

**Федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования**

**УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**Факультет математики и информационных технологий  
Кафедра телекоммуникационных технологий и сетей**

***А.А. Смагин, Е.Г. Чекал, С.В. Липатова***

**ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ  
В ИНФОРМАЦИОННЫХ СРЕДАХ**

Учебное пособие

**Ульяновск  
2012**

УДК 001.89(075.8)  
ББК 87.25 я 73  
С 50

*Печатается по решению Ученого совета  
факультета математики и информационных технологий  
Ульяновского государственного университета*

**Рецензенты:**

доктор технических наук, профессор *К.В. Кумунжиев*;  
кандидат физико-математических наук, доцент *В.Р. Воденин*

**Смагин, А. А.**

**С 50 Основы научных исследований в информационных средах : учебное пособие / А. А. Смагин, Е.Г. Чекал, С. В. Липатова. – Ульяновск : УлГУ, 2012. – 192 с.**

В пособие включены учебно-методические материалы для организации самостоятельного научного исследования в области информационных систем и технологий, даны основные определения, представлены методы научных исследований, подробно описаны этапы исследования, даны рекомендации по использованию программного обеспечения, применяемого в научных исследованиях для поиска научных материалов, их статистической обработки, моделирования и организации вычислительного эксперимента, анализа полученных данных. Материал способствует формированию знаний и практических навыков в области исследования в информационных средах.

Пособие адресовано студентам по направлениям «Информационные системы и технологии» и «Инфокоммуникационные системы и системы связи» и предназначено для бакалавров, магистров факультета математики и информационных технологий, а также может быть полезна аспирантам.

УДК 001.89(075.8)  
ББК 87.25 я 73

© Ульяновский государственный университет, 2012  
© Смагин А.А., Е.Г. Чекал, Липатова С.В., 2012

## Оглавление

Введение.....	5
1. Основные положения научных исследований .....	6
1.1. Наука .....	6
1.2. Знания.....	9
1.3. Задача .....	18
1.4. Научное исследование.....	26
1.5. Научно-техническая информация .....	34
1.6. Организация поиска информации .....	44
1.6.1. Поисковые каталоги. ....	45
1.6.2. Поисковые машины или поисковые указатели. ....	46
1.6.3. Классификационно-рейтинговые системы. ....	53
1.6.4. Метапоисковые системы.....	54
1.6.5. Автономные утилиты.....	55
1.6.6. Приемы поиска информации.....	56
1.6.7. Средства простого поиска. ....	57
1.6.8. Средства расширенного поиска. ....	58
1.6.9. Службы поиска людей. ....	60
2. Методы научного исследования в информационной сфере .....	62
2.1. Научные методы познания в исследованиях.....	62
2.2. Моделирование как метод научного познания .....	69
2.2.1. Модели данных.....	73
2.2.2. Модели знаний .....	85
2.2. Интеллектуальные методы извлечения знаний .....	90
2.2.1. Data Mining .....	91
2.2.1.1. Нейронные сети .....	98
2.2.1.2. Статистические методы .....	102
2.2.2. Эвристическое программирование .....	104
3. Этапы научного исследования в информатике .....	107
3.1. Особенности научного исследования в информатике .....	107
3.2. Этапы научного исследования в информатике.....	114
3.3. Формулировка цели и постановка задачи .....	122
3.3.1. Выявление проблемной ситуации и постановка проблемы.....	122
3.3.2. Выбор объекта и предмета исследования .....	124
3.3.3. Актуальность, научная новизна, практическая значимость .....	126

3.3.4. Техничко-экономическое обоснование.....	129
3.3.5. Формулирование целей, задач исследования .....	130
3.4. Выбор методологии и методов исследования.....	131
3.4.1. Поиск, накопление и обработка научной информации .....	131
3.4.1.1. Научные документы и издания.....	131
3.4.1.2. Работа с источниками научной информации .....	136
3.4.2. Выбор методологий, методов и методик исследования.....	138
3.4.2.1. Понятие метода и методологии научных исследований.....	138
3.4.2.2. Методологии структурного анализа.....	141
3.4.2.3. Пример методики исследования .....	147
3.5. Теоретические и экспериментальные исследования.....	151
3.5.1. Теоретические исследования .....	152
3.5.1.1. Математическое моделирование системы .....	153
3.5.1.2. Основные этапы моделирования систем.....	156
3.5.1.3. Постановка цели моделирования систем .....	157
3.5.1.4. Создание концептуальных моделей систем.....	157
3.5.1.5. Разработка математической модели системы.....	161
3.5.1.6. Методы моделирования систем .....	162
3.5.1.7. Выбор средств моделирования систем.....	163
3.5.1.8. Проверка адекватности и корректировка модели системы.....	164
3.5.2. Экспериментальные исследования.....	165
3.5.2.1. Порядок проведения .....	165
3.5.2.2. Виды экспериментов .....	165
3.5.2.3. Математический эксперимент .....	166
3.5.3. Анализ и интерпретация экспериментальных данных .....	169
3.5.4. Обоснование результатов исследования.....	172
3.6. Оформление результатов научной работы.....	174
3.6.1. Публикация результатов исследований .....	174
3.7. Внедрение и эффективность научных исследований.....	177
3.7.1. Внедрение научных исследований .....	177
3.7.2. Эффективность научной работы.....	178
4. Информационная поддержка научных исследований.....	179
4.1. Программы статистической обработки данных.....	182
4.2. Системы поддержки в процессах принятия решений .....	184
Используемая литература .....	191

## Введение

Дисциплина «Основы научных исследований» направлена на формирование знаний, которые помогают решению типовых задач и является составной частью обучения высококвалифицированного специалиста, важным этапом подготовки ученого к самостоятельному ведению научно-исследовательской деятельности. Типовые задачи следующие:

- математическое моделирование объектов и процессов; определение метода их исследования и разработка алгоритма реализации этого метода;
- построение моделей процессов и объектов с целью их анализа и получения наиболее оптимальных параметров;
- составление программ экспериментальных исследований, реализацию этих программ, включая выбор необходимых технических средств, получение и обработку результатов;
- составление отчетов по результатам, полученным в ходе проводимых исследований.

Процесс изучения дисциплины «основы научных исследований» состоит из следующих основных разделов:

1. Методы научного познания.
2. Методы теоретических и эмпирических исследований.
3. Организация научных исследований и их этапы.
4. Процедуры разработки и проектирования новых технических объектов.
5. Теоретические исследования.
6. Построение моделей физических процессов и объектов.
7. Проведение экспериментальных исследований и обработка их результатов.

Для проведения исследований в различных областях науки используются как общие, так и специальные методы, возможные только в определенных науках. Например, основы научных исследований в агрономии будут коренным образом отличаться от методов, которыми такие

исследования проводятся в точных науках. Однако существующие методы исследований можно классифицировать по единой общей классификации:

1. Философские методы исследований, которые можно определить по подразделам:

- объективность;
- всесторонность;
- конкретность;
- историзм;
- диалектический принцип противоречия;

2. Общенаучные методы и подходы.

3. Частнонаучные методы.

4. Дисциплинарные методы.

5. Методы междисциплинарного исследования.

## **1. Основные положения научных исследований**

### **1.1. Наука**

Наука возникает тогда, когда вопрос становится важнее ответов.

Наука — это непрерывно развивающаяся система знаний объективных законов природы, общества и мышления, получаемых и превращаемых в непосредственную производительную силу общества в результате специальной деятельности людей.

Наука – сложная, развивающаяся упорядоченная система знаний, вбирающая в себя способы, методы, подходы, теории к решению проблем окружающего нас мира, объяснению явлений, исследованию – которые служат человечеству для сохранения и увеличения его прогресса.

Цель науки — познание законов развития природы и общества и воздействие на природу на основе использования знаний для получения полезных обществу результатов. Развитие науки идет от сбора факторов, их изучения и систематизации, обобщения и раскрытия отдельных

закономерностей к связанной, логически стройной системе научных знаний, которая позволяет объяснить уже известные факты и предсказать новые.

Все науки обычно делятся на три группы: естественные науки, социальные и гуманитарные науки, формальные науки (рис. 1.1.).

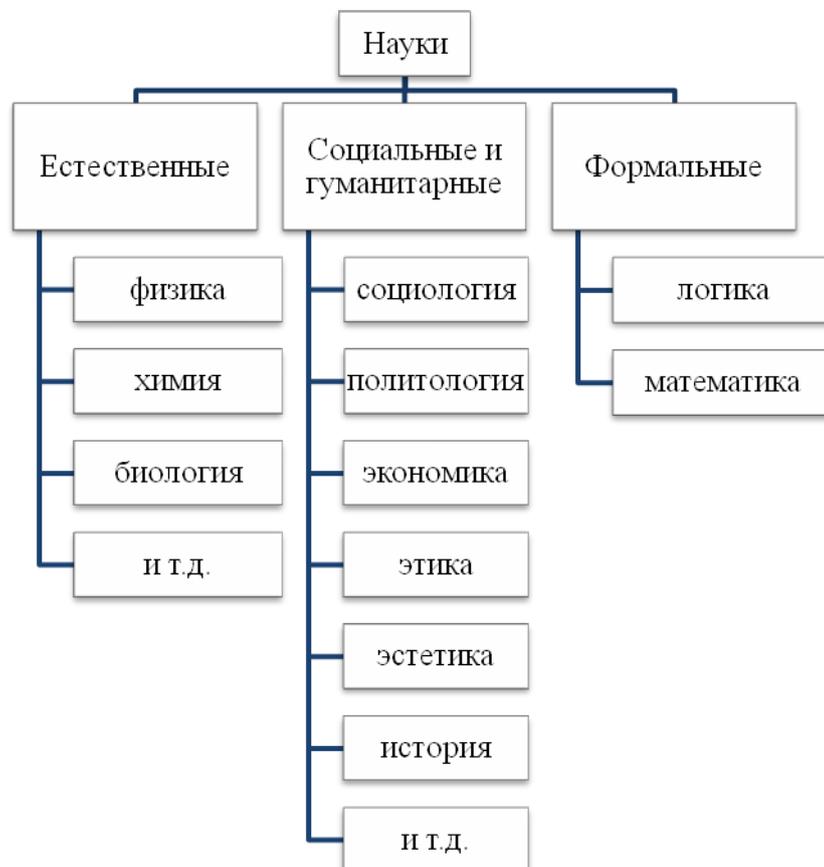


Рис.1.1. Классификация наук

В Российской академии наук несколько иная классификация фундаментальных отраслей науки естествознания. Классификатор имеет три уровня, два из которых приводятся ниже:

- |   |  |
|---|--|
| <p>01. Математика, информатика, механика</p> <p>01-100 Математика</p> <p>01-200 Информатика</p> <p>01-300 Механика</p> <p>02. Физика и астрономия</p> <p>02-100 Ядерная физика</p> <p>02-200 Физика конденсированных сред</p> <p>02-300 Оптика. Квантовая электроника</p> | <p>02-400 Радиофизика, электроника, акустика</p> <p>02-500 Физические основы энергетики</p> <p>02-600 Физика плазмы</p> <p>02-700 Теоретическая физика</p> <p>02-800 Астрономия</p> <p>03. Химия</p> <p>03-100 Органическая химия</p> <p>03-200 Неорганическая химия</p> |
|---|--|

03-300 Высокомолекулярные соединения	05-300 Горные науки
03-400 Физическая химия	05-400 Геофизика
03-500 Динамика и структура атомно-молекулярных систем	05-500 Океанология
03-600 Фундаментальные проблемы формирования материалов	05-600 Физика атмосферы
04. <i>Биология и медицинская наука</i>	05-700 География и гидрология суши
04-100 Общая биология	06. <i>Науки о человеке, природе и обществе</i>
04-200 Физико-химическая биология	06-100 Исторические науки
04-300 Физиология и медицинская наука	06-200 Экономические науки
05. <i>Науки о Земле</i>	06-300 Философские и социальные науки, психология
05-100 Геология	06-400 Лингвистика, филология, культурология, искусствоведение
05-200 Геохимия	

Информатика при этом разделяется на следующие направления:

- 01-201 Математическое моделирование
- 01-202 Искусственный интеллект и принятие решений
- 01-203 Обработка изображений, сигналов и сцен; распознавание образов
- 01-204 Теория оптимизации, исследование операций и оптимальное управление
- 01-205 Математические проблемы теории управления
- 01-207 Вычислительная математика
- 01-208 Теория программирования
- 01-210 Теория игр и статистических решений
- 01-211 Исследование алгоритмов решения прикладных задач
- 01-212 Математическая теория экономического и социального поведения

Процесс познания включает накопление фактов. Без систематизации и обобщения, без логического осмысления фактов не может существовать ни одна наука. Но хотя факты сами по себе еще не наука и факты становятся составной частью научных знаний, когда они выступают в систематизированном, обобщенном виде.

Факты систематизируют и обобщают с помощью простейших абстракций — понятий (определений), являющихся важными структурными элементами науки. Наиболее широкие понятия называют категориями. Это самые общие абстракции. К категориям относятся философские понятия о

форме и содержании явлений, в теоретической экономике — это товар, стоимость и т. д.

Важная форма знаний — принципы (постулаты), аксиомы. Под принципом понимают исходные положения какой-либо отрасли науки. Они являются начальной формой систематизации знаний (аксиомы евклидовой геометрии, постулат Бора в квантовой механике и т. д.).

Важнейшим составным звеном в системе научных знаний являются научные законы, отражающие наиболее существенные, устойчивые, повторяющиеся объективные внутренние связи в природе, обществе и мышлении. Обычно законы выступают в форме определенного соотношения понятий, категорий.

Наиболее высокой формой обобщения и систематизации знаний является теория. Под теорией понимают учение об обобщенном опыте (практике), формулирующее научные принципы и методы, которые позволяют обобщить и познать существующие процессы и явления, проанализировать действие на них разных факторов и предложить рекомендации по использованию их в практической деятельности людей.

## 1.2. Знания

Знания рассматривают

- 1) как специфическую форму общественного сознания;
- 2) как процесс изучения закономерностей объективного мира;
- 3) как определенный вид общественного разделения труда;
- 4) как один из важных факторов общественного.

Знания позволяют адаптироваться и действовать в реальной действительности. Знания можно классифицировать по нескольким критериям (рис. 1.2).

Знание *синтаксического* типа характеризует синтаксическую структуру потока информации, которая не зависит от смысла и содержания

используемых при этом понятий, то есть интеллектуальную систему не образует.

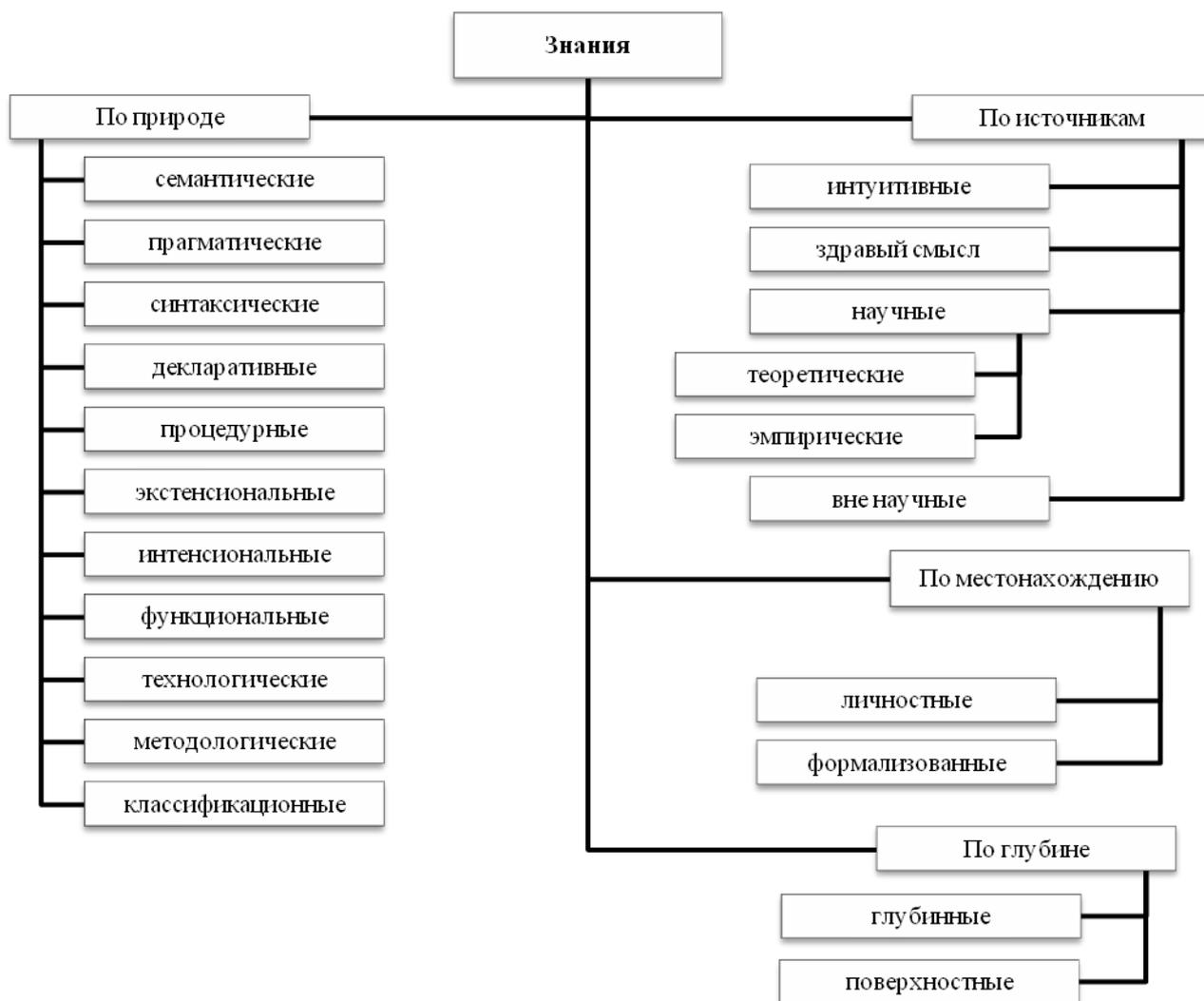


Рис. 1.2. Классификация знаний

*Семантическое* знание рассматривается как структура, образующая текущий контекст. Оно содержит информацию, непосредственно связанную с текущими значениями и смыслом описываемых понятий, и предопределяет состояние связей данных в информационной базе.

*Прагматическое* знание предопределяет наиболее вероятные связи, описывающие данные с точки зрения решаемой задачи (обобщенный или «объективный» контекст), например с учетом действующих в данной задаче специфических критериев и соглашений.

*Декларативные* знания содержат в себе представление о структуре понятий. Эти знания приближены к данным, фактам. Например, высшее учебное заведение есть совокупность факультетов, а каждый факультет в свою очередь есть совокупность кафедр.

*Процедурные* знания имеют активную природу. Они определяют представления о средствах и путях получения новых знаний, проверке знаний. Это алгоритмы разного рода. С развитием информатики все большая часть знаний сосредотачивалась в структурах данных (таблицы, списки, абстрактные типы данных), то есть увеличивалась роль декларативных.

Существенными для понимания природы знаний являются способы определения понятий. Один из широко применяемых способов основан на идее интенционала и экстенционала.

Интенционал понятия – это определение его через соотнесение с понятием более высокого уровня абстракции с указанием специфических свойств.

Экстенционал понятия – это определение понятия через перечисление его конкретных примеров, то есть понятий более низкого уровня абстракции. Интенционалы формируют знания об объектах, в то время как экстенционалы объединяют данные.

Отсюда *интенциональные* знания – это знания о предметной области, которые отражают факты, закономерности, свойства и характеристики, справедливые для любых ситуаций, которые могут возникнуть в этой предметной области.

*Экстенциональные* знания – это знания о предметной области, отражающие факты, закономерности, свойства и характеристики, типичные для конкретных ситуаций или классов однотипных ситуаций, которые могут возникнуть в этой области.

*Функциональные* знания – это знания о выполняемых функциях отдельных предметов и о применении их в реальной действительности.

*Технологические* знания – специализированные знания, обеспечивающие поддержание технологических параметров производства; производственный опыт и навыки, используемые при решении повседневных производственных вопросов. Это может быть знание последовательности операций или знание технологической цепочки, позволяющие достигать поставленные цели в соответствии с принятой технологией.

*Методологические* знания – знания о методах преобразования действительности, научные знания о построении эффективной деятельности. К методологическим знаниям относят знание целей, форм и направлений развития теории, методов и способов эффективного преобразования практики.

*Классификационные* знания применяются главным образом в науке, являются обобщенными, системными знаниями. Пример – система элементов Д. И. Менделеева.

*Интуиция* – это вид знания, специфика которого обусловлена способом его приобретения. Это знание, не нуждающееся в доказательстве и воспринимаемое как достоверное. По способу получения интуиция – это прямое усмотрение объективной связи вещей, не опирающееся на доказательство (интуиция есть усмотрение внутренним зрением; от лат. *intueri* – созерцать).

Под *здравым смыслом* понимают знания, позволяющие принимать правильные решения и делать правильные предположения, основываясь на логическом мышлении и накопленном опыте. В этом значении термин зачастую акцентирует внимание на способности человеческого разума противостоять предрассудкам, заблуждениям, мистификациям.

*Научные* знания в любом случае должны быть основанными на эмпирической или теоретической доказательной основе.

*Теоретические* знания – абстракции, аналогии, схемы, отображающие структуру и природу процессов, протекающих в предметной области. Эти знания объясняют явления и могут использоваться для прогнозирования поведения объектов. Теоретический уровень научного знания предполагает

установление законов, дающих возможность идеализированного восприятия, описания и объяснения эмпирических ситуаций, то есть познания сущности явлений. Теоретические законы имеют более строгий, формальный характер по сравнению с эмпирическими. Термины описания теоретического знания относятся к идеализированным, абстрактным объектам. Подобные объекты невозможно подвергнуть непосредственной экспериментальной проверке.

Эмпирические знания получают в результате применения эмпирических методов познания: наблюдения, измерения, эксперимента. Это знания о видимых взаимосвязях между отдельными событиями и фактами в предметной области. Эмпирические знания, как правило, констатируют качественные и количественные характеристики объектов и явлений. Эмпирические законы часто носят вероятностный характер и не являются строгими.

*Вненаучные* знания могут быть различными. *Паранормальные* знания – знания, несовместимые с имеющимся гносеологическим стандартом. Широкий класс *паранаучного* (пара от греч. около, при) знания включает в себя учения или размышления о феноменах, объяснение которых не является убедительным с точки зрения критериев научности. *Лженаучные* знания – сознательно эксплуатирующие домыслы и предрассудки. В качестве симптомов лженауки выделяют малограмотный пафос, принципиальную нетерпимость к опровергающим доводам, а также претенциозность. Лженаучные знания сосуществуют с научными знаниями.

*Личностные* (неявные, скрытые) знания – это знания людей, полученные из практики и опыта.

*Формализованные* (явные) знания – знания, содержащиеся в документах, на компакт-дисках, в персональных компьютерах, в Интернете, в базах знаний, в экспертных системах. Формализованные знания объективизируются знаковыми средствами языка, охватывают те знания, о которых мы знаем, их можно записать, сообщить другим.

Не всякое знание можно рассматривать как научное. Нельзя признать научными те знания, которые получает человек лишь на основе простого наблюдения. Эти знания играют в жизни людей важную роль, но они не раскрывают сущности явлений, взаимосвязи между ними, которая позволила бы объяснить, почему данное явление протекает так или иначе, и предсказать дальнейшее его развитие.

Правильность научного знания определяется не только логикой, но прежде всего обязательной проверкой его на практике. Научные знания принципиально отличаются от слепой веры, от беспрекословного признания истинным того или иного положения, без какого-либо логического его обоснования и практической проверки. Раскрывая закономерные связи действительности, наука выражает их в абстрактных понятиях и схемах, строго соответствующих этой действительности.

Научное знание оформляется в теории, законы и т.д.

*Научная идея* - интуитивное объяснение явления без промежуточной аргументации, без осознания всей совокупности связей, на основании которой делается вывод. Она базируется на уже имеющемся знании, но вскрывает ранее не замеченные закономерности. Свою специфическую материализацию идея находит в гипотезе.

*Гипотеза* - это предположение о причине, которая вызывает данное следствие. Если гипотеза согласуется с наблюдаемыми фактами, то в науке ее называют теорией или законом. В процессе познания каждая гипотеза подвергается проверке, в результате которой устанавливается, что следствия, вытекающие из гипотезы, действительно совпадают с наблюдаемыми явлениями и, что данная гипотеза не противоречит никаким другим гипотезам, которые считаются уже доказанными. Следует, однако, подчеркнуть, что для подтверждения правильности гипотезы необходимо убедиться не только в том, что она не противоречит действительности, но и в том, она является единственно возможной и с ее помощью вся совокупность наблюдаемых явлений находит себе вполне достаточное объяснение.

С накоплением новых фактов одна гипотеза может быть заменена другой лишь в том случае, если эти новые факты не могут быть объяснены старой гипотезой или ей противоречат. При этом часто старая гипотеза, не отбрасывается целиком, а только исправляется и уточняется. По мере уточнения и исправления гипотеза превращается в *закон*.

*Закон* - внутренняя существенная связь явлений, обуславливающая их необходимое закономерное развитие. Закон выражает определенную устойчивую связь между явлениями или свойствами материальных объектов.

Закон, найденный путем догадки, должен быть затем логически доказан, только тогда он признается наукой. Для доказательства закона наука использует суждения, которые были ранее признаны истинными и из которых логически следует доказываемое суждение. В редких случаях в равной мере оказываются доказуемыми противоречивые суждения. В таких случаях говорят о возникновении *парадокса* в науке, что всегда свидетельствует о наличии ошибок в логике доказательства или несостоятельности исходных суждений в данной системе знаний.

*Парадокс в широком смысле* - это утверждение, резко расходящееся с общепринятым, установившимся мнением, отрицание того, что представляется "безусловно правильным".

*Парадокс в узком смысле* - это два противоположных утверждения, для каждого из которых имеются представляющиеся убедительными аргументы.

Парадоксальность является характерной чертой современного научного познания мира. Наличие парадоксов становится свидетельством несостоятельности существующих теорий, требованием дальнейшего их совершенствования.

Основные пути их разрешения:

- - устранение ошибок в логике доказательств,
- - совершенствование исходных суждений в данной системе знаний.

Для избежания ошибок логика доказательств должна быть подчинена законам формальной логики:

- *закону тождества* (объем и содержание мысли о каком-либо предмете должны быть строго определены и оставаться постоянными в процессе рассуждения о нем),
- *закону противоречия* (в процессе рассуждения о каком-либо определенном предмете нельзя одновременно утверждать и отрицать что-либо в одном и том же отношении, в противном случае оба суждения не могут быть вместе истинными),
- *закону исключения третьего* (в процессе рассуждения необходимо доводить дело до определенного утверждения или отрицания, в этом случае истинным оказывается одно из двух отрицающих друг друга суждений. Этот закон имеет силу лишь при условии соблюдения законов тождества и противоречия),
- *закону достаточного основания* (в процессе рассуждения достоверными следует считать лишь те суждения, относительно истинности которых могут быть приведены достаточные основания).

*Теория* (от лат. *theoreo* - рассматриваю) - система обобщенного знания, объяснения тех или иных сторон действительности. Теория является духовным, мысленным отражением и воспроизведением реальной действительности. Она возникает в результате обобщения познавательной деятельности и практики. Это обобщенный опыт в сознании людей.

Структуру теории формируют *принципы, аксиомы, законы, суждения, положения, понятия категории и факты.*

Под *принципом* в научной теории понимается самое абстрактное определение идеи (начальная форма систематизации знаний). Принцип – это правило, возникшее в результате субъективно осмысленного опыта людей.

*Факт* - в широком смысле может выступать как синоним истины; событие или результат; реальное, а не вымышленное; конкретное и единичное в противоположность общему и абстрактному.

*Фактор* - причина, движущая сила какого-либо процесса, определяющая его характер или отдельные его черты

*Аксиома* (постулат) - это положение, которое берется в качестве исходного, недоказуемого в данной теории, и из которого выводятся все остальные предложения и выводы теории по заранее фиксированным правилам. Аксиомы очевидны без доказательства. В современной логике и методологии науки постулат и аксиома обычно используются как эквивалентные.

Теория является наиболее развитой формой обобщенного научного познания. Она включает в себе не только знания основных законов, но и объяснение фактов на их основе. Теория позволяет открывать новые законы и предсказывать будущее.

*Методология* – философское учение о методах познания и преобразования действительности, применение принципов мировоззрения к процессу познания, духовному творчеству и практике.

В методологии выявляются две взаимосвязанные функции:

- 1) обоснование правил применения мировоззрения к процессу познания и преобразования мира;
- 2) определение подхода к явлениям действительности.

Диалектическая методология всегда опирается на конкретные знания. Исследователь должен иметь определенный запас знаний и уметь применять ее к решению конкретных задач.

*Объект исследования* в общем случае - это структура (подразделение, предприятие, объединение предприятий, отрасль, национальное хозяйство), ее внутренняя и внешняя среда, системы (социальные, экономические, технические, организационные, производственные, научные, политические, культурные, кадровые и др.), совокупность их элементов.

*Внутренняя* среда может характеризоваться составом элементов объекта исследования: ресурсных (материально-техническая база, включающая предметы и средства труда, трудовые ресурсы, информация, финансовые ресурсы), организационных (технология, методы и системы управления, организационная структура), результатов функционирования объекта, например в виде продуктов и услуг. Состояние внутренней среды объекта может оцениваться также ее потенциалом. Информация о внутренней среде необходима исследователям для уточнения целей организации (в том числе социальных), определения внутренних возможностей и потенциала, на которые организация может рассчитывать в конкурентной борьбе.

*Внешняя* среда включает окружение исследуемого объекта, т.е. все то, что не входит непосредственно в него, но с ним взаимодействует и на него влияет.

Среда (мир) – множество связей между объектом и множеством внешних по отношению к исследуемому других объектов (процессов).

### **1.3. Задача**

Процесс решения задачи представляет собой поиск выхода из затруднения или пути обхода препятствия – это процесс достижения цели, которая первоначально не кажется доступной.

Задачи нет в том случае, когда лицо, принимающее решение (ЛПР), не может задать требования, сформулировать критерий достижения цели или неизвестен набор средств достижения цели, т. е. имеет место задача с неопределённостью.

При постановке рассматриваемой задачи может быть учтены не только обязательные основные требования, отражаемые с помощью критерия, но и дополнительные требования, которые могут выступать в качестве ограничений (например, ограничения по затратам).

Корректно поставленная задача.

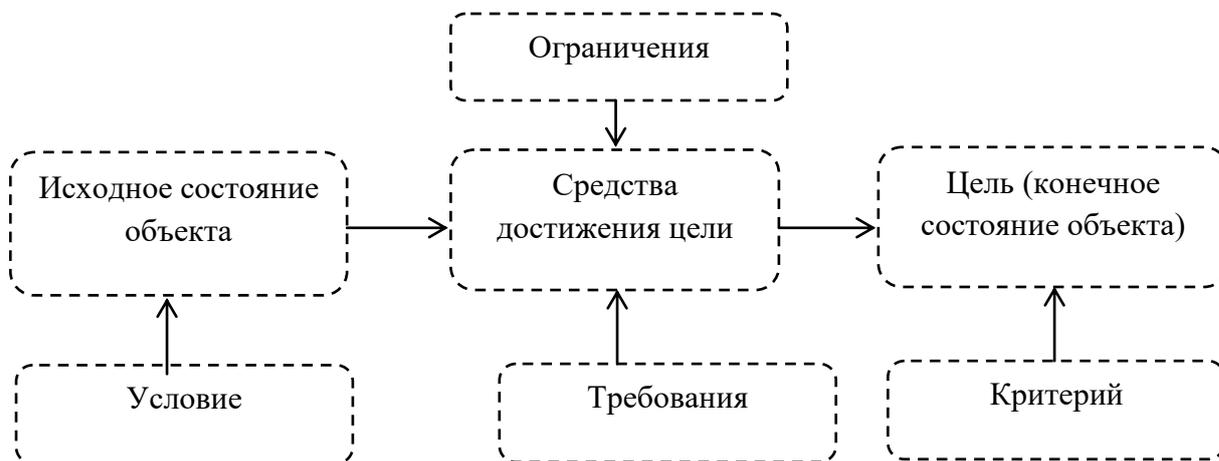
1. Задача имеет решение при любых допустимых исходных данных (существенное решение).
2. Каждым исходным данным соответствует только одно решение (однозначность задачи).
3. Решение устойчивое.

Наиболее существенная черта, отличающая корректно поставленные задачи от некорректно поставленных заключается в том, что первые немедленно сводятся к чисто математическим задачам, вторые не могут быть сведены к ним.

Под решением задачи понимается поиск множества связанных процедур (операций), позволяющих обеспечить достижение поставленной цели (получение результата в рамках заложенных (предписанных) условий при выполнении заданных требований и ограничений).

Под поиском множества связанных процедур (операций) понимается приведение формулировки некоторой задачи к виду, допускающему решение задачи известным способом.

Задача состоит из нескольких элементов (рис. 1.3.).



*Рис.1.3. Состав задачи*

Условие – требуемое соотношение между неизвестными и данными.

Критерии отражают требования и ограничения.

Конечная цель подсказывает средства: изучение цели (неизвестного, заключения) может найти подход к решению задачи.

Задачу можно разделить на 2 типа: на нахождение и на доказательство.

Задача	
На нахождение	На доказательство
некоторого объекта, т. е. неизвестного данной задачи	т. е. установление правильности или логичности некоторого утверждения, подтверждения его или опровержения

*Рис. 1.4. Виды задач.*

При решении задачи необходимо ответить на следующие вопросы:

- К какому типу относится рассматриваемая задача? Не родственна ли она какой-нибудь другой известной задаче? Не похожа ли она на какую-нибудь знакомую задачу?
- Не могли бы вы сформулировать задачу иначе?
- Какие факты являются здесь ключевыми?
- Приняли ли вы во внимание полностью все условия? Испытали ли вы все данные? Понятны ли вам все понятия, являющиеся неотъемлемыми частями задачи?

Задача тесно связана с целью. Задача является шагом к достижению цели.

Цель – предмет мотивации (побуждения к деятельности), направленный на устранение дискомфорта, неопределённости и т. п.

Цель – выражение активности сознания от идеальных устремлений к материальному воплощению, конечному результату деятельности.

Следует отметить, что понятие цели трактуется во многих литературных источниках неоднозначно, например:

- желаемое будущее состояние;
- идеальное представление желаемого результата деятельности;
- то, что представляется в сознании и ожидается в результате определенным образом направленных действий и др.

Таким образом, термин «цель» следует воспринимать как желаемое и выраженное:

- 1) количественно (сколько);

- 2) качественно (что) будущее состояние объекта, имеющего;
- 3) срок достижения (когда);
- 4) ответственного исполнителя (кто);
- 5) ограничения по ресурсам (чем).

Представление целей:

- сетевой график,
- иерархия,
- страты (эшелоны),
- матричная (табличная) форма.

Виды критериев:

- критерий функционирования,
- критерий эффективности,
- целевая или критериальная функция.

Чётко сформулированная задача должна указать (указывать) категорию (множество), к которой принадлежит неизвестное.

Чётко сформулированная задача должна точно устанавливать условие, которому обязано удовлетворять неизвестное.

Во множестве объектов, характеризующихся условием задачи, к которым должно принадлежать неизвестное, содержится подмножество тех объектов, которые должны удовлетворять этому условию, и каждый объект, принадлежащий этому подмножеству называется решением.

Это подмножество может содержать один-единственный объект. И так же решение будет единственным.

*Ситуация* – состояние объекта (объектов) или процесса (процессов) в определённый момент времени. Характеризуется параметрами или свойствами объекта (процесса), зависящими от влияния внешних и внутренних факторов (или воздействий). При этом нет цели, но есть средства достижения цели.

Для разрешения ситуации необходимо:

1. выделить наиболее значимые (релевантные) факторы, влияющие на объект;
2. выделить множество наиболее важных состояний системы и переходов для уточнения цели.

Сложность ситуации:

1. Неполнота и противоречивость данных
2. Неоднозначность восприятия
3. Плохая идентификация ситуации
4. Отсутствие очевидного критерия её оценки

В ситуациях с большой неопределённостью вначале выбирают подход к постановке и решению проблемы:

Подход – синтезированное выражение мысли, обозначающее направление движения к цели и некоторую (связующую) крупнэтапную последовательность действий.

Для решения проблемы необходимо использовать подход «сверху-вниз» (методы структуризации, декомпозиции) (целевой или целенаправленный подход) или «снизу-вверх» (морфологический, лингвистический, тезаурусный, терминальный) – правила композиции, поиска мер близости на пространстве состояний этапов системы

Существует плохая формализуемость задачи: например, из-за того, что критерии достижения цели принципиально не формализуемые (комфорт, нечеткая логика – открытие форточки на некоторый угол, т. е. количество). Задачу, представленную в общем виде, например вербальным образом, нужно формализовать. Одно из определений формализации состоит в уточнении содержания с применением какой-либо устной новой конструкции, и в частности – математического аппарата.

Формализация – это переход от интуитивного ясного представления задачи к строгим математическим выкладкам.

Проблема связана с проблемной ситуацией – известна цель и есть не подходящие для достижения цели средства.

При представлении объекта в виде хорошо детализованной системы проблемная ситуация может быть описана в виде выражений, связывающих цель со средствами.

Возникновение проблем и их решение естественный процесс поступательного развития производительных сил. Он связан с необходимостью постоянного получения новых знаний, требующихся для решения конкретных вопросов материального производства. Таким образом, проблема - это форма выражения необходимого развития научного познания. Она является отражением объективного противоречия между знанием и незнанием.

Проблемы возникают не сами по себе, а как следствие практики, как результат насущной необходимости и обуславливаются определенными условиями развития техники и уровня знаний.

Решение поставленной жизнью задачи (именно так переводится с греческого слово "проблема") требует, в конечном итоге, нового образа мышления, основанного на интеллектуальном подходе и системном анализе явлений.

Альберт Эйнштейн говорил: "Наука должна начинаться с фактов и заканчиваться ими, вне зависимости от того, какие теоретические структуры строятся между началом и концом".

Всякое неизвестное начинается с известного и в процессе исследования неизвестное переходит свою противоположность. В этом суть процесса познания.

Изучая те или иные объекты, в результате взаимодействия с ними мы, прежде всего, собираем информацию, анализ и синтез которой дает нам необходимое знание. Вместе с тем в каждом вопросе или проблеме существует всегда противоречивое отношение двух видов знаний - знание сущности и знание о незнании.

В знании сущности, обычно, ищут ответ на два вопроса:

- что общего у исследуемого объекта с другими, уже известными исследователю объектами?

- чем отличается исследуемый объект от других, уже изученных объектов?

Соотношение этих двух частей знания сущности исследуемого объекта характеризует собой степень проблемности исследования. Математически это понятие может быть представлено следующим соотношением:

$$K = x/(x + a),$$

где  $K$  - так называемый, коэффициент проблемности;  $x$  - неизвестное об объекте;  $a$  - известное об объекте.

Если  $K=0$ , то следует понимать, что проблема отсутствует, а решение поставленной задачи может осуществляться инженерными методами, т.е. в этом случае используется найденная ранее теория.

Если  $0 < K < 1$  - это показывает на существование проблемы. Вряд ли возможна крайность, чтобы  $K = 1$ , т.к. трудно представить себе объект, о котором ничего не было бы известно. (Это противоречит сказанному выше о том, что проблема начинается с практики). Поэтому, крайне важно на первом этапе исследования выделить и систематизировать все известное и неизвестное об изучаемом объекте.

Проблемы в буквальном смысле окружают нас. Одни из них лежат, как говорят, на поверхности, а другие не видимы, их надо искать. Проблемы, как правило, не разделены удобным образом - на технические, экономические, экологические, социальные. В реальной жизни все переплетено и представляет некоторую "мешанину". В ней надо разобраться и суметь выделить проблему в соответствии с профилем собственной подготовки.

Постановка проблемы предполагает организацию исследования, конечной целью которого явилась бы разработка новой закономерности, необходимой для построения инженерных методик решения производственных задач. Способность постановки проблемы является самой яркой характеристикой творческого мышления.

Но наличие проблемы не является достаточным основанием для постановки научных исследований, т.к. не каждая проблема является истинной. Можно выделить следующие критерии для оценки истинности проблемы:

- невозможность дальнейшего развития практики без решения данной проблемы;
- невозможность получения практикой результатов поставленной проблемы;
- получение после решения проблемы результатов, имеющих большую практическую значимость.

Таким образом, главным и основным критерием ценности проблемы является практика. Практика определяет цель научного исследования, а выбор цели в значительной степени зависит от отношения к проблеме.

Изучение проблемы приводит к ее разворачиванию, т.е. возникновению и формированию дополнительных вопросов. Их решение в итоге дает ответ на центральный вопрос проблемы.

В вопросах проявляются различные аспекты проблемы, которые можно представлять как более мелкие научные задачи, относящиеся к конкретной области научного исследования. При этом они могут в некоторых случаях рассматриваться как отдельные темы исследований, а иногда и как самостоятельные проблемы.

Во многих школах на угадывание положено табу, тогда как в любом научном исследовании «сначала угадайте, а потом докажите» - это почти что правило.

Во многих задачах наблюдения, предположения, индуктивные умозаключения, короче – правдоподобные рассуждения играют важнейшую роль.

Метод решения хорош, если с самого начала мы можем предвидеть – и далее подтвердить это, - что следуя этому методу мы достигнем цели.

Сложность принимаемого решения зависит от:

1. Времени, затрачиваемое на его получение (колмогорова сложность),

2. Длины решения,
3. Стоимости реализации (количество ресурсов).

Качество принимаемого решения зависит от:

1. Сложности принимаемого решения,
2. Вариантов (альтернатив) (выбор  $\leftarrow$  результат),
3. Времени реализации решения.

Главным в принимаемом решении – смысл и, особенно, он касается нештатных ситуаций, чем сложнее оно – тем больше требуется новизны (идеи), экстравагантности, неожиданности в принимаемом решении.

