

В. Н. ВОЛКОВА

*Моему учителю, другу и мужу
Денисову Анатолию Алексеевичу
посвящаю*

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ СИСТЕМ

**Санкт-Петербург
Издательство Политехнического университета
2012**

УДК 002.531.5.681.3
ББК 32.97

Рецензенты:

Доктор технических наук профессор Российского государственного университета инновационных технологий и предпринимательства
Л.С. Болотова

Доктор экономических наук, профессор,
заведующий кафедрой «Маркетинг и информационные технологии в экономике» ОАНО АПО «Волжский университет им. В.Н. Татищева»
А.И. Афоничкин

Волкова В.Н. Теоретические основы информационных систем / В.Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2012. – 280 с.

Излагаются теоретические основы информатики и информационного поиска, лежащие в основе создания и функционирования информационных систем различного вида и назначения. Рассматриваются особенности разработки автоматизированных информационных систем фактографического типа как первой стадии автоматизированных систем управления предприятиями и организациями, особенности документальных и документально-фактографических информационно-поисковых систем; структуры и принципы функционирования информационных систем специального назначения. Дается представление о разработке методик организации проектирования и сравнительного анализа готовых программных продуктов при создании ИС для конкретной организации.

Проводится анализ состояния теории информационных систем как научного направления.

© Волкова В.Н., 2012.

© Санкт-Петербургский государственный
политехнический университет, 2012.

ISBN 978-5-7422-3478-4

ВВЕДЕНИЕ

Возрастающая потребность в информации как важнейшем ресурсе современного общества требует от специалистов различного профиля знаний в области информационных систем (ИС) и информационных технологий (ИТ), умения разрабатывать или выбирать готовые программные продукты для информационного обеспечения деятельности предприятий и организаций.

Подготовка специалистов по информационным системам ведется в вузах технического и экономического профиля уже более 30 лет. Существуют несколько Федеральных государственных образовательных стандартов высшего профессионального образования по направлениям и специальностям, связанным с разработкой и эксплуатацией информационных систем, определяющих требования к содержанию и уровню подготовки специалистов, бакалавров и магистров.

Имеются монографии и учебные пособия по различным направлениям информатики. В то же время нет единой трактовки понятия «информационная система», не разработана теория информационных систем как научного направления.

Становление любой науки проходит 4 стадии:

- Обобщение, систематизация и классификация накапливаемых знаний. Определение места науки среди других научных направлений.
- Углубленное изучение основных объектов и процессов предметной области научного направления (структуризация, морфологическое, логико-лингвистическое, семиотическое отображение и исследование).
- Формирование научной методологии и методик исследования научного направления, стремление использовать методы точных наук, применение формальных методов и моделей.
- Развитие теоретического компонента – выработка понятий, категорий, методологических установок, создание теоретических концепций, формирование категориального аппарата научного направления.

Ц е л ь монографии – обобщить сведения об информационных системах; об основных особенностях и принципах построения и проектирования ИС различного вида и назначения; изложить теоретические основы разработки структуры информационной системы для предприятий (организаций), проанализировать перспективы развития теории информационных систем как научного направления.

В соответствии с этой целью излагаются некоторые сведения из теории информатики (гл. 1); дается понятие об информации и инфор-

мационных ресурсах систем (гл. 2); кратко рассматривается история разработки автоматизированных информационных систем различного вида и назначения, предлагается классификация ИС (гл. 3); рассматриваются особенности и теоретические основы разработки автоматизированных информационных систем (АИС) и автоматизированных систем управления (АСУ) предприятий (организаций), включая принципы формирования структур функциональной и обеспечивающей части, принципы выбора готовых программных продуктов для автоматизации деятельности предприятий и организаций (гл. 4); приводятся основные сведения о теории информационного поиска и информационно-поисковых систем (ИПС), рассматриваются принципы функционирования и структура документальных информационно-поисковых систем (ДИПС), включая понятия об информационно-поисковом языке (ИПЯ), классификациях ИПС, системе индексирования (СИ), логике ИПС, критериях смыслового соответствия (КСС), критериях оценки ИПС, релевантности информационного поиска (гл. 5); дается представление о документально-фактографических информационно-поисковых системах (ДФИПС), о системе нормативно-методического обеспечения управления (СНМОУ) предприятия (организации) и ее автоматизированном варианте (АСНМОУ) как информационной системе документально-фактографического типа (гл. 6).

Рассматриваются перспективы развития теории информационных систем на основе исследования закономерностей информетрии (гл. 7), концепций архитектуры информационной системы; ситуационных центров, концепции многоуровневых информационных систем (гл. 8).

В заключении проводится анализ состояния теории информационных систем как научного направления.

Выражаю благодарность рецензентам доктору технических наук профессору *Людмиле Сергеевне Болотовой* и доктору экономических наук, профессору *Александру Ивановичу Афоничкину* за полезные советы при подготовке рукописи; и особенно моему учителю, другу и мужу *Анатолию Алексеевичу Денисову*, который помог мне понять, что информация как структура материи является основой любых преобразований, и развитие организации зависит от того, как ей удастся создать свою информационную систему, информационную инфраструктуру.

В.Н. Волкова

Глава 1. ПОНЯТИЕ ОБ ИНФОРМАТИКЕ КАК НАУЧНОМ НАПРАВЛЕНИИ

1.1. История развития термина «информатика»

Термин «**информатика**» использовался первоначально в зарубежных научных публикациях (нем. *Informatik*, англ. *Information science*, фр. *Informatique*) для названия научно-практического направления, занимающегося определением способов получения, накопления, хранения, преобразования, передачи и использования информации.

Этот термин впервые был введен в 1957 г. в Германии **К. Штейнбухом** (K. Steinbuch)¹. Затем – в 1962 г. во Франции **Ф. Дрейфузом** (F. Dreyfus)², который предложил его как слияние французских слов *information* и *automatique* для названия области, занимающейся автоматизированной переработкой информации. Ф. Дрейфус также сделал переводы термина на ряд других европейских языков.

В отечественной науке термин «информатика» был первоначально использован в 1963 г. профессором Московского энергетического института **Ф.Е. Темниковым**³, который определил информатику как науку об информации вообще, состоящую из 3-х основных частей:

- 1) теория информационных элементов;
- 2) теория информационных процессов;
- 3) теория информационных систем.

В таком значении термин использовался в начальный период создания автоматизированных информационных систем в технической литературе. Однако в последующем это определение долгое время оставалось лишь историческим фактом, и не было оценено должным образом. Возможно, потому что было опубликовано только в специальном журнале «Известия вузов: Электромеханика».

¹ **Steinbuch K.** («Informatik: Automatische Informationsverarbeitung». *SEG-Nachrichten (Technische Mitteilungen der Standard Elektrik Gruppe)*) / **K. Steinbuch.** – *Firmenzeitschrift*, 1957.

² ru.wikipedia.org/wiki/Информатика.

³ **Темников, Ф. Е.** Информатика / Ф. Е. Темников // Известия ВУЗов. Электромеханика, 1963. – № 11.

В 1966 г. независимо термин «информатика» был введен **А.И. Михайловым** вместо термина «научная информация».

Информатикой была названа «... научная дисциплина, изучающая структуру (не конкретное содержание) и свойства научной информации, а также закономерности научно-информационной деятельности, ее теорию, историю, методологию и организацию»¹.

В таком значении в советской научно-технической литературе термин «информатика» стал широко известен, благодаря работам **А.И. Михайлова, А.И. Черного и Р.С. Гиляревского** использовался на протяжении достаточно длительного времени в сфере бурно развивающегося в тот период направления «Научно-техническая информация» [15, 16, 27 и др.].

В последующем термин «информатика» стал использоваться в более узком смысле, применительно, в основном, к техническим и программным средствам хранения и обработки информации на электронно-вычислительных машинах и стал соответствовать англ. *Computer science* – компьютерная наука – в США, англ. *computing science* – вычислительная наука – в Великобритании.

В 1978 г. Международным конгрессом в Японии было принято следующее определение: «*Понятие информатики охватывает области, связанные с разработкой, созданием, использованием и материально-техническим обслуживанием системы обработки информации, включая машины, оборудование, математическое обеспечение, организационные аспекты, а также комплекс промышленного, коммерческого, административного, социального и политического воздействия*»².

В английском толковом словаре по вычислительной технике Dictionary of Computing, изданном в 1983 г., «*Информатика – это наука и технология обеспечения информационного обмена с помощью систем, основанных на применении ЭВМ*»³.

Академик **А.А. Дородницын**⁴ определил информатику как науку о преобразовании информации, которая базируется на вычислительной технике.

¹ **Михайлов А.И.** Информатика – новое название теории научной информации / А.И. Михайлов, А.И. Черный, Р.С. Гиляревский // Научно-техническая информация. – М., 1956. – № 12. – С. 3–8.

