И ДАННЫХ

ОБРАБОТКА [processing]

Действие по глаголу "обработать":

1. Подвергнуть выделке, отделке, какой-нибудь производственной операции.
2. Отделать, привести в законченный вид, сделать более совершенным.
3. Сделать готовым для чего-нибудь.

ПЕРЕРАБОТКА [conversion]

1. Производственный процесс, которому подвергается сырье (в нашем случае — исходные документы, данные и т. п.).
2. Действие по глаголу "переработать": превратить во что-нибудь в процессе работы, обработки (переработка сырья в изделие).
3. Переделать, сделать по-новому, иначе.

Из приведенных значений терминов "*обработка*" и "*переработка*" следует, что в применении к информационным процессам понятие "*переработка*" предполагает внесение существенных преобразований содержания и/или формы исходных информационных продуктов, выполняемых на основе анализа и синтеза содержащихся в них данных. Примерами "*переработки*" могут служить: подготовка нового вида документа (в том числе - реферата, обзора, поискового образа документа и

т.п.), его перевод на другой язык, преобразование данных (например, из текстовой формы в табличную, в том числе их формализация) и др. Процессы "обработки" ассоциируются с "выделкой", "отделкой", "совершенствованием" и т.п. Они составляют производственные операции, которые, не изменяя содержания и формы информационного продукта, делают его пригодным или готовым "для чего-нибудь", например, типографского издания, автоматизированной или машинной обработки, а может быть — и переработки. Существенным отличием обработки документов и данных от их переработки может служить отсутствие в ее процессах использования методов индукции, дедукции и синтеза.

Термины, связанные с процессами обработки и переработки документов и данным:

• **Анализ [analysis]**

1. Метод научного исследования, состоящий в расчленении целого на составные элементы.

2. Разбор, рассмотрение чего-либо.

• **Синтез [synthesis]**

1. Метод исследования какого-либо явления в его единстве и взаимной связи частей.

2. Обобщение, сведение в единое целое данных, полученных в результате анализа.

• **Индукция [induction]**

1. Логический метод исследования, основанный на переходе от частных, одиночных случаев (фактов) к общему выводу ("от частного к общему").

2. Форма мышления, в соответствии с которой мысль наводится на какое-либо общее правило или общее положение, присущее единичным объектам.

• **Дедукция [deduction]**

1. Логический метод исследования, заключающийся в переходе от знания общих положений к частным суждениям или другим общим выводам ("от общего к частному").

2. Форма мышления, в соответствии с которой новое знание выводится чисто логическим путем из предшествующих знаний или мыслей. Дедукция неразрывно связана с индукцией.

• **Формализация [formalization]**

1. Процесс представления информации об объектах (в том числе процессах), явлениях реального мира и мыслительной деятельности человека в формализованном виде (форме). В автоматизированных информационных системах вид формализации задается технологическими инструкциями и другими средствами, например, так называемыми рабочими листами.

2. Метод исследования, основанный на использовании формализованных математических, логических и т.п. систем, в которых вместо естественного языка используется язык специальных символов.

• **Формализация данных [data formalization]** - выбор, разработка и/или реализация заполнения форм входных документов для ввода данных (см.

далее) в систему. Наиболее распространенными формами входных документов в информационных и библиотечных системах являются анкетная и табличная.

- **Цикл обработки [processing cycle]**

1. Постоянно повторяющаяся последовательность операций по обработке документов и/или данных.

2. Последовательность шагов, многократно выполняемых ЭВМ, в процессе работы программы. Центральный процессор ЭВМ непрерывно действует в циклическом режиме, включающем выбор текущей команды программы из памяти, ее исполнение и внесение результатов в память перед обращением к последующей команде нового цикла.

- **Параллельная обработка [parallel processing]** - в автоматизированных системах: одновременное выполнение нескольких программ или разных частей одной программы при работе с данными.

- **Сортировка [sorting]** - упорядочение данных (записей, файлов и т.п.) по какому-либо признаку или группе признаков. Последние выбираются из состава полей данных, связанных с объектом сортировки. Например, сортировка сведений об организациях или фирмах может производиться по признакам: область деятельности (отрасль), страна, город, наименование фирмы и т.д., если существуют заполненные записями соответствующие поля данных, связанные между собой определенными отношениями. Поля данных, содержащие признаки сортировки, называются ключевыми полями.

- **Интерактивный режим, диалоговый режим, онлайн режим [interacting mode, on-line]** - режим работы, при котором взаимодействие человека с ЭВМ (локальной или в вычислительной сети) происходит путем непосредственного и двухстороннего обмена данными, командами или инструкциями между человеком и ЭВМ в "реальном масштабе времени". Режим диалога построен в форме вопросов и ответов, является основным для персонала и пользователей автоматизированных информационных систем всех видов.

- **Цифровая обработка сигнала, ЦОС [Digital Signal Processing, DSP]** - автоматическая или автоматизированная обработка данных, представленных в виде электрических сигналов в двоичном (цифровом) коде.

- **Каталогизация [cataloging]** - вид аналитико-синтетической переработки документов, представляющий собой совокупность процессов, обеспечивающих создание и функционирование библиотечных каталогов, включая и электронные каталоги. Каталогизация включает составление библиографического описания и библиографической записи (в том числе индексирование документов, их систематизацию и предметизацию), организацию и редактирование библиотечных каталогов.

- **Библиографическое описание, БО [bibliographic description]** - совокупность библиографических сведений о документе, приведенных по определенным правилам, устанавливающим наполнение и порядок следования областей и элементов, и предназначенных для идентификации и общей характеристики документа. Объектами БО являются все виды

опубликованных и неопубликованных документов на любых носителях, в том числе: книги, сериальные и другие продолжающиеся ресурсы, нотные, картографические, аудиовизуальные, изобразительные, нормативные и технические документы, микроформы, электронные ресурсы, другие трехмерные искусственные или естественные объекты; составные части документов; группы однородных и разнородных документов.

- **Библиографическая запись, БЗ [bibliographic record]** - включает библиографическое описание, заголовок библиографической записи, термины индексирования, аннотацию или реферат, шифры хранения документа, справки о добавочных БЗ, дату завершения обработки документа, сведения служебного характера.

- **Аннотация [annotation]** - краткая характеристика первичного документа, его части или совокупности документов с точки зрения содержания, назначения, формы и других особенностей. Аннотация носит пояснительный или рекомендательный характер и служит в качестве средства информационного поиска документов.

- **Реферат [abstract]** - краткое изложение содержания первичного документа, его части или совокупности документов, включающее основные сведения и выводы, а также количественные и качественные данные об объектах описания. Реферат предназначен для непосредственного обеспечения пользователей краткой информацией без обращения к первоисточнику, а также для выполнения информационного поиска документов и данных.

- **Библиотечный каталог [library catalogue]** - перечень произведений печати и других документов, имеющих в фонде библиотеки или группы библиотек, выполненный по определенному принципу и раскрывающий состав и/или содержание библиотечных фондов.

ОБРАБОТКА ДОКУМЕНТОВ И ДАННЫХ [data and document processing]

Обобщенное наименование разнородных процессов, связанных с документами и данными. Термин нашел преимущественное применение в контексте вычислительной техники и разного рода автоматизированных системах (информационных, библиотечных, управленческих и др.) и, как правило, относится к рутинным операциям обработки и хранения больших массивов документов и данных. Примерами могут служить процессы и/или отдельные операции форматирования, переформатирования, редактирования (технического), корректуры, измерения, копирования (полного или частичного) записи, перезаписи (переписывания), чтения, сортировки, кодирования и др., выполняемые как человеком, так и автоматизированными средствами.

АВТОМАТИЗИРОВАННАЯ ОБРАБОТКА ДОКУМЕНТОВ И ДАННЫХ [automated data processing]

Выполнение комплекса операций над документами и данными с

помощью ЭВМ. Автоматизированная обработка данных является неотъемлемой частью современной информационной технологии.

В автоматизированных информационных системах процесс обработки может быть условно разделен на два этапа:

1. Получение исходных данных и их первичное преобразование (первичная обработка);

2. Подготовка выходных результатов (вторичная обработка).

Виды автоматизированной обработки данных:

- **Первичная обработка [preprocessing]** - включает операции сбора данных, их первичного учета, индексирования, ввода, перезаписи в формы (формат), удобные для выполнения машинных операций (форматы: предсистемный, рабочий, хранения, поиска), проверку полноты и точности записи данных и их соответствия определенным форматам или правилам представления, проверку на дубль.

- **Вторичная обработка [reprocessing]** - включает внутренние преобразования форматов данных (например, из формата хранения в формат поиска, коммуникативный формат и т.п.), поиск данных, их сортировку, группировку и перегруппировку, редактирование и/или преобразование полученных данных, подготовку и заполнение выходных форм. Частными операциями автоматизированной обработки данных являются также обработка текста, бесклавиатурный ввод и обработка изображения документа.

- **Обработка текста [word processing]** - все виды операций над текстовыми материалами, выполняемые с использованием ЭВМ, включая клавиатурный и бесклавиатурный ввод, редактирование, форматные преобразования, вывод на печать или экран, копирование, хранение, пересылку и др.

- **Обработка изображения документа [DIP, document image processing]** - ввод в ЭВМ документа путем сканирования, т.е. считывания его изображения с использованием сканера, и последующая обработка полученной записи с целью получения требуемого формата ее представления и качества.

- **Диалоговая (интерактивная) обработка [interactive processing]** - обработка данных (в том числе текстовых, табличных и др. материалов), выполняемая с использованием диалогового ввода данных в ЭВМ; является основным режимом работы автоматизированных систем.

- **Пакетная обработка (запросов/данных), пакетный режим (обработки данных) [batch processing]** - обработка, выполняемая путем объединения соответствующих материалов в пакет и его передачи в ЭВМ в виде задания. Пользователь не может влиять на результаты работы ЭВМ до завершения полного цикла обработки задания. Данный режим является эффективным для обработки хорошо отработанных персоналом автоматизированной системы задач, требующих значительных затрат машинного времени на их выполнение.

РАБОЧИЙ ЛИСТ, РЛ [worksheet]

Формализованная анкета, предназначенная для обработки и записи структурированных данных.

В общем случае РЛ содержит состав полей данных, соответствующих виду обрабатываемых документов или данных, и некоторый набор сведений об их содержании и правилах заполнения (например, внутреннее и внешнее имя, длину, структуру записи и т. п.). РЛ может быть реализован на твердом носителе — в виде бумажных бланков, пригодных для заполнения сведениями при так называемой предмашинной обработке документов и данных, а также в виде специально размеченной структуры экрана монитора для клавиатурного ввода.

5.3. ВВОД ДОКУМЕНТОВ И ДАННЫХ В ЭВМ

ВВОД ДОКУМЕНТОВ И ДАННЫХ (в ЭВМ, систему) [input, data (information) input, entry]

Комплекс операций, обеспечивающий считывание и передачу данных от внешнего их носителя в основную память ЭВМ.