#### **1.4. Научное исследование**

Научное исследование представляет собой процесс изучения какого-либо объекта и получения новых знаний. Его также можно рассматривать как научный труд, вид познавательной деятельности; научное изучение некоего предмета, какого-либо явления (объекта) с целью определения законов и закономерностей его возникновения, функционирования, совершенствования, развития, особенностей и тенденций перехода из одного состояния в другое, получения и применения новых знаний в теории и на практике.

Исследование обладает *основополагающими характеристиками*, определяющими его направленность и результаты. К таким характеристикам в первую очередь относят: потребность в исследовании (острота и необходимость решения проблем и задач); его цель, объект и предмет; методология, вид исследования, ресурсы (определенный комплекс средств и возможностей, обеспечивающих успешное проведение исследования и достижения его целей); результаты исследования (как конечный итог и эффективность исследования, определяющая соотношение и соразмерность использованных ресурсов на проведение исследования и достигнутых при этом

Научно-исследовательская деятельность – это такая деятельность, которая направлена на получение новых знаний и их практическое

применение. Несмотря на то, что науки классифицируются в зависимости от сферы познания, предмета и метода познания, основы научных исследований составляют неотъемлемую часть любой науки.

Понятием «научное исследование» определяется деятельность, которая направлена на всестороннее изучение исследуемого объекта, явления или процесса, их внутренней структуры и связей, получение на этой основе и внедрение в практику полезных результатов для человеческого существования. Для того чтобы научные специалисты могли правильно проводить при изучении науки необходимые научные исследования (практически во всех высших целях) и т.п.

При проведении исследования важно правильно и четко определить цель, объект и предмет исследования. Большое значение имеет также сам процесс исследования, представляющий собой совокупность последовательно выполняемых операций.

Любое исследование эффективно лишь тогда, когда достигнуты установленные цели исследовательских работ при соблюдении других условий (сроков и затрат). В связи с этим первостепенное значение имеет тот смысл, что заложен в содержание понятия "цель".

Комплексное обоснование цели в условиях ужесточения конкуренции приобретает ключевое значение, поскольку формулирование цели без необходимых обоснований может привести к потерям на стадии ее реализации, во много раз превышающим экономию, полученную ранее. Кроме того, правильно сформулированные цели могут выступать как эффективный инструмент исследования.

Применительно к исследованию наиболее предпочтительно рассматривать цель как желаемый новый исследовательский результат состояния предмета определенного объекта исследования.

Наука включает в себя также методы. В настоящее время все большее значение приобретает в качестве общего математический метод исследования, т.е. метод количественного изучения явлений и процессов. Это

обусловлено бурным развитием кибернетики, вычислительной математики и ЭВМ.

Когда ученые не располагают достаточным фактическим материалом, то в качестве средства достижения научных результатов они используют гипотезы, выдвигаемые для объяснения какого-либо процесса, которые после проверки могут оказаться истинными или ложными. Гипотеза часто выступает как первоначальная формулировка, черновой вариант открываемых законов.

Научное исследование является формой существования и развития любой науки. Изучение с помощью научных методов явлений и процессов, анализ влияния на них различных факторов, а также изучение взаимодействия между явлениями с целью и позволяет получить убедительно доказанные и полезные для науки и практики решения с максимальным эффектом. Основой разработки каждого научного исследования является методология.

Эмпирические задачи направлены на выявление, точное описание и тщательное изучение различных факторов рассматриваемых явлений и процессов. В научных исследованиях они решаются с помощью различных методов познания - наблюдением и экспериментом.

Наблюдение — это метод познания, при котором объект изучают без вмешательства в него; фиксируют, измеряют лишь свойства объекта, характер его изменения.

Эксперимент — это наиболее общий эмпирический метод познания, в котором производят не только наблюдения и измерения, но и осуществляют перестановку, изменения объекта исследования и т. д. -В этом методе можно выявить влияние одного фактора на другой. Эмпирические методы познания играют большую роль в научном исследовании. Они не только являются основой для подкрепления теоретических предпосылок, но часто составляют предмет нового открытия, научного исследования. Теоретические задачи направлены на изучение и выявление причин, связей, зависимостей,

позволяющих установить поведение объекта, определить и изучить его структуру, характеристику на основе разработанных в науке принципов и методов познания. В результате полученных знаний формулируют законы, разрабатывают теорию, проверяют факты и др. Теоретические познавательные задачи формулируют таким образом, чтобы их можно было проверить эмпирически.

В решении эмпирических и сугубо теоретических задач научного исследования важная роль принадлежит логическому методу познания, позволяющему на основе умозаключительных трактовок объяснять явления и процессы, выдвигать различные предложения и идеи, устанавливая пути их решения. Этот метод базируется на результатах эмпирических исследований.

Результаты научных исследований должны создавать основу для новых научных разработок.

Одним из важнейших требований, предъявляемых к научному исследованию, является научное обобщение, которое позволит установить зависимость и связь между изучаемыми явлениями и процессами и сделать научные выводы. Чем глубже выводы, тем выше научный уровень исследования.

По целевому назначению научные исследования бывают теоретические и прикладные.

Теоретические исследования направлены на создание новых принципов. Это обычно фундаментальные исследования. Цель их — расширить знания общества и помочь более глубоко понять законы природы. Такие разработки используют в основном для дальнейшего развития новых теоретических исследований, которые могут быть долгосрочными, бюджетными и др.

Прикладные исследования направлены на создание новых методов, на основе которых разрабатывают новое оборудование, новые машины и материалы, способы производства и организации работ и др. Они должны удовлетворять потребность общества в развитии конкретной отрасли

производства. Прикладные разработки могут быть долгосрочными и краткосрочными, бюджетными или хоздоговорными.

Каждое научное исследование имеет тему. Темой могут быть различные вопросы науки и техники. Обоснование темы — это важный этап в разработке научного исследования.

Темы могут быть различными, их можно классифицировать следующим образом:

а) по видам связи с общественным производством — научные исследования, направленные на создание новых процессов, машин, конструкций и т. д., полностью используемых для повышения эффективности производства;

- научные исследования, направленные на улучшение производственных отношений, повышение уровня организации производства без создания новых средств труда;
- теоретические работы в области общественных, гуманитарных и других наук, которые используются для совершенствования общественных отношений, повышения уровня духовной жизни людей и др.;

б) по степени важности для народного хозяйства

- — работы, выполняемые по заданию министерств и ведомств;
- исследования, выполняемые по плану (по инициативе) научно-исследовательских организаций;

в) в зависимости от источников финансирования

- госбюджетные, финансируемые из средств государственного бюджета;
- хоздоговорные, финансируемые в соответствии с заключаемыми договорами между организациями-заказчиками, которые используют научные исследования в данной отрасли, и организациями, которые выполняют исследования;

г) по длительности разработки: долгосрочные, разрабатываемые в течение нескольких лет;

- краткосрочные, выполняемые обычно за один год
- долгосрочные.

Выбрать тему зачастую более сложно, чем провести само исследование. К теме предъявляют ряд требований.

Тема должна быть актуальной, т. е. важной, требующей разрешения в настоящее время. Это требование одно из основных. Критерия для установления степени актуальности пока нет. Так, при сравнении двух тем теоретических исследований степень актуальности может оценить крупный ученый данной отрасли или научный коллектив. При оценке актуальности прикладных научных разработок ошибки не возникают, если более актуальной окажется та тема, которая обеспечит большой экономический эффект.

Тема должна решать новую научную задачу. Это значит, что тема в такой постановке никогда не разрабатывалась и в настоящее время не разрабатывается, т. е. дублирование исключается. Дублирование возможно только в том случае, когда по заданию руководящих организаций одинаковые темы разрабатывают два конкурирующих коллектива в целях разрешения важнейших государственных проблем в кратчайшие сроки. Таким образом, оправданное дублирование тем (разработок) иногда может быть одним из требований.

Грань между научными и инженерными исследованиями с каждым годом все более стирается. Однако при выборе тем новизна должна быть не инженерной, а научной, т. е. принципиально новой. Если разрабатывается пусть даже новая задача, но на основе уже открытого закона, то это область инженерно-экономических, не научных разработок. Поэтому необходимо отличать научную задачу от инженерно-экономической. Все то, что уже известно, не может быть предметом научного исследования.

Тема должна быть экономически эффективной и должна иметь значимость. Любая тема прикладных исследований должна давать

экономический эффект в народном хозяйстве. Это одно из важнейших требований.

На стадии выбора темы исследования ожидаемый экономический эффект может быть определен, как правило, ориентировочно. Иногда экономический эффект на начальной стадии установить вообще нельзя. В таких случаях для ориентировочной оценки эффективности можно использовать аналоги (близкие по названию и разработке темы).

При разработке теоретических исследований требование экономичности может уступать требованию значимости. Значимость, как главный критерий темы, имеет место при разработке исследований, определяющих престиж отечественной науки или составляющих фундамент для прикладных исследований, или направленных на совершенствование общественных и производственных отношений и др.

Тема должна соответствовать профилю научного коллектива. Каждый научный коллектив по сложившимся традициям имеет свой профиль, квалификацию, компетентность. Такая специализация, способствующая накоплению опыта исследований, дает свои положительные результаты, повышается теоретический уровень разработок, качество и экономическая эффективность, сокращается срок выполнения исследования. Однако нельзя впадать в крайность, применяя этот принцип. Если допускать монополию в науке, то исключается соревнование идей. Это может снизить эффективность научных исследований. Заказчику будет предоставляться научная продукция, которая не всегда может отражать наилучшие показатели.

Выполняя длительное время работу по узкоспециализированной тематике с устоявшейся методикой, некоторые научные работники теряют к ней интерес. Поэтому в коллективе может быть несколько (до 10%) непрофильных тем, не отличающихся резко от основной тематики коллектива. Это может вызвать энтузиазм, инициативу и прилив творческих сил в коллективе.

Важной характеристикой темы является ее осуществимость или внедряемость. При разработке темы следует оценить возможность ее окончания в плановый срок и внедрения в производственных условиях заказчика. Если это нельзя осуществить вообще или осуществить в сроки, которые не устраивают заказчика, то заведомо планируют разработку бросовых, неэффективных тем.

Существенно упрощается методика выбора тем в научном коллективе, имеющем научные традиции (свой профиль) и разрабатывающем комплексную проблему. В таких коллективах научные исследования выполняют не одиночки, а группы, специализирующиеся на разработке тем или вопросов. Вероятность получить не актуальную, не новую, не эффективную тему исключена. При коллективной разработке научных исследований большую роль приобретают критика, дискуссия, обсуждение проблем и тем. В процессе дискуссии выявляются новые, еще не решенные актуальные задачи разной степени важности, объема, сроков разработки.

Большое значение для выбора тематики имеет четкая формулировка общих задач заказчиком.

Научный руководитель коллектива должен с большим вниманием отнестись к предложениям сотрудников, которые могут выставить ряд тем и вопросов. Перед окончательным решением целесообразно организовать широкую дискуссию.

Важное значение при разработке общей программы исследования имеет выделение долгосрочных и краткосрочных исследований, фундаментальных и прикладных. Соотношение между ними зависит от многих факторов — требований заказчика, научного потенциала коллектива, наличия современного экспериментального оборудования, научного задела коллектива и его работоспособности и т. д.

Приведенные выше требования (критерии), предъявляемые к выбору тем, позволяют всесторонне оценить и установить пригодность их для данной научно-исследовательской организации.

Однако в период бурной НТР в процессе разработки тем, особенно долгосрочных, актуальность их и экономичность иногда могут изменяться в худшую сторону, а так как затраты на выполнение НИР возросли, то очень важным критерием при выборе тем является их перспективность, а следовательно, стабильность.

В данном случае одних субъективных методов оценки недостаточно. Первостепенное значение приобретают численные методы.

Для оценки перспективности тем применяют два метода — математический и экспертных оценок.

Математический метод основан на использовании различных показателей, определяющих перспективность исследований. Наиболее часто в прикладных темах применяют показатель перспективности, в основе которого лежат экономические показатели: В последние годы при выборе тем все шире применяют методы экспертных оценок. Суть этого метода заключается в том, что планируемую тему оценивают специалисты-эксперты. Каждому эксперту выдается оценочная балльная шкала, с помощью которой он устанавливает баллы по теме.

После ответа экспертов на вопросы результаты обрабатывают различными методами. Наиболее простым является метод максимального балла — отдают предпочтение той теме, которая набирает наибольший суммарный балл. В данном случае тема является перспективной, если сумма баллов положительна.

### **1.5. Научно-техническая информация**

Характерной чертой развития современной науки является бурный поток новых научных данных, получаемых в результате исследований. Ежегодно в мире издается более сотни тысяч книг по различным вопросам. Еще больше издается журналов. Но, несмотря на это, огромное количество научно-технической информации остается неопубликованной.

Носителями информации могут быть различные документы:

- книги (учебники, учебные пособия, монографии);
- периодические издания (журналы, бюллетени, труды институтов, научные сборники);
- нормативные документы (стандарты, СНИПы, ТУ, инструкции, временные указания, нормативные таблицы и др.);
- каталоги и прейскуранты;
- патентная документация (патенты, изобретения);
- отчеты о научно-исследовательских и опытно-конструкторских работах;
- информационные издания (сборники НТИ, аналитические обзоры, информационные листки, экспресс-информация, выставочные проспекты и др.);
- переводы иностранной научно-технической литературы;
- материалы научно-технических и производственных совещаний;
- диссертации, авторефераты;
- производственно-техническая документация организаций (отчеты, акты приемки работ и др.);
- вторичные документы (реферативные обзоры, библиографические каталоги, реферативные журналы и др.).

Эти документы создают огромные информационные потоки, темпы которых ежегодно возрастают.

Различают восходящий и нисходящий потоки информации.

Восходящий — это поток информации от пользователей в регистрирующие органы. Вся научно-техническая информация регистрируется в Республиканском институте НТИ.

Исполнитель научно-технической работы (НИИ, вузы и др.) после утверждения плана работ обязан в месячный срок представить информационную карту в Республиканский институт НТИ. К восходящему потоку относят также статьи, направленные в различные журналы.

Нисходящий — это поток информации в виде библиографических обзорных реферативных и других данных, который направляется в низовые организации по их запросам.

Сбор, хранение и выдачу информации осуществляют справочно-информационные фонды (СИФ). В стране имеются отраслевые, республиканские и местные (в НИИ, вузах, ОКБ и т. д.) СИФ.

В СИФ установлен определенный порядок хранения информации. Имеется основной и справочный фонды.

Справочный фонд — это вторичные информационные документы основного фонда. Он представлен в основном библиографическими и реферативными карточками (125x75 мм), хранимыми в каталожных выдвижных ящиках.

Справочный фонд состоит из главной картотеки (содержащей все опубликованные и неопубликованные документы, хранимые в данном СИФ), каталогов и карточек.

По алфавитному каталогу можно отыскать любую информацию в данном СИФ по фамилии автора, редактора или по названию первоисточника.

По систематическому каталогу можно подбирать информацию для различных отраслей знаний. Для ускорения отыскания нужной информации к каталогу прилагается ключ — алфавитный предметный указатель.

В регистрационной картотеке периодических изданий содержатся сведения о журналах, сборниках, бюллетенях, хранимых в данном СИФ (по годам и номерам).

Патенты и авторские свидетельства можно отыскать в картотеке описаний изобретений.

Картотека стандартов содержит различные нормативные документы — стандарты, нормы, ТУ, временные указания и пр.

Поиск нужной информации с каждым годом усложняется. Поэтому все научные работники должны знать основные положения, связанные с информационным поиском.

Информационный поиск — это совокупность операций, направленных на отыскание документов, которые необходимы для разработки темы. Поиск может быть ручной (осуществляется по обычным библиографическим карточкам, картотекам, печатным указателям), механический (носителем информации являются перфокарты), механизированный (основан на применении счетно-перфорационных машин) и автоматизированный (применение ЭВМ).

Информационный поиск осуществляется с помощью информационно-поискового языка (ИПЯ) — семантической (смысловой) системы символов и правил их сочетания. В информационно-поисковой системе применяют различные варианты ИПЯ. В настоящее время наибольшее распространение получила универсальная десятичная классификация документов информации (УДК).

УДК разделяет все области знаний на десять отделов, каждый из которых делится на десять подразделов, а подраздел — на десять частей. Каждая часть детализируется до требуемой степени. Структура УДК состоит из групп основных индексов и определителей. Группы делятся на подгруппы общих и специальных определителей.

УДК просто усваивается работниками издательств и библиотек, удобно шифруется, обладает относительно быстрым поиском информации для узкоспециализированных тем.

В последние годы все чаще применяются механизированная и автоматизированная системы поиска, которые устраняют громоздкость системы УДК.

На российском рынке сведения об информационных ресурсах предоставляются потребителям следующими организациями:

Центр "Информрегистр " ведет следующие электронные каталоги и выпускает справочники:

Электронный каталог Государственного регистра баз данных, который содержит описание зарегистрированных российских баз данных;

- Каталог "Базы данных России" – справочник;
- Каталог "Где найти адрес?";
- Каталог "Российские электронные издания, который содержит описания свыше 300 отечественных электронных изданий, выпускаемых на оптических дисках и дискетах.

ООО "Международное Бюро Информации и Телекоммуникаций"– предлагает на информационном рынке:

- "Российская энциклопедия информации и телекоммуникаций", электронная версия находится на сайте "Инфопартнер";
- Web-каталог деловой информации;
- Web-каталог бизнес-ресурсов по рынку товаров и услуг;
- Информационно-поисковые, аналитические и консультационные услуги.

Обзорную информацию о состоянии информационных ресурсов России публикует ежемесячный журнал "Информационные ресурсы России", выпускаемый Российским объединением информационных ресурсов научно-технического развития.

К государственным информационным ресурсам, ориентированным на внешнего пользователя могут быть отнесены:

- ✓ библиотечная сеть РФ;
- ✓ архивный фонд РФ;
- ✓ государственная система статистики;
- ✓ государственная система научно-технической информации.

Библиотечная сеть РФ насчитывает свыше 150 тыс. библиотек. На федеральном уровне крупнейшими публичными библиотеками являются:

- Российская государственная библиотека (РГБ) – 38 млн. единиц хранения;
- Российская национальная библиотека – 30 млн. единиц хранения;
- Всероссийская государственная библиотека иностранной литературы. Библиотечная сеть вузов включает 500 библиотек – 300 млн. единиц хранения.

Архивный фонд РФ находится в ведении Федеральной архивной службы (Росархив). Объем фондов – 460 млн. единиц хранения.

В единую систему государственной статистики входят Госкомстат России (федеральный орган исполнительной власти, который осуществляет руководство данной системой), его органы в областях и округах, подведомственные учреждения. Это приблизительно 2250 организаций.

Государственная система научно-технической информации (ГСНТИ) была создана в 1960-1970 гг. Государственным комитетом РФ по науке и технологиям, но в результате перестройки эта система, включавшая в себя около 200 специализированных организаций развалилась. В настоящее время в состав ГСНТИ входят следующие научно-технические организации и научно-технические библиотеки:

- Всероссийский научно-технический информационный центр Министерства науки и технологий (ВНТИЦ) - по ведущимся в стране и законченным открытым научно-исследовательским и опытно-конструкторским работам, защищенным диссертациям на соискание ученых степеней, алгоритмам и программам;
- Всероссийский научно-исследовательский институт межотраслевой информации (ВИМИ) – по научно-исследовательским, опытно-конструкторским работам и результатам научно-исследовательской работы оборонного комплекса;
- Российское объединение информационных ресурсов научно-технического развития (Росинформресурс) Министерства науки и

технологий РФ – по использованию результатов научно-технической деятельности предприятий и организаций, а также организации обмена этой информацией между регионами;

- Всероссийский институт научной и технической информации РАН и Министерства науки и технологий РФ (ВИНИТИ);
- Государственная публичная научно-техническая библиотека Министерства науки и технологий РФ, Библиотека РАН, Библиотека по естественным наукам РАН;
- Всероссийский научно-исследовательский институт информации и технико-экономических исследований агропромышленного комплекса Министерства сельского хозяйства и продовольствия РФ;
- Российский федеральный геологический фонд;
- Федеральный фонд государственных стандартов, общероссийских классификаторов технико-экономической информации – по нормативным документам по стандартизации, метрологии и сертификации;
- Российская книжная палата Государственного комитета РФ по печати – по опубликованным в РФ произведениям печати и государственной библиографии<sup>4</sup>
- Российский государственный архив научно-технической документации Федеральной архивной службы РФ;
- Научно-технический центр "Информрегистр" – по электронным изданиям;
- и др.

В современном мире все больше приобретает значение информация не в бумажном, а в электронном виде. В сети интернет появляется множество данных, в том числе и научных. Ресурсы, предоставляющие доступ к ним также разнообразны (см. таблицу )

Таблица 1.1. Научные сетевые ресурсы.

Название	Описание
<b>Научные поисковые системы</b>	
<i>Scirus</i>	Универсальная научная поисковая система. Осуществляет полнотекстовый поиск по статьям журналов большинства крупных иностранных издательств (порядка 17 млн. статей), статьям в крупных архивах статей и препринтов, научным ресурсам Internet (более 250 млн. проиндексированных страниц).
<i>Google Scholar</i>	Поисковая система по научной литературе. Включает статьи крупных научных издательств, архивы препринтов, публикации на сайтах университетов, научных обществ и других научных организаций. Ищет статьи в том числе и на русском языке. Что не маловажно, рассчитывает индекс цитирования публикаций и позволяет находить статьи, содержащие ссылки на те, что уже найдены.
<i>Science Research Portal</i>	Научная поисковая система, осуществляющая полнотекстовый поиск в журналах многих крупных научных издательств, таких как Elsevier, Highwire, IEEE, Nature, Taylor & Francis и др. Ищет статьи и документы в открытых научных базах данных: Directory of Open Access Journals, Library of Congress Online Catalog, Science.gov и Scientific News.
<i>Windows Live Academic</i>	Научная поисковая система от Microsoft. Предназначена для поиска научных статей как в открытых источниках, так и в архивах изданий с платным доступом.
<i>Infotrieve - artical finder</i>	Поиск статей в более чем 35000 журналах по физике, технике, медицине, юриспруденции и др. Возможен поиск только по какой-то определенной области науки.
<i>Medline</i>	Поиск по статьям медицинской тематики. Созданная национальной медицинской библиотекой США эта база данных включает статьи из более 3900 медицинских и биологических журналов, издающихся в 71 стране мира.
<i>High Wire Press</i>	Это большое хранилище научных журналов, предоставляющих бесплатный полнотекстовый доступ к своим статьям (968 журналов, 1.39 млн. статей). Данная поисковая система позволяет осуществлять полнотекстовый поиск в этих журналах + поиск в Medline.
<i>e-Print ArXive</i>	Лос-Аламосский архив электронных публикаций. Это коллекция копий статей по физике, математике, нелинейной динамике, computer science. Цель создания - свободный обмен научной информацией. Сознательные авторы размещают здесь свои статьи до опубликования, а иногда и вовсе без этого. Содержит поисковую систему по тематическим разделам.
<i>ResearchIndex</i>	Научная поисковая система, индексирующая статьи в PostScript и PDF формате с научных веб-сайтов. Многие статьи (по желанию авторов) доступны для бесплатного скачивания. Кроме полнотекстового поиска по статьям система также осуществляет поиск ссылок на данную публикацию или автора.
<i>Scientopica</i>	Научная поисковая система и каталог научных ресурсов.
<i>SciNet - Science search</i>	Рекламирует себя как первая из научных поисковых систем. Совмещена с каталогом научных ресурсов.
<i>Информационн-</i>	Объемная публичная российская информационная система

<i>поисковая система "Артефакт"</i>	коллективного пользования, содержащая политическую, коммерческую, юридическую, научную и адресно-справочную информацию. Разработку, поддержку и сопровождение системы осуществляет информационное агентство "Интегрум-Техно", одно из крупнейших российских агентств, интегрирующих информационные ресурсы на основе современных технологий. Обеспечивает эффективный доступ пользователей к информации в режиме on-line в сети Интернет.
<b>Специализированные поисковые системы</b>	
<i>Zentralblatt MATH database</i>	Поиск математических статей начиная с 1931 года. База данных создана Европейским математическим обществом и Гейдельбергской Академией наук. Содержит более 1.8 млн. документов.
<i>EULER search engine</i>	Поисковая система для математической литературы. Обещает стать крупнейшим поисковиком в этой области. Данные будут предоставлены упоминавшейся выше службой Zentralblatt MATH и рядом других математических библиотек и электронных архивов.
<i>MathSearch</i>	Поисковая система, специализирующаяся на сайтах и статьях по математике и статистике. Создана сиднейским университетом. База данных содержит порядка 300 тыс. документов.
<i>Directory of Mathematics Pre print and e-Print Servers</i>	Каталог ссылок на архивы препринтов и электронных публикаций по математике. Создатели ставят своей целью собрать данные обо всех подобных ресурсах по всему миру.
<i>The worldwide search engine of the chemical industry</i>	Поисковая система по химическим ресурсам.
<i>Неофициальный сервер геологического факультета МГУ</i>	Поисковая система по геологическим сайтам России. В настоящее время проиндексировано более 50 тыс. документов со 127 сайтов.
<i>CHAOS-Bibliography File Search</i>	Ищет статьи по нелинейной динамике.
<i>CiteSeer Publications ResearchIndex</i>	Научная поисковая система, индексирующая статьи в PostScript и PDF формате с научных веб-сайтов. Многие статьи (по желанию авторов) доступны для бесплатного скачивания. Кроме полнотекстового поиска по статьям система также осуществляет поиск ссылок на данную публикацию или автора.
<i>Scholar.ru</i>	Проект был создан для упрощения поиска документов научной тематики на русском языке, в первую очередь – выполненных в России. Основная цель проекта – сбор информации о свободно скачиваемых научных публикациях. Проект не рассчитан на хранение полных текстов статей в том или ином виде, вместо этого используется база ссылок на тексты документов с информацией о самих публикациях (аннотация, авторы и т. д.)
<b>Научные издательства</b>	
<i>Elsevier</i>	ScienceDirect издательства "Elsevier" ("Эльзевир") — крупнейший в мире электронный ресурс информации по науке, технологии и медицине. Разработан для удовлетворения информационных

	потребностей научных, образовательных, коммерческих и правительственных организаций. Часть журналов ScienceDirect находятся в свободном доступе.
<i>Springer</i>	Biological Cybernetics etc.
<i>Blackwell Publishers</i>	Полнотекстовые журналы издательства Blackwell Publishing охватывают почти все области естественных и общественных наук.
<i>Издательство института Иоффе</i>	Журналы: ЖТФ, Письма в ЖТФ, Физика твердого тела, Физика и техника полупроводников. Имеется поисковая система по заголовкам, авторам и аннотациям статей (без словоформ). Статьи доступны для бесплатного скачивания.
<i>World Scientific Publishing</i>	Журнал: International Journal of Bifurcation and Chaos in Applied Science and Engineering
<b>Каталоги и базы научной информации</b>	
<i>Scopus</i>	Крупнейшая в мире единая реферативная база данных, индексирует более 22 000 научно-технических и медицинских журналов примерно 4000 международных издательств. Более 300 журналов на русском языке. В свободном доступе представлен только каталог авторов.
<i>Союз образовательных сайтов</i>	Содержит информацию по естественным наукам
<i>Ingenta</i>	Библиографическая база данных универсального профиля, включающая описания статей из англоязычных журналов и сборников. Ее наполнение ведется с 1988 года. Объем базы составляет более 13 миллионов записей.
<i>Информационное агентство Lexis-Nexis</i>	Содержит несколько млрд. документов с глубоким архивом (до 30 лет по бизнес-информации и более 200 лет по юридическим англоязычным источникам). Его основы закладывались в шестидесятые годы прошлого столетия как собрание баз данных для юристов.
<i>Информационная система Factiva</i>	Это результат взаимодействия крупнейших мировых информационных агентств Reuters и Dow Jones. Два информационных гиганта объединили имеющиеся базы данных и создали колоссальную систему, позволяющую клиентам эффективно искать бизнес-информацию, отслеживать текущие новости о конкурентах и получать обзоры рынков из уникальной коллекции источников. Система предоставляет доступ к оперативным новостям от Dow Jones, Reuters, Associated Press, ИТАР-ТАСС, Прайм-ТАСС и др. региональных агентств. Информация, проходящая по теле- и радио программам (BBC, CNN, ABC, Fox, Эхо Москвы) и электронным СМИ также имеется в этом информационном ресурсе.
<i>Информационная корпорация Questel-Orbit</i>	Включает две подсистемы Questel (Франция) и Orbit (США). Подсистемы определенным образом специализированы как по содержанию баз данных, так и по поисковым возможностям и языкам общения. Корпорация существует более 20 лет и обеспечивает 35 тыс. потребителей по всему миру.
<i>Econlit</i>	База данных по экономике. Производитель - The American Economic Association. Реферативная база данных зарубежных изданий по экономике; реферирует более 400 только экономических журналов, а также труды конференций, книги,

	статьи в коллективных сборниках, диссертации, рецензии.
<i>INSPEC</i>	Ведущая англоязычная реферативная научно-техническая база данных.. Ресурс создается Лондонским Институтом инженеров по электротехнике (The Institution of Electrical Engineers, IEE) и содержит в настоящее время более 8 млн. записей: рефераты публикаций из более 3500 научных журналов по физике, электронике, информатике, компьютерным технологиям и техническим наукам, почти 2 тыс. материалов научных конференций.
<i>GeoRef</i>	База данных по минералогии и кристаллографии. Язык - английский. Производитель - The American Geological Institute. База данных содержит более 2.2 миллионов библиографических описаний из более чем 3000 журналов и других источников с 1933 г. И охватывает такие области геологии, как минералогия и кристаллография.
<i>Agricola</i>	База данных по сельскому хозяйству и смежным наукам. Язык - английский. Производитель - The U.S. Department of Agriculture's National Agricultural Library

### 1.6. Организация поиска информации

Глобальная сеть Internet содержит огромный объем информации, который стремительно увеличивается с каждым днем. В силу этого часто оказывается, что задача нахождения необходимой информации в этом информационном океане является чрезвычайно сложной и нужно уметь эффективно использовать различные поисковые системы (программы поиска информации в сети). Случайно найти нужный адрес в сети можно разными способами: просто путешествуя по узлам (surfing on the web), узнать у знакомых, увидеть в рекламе и т.д.

Целенаправленный поиск явно или неявно требует формулировки цели поиска, понимания того, что является объектом поиска, обоснованного выбора средства поиска и эффективной методики.

#### *Цель поиска*

Цель определяет характеристики объектов поиска, объем и сроки выполнения работы, перечень средств поиска и способы их применения. Например, при подготовке к экзамену требуется методическая литература, учебные курсы, конспекты лекций, для реферата – аналитические обзоры, для доклада – графические материалы, для презентации – файлы мультимедиа, для научного исследования – программное обеспечение, и т.д.

### ***Объект поиска***

В качестве объекта поиска может рассматриваться любая информация, если имеется возможность представления ее в Internet. Это могут быть телефоны и адреса, информация о товарах и услугах, радио и теле трансляции и многое другое. Наиболее распространенными объектами поиска является:

1. Адрес информационного ресурса.
2. Web-страница и включенные в нее элементы: текст, мультимедиа данные, гиперссылки, программы (апплеты) и т.д.
3. Программы в том числе: демонстрационные и тестовые программы, средства улучшения (upgrade), обновления (update) и исправления ошибок (patch) в программах;
4. Сообщения в телеконференциях;
5. Информация из интерактивных баз данных, справочников, каталогов.

### ***Средствами поиска являются:***

- поисковые системы;
- средства локального поиска;
- утилиты автономного поиска.

#### **1.6.1. Поисковые каталоги.**

Каталоги ресурсов – глобальные, локальные, специализированные – представляют собой размещаемые в Сети базы данных с адресами ресурсов. Эти базы данных могут иметь разный объем накопленной информации. Обычно они имеют иерархическую структуру.

Каждая запись в списке категорий – это гиперссылка. Щелчок по ней открывает следующую страницу поискового каталога, на котором выбранная тема представлена более подробно. Продолжая погружаться в тему, можно прийти до списка конкретных Web-страниц и выбрать тот ресурс, который наиболее подходит для решения Вашей задачи. Также в поисковом каталоге можно использовать кнопку Поиск для уточнения поиска нужных страниц.

Поисковые каталоги создаются в основном вручную высококвалифицированными редакторами, которые просматривают пространство WWW, отбирают то, что по их мнению представляет общественный интерес и заносят адреса в каталог.

Главным достоинством тематических каталогов является большая ценность получаемой пользователем информации, что обеспечивается присутствием "человеческого фактора" в процессе анализа и сортировки новых Web-страниц. С другой стороны, тематические каталоги имеют существенный недостаток, связанный опять же с человеческим фактором, ибо из-за ограниченных возможностей человека их базы данных охватывают лишь небольшую часть всего информационного Web-пространства (менее 1 %). Таким образом, несмотря на всю полезность тематических каталогов, использование лишь поисковых систем этого вида часто оказывается явно недостаточным.

Примерами тематических каталогов являются:

- Yahoo ([www.yahoo.com](http://www.yahoo.com)),
- Magellan ([www.mckinley.com](http://www.mckinley.com)),
- List.Ru ([www.list.ru](http://www.list.ru)),
- Созвездие Интернет ([www.stars.ru](http://www.stars.ru)),
- Russia on the Net ([www.ru](http://www.ru)),
- @Rus ([www.atrus.ru](http://www.atrus.ru)),
- Weblist ([www.weblist.ru](http://www.weblist.ru)).

### **1.6.2. Поисковые машины или поисковые указатели.**

Поисковые машины— это сервер с огромной базой данных URL-адресов, который автоматически, в круглосуточном режиме обращается к страницам WWW по всем этим адресам, изучает содержимое этих страниц, формирует и прописывает ключевые слова со страниц в свою базу данных (индексирует страницы). Более того, этот сервер обращается по всем

встречаемым на страницах ссылкам, поэтому при подобной работе поисковая машина в конечном результате теоретически может обойти все сайты Интернет. Для сбора сведений о ресурсах поисковая машина использует специальные программы, которых называют *червяками*, *пауками*, *спайдерами*, *краулерами* и т.п. В каждой найденной странице анализируется заголовок, тема, ключевые слова, текст и состав Web-страницы.

В отличие от тематических каталогов, автоматические индексы охватывают до 25 % общего Web-пространства.

Поисковый сервис выдает список URL адресов, которые указывают на документы, соответствующие запросу пользователя. Общение с поисковым сервисом осуществляется с помощью Web интерфейса.

Поисковый сервис делится на 3 части:

- 1) *Web интерфейс.*
- 2) *Поисковый робот.*
- 3) *СУБД.* В поисковой базе СУБД хранятся ссылки (URL адреса, хранящиеся в Internet). Помимо списка URL адресов в поисковой СУБД также хранится ключевое содержимое соответствующих документов.

***Поисковый робот*** – это программа, осуществляющая автоматическое сканирование Web ресурсов (индексирование) на предмет появления новых, модификаций существующих и удаление старых Web ресурсов.

В результате сканирования обновляется поисковая база.

Существует 2 способа работы поискового робота:

- 1) Сканирование роботом Web ресурсов; заполнение базы данных.
- 2) Обращение пользователя к поисковому сервису через Web интерфейс (используется специальная форма для указания запроса).

Эффективность поисковой системы сводится к объему поисковых баз, а также определяется эффективностью алгоритма ранжирования документов, наличием языка запросов.

**Язык запросов** – это набор команд, позволяющий пользователю уточнить поисковый запрос и получить более точные результаты на свой запрос.

Автоматический индекс имеет отдельную поисковую систему для обеспечения интерфейса с пользователем. Эта система может, просматривая базу данных, по заданному набору ключевых слов находить и выдавать на экран пользовательского компьютера адреса и краткую информацию обо всех Web-страницах, которые содержат данный набор ключевых слов. Таким образом, автоматический индекс состоит из трех частей: программы-робота, собираемой этим роботом базы данных и интерфейса для поиска в этой базе данных. Именно с последней составляющей и работает пользователь. В силу такой организации, автоматический индекс не делает какой-либо классификации или оценивания информации.

Главной задачей любой поисковой машина является поиск информации, соответствующей информационным потребностям пользователя. Очень важно в результате проведенного поиска ничего не потерять, то есть найти все документы, относящиеся к запросу, и не найти ничего лишнего. Поэтому вводится качественная характеристика процедуры поиска - релевантность.

**Релевантность** - это соответствие результатов поиска сформулированному запросу.

Вебмастера желают повысить рейтинг своих страниц и это понятно: ведь на любой запрос к поисковой машине могут быть выданы сотни и тысячи отвечающих ему ссылок на документы. В большинстве случаев только 10 первых ссылок обладают достаточной релевантностью к запросу. Естественно, хочется, чтобы документ оказался в первой десятке, поскольку большинство пользователей редко просматривает следующие за первой десяткой ссылки. Иными словами, если ссылка на документ будет одиннадцатой, то это также плохо, как если бы ее не было вовсе.

Показателями поиска информации являются полнота выдачи (потери информации) и точность выдачи (информационный шум). Введем следующие обозначения:

а – множество релевантных и выданных системой документов;

б – множество нерелевантных, но выданных системой документов;

в – множество релевантных, но не выданных системой документов.

$$\text{Полнота выдачи (ПВ)} = \frac{a}{a+v} * 100\% .$$

$$\text{Точность выдачи (ТВ)} = \frac{a}{a+b} * 100\% .$$

$$\text{Потери информации (ПИ)} = \frac{v}{a+v} * 100\% .$$

$$\text{Информационный шум (ИШ)} = \frac{b}{a+b} * 100\% .$$

Российские поисковые машины:

- Рамблер ([www.rambler.ru](http://www.rambler.ru)),
- Апорт2000 ([www.aport.ru](http://www.aport.ru)),
- Яндекс ([www.yandex.ru](http://www.yandex.ru)),
- Lupa.ru ([www.lupa.ru](http://www.lupa.ru)).

Некоторые зарубежные поисковые машины:

- Alta Vista,
- Excite Search,
- HotBot,
- InfoSeek,
- Lycos,
- Fast Search,
- Google,
- Northern Light.

### **Особенности поисковых машин:**

Каждая поисковая машина обладает рядом особенностей. Эти особенности следует учитывать при изготовлении своих страниц.

*Тип поисковой машины.* "Полнотекстовые" поисковые машины индексируют каждое слово на веб-странице, исключая лишь некоторые стоп-слова. "Абстрактные" поисковые машины создают некий экстракт каждой страницы. Для вебмастеров полнотекстовые машины полезней, поскольку любое слово, встречающееся на веб-странице, подвергается анализу при определении его релевантности к запросам пользователей. Однако для абстрактных поисковых машин может случиться, что страницы проиндексированы лучше, чем для полнотекстовых. Это может исходить от алгоритма экстрагирования, например по частоте употребления в странице одних и тех же слов.

*Размер.* Размер поисковой машины определяется количеством проиндексированных страниц. Приведенные в таблице значения не слишком точны, но могут прояснить некоторые моменты. Например, в поисковой машине с большим размером могут быть проиндексированы почти все ваши страницы, при среднем объеме ваш сервер может быть частично проиндексирован, а при малом объеме ваши страницы могут вообще не попасть в каталоги поисковой машины.

*Период обновления.* Поскольку Веб изменяется непрерывно, поисковые машины индексируют все без учета даты. Однако в каждый момент времени ссылки, выдаваемые в ответ на запросы пользователей, могут быть однодневной давности, а могут быть и месячной давности, а то и больше. Вот некоторые причины, по которым это происходит:

- некоторые поисковые машины сразу индексируют страницу по запросу пользователя, а затем продолжают индексировать еще не проиндексированные страницы
- другие чаще могут "ползать" по наиболее популярным страницам сети, чем по другим.

*Дата индексирования документа.* Некоторые поисковые машины показывают дату, когда был проиндексирован тот или иной документ. Это

помогает пользователю понять, какой "свежести" ссылку выдает поисковая система. Другие оставляют пользователей только догадываться об этом.

*Указанные (submitted) страницы.* В идеале поисковые машины должны найти любые страницы любого сервера в результате прохода по ссылкам. Реальная картина выглядит по-другому. Страницы серверов гораздо раньше появляются в индексах поисковых систем, если их прямо указать (Add URL).

*Не указанные (non-submitted) страницы.* Если хотя бы одна страница сервера указана, то поисковые машины обязательно найдут следующие страницы по ссылкам из указанной. Однако на это требуется больше времени. Некоторые машины сразу индексируют весь сервер, но большинство все-таки, записав указанную страницу в индекс, оставляют индексирование сервера на будущее.

*Глубина индексирования.* Этот параметр относится только к не указанным страницам. Он показывает сколько страниц после указанной будет индексировать поисковая система. Большинство крупных машин не имеют ограничений по глубине индексирования. На практике же это не совсем так. Вот несколько причин, по которым могут быть проиндексированы не все страницы:

- не слишком аккуратное использование фреймовых структур (без дублирования ссылок в управляющем (frameset) файле)
- использование imager без дублирования их обычными ссылками

*Поддержка фреймов.* Если поисковый робот не умеет работать с фреймовыми структурами, то многие структуры с фреймами будут упущены при индексировании.

*Защищенные паролями директории и сервера.* Некоторые поисковые машины могут индексировать такие сервера, если им указать Username и Password. Зачем это нужно? Чтобы пользователи видели, что есть на Вашем сервере. Это позволяет как минимум узнать, что такая информация есть, и, быть может, они тогда подпишутся на Вашу информацию.

*Частота появления ссылок.* Основные поисковые машины могут определить популярность документа по тому, как часто на него ссылаются из других мест Сети. Некоторые машины на основании таких данных "делают вывод" стоит или не стоит тратить время на индексирование такого документа.

*"Способность к обучению".* Если сервер обновляется часто, то поисковая машина чаще будет его реиндексировать, если редко - реже.

*Контроль индексации.* Показывает, какими средствами можно управлять той или иной поисковой машиной. Все крупные поисковые машины руководствуются предписаниями файла robots.txt. Некоторые также поддерживают контроль с помощью META-тегов из самих индексируемых документов.