² **Беликов Е.П.** Об организации в Академии наук СССР работ по информатике, вычислительной технике и автоматизации / Е.П. Беликов // Вестник АН СССР. – 1983. – № 6.

³ **Dictionary of Computing.** Data Communications, hardware and Software. Basics Digital Electronics / John Wiley. – 1983.

⁴ **Дородницын А.А.** Информатика: предмет и задачи / А.А. Дородницын // Кибернетика. Становление информатики: сб. – М.: Наука, 1986.

В 1988г. академик **А.П. Еришов** в Математическом энциклопедическом словаре ¹ дал следующее определение информатики как науки, отрасли промышленности и разновидности человеческой деятельности: *«Информатика – 1) находящаяся в становлении наука, изучающая законы и методы накопления, передачи и обработки информации с помощью ЭВМ; 2) родовое понятие, охватывающее все виды человеческой деятельности, связанные с применением ЭВМ».*

В таком понимании информатика включает дисциплины, относящиеся к обработке информации в вычислительных машинах и вычислительных сетях: как абстрактные, вроде анализа алгоритмов, так и довольно конкретные, например, разработка языков программирования.

Это оказалось удобным как краткое название курса о применении ЭВМ для обработки данных в учебном процессе школы, поскольку школьникам сложно объяснить понятия информационных процессов, информационных систем, поиск, хранение и обработку текстовой информации.

Однако такое сужение смысла термина «информатика» нежелательно в учебном процессе вуза, поскольку уже достаточно давно осознано, что информация – важнейший ресурс социально-экономических организаций, обеспечивающий их развитие, и нужно уметь оценивать содержание, смысл информации как интеллектуального ресурса.

Это осознавалось постепенно.

В 1985 г. академик **А.А. Самарский** обращает внимание на новую научную методологию, возникшую благодаря информатике: «Она основана на развитии в широком применении методов математического моделирования и вычислительного эксперимента и служит ближайшим стратегическим резервом ускорения научно-технического прогресса. Сущность математического моделирования и его главное преимущество состоит в замене исходного объекта соответствующей математической моделью и в дальнейшем ее изучении (экспериментирование с ней) на ЭВМ с помощью вычислительно-логических алгоритмов. Математическое моделирование представляет собой естественное развитие и обобщение методов научного исследования, соединенных с современной информационной технологией. Цикл вычислительного эксперимента объект – модель – алгоритм – программа – ЭВМ – управление объектом отражает основные этапы процесса познания в нынешнем компьютерном воплощении. Здесь органично соединяются сильные стороны теоретических методов и натурального эксперимента. Работа с моделью, а не с объектом, оборачивается оперативным получением подробной и наглядной информации, вскрывающей его внутренние связи, качественные характеристики и количественные параметры. Многократно уменьшаются материальные и трудовые затраты, прису-

¹ *Математический* энциклопедический словарь / Гл. ред. Ю.В. Прохоров.– М.: Сов. энциклопедия, 1988.

щие традиционным экспериментальным подходам, дающим, как правило, лишь крупницы нужной информации. Вычислительный эксперимент не подвластен каким-либо ограничениям – математическая модель может быть безопасно испытана в любых мыслимых и немыслимых условиях»¹.

Академик **Н.Н. Моисеев** считает, что «информатика – это некая синтетическая дисциплина, которая включает в себя и разработку новой технологии научных исследований и проектирование, основанные на использовании электронной вычислительной техники, и несколько крупных научных дисциплин, связанных с проблемой общения с машиной, и, наконец, с созданием машины»².

В ряде определений информатики на первое место ставили собственно «информацию».

Директор института проблем передачи информации **В.И. Сифоров** считал, что: «В основу определения должны быть положены действия над информацией. Информатика развивается под действием потребности общества и согласно внутренней логике развития. В основе этого развития лежат закономерности процессов в ЭВМ, закономерности развития ЭВМ. Информатика имеет дело не с конкретными формами материи, а с категориями: информация, модель и т. п. Информатика – комплексная дисциплина – это наука (фундаментальные исследования) и отрасль производства (опытно-конструкторские работы и совершенствование технологий), а кроме того, и инфраструктурная область (эксплуатация информационных систем)»³.

Заместитель директора Всесоюзного института научной и технической информации в 1970-е гг. **Ю.И. Шемакин** в книге «Введение в информатику» отмечает, что «основной задачей информатики является изучение закономерностей, в соответствии с которыми происходят создание, преобразование, хранение, передача и использование информации всех видов, в том числе с применением современных технических средств»⁴.

Академик **Б.Н. Наумов** в предисловии к сборнику⁵ подчеркивает, что информатика – это «естественная наука, изучающая общие свойства информации, процессы, методы и средства ее автоматизированной обработки». При этом под обработкой информации понимаются процессы ее восприятия, хранения, преобразования, перемещения и вывода (ввода) с применением средств вычислительной техники».

¹ Самарский А.А. Проблема использования вычислительной техники и развитие информатики / А.А. Самарский // Вестник АН СССР, 1985. – № 3.

² Моисеев Н.Н. Информатика: новые пути познания законов природы и общества / Н.Н. Моисеев // Вестник АН СССР 1985. – № 5.

³ Сифоров В.И. Информатика и ее взаимодействие с философией и другими науками / В.И. Сифоров // Философская наука, 1984. – № 2.

⁴ Шемакин Ю.И. Введение в информатику / Ю.И. Шемакин. – М.: Финансы и статистика, 1985.

⁵ Информатика и компьютерная грамотность. ИПИ АН СССР / Отв. ред. академик Б.Н. Наумов. – М.: Наука, 1988.

В документах ЮНЕСКО 1986–1988 гг. термину «информатика» дается широкое толкование ¹.

Указывается, что этот термин охватывает собственно информацию, ее сбор, анализ и обработку, а также соответствующие аппаратные средства, включая микропроцессоры как таковые или же в сочетании с другими электронными системами. Информатика рассматривается как крупное научное направление, заслуживающее активного развития в интересах всего человечества. Она способна (при соответствующем освоении ее методов и средств) помочь человеку полнее использовать информационные ресурсы в интересах научно-технического прогресса и социального развития.

В последнее время некоторыми авторами, в связи с определением современных задач информатики, особый упор делается на обработку знаний.

Роль теории искусственного интеллекта в информатике была обоснована академиком **Г.С.Поспеловым** ².

В.Д. Ильин предложил дать следующее определение информатики: «Предметом информатики как науки будем считать процесс создания, накопления и применения знаний» ³.

К.К. Колин дал следующее определение: «Информатика является общенаучной дисциплиной, которая изучает свойства, закономерности, процессы, методы и средства формирования, хранения и распространения знаний в природе и обществе» ⁴, и в 2000 г. издал учебное пособие ⁵, в котором развивает представление об информатике как междисциплинарной науке о закономерностях и формах движения информации в природе и обществе.

В ряде работ ⁶ на основе сопоставления различных определений информатики предлагается рассматривать ее как науку о формализованном общении.

В связи с бурным развитием компьютеризации в нашей стране некоторые авторы считают, что под информатикой следует понимать профессиональную деятельность по применению и разработке ЭВМ ⁷.

¹ **Свириденко С.С.** Современные информационные технологии / С.С. Свириденко. – М.: Радио и связь, 1989.

² **Поспелов Г.С.** Искусственный интеллект – новая информационная технология / Г.С. Поспелов. – М.: Наука, 1988.

³ **Ильин В.Д.** Система порождения программ / В.Д. Ильин. – М.: Наука, 1989.

⁴ **Колин К.К.** О структуре научных исследований по комплексной программе «Информатика» / К.К. Колин // Науч. труды «Социальная информатика»: сб. – М.: КВШ при ЦК ВЛКСМ, 1990.

⁵ **Колин К.К.** Фундаментальные основы информатики: Социальная информатика / К.К. Колин. – М.: Академический проект; Екатеринбург: Деловая книга, 2000. – 350 с.

⁶ **Белошапка В.О.** О языках моделях и информатике / В.О. Белошапка // Информатика и образование. – 1987. – № 6.

⁷ **Дименштейн Р.П.** Информатика или компьютерное дело / Р.П. Дименштейн, А.Г. Яковлев // Информатика и образование. – 1989. – № 3. Авторское добавление: Компьютер: сб. – Вып. 1. – М.: Финансы и статистика, 1990.

По мнению авторов, в этом случае открывается возможность урегулировать спор со специалистами, которые давно занимаются «старой» информатикой и трактуют ее как науку по организации распространения *научно-технической информации* (именно эту тематику отражает РЖ¹ ВИНТИ «Информатика»).