В зависимости от способа выполнения этих операций различают:

- **автоматический ввод [automatic input]** - метод ввода, в соответствии с которым считывание данных с внешнего носителя и запись их в память ЭВМ осуществляется под управлением специальной программы, которая также настраивает и запускает остальную часть системы;

- **бесклавиатурный ввод [nonkeyboarding input]** - автоматизированный ввод данных в ЭВМ без использования клавиатурной работы;

- **ручной ввод [manual input]** - наименее автоматизированный метод ввода данных с использованием операций клавиатурного или бесклавиатурного ввода (например, при помощи устройств типа планшет);

- **клавиатурный ввод [keyboard input]** - ввод (данных) с клавиатуры (разновидность ручного ввода). Технологическим средством обеспечения клавиатурного ввода является рабочий лист;

- **оптический ввод, сканирование [optical entry, scanning]** - ввод с использованием оптических средств считывания данных (например, сканеров) и поддерживающих их работу программных средств распознавания символов (образов);

- **графический ввод [graphic input]:**

- 1) ввод данных в графической форме (в том числе оптический ввод);

- 2) ввод графических изображений;

- **речевой/голосовой ввод [speech/voice input]** - ввод данных в ЭВМ с голоса человека с использованием программно-технических средств, обеспечивающих перевод аудиоинформации в текстовую форму и запись ее в память ЭВМ. Речевой ввод является одним из важных сервисных средств автоматизации диалога "человек-машина" и развития ЭВМ пятого поколения, начало промышленного выпуска которых может быть отнесено

уже к настоящему времени;

- **распознавание речи, автоматическое распознавание речи [speech recognition]** - процесс и технология, связанные с вводом данных в ЭВМ голосом. Процесс распознавания речи состоит из этапов записи звуковых сигналов, их оцифровки и декодирования. При оцифровке звуковые сигналы разбиваются на фрагменты. Декодирование осуществляется путем сопоставления принятых сигналов с "эталонными" записями с использованием различных моделей и алгоритмов. Последние могут использовать образцы целых слов или их фрагментов (фонем). Считается, что для отображения всех слов любого языка достаточно от 40 до 60 фонем. Наиболее точные с точки зрения распознавания - модели, основанные на распознавании целых слов. Однако более универсальными и гибкими являются модели, основанные на фонемной структуре. При анализе речи, ее записи и последующем воспроизведении "голосом" ЭВМ пятого поколения могут учитывать также особенности тембра говорящего человека. Для этого анализируется и кодируется спектральный состав речи пользователей. Указанное качество может использоваться для ограничения доступа к данным или другим охраняемым объектам;

- **распознавание речи без настройки на пользователя [speaker-independent voice recognition]** - технология преобразования речи в текстовую форму. Обычно используется в системах голосовой электронной почты для обеспечения взаимодействия с голосовыми системами пользователей, имеющих телефоны с импульсным набором номера абонента (в отличие от тонового DTMF (Dual Tone Multi Frequency), при котором цифры набранного номера кодируются сочетанием двух звуковых тонов). Системы распознавания голоса без настройки на пользователя могут работать с весьма ограниченными словарями, например, для произнесения номера абонента вместо набора его цифрами на клавиатуре;

- **диалог [dialog]** - двухсторонний непосредственный обмен информацией, командами или инструкциями между человеком и ЭВМ; основной режим диалога построен в форме вопросов и ответов;

- **диалоговый (интерактивный) ввод [conversational/interactive entry]** - режим ввода данных (с клавиатуры или голосом) в процессе диалога пользователя с системой.

- **прямой ввод** - технологический способ выполнения так называемой безбумажной аналитико-синтетической обработки и переработки первичных документов, завершаемый непосредственным вводом полученных (вторичных) документов без промежуточной их записи на бумажный носитель;

- **поэтапный ввод** - технологический способ выполнения аналитико-синтетической переработки и предмашинной обработки документов, в соответствии с которым различные части итоговых документов (например, библиографического описания, реферата, аннотации, ПОДа и т. п.) готовятся и вводятся по частям разными сотрудниками или группами сотрудников с использованием промежуточных носителей или без них;

- **вход в систему [login]** - процедура, обеспечивающая доступ к системе и регистрацию известных ей, т.е. зарегистрированных пользователей;
- **авторизованный вход [authorized admittance]** - процедура, обеспечивающая защиту входа в систему паролем;
- **ввод-вывод (данных) [I/O, Input-Output]** - совокупность начальных и конечных операций, связанных с обработкой данных средствами вычислительной техники (ЭВМ);
- **выход из системы [log off, log out]** - процедура, при помощи которой пользователь прекращает работу с системой и позволяет ей зарегистрировать это действие.

ВВОДИМЫЕ ДАННЫЕ [input data, input]

Исходные данные, предназначенные для ввода в память ЭВМ, обработки и/или записи их на машиночитаемые носители.

Входной файл [Input file] - файл, содержащий вводимые данные.

Входной поток [Input stream] - последовательность документов и данных, поступающих для ввода в АИС. Документы, составляющие входной поток, подразделяются на информационные документы, предназначенные для формирования баз данных или других информационных массивов, и на запросы на поиск информации в системе.

ЗАПИСЬ ДАННЫХ [data record]

1. Единица структуры данных, зарегистрированная на каком-либо носителе.
2. Машинная операция, состоящая в переводе данных из основной (оперативной) памяти ЭВМ во внешнюю память (например, жесткий магнитный диск, дискету и т.п.). Единицей обмена данными при этом является физическая запись.

ПЕРЕМЕННАЯ [variable]

1. Запись, содержащая сведения о типе или типах информационных элементов, наименование и значение которых может быть получено или изменено с использованием программных средств (программным способом). Простейшим примером переменной может служить поле переменной длины. Переменная характеризуется именем, адресом, типом и описанием.

2. Число или другой элемент программы, имеющий имя и значение, которое может быть выбрано и изменено программой или программистом.

Виды переменных и связанные с ними термины:

- **глобальная переменная [global variable]** - переменная (в программе), доступная для любой команды программы;
- **локальная переменная [local variable]** - переменная (в программе), доступ к которой может осуществляться только при использовании команд в пределах определенной части программы (подпрограммы).
- **имя переменной [variable name]** - короткий символьный код, идентифицирующий переменную;

- **адрес переменной [variable address]** - указатель места в области памяти, в которой хранится запись переменной;
- **тип переменной [type variable]** - обозначение назначения или области распространения ее действия. Существует достаточно большое количество различных типов переменной, например, "метка", "файл", "указатель", "область" и др.;
- **описание переменной [variable declaration]** - определение типа переменной, размера (т.е. длины записи) и структуры записи, идентификатора и других характеристик;
- **поле признака [tag field, variant record]** - подполе вариантной записи, определяющее совокупность имен и типов остальных ее компонентов.

ПАРАМЕТР [parameter]

Переменная, которой в рамках конкретного применения присваивается определенное значение. В программировании при помощи указанного понятия производится описание типизированных объектов, жестко не привязываясь к их конкретному виду. Заданием параметров можно воссоздать необходимую реализацию этих объектов; например, для прямоугольника - длины, толщины и цвета его линий, что позволяет нарисовать любой прямоугольник. В программах текстовых редакторов в качестве одного из параметров используется вид шрифта и т. п.

Различают виды параметров, в частности:

- **динамический параметр [dynamic parameter]** - параметр, значение которого устанавливается во время выполнения программы;
- **заданный параметр [present parameter]** - параметр, значение которого заранее установлено (например, в программе) и изменению не подлежит;
- **ключевой параметр [keyword parameter]** - параметр, значение которого задается с помощью ключевого слова;
- **позиционный параметр [positional parameter]** - параметр, занимающий определенное положение (позицию) в списке параметров. Изменение позиции таких параметров не допускается;
- **необязательный параметр [optional parameter]** - параметр, значение которого соответствует стандартному, и поэтому его можно не указывать.

ВЕРИФИКАЦИЯ [verification]

Проверка, установление подлинности или истинности чего-либо путем сопоставления предъявленных признаков объекта по присвоенному ему коду или имени с записями этих признаков, хранящимися в памяти компьютера. Например, верификация записей производится программными средствами ЭВМ по вызову имени соответствующей главной записи. При автоматизированном распознавании речи верифицируемый пользователь предварительно заявляет свой код, а затем произносит пароль или произвольную фразу, что позволяет автоматизированной системе произвести проверку соответствия этому коду (паролу) голоса эталонной записи,

хранящейся в памяти компьютера.

ИДЕНТИФИКАЦИЯ [identification]

Определение личности или установление соответствия объекта идентификации чему-либо путем сопоставления предъявленных признаков с эталонными записями, хранящимися в памяти компьютера. Например, в системах автоматизированного распознавания речи идентификация голоса производится без предварительного заявления о пользователе путем сравнения его голоса со всеми эталонными записями голосов, по которым система определяет, кому он принадлежит. Идентификация записей или групп записей (файлов, групп файлов) выражается путем сопоставления предъявленных пользователями их наименований с каждым из уникальных или отличных от других имен, находящихся в памяти компьютера, что позволяет автоматизированным системам находить их в массивах других записей (файлов). Одним из активно развиваемых и используемых особенно в последние годы автоматизированных методов является идентификация личности по отпечаткам пальцев.

RFID (Radio Frequency Identification)^ "Радиочастотная идентификация"

Технология идентификации чего-либо (люди, животные, товар, книги в библиотеке, заготовки деталей на производстве и т.п.), предусматривающая использование радиосчитывающего устройства и электронного ярлыка в виде недорогого чипа, несущего уникальные идентификационные данные. Используется в высокоавтоматизированных системах контроля. Одна из реализаций радиочастотной идентификации - RFID-Pallet: технология использования RFID на автоматизированном складе, которая быстро и эффективно учитывает содержимое склада без участия человека.

ХЭШ-СУММА, КОНТРОЛЬНАЯ СУММА [hash total, hashsum]

Метод контроля правильности записи при вводе наборов цифровых данных, например, платежных ведомостей и счетов, при котором происходит автоматический подсчет и сопоставление вводимых данных с контрольной суммой.

АКТУАЛИЗАЦИЯ (данных) [(data) updating]

Поддержание данных в актуальном состоянии, т. е. приведение их в соответствие с состоянием отображаемых объектов предметной области. Актуализация включает в себя операции добавления, исключения, а также редактирования (в том числе правки или исправления) записей.

5.4. ИНФОРМАЦИОННЫЙ ПОИСК

ПОИСК [search]

Совокупность операций, связанных с определением местонахождения объектов с заданными характеристиками или признаками.