*Перенаправление (redirect).* Некоторые сайты перенаправляют посетителей с одного сервера на другой, и этот параметр показывает какой URL будет связан с вашими документами. Это важно, поскольку, если поисковая машина не обрабатывает перенаправление, то могут возникнуть проблемы с несуществующими файлами.

*Стоп-слова.* Некоторые поисковые машины не включают определенные слова в свои индексы или могут не включать эти слова в запросы пользователей. Такими словами обычно считаются предлоги или просто очень часто используемые слова. А не включают их ради экономии места на носителях. Например, Altavista игнорирует слово web и для запросов типа web developer будут выданы ссылки только по второму слову. Существуют способы избежать подобного.

### **Влияние на алгоритм определения релевантности**

Поисковые машины обязательно используют расположение и частоту повторения ключевых слов в документе. Однако, дополнительные механизмы увеличения степени релевантности для каждой машины различны. Этот параметр показывает, какие именно механизмы существуют для той или иной машины.

*Spam-штрафы.* Все крупные поисковые системы "не любят", когда какой-либо сайт пытается повысить свой рейтинг путем, например, многократного указания себя через Add URL или многократного упоминания одного и того же ключевого слова и т. д. В большинстве случаев подобные действия (spamming, stacking) караются, и рейтинг сайта наоборот падает.

*Поддержка META-тегов.* По идее, все поисковые машины должны учитывать метаданные при индексации страниц, однако на практике не все это делают.

*Title.* Этот параметр показывает как поисковые машины генерируют заголовки ссылок для пользователя в ответ на его запрос.

*Description.* Этот параметр показывает как поисковые машины генерируют описания ссылок для пользователя в ответ на его запрос.

*Проверка статуса URL.* Очень полезная для вебмастера черта поисковой машины - можно ли проверить насколько глубоко проиндексирован его сервер и есть ли он вообще в индексе поисковой машины.

*Удаление старых данных.* Параметр, определяющий действия вебмастера при закрытии сервера или перемещении его на другой адрес. Возможны два действия: просто удалить старое содержание и переписать файл robots.txt.

- удаление содержимого: когда поисковая машина попытается реиндексировать документы и не найдет их, старые ссылки в индексе будут удалены. В этом случае все зависит от периода обновления данных для поисковой машины.
- robots.txt: когда поисковая машина запросит этот файл и "увидит", что сервер весь закрыт от индексации, то все ссылки на файлы этого сервера будут удалены из индекса.

*Имя поискового робота.* В этом пункте указаны имена роботов, которыми они отвечают на HTTP-запрос. Полезно для написания robots.txt.

### **1.6.3. Классификационно-рейтинговые системы.**

Классификатор похож на каталог, но в отличие от каталога перед ним не ставится задача собрать как можно больше информации о ресурсах сети. По каждой из категорий, входящих в классификатор, представляются лучшие сайты, а дальше работает счетчик. Выбрав интересующую тему, пользователь получает список Web-узлов, посвященных данной теме, и рейтинг их популярности, которая измеряется в количестве посещений за последние сутки. **Рейтинг** – это сортировка ссылок в порядке их посещаемости. Есть некоторые Web-узлы, специализирующиеся на ведении рейтингов: «Рамблер Top 100» ([www.rambler.ru](http://www.rambler.ru)) -самый крупный классификатор в России, <http://dir.spylog.ru>

#### **1.6.4. Метапоисковые системы**

Метапоисковые системы обеспечивают для каждого запроса одновременный поиск с помощью нескольких поисковых серверов. Такие системы позволяют задавать только простые запросы на поиск. Это сокращает время, но получаемые результаты, как правило, хуже, чем при независимом поиске на каждом поисковом сервере с использованием расширенных возможностей.

Базы данных разных поисковых систем в значительной мере не пересекаются. Поэтому для поиска достаточно редкой информации целесообразно обращаться не к одной, а к нескольким ПС. Однако правила оформления запросов для разных ПС, вообще говоря, отличаются друг от друга. Для того, чтобы не обращаться поочередно к разным поисковым системам и не думать о специфических правилах оформления запроса для каждой из них, были созданы так называемые метапоисковые системы.

Приняв заказ клиента, заданный с помощью ключевых слов в соответствии со своими собственными правилами его оформления, метапоисковая система сама пропишет его в бланках разных поисковых систем, разошлет эти бланки и будет ждать ответа. Когда все поисковые системы пришлют результаты поиска, метапоисковая программа сведет их в один документ и отправит пользователю.

Типичный пример метапоисковой системы **Metabot.ru** ([www.metabot.ru](http://www.metabot.ru)). С ее помощью можно одновременно искать в нескольких наиболее распространенных русскоязычных поисковых системах и каталогах, а также в наиболее известных англоязычных системах и в FTP-архивах.

К таким метапоисковым системам относится **MetaCrawler**, который рассылает запрос на 9 различных поисковых систем (в их число входит: Yahoo, Alta Vista, Lycos, Excite и т. д.). На случай разной интерпретации одних и тех же по смыслу опций в разных поисковых системах MetaCrawler предусматривает даже возможность проверки результатов поиска: прежде чем дать ссылку пользователю, он самостоятельно посмотрит на документ и проверит, соответствует ли он условиям запроса — так как их понимает MetaCrawler. Разумеется, этот режим проверки сильно задерживает получение результатов, но зато позволяет защититься как от неработоспособных ссылок, так и от бессмысленных результатов. Заметим, что на бланке запроса MetaCrawler можно задать время ожидания: в список будут включены только те результаты, которые успеют прийти с различных поисковых систем к этому моменту.

Наиболее удобные метапоисковые системы это:

- [www.360.ru](http://www.360.ru)
- Accufind ([www.accufind.com](http://www.accufind.com));
- Metafind ([www.metafind.com](http://www.metafind.com));
- Metasearch ([www.metasearch.com](http://www.metasearch.com)).

### 1.6.5. Автономные утилиты

Утилиты автономного поиска устанавливаются на компьютере пользователя. Они обеспечивают накопление поисковых запросов, выполняют метапоиск, отслеживают изменения заданных Web-страниц. К подобным программам можно отнести **WebCompass** ([www.quarterdeck.com](http://www.quarterdeck.com)) и **Copernic** ([www.copernic.com](http://www.copernic.com)). Однако применение подобных систем

крайне загружает каналы связи и делает их работу малоэффективной, поэтому при подключении компьютера к Интернету через телефонные линии использовать такие программы не рекомендуется.

Полезными при поиске могут оказаться, так называемые автономные браузеры (off-line browsers), обеспечивающие загрузку заданных Web-узлов без участия пользователя. В таких программах можно задавать “глубину” поиска ссылок внутри узла, тип и предельный размер копируемых файлов, расписание загрузки. Наиболее популярны **WebWhacker** ([www.ftg.com](http://www.ftg.com)) и **Teleport Pro** ([www.tenmax.com](http://www.tenmax.com)).

Ускорить ручной поиск можно с помощью средств анализа структуры Web-узла. Они изображают в удобной форме навигационную карту узла, на которой показаны элементы Web-страниц с аннотациями и их связи. Для этой цели можно применять **WebTurbo** ([www.webturbo.com](http://www.webturbo.com)) или **PersonalCrawler** ([www.vci.co.il](http://www.vci.co.il)).

#### **1.6.6. Приемы поиска информации.**

Несмотря на то, что интерфейс всех поисковых систем различен, все они имеют два общих элемента – это строка ввода запроса пользователя, и кнопка Поиск. Если на основной странице поисковой системы есть тематический каталог, то нужно определиться с тематикой поиска, это значительно ограничит пространство поиска и ускорит работу поисковой системы.

Поиск по одному ключевому слову прост, но приводит к формированию огромного списка страниц, полезную информацию здесь найти трудно.

Гораздо эффективнее поиск по нескольким словам, но здесь важную роль играет связь между словами, которая помогает системе обрабатывать эту группу слов. Для эффективного поиска по нескольким ключевым словам нам нужны специальные команды, которые позволяют связать отдельные слова между собой. Эти команды в поисковых системах образуют язык запросов. В нем всего несколько команд, и освоить его по инструкции, имеющейся на сервере поисковой системы довольно просто. Каждая поисковая машина использует свой язык запросов, поэтому при

использовании разных поисковых систем надо знать особенности каждой. Но есть общий принцип, согласно которому все команды можно разделить на три группы:

- команды простого поиска,
- команды расширенного поиска,
- команды специального поиска.

Рассмотрим приемы формирования запросов на примере поисковой машины Апорт2000 ([www.aport.ru](http://www.aport.ru)).

### 1.6.7. Средства простого поиска.

1. Поиск группы слов. Большинство российских поисковых систем, как и поисковая машина Апорт, воспринимает группу слов так, как будто между ними стоит союз И. Например, *лица в розыске*.
2. Поиск словоформ. В большинстве случаев Апорт позволяет находить разные словоформы, например, запрос *допрос несовершеннолетнего* можно было сформулировать *несовершеннолетний допрос*.
3. Поиск точных форм. Иногда нужно от словоформ отказаться. Допустим нас интересует слово *престол*, а не *престольный* и не *первопрестольный*. В этом случае Апорт использует восклицательный знак, например: *!престол*.
4. Роль заглавных букв. Для большинства поисковых систем заглавные буквы воспринимаются буквально. Так, например поиск по словам укажет только на те страницы, где встречаются слова *Красная Шапочка*. А поиск по запросу *красная шапочка* укажет на страницы, где встречаются слова *Красная Шапочка*, и *Красная шапочка*, и *красная Шапочка*, и *красная шапочка*.
5. Значение подстановочных символов. Апорт позволяет использовать символ \* вместо любого количества символов до конца слова. Например, если нужны документы, в которые входят слова *Татарская республика*, то можно использовать только: *республика Татарс\**.

6. Учет зарегистрированных слов. Зарегистрированными словами (стоп-словами) считаются те, которые не учитываются при поиске. Обычно к ним относятся все слова, которые меньше 4 символов – предлоги, союзы, местоимения, артикли и т.д. Так, например в запросе *все о юристах* будет учитываться только слово *юристах* и его словоформы. В специализированных поисковых системах к зарегистрированным могут относиться слова, характерные для этих систем. Например, если поисковая система ориентирована на поиск книг, то вероятнее всего слово *книга* в ней при обработке запроса не будет учитываться.
7. Средства контекстного поиска. Взятая в кавычки фраза ищется буквально, то есть так как она написана, без словоформ. Это очень полезный вид поиска.

#### **1.6.8. Средства расширенного поиска.**

В расширенном поиске кроме ключевых слов, можно использовать простейшие логические операторы и логические скобки.

1. Оператор И (AND,&,+). С помощью этого оператора объединяются 2 или более слов. В поисковой системе Апорт можно этим оператором не пользоваться, т.к. в ней списки слов и так воспринимаются, как будто между ними стоит оператор И. Но так бывает не всегда (например Alta Vista по умолчанию считает, что ключевые слова связаны соотношением ИЛИ).
2. Оператор ИЛИ (OR,|). Этот оператор обеспечивает поиск по любому из слов группы, например *образование ИЛИ обучение*.
3. Логические скобки. Скобки необходимы, когда важен порядок действия логических операторов, например, *Ломоносов ИЛИ(Михаил И Васильевич)*.
4. Оператор НЕТ (-,NOT). Этот оператор употребляется, когда из результатов поиска нужно исключить какое-либо ключевое слово. Например, по запросу *допрос NOT несовершеннолетний*, будут

возвращены документы, где есть слово *допрос*, но нет слова *несовершеннолетний*.

5. Поиск с указанием расстояния. Этот вид поиска позволяет указать, на каком расстоянии друг от друга могут располагаться слова в документе. Например, *сл5(приостановление следствия)*. В данном случае оба слова должны принадлежать одной группе, длиной не более 5 слов.

В таблице приведены логические операторы для некоторых поисковых машин.

Таблица 1.2. Поисковые операторы.

Оператор	Действие	Rambler	Yandex	Aport
<b>И</b>	Ищет страницы со <i>Словом1</i> <b>и</b> <i>Словом2</i>	<i>Слово1 AND Слово2</i> <i>Слово1 &amp; Слово2</i> <i>Слово1 Слово2</i> (пробел)	<i>Слово1 AND Слово2;</i> <i>Слово1 &amp; Слово2;</i> <i>Слово1 Слово2</i> (пробел); <i>+Слово1 + Слово2</i>	<i>Слово1 AND Слово2;</i> <i>Слово1 &amp; Слово2;</i> <i>Слово1 Слово2</i> (пробел); <i>+Слово1 + Слово2;</i> <i>Слово1 И Слово2</i>
<b>ИЛИ</b>	Ищет страницы со <i>Словом1</i> <b>или</b> <i>Словом2</i>	<i>Слово1 OR Слово2</i> <i>Слово1   Слово2</i>	<i>Слово1 OR Слово2</i> <i>Слово1   Слово2</i>	<i>Слово1 OR Слово2</i> <i>Слово1   Слово2</i> <i>Слово1 ИЛИ Слово2</i>
<b>НЕТ</b>	Ищет страницы, в которых <b>нет</b> <i>Слова</i>	<i>NOT Слово;</i> <i>! Слово</i>	<i>NOT Слово;</i> <i>-Слово;</i> <i>~Слово</i>	<i>NOT Слово;</i> <i>-Слово</i> <b>НЕТ</b> <i>Слово</i>

Поисковые системы могут иметь страницу **Расширенный поиск**, где, задавая параметры поиска, можно построить запрос, не используя напрямую язык запросов. Поэтому перед широким применением любой поисковой системы полезно внимательно ознакомиться с принятыми в ней правилами поиска.

1. Поиск с указанием даты. Оператор *data* позволяет выполнять поиск документов по датам, например *data=01/02/2000 образование*. Поэтому запросу будут выданы документы, содержащие слово *образование* и датированные 1 февраля 2000 г.

2. Поиск по ссылкам на определенный адрес URL. Оператор *url* осуществляет поиск по ссылкам на названный ресурс. Например, по запросу *url=www.ulsu.ru* ,будут выданы все документы проиндексированные на сервере [www.ulsu.ru](http://www.ulsu.ru).
3. Поиск по заголовкам. Команда *title* позволяет находить страницы по их заголовкам. Например, по запросу *title=(Нургалиев Р.Г.)*будут найдены документы, в заголовках которых присутствует фамилия Министра МВД.

### 1.6.9. Службы поиска людей.

Нет единого списка или каталога адресов электронной почты, так же как нет единого печатного телефонного справочника для всего мира. Существует несколько коммерческих и некоммерческих справочных служб, но большинство их не включают в себя какой-либо отдельный регион или дисциплину. Они составляются различными методами. Они могут быть собраны компьютерными программами из сообщений групп новостей Internet или запущены отдельными лицами, которые необязательно являются владельцами адресов. Эти каталоги часто называются «белыми страницами» и включают в себя каталоги адресов электронной почты и почтовых адресов, а также телефонные номера.

Некоторые зарубежные поисковые системы, позволяющие найти людей по фамилии или e-mail:

- [www.bigfoot.com](http://www.bigfoot.com) – осуществляет E-mail Search; Find Friends; White Pages; Find People и т.д.;
- [www.addresses.com](http://www.addresses.com) – имеется каталог E-mail;
- [www.yahoo.com/search/people](http://www.yahoo.com/search/people);

WhoWhere [www.whowhere.com](http://www.whowhere.com) – каталог, производящий поиск даже по схожести звучания или написания фамилий (например, “Kirsanov”, “Kirsanoff” и т. д.).

Некоторые российские поисковые службы:

- <http://whitepages.rin.ru> – Российская Информационная Сеть. При выборе нужного человека, предлагается ему послать письмо со своими личными данными и вопросами к нему. Есть регистрация.
- [www.mates.ru](http://www.mates.ru) – Каталог выпускников школ и вузов разных лет. Есть рубрики: поиск людей, найди меня, одноклассники, однокурсники, регистрация и т.д. О людях содержится много данных, но конкретный e-mail не выдается, предлагается связаться с человеком, и если он пожелает, то ответит Вам.
- <http://poisk.goon.ru> – Поиск людей – одноклассников, родственников и т.д.
- [www.sherlok.ru](http://www.sherlok.ru) – платный поиск людей по фамилиям, телефонам и др. данным.
- [www.centrhoisk.narod.ru](http://www.centrhoisk.narod.ru) - платный поиск людей.

Один из самых надежных способов нахождения информации о персональных контактах, если Вы знаете организацию, к которой относится человек, - это обратиться к «начальной странице» организации (часто есть ссылка на внутренний каталог).

Существуют также обширные каталоги электронных адресов различных людей, которые пользуются электронной почтой (e-mail). Лидерство здесь, по-видимому, принадлежит каталогу Four11 по адресу <http://www.four11.com>.

Совершенно уникальным в настоящее время является всеамериканский адресный справочник по адресу <http://www.databaseamerica.com> , который выдает координаты любого из 90 млн. жителей США и любой из 10 млн. американских компаний.

## 2. Методы научного исследования в информационной сфере

### 2.1. Научные методы познания в исследованиях

В методологии научных исследований выделяют два уровня познания:

- **эмпирический** – наблюдение и эксперимент, а также группировка, классификация и описание результатов эксперимента, наблюдений;
- **теоретический** – построение и развитие научных гипотез, теорий, формулировка законов и выделение из них логических следствий, сопоставление различных гипотез и теорий.

**Метод** (греч.: способ исследования, обучения):

1. Система правил и приёмов подхода к изучению явлений и закономерностей природы, общества, мышления.
2. Способ достижения определённых результатов в познании и практике.
3. Приём теоретического исследования или практического осуществления чего-нибудь, исходящий из знания закономерностей развития объективной действительности.

Наука и ее методы подразумевают, что:

- 1) наблюдение может привести к открытию;
- 2) Наука имеет своей целью обнаружить какой-нибудь регулярно повторяющийся факт, закон или схему;
- 3) Наука имеет больше шансов привести к заслуживающим внимания результатам, если оно напрашивается какой-нибудь удачной мыслью или идеей;
- 4) наука может служить трамплином для обобщений и предложений, но оно не является доказательством;
- 5) проверяйте ваши предположения на частных его случаях и тех фактах, которые из него следуют;

- б) каждый подтвердившийся частный случай или оказавшееся справедливым следствие подкрепляют ваше предположение;
- 7) проводите тщательное различие между намёком на доказательство и самим доказательством, между предположением и фактом;
- 8) не пренебрегайте аналогиями – они могут привести к обнаружению новых фактов.

Исходя из методологии диалектического материализма различают следующие методы научного познания: общенаучные и конкретно-научные (частные).

Методы, используемые в информатике:

1. *Общенаучные:*

- анализ, синтез;
- аналогия, сравнение;
- индукция, дедукция, традукция; наблюдение, экспериментирование; формализация, абстрагирование;
- аксиоматические, гипотетические, эвристические, аналитические;

2. *Специфичные:*

- алгоритмизация, моделирование, программирование (планирование); системные, имитационные;
- поиска, анализа, выбора вариантов, решений;
- обнаружения и исправления ошибок;
- распознавания образов;
- постановки и решения задач;
- естественнонаучные, эмпирические, аксиоматические (математические); инженерные, изобретательские и т.п..

**Анализ** – это метод исследования, который включает в себя изучение предмета путем мысленного или практического расчленения его на составные элементы (части объекта, его признаки, свойства, отношения, характеристики, параметры и т.д.). Каждая из выделенных частей

анализируется отдельно в пределах единого целого. Например, анализ производительности труда рабочих производится по каждому цеху и по предприятию в целом.

**Синтез** – метод изучения объекта в его целостности, в единстве и взаимной связи его частей.

В процессе научных исследований синтез связан с анализом, поскольку он позволяет соединить части предмета, расчлененного в процессе анализа, установить их связь и познать предмет как единое целое (например производительность труда по производственному объединению в целом).

**Индукция** – метод исследования, при котором общий вывод о признаках множества элементов делается на основе изучения этих признаков у части элементов этого множества.

Так, например, изучаются факторы, отрицательно влияющие на производительность труда, по каждому отдельному предприятию, а затем данные обобщаются в целом по производственному объединению, в состав которого входят все эти предприятия как производственные единицы.

**Дедукция** – метод логического умозаключения от общего к частному, когда сначала исследуется состояние объекта в целом, а затем его отдельных элементов.

Применительно к предыдущему примеру сначала анализируется производительность труда в целом по объединению и далее по его производственным единицам.

**Аналогия** – метод научного умозаключения, посредством которого достигается познание одних предметов и явлений на основании их сходства с другими. Он основывается на сходстве некоторых сторон различных предметов и явлений, например, производительность труда в объединении может исследоваться не по каждому предприятию, а лишь по выбранным в качестве аналога, выпускающим однородную с другими предприятиями товарную продукцию и имеющим одинаковые условия для производственной деятельности.

При использовании этого метода полученные результаты распространяются на все аналогичные предприятия. Затраты на такой метод конечно меньше, а вот достоверность полученных выводов оказывается несколько ниже.

**Сравнение** – метод научного изучения, посредством которого устанавливаются сходство и различие предметов и явлений действительности.

**Измерение** – метод научного исследования процесса определения численного значения некоторой величины посредством определенной заранее единицы измерения.

**Исторический подход** – метод научного познания, в процессе которого происходит воспроизведение истории изучаемого объекта, явления во всей ее многогранности с учетом всех случайностей.

**Логический подход** – метод научного умозаключения, посредством которого достигается воспроизведение в мышлении сложного динамического явления в форме исторической теории с отвлечением от случайностей и отдельных несущественных фактов.

**Моделирование** – метод научного познания, основанный на замене изучаемого предмета, явления на его аналог (модель), содержащий существенные черты характеристики оригинала. В экономических исследованиях широко применяется экономико-математическое моделирование, когда модель и ее оригинал описываются тождественными уравнениями и исследуются с помощью ЭВМ (например транспортные маршруты при автомобильных перевозках грузов).

**Абстрагирование** – (от лат. – отвлекать) – метод отвлечения, позволяющий переходить от конкретных предметов к общим понятиям и законам развития.

Он применяется в экономических исследованиях для перспективного планирования, когда на основании изучения работы предприятий за

прошедший период времени прогнозируется развитие отрасли или региона на будущий период.

**Конкретизация** – метод исследования предметов во всей их разносторонности, в качественном многообразии реального существования во времени и пространстве в отличие от абстрактного, отвлеченного изучения предметов. При этом исследуется состояние предметов в связи с определенными условиями их существования и исторического развития.

Так, например, перспективы развития отрасли определяются на основании конкретных расчетов эффективности применения новой техники и технологии, сбалансированности трудовых и материальных ресурсов и др.

**Системный анализ** – изучение объекта исследования как совокупности элементов, образующих систему. В научных исследованиях он предусматривает оценку поведения объекта как системы со всеми факторами, влияющими на его функционирование.

Этот метод широко применяется в экономических исследованиях при комплексном изучении деятельности производственных объединений и отрасли в целом, определении пропорций развития народного хозяйства и т.п.

Единой методики системного анализа в научных исследованиях, к сожалению, пока не имеется. В практике исследований он применяется путем использования следующих методик:

- **процедур теории исследования операций**, позволяющих дать количественную оценку объектам исследования;
- **анализа систем для исследования объектов** в условиях неопределенности;
- **системотехники**, включающей проектирование и синтез сложных систем в процессе исследования их функционирования (проектирование и оценка экономической эффективности АСУ, технологических процессов и др.).

**Комплексный анализ** – метод всестороннего изучения объекта, явления в тесном взаимодействии с представителями самых разных наук и научных направлений.

**Функционально-стоимостный анализ (ФСА)** – метод исследования объекта (явления, изделия, процесса, структуры) по его функции и стоимости, применяемый при изучении эффективности использования материальных и трудовых ресурсов. Важнейшими его принципами являются следующие:

- функциональный подход при исследовании функций объекта и его элементов с целью наиболее полного удовлетворения заданных требований в выборе рациональных путей их реализации;

- народнохозяйственный подход к оценке потребительских свойств и затрат на разработку, производство и использование объекта;

- соответствие полезности функций затратам на их осуществление;

- коллективное творчество, использующее методы поиска и формирования технических решений, качественной и количественной оценок вариантов решений.

Целевой функцией ФСА является достижение оптимального соотношения между потребительской стоимостью объекта и совокупными затратами на его разработку, снижение себестоимости выпускаемой товарной продукции и повышение ее качества, роста производительности труда.

**Формализация** – метод исследования объектов путем представления их элементов в виде специальной символики, например, представление себестоимости продукции специальной формулой (математической зависимостью), в которой при помощи символов изображены статьи затрат.

**Гипотетический метод** (от греч. – основанный на гипотезе) – основан на научном предположении, выдвигаемом для объяснения какого-либо явления и требующем проверки на опыте и теоретического обоснования, чтобы стать достоверно научной теорией. Он применяется при исследовании новых экономических явлений, не имеющих аналогов (изучение

эффективности новых машин и оборудования, телекоммуникационных и мобильных средств связи, себестоимости новых видов товарной продукции и т.п.).

**Аксиоматический** метод предусматривает использование аксиом, являющихся доказанными научными знаниями, которые применяются в научных исследованиях в качестве исходных положений для обоснования новой теории.

Прежде всего, это относится к использованию экономических законов, трудов классиков, научных исследований, являющихся аксиоматическими знаниями научной теории, используемой для дальнейшего развития науки.

**Создание теории** – это метод обобщения результатов исследования, нахождения общих закономерностей в поведении изучаемых объектов, а также распространения результатов исследования на другие объекты и явления, что способствует повышению надежности проведенного экспериментального исследования.

В эмпирических исследованиях применяются наряду с общенаучными также специфические методы формирования эмпирического знания прикладного характера. Это преимущественно чувственные методы человека – ощущения, восприятия и представления.

Оно становится научным тогда, когда определена их причинная связь наблюдением и экспериментом, т.е. выявлены и изучены факторы, вызвавшие превышение издержек, и намечены мероприятия по устранению недостатков.

Наблюдение – метод изучения предмета путем его количественного измерения и качественной характеристики. Применяется при изучении трудоемкости изделий путем хронометражных наблюдений, при контрольном раскрое сырья, расхода материалов, выполнения технологических операций и т.п.

**Эксперимент** – научно поставленный опыт в соответствии с целью исследования для проверки результатов теоретических исследований.

Эксперимент проводится в точно учитываемых условиях, позволяющих следить за ходом явлений и воссоздавать его повторно в заданных условиях, например проведение эксперимента в ряде отраслей народного хозяйства по применению новых систем планирования, управления и стимулирования.

Экспериментальные исследования могут проводиться в научной лаборатории с использованием специальной лабораторной установки или без нее, на предприятиях на действующих образцах продукции с использованием опытно промышленной установки или без нее, в полевых условиях с использованием определенного набора научных средств, специальных научных приборов и оборудования.

Следовательно, общенаучные методы исследования применяются во взаимной связи и обусловленности в теоретических и эмпирических исследованиях.

## **2.2. Моделирование как метод научного познания**

Моделирование в научных исследованиях стало применяться еще в глубокой древности и постепенно захватывало все новые области научных знаний: техническое конструирование, строительство и архитектуру, астрономию, физику, химию, биологию и, наконец, общественные науки. Большие успехи и признание практически во всех отраслях современной науки принес методу моделирования XX в. Однако методология моделирования долгое время развивалась независимо отдельными науками. Отсутствовала единая система понятий, единая терминология. Лишь постепенно стала осознаваться роль моделирования как универсального метода научного познания. Термин "модель" широко используется в различных сферах человеческой деятельности и имеет множество смысловых значений.

**Модель** - объект или описание объекта, системы для замещения (при определенных условиях предложениях, гипотезах) одной системы (т.е.

оригинала) другой системы для изучения оригинала или воспроизведения его каких - либо свойств.

Под **моделированием** понимается процесс построения, изучения и применения моделей. Оно тесно связано с такими категориями, как абстракция, аналогия, гипотеза и др.

**Моделирование** - метод опосредованного познания, в котором объект-модель находится в некотором неполном соответствии с объектом-оригиналом.

**Цель построения модели** – решение некоторой проблемы реального мира (анализ существующего или предполагаемого объекта или явления для принятия решения по его управлению). При построении модели выделяются те аспекты, которые существенны для решения проблемы и игнорируются те аспекты, которые незначительны, которые усложняют проблему, делают анализ реального объекта или явления менее общим или вообще невозможным.

Модели, если отвлечься от областей, сфер их применения, *бывают трех типов*: познавательные, прагматические и инструментальные.

- *Познавательная модель* - форма организации и представления знаний, средство соединения новых и старых знаний. Познавательная модель, как правило, подгоняется под реальность и является теоретической моделью.
- *Прагматическая модель* - средство организации практических действий, рабочего представления целей системы для ее управления. Реальность в них подгоняется под некоторую прагматическую модель. Это, как правило, прикладные модели.
- *Инструментальная модель* - является средством построения, исследования и/или использования прагматических и/или познавательных моделей.

Наиболее **общая классификация моделей** - деление их на *абстрактные* и *физические*.

Применительно к естественным и техническим наукам принято различать следующие виды моделирования:

- *концептуальное моделирование*, при котором совокупность уже известных фактов или представлений относительно исследуемого объекта или системы истолковывается с помощью некоторых специальных знаков, символов, операций над ними или с помощью естественного или искусственного языков;
- *физическое моделирование*, при котором модель и моделируемый объект представляют собой реальные объекты или процессы единой или различной физической природы, причем между процессами в объекте-оригинале и в модели выполняются некоторые соотношения подобия, вытекающие из схожести физических явлений;
- *структурно-функциональное моделирование*, при котором моделями являются схемы (блок-схемы), графики, чертежи, диаграммы, таблицы, рисунки, дополненные специальными правилами их объединения и преобразования;
- *математическое (логико-математическое) моделирование*, при котором моделирование, включая построение модели, осуществляется средствами математики и логики;
- *имитационное (программное) моделирование*, при котором логико-математическая модель исследуемого объекта представляет собой алгоритм функционирования объекта, реализованный в виде программного комплекса для компьютера.

**Требования к модели** определяются ситуацией и решаемой по отношению к ней задачей. Но, вне зависимости от этого, можно сформулировать ряд общих требований:

- 1) полной (т.е. предоставлять пользователю возможность получить ответы на все вопросы, в частности оценить показатели качества разрабатываемой системы);

- 2) гибкой (обеспечивать возможность воспроизведения различных ситуаций во всем диапазоне изменения условий и параметров, возможность отображения различных вариантов структуры системы);
- 3) эффективной (используемые программные и технические средства должны обеспечивать эффективное машинное исследование модели и удобное общение с ними пользователя);
- 4) укладываться в рамках имеющихся ресурсов (длительность разработки и исследования модели должны быть приемлемы в рамках имеющихся ограничений на ресурсы и время);
- 5) наглядной для построения (обозримость основных свойств и отношений);
- 6) доступной и простой для исследования и воспроизводства.

***Свойства модели:***

- 1) конечность: модель отображает оригинал лишь в конечном числе его отношений и, кроме того, ресурсы моделирования конечны;
- 2) упрощенность: модель отображает только существенные стороны объекта;
- 3) приближенность: действительность отображается моделью грубо или приблизительно;
- 4) адекватность: модель успешно описывает моделируемую систему;
- 5) информативность: модель должна содержать достаточную информацию о системе - в рамках гипотез, принятых при построении модели и полученных в результате исследований.

Традиционно под моделированием на ЭВМ понималось лишь имитационное моделирование. Можно, однако, увидеть, что и при других видах моделирования компьютер может быть весьма полезен, за исключением разве физического моделирования, где компьютер вообще-то тоже может использоваться, но, скорее, для целей управления процессом моделирования.

Моделирование на ЭВМ целесообразно использовать в следующих случаях:

- для исследования проектируемой системы на этапах, когда она еще не существует, для анализа и синтеза различных вариантов системы и параметров с целью оценки их и выбора наилучшего;
- в процессе создания и эксплуатации системы для получения информации, дополняющей результаты натуральных испытаний, и помогающей отвечать на вопросы, возникающие на этих этапах;
- для получения прогнозов поведения и эволюции системы на длительных интервалах времени.

Объектами моделирования могут быть процессы, явления, объекты реального мира, создаваемая система и т.д. – т.е. объекты изучения, чаще всего это прикладные модели. Но существует большое множество моделей в информационной сфере, которые являются инструментальными. Если смотреть на моделирование в широком смысле, то любая программа будет являться моделью. Но и в более традиционном подходе к моделированию существуют широко применяемые модели данных, знаний, процессов.

### **2.2.1. Модели данных**

Данные - это информация, представленная в определенном виде, позволяющем автоматизировать ее сбор, хранение и дальнейшую обработку человеком или информационным средством. Для компьютерных технологий данные — это информация в дискретном, фиксированном виде, удобная для хранения, обработки на компьютере, а также для передачи по каналам связи.

В современных информационных системах хранится множество разнообразных данных, которые можно разделить по форме отображения (графические, звуковые, текстовые, численные, видео), постоянно разрабатываются новые форматы организации данных, которые позволяют не только хранить в удобном виде информацию, но и предоставляют средства обработки данных. Технологии управления данных

эволюционировали от ручной обработки до интеллектуального анализа разнородной информации (рис. 2.1) [3].

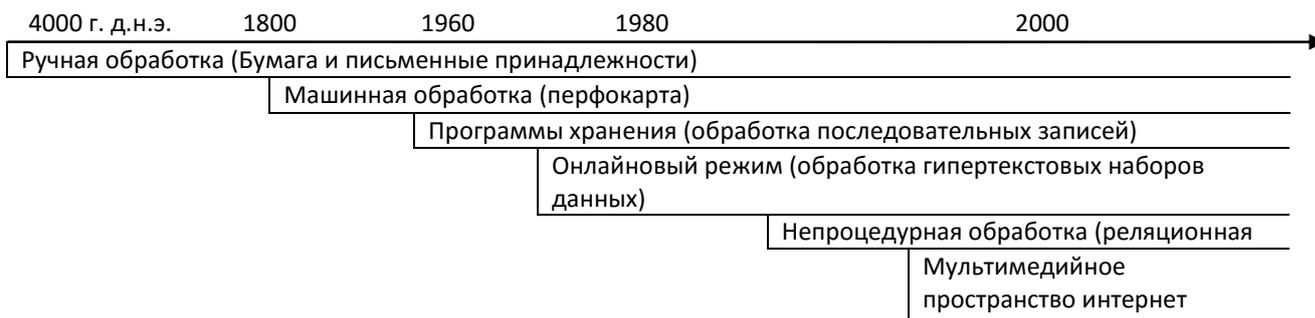


Рис 2.1. Шесть поколений технологий управления данными.

В зависимости от наличия четкой структуры в данных выделяют структурированные данные (рис. 2.2), представимые в виде наборов данных определенных типов, и полуструктурированные и неструктурированные данные, для эффективной работы с которыми требуются специальные программные средства.

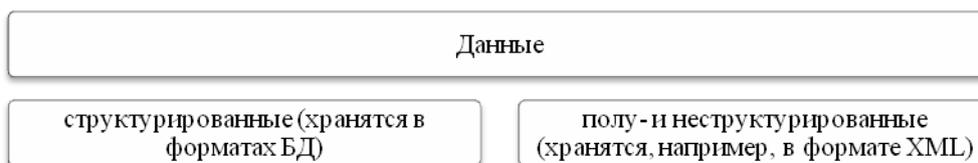


Рис 2.2. Виды данных

Наборы данных, в которых хранятся структурированные данные, можно классифицировать по следующим параметрам (рис. 2.3):

1. *По количеству переменных.* **Одномерные** наборы данных содержат только один признак для каждого объекта. Эти данные позволяют определить типичное значение признака, насколько значения отличаются друг от друга, требуют ли отдельные данные особого внимания. Наборы **двумерных** данных содержат информацию о двух признаках для каждого из объектов. Они дают возможность получить два набора одномерных данных, позволяют установить, существует ли связь между двумя переменными, насколько сильно связаны переменные, можно ли предсказать значение одной переменной по значению другой. **Многомерные** данные содержат информацию о трех или более признаках для каждого объекта. Многомерные данные

можно использовать для получения информации о том, существует ли простая зависимость между этими признаками, насколько они взаимосвязаны (речь идет не только о попарной взаимосвязи признаков, но и о зависимости в совокупности), можно ли предсказать значение одной переменной на основании значений остальных.

2. *По типу.* Значения переменных, которые регистрируются с помощью чисел, имеющих содержательный смысл, называют **количественными данными**. **Дискретная** - это такая переменная, которая может принимать значения только из некоторого списка определенных чисел. **Непрерывной** будем считать любую переменную, не являющуюся дискретной. Она принимает значения из некоторого промежутка. Данные, которые регистрируют определенное качество, которым обладает объект, называются **качественными**. Качественные данные бывают двух типов: **порядковые**, для которых существует имеющий содержательный смысл порядок, и **номинальные**, для которых нет содержательно интерпретируемого порядка.

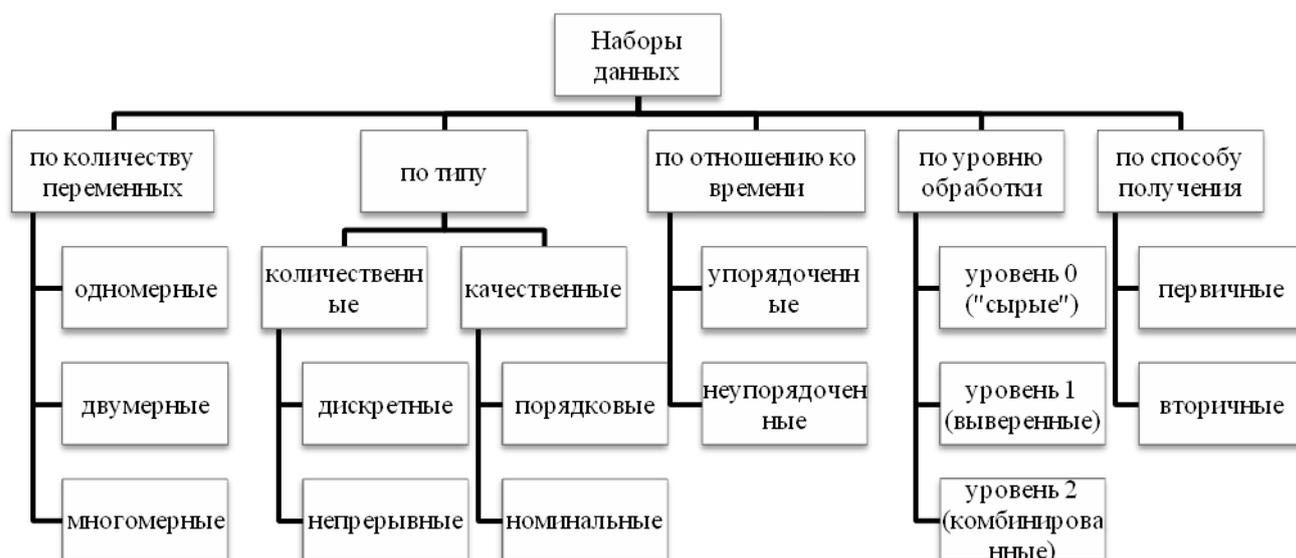


Рис. 2.3. Классификация наборов данных

3. *По отношению ко времени.* Если порядок записи значений данных во времени имеет содержательный смысл, то говорят, что эти данные представляют собой **временной ряд**. Эти данные представляют

информацию об объекте в различные моменты времени. Если порядок записи данных во времени не существен, то говорят об **одном временном срезе**.

4. *По уровню обработки.* "Сырые" данные, полученные от инструментов и в результате моделирования, обрабатываются конвейерами, которые производят стандартные продукты данных. Сырые данные уровня 0 выверяются и исправляются до наборов данных уровня 1, которые комбинируются с другими данными, порождая наборы данных уровня 2. Большая часть аналитической работы происходит с этими наборами данных уровня 2, а в случае обнаружения аномалий производится "спуск" к данным уровня 1.
5. *По способу получения.* Если данные собирались специально для запланированного анализа, то их называют **первичными** данными. Если данные собирались ранее для других задач, то данные **вторичные**.

Особое место среди данных занимают научные данные - логически организованная информация, получаемая в процессе научного познания и отображающая явления и законы природы, общества и мышления.

Научные данные предъявляют определенные требования к технологиям их обработки и хранения, вытекающих из следующих фактов:

- Научные данные обладают большим объемом. (Информация удваивается каждые 2-3 года. Ученые уже обсуждают петабайтные БД для хранения астрономических данных, объемы данных таких компаний как eBay (5 Пбайт), Wal-Mart Stores (2,5 Пбайт), Bank of America (1,5 Пбайт), Dell (1 Пбайт) уже достигли этого порога [58].)
- Научные данные обладают высоким качеством данных. (Новые научные инструменты обладают высокой точностью.)

- Огромные объемы данных невозможно обрабатывать вручную, необходима интеграция средств хранения с интеллектуальными технологиями обработки данных.
- Научные данные должны сопровождаться подробным описанием.

Для обеспечения условий выполнения этих требований создают научные центры данных, обеспечивающие доступ как к данным, так и к приложениям, которые анализируют данные, развиваются как сервисные станции для некоторой научной области. В каждом из этих научных центров производится управление одним или несколькими массивными наборами данных, а также приложениями, обеспечивающими доступ к этим наборам данных, и поддерживается персонал, который понимает данные и постоянно пополняет и совершенствует наборы данных. Это можно видеть на примерах SDSS (Sloan Digital Sky Survey - международный астрономический проект), BaBar в SLAC (Stanford Linear Accelerator Center, BaBar – международный проект в области физики высоких энергий), BIRN (Biomedical Informatics Research Network) в SDSC (San Diego Supercomputer Center), Entrez-PubMed-GenBank в NCBI (National Center for Biotechnology Information.), а также многих других наборов данных [4].

Для описания научных данных используют метаданные. Метаданные – это описательная информация о данных, которая объясняет измеряемые атрибуты, их имена, единицы измерения, точность, формат данных и т.д. Метаданные включают информацию о происхождении данных, описывающую, как измерялись, получались или вычислялись данные. Предполагается, что часть метаданных генерируется автоматически, снижая интеллектуальную нагрузку на ученого.