При этом оказывается, что «старой» информатике соответствует англо-американский термин «information science»². А русскоязычный термин «информатика» объединил в себе (кроме всех прочих значений) наименование двух существенно различных дисциплин «computer science» и «information science». Термин «information science» уместно переводить как «информационное дело», наука о собственно информации. Это позволит, с одной стороны, и в русском языке развести термины «information science» и «computer science», а с другой – отделять науку информатику как «computer science» от науки о собственно информации «information science».

В некоторых более кратких определениях информатика трактуется как особая наука о законах и методах получения и измерения, накопления и хранения, переработки и передачи информации с применением математических и технических средств.

Бытует и такое, самое краткое определение: информатика – это *информация* плюс *автоматика* что сужает представление о современной информатике.

На основе обобщения различных точек зрения в современном представлении можно дать следующее обобщающее определение: *информатика – научная дисциплина, изучающая структуру и общие свойства информации, закономерности процессов обмена информацией от непосредственного устного и письменного общения специалистов до формальных процессов обмена посредством различных носителей информации. Важной сферой информатики является научно-информационная деятельность по сбору, переработке, хранению, поиску и распространению научно-технической информации.*

В ряде работ принята следующая структуризация:

- *теоретическая* информатика,
- *техническая* информатика;
- *прикладная* информатика.

Следует обратить внимание на тот факт, что все имеющиеся определения отражают наличие двух главных составляющих информатики:

- *информация* и
- *средства ее обработки.*

¹ РЖ – реферативный журнал.

² *Dictionary of Information Technology*. MacMillan Press, London, 1982.

Очевидно, что первоочередная функция информатики состоит в разработке методов и средств преобразования информации с использованием компьютера, а также в применении их при организации технологического процесса преобразования информации.

В то же время, выполняя эту функцию, прикладная информатика должна решать более широкий спектр задач:

- исследовать информационные процессы в технических и социально-экономических системах;
- разрабатывать и/или адаптировать технические средства и создавать новые технологии для преобразования информации на основе результатов, полученных в ходе исследования информационных процессов;
- решать научные и инженерные проблемы создания, внедрения и обеспечения эффективного использования компьютерной техники и технологии во всех сферах человеческой деятельности.

При этом информатика исследует разнородные группы вопросов:

- *технические*, связанные с изучением методов и средств надежного сбора, хранения, передачи, обработки и выдачи информации;
- *синтаксические*, связанные с решением задач по формализации и автоматизации некоторых видов научно-информационной деятельности, в частности индексирование, автоматическое реферирование, машинный перевод;
- *семантические*, определяющие способы описания смысла информации, изучающие языки ее описания;
- *прагматические*, т.е. методы анализа и преобразования информации для решения конкретных прикладных задач.

С учетом осознания широкого спектра проблем, включаемых в настоящее время в направление, обобщаемое термином «информатика», ряд ученых¹ считает, что именно определение **Ф.Е. Темникова** следовало бы возродить и принять в качестве обобщающего определения науки об информации как науки об *информационных элементах, информационных процессах и информационных системах*.

Такое определение особенно важно для подготовки бакалавров и магистров по направлениям «Системный анализ и управление», «Прикладная информатика», «Бизнес-информатика», которые ориентированы на необходимость анализа смысла, содержания информации, разработки и выбора соответствующих методов и средств исследования информационных процессов сбора, хранения, обработки и представления информации применительно к конкретным областям и ситуациям принятия решений.

¹ Например, **Юсупов Р.М.** Информатика в системе научного знания XX и XXI веков / Р.М. Юсупов, Б.В. Соколов // Кибернетика и информатика: сб. науч. трудов: – СПб.: Изд-во Политехн. ун-та, 2006. – С. 6–21.

1.2. Информатика – наука об информационных элементах, информационных процессах и информационных системах

Идеи современной информатики возникали в недрах математики и кибернетики, электроники и системотехники, логики и лингвистики. Основные научные направления информатики образуют такие дисциплины как «Теоретические основы вычислительной техники», «Теория информации», «Теория вычислительного эксперимента», «Алгоритмика и программирование», «Искусственный интеллект» и т. п.

Прикладная информатика обслуживает науку, технику, производство и другие виды человеческой деятельности на основе информационных технологий.

Расширение сферы научных методов и практических приложений информатики привело к необходимости ее структуризации. По поводу структуры информатики существуют различные точки зрения.

В 1991 г. научные сотрудники Института проблем информатики АН СССР *И.А. Мизин* и др. рассматривали информатику, развивая приведенное выше определение *В.И. Сифорова*, как комплексную дисциплину¹:

во-первых, это *естественная наука* (фундаментальные и прикладные исследования);

во-вторых, *отрасль промышленности* (опытно-конструкторские работы и производство);

в-третьих, *инфраструктурная область* (профессиональная деятельность и эксплуатация систем информатизации).

Как *естественная наука* информатика изучает общие свойства информации (данных и знаний), методы и системы для ее создания, накопления, обработки, хранения, передачи и распределения с помощью средств вычислительной техники и связи.

Как *отрасль промышленности* информатика занимается проектированием, изготовлением, сбытом и развитием систем информатизации и их компонентов.

Как *инфраструктурная область* информатика занимается сервисом и эксплуатацией систем информатизации, обучением и др.

Как фундаментальная наука информатика связана с философией – через учение об информации и теорию познания; с математикой – через теорию математического моделирования; математическую логику и теорию алгоритмов; с лингвистикой – через учение о знаковых системах (семиотику) и формальных языках. Она также тесно связана с теорией информации и управления.

Важнейшими методологическими принципами информатики являются: изучение объектов и явлений окружающего мира с точки зрения процессов сбора обработки и выдачи информации о них, анализ определенного сходства

¹ *Развитие* определений «информатик» и «информационные технологии»: Препринт Института проблем информатики АН СССР // И.А. Мизин, И.Н. Синицын, Б.Г. Доступов, В.Н. Захаров, А.Н. Красавин; под ред. чл.-корреспондента АН СССР И. А. Мизина. – М.: ИПИ АН СССР, 1991.

этих процессов при их реализации в искусственных и естественных (в том числе в биологических и социальных) системах.

Подобным образом, но в несколько другой терминологии структурируется информатика в учебнике *Н.В. Макаровой* и *В.Б. Волкова* [12] (рис. 1.1), в котором с информатикой связывают одно из следующих понятий: это либо *отрасль производства*, либо *фундаментальная наука*, либо *прикладная дисциплина*. И эта точка зрения широко представлена на различных сайтах в сети Интернет.

При этом обращается внимание на то, что все три направления информатики опираются на средства для преобразования информации) и развивают их.

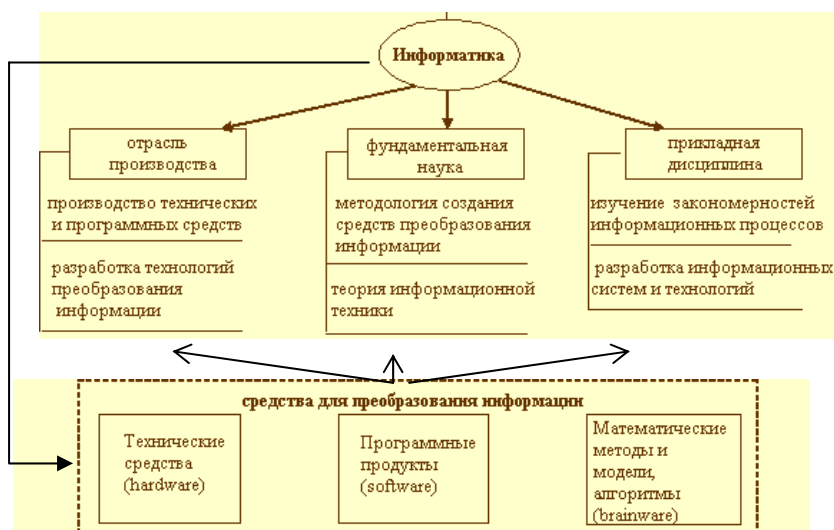


Рис 1.1. Структура информатики

Информатика как *фундаментальная наука* занимается разработкой методов, моделей и алгоритмов, а также связанных с ними математических теорий. Ее прерогативой является исследование процессов преобразования информации и на основе этих исследований разработка соответствующих теорий, моделей, методов и алгоритмов, которые затем применяются на практике.

Практическое использование результатов исследований информатики как фундаментальной науки воплощается в информатике – *отрасли производства*. Широко известны западные фирмы по производству программных продуктов, такие как Microsoft, Lotus, Borland, и технических

средств – IBM, Apple, Intel, Hewlett Packard и другие. Помимо производства самих технических и программных средств разрабатываются также и технологии преобразования информации.

Как *прикладная дисциплина* информатика изучает закономерности протекания информационных процессов в конкретных областях и методологию разработки конкретных информационных систем и технологий.