ИНФОРМАЦИОННЫЙ ПОИСК, ПОИСК ИНФОРМАЦИИ [information retrieval, data search]

Процесс нахождения, отбора и выдачи определенной заранее заданными признаками информации (в том числе документов, их частей и/или данных) из массивов и записей любого вида и на любых носителях. Побудительной причиной осуществления информационного поиска является информационная потребность, выраженная в форме информационного запроса. В зависимости от степени привлечения к информационному поиску технических средств и участия в нем человека, различают "ручной", "машинный" и "автоматизированный" информационные поиски. Последний может производиться в режиме диалога или пакетной обработки запросов.

В автоматизированных информационных системах информационный поиск обеспечивается и осуществляется с привлечением лингвистических, информационных, программно-технических, технологических, организационных средств и составленных из них комплексов.

Непосредственно информационный поиск производится средствами информационно-поисковой системы, являющейся подсистемой АИС. Основными критериями качества результатов информационного поиска являются полнота, точность и оперативность поиска.

ТЕОРИЯ ИНФОРМАЦИОННОГО ПОИСКА [information retrieval theory]

Раздел информатики, изучающий и разрабатывающий вопросы построения информационно-поисковых систем и их использования.

КРИТЕРИЙ [criterion]

Признак, который при оценке объектов, выполняемой путем сравнения, принимается в качестве наиболее существенного для принятия определенного решения, например, отбора и последующей выдачи документов из информационного массива при поиске. При информационном поиске используются разнородные критерии и, в частности, критерии соответствия, смыслового соответствия, формального соответствия и др.

Виды критериев информационного поиска:

- **критерий соответствия [match criterion]** - признак или совокупность признаков, по которым определяется степень соответствия между запросом или поисковым предписанием и поисковым образом документа, самим документом или записью его части для принятия решения о выдаче или не выдаче конкретного документа на информационный запрос, обрабатываемый системой;

- **критерий смыслового соответствия, критерий релевантности (документов, данных) [relevancy criterion]** - признак или совокупность признаков, по которым определяется степень смыслового соответствия между

содержанием поискового предписания и поисковым образом документа, самим документом или записью его части для принятия решения о выдаче или не выдаче конкретного документа на информационный запрос, обрабатываемый системой;

- **критерий формального соответствия [formal match criterion]** - признак или совокупность признаков, по которым определяется степень формального соответствия между поисковым предписанием и поисковым образом документа, самим документом или записью его части для принятия решения о выдаче или не выдаче конкретного документа на информационный запрос, обрабатываемый системой;

- **релевантность [relevance]** - характеристика степени соответствия смыслового содержания документа или данных, найденных в результате поиска, содержанию информационного запроса;

- **пертинентность [pertinence]** - характеристика степени соответствия документа или данных, найденных в результате поиска, информационной потребности пользователя. При правильно сформулированном запросе на поиск пертинентные документы или данные отбираются пользователем из числа релевантных;

- **точность поиска, коэффициент точности поиска/выдачи, ТП [precision ratio, accuracy ratio, precision coefficient, accuracy coefficient]** - количественная характеристика результатов информационного поиска. Точность поиска определяется путем деления количества выданных в результате выполнения поиска релевантных документов на общее количество выданных документов (релевантных и нерелевантных). Данный показатель используется также для оценки качества информационно-поисковых систем;

- **коэффициент информационного шума, коэффициент шума, ИШ [noise, noise ratio, noise coefficient]** - количественная характеристика информационного поиска. Коэффициент информационного шума определяется путем деления количества выданных в результате выполнения поиска нерелевантных документов на общее количество выданных документов (релевантных и нерелевантных). Будучи выраженным в процентах, ИШ = 100%; ТП = 0. Данный показатель используется также для оценки качества информационно-поисковых систем;

- **полнота поиска, коэффициент полноты поиска, ПП [recall, recall ratio, recall coefficient]** - количественная характеристика результатов информационного поиска, которая определяется путем деления количества выданных в результате выполнения поиска релевантных документов на общее количество релевантных документов, имеющих в информационно-поисковой системе (выданных и не выданных). Этот показатель используется также для оценки качества информационно-поисковых систем;

- **коэффициент потерь, КП [losses coefficient]** - количественная характеристика результатов информационного поиска. Коэффициент потерь определяется путем деления количества не выданных в результате выполнения поиска релевантных документов на общее количество релевантных документов, имеющих в информационно-поисковой системе

(выданных и не выданных). Будучи выраженным в процентах, КП = 100% - ПП. Этот показатель используется также для оценки качества информационно-поисковых систем;

- **оперативность (время) поиска [seek time]**

1. Время, затрачиваемое на выполнение поиска, с момента формирования задания и выдачи команды на производство поиска до момента получения его результатов.

2. В ЭВМ: время, затрачиваемое головкой чтения/записи дисководов (накопителя), для того чтобы достичь определенного элемента данных, записанного на диске, дискете или другом носителе данных внешней памяти ЭВМ.

Процессы, завершающие поиск:

- **отбор [selection]** - процесс выделения из небольшого массива документов или данных, отвечающих дополнительным признакам, не учтенным или частично учтенным при информационном поиске. Отбор может составлять часть информационного поиска или выполняться как самостоятельная операция. В зависимости от степени участия в этом процессе человека различают "автоматический" и "машинный" (т.е. с привлечением технических средств) или "ручной" отбор;

- **выдача/вывод информации [information output]** - в автоматизированных системах: процесс выведения документов или данных на экран монитора ЭВМ или в виде твердой копии при помощи печатающего устройства или планшетного графопостроителя;

- **отображение [display, imaging, mapping]**

1. Процесс визуального представления данных, например, на экране монитора ЭВМ.

2. Установление соответствия между элементами двух множеств.

3. Установление соответствия между двумя структурами.

4. Множество значений, обладающих определенным соответствием со значениями и величинами другого множества.

5.5. ИНДЕКСИРОВАНИЕ, ПОИСКОВЫЙ ОБРАЗ ДОКУМЕНТОВ И ЗАПРОСОВ

ИНДЕКС [index, code, notation, mark, symbol]

1. Условный знак (в том числе слово, словосочетание, цифра, буквенный или буквенно-цифровой код и т.п.), обозначающий определенное понятие и используемый для записи результатов классифицирования, а также идентификации объектов поиска в информационно-поисковых массивах.

2. Указатель адреса.

3. Таблица в электронном каталоге, определяющая местоположение набора данных.

4. Совокупность указателей, при помощи которых можно найти запись в файле данных.

5. Уровень в иерархической структуре данных системы многоуровневого

индекса.

Виды индексов, связанных с их структурой:

- **многоуровневый индекс [multilevel index]** - индекс, относящийся к структуре индексов, которая построена в виде иерархического дерева, имеющего два и более уровней иерархии. Корень этого дерева называется индексом высшего уровня или главным индексом [**master index**];

- **одноуровневый индекс** - индекс, относящийся к структуре, которая не отвечает признакам многоуровневости.

В зависимости от характера используемой системы знаков различают:

- **буквенный индекс [alphabetic code, alphabetic notation]** - индекс, использующий отдельные буквы или сочетание букв алфавита;

- **цифровой индекс [numerical code, numerical notation]** - индекс, использующий отдельные цифры, числа, сочетания цифр или их комбинации;

- **десятичный индекс [decimal code, decimal classification code]** - цифровой индекс, составленный на основе десятичной системы счисления;

- **алфавитно-цифровой индекс [alphanumeric code]** - смешанный индекс, состоящий из букв и цифр;

- **смешанный индекс [mixed code, mixed notation]** - индекс, состоящий из разнородных знаков, например, из букв различных алфавитов, букв и цифр и т. п.

В зависимости от структуры и организации записи различают:

- **простой индекс [unitary code]** - индекс, имеющий вид одной законченной записи и отображающий одно понятие, один признак, одно множество и т. п.;

- **сложный индекс [compound code, compound classification number]** - индекс, представляющий собой сложное понятие и составленный путем объединения нескольких других (простых и/или сложных) индексов (см. ранее);

- **составной индекс [combined index (code), combined classification number]**

1. Индекс, образованный в результате объединения нескольких индексов, каждый из которых сохраняет свое основное значение;

2. Индекс, в качестве которого используется комбинация атрибутов;

- **многоуровневый (ступенчатый, иерархический) индекс [multilevel index, hierarchically structured classification number]** - простой или сложный индекс, в котором отдельные его составные части расположены в форме "дерева" в последовательности от общего (корневого) значения к частному, например, индексы УДК, ГРНТИ и др. Корневым в этой структуре (дереве) является индекс высшего уровня;

- **интервальный индекс [interval index]** - индекс, значения которого определяются некоторой областью, например, диапазоном от 3 до 12.

В зависимости от уровня приоритетности различают:

- **гипериндекс [hyperindex]** - высший уровень индекса индексной организации баз данных, принятый в некоторых СУБД (наряду с главным и

нормальными индексами);

- **главный (основной, первичный, старший) индекс [master index, primary index, main subject code, main classification number]**

1. Индекс высшего уровня в иерархической системе организации данных;

2. Индекс, отражающий главную тему содержания индексируемого текста, документа и т. п. и относящийся к основной принятой системе классификации;

- **нормальный индекс [normal index]** - подмножество ключей базы данных, соответствующих конкретному значению поля, объявленного дескриптором (признаком поиска); используется в четырехуровневой системе индексов СУБД, например, ADABAS;

- **вспомогательный (дополнительный) индекс [additional index, auxiliary code, auxiliary classification number]** - индекс, являющийся дополнением к главному (основному) индексу и отражающий дополнительные признаки индексируемого текста, документа и т.п. или относящийся к вспомогательной системе классификации.

В зависимости от характера индексируемых объектов и/или назначения индекса различают, например:

- **каталожный индекс [catalog classification mark, catalog classification number]** - индекс, используемый для расстановки и поиска карточек в каталоге;

- **индекс массива [array index]** - индекс, присваиваемый массиву документов или данных для его идентификации;

- **индекс файла [index number]** - в некоторых операционных системах (например, UNIX) — номер индексного дескриптора файла и др.

ИНДЕКСАЦИЯ [subscripting, notation system, indexing]

1. Метод, обеспечивающий возможность обращения к элементу массива с помощью указания массива и выражений, определяющих местоположение этого элемента в массиве.

2. Система (совокупность) индексов, используемая для индексирования и соответствующая определенной системе классификации.

Примечание: В указанных выше определениях понятие "индексация" никоим образом не обозначает процесс; его нельзя смешивать с понятием "индексирование"!

3. Автоматическое определение истинного адреса путем сопоставления содержимого индексного регистра с адресной частью команды.

Различаются следующие виды индексации:

- **кумулятивная индексация [cumulative indexing]** - индексация, предусматривающая присвоение одному адресу несколько индексов;

- **однорядная индексация [uniserial indexing]** - индексация, в которой использованы так называемые однорядные знаки: буквы одного алфавита, цифры одной системы счисления и т. п.;

- **одноуровневая индексация [single-level indexing]** - индексация с

использованием одноуровневых индексов;

- **многоуровневая индексация [multi level indexing]** - индексация с использованием многоуровневых индексов;

- **смешанная индексация [mixed indexing]** - индексация, в которой использованы различные знаки: буквы, цифры и т. п.