Большинство научных данных представлены в виде массива и, при использовании специализированной технологии, ориентированной на работу с массивами, можно превзойти систему, в которой массивы моделируются с помощью таблиц, поэтому для научных данных создают свои форматы [1].

Кроме того, научные пользователи исторически избегают применения коммерческих продуктов управления базами данных, предпочитая использовать специализированные решения, такие как HDF-5, MatLab и NetCDF [5]. Существуют специализированные форматы представления научного знания (см. таблицу 2.1) [4]. Научные форматы файлов обеспечивают лишь минимальные инструментальные средства для поиска в табличных данных и их анализа.

Основная цель этих стандартов и инструментов состоит в том, чтобы обеспечить возможность размещения таблиц и подмассивов в адресном пространстве программных сред, в которых можно манипулировать данными с использованием языка программирования.

В каждой дисциплине появляются определения общей терминологии (измерения и единицы измерений, например UniversalContentDescriptors в астрономии) в виде онтологии (называемой также управляемым словарем). В дополнение к стандартизации, пригодные для использования компьютерами онтологии помогут построить SemanticWeb: приложения станут совместимыми на семантическом уровне, а не только синтаксически совместимыми, как это возможно при использовании текущего поколения Web-сервисов на основе типизированных интерфейсов. Схему БД можно рассматривать как фрагмент онтологии. Интеграция XML в современных системах управления базами данных открывает двери для существующих стандартов RDF и OWL. Форматы научного хранения данных разрозненны и появляются проблемы при использовании их в междисциплинарных исследованиях, при построении общих онтологий и интеграции научного знания. Инструменты обработки таких форматов также ограничены.

Коммерческие системы используют стандартные модели представления данных, снабженные достаточно широким диапазоном их обработки. Таких моделей также существует множество, но они более универсальны.

Таблица 2.1. Форматы хранения научных данных

Формат	Описание	Область применения	Разработчик
NetCDF (Network Common Data Form)	Машиннонезависимый двоичный формат файлов, заголовок формата содержит информацию о содержимом файла, является открытым стандартом.	Формат самоописываемых расширяемых файлов. Применяется в климатологии, ГИС.	University Cooperation for Atmospheric Research, UCAR)
CERIF (Common European Research Information Format)	Определяет набор обязательных и дополнительных полей, которые должны использоваться для описания научных проектов.	Базы данных общего назначения для обмена информацией о научных проектах.	Европейская комиссия (EuropeanCommission)
CIF (Crystallographic Information File)	Файл CIF состоит из имени поля данных, собственно значения и утилиты для повторяющихся значений. Имена полей данных иерархические, образуют категории данных. Отсортированный список имен полей вместе с их точными определениями образуют словарь CIF.	Передача кристаллографических данных между лабораториями, журналами и базами данных.	International Union of Crystallography (IUCr)
GRIB (Gridin Binary)	Каждая запись GRIB предназначена для передачи или хранения единственного параметра вместе со значениями, которые представляют собой массив точек или набор спектральных коэффициентов одного уровня.	Хранение и обмен информацией о погоде в бинарной форме между элементами метео-грид.	World Meteorological Organization (WMO)
HDF (Hierarchical Data Format)	Самоописывающийся формат файлов для переноса различных типов данных между различными компьютерами. Библиотека HDF содержит интерфейсы для хранения и поиска, сжатых и несжатых растровых изображений с палитрами цвета, интерфейс для хранения и поиска n-размерных наборов научных данных вместе с информацией о данных.	Обмен научными базами данных.	National Center for Supercomputing Applications (NCSA)
JCAMP-DX (Joint Committee on Atomic and Molecular Physical Data Exchange Standards)	JCAMP-DX представляет собой открытый протокол, основанный на ASCII и предназначенный для обмена данными между компьютеризированными лабораторными инструментами. Протокол машиннонезависим, самодокументируется и открыт для завершения. JCAMP-DX определяет структуру файлов с данными и не является программой.	Обмен спектральной информацией в машиночитаемой форме.	International Union of Pure and Applied Chemistry Working Party on Spectroscopic Data Standards (JCAMP-DX)
SEDRIS (Synthetic Environment Data Representation and Interchange Specification)	SEDRIS состоит из модели данных, набора API и отдельных программных утилит. Модель предоставляет средства для определения всех элементов данных и их взаимосвязей, необходимых для создания искусственной среды. API используются на различных уровнях сложности. Утилиты используются для поиска и проверки в посланиях SEDRIS	Обмен данными о моделях в симуляторах военного назначения.	US Defense Modeling and Simulation Office и ISO/IEC JTC1/SC24 WG8
VICAR (Video Image Communication And Retrieval).	VICAR состоит из двух главных частей: меток, описывающих данный файл и области изображений, содержащей реальные данные. Метки делятся на две группы, одна в начале, другая в конце файла, хотя в норме достаточно меток в начале.	Работа с изображениями со спутников.	Multimission Image Processing Laboratory (MIPL) at the Jet Propulsion Laboratory

Модель данных - это некоторая абстракция, которая будучи применительно к конкретным данным, позволяет пользователям и разработчикам трактовать их как информацию, то есть сведения, содержащие не только данные, но и взаимосвязь между ними.

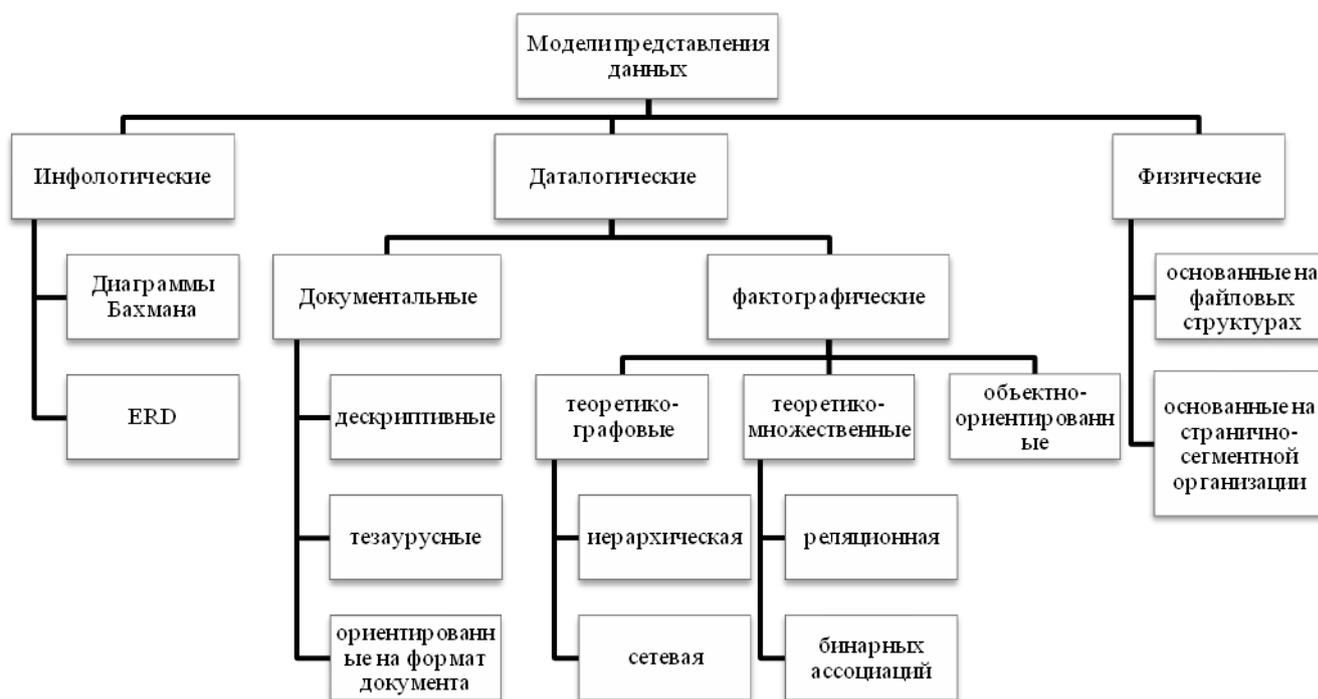
В модели данных описывается некоторый набор родовых понятий и признаков, которыми должны обладать все конкретные СУБД и управляемые ими базы данных, если они основываются на этой модели. Наличие модели данных позволяет сравнивать конкретные реализации, используя один общий язык.

Модели данных разделяют по уровням абстракции (рис. 2.4) верхний, понятный пользователям и IT-специалистам (инфологический), независимый от реализации и конкретной СУБД. Модели этого уровня называют концептуальными. Нижний уровень, определяющий расположение и кодирование данных в файловой системе под управлением конкретной операционной системы. Этот уровень скрыт полностью от пользователя и частично от разработчика (необходимы настройки модели), модели этого уровня называются физическими. Промежуточный уровень – логический – содержит отображение данных на конкретную СУБД, учитывает особенности ее реализации.

*Инфологические модели* отображают представления о содержимом базы данных пользователей и разработчиков, выполненное с использованием естественного языка, математических формул, таблиц, графиков и других средств, понятных всем людям, работающим над проектированием базы данных. Инфологические модели данных используются на ранних стадиях проектирования для описания структур данных в процессе разработки приложения [32]. Существует множество подходов к построению таких моделей: графовые модели, семантические сети, модель "сущность-связь" и т.д.

*Даталогические модели.* Документальные модели данных соответствуют представлению о слабоструктурированной информации,

ориентированной в основном на свободные форматы документов, текстов на естественном языке.



*Рис 2.4. Классификация моделей представления данных по уровням представления.*

Дескрипторные модели — самые простые из документальных моделей, они широко использовались на ранних стадиях использования документальных баз данных. В этих моделях каждому документу соответствовал дескриптор — описатель. Обработка информации в таких базах данных велась исключительно по дескрипторам, то есть по тем параметрам, которые характеризовали документ, а не по самому тексту документа.

Тезаурусные модели основаны на принципе организации словарей, содержат определенные языковые конструкции и принципы их взаимодействия в заданной грамматике. Принцип хранения информации в этих системах и подчиняется тезаурусным моделям.

Модели, ориентированные на формат документа, основаны на языках разметки документов и связаны прежде всего со стандартным общим языком разметки -- SGML (Standard Generalised Markup Language). Гораздо более простой и удобный, чем SGML, язык HTML позволяет определять

оформление элементов документа и имеет некий ограниченный набор инструкций -- тэгов, при помощи которых осуществляется процесс разметки. В качестве элемента гипертекстовой базы данных, описываемой HTML, используется текстовый файл, который может легко передаваться по сети с использованием протокола HTTP. XML (ExtensibleMarkupLanguage) -- это язык разметки, описывающий целый класс объектов данных, называемых XML-документами. Он используется в качестве средства для описания грамматики других языков и контроля за правильностью составления документов. То есть сам по себе XML не содержит никаких тэгов, предназначенных для разметки, он просто определяет порядок их создания.

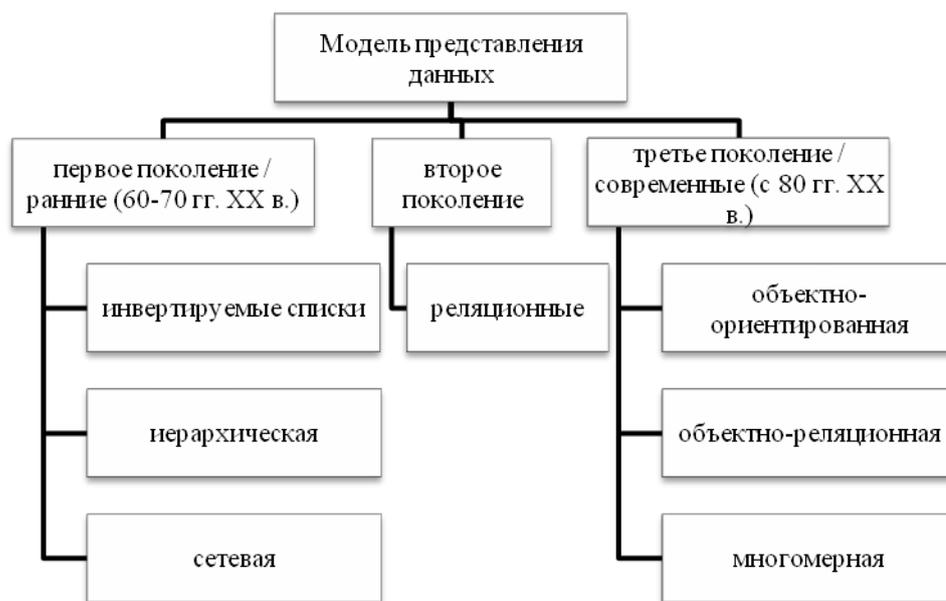
*Фактографические модели.* Теоретико-графовые модели данных -- отражают совокупность объектов реального мира в виде графа взаимосвязанных информационных объектов. В зависимости от типа графа выделяют иерархическую или сетевую модели.

*Физические модели данных* оперируют с категориями, касающимися организации внешней памяти и структур хранения, используемых в данной операционной среде. В настоящий момент в качестве физических моделей используются различные методы размещения данных, основанные на файловых структурах: это организация файлов прямого и последовательного доступа, индексных файлов и инвертированных файлов, файлов, использующих различные методы хеширования, взаимосвязанных файлов. Кроме того, современные СУБД широко используют страничную организацию данных. Физические модели данных, основанные на страничной организации, являются наиболее перспективными.

Такое разделение на уровни позволяет обеспечить логическую и физическую независимость данных.

Модели представления данных уровня логики обычно классифицируют и по временному параметру (рис. 5), выделив поколения БД, которые отражают эволюцию систем управления данными. Среди рассматриваемых (см. таблицу 2) моделей наиболее распространена реляционная (за счет

простоты понимания и реализации, отлаженного механизма поддержки целостности и развитого языка манипулирования данными).



*Рис. 2.5. Классификация моделей представления данных по времени появления.*

В современных сценариях управления данными редки случаи, когда все данные могут находиться под управлением традиционной реляционной СУБД или какой-либо другой модели данных или системы. Вместо этого разработчики часто сталкиваются с набором слабо связанных источников данных и поэтому вынуждаются каждый раз решать повторяющиеся низкоуровневые задачи управления данными в разнородных коллекциях. В число этих задач входят обеспечение возможностей поиска и запрашивания данных; соблюдение правил, ограничений целостности, соглашений об именовании и т.д.; отслеживание происхождения данных; обеспечение доступности, восстановления и контроля доступа; управляемое развитие данных и метаданных. [2]

Поэтому постоянно ведутся работы по разработки новых моделей и концепций представления данных.

Таблица 2.2. Основные модели представления данных [35]

Модель представления данных	Недостатки	Достоинства	СУБД, поддерживающие модель
Инвертированные таблицы	Ограничено количество файлов для хранения данных, количество связей между ними, длина записи и количество ее полей. Отсутствуют общие правила поддержания целостности.	Простота организации и наличие удобных языков манипулирования данными.	Datacom/DB, ADAptableDatabase System
Иерархическая	Громоздкость, сложные логические связи, сложное понимание.	Эффективное использование памяти ЭВМ, малое время выполнения операций над данными.	InformationManagementSystem, IBM
Сетевая	Высокая сложность и жесткость БД, сложность понимания и выполнения обработки информации в БД, ослаблен контроль целостности связей.	Эффективная реализация по показателям затрат памяти и оперативности; Большие возможности для образования произвольных связей.	IDMS
Реляционная	Атомарность значений атрибутов	Простота и наглядность	IBM System R, MySQL
Постреляционная	Сложное решение проблемы обеспечения целостности и непротиворечивости данных.	Возможность предоставления связанных реляционных таблиц одной постреляционной. Высокая наглядность информации и повышенная эффективность ее обработки.	Cache
Многомерная	громоздкость для простейших задач обычной оперативной обработки информации	удобство и эффективность аналитической обработки больших объемов данных, связанных со временем.	Essbase, Universe, Unidata, SciDB
Объектно-ориентированная	Высокая понятийная сложность, неудобство обработки, низкая скорость выполнения запросов.	Возможность отображения информации о сложных взаимосвязях объектов; объектно-ориентированная модель БД позволяет идентифицировать отдельную запись БД и определить функции их обработки.	GemStone, Objectivity/DB, Versant, ObjectStore
Объектно-реляционная	Сохраняет достоинства реляционного подхода	база данных сохраняет свою табличную структуру, но способ обработки некоторых полей таблиц определяется извне, т.е. программистом.	Postgres, IBM DB2 Universal Database

### 2.2.2. Модели знаний

**Знание** – форма существования и систематизации результатов познавательной деятельности человека. Знание помогает людям рационально организовывать свою деятельность и решать различные проблемы, возникающие в ее процессе; субъективный образ объективной реальности, то есть адекватное отражение внешнего и внутреннего мира в сознании человека в форме представлений, понятий, суждений, теорий.

Знание (*с точки зрения представления знаний в интеллектуальных системах*) – это связи и закономерности предметной области (принципы, модели, законы), полученные в результате практической деятельности и профессионального опыта, позволяющего специалистам ставить и решать задачи в данной области.

Знания от данных отличаются рядом **свойств**:

- внутренняя интерпретируемость;
- структурированность;
- связность;
- семантическая метрика;
- активность.

Совокупность знаний, нужных для принятия решений, принято называть предметной областью или **знаниями о предметной области**. В любой предметной области есть свои понятия и связи между ними, своя терминология, свои законы, связывающие между собой объекты данных предметной области, свои процессы и события. Кроме того, каждая предметная область имеет свои методы решения задач.

При представлении знаний в памяти интеллектуальной системы традиционные языки, основанные на численном представлении данных, являются неэффективными. Для этого используются специальные языки представления знаний, основанные на символьном представлении данных. Они делятся на типы по формальным моделям представления знаний.

Наиболее часто используется на практике классификация моделей представления знаний, приведенная на рис. 2.6, где модели представления знания делятся на детерминированные (жесткие) и мягкие.



Рис. 2.6. Классификация моделей представления знаний

Собственно мягкие модели используются для представления знаний, при этом являясь методами решения других задач, например кластеризации, классификации, оптимизации и т.д. Детерминированные модели разрабатывались непосредственно для задачи представления знаний.

Таблица 2.3. Характеристика традиционных моделей представления знаний

Модель	Достоинства	Недостатки
<b>Продукции</b>	Наглядность, высокая модульность, легкость внесения дополнений и изменений, простота механизма логического вывода, простота интерпретации.	При накоплении большого числа (нескольких сотен) продукций они начинают противоречить друг другу, возникают трудности при добавлении правил, зависящих от уже имеющихся в базе знаний, отсутствует целостный образ знаний, неясна взаимосвязей между правилами.
<b>Семантические сети</b>	Наглядность, соответствует представлениям об организации	Представляют собой пассивные структуры, для обработки которых необходим специальный аппарат формального вывода и планирования, произвольная структура и различные типы

	долговременной памяти человека, позволяет снизить объем хранимых данных.	вершин и связей усложняют процедуру обработки информации, сетевая модель не дает ясного представления о структуре предметной области.
<b>Фреймы</b>	Гибкость, наглядность, удобный способ включения процедурных знаний, сводимость к другим моделям, модульность.	Отсутствие универсальной процедуры управления выводом кроме механизма наследования, является идеологической концепцией.

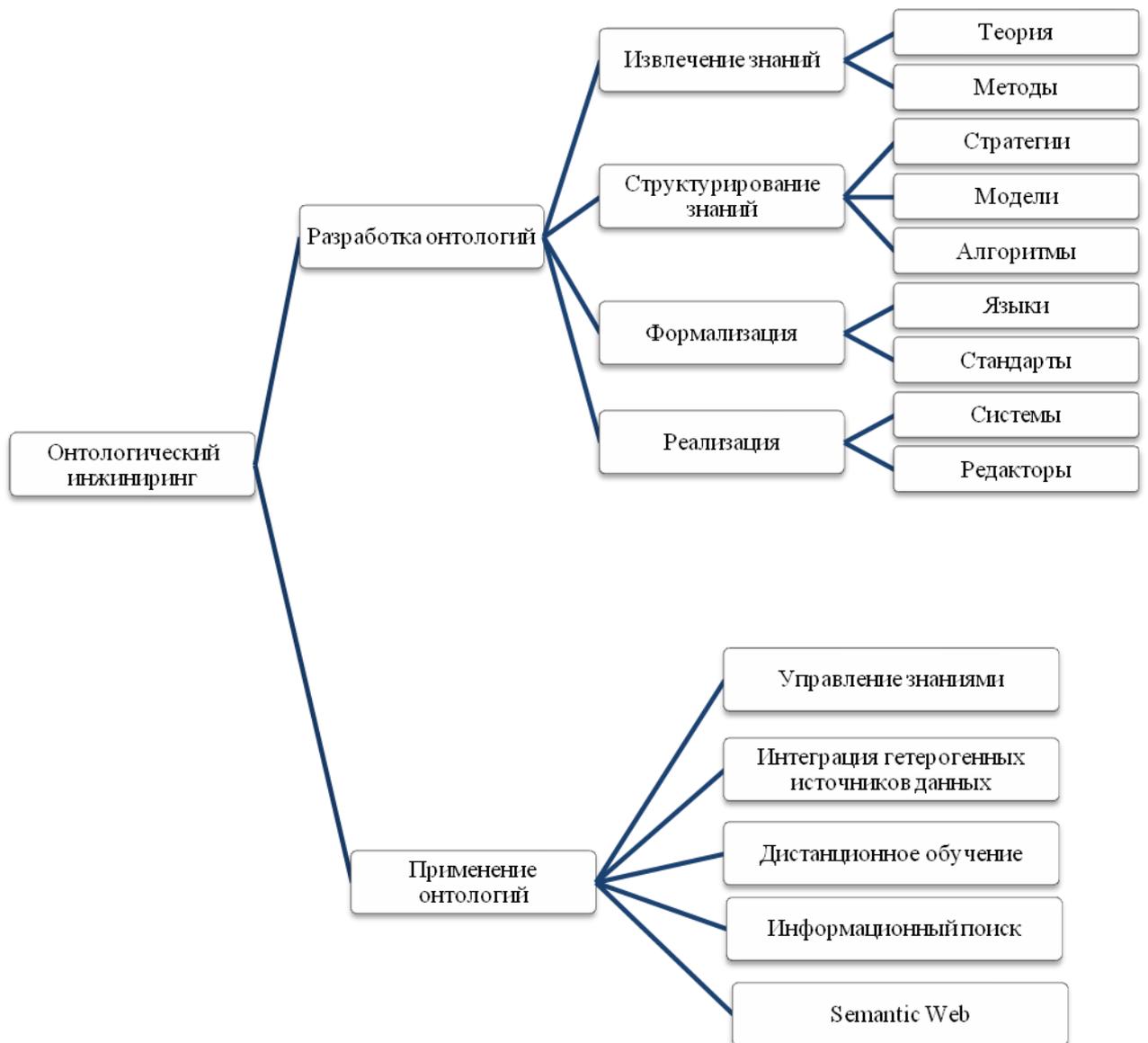
К моделям представления знаний можно отнести и онтологии.

Одно из самых известных определений онтологии, сформулированное Грубером следующее: **онтология** – это спецификация концептуализации. Под это определение подпадает ряд структур, отличающихся разной степенью формализованности: глоссарий, простая таксономия, тезаурус, полностью аксиоматизированная теория и т.д. Поэтому онтологии разделяют по степени формальности, содержанию и цели создания и ряду других признаков.

Сейчас появилось целое направление, объединяющие задачи, связанные с онтологиями – онтологический инжиниринг (рис. 2.7). В проектировании (разработке) онтологий условно можно выделить два направления, в которых по-разному рассматривались онтологии:

- онтология - формальная система, основанная на математически точных аксиомах;
- онтология - система абстрактных понятий, существующих только в сознании человека, которая может быть выражена на естественном языке (или какой-то другой системой символов).

В настоящее время данные подходы тесно взаимодействуют. Идет поиск связей, позволяющих комбинировать соответствующие методы.



*Рис.2.7. Онтологический инжиниринг*

К онтологиям можно отнести различные структуры, которые использовались и существовали и до появления онтологического инжиниринга (рис. 2.8.), а также новые форматы, разработанные на специальных языках описания онтологий, разработанных в последнее время.

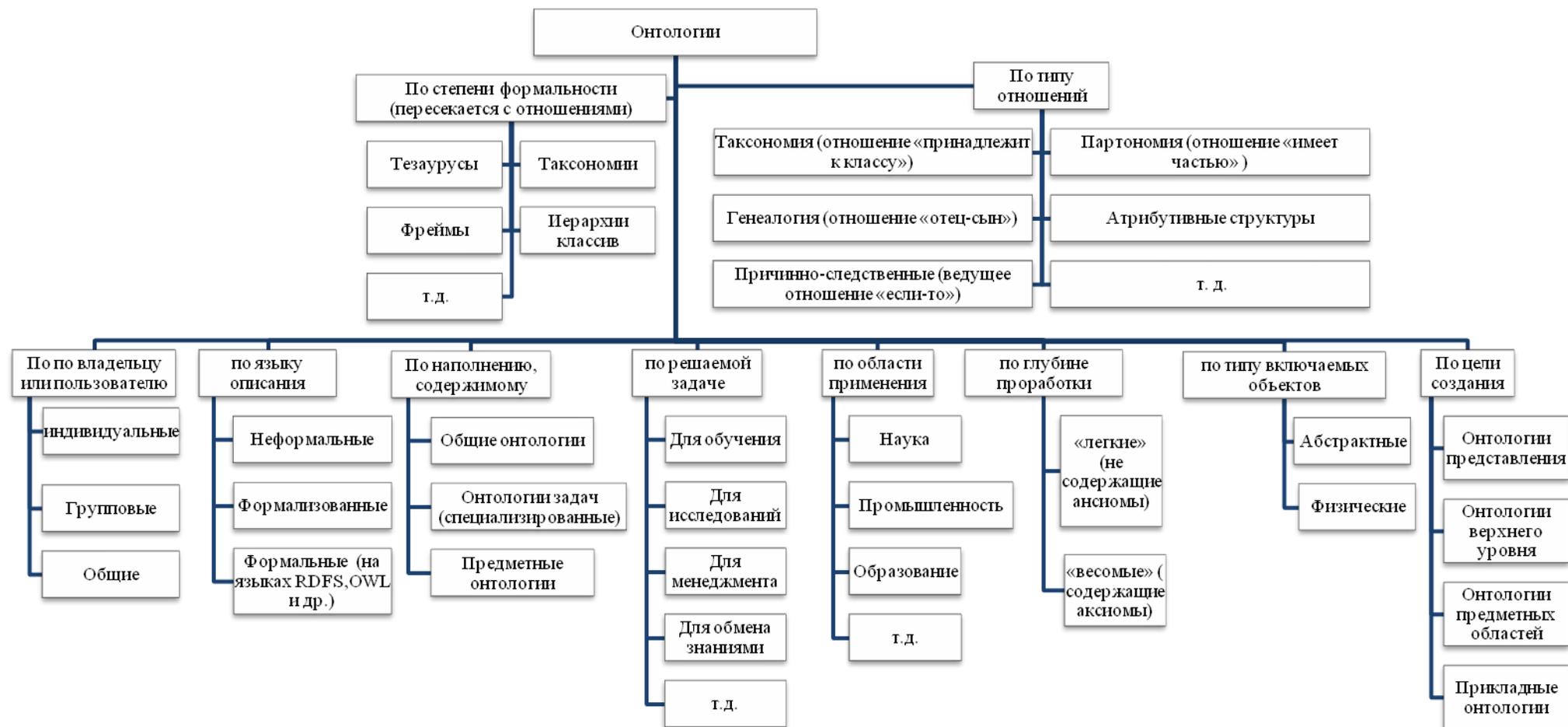


Рис. 2.8. Классификация онтологий.

## 2.2. Интеллектуальные методы извлечения знаний

Термин *приобретение знаний* носит обобщенный характер и совершенно нейтрален к способу передачи знаний. Приобретение знаний – процесс, который может проводиться человеком-специалистом (когнитологом, и тогда это процесс *извлечения знаний*) или автоматически. При работе с большими объемами накапливаемой информации необходимо постоянно оперативно отслеживать динамику процессов, а это практически невозможно без автоматизации аналитической деятельности.

В первом случае – это процесс извлечения знаний из доступных источников: специальной литературы, документации, специалистов-экспертов. Существует множество методов работы с экспертами (рис. 2.9), их оценки и представления полученных знаний в формальном виде.



Рис. 2.9. Классификация методов работы с экспертами

При автоматизированном приобретении знаний база знаний строится посредством диалога эксперта и специальной программы, изначально ориентированной на конкретный способ представления знаний. В роли таких программ могут выступать оболочки экспертных систем с развитым интерфейсом, редакторы онтологий, системы поддержки принятия решений с возможностью наполнения базы знаний и т.д.

### 2.2.1. Data Mining

Кроме этого, автоматизация процесса приобретения знаний использует методы **интеллектуального анализа данных** или **data mining** — это процесс выявления значимых корреляций, образцов и тенденций в больших объемах данных.

Data Mining использует множество существующих методов из других смежных дисциплин для своих задач. Их можно классифицировать следующим образом (рис. 2.10 ):

- **Непосредственное использование данных, или сохранение данных.** В этом случае исходные данные хранятся в явном детализированном виде и непосредственно используются на стадиях прогностического моделирования и/или анализа исключений. Проблема этой группы методов - при их использовании могут возникнуть сложности анализа сверхбольших баз данных.
- **Выявление и использование формализованных закономерностей, или дистилляция шаблонов.** При технологии дистилляции шаблонов один образец (шаблон) информации извлекается из исходных данных и преобразуется в некие формальные конструкции, вид которых зависит от используемого метода Data Mining. Этот процесс выполняется на стадии свободного поиска, у первой же группы методов данная стадия в принципе отсутствует. На стадиях прогностического моделирования и анализа исключений используются результаты стадии свободного поиска, они значительно компактнее самих баз данных. Напомним, что конструкции этих моделей могут быть трактуемыми аналитиком либо нетрактуемыми ("черными ящиками").

Другая классификация разделяет все многообразие методов Data Mining на две группы:

- Статистические методы, основанные на использовании усредненного накопленного опыта, который отражен в ретроспективных данных;

- Кибернетические методы, включающие множество разнородных математических подходов.

Эта схема разделения основана на различных подходах к обучению математических моделей.

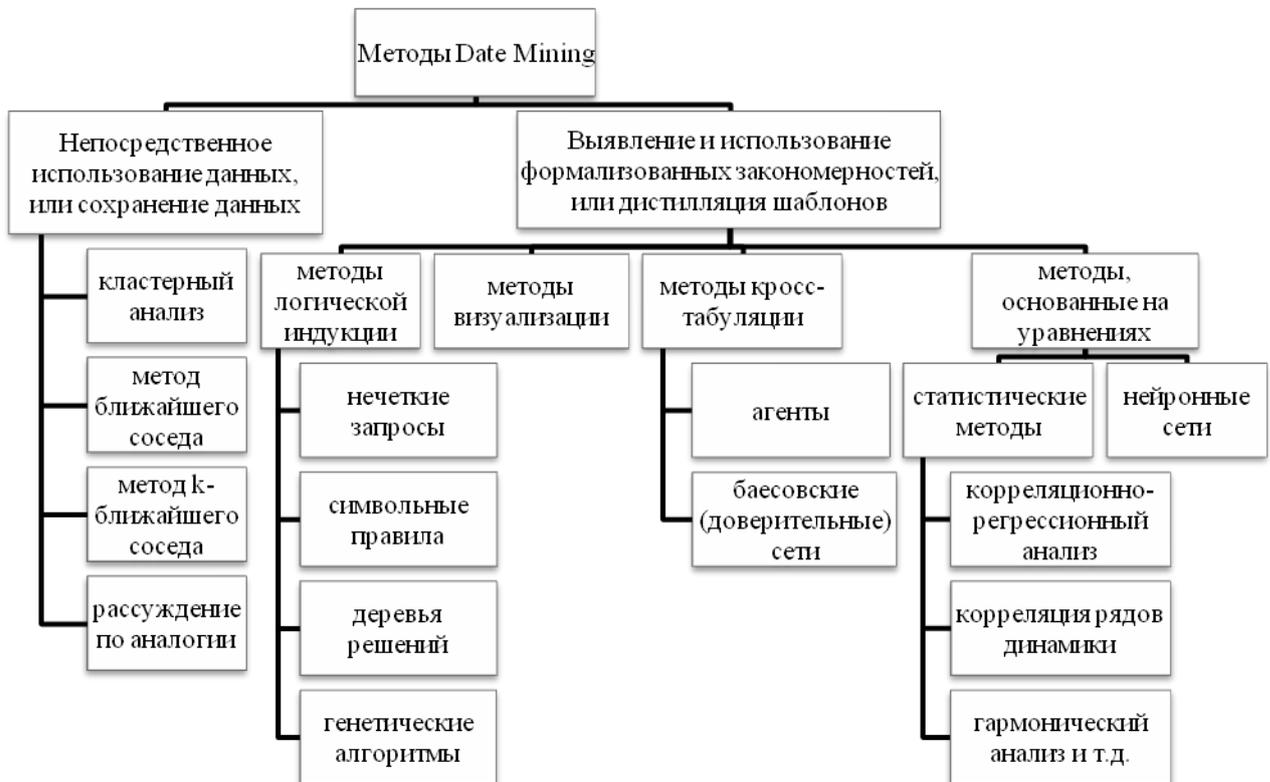


Рис. 2.10. Классификация методов приобретения (добычи) данных

Data Mining может состоять из двух или трех стадий.

**Стадия 1.** Выявление закономерностей (свободный поиск).

**Стадия 2.** Использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование). В дополнение к этим стадиям иногда вводят стадию валидации, следующую за стадией свободного поиска. Цель валидации - проверка достоверности найденных закономерностей. Однако, мы будем считать валидацию частью первой стадии, поскольку в реализации многих методов, в частности, нейронных сетей и деревьев решений, предусмотрено деление общего множества данных на обучающее и проверочное, и последнее позволяет проверять достоверность полученных результатов.

**Стадия 3.** Анализ исключений - стадия предназначена для выявления и объяснения аномалий, найденных в закономерностях.

Стандарты, имеющие прямое или опосредованное отношение к Data Mining, можно объединить в группы:

- стандарты, базирующиеся на услугах Data Mining (услуги создания модели управления, скоринговые услуги, услуги анализа данных, услуги исследования данных, статистические услуги моделирования);
- стандарты web-службы (SOAP/XML, WSRF, и т.д.), Grid-Услуги (OGSA, OGSA/DAI, и т.д.), Семантические Стандарты Web (RDF, OWL, и т.д.);
- стандарты, которые должны появиться в ближайшее время: стандарты для технологического процесса, стандарты для преобразований данных, стандарты для оперативного (real time) Data Mining, стандарты для сетей данных (data webs).

**CRISP-DM** (The Cross Industrie Standard Process for Data Mining – Стандартный межотраслевой процесс Data Mining) является наиболее популярной и распространенной методологией. Членами консорциума CRISP-DM являются NCR, SPSS и DaimlerChrysler. В соответствии со стандартом CRISP, Data Mining является непрерывным процессом со многими циклами и обратными связями. Data Mining по стандарту CRISP-DM включает следующие фазы:

1. Осмысление бизнеса (Business understanding).
2. Осмысление данных (Data understanding).
3. Подготовка данных (Data preparation).
4. Моделирование (Modeling).
5. Оценка результатов (Evaluation).
6. Внедрение (Deployment).

При помощи методологии CRISP-DM Data Mining превращается в бизнес-процесс, в ходе которого технология Data Mining фокусируется на решении конкретных проблем бизнеса.

CRISP-DM - это не единственный стандарт, описывающий методологию Data Mining. Помимо него, можно применять такие известные методологии, являющиеся мировыми стандартами, как Two Crows, SEMMA, а также методологии организации или свои собственные.

**SEMMA** методология реализована в среде SAS Data Mining Solution (SAS). Ее аббревиатура образована от слов Sample ("Отбор данных", т.е. создание выборки), Explore ("Исследование отношений в данных"), Modify ("Модификация данных"), Model ("Моделирование взаимосвязей"), Assess ("Оценка полученных моделей и результатов").

Подход SEMMA подразумевает, что все процессы выполняются в рамках гибкой оболочки, поддерживающей выполнение всех необходимых работ по обработке и анализу данных. Подход SEMMA сочетает структурированность процесса и логическую организацию инструментальных средств, поддерживающих выполнение каждого из шагов.

**PMML (Predictive Modeling markup Language)** - языке описания предикторных (или прогнозных) моделей или языке разметки для прогнозного моделирования. PMML относится к группе стандартов по хранению и передаче моделей Data Mining. Разработка и внедрение этого стандарта ведется IT-консорциумом DMG (Data Mining Group). DMG - группа, в которую входят все лидирующие компании, разрабатывающие программное обеспечение в области анализа данных. Основа этого стандарта - язык XML. Примером другого стандарта, также основанного на языке XML, является стандарт обмена статистическими данными и метаданными.

Стандарт PMML используется для описания моделей Data Mining и статистических моделей. Основная цель стандарта PMML - обеспечение возможности обмена моделями данных между программным обеспечением разных разработчиков.

В организации MDC (MetaData Coalition, [www.mdcinfo.com](http://www.mdcinfo.com)) и OMG (Object Management Group, [www.omg.org](http://www.omg.org)), разрабатывающие два конкурирующих стандарта – в области интеллектуальных технологий для бизнеса - **OIM** (Open Information Model) и **CWM** (Common Warehouse Metamodel) - общую метамоделю хранилищ данных решили объединить свои достижения и усилия под управлением OMG.

Стандарт CWM включает описание базовых элементов объектной модели, реляционных отношений, языка XML, структуры семантики предметной области, архитектуры OLAP, добычи данных, технологии перегрузки данных и некоторых расширений.

**JDM** (The Java Data Mining standard - Java Specification Request 73, JSR-73). Стандарт, разработанный группой JSR 73, Java Data Mining API (JDM) - это первая попытка создать стандартный Java API (программный интерфейс приложения) для получения доступа к инструментам Data Mining из Java-приложений. Вторая группа стандартов направлена на разработку надстройки над языком SQL, которая позволяла бы обращаться к инструментарию Data Mining, встроенному непосредственно в реляционную базу данных. К этой группе можно отнести следующие стандарты: SQL/MM, OLE DB for Data Mining.

**Стандарт SQL/MM** представляет собой набор определенных пользователем SQL процедур для возможностей вычислений и использований моделей Data Mining.

**The OLE DB for Data Mining standard of Microsoft.** Этот стандарт позволяет, подобно SQL/MM, применять методы Data Mining в структуре реляционных баз данных. Этот стандарт является расширением OLE DB.

Сейчас существует множество различных систем, ориентированных на методы Data Mining, кроме того, разработчики универсальных статистических пакетов в дополнение к традиционным методам статистического анализа, включают в пакет определенный набор методов Data Mining. Это такие пакеты как SPSS (SPSS, Clementine), Statistica

(StatSoft), SAS Institute (SAS Enterprise Miner). Некоторые разработчики OLAP-решений также предлагают набор методов Data Mining, например, семейство продуктов Cognos. Есть поставщики, включающие Data Mining решения в функциональность СУБД: это Microsoft (Microsoft SQL Server), Oracle, IBM (IBM Intelligent Miner for Data).

Системы можно разделить по назначению: инструменты общего и специфического- а также по цене- бесплатные и коммерческие инструменты.

Наиболее популярная группа инструментов содержит следующие категории:

- наборы инструментов;
- классификация данных;
- кластеризация и сегментация;
- инструменты статистического анализа;
- анализ текстов (Text Mining), извлечение отклонений;
- инструменты визуализации.