На основе первоначального определения информатики, предложенного в 1963 г. **Ф.Е. Темниковым**, в качестве способа систематизации разрозненно возникающих сведений об информации и информатике можно принять выделение 3-х основных частей: (рис. 1.2):

- 1) *информационные элементы*,
- 2) *информационные процессы*,
- 3) *информационные системы*.

Интерпретируем эти части применительно к настоящему периоду развития наук об информации.

Информатика – наука об информационных элементах

Информация – это отражение материи, т.е. объектов и их свойств.

Следовательно, *информационный элемент* – это отражение материально существующих элементов, элементарных объектов, учитываемых при решении прикладных задач.

В теории систем *элемент* определяется как предел членения системы с точки зрения аспекта ее рассмотрения, цели исследования или создания, решаемой конкретной задачи [1, 21, 25].

При определении состава элементов информатики следует иметь в виду наличие двух главных составляющих информатики – *информации и средств ее обработки*.

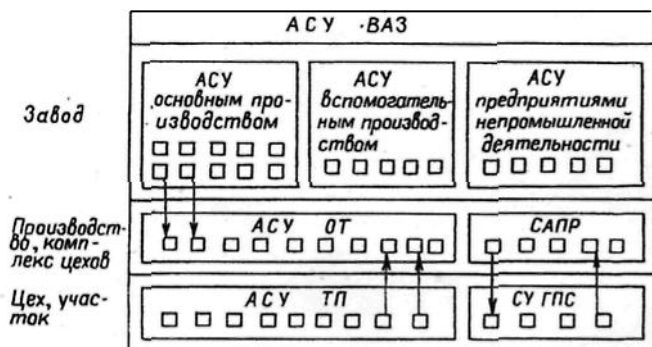
Более подробно понятие «информация», ее особенности, способы измерения, вопросы исследования информации как ресурса предприятия (организации) рассматриваются в гл. 2.

В качестве *информационных элементов* естественно, прежде всего, рассматривать *данные*.

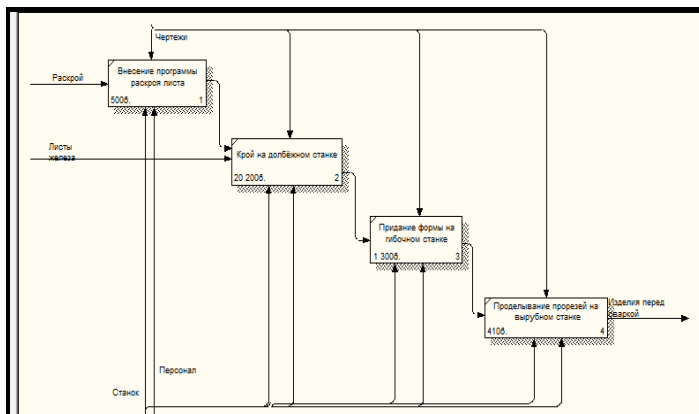
В технических науках информатика первоначально развивалась как наука о передаче и обработке данных, и в качестве элементов рассматривались символы: *буквы, цифры, слова*. Затем – *сведения* о деталях и других компонентах изделий.

В экономике в качестве данных рассматриваются сведения об объектах производственного или обслуживающего процесса, *показатели*, характеризующие состояние объекта, экономического процесса, происходящие в системе изменения.

Информационные системы



Информационные процессы



Информационные элементы

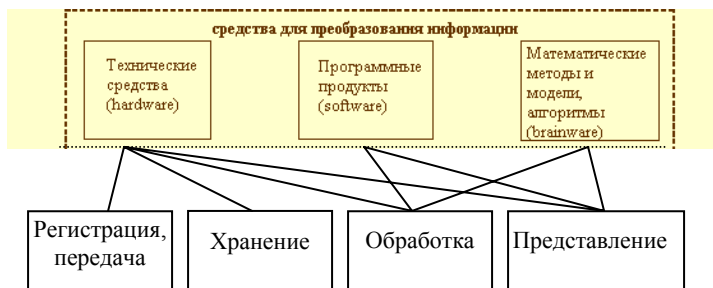


Рис. 1.2 Структура информатики по Ф.Е. Темникову

Для названия совокупностей данных вводят термины *информационный массив, база данных (БД), хранилище данных*.

Для позиционирования организации в системе экономики необходима нормативно-правовая информация, представляемая в виде *документальной информации*, т.е. *текстов*, в которых нужно искать элементы при этом элементы – *нормы, статьи*, т.е. *фрагменты текста*.

Для организации управления предприятием (организацией) нужна нормативно-техническая, нормативно-методическая, маркетинговая, мониторинговая информация, являющаяся *документально-фактографической*, т.е. представляемой в форме *текстов*, из которых нужно извлекать конкретные *данные* о производственных нормативах, состоянии среды, сведения о наличии и ценах товаров на рынках и т.п.

В каждом конкретном приложении нужно выделять информационные элементы с учетом поставленной задачи и необходимой степени детализации.

Для сбора, передачи, хранения, обработки и представления информации лицам, принимающим решения, нужны технологии, *технические и программные средства*, которые также являются *элементами информатики*.

В состав технических средств входят компьютеры и связанные с ними периферийные устройства (мониторы, клавиатуры, принтеры и плоттеры, модемы и т.д.), линии связи, средства оргтехники и т.п., т.е. те материальные ресурсы, которые обеспечивают преобразование информации. Важнейшей задачей информатики является изучение и обеспечение «дружественного» интерфейса между человеком и аппаратно-программными средствами обработки информации.

Диалог компьютера с человеком осуществляется с помощью искусственных языков. Языки, предназначенные для реализации формальных алгоритмов (вычислительных), называют *алгоритмическими*, или языками программирования. К настоящему времени создано немало алгоритмических языков для описания задач различных классов. По мере усложнения алгоритмов разрабатывают специализированные языки – *процедурно-ориентированные, проблемно-ориентированные, языки логического программирования*, которые также являются элементами информатики.

Общение человека и ЭВМ существенно упрощается, если имеются готовые *программы*, соответствующие характеру решаемых задач. Вызов готовых программ и их выполнение осуществляются в соответствии с инструкциями по эксплуатации ЭВМ. К *программным средствам (продуктам)* относятся операционные системы, интегрированные оболочки, системы программирования и проектирования программных продуктов, различные прикладные пакеты, такие, как текстовые и графические редакторы, бухгалтерские и издательские системы и т.д.

При проектировании и изготовления любого программного или технического средства в силу их сложности применяют *математические методы, модели и алгоритмы*

Методы и средства обработки информации детализируют:

1) методы и алгоритмы автоматизированного сбора, хранения, поиска и передачи информации;

2) методы и алгоритмов обработки и преобразования информации;

3) технологии и электронно-вычислительной техники, позволяющих развивать первые два направления;

4) методы, модели и алгоритмы для проектирования и производства новых и совершенствования старых технических и программных систем. В этом случае можно сказать, что средства преобразования информации используются для производства себе подобных. Тогда их пользователем является специалист в области информатики, а не конечный пользователь.

Информатика – наука об информационных процессах

Процессы – это последовательности операций во времени при проектировании, производстве изделий, в экономике – бизнес-процессы, организационно-технологические процедуры (ОТП) подготовки и реализации управленческих решений по организации производства, обслуживания, торговли, и т. п.

Информационные процессы – процессы, связанные с поиском, хранением, передачей, обработкой, использованием информации, последовательность операций при прохождении информации, сопровождающей последовательность операций во времени при проектировании, производстве изделий, при подготовке и реализации управленческих решений.

В технических дисциплинах (технической информатике) изучаются процессы *передачи* информации, в реализации которых участвуют *источник* и *приемник* информации (получатель) и канал передачи информации – *канал связи*, представляющий собой совокупность технических устройств, обеспечивающих передачу сигнала от источника к получателю (рис. 1.3).

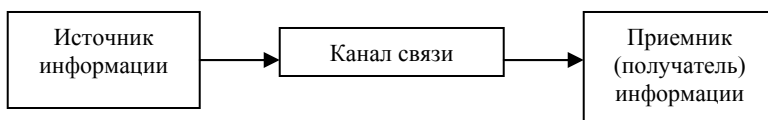


Рис. 1.3. Процесс передачи сигнала

При этом применяют *кодирующее* и *декодирующее* устройства. В процессе передачи информация может теряться и искажаться: Помехи, или, как их называют специалисты, шумы, искажают информацию. Существует наука, разрабатывающая способы защиты информации – *криптология*.