ИНДЕКСИРОВАНИЕ [indexing]

Процесс, связанный с описанием содержания документа, факта или информационного запроса на информационно-поисковом языке.

Процесс индексирования включает:

- 1) анализ содержания индексируемого материала и выбор из него так называемых номинативных лексических единиц, существенных для его понимания;

- 2) формирование перечня ключевых слов, используемых при свободном индексировании;

- 3) нормализацию ключевых слов по форме и содержанию;

- 4) избыточное индексирование;

- 5) заполнение рабочего листа с введением в него грамматических средств.

В зависимости от объекта и содержания процесса индексирования его результатами являются: поисковый образ документа (ПОД), поисковый образ лексической единицы (ПОЛЕ), поисковый образ запроса (ПОЗ) или поисковое предписание (ПП).

ПРЕДКООРДИНАТНОЕ ИНДЕКСИРОВАНИЕ, КЛАССИФИКАЦИОННОЕ ИНДЕКСИРОВАНИЕ, СИСТЕМАТИЗАЦИЯ [pre-coordinate indexing, classifying]

1. Система индексирования документов на основе предварительной (до поступления запроса) организации лексических единиц, наименований рубрик и их классификационных индексов.

2. Индексирование документов и других текстовых материалов с использованием классификационных словарей (рубрикаторов). Указанный вид индексирования в соответствии с ГОСТ 7.74-96 носит также наименования библиографической или классификационной систематизации.

К предкоординатному индексированию также относят:

- **систематизацию [ordering, systematization]** - процесс упорядоченного распределения каких-либо объектов (например, библиографических записей), осуществляемый по сходству или различию присущих им признаков;

- **предметизацию [subject cataloguing, subject indexing, featuring, subject analysis]**

1. Индексирование предметного содержания документов средствами языка предметных рубрик.

2. Вид систематизации, связанный с распределением объектов по предметным рубрикам - некоторым классификационным признакам.

3. Определение предметной рубрики для произведения печати или

другого документа в соответствии с его содержанием.

КООРДИНАТНОЕ (ПОСТКООРДИНАТНОЕ) ИНДЕКСИРОВАНИЕ [coordinate (post-coordinate) indexing]

Индексирование документов и фактов на естественном языке с использованием дескрипторов и/или ключевых слов, отражающих в семантическом (понятийном) плане следующие лексикографические категории:

- 1) термин, выражающий конкретное научно-техническое понятие или объект описания;
- 2) имя собственное, идентификатор;
- 3) наименование параметра;
- 4) значение параметра (выраженное текстом или именованной величиной);
- 5) числовое выражение;
- 6) обозначение единицы величины.

В соответствии с ГОСТ 7.66-92 цель координатного индексирования состоит во всестороннем отражении содержания документа или запроса путем включения в поисковый образ всех необходимых для этого терминов индексирования.

При координатном индексировании документов и фактов дескрипторы и/или ключевые слова связываются или комбинируются для указания любых отношений, необходимых для организации наиболее полного и точного поиска.

Координатное индексирование предполагает деление длинных словосочетаний на короткие, содержащие, как правило, не более 2-3 слов. Допускается использование общепринятой или утвержденной (в соответствующей АС) аббревиатуры. При свободном индексировании ключевыми словами, взятыми из текста документа, последние должны быть приведены к канонической форме по ГОСТ 7.25. Количество лексических единиц, используемых в ПОДе, должно соответствовать требованию полного отображения важных для пользователей аспектов содержания индексируемого документа.

По разным признакам различают также следующие виды индексирования:

• **контролируемое индексирование [controlled indexing]** - индексирование, контролируемое машинными словарями системы;

• **свободное индексирование [free indexing]:**

1) индексирование, технология которого не предусматривает замену ключевых слов текста в соответствии с рекомендациями специального словаря;

2) координатное индексирование текста документа ключевыми словами, выбираемыми непосредственно из самого текста или добавляемыми в ПОД без использования какого-либо нормативного словаря;

• **избыточное индексирование [redundant indexing]** - дополнение ПОДа и/или ПОЗа лексическими единицами ИПЯ, связанными сильными парадигматическими отношениями с лексическими единицами исходного ПОДа, в том числе индексами других уровней при использовании системы многоуровневой индексации. Избыточное индексирование служит средством обеспечения повышения полноты поиска. Различают два вида избыточного индексирования - восходящее и нисходящее;

• **автоматизированное индексирование [computer-aided indexing, automated indexing]** - индексирование с использованием формальных процедур, выполняемых вычислительной техникой, и интеллектуальных процедур, выполняемых человеком при формировании ПОД [14];

• **автоматическое индексирование, автоиндексирование [automatic indexing]** - индексирование, выполняемое полностью программно-техническими средствами (ЭВМ);

ЛОГИЧЕСКАЯ ОПЕРАЦИЯ [logical operation]

Машинная операция, реализующая действия над логическими величинами с получением логического значения. В языках программирования и при информационном поиске чаще всего используются три логические операции: конъюнкция, дизъюнкция и отрицание (см. далее).

• **Конъюнкция [konjaction]** - логическая операция умножения, требующая совпадения двух или нескольких признаков искомого документа или факта. Соответствующий ей логический оператор - "И"; обозначается знаками "&" и "и"; используется для сужения области поиска.

• **Дизъюнкция [disjaction]** - логическая операция сложения, разделяющая два или несколько признаков искомого документа или факта. Соответствующий ей логический оператор "ИЛИ"; обозначается знаками ",", "и" и "или"; используется для расширения области поиска.

• **Отрицание [negatio]** - логическая операция "импликации" (вида логической связи), требующая исключения из области поиска тех документов или фактов, которые имеют "нежелательные" признаки. Соответствующий ей логический оператор - "НЕ" ("NOT"). Обозначается знаками "!", "NOT".

ИНДЕКСНЫЙ ФАЙЛ [index file]

1. Файл, обеспечивающий доступ к записи по ключу.
2. Список указателей записей базы данных, упорядоченный по значениям одного или нескольких ее полей. В качестве указателей используются логические или физические адреса записей.

ИНДЕКСИРОВАННЫЙ ФАЙЛ [indexed file]

Файл данных, в котором обращение к записям производится при помощи системы индексов, обеспечивающей быстрый доступ. В автоматизированных информационных и библиотечных системах примером последнего является режим быстрого поиска. Если одно и то же поле используется в индексе и для

упорядочения записей файла, то индекс называется основным, а файл — индексно-последовательным. В другом случае индекс называется вторичным.

ИНВЕРТИРОВАННЫЙ ФАЙЛ [inverted file]

1. В информационно-поисковых системах - файл или каталог, организованный таким образом, что каждый их элемент (запись) идентифицирован соответствующим индексом, причем все записи (в файле или каталоге) упорядочены в соответствии с числовым или алфавитным возрастанием индексов.

2. Файл, в котором последовательность расположения данных, изменена на обратную.

Разновидности инвертированных файлов:

- **индексно-последовательный набор данных [indexed-sequential data set]** - набор (массив) данных, в котором каждая запись содержит ключ, определяющий местоположение этой записи. Место размещения каждой записи в индексно-последовательном наборе данных вычисляется при помощи индекса, а сами записи упорядочены в возрастающей последовательности значений индекса;

- **индексно-последовательный файл [indexed-sequential file]** - файл, каждая из записей которого снабжена своим ключом, в результате чего обеспечивается как прямой доступ к каждой записи по ее ключу, так и последовательный доступ в соответствии с реализованным принципом упорядочения записей по ключам;

- **файл прямого (произвольного) доступа [direct access file]** - файл, доступ к записям в котором производится по адресу либо последовательно путем поиска по ключу.

6. БЕЗОПАСНОСТЬ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ

6.1. ОБЩИЕ ТЕРМИНЫ И ПОНЯТИЯ

БЕЗОПАСНОСТЬ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ [information technology security, information security, data security]

1. Признак, характеризующий отсутствие угрозы человеческому сообществу, определенной группе людей, конкретному индивидууму и/или окружающей среде от использования средств или продуктов современной информационной технологии. В конкретных случаях применения термина принято дополнять его сведениями об области распространения или использования, например, "безопасность информационной системы", "информационная безопасность личности", "информационная безопасность фирмы" и т. п. (см. также "ITSEC").

2. Состояние защищенности данных (информации), обрабатываемых и сохраняемых средствами вычислительной техники или автоматизированной системы от внутренних и внешних угроз.

3. Совокупность правовых, организационных, технологических мер, методов и средств, предназначенных для предотвращения потерь или неправильного (случайного или сознательного) использования информации. Частными средствами и мерами реализации информационной безопасности являются **защита данных, защита от несанкционированного доступа, защитный экран**, средства и способы борьбы с компьютерными вирусами, резервное копирование и **архивация** данных и др.

В соответствии с действующим в Российской Федерации законодательством и, в частности, Федеральным законом РФ "Об информации, информатизации и защите информации" от 20 февраля 1995 г. №24-ФЗ, информационная безопасность обеспечивается с учетом характера конкретных видов документов и данных.

Ниже приведены **классификация информационных источников по признакам ограничения их распространения и степени защиты**, а также основания для отнесения к той или иной категории:

1. **Открытая информация [public data (information)]** - информация во всех областях знаний и деятельности, относящаяся как к государственным, так и негосударственным информационным ресурсам, доступ к которой не может быть ограничен.

2. **Информация с ограниченным доступом [information with restricted access]** - информация, которая в соответствии с российским законодательством не отнесена к разряду открытой. Информацию с ограниченным доступом составляют:

а) **Информация, отнесенная к государственной тайне [State secret information]** - документы, данные и сведения в любой форме, которым на основании Закона РФ "О государственной тайне" уполномоченными органами присвоены грифы "Особой важности", "Сов. Секретно" и "Секретно".

б) **Конфиденциальная информация [confidential information, privileged information]** - информация, отнесенная Российским законодательством к "Персональным данным", "Сведениям о коммерческой деятельности", "Сведениям о профессиональной деятельности" и/или "Сведениям служебного характера". Режимы защиты конфиденциальной информации устанавливаются ее собственниками - государственными органами, учреждениями, организациями, предприятиями, фирмами, физические лица.

в) **Персональные данные [personal information]** - сведения, защищенные Ст. 23 Конституции РФ "о праве на неприкосновенность частной жизни, личную и семейную тайну".

г) **Сведения о коммерческой деятельности [commercial data (information)]** - сведения, защищенные Гражданским кодексом РФ и федеральными законами, как "Коммерческая тайна" (определение см. в Гражданском кодексе РФ, Ст. 139).