*Таблица 2.4. Примеры систем Data Mining*

<b>Название</b>	<b>Описание</b>
Apriori	инструмент для нахождения ассоциативных правил при помощи алгоритма Apriori
ARtool ( <a href="http://www.cs.umb.edu/">http://www.cs.umb.edu/</a> )	инструмент содержит набор алгоритмов для поиска ассоциативных правил в бинарных базах данных (binary databases)
Azmy SuperQuery ( <a href="http://www.azmy.com/">http://www.azmy.com/</a> )	поисковик ассоциативных правил
Clementine ( <a href="http://www.spss.com/clementine">http://www.spss.com/clementine</a> ).	Data Mining с использованием Clementine является бизнес-процессом, разработанным для минимизации времени решения задач. Clementine поддерживает процесс Data Mining: доступ к данным, преобразования, моделирование, оценивание и внедрение. При помощи Clementine Data Mining выполняется с методологией CRISP-DM.
ClustanGraphics3, ( <a href="http://www.clustan.com/">http://www.clustan.com/</a> )	· иерархический кластерный анализ "сверху вниз", поддерживаются мощные графические возможности
CViz Cluster Visualization, ( <a href="http://www.alphaworks.ibm.com/tech/cviz">http://www.alphaworks.ibm.com/tech/cviz</a> )	продукт для анализа наборов данных с большой размерностью, обеспечивает визуализацию наполнения кластеров объектами
DBMiner 2.0 Enterprise ( <a href="http://www.dbminer.com">http://www.dbminer.com</a> )	инструмент для исследования больших баз данных; использует Microsoft Сервер SQL 7.0 Plato.
DM-II system ( <a href="http://www.comp.nus.edu.sg/">http://www.comp.nus.edu.sg/</a> )	инструмент включает алгоритм СВА для выполнения классификации на основе ассоциативных правил и некоторых других характеристик

FIMI, Frequent Itemset Mining Implementations ( <a href="http://fimi.cs.helsinki.fi/">http://fimi.cs.helsinki.fi/</a> )	является репозиторием, включающим программное обеспечение и базы данных
IBM Intelligent Miner for Data ( <a href="http://www.ibm.com/software/data/iminer/fordata/">http://www.ibm.com/software/data/iminer/fordata/</a> ).	Data Mining процесс: от подготовки данных до презентации результатов. Поддержка языков XML и PMML.
IREX ( <a href="http://www.giwebb.com">http://www.giwebb.com</a> )	сегментирование данных с целью оптимизации числовых результатов, например, прибыли
KXEN (Knowledge eXtraction ENgines).	· Инструмент, работающий на основе теории Вапника (Vapnik) SVM. Решает задачи подготовки данных, сегментации, временных рядов и SVM-классификации.
Magnum Opus ( <a href="http://www.rulequest.com/MagnumOpus-info.html">http://www.rulequest.com/MagnumOpus-info.html</a> )	является быстрым инструментом поиска ассоциативных правил в данных, поддерживается операционными системами Windows, Linux и Solaris
Megaputer Polyanalyst Suite ( <a href="http://www.megaputer.com/">http://www.megaputer.com/</a> )	включает машину поиска ассоциативных правил
Nuggets ( <a href="http://www.data-mine.com/">http://www.data-mine.com/</a> )	это набор, включающий поиск ассоциативных правил и другие алгоритмы
Oracle Data Mining (ODM) ( <a href="http://otn.oracle.com/products/bi/9idmining.html">http://otn.oracle.com/products/bi/9idmining.html</a> ).	Инструмент обеспечивает GUI, PL/SQL-интерфейсы, Java-интерфейс. Используемые методы: байесовская классификация, алгоритмы поиска ассоциативных правил, кластерные методы, SVM и другие.
Polyanalyst ( <a href="http://www.megaputer.com/">http://www.megaputer.com/</a> ).	· Набор, обеспечивающий всесторонний Data Mining. Сейчас, помимо методов прежних версий, также включает анализ текстов, лес решений, анализ связей. Поддерживает OLE DB for Data Mining и DCOM-технологию.
Purple Insight MineSet	является набором визуального Data Mining, включающим визуализатор ассоциативных правил
SAS Enterprise Miner ( <a href="http://www.sas.com/">http://www.sas.com/</a> ).	· Интегрированный набор, который обеспечивает дружественный GUI. Поддерживается методология SEMMA.
SPSS ( <a href="http://www.spss.com/clementine/">http://www.spss.com/clementine/</a> ).	· Один из наиболее популярных инструментов, поддерживается множество методов Data Mining.
Statistica Data Miner ( <a href="http://www.StatSoft.com/">http://www.StatSoft.com/</a> ).	Инструмент обеспечивает всесторонний, интегрированный статистический анализ данных, имеет мощные графические возможности, управление базами данных, а также приложение разработки систем.
The LPA Data Mining Toolkit ( <a href="http://www.lpa.co.uk/dtm.htm">http://www.lpa.co.uk/dtm.htm</a> )	поддерживает поиск ассоциативных правил в реляционных базах данных.
Weka ( <a href="http://www.cs.waikato.ac.nz/ml/weka/index.html">http://www.cs.waikato.ac.nz/ml/weka/index.html</a> ).	пакет Weka представляет собой набор алгоритмов машинного обучения для решения реальных Data Mining-проблем. Weka написана на Java и запускается практически со всех платформ.
Wizsoft модуль WizRule	нахождение ассоциативных правил и потенциальных ошибок данных
XAffinity(TM)	используется для идентификации сходств или шаблонов в транзакциях

### 2.2.1.1. Нейронные сети

Искусственная нейронная сеть – математическая модель, реализуемая программно или аппаратно, построенная по подобию естественных нейронных сетей (сетей нервных клеток живого организма), представляющая собой соединение простых взаимодействующих между собой процессоров - искусственных нейронов.

Схема искусственного нейрона представлена на рис. 2.11, где  $X_1..X_N$  – входы нейрона,  $W_1..W_N$  – синаптические веса связей нейрона,  $S$  – взвешенная сумма входных значений нейрона,  $F(S)$  – функция активации, значением которой является  $Y$  – выходное значение нейрона.

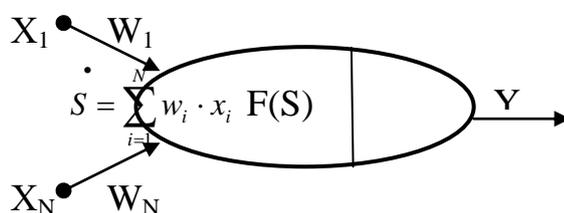
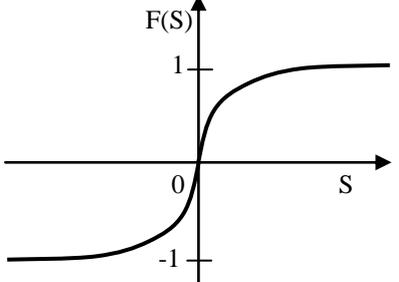


Рис. 2.11. Формальный нейрон (математическая модель)

Функции активации могут различны, наиболее часто используемые представлены в таблице 2.5.

Таблица 2.5. Основные функции активации.

Название	Формула	График (пример)
Пороговая	$F(S) = \begin{cases} 1, & \text{при } S \geq T, \\ 0, & \text{при } S < T, \end{cases}$ $T = \text{const.}$	
Линейная	$F(S) = k \cdot S,$ $k = \text{const.}$	
Сигмоидальная	$F(S) = \frac{1}{1 + e^{-S \cdot k}},$ $k = \text{const.}$	

Гиперболический тангенс	$F(S) = th\left(\frac{S}{k}\right) = \frac{e^{\frac{S}{k}} + e^{-\frac{S}{k}}}{e^{\frac{S}{k}} - e^{-\frac{S}{k}}},$ $k = const.$	
-------------------------	---	---

Существует множество нейронных сетей, которые классифицируются по нескольким признакам (таблица 2.6). Наибольшее распространение получили слоистые сети прямого распространения.

Таблица 2.6. Типы искусственных нейронных сетей.

Тип	Описание
<b>По топологии</b>	
Полносвязные	Каждый нейрон связан с другим нейроном в сети (из-за высокой сложности обучения не используется).
Слоистые	Нейроны располагаются слоями, каждый нейрон последующего слоя связан с нейронами предыдущего. Есть однослойные и многослойные сети.
<b>По типу связей</b>	
Прямого распространения	Все связи между нейронами идут от выходов нейронов предыдущего слоя к входам нейронов последующего.
Рекуррентные	Допускаются связи выходов нейронов последующих слоев с входами нейронов предыдущих.
<b>По организации обучения</b>	
С учителем	При обучении используются обучающие выборки, в которых определены требуемые от сети выходные значения, такие сети используют для решения задач классификации.
Без учителя	Нейронная сеть сама в процессе работы выделяет классы объектов и относит объект к определенному классу, такие сети используют для задач кластеризации.
<b>По типу сигнала</b>	
Бинарные	На вход нейронных сетей подают только нули или единицы.
Аналоговые	Подаваемые на входы нейронов сигналы могут быть произвольными (вещественными числами).
<b>По типу структур</b>	

Однородная	Все нейроны в нейронной сети используют одну функцию активации.
Неоднородная	Нейроны в нейронной сети имеют разные функции активации.

Для решения конкретной задачи нужно выбрать подходящую нейронную сеть. При этом нужно учитывать не только перечисленные в таблице критерии, но и архитектуру сети. Выбор архитектуры подразумевает определение количества слоев и нейронов в этих слоях. Не существует формального алгоритма по определению нужной архитектуры, поэтому на практике выбирают или заведомо маленькую сеть и постепенно ее наращивают или заведомо большую и постепенно выявляют неиспользуемые связи и сокращают сеть.

Нейронная сеть, прежде чем использоваться на практике для решения какой-либо задачи, должна быть обучена. Обучение нейронной сети - это процесс настройки синаптических весов. Существует множество алгоритмов, ориентированных на определенные типы сетей и на конкретные задачи, рассмотрим алгоритмы для однослойной и многослойной сетей.

Задачи, решаемые нейронными сетями

1. **Классификация образов.** Задача состоит в определении принадлежности входного образа (например, языкового сигнала или рукописного символа), представленного вектором признаков, к одному или нескольким предварительно определенным классам. К известным приложениям относятся распознавание букв, распознавание языка, классификация сигнала электрокардиограммы, классификация клеток крови.

2. **Кластеризация/категоризация.** При решении задачи кластеризации обучающее множество не имеет меток классов. Алгоритм кластеризации основан на подобии образов и помещает похожие образы в один кластер. Известны случаи применения кластеризации для добычи знаний, сжатия данных и исследования свойств данных.

3. **Аппроксимация функций.** Предположим, что есть обучающая выборка  $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$  (пары данных вход-выход), которая

генерируется неизвестной функцией  $F$ , искаженной шумом. Задача аппроксимации состоит в нахождении неизвестной функции  $F$ . Аппроксимация функций необходима при решении многочисленных инженерных и научных задач моделирования.

4. **Предвидение/прогноз.** Пусть заданы  $n$  дискретных отсчетов  $\{y(t_1), y(t_2), \dots, y(t_n)\}$  в последовательные моменты времени  $t_1, t_2, \dots, t_n$ . Задача состоит в предвидении значения  $y(t_{n+1})$  в следующий момент времени  $t_{n+1}$ . Предвидение/прогноз имеют большое значение для принятия решений в бизнесе, науке и технике (предвидение цен на фондовой бирже, прогноз погоды).

5. **Оптимизация.** Многочисленные проблемы в математике, статистике, технике, науке, медицине и экономике могут рассматриваться как проблемы оптимизации. Задачей алгоритма оптимизации является нахождение такого решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию.

6. **Память, адресуемая по содержанию.** В традиционных компьютерах обращение к памяти доступно только с помощью адреса, не зависящего от содержания памяти. Более того, если допущена ошибка в вычислении адреса, то может быть найдена совсем другая информация. Ассоциативная память или память, адресуемая по смыслу, доступна по указанию заданного содержания. Содержимое памяти может быть вызвано даже по частичному входу или при поврежденном содержании. Ассоциативная память может быть использована в мультимедийных информационных базах данных.

7. **Управление.** Рассмотрим динамическую систему, заданную совокупностью  $\{u(t), y(t)\}$ , где  $u(t)$  – входное управляющее воздействие, а  $y(t)$  – выход системы в момент времени  $t$ . В системах управления с эталонной моделью целью управления является расчет такого входного воздействия  $u(t)$ , при котором система действует по желательной траектории, заданной эталонной моделью. Примером является оптимальное управление двигателем.

### 2.2.1.2. Статистические методы

Целесообразно выделить три вида научной и прикладной деятельности в области статистических методов анализа данных (по степени специфичности методов, сопряженной с погруженностью в конкретные проблемы):

- а) разработка и исследование методов общего назначения, без учета специфики области применения;
- б) разработка и исследование статистических моделей реальных явлений и процессов в соответствии с потребностями той или иной области деятельности;
- в) применение статистических методов и моделей для статистического анализа конкретных данных.

Статистических методов множество, некоторые применяются широко, другие ориентированы на решение конкретной прикладной задачи. В таблице 2.7. приведены примеры наиболее распространенных методов.

Таблица 2.7. Некоторые статистические методы обработки данных

Метод	Описание
<i>Дискриминантный анализ</i>	Статистический метод, используемый для прогнозирования вероятности какого-либо события. Относится к методам классификации с обучением. Используется для разделения респондентов в различающиеся между собой группы на основе некоторых характеристик. Обычно зависимая переменная номинальная или порядковая, а независимые переменные (предикторы) - метрические (интервальные).
<i>Дисперсионный анализ</i>	Метод статистического анализа, позволяющий определить достоверность гипотезы о различиях в средних значениях на основании сравнения дисперсий распределений. Этот метод имеет смысл только лишь для интервальных переменных с наложенными дополнительными ограничениями.
<i>Кластерный анализ</i>	Кластерный анализ - это совокупность методов, позволяющих классифицировать многомерные наблюдения, каждое из которых описывается неким набором переменных. Целью кластерного анализа является образование групп схожих между собой объектов, которые принято называть кластерами. Слова кластер английского происхождения (cluster), переводится как сгусток, пучок, группа. Родственные понятия, используемые в литературе, - класс, таксон, сгущение. В отличие от комбинационных группировок кластерный анализ приводит к разбиению на группы с учетом всех группировочных признаков одновременно.
<i>Регрессионный</i>	Статистический метод установления зависимости между

<b>анализ</b>	независимыми и зависимыми переменными. Регрессионный анализ на основе построенного уравнения регрессии определяет вклад каждой независимой переменной в изменение изучаемой (прогнозируемой) зависимой переменной величины. Выделяют два вида регрессионного анализа - парный регрессионный анализ и анализ на основе множественной регрессии. В маркетинге часто используется для определения линейной связи интервальных переменных. Если зависимая переменная является дихотомической или категориальной, необходимо использовать логистическую регрессию.
<b>Совместный анализ</b>	Метод анализа, предназначенный для оценки и сравнения атрибутов продуктов с целью выявления тех из них, которые оказывают наибольшее влияние на покупательские решения. Потребителю обычно нелегко сформулировать свою систему ценностей. Поэтому вместо того, чтобы вынуждать респондентов думать о каждом атрибуте по отдельности, в рамках С.а. потребители выносят свои суждения о продукте в целом. Затем по результатам анализа определяется система ценностей, лежащая в основе их выбора. Ключевым понятием С.а. является "полезность" (utility) - способность товара удовлетворять потребности потребителя.
<b>Факторный анализ</b>	Совокупность методов, которые на основе реально существующих связей признаков (или объектов) позволяют выявлять латентные (или скрытые) обобщающие характеристики структуры и механизма развития изучаемых явлений и процессов. Понятие латентности в определении ключевое. Оно означает неявность характеристик, раскрываемых при помощи методов факторного анализа. Толчком для развития методов факторного анализа изначально послужили работы в области психологии и практически все первые работы по факторному анализу были напечатаны в журналах "Психология и "Психометрика". Позже методы факторного анализа стали активно использоваться в социологических исследованиях, медицине, военной науке и экономике. В маркетинге исследователи обратили внимание на методы факторного анализа в связи с углублением анализа потребительского поведения, развитием психографики и т.п. задач, в которых необходимо выявление явно не наблюдаемых факторов. Необходимым условием расширения сферы применения факторного анализа является математизация исследуемого процесса и применение компьютерных программ обработки данных.
<b>Анализ временных рядов</b>	Один из методов прогнозирования спроса; основан на разбивке данных об объеме продаж в прошлом на компоненты, характеризующие тренды, циклические (например, циклы деловой активности) и сезонные колебания, а также случайные изменения. В.р.а. предполагает выявление причин изменения спроса в прошлом для прогнозирования спроса в будущем.
<b>Корреляционный анализ</b>	метод обработки статистических данных, с помощью которого измеряется теснота связи между двумя или более переменными. Корреляционный анализ тесно связан с регрессионным анализом (также часто встречается термин «корреляционно-регрессионный

	<p><i>анализ</i>», который является более общим статистическим понятием), с его помощью определяют необходимость включения тех или иных факторов в уравнение множественной регрессии, а также оценивают полученное уравнение регрессии на соответствие выявленным связям (используя коэффициент детерминации).</p>
--	--

### 2.2.2. Эвристическое программирование

Эвристика (от др.-греч. «нахожу», «открываю») — отрасль знания, изучающая творческое, неосознанное мышление человека. Эвристика связана с психологией, физиологией высшей нервной деятельности, кибернетикой и другими науками, но сама как наука ещё полностью не сформировалась.

Эвристические методы предполагают интеллектуальный анализ задач с выявлением основных особенностей, приводящих к сокращению перебора вариантов решения без строгого доказательства. Такие методы увеличивают вероятность получения работоспособного - но не всегда оптимального - решения творческой задачи, возникшей, например, из-за неразработанности конкретной теории, неполноты или недостоверности исходных данных.

К эвристическим методам относятся:

- метод эвристических вопросов;
- метод многомерных матриц;
- метод свободных ассоциаций;
- метод инверсии;
- метод эмпатии (метод личной аналогии);
- упорядоченный поиск (применение теории решений);
- система обучения, берущую свои истоки от сократовской майевтики;
- метод коллективного поиска оригинальных идей, базирующийся на следующих психолого-педагогических закономерностях и соответствующих им принципах;
- интервьюирование потребителей;
- анкетный опрос;
- исследование поведения потребителей;

- методика ТРИЗ;
- мозговая атака, мозговой штурм;
- морфологический анализ (морфологические карты, морфологический ящик);
- анализ взаимосвязанных областей решения (AIDA) и т.д.

Примерами эвристических моделей являются:

- модель слепого поиска, которая опирается на метод проб и ошибок;
- лабиринтная модель, в которой решаемая задача рассматривается как лабиринт, а процесс поиска решения — как блуждание по лабиринту;
- структурно-семантическая модель, которая исходит из того, что в основе эвристической деятельности по решению задачи лежит принцип построения системы моделей, которая отражает семантические отношения между объектами, входящими в задачу.

Научно-методологической основой эвристического принципа является *эвристическое программирование*, которое представляет собой раздел прикладной математики, разрабатывающий специфические методы решения таких задач, для которых вообще не существует или заранее не известны какие-либо строгие критерии применимости современных научных методов.

Сточки зрения эвристического программирования эвристика - это правило, теоретически не обоснованное, которое позволяет сократить количество переборов в пространстве поиска.

Цель *эвристического программирования* состоит в том, чтобы, с одной стороны, интенсифицировать мыслительные процессы творческой человеческой деятельности, а с другой - глубже изучить и познать их сущность. В соответствии с этой концепцией в принципе любой вид умственной деятельности человека может быть формализован в виде некоторой динамической модели на ЭВМ, ибо ЭВМ позволяет выполнять простые логические преобразования, определенная последовательность

которых дает любые правила перехода от одной элементарной информационной модели к другой.

Обычная процедура построения моделей методом *эвристического программирования* строится следующим образом. Испытуемым предлагается решать некоторую задачу, сопровождая свои размышления устными комментариями хода своих рассуждений. Все высказывания испытуемых тщательно протоколируются. Полученный в ходе анализа материал используется при составлении компьютерной программы - модели данного вида поведения. Таким образом, программа является моделью не испытуемого, а протокола. Такая модель должна выполнять то, что делает испытуемый и так, как это делает он.

### **3. Этапы научного исследования в информатике**

В самом обобщенном виде можно выделить следующие основные *этапы научного исследования*:

- выявление проблемной ситуации и постановка проблемы;
- выделение объекта и предмета исследования;
- выдвижение первоначальной гипотезы;
- теоретическое исследование;
- экспериментальные исследования;
- анализ и интерпретация экспериментальных данных;
- обоснование гипотезы (теории);
- оформление и внедрение результатов.

В зависимости от специфики науки этапы исследования уточняются и варьируются.

#### **3.1. Особенности научного исследования в информатике**

Информатика триединая и представляет собой, с одной стороны, *фундаментальную научную дисциплину*, с другой стороны, *прикладную научную дисциплину*, с третьей – *область практической деятельности* человечества.

Объединяющими эти три стороны информатики являются следующие ее *объекты* (предметы) изучения, исследования, создания или использования:

- а) собственно информация;
- б) информационные модели, процессы, технологии, средства, структуры, ресурсы, отношения с другими “информационными” объектами (включая субъекты) Вселенной.

В общем виде *предметами* (объектами) *исследования* (внимания) информатики как науки являются познание, изучение сущности и

закономерностей строения, функционирования, поведения, развития указанных выше объектов (предметов) исследования.

*Целями изучения информатики, как и любой другой фундаментальной науки, являются познание, описание, объяснение и предсказание (на основе выявленных ею законов, закономерностей и построенных теорий) сущности, свойств, особенностей и закономерностей строения и “жизни” предметов (объектов) исследования по отношению к объектам (предметам) ее изучения.*

*Целью и объектом исследования информатики как прикладной науки является разработка теоретических основ создания и эффективного применения информационных моделей, технологий, систем, других информационных средств и ресурсов.*

Наконец, *предметами (объектами) информатики как области практической деятельности человечества являются разработка, изготовление, выпуск и разностороннее эффективное использование информационных моделей, технологий, систем, ресурсов, услуг, а также получение, изучение, анализ, передача, использование информации в нужном объеме, заданного качества, в требуемые сроки при минимальной себестоимости.*

С другой стороны, информатику можно рассматривать с точки зрения решаемых в ней задач, предметов, областей и методов исследования. Тогда ее укрупненно можно считать состоящей из следующих четырех частей: *формальной, технической, технологической и прикладной*, которые, в свою очередь, как подсистемы делятся на отдельные взаимосвязанные, пересекающиеся, недостаточно четко очерченные, отделяемые друг от друга части (рис. 3.1).

*Формальная информатика* изучает и разрабатывает прежде всего теоретические основы сущности, закономерностей информации, теорию и практику методов и средств моделирования объектов разной природы, выявления, представления, запоминания, использования знаний; сбора, обработки, анализа и интерпретации данных, обнаружения закономерностей,

распознавания образов; управления, диагностики, предсказания поведения и проектирования объектов; алгоритмизации и программирования и т.д.

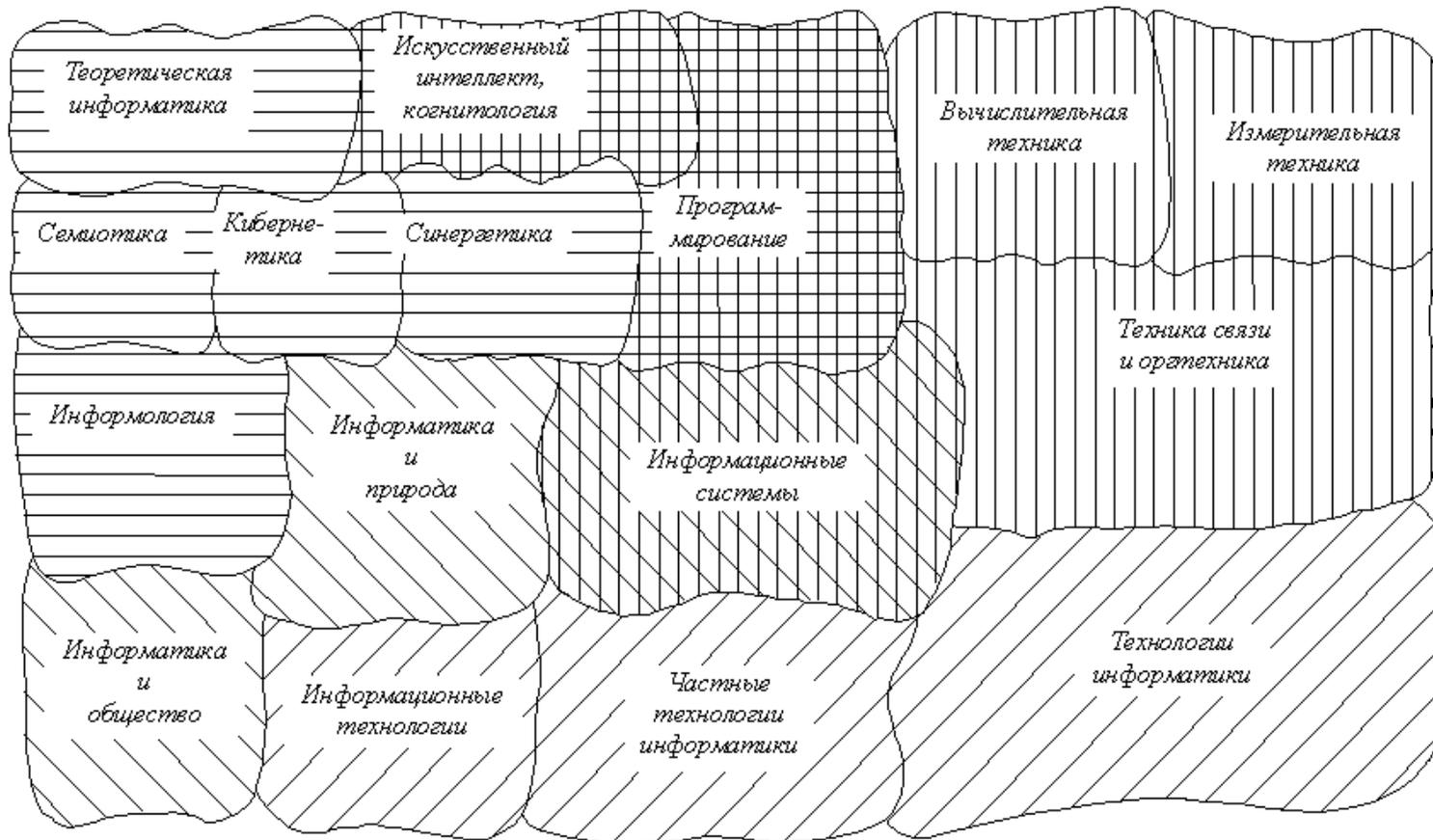


Рис. 3.1. Условное изображение структуры информатики:

- ════════ — формальная информатика | | — техническая информатика;  
 ════════ — технологическая информатика ════════ — прикладная информатика

*Техническая информатика* занимается теорией и практикой создания, применения и сопровождения аппаратных (технических) средств информатики – средств вычислительной и измерительной техники, связи, оргтехники, техники сбора, обработки, анализа и интерпретации данных и т.п. В определенном смысле на стыке технической информатики с формальной находится программирование, а технической и прикладной информатики – информационные системы (рис. 3.1).

*Технологическая информатика* включает в себя информационные технологии решения различных теоретических и прикладных исследовательских и практических задач, технологии отдельных компонентов информатики в целом, причем технологии как фундаментальные и прикладные научные дисциплины и как реально действующие области деятельности, технологические процессы и обеспечивающие их средства.

*Прикладную информатику* составляют научные дисциплины и области практической деятельности, занимающиеся исследованием закономерностей функционирования, основами построения, разработкой и применением информационных систем различного назначения, рассмотрением информационных процессов в природе и обществе, их сущности, структур и закономерностей, возможностей модификации и применения, а также появления и развития новых интегрированных научных фундаментальных и прикладных дисциплин, таких как биокибернетика, бионика, биогеоценология и др.

Рассматривая информатику с точки зрения методов ее исследования, в ней можно выделить следующие методы [24, 25]: *естественнонаучные, аксиоматические (математические), инженерные, эмпирические (экспериментальные), информационные, системные, изобретательские, моделирующие, эвристические* и др.

По основным группам операций информационных технологических процессов можно выделить *сбор, обработку, анализ, интерпретацию и применение информации*.

С точки зрения прикладных задач и обеспечивающих их средств, технологий в информатике можно выделить следующие задачи: гносеологические (познавательные, исследовательские), логостические (обучающие), управленческие, производственные, коммуникационные, сервисные и обеспечивающие их функционирование.

Эти задачи, в свою очередь, делятся на подзадачи, зачастую имеющие самостоятельное значение. Например, из управленческих можно выделить такие задачи, как планирование, регулирование, учет, прогнозирование, контроль, принятие решений, исполнение и т.п., из исследовательских – диагностика, генезис, прогнозирование, моделирование, познание, из производственных – проектирование, изготовление, испытание, сопровождение, маркетинг и т.д.

Наконец, среди *средств информатики*, обеспечивающих функционирование подсистем информационных систем (ИС), можно выделить следующие: аппаратные (технические), математические, программные, лингвистические, информационные, технологические, логостические, метрологические, организационно-экономические, правовые (юридические), эргономические.

Рассмотрим эти средства (подсистемы) на примере информационных систем.

*Средства аппаратного (технического) обеспечения* – это комплекс технических средств сбора, обработки, анализа, интерпретации и применения данных и информации, используемых для функционирования ИС.

*Средства математического обеспечения* – это совокупность средств модельного (различные модели, в том числе математические) и алгоритмического (методы, алгоритмы, правила и т.п.) обеспечения,

позволяющих выполнять формализацию и решение задач согласно назначению ИС.

*Средства программного обеспечения* – комплекс взаимосвязанных системных и прикладных программ регулярного применения, которые управляют работой технических средств и других подсистем в требуемых режимах, реализующих решение функциональных задач (обеспечивающих функционирование подсистем), а также осуществляют взаимодействие человека со всеми средствами ИС.

*Средства лингвистического обеспечения* – совокупность языковых средств и трансляторов, используемых для повышения эффективности функционирования ИС и обеспечивающих общение человека с другими подсистемами ИС.

*Средства информационного обеспечения* – совокупность всех унифицированных систем (баз, массивов) данных (переменных, постоянных и полупостоянных), документов, единой системы классификации, кодирования, методологии построения и управления ими, схем информационных потоков, необходимых для выполнения ИС своих функций.

*Средства технологического обеспечения* – совокупность методических и руководящих материалов, охватывающих общесистемные методики, руководящие материалы, нормативно-справочные документы, стандарты, инструкции, приемы, методы и вспомогательные средства для выполнения всех операций технологического процесса постановки и решения задач пользователя с помощью ИС, а также устанавливающих состав, правила отбора и эксплуатации ИС.

*Средства логистического обеспечения* – совокупность знаний, примеров, задач, методик и т.п., обеспечивающих обучение пользователя умению работать и приобретение им первоначальных навыков работы с ИС.

*Средства метрологического обеспечения* – совокупность стандартов, методик и средств тестирования, определения работоспособности ИС, точности и достоверности получаемых с ее помощью результатов.

*Средства организационно-экономического обеспечения* – совокупность: а) методов и средств, регламентирующих взаимодействие пользователя и работников с подсистемами ИС и между собой в процессе разработки и эксплуатации ИС; б) экономических параметров, методов организации исследования производства, технологического процесса и труда с использованием ИС, схем взаимодействия задач функционирования системы объект – ИС – субъекты – пользователи, определяющих эффективность, достижение объектом и/или субъектом заданных целей при применении ИС; в) документов, устанавливающих порядок, форму выдачи, требования к документальному оформлению выходных результатов ИС.

*Средства правового обеспечения* – совокупность правовых норм и нормативных документов, определяющих юридический статус и функционирование ИС и юридически регламентирующих правила, порядок получения, преобразования и использования результатов, вырабатываемых ИС и в ИС.

*Средства эргономического обеспечения* – совокупность методов и средств, создающих оптимальные условия для эффективного взаимодействия пользователя с различными подсистемами ИС [24, 25].

### **3.2. Этапы научного исследования в информатике**

В зависимости от используемых методов последовательность этапов исследования может быть различной. Так, при применении информационного подхода схема исследования приведена на рис. 3.2, моделирования – рис. 3.3, 3.4, аксиоматических методах исследования – рис. 3.5, теории решения изобретательских задач (ТРИЗ) – рис. 3.6, системных методах исследования – рис.3.7, естественнонаучных методах - рис.3.8., инженерных методах – рис. 3.9 [24, 25].

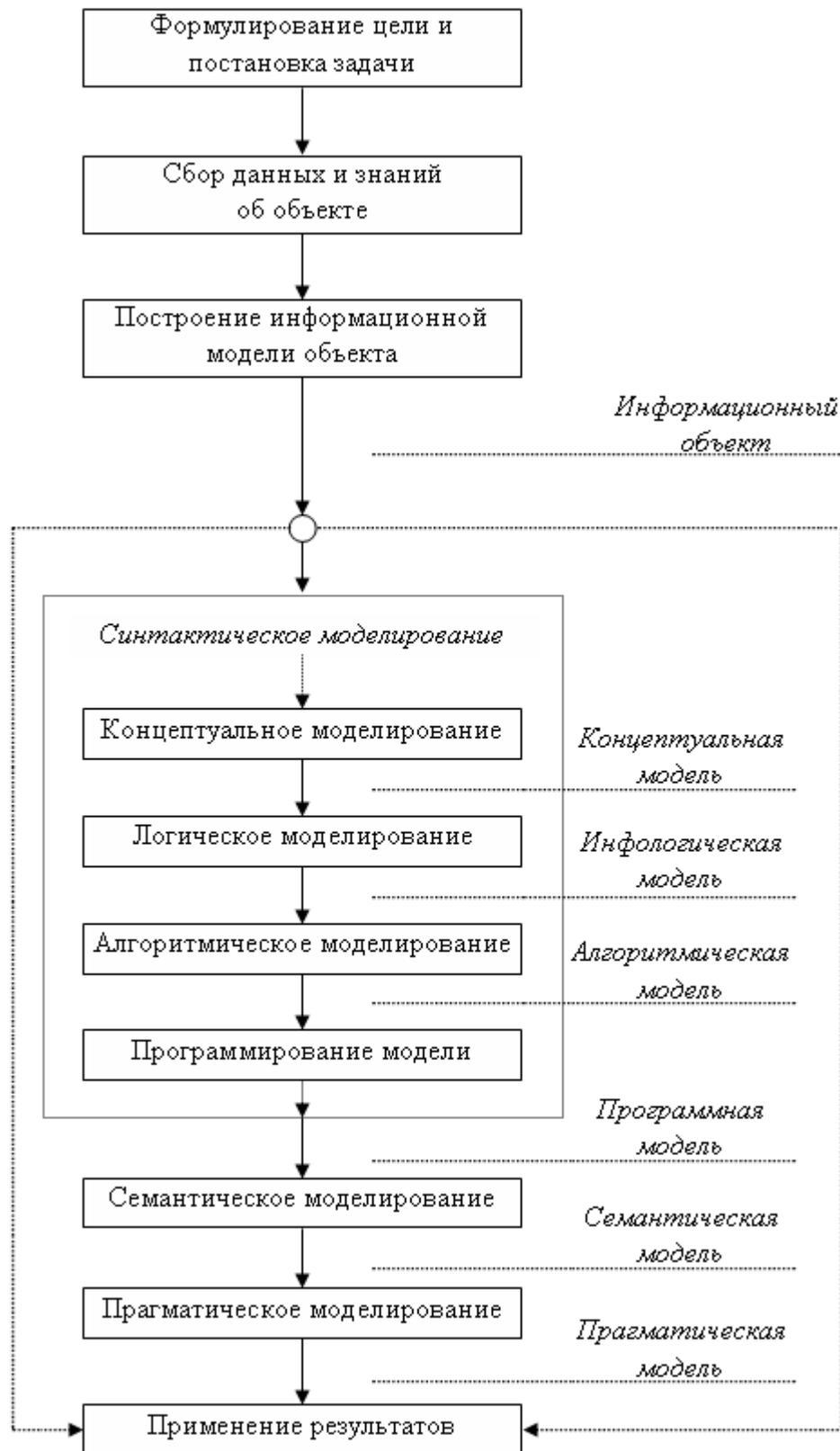


Рис.3.2. Схема исследования в информационном подходе



Рис.3.3. Укрупненная схема исследования моделирования физического объекта (стадия идентификации)

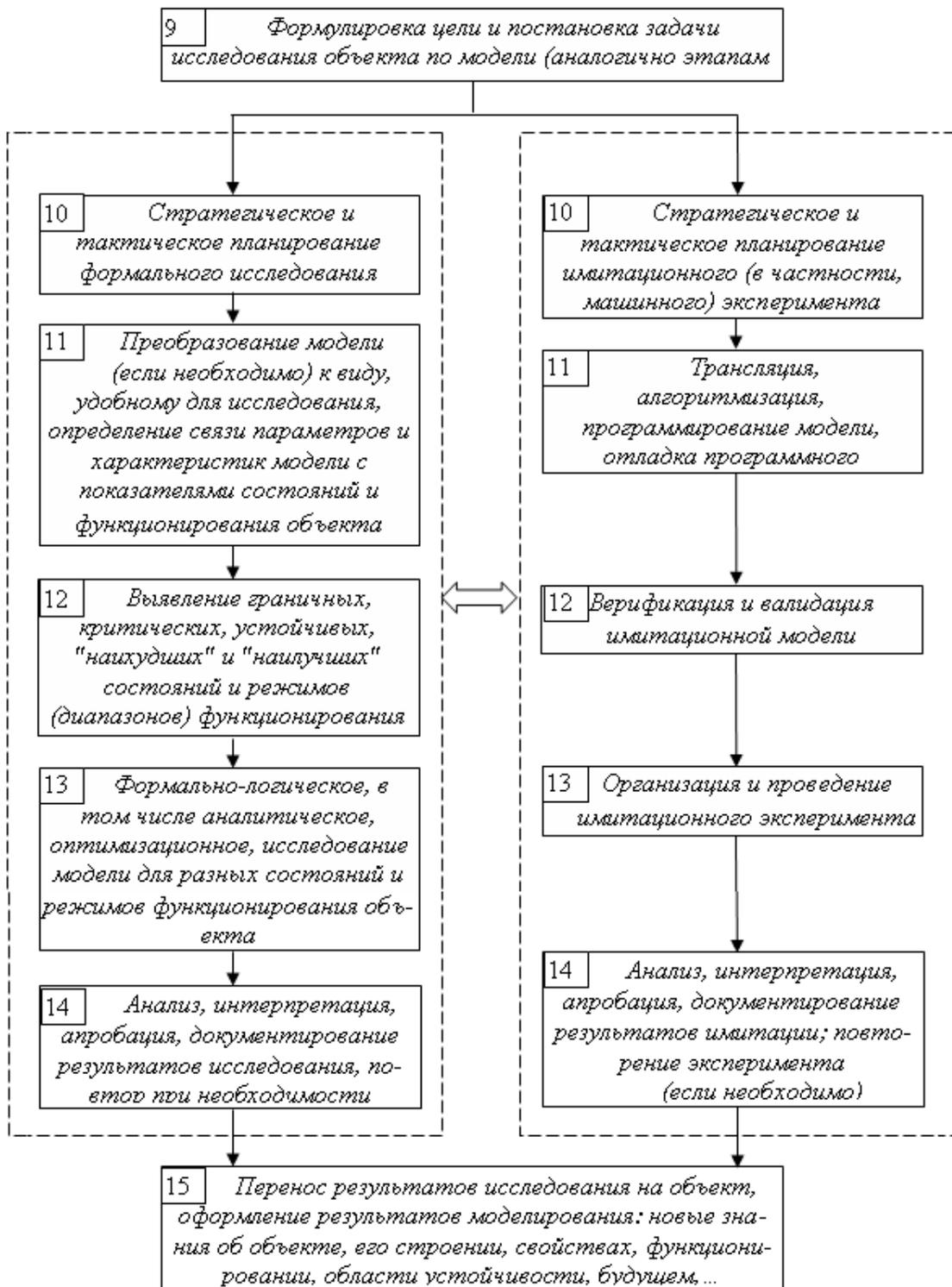


Рис.3.4. Окончание (стадия исследования объекта по модели: работа с моделью и перенос полученных по ней результатов на объект)

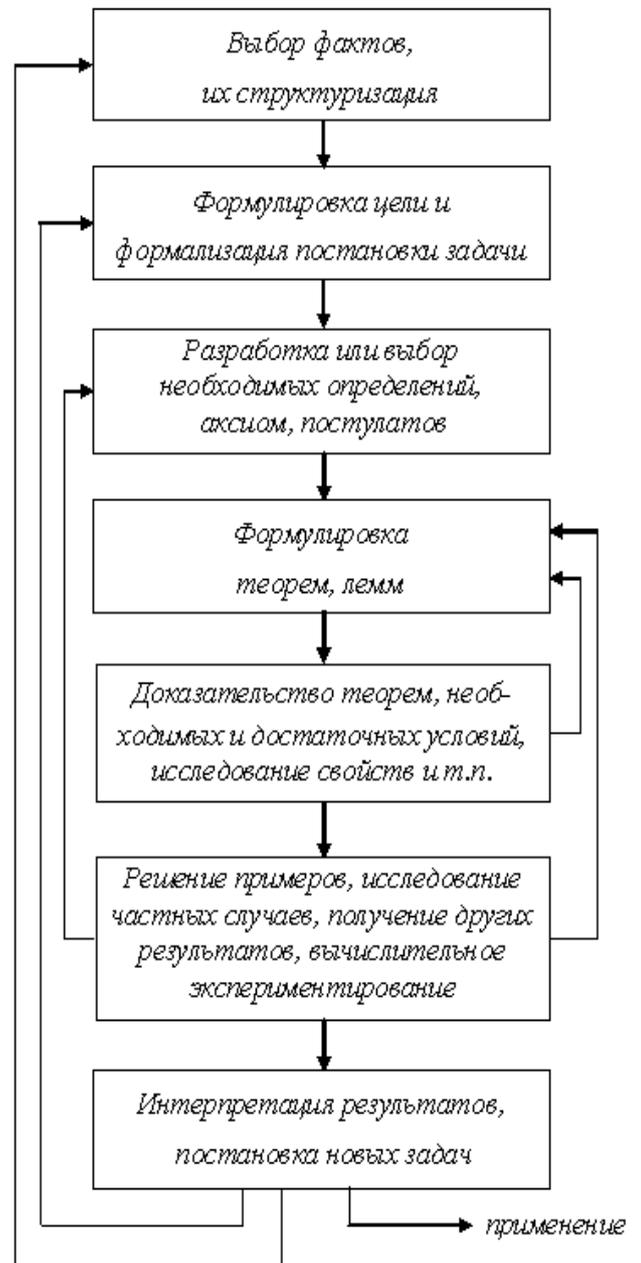
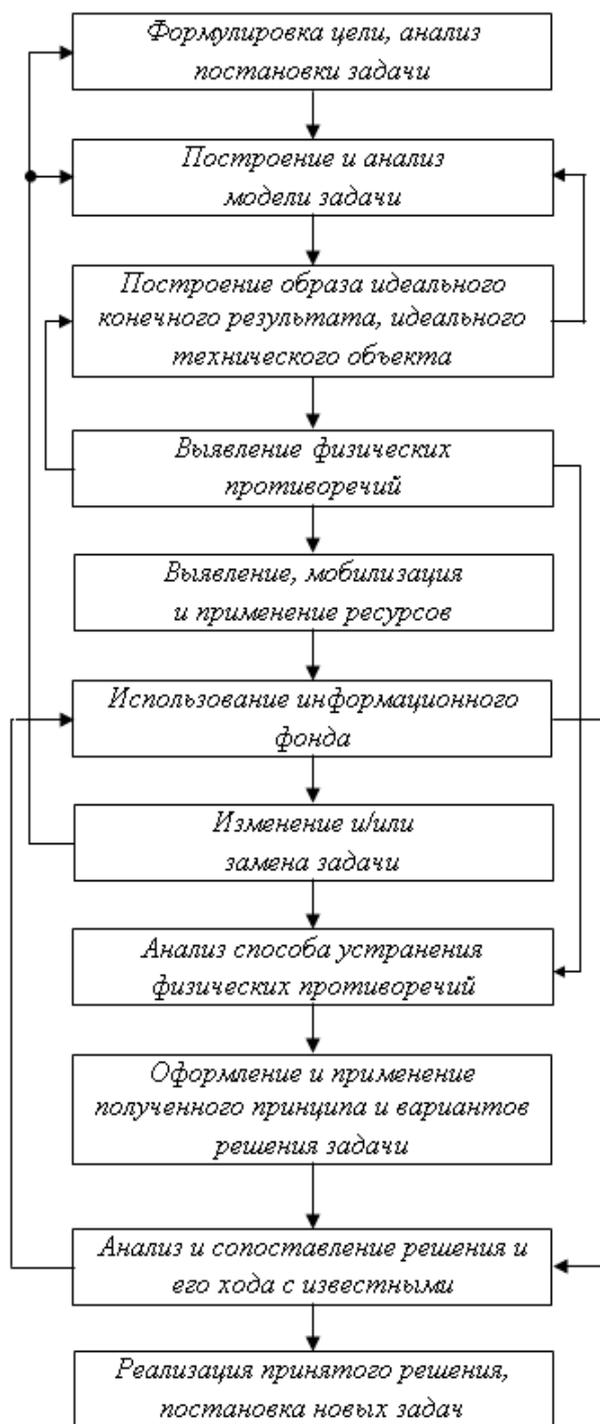