Каналы передачи сообщений характеризуются *пропускной способностью* и *помехозащищенностью*. Каналы делятся на *симплексные*, с передачей информации только в одну сторону (телевидение), и *дуплексные*, по которым возможно передавать информацию в оба направления (телефон, телеграф). По каналу могут одновременно передаваться несколько сообщений. Каждое из этих сообщений выделяется (отделяется от других) с помощью специальных фильтров. Например, возможна фильтрация по частоте передаваемых сообщений, как это делается в радиоканалах. Пропускная способность канала определяется максимальным количеством символов, передаваемых ему в отсутствие помех. Эта характеристика зависит от физических свойств канала. Для повышения помехозащищенности канала используются специальные методы передачи сообщений, уменьшающие влияние шумов.

В социально-экономических системах информационные процессы определяются *компьютерной технологией*, которая включает в себя последовательное выполнение определенных этапов работы с информацией: *ввод, хранение, обработка, поиск, предоставление*.

На подготовительных этапах осуществляется содержательный и формализованный анализ решаемой задачи, выбор метода и математической модели ее решения, определяется последовательность и порядок решения, его алгоритмическое описание, составляются программы на каком-либо доступном для машины языке. Затем программы вводятся в ЭВМ, отлаживаются, редактируются и записываются для хранения на внешних носителях.

Содержание исполнительных этапов зависит от характера задачи и типа используемой ЭВМ. Процесс сводится к автоматическому выполнению программы, причем часть программы может выполняться с участием человека. Завершающим этапом является анализ, оценка полученных результатов для их практического использования и совершенствования разработанных алгоритмов и программ.

Подготовительные этапы выполняются человеком, исполнительные – машиной или машиной с участием человека (диалоговые режимы работы ЭВМ).

Для исследования информационных процессов при проектировании организационных структур и автоматизированных информационных систем предприятий (организаций) в 1970-е гг. в теории систем был предложен *функционально-технологический подход*. Однако его в тот период практически не удалось реализовать из-за ограниченных возможностей вычислительной техники.

Для представления и исследования процессов при проектировании информационных систем в 1990-е гг. разработана методология SADT¹ и семейство IDEF-технологий, которые служат полезным инструментом для анализа процессов в информационных системах, что позволяет развить функционально-технологический подход, называемый в настоящее время *процессным*.

Изучение IDEF-технологий и других автоматизированных средств для реализации методологии SADT обеспечивается специальными дисциплинами –

¹ **Ross D.** Applications and extension of SADT // IEEE. Computer. – April, 1995.

«Информационные системы и технологии», «Проектирование информационных систем», «Моделирование бизнес-процессов».

Кроме того, информационные процессы – это не только отображение реально протекающих материальных процессов на предприятиях. Существуют процессы возникновения и распространения научно-технической информации. В теории научно-технической информации о таких процессах принято говорить как об *информационных потоках* и исследовать закономерности их строения (законы *Г. Циффа*, *С. Брэдфорда*, *Б. Викери*, закономерность концентрации – рассеяния *В. И. Горьковой*) [5].

Для развития предприятия, формирования ядра бизнеса исследование информационных потоков научно-технической, мониторинговой информации – весьма важная задача, которой начинают уделять все больше внимания, в том числе в материалах, представленных в сети Интернет.

Закономерности информетрии и возможности их применение при исследовании информационных потоков, процессов возникновения и распространении информации рассматриваются в гл. 7.

Информатика – наука об информационных системах

Термин «система обработки данных» (СОД) был введен при создании систем радиоуправления ракетами и другими искусственными космическими объектами, широко использовался при измерении и передаче информации о состоянии параметров организма космонавтов на борту космического корабля (системы типа «ТРАЛ»), при сборе и обработке статистической информации о состоянии атмосферы. В дальнейшем термин СОД стал использоваться и применительно к социально-экономическим объектам. Этот термин подразумевал в том числе сбор и обработку совокупности данных. В то же время понятие «система» – более сложная категория.

В истории становления понятия *информационная система* относительно независимо развивались несколько направлений. При этом, поскольку понятие информационной системы начало переосмысливаться и формироваться в связи с автоматизацией хранения и поиска информации, это понятие стало практически неотделимым от понятий – «*Автоматизированная информационная система*» и «*Информационно-поисковая система*».

Начиная с 1960-х гг. в истории развития информационных систем нашей страны относительно независимо сформировались два направления:

- разработка *автоматизированных информационных систем* – АИС как первой очереди *автоматизированных систем управления* – АСУ;
- разработка *автоматизированных систем научно-технической информации* – АС НТИ.

Работы проводились на основе системно-целевого подхода. АСУ и АСНТИ создавались для всех уровней и направлений управления страной:

государственного, отраслевого, регионального, предприятий и организаций. АИС создавались как *фактографические* информационные системы с представлением информации пользователям в виде регламентированных форм, в которых данные были сгруппированы в соответствии с алгоритмами решаемых на ее основе прикладных задач управления. АСНТИ разрабатывались как *документальные* информационно-поисковые системы (ДИПС).

Для управления разработками были подготовлены и утверждены соответствующие руководящие методические документы, выполняющие роль стандартов.

На предприятиях создавались *системы нормативно-методического обеспечения управления* (СНМОУ) и их автоматизированный вариант – АСНМОУ.

В связи с политическими и экономическими преобразованиями 1990-1991 гг. работы были приостановлены. В течение этого периода, вплоть до 1999 г., в основном, приобретались зарубежные локальные информационные системы и базы данных, которые внедрялись на предприятиях и в различных фирмах, создавались отдельные страницы документальной и фактографической информации в сети Интернет.

Но затем происходит постепенная адаптация к новым экономическим условиям, и в частности, информационные ресурсы и некоторые принципы организации Государственной системы научно-технической информации (ГСНТИ) сохранились в системе, которую в настоящее время условно называют Российской ГСНТИ (РГСНТИ).

При переходе к рыночной экономике, правовому государству возросла роль еще одного важного вида информации – *нормативно-правовой*. Начали создаваться автоматизированные системы *нормативно-правовой документации* (АС НПД) или *справочно-правовые системы* (СПС) разного рода и назначения (системы «Консультант+», «Кодекс», «Гарант» и др.).

На новой технической базе с использованием современных технологий развиваются информационные системы (ИС) государственного управления на федеральном и муниципальном уровнях, в частности, ИС государственной статистики; ИС природных ресурсов, явлений и процессов; государственные системы экономической, финансовой и научно-технической информации, внешнеэкономической деятельности; ИС библиотечной сети РФ и другие ИС специального назначения.

Виды и классификации информационных систем, определения информационных систем разного вида, принципы их разработки рассматриваются в главах 3–6.

Глава 2. ПОНЯТИЕ ОБ ИНФОРМАЦИИ И ИНФОРМАЦИОННЫХ РЕСУРСАХ ПРЕДПРИЯТИЙ И ОРГАНИЗАЦИЙ

2.1. Информация: понятие и основные свойства

Строгого определения информации нет.

Примем вначале наиболее распространенное определение и рассмотрим особенности информации.

Информация – это сведения об окружающем мире (объекте, процессе, явлении, событии), которые являются объектом преобразования (включая хранение, передачу и т.д.) и используются для выработки поведения, для принятия решения, для управления или для обучения.

Характерные черты информации:

- 1) представляет собой наиболее важный ресурс современного производства: снижает потребность в земле, труде, капитале, уменьшает расход сырья и энергии;
- 2) вызывает к жизни новые производства;
- 3) является товаром, причем продавец информации ее не теряет после продажи;
- 4) придает дополнительную ценность другим ресурсам, в частности, трудовым: действительно, работник с высшим образованием ценится больше, чем со средним;
- 5) может накапливаться.

С информацией всегда связывают три понятия (их взаимосвязь показана на рис. 2.1):

- *источник информации* – тот элемент окружающего мира, сведения о котором являются объектом преобразования;
- *потребитель информации* – тот элемент окружающего мира, который использует информацию;
- *сигнал* – материальный носитель, который фиксирует информацию для переноса ее от источника к потребителю и *код*, в форме которого информация передается.

Так, источником информации, которую получает читатель, является информатика как сфера человеческой деятельности; потребителем – сам читатель, а сигналом – бумага с текстом (в этом случае говорят, что

информация имеет бумажный носитель). Будучи прочитанной и освоенной читателем, информация приобретет еще один носитель – биологический, когда она «записывается» в память человека. Очевидно, что источник и потребитель в этом случае не меняются.

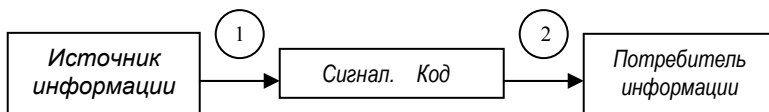


Рис. 2.1. Схема взаимосвязи основных понятий информации

Исследованием феномена информации, ее видов и особенностей занимаются философы, специалисты по техническим наукам, экономисты, психологи. Обобщая и преломляя результаты этих исследований применительно к проблеме использования информации как ресурса развития предприятий и организаций, выделяют ряд специфических особенностей информации, обуславливающие ее отличие от других видов ресурсов: *практически неубывающая потенциальная эффективность информации, тиражируемость и многократность использования*, зависимость фактической *реализуемости и эффективности* от степени использования информации, *абстрактный характер* и т.п.