д) **Сведения о профессиональной деятельности [data (information) on the professional activity]** - к ним относятся следующие виды сведений:

- о сущности изобретения;
- составляющие тайну следствия;
- отнесенные к "Банковской тайне" (Ст. 26 Федерального закона "О банках и банковской деятельности");
- о предмете Договора и результатах его выполнения (Ст. 771 Гражданского кодекса РФ);

е) **Сведения служебного характера [office data (information)]**:

- документы с грифом ДСП в федеральных органах и их структурах (Постановление Правительства РФ от 3 ноября 1994 г. №1233);
- служебная информация о деятельности Банка России (Ст. 92 Федерального закона "О Центральном банке РФ");
- служебные сведения, отнесенные Гражданским кодексом и федеральными законами к "Служебной тайне" (см. п. 3 Перечня 15);
- служебные сведения, имеющие коммерческую ценность (Ст. 139 Гражданского кодекса РФ);
- сведения о вкладах физических лиц (Ст. 108 Закона РФ "О федеральных органах налоговой полиции").

В Российской Федерации ИБ регулируют следующие стандарты:

- **ГОСТ Р ИСО/ШК 15408-1-2002.** Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Ч.1. Введение и общая модель. Госстандарт России;
- **ГОСТ Р ИСО/ШК 15408-2-2002.** Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Ч.2. Функциональные требования безопасности. Госстандарт России;
- **ГОСТ Р ИСО/ШК 15408-3-2002.** Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Ч.3. Требования доверия к безопасности. Госстандарт России;
- **ГОСТ Р 50739-95.** Средства вычислительной техники. Защита от несанкционированного доступа к информации. Общие технические требования. Госстандарт России;
- **ГОСТ Р 50922-96.** Защита информации. Основные термины и определения;
- **ГОСТ Р 51188-98.** Защита информации. Испытания программных средств на наличие компьютерных вирусов. Типовое руководство. Госстандарт России;
- **ГОСТ Р 51275-99.** Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения. Госстандарт России;
- **ГОСТ Р ИСО 7498-1-99.** Информационная технология. Взаимосвязь открытых систем. Ч.1. Базовая модель. Госстандарт России;
- **ГОСТ Р ИСО 7498-2-99.** Информационная технология. Взаимосвязь открытых систем. Базовая эталонная модель. 4.2. Архитектура защиты информации. Госстандарт России.

MSS (Managed Security Services) - "Услуги по управлению информационной безопасностью" составляют значительную и постоянно развивающуюся часть рынка информационных технологий. Этот рынок состоит из ряда сегментов, определяемых концепцией 3А: управляемых защитных экранов; VPN; IDS; услуг управляемого мониторинга безопасности; ПО управлением безопасностью контента (Secure Content Management SW) и антивирусными решениями; аппаратных систем безопасности ИТ (Security Hardware) и др.

ЗАЩИТА ДАННЫХ, ЗАЩИТА ИНФОРМАЦИИ, БЕЗОПАСНОСТЬ ДАННЫХ [data protection]

Совокупность мер, обеспечивающих защиту прав собственности владельцев информационной продукции, в первую очередь, программ, баз и банков данных от несанкционированного доступа, использования, разрушения или нанесения ущерба в какой-либо иной форме.

Существуют следующие принципы защиты информации (ЗИ):

- ЗИ должна быть комплексной и включать правовые, административные и программно-аппаратные средства;
- ЗИ должна строиться адаптивно с учетом постоянно изменяющихся условий;
- абсолютно надежной ЗИ не существует, поэтому она должна строиться исходя из характера потенциальной угрозы и ценности защищаемой информации, что определяет в комплексе характер выделяемых на нее сил и средств.

В зависимости от характера ЗИ различают:

- защиту от несанкционированного доступа ресурсов автономно работающего ПК (реализуется преимущественно программными и программно-аппаратными средствами);
- защиту серверов и отдельных пользователей сети Интернет от хакеров (для этого используются межсетевые экраны - брандмауэры);
- защиту секретной, конфиденциальной и личной информации от чтения и использования посторонними лицами (применяются программные, в том числе криптографические, и аппаратные средства);
- ЗИ от утечки по побочным каналам, например, радио, электромагнитного излучения, цепям питания и т.п. (применяются разнородные средства - экранирование рабочих помещений, генераторы шума, специальные составы оборудования и комплектующих средств, имеющие минимальный уровень излучения и т. п.);
- защиту программного обеспечения от копирования (использование электронных ключей);
- ЗИ от шпионских устройств, устанавливаемых непосредственно в комплектующие изделия ПК (выполняется специальными средствами компетентных органов);
- ЗИ от сетевых атак, а также атак методами социальной инженерии (см. далее);

- Борьбу с киберсквоттингом.

ITSEC (Information Technology Security Evaluation Criteria) - принятые в 1991 г. сообществом четырех европейских стран (Франция, Германия, Нидерланды и Великобритания) "Критерии оценки безопасности информационных технологий". Указанные в документе критерии рассматривают следующие составные части информационной безопасности:

- **конфиденциальность информации [information sensitivity]** - защита от несанкционированного получения информации;
- **целостность информации [information integrity]** - защита от несанкционированного изменения данных (информации);
- **доступность информации [information accessibility]** - защита от несанкционированного удержания информации и ресурсов.

Обеспечение гарантированных функций безопасности в соответствии с указанными критериями достигается комплексом мер, распределенных по своему содержанию в следующих разделах выпущенного ITSEC нормативного акта: "Идентификация и аутентификация", "Управление доступом", "Подотчетность", "Аудит", "Повторное использование объектов", "Точность информации", "Надежность обслуживания" и "Обмен данными". Для облегчения задач оценки и сертификации информационных систем и продуктов "Критерии ..." содержат описания десяти типизированных классов функций безопасности для правительственных и коммерческих систем.

Конфиденциальная информация [sensitive information]

1. Сведения о личности, которые охраняются законодательством в соответствии с правом гражданина на защиту от тайного надзора или нанесения ему ущерба со стороны государства, каких-либо юридических, а также других частных лиц. Сохранение конфиденциальности находится в определенном противоречии с требованиями обеспечения государственной безопасности и поддержки других государственных интересов. Оно еще более усложняется в условиях действия современных информационных технологий, в том числе создания и развития правительственных и неправительственных банков данных, которые содержат или могут содержать конфиденциальную информацию, и средств теледоступа к ним.

2. Любая информация, требующая защиты от несанкционированного доступа.

Копирайт [copyright] - право автора или собственника какого-либо продукта интеллектуальной деятельности (например, литературного произведения, программы, базы данных и т.п.) запретить его несанкционированное использование и тиражирование. Защищается государственным и международным законодательством. В России такими законодательными актами являются Закон "Об информации, информатизации и защите информации" (базовый) и Закон "О правовой охране программ для электронных вычислительных машин и баз данных", выпущенные соответственно в 1995 г. и 1992 г. В 1981 г. Совет Европы одобрил Конвенцию по защите данных. В Великобритании аналогичный

закон принят в 1984 г. Указанные законы устанавливают нормы, регулирующие отношения в области формирования и потребления информационных ресурсов, создания и применения информационных систем, информационных технологий и средств их обеспечения, защиты информации и защиты прав граждан в условиях информатизации общества (см. также "ITSEC").

Система защиты данных [security system] - комплекс программных, технических, криптографических и организационных средств, обеспечивающих защиту данных от несанкционированного доступа к ним, их использования, а также преднамеренного или случайного разрушения и искажения.

АТАКА, НАПАДЕНИЕ [attack]

Всякое действие, связанное с несанкционированным доступом в вычислительную сеть (а также его попыткой) и/или преднамеренным нанесением ущерба как сети в целом, так и любым ее составным частям, включая условия или результаты их функционирования.

Виды сетевых атак (нападений):

- **Внутренняя атака [insider attack]** - нападение на защитный экран изнутри защищаемой сети.

- **Перехват IP [IP spacing/hijacking]** - нападение, при котором атакующий перехватывает или корректирует установленный канал доступа. Нападающий в такой атаке маскируется под законного пользователя, прошедшего процедуру аутентификации.

- **Подделка DNS [DNS spoofing]** - подделка идентификаторов ЭВМ сети, например, путем перехвата ответов DNS и замены их фальшивыми.

- **Подделка IP, IP-спуфинг, спуфинг [IP spoofing]** - вид атаки на сеть, при которой хакер, находящийся внутри корпоративной сети или вне ее, выдает себя за добросовестного пользователя, поддельывая санкционированные внешние или внутренние IP-адреса системы. Атаки IP-спуфинга часто являются исходными для производства атак другого рода, например DoS, для сокрытия личности хакера. Одной из разновидностей спуфинга является обман дактилоскопических сканеров за счет предъявления им так называемого "ложного пальца", выполненного из какого-либо пластического материала.

- **Brute-force attack** - атака методом "грубой силы": атака на зашифрованный текст или электронную цифровую подпись, использующая прямой перебор всех возможных вариантов ключей шифрования или ЭЦП.

- **CSS (Cross-Site Scripting)** - "Межсайтовый скриптинг": вид атаки, при котором в загружаемую Web-страницу внедряется созданный атакующим посторонний код. Целью атаки может являться запуск скрипт-вируса для воровства чего-либо, модификация содержимого Web-страницы, перехват сессии и т.п.

- **DoS (Denial of Service), DDoS (Distributed Denial of Service)** - наименование одного из наиболее распространенных видов атак на

отдельные ПК и/или вычислительные сети (например, корпоративные) в Интернете, вызывающего "отказ в обслуживании", т.е. в выполнении вполне законных соединений, в связи с тем, что все ресурсы сети, ее программного обеспечения и/или ПК адресата заняты обслуживанием большого числа запросов, искусственно созданных нападающей стороной. В Интернете в последние годы для осуществления DoS-атак (особенно на банки, казино и другие финансовые учреждения) стали широко использоваться так называемые боты.

- **Flood** - "Наводнение": DoS-атака, выполняемая отправкой жертве большого числа TCPTM, UDPTM или UMP-пакетов.

- **Land** - ВНfl DDoS-атаки, выполняемая с целью "зацикливания" IP-пакета с совпадающими адресами отправителя и получателя; использует некорректную реализацию сетевых сервисов некоторых операционных систем.

- **Ping of Deaths** - "Свист смерти": вид Dos-атаки, осуществляемой при помощи некорректно фрагментированных UMP-пакетов.

- **Teardrop** - "Слеза": вид DoS-атаки, провоцирующей компьютер на сборку пакета отрицательной длины.

- **Tiny fragment** - "Крошечный фрагмент": атака, позволяющая проникнуть внутрь защищаемой сети с помощью специальным образом фрагментированных TCP-пакетов.

- **UBE (Unsolicited Bulk E-mail)** - "Нежелательный большой объем электронной почты": массовая рассылка нежелательных сообщений электронной почты (спам).

- **URL spoofing** - вид атаки, при котором искажается адрес загруженного сайта, производится модификация заголовка окна или данных в панели статуса.