Рис.3.6. Схема исследования, реализующего теорию решения изобретательских задач (ТРИЗ)

Рис.3.5. Схема исследования в аксиоматических методах исследования

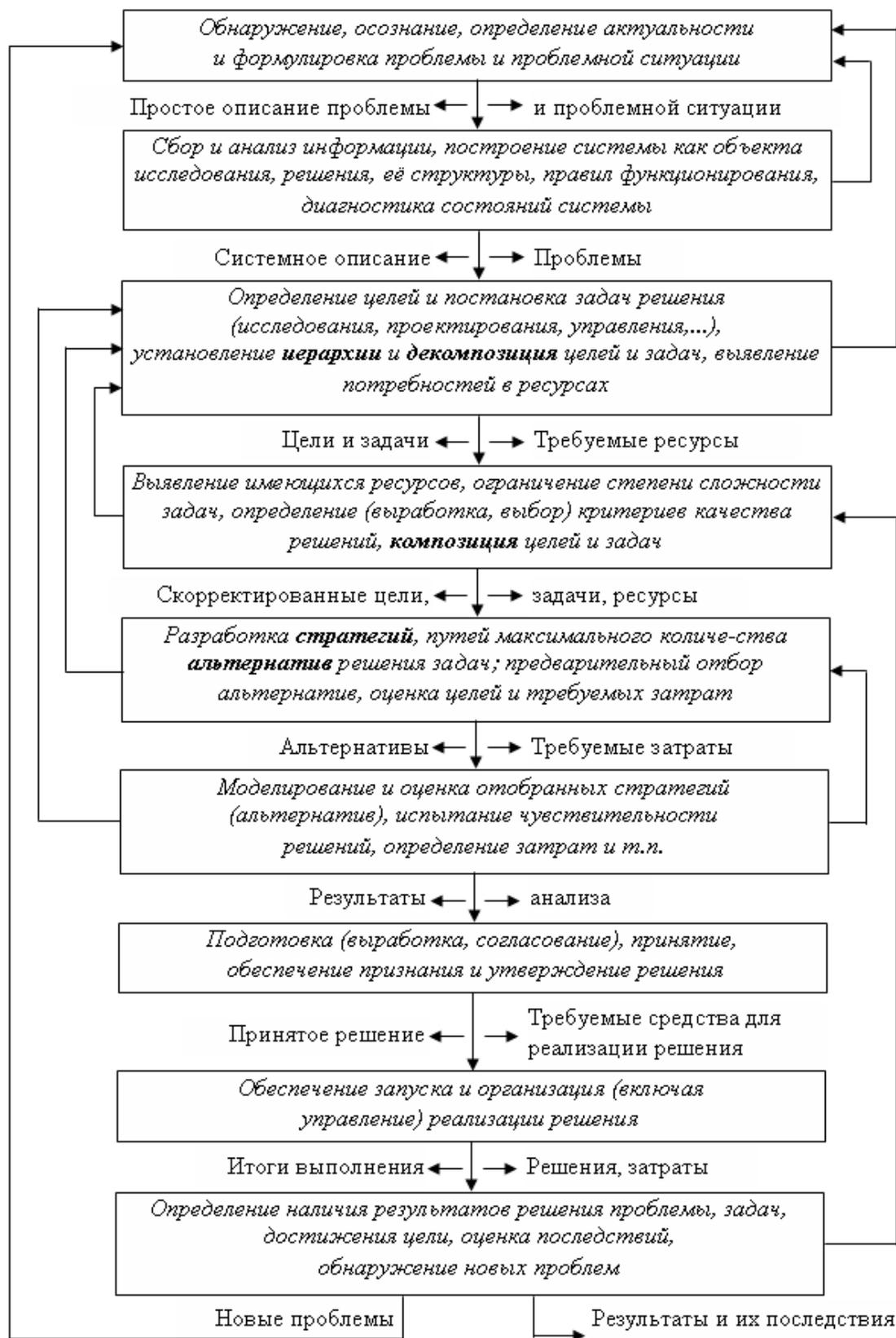


Рис.3.7. Схема исследования в системных методах исследования

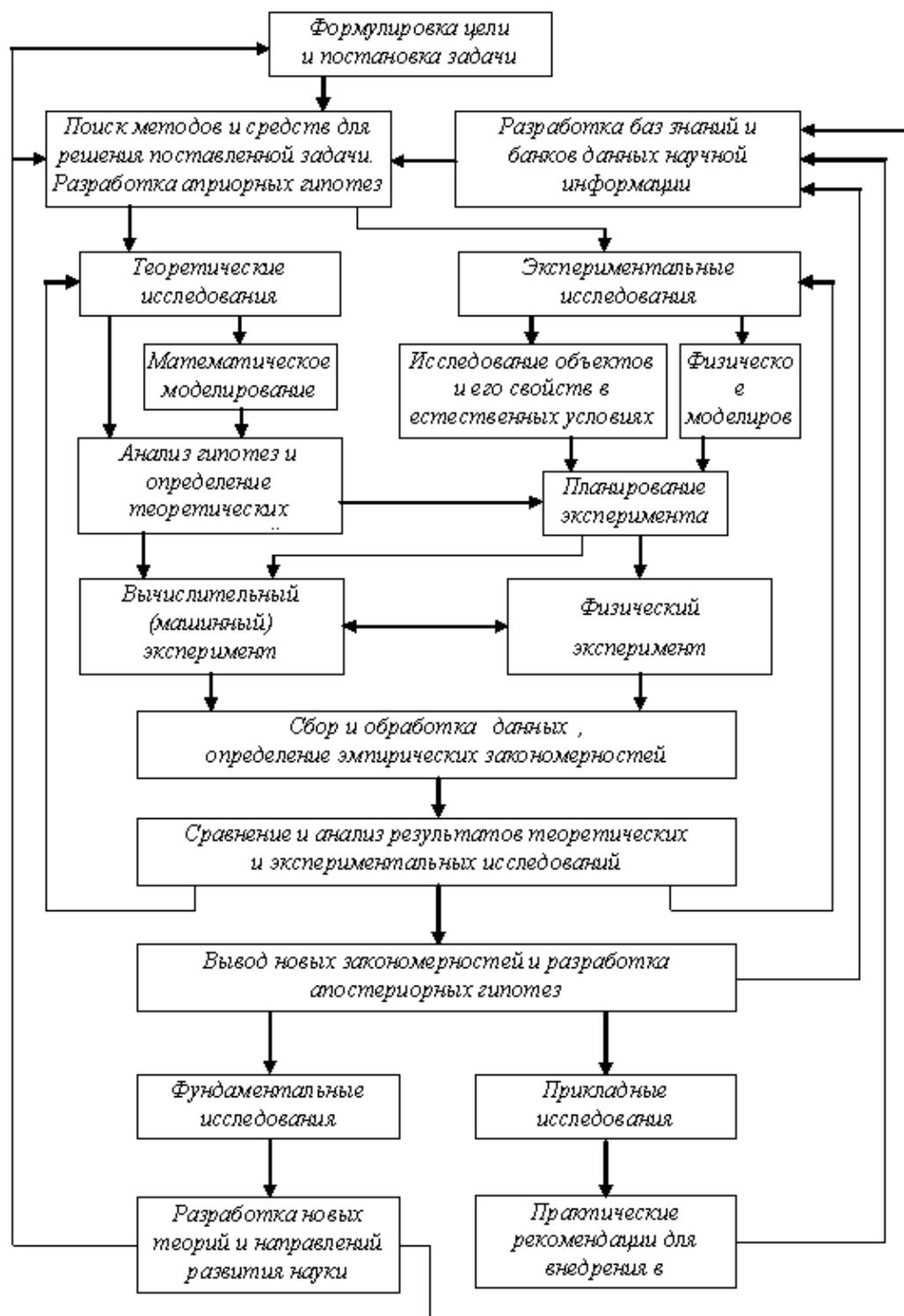


Рис.3.8. Схема исследования в естественнонаучных методах исследования



Рис.3.9. Схема исследования в инженерных методах

### 3.3. Формулировка цели и постановка задачи

Как видно из представленных выше схем процессов исследования, и как считают большинство методологов выявление проблемной ситуации и постановка проблемы является началом исследовательского поиска.

#### 3.3.1. Выявление проблемной ситуации и постановка проблемы

*Проблемная ситуация* – это объективное состояние рассогласования и противоречивости научного знания, возникающее в результате его неполноты и ограниченности. В зависимости от того, какие элементы знания приходят к рассогласованию или конфронтации, вычленяются следующие типы проблемных ситуаций:

- расхождение теорий с некоторыми экспериментальными данными, по разным параметрам;

- конфронтации теорий, применяемых к одной предметной области

(=эквивалентных как в эмпирическом (одинаково хорошо согласуются с эмпирическими данными), так и семантическом плане (несут одно и то же содержание), но различающиеся лингвистически (средствами описания),

= либо эквивалентных в эмпирическом плане, но не эквивалентных в семантическом отношении, когда на основании одних и тех же экспериментальных данных строятся различные гипотезы, вычленяется разное теоретическое содержание,

= неэквивалентных в эмпирическом плане и обладающих различной семантикой);

- столкновение парадигм, исследовательских программ, стилей научного мышления

(= несовпадение онтологических схем (картин мира), лежащих в основе конкурирующих теорий,

= противоречие между теорией и методологическими установками научного сообщества,

= противоречие между теорией и тем или иным мировоззрением, считающееся для теории более серьезным испытанием, чем эмпирические аномалии).

Проблемная ситуация как объективное состояние научного знания фиксируется в системе высказываний – тем самым формулируется *проблема*, в которой рассогласованность, противоречие и неполнота, неявно содержащиеся в ситуации, принимают явную и определенную форму. Проблема рассматривается как «знание о незнании».

*Постановка и выбор научных проблем* в логическом плане определяются такими предпосылочными структурами как парадигма, исследовательская программа и научная картина мира, которые могут стимулировать решение одних или запрещать, как не имеющую смысла, постановку других проблем. Кроме того, постановка новой проблемы существенно зависит от такого объективного фактора как степень зрелости или развитости предмета исследования. [40]

Любая проблема принадлежит некоторому научному направлению. Под *научным направлением* понимается наука или комплекс наук, в области которых ведутся исследования. Различают несколько структурных единиц научного направления: *комплексные проблемы; проблемы, темы и научные вопросы*.

*Комплексная проблема* представляет собой совокупность проблем, объединенных единой целью. Понятие проблемы дано выше. Можно также сказать, что *проблема* - это совокупность сложных теоретических и практических задач.

*Тема научного исследования* является составной частью проблемы. В результате исследований по теме получают ответы на определенный круг научных вопросов, охватывающих часть проблемы. Обобщение результатов ответов по комплексу тем может дать решение научной проблемы.

Под *научными вопросами* обычно понимаются мелкие научные задачи, относящиеся к конкретной теме научного исследования.

Комплексными проблемами и проблемами занимаются научные коллективы. Научные темы и вопросы могут исследовать отдельные научные работники, специалисты, аспиранты, студенты [33].

Выбрать тему научного исследования могут помочь:

- беседы с научным руководителем, со специалистами-практиками;
- предыдущий «задел» в научном исследовании;
- стаж в избранной области знания;
- просмотр каталогов защищенных диссертаций и ознакомление с уже выполненными в научном коллективе диссертационными работами;
- ознакомление с новейшими результатами исследований в смежных, пограничных областях науки и техники;
- оценка состояния разработки методов исследования, технологических приемов применительно к конкретной области экономики;
- пересмотр известных научных решений при помощи новых методов, с новых теоретических позиций, с привлечением новых, существенных фактов;
- ознакомление с аналитическими обзорами и статьями в специальной периодике [34,17].

Окончательная формулировка темы может быть впоследствии уточнена и откорректирована.

### **3.3.2. Выбор объекта и предмета исследования**

Обязательным элементом научного исследования является формулировка *объекта и предмета исследования*.

*Объект* в гносеологии, теории познания – это то, что входит в сферу деятельности и в классической интерпретации даже противостоит познающему его субъекту, это сторона, аспект, та точка зрения, проекция, порождающая проблемную ситуацию, по которой исследователь познает целостный объект, выделяя наиболее существенные с его точки зрения

признаки объекта. Один и тот же реальный объект может быть объектом разных исследований или научных направлений.

*Предмет исследования* идеален, он формулируется уже не в эмпирическом, а теоретическом познании, в методологических конструкциях и концепциях автора, он связан с поиском закономерностей, зависимостей и корреляций, которые можно объяснить и предсказать на уровне теории.

Некоторые полагают, что объект и предмет исследования как категории научного процесса соотносятся между собой как общее и частное, а в объекте якобы выделяется та его часть, которая служит предметом исследования. Именно на него и направлено основное внимание исследователя.

Однако более глубокое рассмотрение вопроса показывает, что объект располагается скорее в сфере эмпирического знания, тогда как предмет исследования может быть определен достаточно адекватно лишь в сфере теории и метода. Объект – это не просто какие-то стороны реального объекта. Сам объект науки конструируется в процессе научного познания и представляет стороны, особенности, характеристики и параметры процесса или явления, порождающие проблемную ситуацию и избранные для научного анализа.

В объекте исследования должны быть в достаточной степени представлены те исследуемые особенности или закономерности, которые определены как предмет исследования. При определении объекта научного исследования принципиально решается вопрос о выборке, инструментарии, о том, каким путем они могут быть вовлечены исследователем в научный оборот [59].

Предмет – это не то, что якобы находится в границах самого объекта, а лишь та его проекция, которая обозначает границы теории объекта. Предмет исследования безусловно конструируется с целью установления закономерностей изменения объекта. Предмет исследования определяется в сфере теории и методологии как новое научное знание, метод и инструмент получения этого научного знания об объекте [59].

### 3.3.3. Актуальность, научная новизна, практическая значимость

Выбранную научную тему нужно проанализировать и зафиксировать в четких формулировках, определить актуальность темы, научную новизну положений, их практическую и теоретическую значимость.

*Актуальность* предполагает соответствие научной работы состоянию науки на сегодняшний день, ее реальным потребностям и пригодность для решения ее насущных проблем. Обоснование актуальности темы должно соответствовать следующим конкретным требованиям: во-первых, исследователь должен кратко осветить причины обращения именно к этой теме; во-вторых, он должен раскрыть актуальность обращения к этой теме применительно к внутренним потребностям науки - объяснить, почему эта тема назрела именно сейчас, что препятствовало адекватному раскрытию ее раньше, показать, как обращение к ней обусловлено собственной динамикой развития науки, накоплением новой информации по данной проблеме, недостаточностью ее разработанности в имеющихся исследованиях, необходимостью изучения проблемы в новых ракурсах, с применением новых методов и методик исследования и т.д.

Приведем некоторые приемы по обоснованию актуализации темы диссертации, предлагаемые в различных пособиях:

- задачи фундаментальных исследований требуют разработки данной темы для теоретического объяснения новых фактов,
- уточнение, развитие и решение проблемы диссертации возможны и остро необходимы в современной российской ситуации;
- теоретические положения диссертации позволят снять разногласия в понимании процесса или явления,
- гипотезы и закономерности, выдвинутые в работе, позволяют обобщить известные ранее и полученные соискателем эмпирические данные [59].

*Научная новизна* - это признак, наличие которого дает исследователю право на использование понятия «впервые» при характеристике полученных им результатов и проведенного исследования в целом. Понятие «впервые» в науке означает факт отсутствия подобных результатов до их публикации.

Различают методологическую, теоретическую, технологическую и практическую новизну.

Можно выделить следующие элементы *методологической новизны*, которые могут быть представлены в исследовании:

- усовершенствование имеющихся или разработка новых методик;
- новый метод решения;
- новое применение известного решения или метода;
- новые результаты теории эксперимента, их следствия.

К *новым теоретическим* результатам можно отнести:

- новый объект исследования, то есть задача поставлена и рассматривается впервые;
- новую постановку известных проблем или задач;
- новые или усовершенствованные критерии, показатели;
- разработку оригинальных математических моделей процессов и явлений;
- новый метод решения.

*Технологическая новизна* результатов исследования проявляется, например, в разработке:

- авторского инструментария;
- устройств и способов на уровне изобретений и полезных моделей.

К *новым практическим* результатам относится то, что создается на основе положений исследования для возможного использования в практике, например:

- разработка проектов и их внедрение;
- достижения в результате внедрения инновационных разработок;

- обобщение на основе положений исследований с разработками прикладного характера (элемент новизны состоит в создании новых теоретических положений в приложении их к практике);
- рекомендации, позволяющие повысить эффективность использования разработки, производительность процесса, снизить себестоимость и т.п.

*Теоретическая значимость* исследования показывает роль, значение и место достигнутых в ходе научного исследования результатов в какой-либо научной теории, методологии.

*Практическая значимость* исследования во многом определяется характером выполняемого исследования. Исследования теоретического и методологического характера обладают более опосредованной практической значимостью, чем работы методического или прикладного характера.

Если исследование будет носить *теоретический или методологический характер*, то его практическая значимость может проявиться в публикациях основных результатов исследования в научных статьях, монографиях, учебниках, в наличии авторских свидетельств, актов о внедрении результатов исследования в практику; апробации результатов исследования на научно-практических конференциях и симпозиумах; в использовании научных разработок в учебном процессе высших и средних специальных учебных заведений, в участии в разработке федеральных и региональных программ; использовании результатов исследования при подготовке новых нормативных и методических документов.

Если исследование будет носить *методический характер*, то его практическая значимость может проявиться в наличии научно обоснованных и апробированных в ходе экспериментальной работы системы методов и средств совершенствования технического развития страны. Сюда же относят исследования по научному обоснованию новых и развитию действующих систем, методов и средств того или иного вида деятельности.

Если предполагается, что будущее исследование будет обеспечивать научное обоснование путей оптимизации трудовых и материальных ресурсов или производственных процессов, то есть носить сугубо *прикладной характер*, то его практическая значимость может проявляться в следующих формах:

- научного обоснования вариантов направлений, способов совершенствования условий и повышения эффективности практической деятельности;
- экономического обоснования мероприятий по использованию научно-технических достижений в различных областях науки и практики;
- обоснование предложений по использованию достижений научных разработок в практической деятельности предприятий и организаций;
- решение отдельных научных вопросов при разработке научно-исследовательских тем, выполняемых госбюджетных и хоздоговорных научных работ;
- использование результатов исследования в разработках научно-исследовательских, проектных институтах и других организаций [10, 34].

#### **3.3.4. Технико-экономическое обоснование**

При выполнении научно-исследовательских (НИР) и опытно конструкторских работ (ОКР) может потребоваться разработка основного исходного предпланового документа - технико-экономическое обоснование (ТЭО) темы. Только при наличии такого обоснования возможно дальнейшее планирование и финансирование темы заказчиком.

В первом разделе ТЭО темы указываются причины разработки (ее обоснование), приводится краткий литературный обзор, в котором описываются уже достигнутый уровень исследований и ранее полученные результаты. Особое внимание уделяется еще не решенным вопросам, обоснованию актуальности и значимости работы. Такой обзор позволяет

наметить методы решения, задачи и этапы исследования, определить конечную цель выполнения темы. Сюда входят патентная проработка темы и определение целесообразности закупки лицензий.

На стадии составления ТЭО устанавливается область использования ожидаемых результатов НИР, возможность их практической реализации в данной отрасли, определяется предполагаемый (потенциальный) экономический эффект за период применения новой техники (зависящей от продолжительности разработки НИР и ОКР, этапов завершения и внедрения отдельных вопросов). Кроме экономического эффекта в ТЭО указываются предполагаемые социальные результаты (рост производительности труда, качества продукции, повышение уровня безопасности и производственной санитарии, обеспечение охраны природы и окружающей среды). В результате составления ТЭО делается вывод о целесообразности и необходимости выполнения НИР и ОКР. Технико-экономическое обоснование утверждается заказчиком [33].

### **3.3.5. Формулирование целей, задач исследования**

На основе сконструированных исследователем объекта и предмета исследования, сформулированной проблемы устанавливается центральный момент исследовательской работы – *цель исследования*, то есть что в самом общем виде нужно достигнуть в итоге всей работы, желаемый конечный результат конкретной исследовательской деятельности.

Цель исследовательской работы состоит в раскрытии ее темы, получении новых положительных результатов для человека и общества на основе новых знаний. Выбор цели осуществляют в рамках научной проблемы. Характеризуя цель исследований, следует отмечать, в какие этапы исследования предмета предполагается сделать свой основной вклад: в постановку проблемы, разработку или развитие теоретической концепции, методические решения проблемы, совершенствование инструментария и средств исследования.

Сформулированная цель и гипотеза исследования логически определяют его задачи. *Задачи исследования* – это частные, сравнительно самостоятельные цели исследования в конкретных условиях проверки сформулированной гипотезы применительно к выбранному объекту и предмету исследования. Количество, содержание и глубина задач выбирается индивидуально, но среди них можно выделить статистические задачи оценки существующих и перспективных требований к объекту исследования, методологические и теоретические задачи раскрытия предмета исследования, экспериментальные задачи подтверждения правильности теории, практические задачи выявления эффективности решений. Задачи исследования формулируются как относительно самостоятельные, законченные части. Это обычно делается в форме перечисления: изучить..., описать..., установить..., выявить..., вывести формулу...

### **3.4. Выбор методологии и методов исследования**

#### **3.4.1. Поиск, накопление и обработка научной информации**

##### **3.4.1.1. Научные документы и издания**

Под *источником информации* понимается документ, содержащий какие-либо сведения. К документам относят различного рода издания, являющиеся основным источником научной информации. *Издание* – это документ, предназначенный для распространения содержащейся в нем информации, прошедший редакционно-издательскую обработку, полученный печатанием или тиснением, полиграфически самостоятельно оформленный, имеющий выходные сведения.

Издания классифицируют по различным основаниям:

- по целевому назначению (официальное, научное, учебное, справочное и др.);
- степени аналитико-синтетической переработки информации (информационное, библиографическое, реферативное, обзорное);

- материальной конструкции (книжное, журнальное, листовое, газетное и т.д.);
- знаковой природе информации (текстовое, нотное, картографическое, изоиздание);
- объему (книга, брошюра, листовка);
- периодичности (непериодическое, сериальное, периодическое, продолжающееся); составу основного текста (моноиздание, сборник);
- структуре (серия, однотомное, многотомное, собрание сочинений, избранные сочинения).

Кроме того, документы подразделяются на *первичные* (содержащие непосредственные результаты научных исследований и разработок, новые научные сведения или новое осмысление известных идей и фактов) и *вторичные* (содержащие результаты аналитико-синтетической и логической переработки одного или нескольких первичных документов или сведения о них).

Как первичные, так и вторичные документы подразделяются на опубликованные (издания) и непубликуемые.

В числе опубликованных первичных документов - *книги* (непериодические текстовые издания объемом свыше 48 страниц); *брошюры* (непериодические текстовые издания объемом свыше четырех, но не более 48 страниц). Среди книг и брошюр важное научное значение имеют *монографии*, содержащие всестороннее исследование одной проблемы или темы и принадлежащие одному или нескольким авторам, и затем *сборники научных трудов*, содержащие ряд произведений одного или нескольких авторов, рефераты и различные официальные или научные материалы.

Для учебных целей издаются *учебники и учебные пособия* (учебные издания). Это непериодические издания, содержащие систематизированные сведения научного и прикладного характера, изложенные в форме, удобной для преподавания и изучения.

Некоторые издания, публикуемые от имени государственных или общественных организаций, учреждений и ведомств, называются *официальными*. Они содержат материалы законодательного, нормативного или директивного характера.

Наиболее оперативным источником НТИ являются периодические издания, выходящие через определенные промежутки времени, постоянным для каждого года числом номеров. Традиционными видами периодических изданий являются газеты и журналы. К периодическим относятся также продолжающиеся издания, выходящие через неопределенные промежутки времени, по мере накопления материала. Обычно это сборники научных трудов институтов, вузов, научных сообществ, публикуемых без строгой периодичности под общим заглавием "Труды", "Ученые записки", "Известия" и др. и

К специальным видам технических изданий принято относить нормативно-техническую документацию, регламентирующую научно-технический уровень и качество выпускаемой продукции (стандарты, инструкции, типовые положения, методические указания и др.).

Стандарт - нормативно-технический документ, устанавливающий комплекс норм, правил, требований к объекту стандартизации и утвержденный компетентным органом.

К основным видам непубликуемых первичных документов относятся научно-технические отчеты, диссертации, депонированные рукописи научные переводы, конструкторская документация, информационные сообщения о проведенных научно-технических конференциях, съездах, симпозиумах, семинарах.

*Автореферат диссертации* – научное издание в виде брошюры, содержащее составленный автором реферат проведенного им исследования, представляемого на соискание ученой степени.

*Препринт* – научное издание, содержащее материалы предварительного характера, опубликованные до выхода в свет издания, в котором они могут быть помещены.

*Материалы научной конференции* – научный неперіодический сборник, содержащий итоги научной конференции (программы, доклады, рекомендации, решения).

*Тезисы докладов (сообщений) научной конференции* – научный неперіодический сборник, содержащий опубликованные до начала конференции материалы предварительного характера (аннотации, рефераты докладов и (или) сообщений).

Вторичные документы и издания подразделяют на справочные, обзорные, реферативные и библиографические.

В *справочных изданиях* (справочники, словари) содержатся результаты теоретических обобщений, различные величины и их значения, материалы производственного характера.

В *обзорных изданиях* содержится концентрированная информация, полученная в результате отбора, систематизации и логического обобщения сведений из большого количества первоисточников по определенной теме за определенный промежуток времени. Различают обзоры аналитические (содержащие аргументированную оценку информации, рекомендации по ее использованию) и реферативные (носящие более описательный характер). Кроме того, работники библиотек часто готовят библиографические обзоры, содержащие характеристики первичных документов как источников информации, появившихся за определенное время или объединенных каким-либо общим признаком.

*Реферативные издания* (реферативные журналы, реферативные сборники) содержат сокращенное изложение первичного документа или его части с основными фактическими сведениями и выводами. Реферативный журнал - это периодическое издание журнальной или карточной формы, содержащее рефераты опубликованных документов (или их частей).

Реферативный сборник - это периодическое, продолжающееся или непериодическое издание, содержащее рефераты непубликуемых документов (в них допускается включать рефераты опубликованных зарубежных материалов).

*Библиографические указатели* являются изданиями книжного или журнального типа, содержащими библиографические описания вышедших изданий. В зависимости от принципа расположения библиографических описаний указатели подразделяются на систематические (описания располагаются по областям науки и техники в соответствии с той или иной системой классификации) и предметные (описания располагаются в порядке перечисления важнейших предметов в соответствии с предметными рубриками, расположенными в алфавитном порядке).

Вторичные непубликуемые документы включают регистрационные и информационные карты, учетные карточки диссертации, указатели депонированных рукописей и переводов, картотеки "Конструкторская документация на нестандартное оборудование", информационные сообщения. К ним принято относить также вторичные документы, которые публикуются, но рассылаются по подписке (Бюллетени регистрации НИР и ОКР, сборники рефератов НИР и ОКР и др.). [50, 33]

В настоящее время в России выпуском информационных изданий занимаются институты, центры и службы научно-технической информации (НТИ). Они объединяются в Государственную систему научно-технической информации (ГСНТИ), осуществляющую централизованный сбор и обработку основных видов.

*ВИНИТИ* (Всероссийский институт научной и технической информации) выполняет функции головного научно-исследовательского информационного органа страны и обрабатывает поток опубликованной научно-технической литературы в области точных, естественных и технических наук, подготавливает и издает библиографические указатели,

реферативные журналы, сигнальную информацию, экспресс-информацию, обзорную информацию.

*ВНИИЦ* (Всероссийский научно-технический информационный центр), осуществляет сбор, накопление и обработку информации по всем видам публикуемых исследовательских работ (отчеты о НИР и ОКР, защищенные диссертации), проводимых в стране, и издает по ним информационные издания реферативного и сигнального типов.

*ВНИИКИ* (Всероссийский научно-исследовательский институт технической информации, классификации и кодирования), издает информационные указатели литературы.

*ВНИИПИ* (Всероссийский научно-исследовательский институт патентной информации), выпускает оригинальные и собственные информационные издания по различным направлениям изобретательства, в том числе сигнальные, библиографические и реферативные издания.

*ГПНТБ* (Государственная публичная научно-техническая библиотека) выполняет функции общесоюзного научно-методического центра для научных, специальных и технических библиотек и Всероссийского центра координации библиографической работы в области техники.

#### **3.4.1.2. Работа с источниками научной информации**

К традиционным источникам научной информации относятся:

- диссертации и авторефераты диссертаций;
- периодика: научные и реферативные журналы, академические вестники, сборники научных статей, материалы научных конференций;
- книги: научные монографии, брошюры, учебники, научно-методические издания, учебные пособия;
- документы: сборники нормативных документов, стандарты, нормативные акты, инструкции;
- энциклопедии, словари и справочники;
- отчеты о научно-исследовательской работе;

- патенты и авторские свидетельства;
- информационные издания: аналитические обзоры, информационные выпуски и бюллетени, выставочные проспекты;
- переводы научной литературы;
- оригинальные источники зарубежной научной литературы.

Для информационного поиска рекомендуется также обращаться к автоматизированным информационно-поисковым системам, базам и банкам данных, электронным библиотечным каталогам. Среди научных журналов наибольший рейтинг имеют журналы, рекомендованные ВАК (Высшей Аттестационной Комиссии).

Целью работы с источниками научной информации является анализ степени разработанности проблемы, состояния теории и практики по рассматриваемому вопросу. При этом проводится конструктивная критика положений других авторов, с которыми не согласен исследователь. Указываются причины, вследствие которых ранее полученные результаты не удовлетворяют потребностям практики. Авторы группируются по аспектам и концепциям, в русле которых они проводят исследование данной проблематики. Освещение состояния вопроса исследований заканчивается краткими выводами о состоянии и степени разработанности проблемы. Перечисляется круг проблемных вопросов и задач, которые недостаточно разработаны в научной литературе и которые необходимо исследовать в работе [59].

Работу с источниками рекомендуется разбить на несколько этапов:

- предварительный выбор по каталогам, реферативным обзорам, спискам литературы;
- поиск самих источников в соответствии со сформированным списком;
- анализ первоисточников, их отбор, ксерокопирование, конспектирование.

Способы изучения, анализа научной информации довольно индивидуальны для каждого исследователя. Можно сделать несколько замечаний.

При изучении литературы используется не вся информация, а только та, которая имеет непосредственное отношение к исследованию. Причем, поскольку научные исследования в информатике могут касаться любой области человеческой деятельности, то источниками научной информации в данном случае могут быть издания разнообразных научных направлений.

В докторских диссертациях можно найти описание методологий, методов и методик, применяемых в научных исследованиях подобных направлений.

Монографии дают полный, систематизированный обзор научной информации по конкретной научной теме, актуальный на момент выхода издания.

Для изучения нового понятия или метода целесообразнее использовать учебники, где в доступной форме излагаются эти вопросы.

Следует обязательно писать полные библиографические данные на сделанных ксерокопиях. Пример того, какие данные следует записывать, можно посмотреть в разделе оформления списка литературы данного пособия.

### **3.4.2. Выбор методологий, методов и методик исследования**

#### **3.4.2.1. Понятие метода и методологии научных исследований**

*Метод научного исследования* – это способ познания объективной действительности. Способ представляет собой определенную последовательность действий, приемов, операций. В зависимости от содержания изучаемых объектов различают методы естествознания и методы социально-гуманитарного исследования. Методы исследования классифицируют по отраслям науки: математические, биологические, медицинские, социально-экономические и т.д.

В зависимости от уровня познания выделяют методы эмпирического, теоретического и метатеоретического уровней. К методам *эмпирического* уровня относят наблюдение, описание, сравнение, счет, измерение, анкетный опрос, собеседование, тестирование, эксперимент, моделирование и т.д. К методам *теоретического* уровня причисляют аксиоматический, гипотетический (гипотетико-дедуктивный), формализацию, абстрагирование, общелогические методы (анализ, синтез, индукцию, дедукцию, аналогию) и др. Методами *метатеоретического* уровня являются диалектический, метод системного анализа, метафизический, герменевтический и др.

В зависимости от сферы применения и степени общности различают методы:

- 1) всеобщие (философские), действующие во всех науках и на всех этапах познания;
- 2) общенаучные, которые могут применяться в гуманитарных, естественных и технических науках;
- 3) частные – для родственных наук;
- 4) специальные – для конкретной науки, области научного познания.

От рассматриваемого понятия метода следует отличать понятия техники, процедуры и методики научного исследования.

Под *техникой исследования* понимают совокупность специальных приемов для использования того или иного метода, а под *процедурой исследования* – определенную последовательность действий, способ организации исследования.

*Методика* – это совокупность способов и приемов познания.

Любое научное исследование осуществляется определенными приемами и способами, по определенным правилам. Учение о системе этих приемов, способов и правил называют *методологией*. Понятие «методология» в литературе употребляется в двух значениях:

- 1) совокупность методов, применяемых в какой-либо сфере деятельности (науке, политике и т.д.);

2) учение о научном методе познания [44].

Существуют следующие уровни методологии:

1) всеобщая методология, которая является универсальной по отношению ко всем наукам и в содержание которой входят философские и общенаучные методы познания;

2) частная методология научных исследований для группы родственных наук, которую образуют философские, общенаучные и частные методы познания;

3) методология научных исследований конкретной науки, в содержание которой включаются философские, общенаучные, частные и специальные методы познания.

В ходе исследования применяются основные методы информатики или иных фундаментальных и прикладных наук.

К основным методам информатики, используемых в научных исследованиях, относят:

- всеобщие (диалектический и метод системного анализа);

- общенаучные (сравнение, экспериментирование, обобщение, абстрагирование, формализация, анализ, синтез, индукция, дедукция, традиция, аналогия, моделирование, идеализация, аксиоматические, гипотетические, исторический, системные);

анализ, синтез; аналогия, сравнение; индукция, дедукция, традиция; наблюдение, экспериментирование; формализация, абстрагирование; аксиоматические, гипотетические, эвристические, аналитические);

- специфичные (алгоритмизация, моделирование, программирование (планирование); имитационные; поиска, анализа, выбора вариантов, решений; обнаружения и исправления ошибок; распознавания образов; постановки и решения задач; естественнонаучные, эмпирические, аксиоматические (математические); инженерные, изобретательские, моделирующие, информационные и т.п.) [25].

Следует отметить, что методы системного анализа превалирует в научных исследованиях по информатике. Далее приводятся некоторые из них.

### **3.4.2.2. Методологии структурного анализа**

Для описания и анализа системы разработано большое количество различных подходов и методов. В истории развития методологий моделирования систем можно выделить следующие основные этапы [13, 17].

- 40-60-е гг. Появление алгоритмических языков описания.
- 60-е гг. Появление методологии SADT (структурного анализа и проектирования).
- 70-80-е гг. Появление методологий серии IDEF, DFD, ERD.
- 90-е гг. Появление методологии ARIS (архитектура интегрированных информационных систем), UML (универсальный язык моделирования), методологий компаний Oracle, Baan, Rational
- 2000 г. Принятие МС ИСО серии 9000 версии 2000 г., содержащее четкое определение процессного подхода к управлению организацией.

Основу многих современных методологий моделирования систем составила методология SADT (Structured Analysis and Design Technique – метод структурного анализа и проектирования, предложенный Д. Россом) [37]. В основе этого подхода лежит графический язык описания и моделирования систем.

В середине 70-х ВВС США реализовало программу интегрированной компьютеризации производства ICAM (Integrated Computer Aided Manufacturing) [23]. В рамках этой программы была разработана методология IDEF (ICAM Definitions), позволяющая представить и исследовать структуру, параметры и характеристики производственно-технических и организационно-экономических систем.

В настоящее время к семейству IDEF можно отнести следующие стандарты:

IDEF0 - Function Modeling - методология функционального моделирования. С помощью наглядного графического языка IDEF0, изучаемая система предстает перед разработчиками и аналитиками в виде набора взаимосвязанных функций (функциональных блоков - в терминах IDEF0). Следует отметить, что модели в нотации IDEF0 предназначены для высокоуровневого описания функционирования организаций. Их основное преимущество состоит в возможности описывать управление процессами организации;

IDEF1 - Information Modeling – методология моделирования информационных потоков внутри системы, позволяющая отображать и анализировать их структуру и взаимосвязи;

IDEF1X (IDEF1 Extended) - Data Modeling – методология построения реляционных структур. IDEF1X относится к типу методологий "Сущность-взаимосвязь" (ER – Entity-Relationship) и, как правило, используется для моделирования реляционных баз данных, имеющих отношение к рассматриваемой системе;

IDEF2 - Simulation Modeling – методология динамического моделирования развития систем. В связи с весьма серьезными сложностями анализа динамических систем от этого стандарта практически отказались, и его развитие приостановилось на самом начальном этапе. Однако в настоящее время присутствуют алгоритмы и их компьютерные реализации, позволяющие превращать набор статических диаграмм IDEF0 в динамические модели, построенные на базе "раскрашенных сетей Петри" (CPN – Color Petri Nets);

IDEF3 - Process Description Capture – методология документирования процессов, происходящих в системе. С помощью IDEF3 описываются сценарий и последовательность операций для каждого процесса. IDEF3 имеет прямую взаимосвязь с методологией IDEF0 – каждая функция

(функциональный блок) может быть представлена в виде отдельного процесса средствами IDEF3. IDEF3 лежит в основе популярной в настоящее время методологии ARIS.

IDEF4 - Object-oriented Design – методология построения объектно-ориентированных систем. Средства IDEF4 позволяют наглядно отображать структуру объектов и заложенные принципы их взаимодействия, тем самым, позволяя анализировать и оптимизировать сложные объектно-ориентированные системы;

IDEF5 - Ontology Description Capture – методология онтологического исследования сложных систем. С помощью методологии IDEF5 онтология системы может быть описана при помощи определенного словаря терминов и правил, на основании которых могут быть сформированы достоверные утверждения о состоянии рассматриваемой системы в некоторый момент времени. На основе этих утверждений формируются выводы о дальнейшем развитии системы и производится её оптимизация;

IDEF6 - Design Rationale Capture;

IDEF7 - Information System Audit Method;

IDEF8 - User Interface Modeling;

IDEF9 - Scenario-driven Info Sys Design Spec – методология моделирование требований;

IDEF10 - Implementation Architecture Modeling;

IDEF11 - Information Artifact Modeling;

IDEF12 - Organization Modeling;

IDEF13 - Three Schema Mapping Design;

IDEF14 - Network Design.

В 1993 году методология IDEF0 была утверждена в качестве федерального стандарта США для функционального моделирования. В 2001 году Госстандарт России принял ГОСТ Р ИСО 9004-2001 «Системы менеджмента качества. Рекомендации по улучшению деятельности» на основе IDEF0. В настоящее время методология IDEF0 рассматривается ИСО

на предмет международного стандарта IPS (стандарты по обработке информации).

Нотации DFD (Data Flow Diagramming) предназначены для описания потоков данных. Они позволяют отразить последовательность работ, выполняемых по ходу процесса, и потоки информации, циркулирующие между этими работами. Кроме того, нотация DFD предоставляет возможность описывать потоки документов (документооборот) и материальных ресурсов (например, движение материалов от одной работы к другой). С помощью схемы процессов в DFD выявляют основные потоки данных, что важно для последующего создания моделей структуры данных и разработки требований к информационной системе организации.

Унифицированный язык моделирования (Unified Modeling Language, UML) является графическим языком для визуализации, специфицирования, конструирования и документирования систем на основе объектно-ориентированной технологии. С помощью UML можно разработать детальный план создаваемой системы, отображающий не только ее концептуальные элементы, такие как системные функции и процессы, но и конкретные особенности реализации, в том числе классы, написанные на специальных языках программирования, схемы баз данных и программные компоненты многократного использования [11,12].

Описанные выше методы анализа и описания касаются процессов [20], заранее определенных и по составу, и по последовательности действий, и по времени, и др. Для описания неопределенных процессов рекомендуется диаграмма процесса принятия решения (Process Decision Program Chart - PDPC, представляющая собой иерархическую структуру в виде дерева. Диаграмма процесса принятия решения позволяет спланировать возможные варианты событий и действий, которые должны произойти, чтобы достигнуть целей процесса.

Метод описания процесса как системы объектов и прерываний полезно использовать в тех случаях, когда известны действия (возможности

процесса), но не известно, какие и в какой последовательности необходимо совершить для получения запланированного результата.

Многие процессы состоят исключительно из действий в ответ на событие. Для таких процессов рекомендуется [20] объектно-событийное описание, при котором рассматриваются объекты, участвующие в процессе, и события, которые порождают действия, вносящие изменения в объекты. Весь процесс, таким образом, представляется как наблюдение (мониторинг) за событиями и изменение объектов в соответствии с этими событиями.