Основные из них охарактеризованы в табл. 2.1.

Таблица 2.1

Основные особенности информации

Особенности информации	Краткая характеристика
<i>Практически неубывающая потенциальная эффективность информации</i>	Потенциальная эффективность информационных ресурсов не исчезает после не только однократного, но и многократного использования одной и той же информации (уменьшение эффективности может быть связано только с устареванием информации). В то время как основная характеристика материальных ресурсов ((минеральных ресурсов, сырья и т.п.)) – их потенциальная эффективность – уменьшается по мере их использования. В случае возобновляемых ресурсов их потенциальная эффективность может восстанавливаться, но для этого требуется определенный (часто значительный) период. Потенциальная эффективность искусственно создаваемых технических средств, в том числе оборудования, также имеет предел, определяемый их сроком службы, и исчезает после списания станка или другого технического средства.

Особенности информации	Краткая характеристика
	Эта особенность информационных ресурсов часто реализуется не сразу, а спустя многие годы (как бывало, например, с рядом открытий, изобретений или новых идей, не понятых поколениями при жизни их авторов).
<i>Независимость информации от ее создателей</i>	Свойство неубываемости информации является одним из условий неубывающей потенциальной эффективности информации и приводит к относительной <i>независимости информации от ее создателей</i> .
<i>Тиражируемость и многократность использования</i>	Иногда говорят и о тиражируемости технических средств, например, станка. Но в этом случае речь идет о тиражируемости конструкции станка, идеи, т. е. информации, а на создание каждого экземпляра станка одного и того же типа снова нужно затратить труд, в то время как при тиражировании информации труд (умственный) на ее создание больше не тратится, а труд, затрачиваемый на размножение информации, незначителен и им, как правило, можно пренебречь при оценке эффективности использования информации.
<i>Неаддитивность, некоммутативность и неассоциативность информации</i>	Отсутствие свойств <i>аддитивности, коммутативности и ассоциативности</i> у информации означает, что содержащаяся в каком-либо сообщении информация не есть просто арифметическая сумма составляющих это сообщение элементов информации, что эти элементы нельзя без искажения смысла располагать в сообщении в любой произвольной последовательности и группировать в произвольные сочетания. Эта особенность информации относится к научной и социальной информации.
<i>Кумулятивность информации</i>	Это свойство, характеризующее также в основном научную и социальную информацию, связано с одной из основных закономерностей развития науки и общества – ее преемственностью. Предшествующие достижения науки и культуры являются фундаментом для дальнейшего развития научной мысли и общества. Поэтому необходимы не только новые открытия, но и специальная работа по систематизации, оценке и обобщению информации. Со свойством кумулятивности тесно связано и явление концентрации информации во времени, т. е. переход знания о мире на все более высокие уровни абстракции.

Продолжение табл.2.1

Особенности информации	Краткая характеристика
Зависимость фактической <i>реализуемости</i> и <i>эффективности</i> от степени использования информации	В случае материальных ресурсов эффективность можно оценить коэффициентом использования материалов, сырья, электроэнергии и т. п., коэффициентом полезного действия оборудования и др. технических средств. Оценка реализуемости и эффективности информации – более сложная проблема, которая кратко будет рассмотрена ниже.
Необходимость наличия <i>источника, переносчика и приемника</i> (потребитель информации)	Сообщение становится информацией только в случае, когда есть <i>источник, переносчик</i> (в том числе передатчик, носитель) и <i>приемник</i> (потребитель), который должен хотеть воспринять информацию и быть способным ее понять и использовать. Эта особенность является определяющей при оценке как потенциальной, так и фактически реализуемой эффективности информационных ресурсов.
<i>Материя и информация</i> – парные философские категории	Появление новой информации всегда сопутствует появлению (созданию) новых форм существования материальных объектов и процессов, независимо от того, осознают ли этот факт их создатели. Поэтому вклад в формирование информационных ресурсов вносят практически все профессиональные группы работников предприятия: рабочие, которые создают новые образцы изделий и принимают участие в совершенствовании технологии (рационализаторская и изобретательская деятельность), инженерно-технические работники, которые проектируют изделия, системы, новые технологические процессы и т. п., ученые и научные работники, которые изучают явления и процессы естественного и искусственного (созданного руками человека) мира, делают открытия, разрабатывают фундаментальные основы будущих новых изделий, систем, технологий, а также руководители предприятий (организаций) и управленческие работники, которые занимаются непосредственно регистрацией, хранением и обработкой производственной и управленческой информации разного рода.
<i>Наличие ценности</i>	Ценность или полезность информации влияет на поведение получателя информации, на принятие им управленческих решений. Информация имеет тем большую ценность, чем в большей мере она <i>влияет на достижение целей</i> , стоящих перед получателем информации.

Имеется и ряд других особенностей информации, специфичных для научно-технической информации.

Научно-техническая информация (НТИ) или в широком смысле научная информация – это «получаемая в процессе познания логическая информация, которая адекватно отображает явления и законы природы, общества и мышления и используется в общественно-исторической практике»¹.

Это определение отражает четыре наиболее важных признака, необходимых для раскрытия понятия научная информация:

1. Научной может быть лишь информация, полученная в процессе познания объективных закономерностей природы, общества и мышления. При этом основу процесса познания составляет практика, производственная деятельность людей. К практике относятся материальное производство, научный эксперимент, сельскохозяйственная деятельность, деятельность по преобразованию природы и т. п.

2. Не всякая информация, полученная в процессе познания, есть научная. Чувственное познание дает представление лишь об отдельных, внешних сторонах вещей. Для того чтобы стать руководством к практической деятельности, необходимо знание сущности вещи, законов природы и общественной жизни. А для этого необходимо абстрактно-логическое мышление, облеченное в языковую форму. Именно этим *научная информация* отличается от *сведений* или *данных*, получаемых в процессе чувственного (эмпирического) познания. Данные и сведения – это «сырье» для создания научной информации.

3. Необходимо адекватное отображение явлений и законов природы, общества и мышления. При этом определение адекватности новых гипотез или теорий является сложной проблемой, по-разному решаемой на различных этапах развития общества с учетом конкретно-исторических условий.

4. Научная информация непременно должна использоваться в общественно-исторической практике. Это обеспечивает проверку истинности логической информации.

А.Д. Урсул уточняет приведенное определение, дополняя его возможностью *использования информации для предвидения и преобразования действительности*².

Кроме того, *А.Д. Урсул* включает в определение необходимость *фиксации получаемой в процессе познания информации в системе точных понятий* (суждений, умозаключений, теорий, гипотез).

В то же время в [15, 16] научную информацию отличают от научных знаний. В *знаниях* информация представлена уже в наиболее обобщенном и систематизированном виде и выражается в *системе понятий, суждений, умозаключений, в теориях*. Таким образом, научные

¹ *Словарь* современного русского литературного языка. – М. – Л., Изд-во АН СССР, 1956. – Т. 5. – С. 418.

² *Урсул А.Д.* Информация / *А.Д. Урсул.* – М.: Наука, 1971. – 295 с. *Урсул А.Д.* Отражение и информация / *А.Д. Урсул.* – М.: Мысль, 1973. – 231 с.

знания – не вся научная информация, а лишь ее определенная часть. Предлагается различать ряд уровней, на которых совершаются акты мышления: сбор научных фактов, систематизация и обобщение, формулирование и проверка гипотез, концепций, законов, построение теорий.

Первоначально различали *научную* и *научно-техническую* информацию, относя последний термин к области «наука и техника» [15, 16]. Однако в последующем эти термины используются как синонимы, поскольку информация, полученная и используемая в технике (как и в любой другой области человеческой деятельности) обязательно относится к сфере науки, получена в результате научных исследований. Поэтому в настоящее время эти термины используются как синонимы, и даже в большей мере принят термин «научно-техническая информация».

Основные специфические особенности научно-технической информации [15, 16, 18] приведены в табл. 2.2.