- **WEPAAttack** - "WEP-атака": вид атак на беспроводные сети, связанных с вскрытием WEP-шифрования.

Способы сетевого мошенничества:

- **Социальная инженерия, социальный инжиниринг [social engineering]**

1. В обычном смысле: действия злоумышленников, связанные с получением конфиденциальной информации лично или с использованием средств связи (например, телефона) путем обмана, в том числе выдачи себя за другое лицо.

2. В широком значении: психологическое воздействие на человека с использованием современных телекоммуникационных и вычислительных средств (в том числе через Интернет) путем введения в заблуждение для достижения каких-либо (часто преступных) целей (в том числе проникновение в охраняемый объект, базу данных, получение конфиденциальной информации и т.п.).

- **Киберсквоттинг, хапперство [cybersquatting]** - вид деятельности, связанный с захватом доменных имен с целью их дальнейшей перепродажи или недобросовестного использования. Наибольший интерес у лиц и

организаций, занимающихся киберсквоттингом (они известны под именами "кибереквоттеры", "хапперы" и "сквоттеры"), представляют доменные имена, имеющие следующие признаки:

- 1) содержат весьма распространенные слова и словосочетания, обозначающие определенные понятия и виды деятельности;
- 2) совпадают (полностью или частично) с наименованиями широко известных торговых марок и фирм;
- 3) совпадают с именами и фамилиями знаменитых людей;
- 4) имеют в своем составе географические названия (стран, городов, районов и т. п.).

• **Получение нелегальных полномочий [abuse of privilege]** - действия пользователя, связанные с получением не разрешенных ему прав, противоречащие закону или политике безопасности организации-владельца сети.

• **Фишинг [fishing, phishing]** - вид сетевого мошенничества, связанный с намерением получить номера счетов, паролей, PIN-кодов пластиковых карт и других данных с целью незаконного обогащения нападающей стороны. Этот "рыболовный промысел" производится путем "забрасывания сетей", при котором намеренной жертве от имени банка, Интернет-магазина, онлайн-аукциона или других организаций и фирм под каким-либо предлогом (необходимость обновления БД, утери персональных данных вследствие хакерской атаки и т.п. причин) направляется предложение сообщить о себе те или иные данные.

Технологии и способы технической защиты информации:

• **CSMS (Cisco Security Management Suite)** - выпущенный в начале 2006 года компанией Cisco набор приложений для управления безопасностью вычислительных сетей. В состав CSMS входят: новый модуль CSM (Cisco Security Manager) - консоль централизованного управления средствами защиты и Cisco Security MARS (Monitoring, Analysis and Response System) - система мониторинга, анализа и реагирования. CSM состоит из наборов средств отображения (views) состояния защиты сети, которые могут быть созданы на трех уровнях: device view - устройств безопасности (позволяет посмотреть полный перечень устройств безопасности и применимых к ним правил); topology view - топологии (отображает структуру системы безопасности с использованием масштабируемых карт сети) и policy view - наборов правил (позволяют создавать и настраивать правила безопасности в зависимости от текущих потребностей организации). Программно-аппаратный комплекс Cisco Security MARS предназначен для управления безопасностью сети с учетом характера текущих угроз.

• **DRM (Digital Right Management)**- "Управление цифровыми правами": технология, обеспечивающая ограничение на аппаратном уровне несанкционированного их владельцами использования оцифрованных интеллектуальных продуктов (аудио-, графика, тексты, мультимедиа и т.п.).

• **LT (LaGrande)** - условное наименование программы и технологии, направленных на усиление защиты хранящихся на ПК данных. Содержание

LT объявлено на Форуме Intel (IDF) в 2003 г. в калифорнийском городе Сан-Хосе и Москве. Суть технологии заключается в том, что ПК должен стать защищенным хранилищем данных, находящихся в разных "отсеках" памяти ПК в зашифрованной форме. Допуск к ним пользователя производится путем запуска специальной программы, содержащей криптографические ключи. По окончании работы пользователя данные снова шифруются. Реализация технологии LT требует изменения архитектуры центрального процессора, контроллера памяти, контроллеров ввода-вывода и добавления модуля TPM (Trusted Platform Module), который служит безопасным хранилищем данных, реализует ряд криптографических функций, а также содержит и сообщает сведения о целостности данных. Одновременно существенной переработки требуют BIOS и операционная система. Кроме того, должны использоваться специальные конструкции видеокарты, клавиатуры и мыши. Выработаны все необходимые спецификации LT, получившие поддержку крупнейших участников рынка ИТ. В частности, корпорации Intel и AMD поддерживают реализацию LT в своих процессорах и чипсетах.

• **MCP (Microsoft Client Protection)** - система программных продуктов и технологий, разрабатываемых корпорацией Microsoft, для защиты настольных ПК, ноутбуков и файловых серверов в информационных системах предприятий от вирусов, шпионского ПО и других видов внешних угроз. В ней предполагается:

1) усилить внутренние средства защиты технических и программных средств;

2) создать специальные средства, снижающие уязвимость от различных вредоносных программ;

3) повысить уровень управления идентификацией пользователей и их доступом к различным ресурсам.

• **MILS (Multiple Independent Levels of Security/Safety)** - "Многоуровневая система безопасности с изолированными разделами": архитектура построения MLS-систем (включая и сетевые решения), основанная на реализации требований NEAT. MILS-архитектура строится на базе компактных компонентов, которые обеспечивают защиту отдельных приложений индивидуальным способом, что собственно и позволяет построить многоуровневую защиту. MILS-компоненты можно по отдельности сертифицировать и затем использовать в других проектах.

Синхронизацию работы MILS-компонентов осуществляет ядро безопасности, которое, в отличие от решений защиты, использующих объемную автономную систему, обеспечивает выполнение только функций NEAT и занимает всего 5% процессорных ресурсов. Основные задачи, решаемые ядром безопасности MILS-архитектуры:

1) разделение адресного пространства;

2) изоляция разделов, в которых работают приложения (сбой в одном разделе не приводит к сбоям в других);

3) разделение интервалов работы процессора;

4) защищенная и авторизованная связь между MILS-разделами (каждое

приложение выполняется в собственном пространстве памяти и может взаимодействовать только с разделами своего или более низких уровней защиты, используя для этого специальные MILS-сокеты).

Особенностью MILS-архитектуры является тот факт, что она не требует переделки ядра ОС.

- **MLS (Multiple Levels of Security)** - "Многоуровневая защита (информации)": принципы и технология построения автоматической системы защиты, заключающиеся в разделении виртуальной памяти на сегменты кода и данных. Взаимодействие между этими сегментами осуществляется с учетом уровня их секретности.

- **NEAT (Non-bypassable, Evaluatable, Always invoked, Tamperproof)** - концепция защиты информации, развивающая принципы MLS следующими требованиями к функциям средств защиты:

- 1) их нельзя обойти или отключить;
- 2) они должны быть невелики и математически выверены;
- 3) должны вызываться регулярно и работать постоянно при любых действиях прикладных программ;
- 4) они не могут быть ни повреждены, ни модифицированы плохим или зловредным кодом защищаемого приложения.

6.2. КОДИРОВАНИЕ И ДЕКОДИРОВАНИЕ ДОКУМЕНТОВ И ДАННЫХ

КОДИРОВАНИЕ [coding, encode]

1. В информационных системах: процесс преобразования текстовых и других материалов из одной системы символов в другую.

2. Процесс отображения дискретных сообщений сигналами в виде определенных сочетаний символов.

3. Процесс преобразования детальной спецификации из одной программы в другую программу.

4. Процесс шифрования.

ДЕКОДИРОВАНИЕ [decoding, code translation, interpretation]

В информационных системах процесс преобразования кодированных данных в исходную или другую пригодную для чтения форму UU-кодирование/декодирование [UUencode/UUdecode] - самый распространенный набор алгоритмов кодирования и декодирования, благодаря которому становится возможным обмен файлами между различными системами электронной почты. ЭВМ отправителя преобразует текст файла в ASCII-код, который может быть декодирован получателем в процессе восстановления исходных данных.

Кодер [coder] - устройство, преобразующее исходный текст в какой-либо код; кодирующее устройство.

ДЕКОДЕР [decoder]

1. Устройство, преобразующее закодированный текст в исходную форму (вид); декодирующее устройство.

2. Электронное устройство, предназначенное для выбора и реализации одного из способов передачи данных, например, для направления данных в индивидуальные ячейки памяти внутри быстродействующего запоминающего устройства ЭВМ.

КОД (от франц. code)

1. В широком смысле: система условных обозначений или сигналов, например, телеграфный код.

2. Используемый для кодирования набор символов, которому приписан некоторый смысл.

3. Шифр.

4. Множество слов (кодовых комбинаций), букв некоторого алфавита, цифровых или других знаков и их сочетаний, поставленное во взаимно однозначное соответствие другому (кодируемому) множеству.

5. Программа на машинном языке.

6. Язык, на котором написана программа.

Наиболее используемые в информационной технологии коды:

• **Двоичный (цифровой) код [binary number code]** - код, основанный на двоичной системе счисления [binary number system], использующей для представления буквенно-цифровых и других символов наборы комбинаций цифр 1 и 0. Является основой всех используемых в цифровых ЭВМ кодов (например ASCII, шрифтов Windows и др.).

• **ASCII (American Standard Code for Information Interchange)** - американский стандартный код для обмена текстовой информацией. Первая ("нижняя") половина таблицы ASCII (коды 0-126) содержит символы английского алфавита, знаки препинания и арабские цифры. Она является общепринятой во всем мире. Для национальных языков используются "верхняя" половина таблицы ASCII (коды 127-255, или "расширенные" ASCII-коды). В ней находятся буквы национальных алфавитов и специальные символы. ASCII представляет собой систему кодирования, в которой буквам, цифрам и знакам присвоены определенные числовые (десятичные и двоичные) значения. Например, десятичное число 45 (двоичное - 0101101) соответствует знаку "-", а 65 (двоичное - 1000001) - заглавной букве "A". Первые 32 числа кода используются для функций управления (например, возврата каретки или возврата на один символ). Восьмой бит используется для представления дополнительных символов или для целей контроля правильности передачи данных.

• **EBCDIC (Extended Binary-Coded Decimal Interchange Code)** - расширенный двоично-десятичный 8-разрядный код, способный поддерживать 256 буквенно-цифровых символов.

• **Unicode** - 16-разрядная система кодирования, совместимая с системой ASCII, которая охватывает символы разных письменностей: латинской, кириллической, греческой, а также языков, использующих иероглифы.

Система присваивает уникальный код любому символу, независимо от платформы, программы или языка. Для представления символа используются два байта. Коды разделены на несколько областей: область с кодами от 0000 до 007F содержит символы набора Latin 1 (младшие байты соответствуют кодировке ISO 8859-1), далее идут области, в которых расположены символы различных письменностей, а также знаки пунктуации и технические символы. Часть кодов зарезервирована для применения в будущем. Символам кириллицы выделены коды в диапазоне от 0400 до 0451. Для работы с документами Unicode необходимы соответствующие шрифты. Как правило, файл шрифта Unicode содержит начертания не для всех символов, определенных в стандарте.