Для систем управления выявление и уточнение функций путем структуризации целей-функций наиболее часто проводится по разработанным ранее методикам [16]: ПАТТЕРН (PATTERN), С.А. Валуева, Б.Д. Кошарского, А.И. Уёмова, Р. Акоффа, Ф. Эмери, А.Н. Леонтьева, Ф.И. Перегудова, В.З. Ямпольского, Л.В. Кочнева, Ю.И. Черняка, Е.П. Голубкова и др. Обобщением этих методик является методика Е.Н. Волковой и А.А. Денисова.

Методика ПАТТЕРН была первой методикой системного анализа, в которой определялись порядок, методы, формирования и оценки приоритетов элементов структур целей (названных в методике "деревьями целей"). Инициатором создания методики считается Ч. Дэвис. Главное достоинство этой методики состоит в том, что в ней определены классы критериев оценки относительной важности, взаимной полезности, состояния и сроков разработки. Однако логика формирования структуры, принципы и приемы структуризации недостаточно в ней отработаны.

Первыми работами, в которых предложены не только принципы формирования "дерева целей", но и признаки структуризации были работы Ю.И. Черняка.

Для структуризации функций именно систем организационного управления одной из первых была предложена методика С.А. Валуева. В основу этой методики положены принципы анализа характеристик организационной системы, определение функций, раскрывающих

содержание процесса управления, и впервые было предложено учитывать этапы цикла принятия решения.

Методики Е.П. Голубкова уже ориентированы на структуризацию не только целей и функций, но и на анализ системы в целом.

Методика Б.Д. Кошарского и А.И. Уёмова, базирующаяся на двойственном определении системы, нашла широкое применение в различных отраслях при структуризации целей и функций предприятий в процессе разработки структуры функциональной части их автоматизированных систем управления (общеизвестными стали признаки "цикл управления" и "объект управления"). Недостатком методики является отсутствие средств для определения новых объектов, новых видов деятельности и функций, связанных с развитием предприятий.

Для этапов развития системы, пересмотра производственной и организационной структур, при проектировании новых предприятий показана методика, основанная на концепции системы, учитывающей среду и целеполагание (Ф.И.Перегудов, В.З. Ямпольский, Л.В. Кочнев).

В случаях постановки новых проблем, структуризации целей развития новых видов деятельности применяется методика, базирующаяся на концепции деятельности (А.Н. Леонтьева).

Методика Р. Акоффа и Ф. Эмери (структуризации целей системы, стремящейся к идеалу) позволяет обеспечить полноту выявления подцелей и функций для систем управления предприятий социальной сферы. В этой методике предусматривается принцип фрактальности, т.е. структуризации каждой ветви нижележащего уровня с использованием составляющих соседних уровней.

Обобщенная методика анализа целей и функций систем управления (Е.Н.Волковой, А.А. Денисова) включает несколько методик структуризации, методов оценки структур и предусматривает возможность выбора методики получения первоначального варианта структуры и методов его оценки, наиболее подходящих для соответствующего периода развития предприятия

и его системы организационного управления, с учетом особенностей создаваемого предприятия. Указанная методика используется как основа для формирования конкретных методик проектируемых или преобразуемых организаций.

В перечисленных методиках этап оценки структуры в явном виде не всегда присутствует. Объясняется это тем, что, формируя каждый уровень "дерева целей", ЛПР оценивает его составляющие на основе оценок "включить - не включить" в структуру. Причем, чем больше уровней структуризации, тем более дифференцированными являются оценки вышестоящих уровней этой структуры. Таким образом, оценка структуры фактически осуществляется в процессе формирования структуры.

Ниже приведен пример методики исследования, основывающейся на описанных выше методиках, для объектов исследования с достаточно большой степенью неопределенности.

#### **3.4.2.3. Пример методики исследования**

Объект исследования представляется как система. Исследование каждой системы проводится с использованием любых необходимых методов и операций системного анализа, а их последовательность определяется ведущим исследование системным аналитиком и во многом носит индивидуальный, приспособленный к данному случаю характер. Причем, для полного исследования системы рекомендуется начать анализ "с нуля" [36].

Для обеспечения полноты анализа проблемы рекомендуется использовать полную формальную модель проблемной ситуации, состоящей из трех взаимодействующих систем [46]: проблемосодержащей системы, проблеморазрешающей системы, окружающей среды. Модель проблемной ситуации позволяет выделить проблеморазрешающую систему, определить ее взаимодействие с внешней средой, представить модель структуры внешней среды предмета исследования.

Для разрешения проблемы ставится цель исследования и определяются критерии достижения цели. Далее цель представляется в виде совокупности целей-функций. Достигается это путем декомпозиции цели, то есть, строится дерево целей-функций, преломленное в потребностях проблемосодержащей системы, возможностях проблеморазрешающей системы и ограничениях окружающей среды.

В качестве принципов декомпозиции можно принять следующие:

- полнота охвата функций,
- равномерность структуризации,
- учет гипотезы Миллера,
- фрактальность,
- простота дерева целей-функций.

Если проблеморазрешающая система является социотехнической системой, то можно воспользоваться следующими признаками структуризации [18]:

- виды деятельности,
- целеполагающие системы,
- цикл управления,
- виды сред.

Основанием всякой декомпозиции является содержательная модель-основание [46]. Полнота декомпозиции обеспечивается полнотой модели-основания. Необходимым условием полноты модели-основания является полнота формальной модели (фрейма), наполнение которой конкретным содержанием приводит к содержательной модели-основанию. Включение в модель-основание только компонент, существенных к цели анализа, позволяет достичь компромисса между противоречивыми принципами полноты и простоты.

Разработанное с учетом полноты охвата функций, равномерности структуризации, учета гипотезы Миллера и фрактальности, и охватывающее основные виды деятельности, целеполагающие системы, цикл управления и

виды сред - дерево целей-функций предмета исследования представляет собой достаточно полную и непротиворечивую иерархию целей-функций, удовлетворяющую критерию функциональной полноты проектируемых систем.

Определяемые множество входных, выходных данных проблеморазрешающей системы, и оператор, отображающий входные данные в выходные в виде дерева целей-функций, позволяют представить функциональную модель системы.

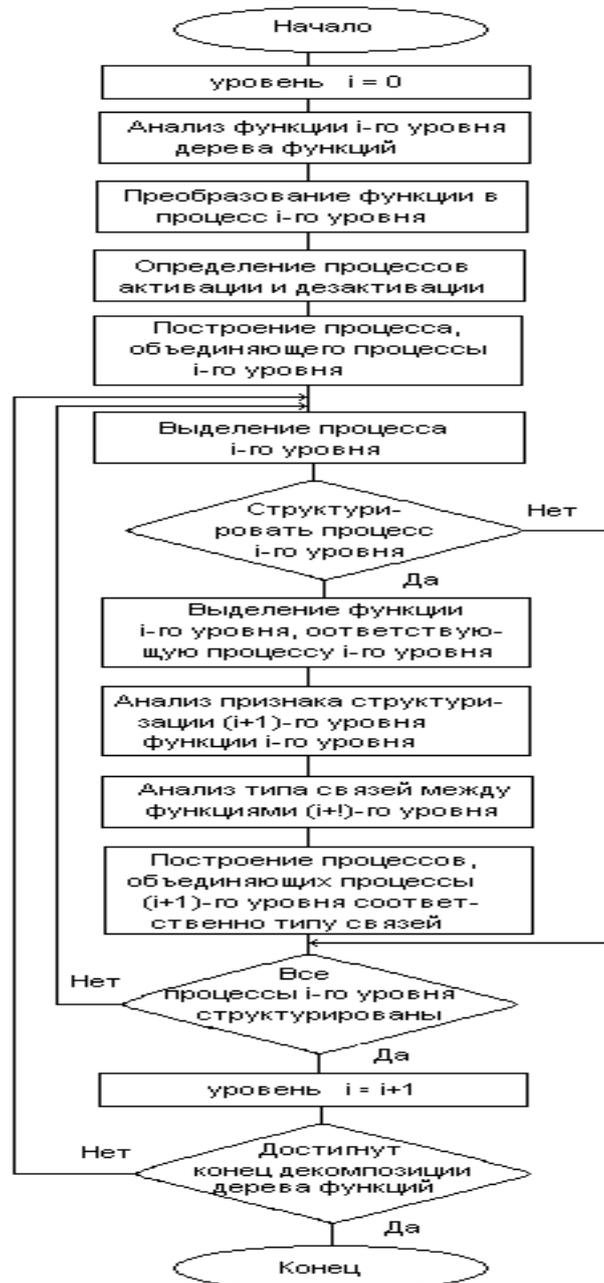
Обозначенные цели в дереве целей-функций, по существу, являются "планируемыми выходными" характеристиками начальных и промежуточных процессов, происходящих в системе.

Алгоритм перехода от дерева целей к взаимодействующим процессам представлен на рис. 3.10. Построение процессов системы проводится, начиная с вершины дерева (нулевой уровень декомпозиции). Анализируется цель-функция 0-го уровня дерева целей. Затем осуществляется преобразование цели-функции в процесс 0-го уровня. Для нулевого шага определяются процессы активации системы (начало функционирования) и деактивации системы (конец функционирования). Затем строится процесс, объединяющий процессы 0-го уровня.

В общем,  $i$ -шаг декомпозиции процессов заключается в выделении структурируемого процесса; в определении функции  $i$ -го уровня, соответствующей этому процессу; в анализе признака структуризации  $(i+1)$ -го уровня дерева целей для функции  $i$ -го уровня и типам связей, порождаемым этим признаком; преобразовании функций  $(i+1)$ -го уровня в процессы  $(i+1)$ -го уровня, которые объединяются в процессы, соответствующие типам связей.

Структуризация процессов заканчивается при достижении последнего уровня декомпозиции дерева целей-функций. Таким образом, дерево целей-функций является основой для построения модели взаимодействующих процессов.

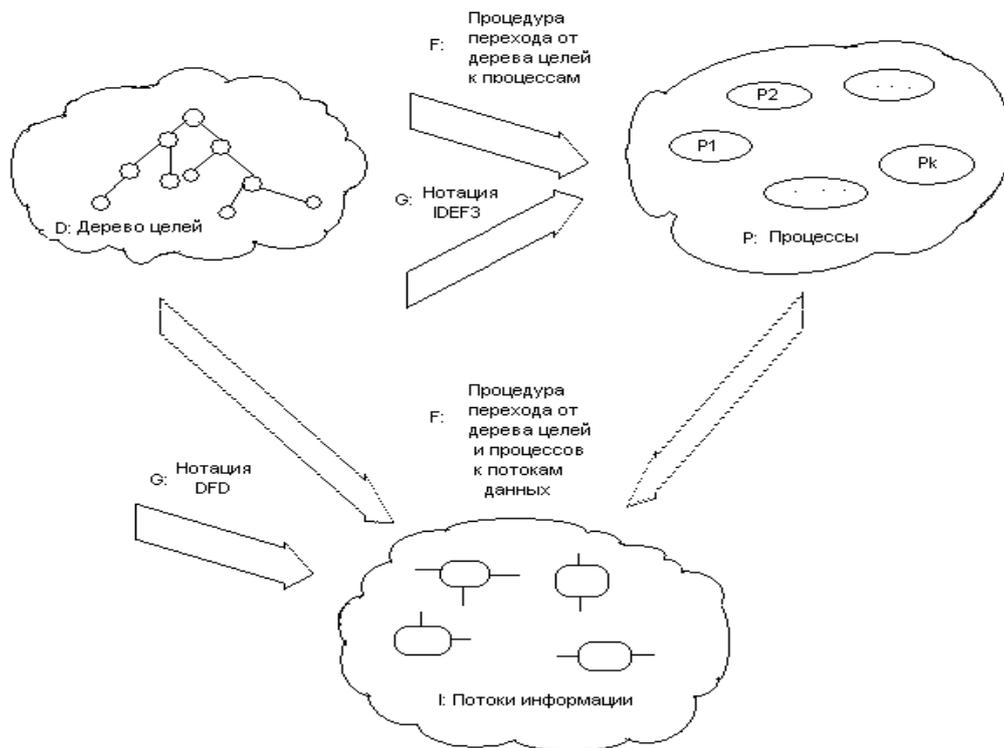
Построенные последовательности процессов определяют движение информации посредством своих входных и выходных данных, т.е. может быть построена модель информационных потоков.



*Рис. 3.10. Алгоритм перехода от дерева целей к взаимодействующим процессам*

Особенностью данной методики, в отличие от существующих методик (IDEF, UML, BPwin и др.), является следующая последовательность шагов моделирования, основывающаяся на дереве целей (рис. 3.11).

ФУНКЦИИ (дерево целей) -> ПРОЦЕССЫ (дерево целей, IDEF3)->  
-> ПОТОКИ ДАННЫХ (процессы, дерево целей, DFD) ->  
-> УЛУЧШЕНИЕ (оптимизация) ПРОЦЕССОВ И ПОТОКОВ



*Рис. 3.11. Последовательность шагов моделирования системы*

Предлагаемая последовательность моделирования близка к реализованной в системе AllFusion Process Modeler, но, в отличие от последней, на первом шаге предлагает построение дерева целей-функций, вместо IDEF0-моделей [56].

### **3.5. Теоретические и экспериментальные исследования**

Теоретические исследования могут предшествовать экспериментальным исследованиям или завершать их.

В первом случае, на этом этапе осуществляется формализация предметной области, выдвижение рабочей гипотезы, построение моделей исследуемого объекта. При исследовании систем, на этом этапе может проводиться проектирование систем. Затем проводятся экспериментальные исследования. После завершения теоретических и экспериментальных исследований проводится общий анализ полученных результатов, осуществляется сопоставление гипотезы с результатами экспериментов. В

результате анализа расхождений уточняются теоретические модели. В случае необходимости проводятся дополнительные эксперименты.

В ряде случаев порядок проведения исследования меняется. Случается, что теоретический и экспериментальный этапы научного исследования меняются местами. Тогда эксперимент может предшествовать теоретической части. Такая последовательность характерна для поисковых исследований, когда теоретическая база является недостаточной для выдвижения гипотез. В таком случае теория призвана обобщать результаты экспериментальных исследований.

### **3.5.1. Теоретические исследования**

Целью теоретических исследований является выделение в процессе синтеза знаний существенных связей между исследуемым объектом и внешней средой, обобщение и объяснение результатов эмпирического исследования, выявление общих закономерностей и их формализация.

Задачами теоретического исследования являются:

- обобщение результатов исследования;
- нахождение общих закономерностей путем обработки и интерпретации опытных данных;
- расширение результатов исследования на ряд подобных объектов без повторения всего объема исследований;
- изучение объекта, недоступного для непосредственного исследования;
- повышение надежности экспериментального исследования объекта (обоснования параметров и условий наблюдения, точности измерений).

Теоретическое исследование завершается формированием теории, необязательно связанной с построением ее математического аппарата. Теория проходит в своем развитии различные стадии от качественного объяснения и количественного измерения процессов до их формализации, и в

зависимости от стадии может быть представлена как в виде качественных правил, так и в виде математических соотношений.

### **3.5.1.1. Математическое моделирование системы**

При определении понятий математической модели и моделирования нужно исходить из общих понятий.

Моделирование представляет собой воспроизведение определенных свойств и связей исследуемого объекта (или объектов, порождающих явления, процессы) в другом, специально созданном объекте - в модели с целью их более тщательного изучения. Поскольку при моделировании в большинстве случаев имеют дело с частичным сходством или аналогией между оригиналом и моделью, то для описания процесса моделирования наиболее существенное значение имеют понятия и методы теории подобия и аналогии, которые опираются на понятие гомоморфизма. С этой точки зрения модель можно определить как такую систему объектов (материальных или идеальных), отношения между которыми гомоморфно отображают отношения между объектами оригинала.

В научном познании используют концептуальные, или понятийные, модели. Математическая модель – это разновидность концептуальной знаковой модели, так как, во-первых, выражает с помощью идей понятий, теорий некоторые существенные свойства и отношения оригинала; во-вторых, выступая как математическое описание действительности либо как интерпретация некоторой абстрактной системы, - она в известной мере замещает оригинал и служит для изучения последнего; в-третьих, исследование математической модели любого типа дает не просто новую информацию, а информацию, отличающуюся значительно большей точностью, ясностью и надежностью, которая присуща математическому методу. Но для этого, конечно, необходимо, чтобы сама модель была правильно построена.

Построение математических моделей скорее искусство, чем регламентируемая деятельность. Но определенные эвристические принципы, рекомендации, последовательности фаз создания модели, ее анализа и проверки – могут помочь в достижении этой цели.

Первая фаза начинается с установления существенных для изучаемого объекта или процесса свойств, отношений и закономерностей, которые должны быть выделены, абстрагированы от всех второстепенных, случайных, несущественных факторов и моментов, затрудняющих исследование. Вторая фаза математического моделирования сводится к формированию найденных на первом этапе качественных зависимостей на точном количественном языке математики, используя все теории и методы современной математики. Поскольку в самом начале исследования математическая модель выступает как некоторая гипотеза, то дальнейший шаг в исследовании состоит в тщательной разработке этой гипотезы. Третья фаза связана с логико-математическим анализом принятой модели, когда сначала математическая модель проверяется на противоречивость, а затем – из нее выводятся различные следствия. Четвертая фаза математического моделирования связана с интерпретацией полученных следствий с помощью известного эмпирического материала. И последняя, пятая фаза состоит в проверке полученных ранее эмпирически интерпретируемых следствий путем сопоставления их с результатами наблюдения или специально поставленного эксперимента.

Построение математической модели должно удовлетворять двум важнейшим, но противоречивым критериям: быть адекватной, но достаточно простой.

Математические модели, если классифицировать их по разным основаниям, бывают: поведенческие и структурные, эвристические и прогностические, функциональные, однозначные (динамические) и вероятностные (стохастические).

Математическая модель может выступать как описание и как интерпретация. Играя в познании разную роль, эти два понятия математической модели дополняют друг друга. Когда в естествознании и других науках применяются математические методы, то математическая модель выступает как описание реальных свойств и отношений мира, при анализе же формальных систем она служит для правильной интерпретации.

В новейших системных исследованиях роль математического моделирования настолько велика, что нередко общую теорию систем отождествляют с теорией, разрабатывающей принципы математического моделирования сложных систем. Единство процессов, протекающих в неживой материи биосферы, процессов биотического характера и процессов общественного развития, их взаимное влияние и обусловленность требуют сегодня по-иному взглянуть на смысл и цели исследовательской деятельности и предъявляют новые требования к методам этих исследований. Одно из них – это развитие принципов организации междисциплинарных системных исследований, позволяющих связывать в единое целое исследования разной природы. В этих условиях метод математических моделей приобретает особое значение, это не только способ исследования, но и язык, позволяющий преодолеть эффект «вавилонской башни» в тех исследованиях, которые ведут специалисты разных направлений. Система моделей начинает играть роль архитектурного проекта, каркаса всего исследования. Она позволяет сформулировать требования к объему и уровню детализации информации, выявить лакуны в наших знаниях, направления главных усилий и т.д. В этой ситуации очень важно стремиться к общему уровню математического описания.

Математическое моделирование является необходимой предпосылкой математического (вычислительного) эксперимента.

### 3.5.1.2. Основные этапы моделирования систем

Для моделирования необходимо создать модель и провести ее исследование. Перед созданием модели требуется конкретизировать цели моделирования. После исследования надо выполнить анализ результатов моделирования. Процесс создания модели проходит несколько стадий. Он начинается с изучения системы и внешних воздействий, и завершается разработкой или выбором математической модели или программы для ЭВМ, если моделирование будет проводиться с ее помощью. Некоторые математические модели могут быть исследованы без применения средств вычислительной техники.

Моделирование на ЭВМ предполагает наличие следующих укрупненных этапов:

- формулирование цели моделирования;
- разработки концептуальной модели;
- разработки математической модели;
- выбора метода моделирования;
- выбора средств моделирования;
- разработки программной модели;
- проверки адекватности и корректировки модели;
- планирования экспериментов на ЭВМ;
- моделирования на ЭВМ;
- анализа результатов моделирования.

Узловыми задачами моделирования являются:

- проверка адекватности модели,
- построение и преобразование алгоритмических моделей,
- упрощение и усложнение изучаемых моделей,
- повышение точности и достоверности результатов моделирования,
- обеспечение возможности исследования и получения характеристик для классов моделей,

- организация направленных вычислительных экспериментов,
- возможности структурного анализа моделей и т.д.

Перечисленные этапы вытекают из общей методологии моделирования систем. При моделировании различных систем трудоемкости одних и тех же этапов могут быть разными. В процессе моделирования конкретной системы могут иметь место некоторые изменения технологии. В частности, может быть заранее predetermined метод моделирования или выбрано конкретное средство моделирования. Математическая модель окажется настолько простой, что не потребуются проведения машинных экспериментов, разработка программной модели исключит необходимость создания математической модели.

#### **3.5.1.3. Постановка цели моделирования систем**

Для одной и системы можно составить множество моделей. Они будут отличаться степенью детализации и учета тех или иных особенностей и режимов функционирования, отражать определенную грань сущности системы, ориентироваться на исследование определенного свойства или группы свойств системы в зависимости от целей исследований.

Если может быть создано несколько вариантов системы, то из них требуется выбрать наилучший. Для оценки вариантов (для выбора лучшего) можно использовать критерий эффективности вариантов системы. Решение этих задач и является основной целью моделирования (в том числе, в выявлении вида критерия эффективности исследуемой системы).

Все последующие этапы моделирования проводятся с ориентацией на определенную цель моделирования.

#### **3.5.1.4. Создание концептуальных моделей систем**

Концептуальная (содержательная) модель - это абстрактная модель, определяющая состав и структуру системы, свойства элементов и причинно-следственные связи, присущие исследуемой системе и существенные для достижения цели моделирования. В концептуальной модели обычно в

словесной форме приводятся сведения о природе и параметрах элементарных явлений исследуемой системы, о виде и степени взаимодействия между ними, о месте и значении каждого элементарного явления в общем процессе функционирования системы.

Основная проблема при создании модели заключается в нахождении компромисса между простотой модели и ее адекватностью с исследуемой системой. Процесс создания концептуальной модели, очевидно, никогда не может быть полностью формализован. Именно в связи с этим иногда говорят, что моделирование является не только наукой, но и искусством.

На пути создания концептуальной модели выделяют: выбор уровня детализации, детализацию, локализацию, структуризацию, управление, выделение процессов, отражение состояний.

Проблема выбора уровня *детализации* может быть разрешена, путем построения иерархической последовательности моделей. Система представляется семейством моделей, каждая из которых отображает ее поведение на различных уровнях детализации. На каждом уровне существуют характерные особенности системы, переменные, принципы и зависимости, с помощью которых описывается поведение системы. Уровни детализации иногда называются стратами, а процесс выделения уровней - стратификацией. Выбор страт зависит от целей моделирования и степени предварительного знания свойств элементов. При построении ориентированной и стратифицированной концептуальной модели необходимо руководствоваться следующим. В модель должны войти все те параметры системы и, в первую очередь, параметры, допускающие варьирование в процессе моделирования, которые обеспечивают определение интересующих исследователя характеристик при конкретных внешних воздействиях на заданном временном интервале функционирования системы. Остальные параметры должны быть, по возможности, исключены из модели.

*Детализация* (расчленение системы на элементы) системы должна производиться до такого уровня, чтобы для каждого элемента были известны или могли бы быть получены зависимости параметров выходных воздействий элемента, существенных для функционирования системы и определения ее выходных характеристик, от параметров воздействий, которые являются входными для этого элемента. В модели должны присутствовать все элементы, которые реализуют выполнение всех технологических процессов. Кроме них в модель могут быть включены элементы, которые служат для управления ресурсами и процессами и для хранения объектов преобразования в промежутках времени между выполнением элементарных операций, а также для хранения информации, необходимой для управления.

*Локализация*, которая осуществляется путем представления внешней среды в виде генераторов внешних воздействий, включаемых в состав модели в качестве элементов.

Завершается построение *структуры* модели указанием связей между элементами. Связи могут быть подразделены на вещественные и информационные. Вещественные связи отражают возможные пути перемещения продукта преобразования от одного элемента к другому. Информационные связи обеспечивают передачу между элементами управляющих воздействий и информации о состоянии. Отметим, что как информационные, так и вещественные связи не обязательно должны быть представлены в системе некоторым материальным каналом связи. В простых системах, составленных из однофункциональных элементов, имеющих не более чем по одной выходной вещественной связи, информационные связи могут вообще отсутствовать.

*Управление* процессом функционирования в таких системах определяется самой структурой, т. е. в них реализован принцип структурного управления. Примерами таких систем могут служить логические элементы и аналоговые вычислительные машины. В более сложных системах,

включающих многофункциональные элементы или элементы, которые имеют больше чем по одной выходной вещественной связи, имеются управляющие средства (решающие элементы) и соответствующие информационные связи. Управление требуется для указания, какому элементу какой исходный объект когда и откуда взять, какую операцию по преобразованию выполнить и куда передать. О таких системах можно говорить, что они функционируют в соответствии с программным или алгоритмическим принципом управления. В концептуальной модели должны быть конкретизированы все решающие правила или алгоритмы управления рабочей нагрузкой, элементами и процессами.

Рассмотренные выше действия направлены на создание модели, отражающей статику системы — состав и структуру. Поскольку основной интерес представляют динамические системы, следует дополнить эту модель описанием работы системы, т.е. *выделением процессов*.

Функционирование системы заключается в выполнении технологического процесса преобразования вещества, энергии или информации. В сложных системах зачастую одновременно протекает несколько технологических процессов. Задается технологический процесс одним из видов изображения алгоритма.

Алгоритм однозначно определяет, какие ресурсы системы, в какой последовательности и какие операции должны выполнить для достижения некоторого целевого назначения системы. Совокупность алгоритмов управления совместно с параметрами входных воздействий и элементов отражают динамику функционирования системы.

Обычно алгоритмы преобразовываются к виду удобному для моделирования. Данный подход к описанию динамики работы системы особенно удобен для имитационного моделирования и является естественным способом определения множества характеристик системы:

$$V = \Phi (X, Q, A, T),$$

где  $\Phi$  — множество операторов вычисления выходных характеристик,

Q – параметры входных воздействий,

A – алгоритмы управления,

T – время,

X – элементы системы.

Созданная концептуальная модель должна быть проверена на адекватность исследуемому объекту. Поскольку на данном этапе возможен только умозрительный анализ и эксперимент, желательно, чтобы такую проверку выполняли эксперты, а не разработчик модели.

### **3.5.1.5. Разработка математической модели системы**

Под разработкой математической модели подразумевается создание полностью формализованного описания динамики функционирования системы.

Концептуальная модель и количественные исходные данные служат основой для разработки математической модели. Создание математической модели преследует две основные цели: 1) дать формализованное описание структуры и процесса функционирования системы для однозначности их понимания; 2) попытаться представить процесс функционирования в виде, допускающем аналитическое исследование системы.

Разработка единой методики создания математических моделей, очевидно, не представляется возможной. Это обусловлено большим разнообразием классов систем. В качестве формализованных схем, позволяющих описать функционирование системы, могут быть использованы: агрегативные системы, кусочно-линейные агрегаты, стохастические сети, сети Петри, системы массового обслуживания, непрерывные детерминированные системы, автоматы, модели системной динамики и др.[6]

### 3.5.1.6. Методы моделирования систем

Существующие методы моделирования систем [18] можно свести к такой обобщенной схеме на рис. 3.12.



Рис. 3.12. Обобщенная схема методов моделирования систем

### 3.5.1.7. Выбор средств моделирования систем

В качестве программных средств моделирования систем могут быть использованы процедурно-ориентированные алгоритмические языки, проблемно-ориентированные языки или автоматизированные системы моделирования (такие как Anylogic, GPSS).

Система GPSS World(tm) - это мощная среда компьютерного моделирования общего назначения, разработанная для профессионалов в области моделирования. Это комплексный моделирующий инструмент, охватывающий области как дискретного, так и непрерывного компьютерного моделирования, обладающий высоким уровнем интерактивности и визуального представления информации.

Более современная система AnyLogic обеспечивает поддержку всех этапов имитационного моделирования для различных типов динамических моделей – дискретных, непрерывных и гибридных, детерминированных и стохастических. Создание модели, ее выполнение, оптимизация параметров, анализ полученных результатов, верификация модели – все эти этапы реализованы в среде AnyLogic. AnyLogic используется для разработки имитационных исполняемых моделей и последующего их прогона с целью их анализа. Разработка модели выполняется в графическом редакторе AnyLogic с использованием многочисленных средств поддержки, упрощающих работу. Построенная модель затем компилируется встроенным компилятором AnyLogic и запускается на выполнение. В процессе выполнения модели пользователь может наблюдать ее поведение, изменять параметры модели, выводить результаты моделирования в различных формах и выполнять разного рода компьютерные эксперименты с моделью.

Графический редактор AnyLogic позволяет разработчику модели описывать структуру моделируемой системы, ее подсистемы и поведение объектов в модели графически, выполняя генерацию соответствующего

программного кода автоматически. Графическая разработка легче, быстрее, понятнее, чем написание программных текстов.

### **3.5.1.8. Проверка адекватности и корректировка модели системы**

Проверка адекватности проводится в том или ином виде, начиная от разработки концептуальной модели и кончая анализом результатов.

Проверка адекватности модели системе заключается в анализе ее соразмерности с исследуемой системой, а также равнозначности системе.

Адекватность моделей проверяется путем:

- проверки законов, справедливых для изучаемых моделей,
- либо путем соответствия структуры математической и машинной модели,
- соответствия поведения моделируемой и реальной системы,
- правильности интерпретации результатов.

На практике оценка адекватности обычно проводится путем экспертного анализа разумности результатов моделирования. Можно выделить следующие виды проверок:

- проверка моделей элементов (в сомнительных случаях следует детализировать элемент или провести дополнительный анализ);
- проверка модели внешних воздействий (принятые предположения, аппроксимации и гипотезы необходимо оценить математическими методами);
- проверка концептуальной модели функционирования системы (выявляются ошибки постановки задачи);
- проверка формализованной и математической модели;
- проверка способов измерения и вычисления выходных характеристик (выявляются ошибки решения);
- проверка программной модели (анализируется соответствие операций и алгоритмов функционирования программной и математической модели, приводятся контрольные расчеты при типовых и предельных

- значениях переменных, выявляются инструментальные ошибки программирования),
- проверка соответствия поведения моделируемой и реальной системы,
  - проверка правильности интерпретации результатов.

### **3.5.2. Экспериментальные исследования**

#### **3.5.2.1. Порядок проведения**

Экспериментальные исследования продолжают теоретический этап. Эксперимент как поставленный научно опыт, является наиболее сложной и трудоемкой частью исследования. Его цели могут быть различными, поскольку зависят они от характера всего исследования, а также последовательности его проведения.

Перед организацией экспериментальных исследований разрабатываются задачи, выбираются методика и программы эксперимента. Его эффективность существенно зависит от выбора средств измерений. При решении этих задач необходимо руководствоваться инструкциями и ГОСТами.

Принимаемые методические решения формулируются в виде методических указаний на проведение эксперимента.

После разработки методик исследования составляется рабочий план, в котором указываются объем экспериментальных работ, методы, техника, трудоемкость и сроки.

#### **1.5.2.2. Виды экспериментов**

Постановка и организация эксперимента определяется его назначением. Эксперименты различаются по способу формирования условий (естественных, искусственных), по целям исследования (преобразующие, констатирующие, контролирующие, поисковые, решающие), по организации проведения (лабораторные, натурные, производственные и т.п.), по структуре изучаемых объектов (простые, сложные), по характеру внешних воздействий на объект исследования (вещественные, энергетические, информационные),

по характеру взаимодействия средств экспериментального исследования с объектом исследования (обычный, модельный), по типу моделей (материальный, мысленный), по контролируемым величинам (пассивный, активный), по числу варьируемых параметров (однофакторный, многофакторный), по средствам исследования (математический, физический и т.п.) и т.п.

В научных исследованиях по информатике особая роль отводится математическому эксперименту.

### **1.5.2.3. Математический эксперимент**

Математический (вычислительный) эксперимент – это метод наиболее совершенного мысленного экспериментирования с моделями сложных реальных процессов и технических систем, выраженных на языке математики, имеющей своей целью проверку эмпирических обобщений, гипотез, теорий в науке и эвристическое выявление некоторых закономерностей природы, и достигающий своей цели посредством ЭВМ и математических моделей.

Возникновение математического эксперимента стало возможным благодаря появлению компьютеров, работающих в режиме диалога, усовершенствованию теории и практики программирования и разработки теории численных методов, развитию и усовершенствованию методов построения математических моделей, использованию в этих целях языка не только классической, но и современной математики.

Математический эксперимент подобен математическому моделированию тем, что здесь также идет построение математической модели, затем нахождение приближенного численного метода решения задачи, программирование, расчет на ЭВМ, анализ и интерпретация результатов, полученных в ходе исследования построенной математической модели. Математический эксперимент отличается от математического моделирования тем, что на этом он не кончается, а в случае

неудовлетворительного результата все расчеты вновь повторяются, но уже с новой моделью, пока не будет найдена математическая модель, наилучшим образом описывающая исследуемые процессы. То есть математический эксперимент – это эксперимент над моделями.

Цели моделирования достигаются путем исследования разработанной модели. Исследования заключаются в проведении экспериментов, в результате которых определяются выходные характеристики системы при разных значениях управляемых переменных параметров модели. Эксперименты следует проводить по определенному плану. Особую важность приобретает планирование экспериментов при численном и статистическом имитационном моделировании на ЭВМ. Это обосновывается большим числом возможных сочетаний значений управляемых параметров, а каждый эксперимент проводится при определенном сочетании значений параметров. При ограниченных вычислительных и временных ресурсах обычно не представляется возможным провести все эксперименты. Возникает необходимость в выборе определенных сочетаний параметров и последовательности проведения экспериментов. Это называется стратегическим планированием.

Разработка плана начинается на ранних этапах создания модели, когда выявляются характеристики качества и параметры, с помощью которых предполагается управлять качеством функционирования системы. Эти параметры называют в теории планирования экспериментов факторами. Затем выделяются возможные значения количественных параметров и варианты качественных (функциональных) параметров. Их называют уровнями.

Если число факторов велико, то для проведения исследований системы используется один из методов составления плана по неполному факторному анализу. Эти методы хорошо разработаны в теории планирования экспериментов. Особую важность приобретает тщательное планирование экспериментов при исследовании нестационарных систем в связи с

необходимостью существенного увеличения общего количества экспериментов.

Совокупность методов уменьшения длительности машинного эксперимента при обеспечении статистической достоверности результатов имитационного моделирования получила название тактического планирования.

На длительность одного эксперимента (периода моделирования  $T$ ) влияет степень стационарности системы, взаимозависимости характеристик и значения начальных условий моделирования.

Большинство имитационных моделей используется для изучения установившихся равновесных режимов функционирования. Но в начальный период работы системы или модели существует переходный режим даже при неизменных значениях параметров входных воздействий. Причем длительность переходного режима может быть весьма большой. Значения выходных характеристик, измеренные в переходный период, смещают их общие оценки.

Существует три основных метода уменьшения ошибки, обусловленной начальными условиями.

Первый состоит в достаточном увеличении периода моделирования. С увеличением числа замеров влияние начального смещения на статистическую оценку стремится к нулю.

Второй метод состоит в том, чтобы начинать сбор статистики не с начального момента, а по истечении некоторого времени.

Третий метод заключается в инициализации модели не с «нулевого», а специально заданного состояния, близкого к установившемуся.

Первые два метода приводят к увеличению длительности эксперимента и не дают гарантии уменьшения ошибки, так как априорно неизвестна длительность переходного режима. Третий метод можно применять при наличии информации о подходящем начальном состоянии. В последующих

экспериментах для задания начальных состояний могут использоваться уточненные сведения из предшествующих экспериментов.

При моделировании нестационарных систем установившийся режим может полностью отсутствовать. Естественным методом определения характеристик имитационного моделирования нестационарных систем является метод повторных экспериментов. В этом случае число экспериментов существенно увеличивается, что приводит к особым требованиям по их планированию.

Вычислительный эксперимент объединяет в себе две составляющие:

- аналитическую,
- имитационную.

Первая связана с реализацией выявленных закономерностей объекта в цифровой модели, и проведением вычислительных операций.

Вторая связана с имитацией неизученных сторон объекта, и имитацией внешней взаимодействующей среды объекта.

Организация экспериментов носит итеративный характер. По мере проведения имитационных и аналитических экспериментов алгоритмически-цифровая модель объекта (явления, системы) уточняется и порождает ряд новых экспериментов и анализы с обработкой результатов.

Математический эксперимент играет важную роль как принципиально новый метод научного познания, занимающий промежуточное место между классическим дедуктивным и классическим экспериментальным методом исследования, сочетающий теорию и практику, дедукцию и индукцию.

### **3.5.3. Анализ и интерпретация экспериментальных данных**

После завершения теоретических и экспериментальных исследований проводится общий анализ полученных результатов, осуществляется сопоставление гипотезы с результатами эксперимента. В результате анализа расхождений уточняются теоретические модели. В случае необходимости проводятся дополнительные эксперименты.

Этот этап подразумевает необходимость сравнения теоретического и экспериментального этапов научного исследования для окончательного подтверждения гипотезы и дальнейшего формулирования выводов и вытекающих из нее следствий. Иногда результат бывает и отрицательным, тогда гипотезу приходится отвергать.

При статистическом моделировании в ходе имитационного эксперимента измеряются множества значений по каждой выходной характеристике. Эти выборки необходимо обрабатывать для удобства последующего анализа и использования. Поскольку выходные характеристики зачастую являются случайными величинами или функциями, обработка заключается в вычислении оценок математических ожиданий, дисперсий и корреляционных моментов. Оценки, полученные в результате статистической обработки измерений, должны быть состоятельными, несмещенными и эффективными.

Для того чтобы исключить необходимость хранения в машине всех измерений, обработку проводят по рекуррентным формулам, когда оценки вычисляют в процессе эксперимента методом нарастающего итога по мере появления новых измерений.

Для стохастических характеристик можно построить гистограмму относительных частот — эмпирическую плотность распределения. С этой целью область предполагаемых значений характеристики разбивается на интервалы. В ходе эксперимента по мере измерений определяют число попаданий характеристики в каждый интервал и подсчитывают общее число измерений. После завершения эксперимента для каждого интервала вычисляют отношение числа попаданий характеристики к общему числу измерений и длине интервала. Для построенной гистограммы можно попытаться подобрать теоретический закон распределения. Делается это так же, как и при подготовке исходных данных моделирования.

Если искомая характеристика является стационарной случайной функцией времени и обладает свойством эргодичности, то для ее оценки вычисление среднего по времени заменяется вычислением среднего по

множеству измерений при одном достаточно продолжительном эксперименте.

Для случайных нестационарных характеристик период моделирования  $T$  разбивается на отрезки с постоянным шагом (прогоны или сечения), и запоминаются значения характеристики в конце каждого прогона. Проводится серия экспериментов с разными последовательностями случайных параметров модели. Затем измерения каждого сечения обрабатываются как при оценке случайных величин. Процессы обработки измерений имитационного эксперимента направлены на получение интегральных характеристик, т. е. на сжатие данных.

По результатам статистического моделирования может быть проведен анализ зависимостей характеристик от параметров системы и внешних воздействий. Для этого можно воспользоваться корреляционным, дисперсионным или регрессионным методами.

С помощью корреляционного анализа можно установить наличие связи между двумя или более случайными величинами. Оценкой связи служит коэффициент корреляции при наличии линейной связи между величинами и нормальном законе их совместного распределения. Коэффициент корреляции, равный единице по абсолютной величине, свидетельствует о наличии функциональной нестохастической линейной связи между анализируемыми величинами. При равенстве нулю коэффициента корреляции связь отсутствует. Промежуточные значения коэффициента корреляции соответствуют наличию линейной связи с рассеянием или нелинейной корреляции.

Дисперсионный анализ можно использовать для установления относительного влияния различных факторов на значения выходных характеристик. При этом общая дисперсия характеристики разлагается на компоненты, соответствующие рассматриваемым факторам. По значениям отдельных компонентов делают вывод о степени влияния того или другого фактора на анализируемую характеристику.

Когда все факторы в эксперименте являются количественными, можно найти аналитическую зависимость между характеристиками и факторами. Для этого используются методы регрессионного анализа. Найденная зависимость называется эмпирической моделью. Регрессионный анализ заключается в том, что выбирается вид соотношения между зависимыми и независимыми переменными, по экспериментальным данным вычисляются параметры выбранной зависимости и оценивается качество аппроксимации экспериментальных данных моделью. Если качество неудовлетворительное, берется зависимость другого вида, и процедура повторяется.

К анализу результатов моделирования можно отнести задачу анализа чувствительности модели к вариациям ее параметров. Под анализом чувствительности понимают проверку устойчивости характеристик процесса функционирования системы к возможным отклонениям значений параметров.

Анализ результатов моделирования позволяет уточнить множество информативных параметров модели, что может привести к существенному изменению первоначального вида концептуальной модели; позволяет найти функциональные зависимости характеристик и параметров, что иногда дает возможность создать аналитические модели системы, или определить весовые коэффициенты критерия эффективности.

#### **3.5.4. Обоснование результатов исследования**

Результатом научных исследований являются новые научные знания.

*Научное знание* – это продукт специализированной, профессиональной формы человеческой деятельности, которая предполагает существование особой цели его приложения, а также применение научных методов.

Методологические требования к научному знанию – быть объективным, доказательным, точным, принципиально критичным, ориентированным на адекватное постижение реальности [40].

Важнейшими способами обоснования полученного знания являются:

- многократные проверки результатов исследования наблюдениями и экспериментами;

- обращение к статистическим данным, которые получены разными учеными независимо друг от друга.

При обосновании теоретических концепций обязательными являются следующие требования:

- непротиворечивость;
- соответствие эмпирическим данным;
- возможность описывать известные явления;
- способность предсказывать новые явления, факты.