Таблица 2.2:

Специфические особенности научно-технической информации

Особенности НТИ	Краткая характеристика
1. <i>Неотрывность научной информации от физического носителя</i>	Несмотря на то, что научная информация по своей природе идеальна (нематериальна), она не может существовать вне той или иной материальной оболочки (носителя). Эта особенность отражает более общую: материя и информация – парные категории. Информация не существует без отражения материи, как и без других атрибутов материи – движения, пространства, времени и т. д.; она является содержанием отражения.
2. <i>Неаддитивность, некоммутативность и неассоциативность</i>	Отсутствие свойств <i>аддитивности, коммутативности и ассоциативности</i> у научной информации, на что впервые обратил внимание В.И. Сифоров ¹ означает, что содержащаяся в каком-либо сообщении информация не есть просто арифметическая сумма составляющих это сообщение элементов информации (например, слов), что эти элементы нельзя без искажения смысла располагать в сообщении в любой произвольной последовательности и группировать в произвольные сочетания.
3. <i>Наличие ценности субъективно</i>	<i>Ценность</i> научной информации <i>субъективна</i> . То, что очень важно для одного субъекта, может быть совсем не важно для другого и наоборот. Производство информации всегда должно быть <i>адресным</i> , иначе оно бессмысленно.

¹ **Сифоров В.И.** Научная информация и повышение эффективности научных исследований / В.И. Сифоров // Управление, планирование и организация научных и технических исследований: кн. – М.: 1970. – Т. 3. – С. 291–292.

Особенности НТИ	Краткая характеристика
4. <i>Общественная природа</i>	<p>Источником научно-технической информации является познавательная деятельность людей. Явления и законы природы познаются совокупностью индивидуумов в их взаимодействии, обществом. Общество обладает памятью, более емкой и разносторонней, чем память любого индивидуума.</p> <p>Совокупность научных знаний о явлениях и законах мира и деятельности людей, накопленная обществом, называют общечеловеческим <i>тезаурусом</i>.</p> <p>В этом смысле совокупность понятий и закономерностей функционирования предприятия иногда называют <i>тезаурусом предприятия</i> [29].</p>
5. <i>Семантический характер и языковая природа</i>	<p>Термин «семантический» характеризует научную информацию в плане содержания, а не формы (план выражения). Научная информация является смысловой, семантической, т. е. имеет понятийный характер. Именно понятия составляют смысл слов, в них обобщаются существенные признаки явлений и процессов.</p> <p>При этом понятия формируются и существуют в языковой оболочке. Под языком понимается знаковая система любой физической природы, служащая средством человеческого общения, мышления и выражения.</p> <p>Язык может быть естественным и искусственным, т. е. созданным для узких потребностей (математическая символика, язык структур, информационно-поисковый язык и т. п.). Язык – явление социальное. Термин «языковый» (от греч. <i>λογος</i>, <i>logos</i> – слово, понятие, смысл, разум). Поэтому научную информацию можно считать логической.</p> <p>В то же время семантика НТИ значительно более изменчива, чем форма ее выражения – язык.</p>
6. <i>Независимость от языка и носителя информации</i>	<p>Инвариантность научной информации относительно языка, на котором она выражается, и от ее физических носителей, характерна для научной информации, но не является абсолютно универсальным свойством. Например, эстетическая информация существенно зависит от языка, и даже от физического носителя. Аналогично конструкторская информация не может быть адекватно отображена в вербальной форме.</p>
7. <i>Дискретность</i>	<p>Дискретность является диалектически противоречивым свойством. Семантическая научная информации: в плане языкового выражения представляется единицами языка (словами, предложениями, абзацами, т. е. дискретными носителями); а в плане содержания – понятиями, описанием фактов, законов, гипотез, концепций, теорий, которые можно считать превращением дискретной информации в непрерывную, целостную.</p>

Особенности НТИ	Краткая характеристика
	<p>В то же время «научные произведения» имеют логическую завершенность и отчуждение этих произведений от их творцов происходит не непрерывно, а дискретно, в виде относительно законченных творческих актов, получающих материальное воплощение, и их можно считать «квантом» научной информации.</p> <p>В случае устных научных произведений (лекции, доклады, выступления) содержательная дискретность проявляется в виде смысловых фрагментов, а материальная дискретность – в виде риторических периодов, характеризующихся паузами.</p> <p>Свойством дискретности обладают не все виды научной информации. Например, данные, получаемые с помощью некоторых измерительных приборов, могут быть и непрерывными (аналоговые устройства).</p>
8. <i>Независимость от создателей</i>	<p>Это свойство характерно в большей мере для научной социальной информации, и особенно для информации, используемой при организационном управлении. В этой сфере передача опыта часто осуществляется без фиксации авторов нововведений.</p> <p>Для регистрации авторства в области научно-технической информации существуют центры регистрации и выдачи патентов и авторских свидетельств, в которых проводится работа по установлению патентной чистоты на основе сопоставления предлагаемых изобретений и открытий с предшествующими; патентные библиотеки; центры регистрации научных отчетов, диссертаций и др. неопубликованных материалов, специальные издания, представляющие обобщенные сведения о новых научно-технических материалах и их авторах.</p> <p>В нашей стране для обеспечения работы по поиску авторов научно-технической информации в отраслях науки и техники существуют специальные издания – Бюллетень изобретений и открытий, сигнальные, обзорные и экспресс-информации, библиографические указатели, реферативные журналы и т. д.</p> <p>По мере развития науки, увеличения числа открытий и изобретений, возникла проблема авторства и в сфере научных исследований. Все больше наука становится не индивидуальной, а общечеловеческой. Возрастает число крупномасштабных коллективных научных работ. Усложняется проблема выявления и учета цитируемости опубликованных материалов, особенно статей и тезисов докладов. В XIX в. этой работой активно занимался Всесоюзный институт научно-технической информации (ВИНИТИ).</p>

Особенности НТИ	Краткая характеристика
	<p>Однако эта работа может проводиться в основном за счет госбюджетного финансирования, и в современных условиях рынка и сети Интернет становится все более затруднительной. Существуют и проблема охраны авторских прав (особенно в сети Интернет).</p> <p>Это свойство связано с одной из основных закономерностей развития науки и общества – ее преемственностью и интернациональным характером.</p> <p>Предшествующие достижения науки и культуры являются фундаментом для дальнейшего развития научной мысли и общества.</p>
<p>9. <i>Кумулятивность</i></p>	<p>Необходимы не только новые открытия, но и специальная работа по систематизации, оценке и обобщению информации. Со временем все второстепенное, частное отсеивается, а существенное, основное получает простое и краткое выражение. Со свойством кумулятивности тесно связано и явление <i>концентрации</i> информации во времени, переход знания на более высокие уровни абстрагирования.</p>
<p>10. <i>Концентрация и рассеяние (интеграция и дифференциация) информации</i></p>	<p>Явление <i>концентрации</i> информации во времени – это переход знания о мире на все более высокие уровни абстракции: А. Научные факты. ⇒ В. Научные гипотезы, концепции. ⇒ С. Совокупность гипотез, концепций теорий, законов, образующие основы определенной науки, парадигмы, отрасли знаний. ⇒ D. Научная информация мировоззренческого характера.</p> <p>Условно явление концентрации представлено на рис. 2.2.</p> <div data-bbox="434 1008 934 1262" data-label="Figure"> </div> <p>Рис. 2.2</p> <p>Это явление тесно связано со свойством кумулятивности.</p>

Особенности НТИ	Краткая характеристика
	<p>Свойство научной информации <i>рассеиваться</i> по многим научным произведениям непосредственно связано с рассмотренными выше свойствами дискретности, кумулятивности, независимости от создателей и свойством старения. Оно проявляется в том, что одни и те же содержательные единицы научной информации (факты, понятия, гипотезы, теории и т.п.) одних авторов могут по-разному и в разных контекстах использоваться в произведениях других авторов, могут менять языковое выражение.</p> <p>Это неизбежно, способствует развитию научной мысли. Заимствованная информация, располагаясь в новых произведениях в соответствии с замыслом их авторов, формирует новый целостный смысл, что и обеспечивает развитие науки.</p> <p>Явления <i>концентрации</i> и <i>рассеяния</i> представляют собой проявление явлений <i>интеграции</i> и <i>дифференциации</i>, которые являются закономерностями развития науки. Изучения этих явлений – важнейшая проблема информатики. В результате исследования этой проблемы были открыты законы <i>Г. Цунфа</i>, <i>С. Брэдфорда</i>, <i>Б. Викери</i>, закономерность концентрации – рассеяния <i>В.И. Горьковой</i> (см. гл. 7).</p>
11. <i>Старение</i>	<p>Старение научно-технической информации не связано непосредственно только со временем ее возникновения. В строгом понимании устаревают только те факты, гипотезы теории и т.п., которые с появлением новой научной информации оказываются неверными, перестают адекватно отображать явления и закономерности материального мира, общества и мышления, т.е. перестают быть научной информацией. Примером устаревания информации служит замена геоцентрической концепции Аристотеля – Птолемея на гелиоцентрическую концепцию Коперника – Бруно.</p> <p>Наибольшему старению подвержена информация класса В, в наименьшей – классов А и С.</p> <p>К проблеме старения информационных изданий нужно относиться аккуратно. Критерием здесь не может служить год издания, поскольку первоисточник может содержать более точные и глубокие сведения по сравнению с его пересказом в других изданиях.</p>

Особенности НТИ	Краткая характеристика
	В то же время старение информационных изданий может быть связано не с устареванием фактов или теорий, а с более сжатым их изложением, с изменением формализованного языка представления этих теорий (что происходило, например, с алгеброй, статистикой и др. науками), т.е. со свойством кумулятивности научной информации. Старение может быть связано с изменением господствующей концепции, парадигмы на том или ином этапе развития научного знания.
12. <i>Зависимость получаемой информации от априорной информации</i>	<i>Потребление информации требует определенных навыков и усилий: чем сложнее информация, тем большим интеллектом должен обладать ее потребитель. Усвоение многих видов информации требует специальных знаний, расширения тезауруса потребителя информации. Проводятся исследования зависимости получаемой информации от априорной, т.е. от тезауруса потребителя. Например, исследования Ю.А. Шрейдера (см. раздел 2.2).</i>

Особенности научно-технической информации необходимо учитывать при разработке информационно-поисковых систем, при комплектовании и актуализации информационных фондов.