- **Cyrillic KOI8-R (КОИ — Код обмена информацией)** - 8-разрядный код, ориентированный на обмен сообщениями в Интернете, написанными кириллицей (на русском языке). Его организация обеспечивает смещение русских букв относительно латинских на 128 бит (см. ранее "ASCII"). В случае "срезания" восьмого бита при прохождении сообщений по сети, исходный русский текст транслитерируется латинским шрифтом.

- **Штриховой код [bar code]** - код, сформированный в виде набора коротких линий (штрихов) и пробелов различной ширины. Автоматизированное кодирование и декодирование производится с использованием средств вычислительной техники, однако расшифровка штрихового кода может быть произведена и человеком, знакомым с принципом кодирования. Автоматизированная запись и считывание штриховых кодов производятся при помощи специализированных принтеров и сканеров. Штриховые коды широко используются в торговле, промышленности и библиотечном деле.

- **PIN (Personal Identification Number)** - "Персональный идентификационный номер":

1. В банковском деле: уникальный цифровой код, используемый в качестве пароля для идентификации клиента при пользовании автоматическим банкоматом. PIN записывается на магнитной полоске банковской карточки клиента в зашифрованной форме. Перед выдачей денег банкомат сопоставляет считанную с карточки запись с ручным набором, выполненным клиентом при помощи клавиатуры.

2. Аналог пароля в различных механизмах аутентификаций.

ШИФР [cipher]

Код, значение и правила использования которого известны ограниченному кругу лиц. Шифр предназначен для защиты информации и других объектов от несанкционированного доступа.

ПАРОЛЬ [password]

Секретная комбинация символов или слово, предъявляемые пользователем для получения доступа в автоматизированную систему. Служит для защиты программ и данных от несанкционированного доступа.

БИОМЕТРИЧЕСКИЙ ПАРОЛЬ [biometric password]

Средство аутентификации, основанное на предварительной записи и сопоставлении с предъявляемыми при контроле некоторых биологических характеристик пользователей или лиц, имеющих доступ к чему-либо в охраняемой системе. Такими биологическими характеристиками могут быть изображение (лица, радужной оболочки глаза, отпечатков пальцев и т.д.), тембр голоса, ДНК, геометрическая форма руки, ритм работы на клавиатуре и т.д.

Biometrics - средства обеспечения безопасности, использующие биометрический контроль.

ИДЕНТИФИКАТОР ПОЛЬЗОВАТЕЛЯ [user identifier]

Присвоенный пользователю кодовый шифр (пароль), с которым зарегистрированный пользователь входит в систему и который используется автоматизированной системой аутентификации защищаемого объекта или ресурса для определения уровня его прав, а также для регистрации факта доступа и характера выполненных им работ или предоставленных ему услуг. В некоторых автоматизированных системах (в первую очередь, теледоступа к информационным ресурсам) предусматривается возможность оказания ограниченного объема услуг незарегистрированным пользователям, входящим с гостевым идентификатором (паролем) — guest.

В последние годы в качестве идентификаторов стали все чаще применяться биометрические средства, основанные на использовании в качестве паролей записанных в памяти ПК образцов изображений зарегистрированных пользователей (например, лица, сетчатки глаза, отпечатков пальцев и т. п.) и/или тембров их голосов, а также сочетания указанных признаков.

Token - "Токен": электронный идентификатор пользователя, предназначенный для хранения паролей, шифровальных ключей и/или цифровых сертификатов с целью обеспечения защищенного доступа к информационным ресурсам предприятия, физического доступа в здания или помещения и т. п. В основе конструкции токенов лежит использование USB-ключей и смарт-карт. Токены применяют двухступенчатый механизм аутентификации: пользователь вводит свой PIN, с помощью которого система опознает владельца токена. В свою очередь, токен, получив сигнал от системы, высвечивает на своем дисплее идентификационное число (оно меняется для каждого пользователя через каждые несколько минут), ввод которого и позволяет произвести аутентификацию. Таким образом, создается двойная защита пароля.

6.3. КРИПТОЛОГИЯ И СВЯЗАННЫЕ С НЕЮ ПОНЯТИЯ

КРИПТОЛОГИЯ [cryptologic]

Область деятельности, связанная с защитой информации. По мнению

одного из основоположников криптологии как науки Джеймса Иэсси (университет ЕТН, Швейцария), она делится на два направления - криптографию и криптоанализ.

Криптография [cryptography] - поиск и использование методов, гарантирующих надежное засекречивание сообщений и/или подтверждения их подлинности. Результат криптографической обработки исходных (открытых) сообщений носит наименование криптограммы или шифр-текста. Управление процессами шифрования и дешифрирования производится при помощи специальных алгоритмов и ключей. При этом алгоритмы могут быть известны (не засекречены), а основная нагрузка на защиту сообщений ложится на секретный ключ. Степень способности криптосистемы противостоять раскрытию шифра называют ее стойкостью.

Криптоанализ [cryptanalysis] - деятельность, ориентированная на вскрытие шифров ("раскрытие шифра текста" - не путать с "дешифрированием!"), а также подделку кодов сигналов так, чтобы их можно было принять за подлинные.

Шифрование [ciphering, coding, encryption] - процесс перевода текстов, цифровых и других данных в зашифрованную форму с целью ограничения доступа к их содержанию нежелательных лиц, организаций.

Дешифрирование [deciphering, decoding] - процесс преобразования закодированных текстов, цифровых и других данных в их первоначальную или другую удобную для чтения форму.

ЭЛЕКТРОННЫЙ КЛЮЧ, ШИФРОВАЛЬНЫЙ КЛЮЧ, КЛЮЧ [cipher code, key source, key]

1. Цифровой код, применяемый для шифрования и/или дешифрирования документов и, в частности, при их пересылке или предоставлении доступа в Интернет.

2. Программно-аппаратное устройство, донгл [dongle], защищающее прикладные программы от несанкционированного использования. Обычно подсоединяется к последовательному или параллельному порту ЭВМ (например, между портом и кабелем принтера или модема).

Виды шифровальных ключей:

- **открытый ключ [public-key]** - ключ, код которого известен всем участникам передачи и/или приема сообщений. При шифровании открытым ключом для дешифрирования сообщений используется секретный (закрытый) ключ;

- **секретный (закрытый) ключ [private-key]** - ключ, код которого известен только одной стороне (передающей или принимающей документы и данные);

- **симметричные ключи [symmetric(al) key]** - одинаковые ключи для шифрования и дешифрирования данных;

- **асимметричные ключи [asymmetric(al) key]** - разные ключи для шифрования и дешифрирования данных;

- **сеансовый ключ [session key]** - действующий только в одном сеансе

передачи сообщений, симметричный (секретный) ключ шифрования, с помощью которого производится как шифрование, так и дешифрирование с использованием одного и того же алгоритма симметричного шифрования. Этот ключ передается безопасным образом обеим взаимодействующим сторонам для передачи и приема зашифрованных данных.

Методы шифрования:

- **Шифрование секретным (закрытым) ключом** - посылающая и принимающая сообщения стороны имеют по одному разному секретному ключу соответственно для шифрования и дешифрирования данных. Этот вид ключей используется принятым в 1977 г. стандартом CLUADES (Data Encryption Standard), который предназначен для защиты конфиденциальной информации.

- **Шифрование двойным ключом** - обе стороны используют два одинаковых ключа, один из которых секретный, а другой - открытый. В этой группе методов шифрования наиболее широко известен метод RSA (Rivest – Shamir - Adleman). Пользователь, желающий отправить секретное послание адресату, шифрует данные с использованием открытого ключа конкретного адресата. При получении данных адресат дешифрирует их с помощью секретного ключа, известного только ему. Использование в Интернете этого способа шифрования сообщений считается достаточно надежным. Его недостаток заключается в том, что необходимо быть уверенным, что используется действительно открытый ключ данного конкретного индивида. Обычный способ решения - передача функций генерирования открытых ключей специальной организации, которую называют Certification Authority (Управление сертификации).

- **Симметричное шифрование [symmetric(al) encryption (coding)]** - способ шифрования и дешифрирования отправителем и получателем передаваемых по сетям Интернет/Интранет сообщений, в соответствии с которым обеими сторонами применяется один и тот же ключ, об использовании которого они заранее договорились. Если ключ не был скомпрометирован, то при дешифрировании автоматически выполняется аутентификация отправителя. Так как отправитель и получатель являются единственными сторонами, которые знают этот симметричный ключ, при компрометации ключа будет нанесен ущерб только взаимодействию этих двух пользователей. Общей проблемой, актуальной и для других криптосистем, является вопрос безопасного распространения секретных ключей. Алгоритмы симметричного шифрования используют ключи не очень большой длины и могут быстро шифровать большие объемы данных.

- **Асимметричное шифрование, шифрование с открытым ключом [public key encryption (cryptography), asymmetric(al) encryption (coding)]** - в этой методологии ключи для шифрования и дешифрирования передаваемых в сетях Интернет/Интранет сообщений разные, хотя и создаются они одновременно. Один ключ делается открытым, а другой сохраняется в тайне. Данные, зашифрованные одним ключом, могут быть расшифрованы только другим ключом. Все асимметричные криптосистемы являются объектом атак

путем прямого перебора ключей, и поэтому в них для обеспечения высокого уровня защиты должны использоваться гораздо более длинные ключи, чем в симметричных криптосистемах. Это требует увеличения вычислительных ресурсов для шифрования и дешифрирования сообщений, расшифровки сеансового ключа, а полученный сеансовый ключ - для расшифровки.

• **PKI (Public Key Infrastructure) - "Инфраструктура открытых ключей"**: набор средств и правил управления криптографическими ключами - электронными сертификатами. Последние обеспечивают пользователям возможность доступа к защищенным ими ресурсам через Интернет с любого ПК. Многие годы PKI была прерогативой правительственных учреждений, крупных компаний и финансовых структур; в настоящее время используется более широким кругом организаций разного рода, в частности, создающих и эксплуатирующих так называемые виртуальные частные сети (VPN). В связи с растущим спросом на PKI в мире развивается рынок услуг, обеспечивающий регистрацию клиентов и получение ими сертификатов через Интернет.

ЦИФРОВАЯ ПОДПИСЬ, ЭЛЕКТРОННАЯ ПОДПИСЬ, ЭЛЕКТРОННАЯ ЦИФРОВАЯ ПОДПИСЬ, ЭЦП [digital signature]

Криптографическое средство аналог подписи, позволяющий подтвердить подлинность электронного документа. ЭЦП представляет собой уникальную последовательность битов, жестко связанную с конкретным документом. Ее структура зависит от двух аргументов: текста документа и так называемого **секретного ключа**, известного только его владельцу. Всякое несанкционированное изменение текста документа ведет к искажению цифровой подписи. Считается, что из-за сложной привязки к указанным аргументам электронную подпись невозможно подделать. Наиболее распространенные принципы реализации ЭЦП основаны на использовании **хэш-функций** и, в частности MD5, которая выполняет специальное преобразование битов в файле и генерирует в его конце псевдослучайное 128-битное число, соответствующее количеству бит в записи.