Следует строго соблюдать один из законов логики - закон достаточного основания

- всякая мысль, чтобы стать достоверной, должна быть обоснована другими мыслями, истинность которых доказана или самоочевидна.

Обоснованность результатов диссертационного исследования достигается:

- базированием на строго доказанных и корректно используемых выводах фундаментальных и прикладных наук, положения которых нашли применение в работе;

- комплексным использованием известных, проверенных практикой теоретических и эмпирических методов исследования;

- указанием на то, что решение ряда новых задач стало возможным благодаря известным достижениям определенных научных дисциплин и не противоречит их положениям, а методики согласуются с опытом их создания и дальнейшего совершенствования;

- устранением противоречий между теоретическими положениями, развитыми исследователем и известными законами эволюции науки, техники, знания. Обоснованием результатов с помощью известных процедур проектирования, методов поиска технических решений, а также физического и математического моделирования;

- проверкой теоретических положений и новых решений, идей экспериментальными исследованиями, которые должны быть метрологически обеспечены;

- сопоставлением результатов эксперимента и испытаний, проведенных исследователем, с известными экспериментальными данными других исследователей по тем же проблемам.

Необходимая полнота решения проблемы достигается с помощью экспериментальной проверки теоретических положений исследования, а также согласованностью собственных экспериментальных данных с экспериментальными данными других исследователей.

Достаточность решения заключается в согласованности полученных исследователем экспериментальных данных с известными теоретическими положениями других авторов и с обоснованными и согласованными теоретическими решениями, полученными лично [10].

Истинным (верным) считается то, что получено с помощью корректных логических выводов и доказательств, подтверждено корректным физическим и/или машинным экспериментом, а также принципиально другими вариантами решений [24].

### **3.6. Оформление результатов научной работы**

#### **3.6.1. Публикация результатов исследований**

Для доведения до научной общественности результатов научных исследований последние оформляются в виде отчета, статьи, сообщения, и публикуются в научных журналах и/или докладываются на научных конференциях.

*Отчет по НИР*, согласно ГОСТ 7.32-81, содержит следующие разделы:

- аннотация и реферат с кратким изложением задач исследования и полученных результатов,

- введение с характеристикой отечественных и зарубежных достижений по исследуемой теме,

- постановка задачи,
- формулировка технического задания,
- анализ известных методов и способов решения задачи,
- обоснование выбранного метода решения,
- расчеты и результаты экспериментов приводятся в форме, дающей возможность проанализировать справедливость полученных результатов,
- выводы с сопоставлением и анализом полученных в процессе исследования теоретических и экспериментальных данных,
- заключение с оценкой результатов и указанием путей их использования,
- приложения, куда выносятся вся доказательная база исследования (подробные доказательства, таблицы с результатами экспериментов, графики, схемы и т.п.).

*Аннотация* – это краткая характеристика отчета либо другого документа сточки зрения содержания, назначения, формы и других особенностей. Аннотация выполняет сигнальную функцию и отвечает на вопрос: «О чем говорится в первичном документе?». Согласно ГОСТ 7.9-77 она включает характеристику типа документа, основной темы, объекта, цели работы и ее результаты. Средний объем аннотации 600 знаков.

*Реферат* представляет собой сокращенное изложение содержания первичного документа с основными фактическими сведениями и выводами. Реферат выполняет познавательную функцию, отвечая на вопрос: «Что говорится в первичном документе?». Согласно ГОСТ 7.9-77 реферат включает заглавие, тему, предмет, характер и цель работы, методы проведения работы, конкретные результаты работы. Средний объем реферата от 500-2500 знаков [33].

Для *научной статьи* рекомендуется придерживаться следующего общего плана изложения, включающего: оглавление, предисловие, введение, краткий обзор литературы, основное содержание, выводы, заключение, перечень литературных источников.

В предисловии излагаются основания для проведения НИР, когда и где была выполнена работа, перечисляются организации и лица, оказывавшие содействие при ее выполнении. Во введении дается постановка основной темы исследования, определяется ее значение, актуальность, цели и задачи, состояние на данный момент. Далее дается краткий обзор литературы для выяснения состояния разработки темы.

В основное содержание работы включаются методы, алгоритмы решения, экспериментальные данные, обобщения и выводы самого исследования. Каждая часть основного содержания заканчивается выводами в виде кратко сформулированных тезисов, в которых формулируются результаты исследования, описанные в данной части.

В заключении дается обобщение наиболее существенных положений проведенного научного исследования, подводятся его итоги, формулируются вопросы, которые еще требуют разрешения.

Конкретные требования к оформлению научной статьи публикуются в каждом научном журнале. Объем статьи колеблется от 3 до 10 страниц. При превышении этого объема статья, как правило, депонируется. Статья меньшего размера относится к тезисам, сообщениям.

*Научный доклад* должен содержать [44]: преамбулу, введение, методы и методики исследования, полученные результаты, выводы и заключение.

В преамбуле следует показать, что исследуемые вопросы интересны и теоретически разрешимы, предлагаемые решения оригинальны и эффективны, а исследование имеет признаки новизны.

Во введении обосновывается актуальность исследования путем определения важности предметной области, указания на недостаточно удовлетворительные предыдущие решения проблемных вопросов, и, как следствие, формулируются цели и задачи нового исследования.

Далее обосновывается применяемая методика исследования. Указываются особенности построенных моделей исследуемого объекта, основные оцениваемые параметры моделей. Демонстрируются результаты

математических экспериментов, оформленные в виде обобщенных функциональных зависимостей.

Затем формулируются выводы в виде кратких нумерованных тезисов результатов исследования. В заключении приводится оценка эффективности полученных результатов.

### **3.7. Внедрение и эффективность научных исследований**

#### **3.7.1. Внедрение научных исследований**

Внедрение фундаментальных и прикладных научных исследований в производство осуществляется через разработки, проводимые, как правило, в опытно-конструкторских бюро, проектных организациях, опытных заводах и мастерских. Разработки оформляются в виде опытно-технологических или опытно-конструкторских работ, включающих формулировки темы; цели и задачи разработки; изучение литературы; подготовку к техническому проектированию экспериментального образца; техническое проектирование (разработка вариантов технического проекта с расчетами и разработкой чертежей); изготовление отдельных блоков, их объединении в систему; согласование технического проекта и его технико-экономическое обоснование. После этого выполняется рабочее проектирование (детальная проработка проекта); изготавливается опытный образец; производятся его опробование, доводка и регулировка; стендовые и производственные испытания. После этого осуществляется доработка опытного образца (анализ производственных испытаний, переделка и замена отдельных узлов).

Успешное выполнение перечисленных этапов работы дает возможность представить образец к государственным испытаниям, в результате которых образец запускается в серийное производство. Разработчики при этом осуществляют контроль и дают консультации. Внедрение завершается оформлением акта экономической эффективности результатов исследования.

### 3.7.2. Эффективность научной работы

Эффект от внедрения НИР – это результат, который находит отражение в сокращении живого и овеществленного труда на производство продукции.

Эффект научных исследований может иметь различную природу: *экономический, социально-экономический, укрепление обороноспособности* и т.д.

Фактическую годовую экономию живого и овеществленного труда, выраженную в денежных единицах, называют *годовым экономическим эффектом*, который, в зависимости от стадии завершения работы, может быть предварительным, ожидаемым, фактическим, потенциальным.

*Предварительный экономический эффект* рассчитывается на стадии технико-экономического обоснования целесообразности постановки исследований по укрупненным показателям на ожидаемый объект внедрения.

*Ожидаемый экономический эффект* рассчитывается на стадии завершения научных исследований по результатам НИР и нормативно-справочным показателям на планируемый объем внедрения. Ожидаемый экономический эффект может быть выражен в процентном виде на основе экспертных оценок.

*Фактический экономический эффект* рассчитывается после внедрения разработки по фактическим показателям отчетного года и действующим нормам организации, осуществившей внедрение.

*Потенциальный экономический эффект* – это сумма, рассчитанная по укрупненным показателям на возможный объем внедрения.

Эффективность исследования коллектива и отдельного работника оценивается по-разному. Эффективность работы одного научного работника оценивают числом публикаций, новизной разработок и др. Эффективность научно-исследовательской группы оценивается показателями экономической эффективности и производительности труда, количеством внедренных тем, авторских свидетельств и патентов и т.п. [33].

#### 4. Информационная поддержка научных исследований

Информационная поддержка научных исследований возможна на всех этапах от сбора данных до оформления полученных результатов. Особое место занимают технологии поддержки принятия решений.

Профессиональные математические пакеты - это программы (пакеты программ), обладающие средствами выполнения различных численных и аналитических (символьных) математических расчетов, от простых арифметических вычислений, до решения уравнений с частными производными, решения задач оптимизации, проверки статистических гипотез, средствами конструирования математических моделей и другими инструментами, необходимыми для проведения разнообразных технических расчетов. Все они имеют развитые средства научной графики, удобную справочную систему, а также средства оформления отчетов.

Сейчас существует множество программ, которые способны осуществить поддержку научных исследований и помочь в построении эксперимента, получении данных и их обработке (см. таблицу 4.1).

Таблица 4.1. Информационный инструментарий для научных исследований.

Программа	Описание
<i>Математические программы</i>	
<b>MATLAB</b>	Система MatLab относится к среднему уровню продуктов, предназначенных для символьной математики, но рассчитана на широкое применение. Изначально MatLab предназначалась исключительно для вычислений, в процессе эволюции (а сейчас выпущена уже версия 7), в дополнение к прекрасным вычислительным средствам, у фирмы Waterloo Maple по лицензии для MatLab было приобретено ядро символьных преобразований, а также появились библиотеки, которые обеспечивают уникальные для математических пакетов функции. Например, широко известная библиотека Simulink. В системе MatLab также существуют широкие возможности для программирования. Ее библиотека C Math (компилятор MatLab) является объектной и содержит свыше 300 процедур обработки данных на языке C. Для визуализации моделирования система MatLab имеет библиотеку Image Processing Toolbox, которая обеспечивает широкий спектр функций, поддерживающих визуализацию проводимых вычислений непосредственно из среды MatLab, увеличение и анализ, а

	также возможность построения алгоритмов обработки изображений.
<b>Octave</b>	Это программа числовых вычислений, хорошо совместимая с MatLab. Распространяется Octave на условиях открытой лицензии с исходным кодом .
<b>KOctave</b>	Программа KOctave по сути представляет собой более продвинутый графический интерфейс для системы Octave. В результате использования KOctave система Octave становится полностью похожей на MatLab.
<b>Genius</b>	Небольшая математическая программа Genius. Распространяется Genius тоже на условиях открытой лицензии с исходным кодом (OpenSource). Она имеет собственный язык GEL, развитый инструментарий Genius Math Tool и хорошую систему подготовки документов для публикации (с использованием таких языков оформления, как LaTeX, Troff (eqn) и MathML).
<b>Maple</b>	Пакет Maple — совместная разработка Университета Ватерлоо (шт. Онтарио, Канада) и Высшей технической школы (ETHZ, Цюрих, Швейцария). Программа Maple является одним из лидеров среди универсальных систем символьных вычислений. Она предоставляет пользователю удобную интеллектуальную среду для математических исследований любого уровня и пользуется особой популярностью в научной среде.
<b>Derive</b>	маленькая коммерческая математическая система Derive привлекательна своей нетребовательностью к аппаратным ресурсам ПК. Derive является больше учебной системой компьютерной алгебры начального уровня.
<b>Scientific Workplace</b>	Scientific WorkPlace развивалась как редактор научных текстов. Однако со временем компания MacKichan Software, Inc. Лицензировала символьный движок Maple у компании Waterloo Maple, Inc., и теперь эта программа объединяет простой в использовании текстовый процессор, обеспечивающий создание математических текстов и систему компьютерной алгебры в одной среде.
<b>YaCaS</b>	YaCaS (аббревиатура от выражения Yet Another Computer Algebra System — еще одна система компьютерной алгебры), то это бесплатная кроссплатформенная альтернатива Maple, построенная на тех же принципах. Мощный и высокоэффективный движок YaCaS полностью реализован на C++ на условиях открытой лицензии (OpenSource).
<b>Mathematica</b>	Компания Wolfram Reseach, Inc., разработавшая систему компьютерной математики Mathematica, по праву считается старейшим и наиболее солидным игроком в этой области. Система Mathematica сегодня рассматривается как мировой лидер среди компьютерных систем символьной математики для ПК, обеспечивающих не только возможности выполнения сложных численных расчетов с выводом их результатов в самом изысканном графическом виде, но и проведение особо трудоемких аналитических преобразований и вычислений.
<b>Maxima</b>	Отметим, что система Maxima — это некоммерческий проект с открытым кодом. В программе Maxima для математической работы используется язык, сходный с языком в пакете Mathematica, а графический интерфейс построен по тем же принципам. Изначально программа называлась Xmaxima и создавалась для UNIX-систем. Кроме того, сейчас у системы Maxima есть еще более мощный, эффективный и дружелюбный кроссплатформенный графический

	интерфейс.
<b><i>Kalamaris</i></b>	Что касается программы Kalamaris, то это также новый проект, который имеет подход и идеологию, схожие с системой Mathematica. Проект еще не завершен, но тоже является неплохой бесплатной альтернативой такому коммерческому монстру, как Mathematica.
<b><i>MathCAD</i></b>	Простой, продвинутый редактор математических текстов с широкими возможностями символьных вычислений и прекрасным интерфейсом. MathCad не имеет языка программирования как такового, а движок символьных вычислений заимствован из пакета Maple. Пакет имеет удобные возможности импорта/экспорта данных. Например, можно работать с электронными таблицами Microsoft Excel прямо внутри MathCad-документа.
<b><i>MuPAD</i></b>	Представляет собой современную интегрированную систему математических вычислений, при помощи которой можно производить численные и символьные преобразования, а также чертить двумерные и трехмерные графики геометрических объектов.
<b><i>KmPlot</i></b>	Программа KmPlot распространяется на условиях открытой лицензии с исходным кодом. Имеет простой интерфейс.
<b><i>Программы моделирования</i></b>	
<b><i>Electronics Workbench</i></b>	Предназначено для проектирования и моделирования электрических схем и процессов, однако возможно моделирование и на уровне простейших функциональных блоков. Продукт позволяет создавать схемы любой сложности из библиотек элементов и проводить их полное тестирование при помощи стандартных инструментов. Программа удобна тем, что пользователь может снять различные характеристики в любой точке схемы, а также получить их представление в цифровой и аналоговой форме, в виде графиков.
<b><i>Commsim</i></b>	Применяется для разработки и имитации коммуникационных систем. Он позволяет моделировать как аналоговые, так и цифровые коммуникационные системы на уровне функциональных блок-диаграмм. Данный инструмент существенно сокращает время на конструирование и разработку новых систем или сетей. Достаточно сказать, что типичное приложение, выполненное в этом средстве моделирования, — ретранслятор, мобильный телефон, радиочастотный модем.
<b><i>Model Vision Studium (MVS)</i></b>	Компьютерная лаборатория для моделирования и исследования сложных динамических систем, интегрированная графическая оболочка для быстрого создания интерактивных визуальных моделей сложных динамических систем и проведения вычислительных экспериментов с ними. Пакет MVS позволяет описать модель на специальном графическом языке, а затем автоматически построить программу для воспроизведения ее поведения, использующую для этого современные численные методы. Пакет предназначен для исследования гибридных, обладающих одновременно «непрерывными» и «дискретными» свойствами систем.
<b><i>AnyLogic</i></b>	Программное обеспечение для имитационного моделирования бизнес-процессов, разработанное российской компанией «Экс Джей Текнолоджис» (англ. <i>XJ Technologies</i> ). Инструмент обладает современным графическим интерфейсом и позволяет использовать язык Java для разработки моделей/

<i><b>Статистические программы</b></i>	
<i><b>SPSS</b></i>	Аббревиатура от Statistical Package for the Social Sciences, относительно недавно стала принадлежать компании IBM и сменила название на PASW (Predictive Analytics SoftWare) Statistics. Для работы SPSS необходимо не менее 1 Гб оперативной памяти. Операционные системы, в которых можно запускать SPSS: Windows, MacOS и Linux.
<i><b>R</b></i>	Среда статистических вычислений и язык программирования. Стандартная комплектация R не предполагает графического интерфейса. В системе R имеются широкие возможности для статистической обработки данных, в том числе и для работы с графикой, а оконный интерфейс можно установить как дополнительное приложение.
<i><b>Statistica</b></i>	Statistica разрабатывается компанией StatSoft. Statistica разработана под Windows. Имеет широкий набор методов.
<i><b>Stadia</b></i>	Система контекстной экранной помощи включающей объемный справочник-гипертекст и экспертную систему по выбору метода статистического анализа, является российской разработкой.
<i><b>StatGraphics</b></i>	универсальный, многопрофильный пакет с хорошо методически продуманным меню-ориентированным интерфейсом пользователя.
<i><b>SAS</b></i>	зародилась в 60-х годах XX века в Университете Северной Каролины как приложение для анализа результатов сельскохозяйственных исследований. На сегодняшний день система продолжает развиваться в компании SAS Institute. Система состоит из модулей, каждый из которых выполняет определенный круг задач. Наиболее часто при статобработке используются модули BASE и STAT. В системе SAS реализован собственный язык программирования, который по своему синтаксису ближе к бэйсику.
<i><b>Stata</b></i>	Разрабатывается корпорацией StataCorp. Приложение может работать на операционных системах семейства Windows, в MacOS и Linux. Ввод данных здесь возможен как путем загрузки из внешних файлов, так и с использованием встроенного табличного редактора, который довольно прост, но позволяет выполнять все необходимые манипуляции с таблицами.
<i><b>Мезозавр</b></i>	В пакете реализована оригинальная система экспертной оценки сложных моделей временных рядов.
<i><b>MedCalc</b></i>	Разрабатывается бельгийской компанией MedCalc Software, позиционируется как полноценное статистическое приложение, созданное в соответствии с потребностями исследователей биомедицинского направления. Разработчики акцентируют внимание исследователей на удобстве использования MedCalc для анализа ROC-кривых.

#### **4.1. Программы статистической обработки данных**

Системы статистической обработки данных можно разделить на:

- интегрированные методо-ориентированные пакеты общего назначения;
- специализированные методо-ориентированные пакеты;

- обучающие программы;
- предметно-ориентированные пакеты.

В *универсальных пакетах*, предлагающих широкий диапазон статистических методов, отсутствует ориентация на конкретную предметную область. Из зарубежных универсальных пакетов наибольшую известность получили компьютерные системы SAS, SPSS, SYSTAT, Minitab, Statgraphics, *Statistica*.

*Специализированные пакеты*, как правило, содержат несколько статистических методов или методы, применяемые в конкретной предметной области. Чаще всего это системы, ориентированные на анализ временных рядов, корреляционно-регрессионный, факторный или кластерный анализ. *«Полуспециализированными» и «полууниверсальными»* можно считать российские пакеты STADIA, ОЛИМП и белорусский пакет РОСТАН.

К этому же классу следует отнести и американские пакеты ODA, WinSTAT, Statit, UNISTAT, Multivariance 7, JMP, SOLO, STATlab. К *специализированным пакетам по классификации и снижению размерности* можно отнести такие отечественные системы, как КЛАСС-МАСТЕР, КВАЗАР, PALMODA, Stat-Media, STARC, а также ряд зарубежных пакетов, например MVSP.

Широко известны пакеты, решающие смежные с классификацией задачи: американские системы BMDP/W, SigmaStat, Statistix, TURBO Spring-Stat-Win, а также отечественный пакет «Статистик-Консультант для Windows». Кроме того, на рынке имеются *статистические экспертные системы*, например СТАТЭКС, Statistical Navigator Pro. Среди нестатистических пакетов, решающих задачи классификации, можно отметить пакеты PolyAnalyst, ДА-система, АРГОНАВТ, ЛОРЕГ, пакет ОТЭКС и разнообразные нейросетевые пакеты.

## 4.2. Системы поддержки в процессах принятия решений

Кроме этих систем, можно использовать инструменты поддержки принятия решений. При выборе альтернативы решения могут использоваться три подхода: исследование и анализ, прошлый опыт и проведение эксперимента. Подход к принятию решения выбирается в соответствии с имеющимися ресурсами и поставленной задачей, каждый из подходов имеет свой инструментарий (см. таблицу 4.2.). Подходы могут применяться одновременно и дополнять друг друга.

Информационные системы характеризуются рядом параметров: архитектура (например двух-, трехуровневая, сервис ориентированная), режим работы (в режиме реального времени или нет), платформа, количество пользователей, защищенность, технические требования, используемые методы, проблемная и предметная область, вид интерфейса, уровень автоматизации. и т.д.

Таблица.4.2. Примерное распределение систем по подходам к принятию решения.

	Исследование и анализ	Прошлый опыт	Проведение эксперимента
Условия использования	наличие достаточного количества статистической информации	наличие экспертов в заданных проблемной и предметной областях	возможность построения физической, математической, имитационной или другой модели
	наличие данных о решении аналогичных задач и /или подзадач		
Инструментарий	информационно-справочные системы	экспертные системы	системы моделирования
	хранилища данных		
	системы поддержки принятия решений	базы прецедентов	расчетно-вычислительные системы

При выборе инструмента важно знать на основе каких *методов* работает система, т.к. каждый из методов имеет свою область применения, ограничения, требования к исходным данным, погрешность. Методы,

используемые в системах оказания помощи разнообразны, их можно разделить на три группы:

- методы искусственного интеллекта (нейронные сети, генетические алгоритмы, нечеткая логика, модели представления знаний и т.д.);
- традиционные (математические) методы принятия решения (линейное, нелинейное, динамическое программирование, методы решения экстремальных задач и т.д.);
- методы, относящиеся к предметной области (физические законы термодинамики, кинематики, электричества и т.д.).

*Проблемная область* определяет задачи, которые способна решать система. Задачи могут быть следующие: распределение ресурсов, планирование, оптимизация, управление и т.д. Проблемная область и методы тесно связаны, т.к. каждый метод имеет свою сферу применения. В теории принятия решения эта связь явно прослеживается (см. таблицу 4.3.).

*Таблица 4.3. Классы и методы решения задач теории принятия решений.*

<b>Классы задач</b>	<b>Методы решения</b>
Поисковые	Нелинейное программирование
Распределенные	Линейное программирование
Управление запасами	Теория управления запасами
Массовое обслуживание	Теория массового обслуживания
Календарное планирование	Теория расписания
Состязательные задачи	Теория игр

Задачи систем разнообразны, наиболее распространенными являются:

- интерпретация данных;
- диагностика;
- мониторинг;
- проектирование;
- прогнозирование;
- планирование;
- обучение;
- управление.

Схематично взаимосвязь этих задач можно представить так (рис. 4.1). Интерпретация данных требует отнесения информации к тому или иному классу, эти классы могут быть заданы заранее (задача классификации) или определяться на основе поступивших данных (задача кластеризации).

Основная задача мониторинга — непрерывная интерпретация данных в реальном масштабе времени и сигнализация о выходе тех или иных параметров за допустимые пределы. Планирование, прогнозирование, обучение невозможны без предварительного накопления информации и ее анализа, поэтому задача мониторинга является их составляющей.

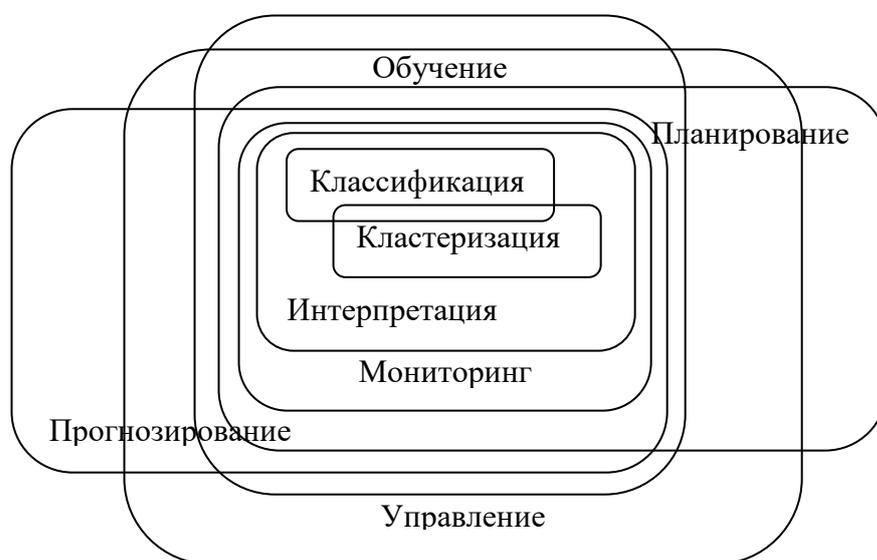


Рис. 4.1. Схема отношения задач.

Управление должно предваряться этапами сбора и анализа информации, разработки плана организации воздействий на систему, с учетом предыдущего опыта и прогноза поведения системы, поэтому задача управления пересекается с остальными задачами.

Именно поэтому разные системы ориентированные на различные задачи используют одни и те же методы, что доказывает их общность. Например, применение нейронных сетей и нечеткой логики достаточно разнообразно (см. таблицу 4.4).

Таблица 4.4. Примеры использования нейросетей и нечеткой логики в различных системах

Вид систем	Нейронные сети	Нечеткие системы
------------	----------------	------------------

<i>Поддержка принятия решений</i>	Пакет BrainMaker для задач прогнозирования, оптимизации и моделирования ситуаций.	Пакет FuziCalc для обработки табличных данных.
<i>Экспертные и гибридные системы</i>	Гибридная экспертная система RAISON.	Нечеткая оболочка ЭС FuzzyClips.
<i>Моделирование</i>	Любая искусственная нейронная сеть является примитивной моделью естественных нейросетей.	Любая нечеткая лингвистическая переменная, нечеткая база знания являются моделью представлений эксперта о реальном мире.
<i>Информационно-справочные системы</i>	Семантические нейронные сети применяют в web-системах для описания ресурсов и их поиска.	Советующие и рекомендательные системы на базе нечеткой логики (хотя они больше относятся к экспертным системам).
	Нейросетевой продукт SelectCast позволял определять область интересов пользователей Интернета и предлагал им рекламу соответствующей тематики	
<i>Автоматизированные системы</i>	Автопилотируемый гиперзвуковой самолет-разведчик LoFLYTE.	Система управления транспортным средством на базе пакета CubiCalc
	Продукт Falcon выявляет в реальном времени подозрительные сделки по краденым кредитным и дебетным картам.	Нечеткая система управления пригородными поездами японского города Сендай.

Задачи могут решаться несколькими способами, и методы применяются для решения ряда доступных им задач.

*Предметная область* определяет сферу жизнедеятельности, в которой будут решаться задачи проблемной области: экономика, медицина, промышленность, бизнес и т.д. Если пользователь может использовать систему в нескольких предметных областях и/или предметная область не зафиксирована в алгоритмах системы, то такая система считается универсальной.

Методы, используемые в системах, зависят от конкретной реализации и могут быть разнообразны, так же как и интерфейс может быть различным у одного класса систем. Но другие атрибуты типичны для всего класса систем

(см. таблицу 4.5), и именно их нужно в первую очередь учитывать при выборе класса системы.

Каждый класс систем имеет свою классификацию, свои особенности и уже их нужно учитывать при дальнейшем выборе конкретной системы. Каждая из систем имеет свои плюсы и минусы, поэтому кроме при выборе любой информационной системы нужно учитывать и такие свойства как:

- гибкость;
- масштабируемость;
- цена системы;
- модификационные возможности (связанные с конкретными потребностями и пожеланиями предприятий);
- открытость системы;
- наличие у компании-производителя успешных внедрений в фирмах подобного плана и имиджа, располагающего к доверию;
- соотношение качества продукта, функциональности и его стоимости и т.д.

Таблица 4.5. Обобщенные значения атрибутов

	<b>Автоматизированные</b>	<b>Экспертные</b>	<b>СППР</b>	<b>Моделирования</b>
<i>Проблемная область</i>	Ориентированы на конкретную задачу	Поддержка решений	принятия	Оптимизация, сбор статистической информации, управление, идентификация
<i>Предметная область</i>	Строго определена и связана с конкретной реализацией системы	Может быть универсальной (настраиваемой под область) или задана для конкретной реализации		
<i>Уровень автоматизации</i>	Высокий (может заменить ЛПР)		Средний (носит рекомендательный характер)	Низкий (не генерируют решений)

Сейчас на рынке информационных технологий представлено множество систем (примеры систем см. таблица 4.6), каждая из которых имеет свои особенности уже в рамках своего класса систем.

У рассматриваемых систем много общего, но все же это разные классы систем и заменить друг друга они не могут. Но они дополняют друг друга. Основная тенденция современного развития информационных технологий - это интеграция.

Таблица 4.6. Примеры систем оказания помощи

<b>Автоматизированные</b>	<b>Экспертные</b>	<b>СППР</b>	<b>Моделирования</b>
АСУ конвейерным транспортом склада заполнителей бетонного хозяйства	<i>В военном деле</i> ACES ASTA DART HANNIBAL. I&W RUBRIC.	Система поддержки принятия управленческих решений "КУРС"	<i>Программы с поточной моделью управления</i>  Jigrein VisSim MBTY Simulink Easy5
ГАСНТИ (Государственная автоматизированная система научно-технической информации)	<i>В электронике</i>  ACE.	СВОSSdss	<i>Интерпретаторы систем уравнений</i>  Dynast
АСУ «Морфлот»	<i>В информатике</i>  CODES	DSS (EPAM Systems)	20-sim Dymola Simplorer
Аппаратура системы управления насосной станции водоотлива рудника (шахты)	<i>В компьютерных системах</i>  MIXER		ITI-sim Pspice Multisim Micro-Cap

Различные системы интегрируют между собой, и рассматривать в таком случае нужно не отдельную систему, а весь информационный комплекс. Но общность методов и средств позволяет не только интеграцию, но и гибридизацию систем. При этом можно использовать уже готовые и проверенные методы и системы, дополняя и модифицируя их. Именно на это должны обращать внимание специалисты (разработчик, администратор, специалист по развертыванию системы и т.д.), составляя информационный комплекс для своего научного исследования.



## Используемая литература

1. A Conversation with Michael Stonebraker and Margo Seltzer [Text] // ACM Queue.2007. - Volume 5, № 4. May/June.
2. Franklin, M. From Databases to Dataspaces: A New Abstraction for Information Management [Text] / M. Franklin, A. Halevy and D. Maier // SIGMOD Record. – 2005. - Vol. 34, № 4, Dec.
3. Gray, J. Data Management: Past, Present, and Future [Text]/ J. Gray // IEEE Computer. – 1996. - Vol. 29, № 10, October.
4. Gray, J. Scientific Data Management in the Coming Decade [Text] / Jim Gray, David T. Liu, Maria Nieto-Santisteban, Alex Szalay, David J. DeWitt, Gerd Heber // SIGMOD Record. -2005. - Vol. 34, №. 4, Dec.
5. Stonebraker, M. One Size Fits All? – Part 2: Benchmarking Results [Text]/ Michael Stonebraker, Chuck Bear, Uğur Çetintemel, Mitch Cherniack, Tingjian Ge, Nabil Hachem, Stavros Harizopoulos, John Lifter, Jennie Rogers, and Stan Zdonik // Proceedings of the 3rd Biennial Conference on Innovative Data Systems Research (CIDR). – 2007.- January 7-10, Asilomar, California, USA.
6. Аврамчук, Е.Ф. Технология системного моделирования [Текст] / Под общей редакцией С.В. Емельянова. — М.: Машиностроение; Берлин: Техник, 1988. — 520 с.
7. Алабужем, П.М. Теория подобия и размерностей. Моделирование. [Текст] / П.М. Алабужем, В.Б. Геронимус и др.— М.: Высш. шк., 1968. — 320 с.
8. Бглуха, М.Т. Основи наукових досліджень. [Текст] / М.Т. Бглуха — К.: Вища шк., 1997. — 271 с.

9. Блехман, И.И. Прикладная математика: предмет, логика, особенности подхода. [Текст] / И.И. Блехман, А.Д. Мышкинс, А.Г. Пановка — К.: Наук, думка, 1976. — 270 с.
10. Бондаренко, А.В. Практические советы по написанию диссертации: Учебно-методическое пособие для подготовки аспирантов и соискателей ученых степеней кандидата и доктора наук [Текст] / Под ред. А.В. Бондаренко. — Уфа: УГНТУ, Нефтегазовое дело, 2011. — 24 с.
11. Буч, Г. Язык UML. Специальный справочник [Текст] / Г. Буч, Д. Рамбо, А. Джекобсон. Пер. с англ. — М.: Питер, 2002. — 656 с.
12. Васильев, М. Объектно-ориентированный подход к моделированию информационных систем [Электронный ресурс] / М. Васильев, И. Хомков, С. Шаповаленко — Режим доступа: <http://www.termika.ru> 10.02.2011
13. Вендров, А.М. CASE-технологии. Современные методы и средства проектирования информационных систем. [Электронный ресурс] / А.М. Вендров - Режим доступа: <http://www.citforum.ru> 23.02.2012
14. Венецкий, И.Г. Теория вероятностей и математическая статистика. [Текст] / И.Г. Венецкий, Г.С. Кильдишев — М.: Статистика, 1975. — 264 с.
15. Вентцель, Е.С. Теория вероятностей. [Текст] / Е.С. Вентцель— М.: Наука, 1969. — 432 с.
16. Волков, О. Стандарты и методологии моделирования бизнес-процессов [Электронный ресурс] / О. Волков — Режим доступа: <http://www.connect.ru> 03.12.2011
17. Волков, Ю.Г. Диссертация. Подготовка, защита, оформление. [Текст]: практическое пособие. / Ю.Г. Волков — М.: Альфа-М, Инфра-М, 2011. — 176 с.
18. Волкова, Е.Н. Основы теории систем и системного анализа [Текст] / Е.Н. Волкова — СПб.: ГТУ, 2001. — 512 с.

19. Гаврилов, А.В. Гибридные интеллектуальные системы. [Электронный ресурс] / А.В. Гаврилов, Ю.В. Новицкая // Тезисы Международной конференции "Информационные системы и технологии" 2003 г. Режим доступа: [http://ermak.cs.nstu.ru/ist2003/papers/gavrilov\\_novitskaya.pdf](http://ermak.cs.nstu.ru/ist2003/papers/gavrilov_novitskaya.pdf). 12.03.2011 г.
20. Галлеев, В.И. Кухня процессного подхода [Текст] / В.И. Галлеев, К.В. Пичугин // Методы менеджмента качества, 2003. — Вып. 4.
21. Галушко, В.Г. Вероятностно-статистические методы на автотранспорте. [Текст] / В.Г. Галушко— К.: Вища шк., 1976. — 232 с.
22. Гнеденко, Б.В. Курс теории вероятностей. [Текст] / Б.В. Гнеденко— М.: Наука, 1969. — 400 с.
23. ГОСТ Р ИСО 9004-2001. Системы менеджмента качества. Рекомендации по улучшению деятельности. [Текст] – Введ. 15.08.2001. - М.: Госстандарт РФ, 2001. – 49 с.
24. Губарев, В.В. Информатика в рисунках и таблицах (Фрагменты системного путеводителя по концептуальным основам) [Текст]: учебное пособие / В.В. Губарев — Новосибирск: НГТУ, 2002. — 156 с.
25. Губарев, В.В. Концептуальные основы информатики. Часть 1. Сущностные основы информатики [Текст]: учебное пособие / В.В. Губарев — Новосибирск: НГТУ, 2005. — 149 с.
26. Доблаев, Л.П. Психологические основы работы над книгой. [Текст] / Л.П. Доблаев — М.: Книга, 1970. — 72 с.
27. Драйпер, Н. Прикладной регрессионный анализ. [Текст] / Н. Драйпер, Г. Смит— М.: Статистика, 1973. — 210 с.
28. Ермаков, С.М. Курс статистического моделирования. [Текст] / С.М. Ермаков, Г.Л. Михайлов— М.: Наука, 1976. — 168 с.
29. Зельдович, Я.Б. Элементы прикладной математики. [Текст] / Я.Б. Зельдович, А. Д. Мышкин — М.: Наука, 1965.

30. Кальченко, Д. Нейронные сети: на пороге будущего. [Электронный ресурс] / Д. Кальченко // КомпьютерПресс, № 1, 2005 г. Режим доступа: <http://alt-future.narod.ru/Ai/neiro.htm>. 12.03.2011 г.
31. Капица, П. Л. Эксперимент, теория, практика. [Текст] / П.Л. Капица— М.: Наука, 1977. — 420 с.
32. Кириллов, В.В. Основы проектирования реляционных баз данных [Электронный ресурс] / В.В. Кириллов – М.: Центр информационных технологий - Режим доступа: <http://www.citforum.ru/database/dbguide/index.shtml>. Последнее посещение: 26.10.2011 г.
33. Крутов, В.И. Основы научных исследований [Текст]: учебник для технических вузов / Под ред. В.И. Крутова, В.В. Попова. — М.: Высшая школа, 1989. — 400 с.
34. Кузин, Ф.А. Кандидатская диссертация. Методика написания, правила оформления и порядок защиты [Текст]: практическое пособие для аспирантов и соискателей ученой степени / Ф.А. Кузин — М.: Ось-89, 1998. — 208 с.
35. Кузнецов, С. Д. Объектно-реляционные базы данных: прошедший этап или недооцененные возможности? [Текст] / С.Д. Кузнецов // Труды Института системного программирования, т. 13, часть 2, М., ИСП РАН, 2007. - С. 115-140.
36. Кумунжиев, К.В. Теория систем и системный анализ. [Текст] / К.В. Кумунжиев— Ульяновск: УлГУ, 2003. — 240 с. Часть 2. —119 с.
37. Марка, Д. Методология структурного анализа и проектирования [Текст] / Д. Марка, К. МакГоуэн. Пер. с англ . М.: — 1993, 240 с ., ISBN 5-7395-0007-9.
38. Масалович А. Этот нечеткий, нечеткий, нечеткий мир. [Электронный ресурс] / А. Масалович // Компьютерная неделя, №16, 1995 г. Режим доступа: <http://www.pcweek.ru/themes/detail.php?ID=78913>. 12.03.2011 г.
39. Масалович, А. Нейронная сеть-оружие финансиста. [Электронный ресурс] / А. Масалович // Компьютерная неделя, №01, 1995 г. Режим

- доступа: [http://www.pcweek.ru/themes/detail.php?ID=77372&THEME\\_ID.](http://www.pcweek.ru/themes/detail.php?ID=77372&THEME_ID.)  
12.03.2011 г.
40. Микешина, Л.А. Философия науки: Современная эпистемология. Научное знание в динамике культуры. Методология научного исследования [Текст]: учебное пособие / Л.А. Микешина — М.: Прогресс-Традиция: МПСИ: Флинта, 2005. — 464 с.
41. Налимов, В.В. Теория эксперимента. [Текст] / В.В. Налимов — М.: Наука, 1971. — 218с.
42. Новиков, А.М. Как работать над диссертацией. Пособие для начинающего педагога-исследователя [Текст] / А.М. Новиков — М.: Эгвес, 2003. — 104 с.
43. Ожегов, С.И. Современный словарь иностранных слов [Текст] / С.И. Ожегов, Н.Ю. Шведова.— СПб.: 1994. — 376 с.
44. Ожегов, С.И. Толковый словарь русского языка [Текст] / С.И. Ожегов, Н.Ю. Шведова. — М.:1999. — 354 с.
45. Отоцкий, Л. Семь критериев выбора ERP-систем для России. [Электронный ресурс] / Л. Отоцкий//«Открытые системы», № 4, 1998 г. Режим доступа: <http://www.osp.ru/os/1998/04/179564/>. 12.03.2011 г.
46. Перегудов, Ф.И. Введение в системный анализ [Текст] : учебное пособие для вузов / Ф.И. Перегудов, Ф.П. Тарасенко — М.: Высшая школа, 1989. — 367 с.
47. Пустыльник, Е.И. Статистические методы анализа и обработки наблюдений. [Текст] / Е.И. Пустыльник — М.: Наука, 1968. — 270 с.
48. Рачков, П.Л. Науковедение. [Текст] / П.Л.Рачков — М.: Изд-во Моск. ун-та, 1974. — 242 с.
49. Румшинский, Л.З. Математическая обработка результатов эксперимента. [Текст] / Л.З. Румшинский — М.: Наука, 1971. — 192 с.
50. Сабитов, Р.А. Основы научных исследований. [Текст]: учебное пособие / Р.А. Сабитов — Челябинск: ЧГУ, 2002. — 138 с.

51. Сайт компании «ERAM Systems». [Электронный ресурс] Режим доступа: <http://www.eram-group.ru/> 12.03.2011 г.
52. Сайт компании «Курс». [Электронный ресурс]. Режим доступа: <http://www.curs.ru/> 12.03.2011 г.
53. Сайт фирма «СиБОСС». [Электронный ресурс]. Режим доступа: <http://www.cboss.ru/main/index.html> 12.03.2011 г.
54. Семушин, И.В. Общие правила разработки презентации и макет для ее изготовления в системе LATEX2 [Электронный ресурс] / И.В. Семушин. Режим доступа: <http://www.staff.ulsu.ru/semoushin> 12.05.2012 г.
55. Сиденко, В.М. Основы научных исследований. [Текст] / В.М. Сиденко, И.М. Грушко — Харьков: Вища шк., 1979. — 200 с.
56. Чекал, Е.Г. Методика исследования и моделирования информационного центра системы наблюдения надводного судна [Текст] / Е.Г. Чекал, А.А. Смагин, А.А. Чичев // Ученые записки УлГУ, серия "Информационные технологии", 2007. — Вып.1. — Ульяновск: УлГУ. — С.101-106
57. Чкалова, О.Н. Основы научных исследований. [Текст] / О.Н. Чкалова — К.: Вища шк.
58. Яковлев, А. Хранилища данных - для всех и повсюду [Текст] / Александр Яковлев // Computerworld Россия. — 2008. - № 39. — ISSN 1560-5213.
59. Ярская, В.Н. Методология диссертационного исследования: Методическое пособие. [Текст] / В.Н. Ярская — Саратов: ПМУЦ, 2002. — 189 с.