MD5 может использоваться для файлов любой длины. Если два файла отличаются хотя бы на один бит, значение числа MD5 будет иным. И наоборот, если два файла имеют одно и то же характеризующее их число, вероятность того, что это разные файлы - бесконечно мала. После вычисления этого числа, оно шифруется с использованием **открытого ключа** собственника записи в файле, что и представляет собой цифровую подпись. Перед тем, как проверить подпись, пользователь дешифрирует ее с использованием своего открытого ключа. Если полученное значение числа совпадает с исходным, это означает, что запись в файле не изменялась, и цифровая подпись считается аутентичной. Получателю документа для его дешифрирования передается значение открытого ключа.

Достоинством использования MD5 в качестве цифровой подписи является высокая скорость вычислений для больших файлов, однако стопроцентной гарантии защиты дать невозможно. Надежность ЭЦП

находится в прямой зависимости от длины кода, используемого в ключе. Существуют различные алгоритмы ЭЦП. К ним следует отнести отечественные стандарты ГОСТ Р34.10-94 и ГОСТ 28147-89 (обязательные для применения в государственных организациях России), новый ГОСТ Р34.10-2001 (с июля 2002 г. он заменяет ранее действовавшие стандарты), ряд общеизвестных алгоритмов, включая RSA (Rivest – Shamir - Adleman), DSA (Digital Signature Algorithm), Эль Гамала и др.

Криптографическая контрольная сумма [cryptographic checksum] - контрольная сумма, предназначенная для проверки неизменности данных в файлах.

Некоторые термины, связанные с криптографией:

- **Алгоритм шифрования [encryption (cryptography) algorithm]** - математическая формула использования ключей для шифрования документов и сообщений. Для симметричных алгоритмов требуются симметричные ключи, для асимметричных алгоритмов - асимметричные ключи.

- **AES (Advanced Encryption Standard) - "Усовершенствованный стандарт шифрования"**: алгоритм шифрования, предложенный Американским институтом стандартов в качестве общемирового стандарта вместо другого более раннего стандарта - **DES (Data Encryption Standard)**. Он отличается от предыдущего, в частности, более длинным ключом, что затрудняет его раскрытие.

- **CA (Certification Authority) - "Центр сертификатов"**: служба или организация, осуществляющая выпуск криптографических ключей (сертификатов) для индивидуальных пользователей или организаций.

- **DES (Data Encryption Standard) - "Стандарт шифрования данных"**, наиболее распространенный стандарт шифрования данных, разработанный в США в 1977 г.

- **ECB (Electronic Codebook) - "Электронная кодовая книга"**, простейший из режимов работы алгоритмов симметричного шифрования путем отдельного шифрования блоков данных; называется также режимом "простой замены".

- **EFS (Encrypted File System) - "Шифрующая файловая система"**: файловая система (поддерживается в Microsoft Windows 2000 и XP), позволяющая прозрачно шифровать содержимое файлов и каталогов.

- **FEK (File Encryption Key) - "Ключ шифрования файла"**: случайный ключ для шифрования файлов, аналог дискового ключа.

- **Fortezza** - набор стандартов безопасности, разработанный Агентством национальной безопасности США (АНБ); также название аппаратных шифраторов, реализующих данные стандарты.

- **IDEA (International Date Encryption Algorithm) - "Международный алгоритм шифрования данных"**: алгоритм симметричного шифрования, разработан в 1992 г.; несмотря на столь звучное название, имеет существенно меньшее распространение, чем DES.

- **OFB (Output Feedback) - "Обратная связь по выходу"**: один из режимов работы алгоритмов симметричного шифрования; называется также

"гаммированием".

- **PGP (Pretty Good Privacy)** - "Довольно хорошая секретность": алгоритм шифрования данных, разработанный в 1996 г. Филипом Р. Зиммерманном (Philip R. Zimmermann), а также семейство средств стандартного программного обеспечения, используемых для разных операционных систем и приложений, в частности, для шифрования электронной почты и защиты данных, записанных на диски, от несанкционированного чтения (в локальном и интерактивном режимах, в том числе через Интернет). PGP получил широкое распространение и стал международным криптографическим стандартом.

- **PKC (Public Key Cryptography)** - шифрование открытым ключом.

- **PKS (Public Key System)** - криптографическая система с открытым ключом.

- **SCR (Smart Card Reader)** - устройство, осуществляющее запись и считывание данных со смарт-карт.

- **SET (Secure Electronic Transaction)** - "Защищенная электронная транзакция": распространенный протокол защиты электронных платежей.

- **SSET (Simplified Secure Electronic Transaction)** - "Упрощенная защищенная электронная транзакция": протокол защиты электронных платежей, упрощенный по сравнению с SET для ускорений вычислений и возможности использования при ограниченных ресурсах.

- **SSH (Secure Shell)** - "Защищенная оболочка": протокол и программа, позволяющие выполнять Telnet-сеансы с передачей информации в зашифрованном виде.

- **TGS (Ticket-Granting Service)** - "Служба предоставления билета": одна из служб KDS (Центра распределения ключей), предоставляющая клиенту билет, разрешающий доступ к определенному серверу.

- **TGT (Ticket-Granting Ticket)** - "Билет предоставления билета": билет, разрешающий доступ к Службе предоставления билета - TGS.

- **TM (Touch Memory)** - чип, содержащий микросхему энергонезависимой памяти, используется для хранения криптографических ключей.

- **XOR (Exclusive-OR)** - "Исключающее ИЛИ": логическая операция; один из основных криптографических примитивов, используемых в алгоритмах шифрования.

Библиографический список

1. Толковый словарь по вычислительным системам / под ред. В. Иллинуорта и др.; пер. с англ. А. К. Белоцкого и др. - М.: Машиностроение, 1991. - 560 с.

2. Борковский А.Б. Англо-русский словарь по программированию и

Информатике (с толкованием). - М.: Рус. яз., 1990. - 333 с.

3. Толковый словарь по вычислительной технике / пер. с англ. — М.: Рус. редакция. 1995.-478 с.

4. Першиков В. И., Савинков В.М. Толковый словарь по Информатике.- 2-е изд, доп. - М.: Финансы и статистика, 1995. - 543 с.

5. Воройский Ф. С. Информатика. Энциклопедический словарь-справочник: введение в современные информационные и телекоммуникационные технологии в терминах и фактах. - М.: Физмалит, 2006. - 768 с.

6. ГОСТ 15971-90. Система обработки информации. Термины.-Взамен ГОСТ 15971-84; введ. 01.01.92.

7. ГОСТ 34.601-90. Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Стадии создания. - Переизд. 2002 г.-Взамен ГОСТ 24.601-86 и ГОСТ 24.602-86; введ. 01.01.92.

8. ГОСТ 16487-83. Делопроизводство и архивное дело. Термины и определения.- Взамен ГОСТ 16487-70; введ. 01.01.85.

9. ГОСТ 6.10.1-88. Унифицированные системы документации. Основные положения.- Взамен ГОСТ 6.10.2-83; введ. 01.07.89.

10. ГОСТ 7.59-2003. (ИСО 5963-85). Индексирование документов. Общие требования к систематизации и предметизации. - Взамен ГОСТ 7.59-90; введ. 01.01.2004.

11. ГОСТ 7.60-2003. (ИСО 5127-2-83). Издания. Основные виды. Термины и определения.-Взамен ГОСТ 7.60-90; введ. 01.07.2004.

12. ГОСТ 7.27-80. Научно-информационная деятельность. Основные термины и определения; введ. 01.01.82.

13. ГОСТ 7.9-95 (ИСО 21 4-76). Реферат и аннотация. Общие требования.— Взамен ГОСТ 7.9-774; введ. 0107.97.

14. Федеральный закон РФ от 23.09.92 г. №3523-1 "О правовой охране программ для электронных вычислительных машин и баз данных".

15. Федеральный закон РФ от 10.01.2002 г. №1-ФЗ "Об электронной цифровой подписи".

16. Федеральный закон РФ от 21.07.1993 г. №5485-1 "О государственной тайне" (с изменениями от 06.10.1997 г.).

17. Доктрина информационной безопасности Российской Федерации от 09.09.2000 г. №ПР-1895.

18. Тиори Т., Фрай Дж. Проектирование структур баз данных. Кн. 1.- М.: Мир, 1985. - 348 с.

19. Воройский Ф. С. Аналитико-синтетическая обработка и переработка информации в автоматизированных системах НТИ: основы организации и технологии. - М.: ИПКИР, 1991. - 217 с.

20. Бакленд М. Реконструкция (перестройка) библиотечного обслуживания / пер. с англ. И.Ю. Багровой. - М.: Мир , 1996. -129 с.

21. Библиотечные компьютерные технологии: сб. ст. /; сост. О. А. Лавренова. -М. : Рос. гос. б-ка, 1996. - Вып. 1-2.

СОДЕРЖАНИЕ

О пользовании словарем	3
1. Данные, информация, знания, логика	5
2. Информационные ресурсы, теория информации, информатика	10
3. Носители информации, документы, документация, издания	13
4. Принципы структурированного представления документов и данных	19
4.1. Информационные элементы и их виды	-
4.2. Запись, файл, массив, ключ	21
4.3. Структуры, модели данных	26
4.4. Формат, поле данных	40
5. Информационная технология	44
5.1. Общие понятия и термины	-
5.2. Обработка и переработка документов и данных	46
5.3. Ввод документов и данных в ЭВМ	51
5.4. Информационный поиск	55
5.5. Индексирование, поисковый образ документов и запросов	58
6. Безопасность информационной технологии	65
6.1. Общие термины и понятия	-
6.2. Кодирование и декодирование документов и данных	74
6.3. Криптология и связанные с нею понятия	77
Библиографический список	83

Учебное издание

Дмитрий Георгиевич Подобед
Ольга Валентиновна Подобед

ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
(КРАТКИЙ СЛОВАРЬ - СПРАВОЧНИК)

Учебное пособие

Редактор и корректор Н.П.Новикова

Техн. редактор Л.Я.Титова

Тем. план 2010 г., поз. 37

Подп. к печати 24.03.2010 . Формат 60x86/16. Бумага тип. № 1.

Печать офсетная. 5,25 уч.-изд. л.; 5,25 усл. печ. л.

Тираж 50 экз. Изд. № 37 Цена «С». Заказ _____.

Ризограф ГОУВПО Санкт-Петербургского государственного
технологического университета растительных полимеров, 198095, СПб.,
ул. Ивана Черных, 4.