Виды и формы научно-технической информации будут рассмотрены при характеристике систем научно-технической информации (гл. 5).

Рассмотренные особенности научно-технической информации позволяют охарактеризовать ее соотношение с другими видами экономической и социальной информации. А это, в свою очередь, помогает решать вопросы сбора, хранения, предоставления, защиты информации.

В настоящее время достаточно широко распространено мнение экономистов о том, что научно-техническая информация является не только важным ресурсом экономической системы, но и товаром. В то же время есть точки зрения, в соответствии с которыми научно-техническая информация не является товаром в классическом смысле, а если и имеет ценность на рынке, то как товар особого свойства.

Действительно, научно-техническую информацию можно отнести к категории товара на основе следующих свойств: 1) НТИ есть продукт труда; 2) НТИ продается и покупается, т.е. имеет цену на рынке. При этом товар есть продукт труда, который удовлетворяет какую-либо потребность людей и производится не для собственного потребления, а для продажи. И этим свойствам тоже в принципе может удовлетворять научная информация.

Правда, по Марксу «Для того, чтобы стать товаром, продукт должен быть передан в руки того, кому он служит в качестве потребительной стоимости

посредством обмена»¹. В то же время, хотя изначально научно-техническая информация создается не для продажи, она может использоваться посредниками (реализация книг фирмами-посредниками, продажа патентов и т. п.).

Стремлению рассматривать научно-техническую информацию как товар способствуют такие ее свойства, как практически неубывающая потенциальная эффективность информации, тиражируемость и многократность использования, зависимость фактической реализуемости и эффективности от степени использования информации, наличие ценности.

Если научно-техническую информацию признать товаром, то следует изменить отношение к ее распространению, к ее представлению в сети Интернет.

В то же время научно-техническая информация – особый товар, и нужно, видимо, по-разному относиться к хранению и организации доступа к различным ее видам, учитывать, что НТИ является основой развития общества в целом, основой развития науки и образования.

По мере исследования свойств информации разного вида стали осознавать, что существуют различные виды информации.

Так, **Е.К. Войшвилло**² предложил различать информацию *восприятия* (знаки) и *информацию-значение*, смысловое содержание для потребителя информации (означаемое знака).

В дальнейшем с точки зрения пользователя информацию стали рассматривать в нескольких аспектах:

синтактический (или материально-энергетический, знаковый, технологический) – с точки зрения техники передачи информации;

семантический – с точки зрения смыслового содержания и правильности толкования;

прагматический – для достижения целей потребителя.

При этом считалось, что наиболее общим является понятие *прагматической информации*.

Рассматривали эти аспекты и как уровни существования информации и вводили для разных уровней разные оценки (подробнее см. в следующем разделе).

Развивалось понятие *информации* и как философской категории. В частности, **А.Д. Урсул**³ стал рассматривать информацию как форму существования материи, подобно массе и энергии.

¹ **Маркс К., Энгельс Ф.** Соч. – Т. 23. – С. 49.

² **Войшвилло Е.К.** Понятие / Е.К. Войшвилло. – М.: МГУ, 1967. – 286 с.; **Войшвилло Е.К.** Символическая логика: Классическая и релевантная: Философско-методологические аспекты: учеб. пособие / Е.К. Войшвилло. – М.: Высшая школа, 1989. – 149 с.

³ **Урсул А.Д.** Информация / А.Д. Урсул. – М.: Наука, 1971. – 295 с.; **Урсул А.Д.** Отражение и информация / А.Д. Урсул. – М.: Мысль, 1973. – 231 с.

В теории информационного поля *А.А.Денисова*¹ информация рассматривается как *структура материи*, как категория, парная по отношению к материи. В этой теории вводятся понятия *чувственной* и *логической* (семантической и прагматической) информации, которые для удобства практических приложений названы *информацией восприятия* и *информационным потенциалом* (см. [8], и разделы, подготовленные А.А. Денисовым в [1, 21, 25]).

Долгое время в предыдущих теориях эти информации рассматривались как отдельно существующие, или как уровни информации, и поэтому единое определение понятия *информации* являлось предметом дискуссий.

А.А. Денисов показал, что *информацией* для потребителя является пересечение информации восприятия или *чувственной информации* и ее потенциала или *логической информации*, в результате чего формируется единое понятие – *информационный смысл, информационная сложность*.

При этом он определяет информацию следующим образом:

«... информация – это понятие, не поддающееся анализу средствами формальной логики и требующее применения к нему диалектической логики, которая обеспечивает возможность анализа не только абсолютно, но и относительно истинных высказываний» [1, с. 239].

2.2 Подходы к измерению и оценке ценности информации

В различных науках об информации предпринимались попытки ее измерения.

Первое научное осмысление понятия *количество информации* началось в теории связи. Были введены меры информации.

В 1924 г. *Г. Найквист* показал, что скорость передачи сигналов W по каналу связана с числом n различных кодовых символов зависимостью

$$W = k \log n, \quad (2.1)$$

где k – константа, определяемая скоростью передачи последовательных символов. Он первым предложил логарифмическую меру информации.

В 1928 г. *Р. Хартли*² определил информацию сообщения как

$$H = m \log n, \quad (2.2)$$

¹ *Денисов А.А.* Теоретические основы кибернетики. Информационное поле / А.А. Денисов. – Л.: ЛПИ, 1975. – 40 с.

² *Хартли Р.* Передача информации / Р. Хартли // Теория информации и ее приложения. – М.: 1959. – С. 5–35.

где m – количество символов в сообщении, а n – количество доступных для использования различных символов.

С точки зрения современной теории информации эта мера применима только тогда, когда символы выбираются независимо один от другого, и выбор любого из символов равновероятен.

Наиболее полное развитие теории информации применительно к проблемам ее передачи было осуществлено в сороковые годы **К. Шенноном**¹.

Шеннон связал понятие информации с понятием *энтропии* и предложил следующую энтропийную меру:

$$H = - \sum_{i=1}^n p_i \log p_i, \quad (2.3)$$

где p_i – вероятность того, будет или нет выбран i -й символ из полного набора в n символов, которые вырабатывает источник сообщений; величина H (энтропия) измеряется в битах (от англ. *binary digit* – двоичная система).

Работы Шеннона и его последователей нашли широкое применение на практике – при оптимизации каналов связи, в системах обработки данных, при разработке электронно-вычислительной техники и т. п.

Важные аспекты теории информации применительно к теории связи были решены **А.Н. Колмогоровым, Н. Винером**. В частности, они независимо решили задачу *фильтрации* (выделения полезного сигнала из комбинации «сигнал–шум» при заданных статистических характеристиках сигнала) и *предсказания* значения полезного сигнала.

Эта задача находит широкое использование при проектировании непрерывных динамических систем, в системах управления технологическими процессами.

С развитием кибернетики и теории систем понятие информации развивалось.

Для оценки возможностей различных технических средств сбора, передачи и обработки информации были предложены: теоретико-игровой подход к оценке информации **Р.Л. Стратоновича**² алгоритмический подход **М.М. Бонгарда**³.

В дальнейшем информацию стали связывать не только с *источником* и *приемником* сигналов, но и с ее *значением* для потребителя.

¹ **Шеннон К.Э.** Работы по теории информации и кибернетике: сб. статей / К.Э. Шеннон. Пер. с англ. с предисловием А.Н. Колмогорова / Под ред Р.Л. Добрушина и О.Б. Лупанова. – М.: Иностран. лит., 1963. – 829 с.

² **Стратонович Р.Л.** О ценности информации / Р.Л. Стратонович // Изв. АН СССР: Техническая кибернетика. – 1965. – № 5. – С. 25–38.

³ **Бонгард М.М.** О понятии «полезность информации» / М.М. Бонгард // Проблемы кибернетики: сб. – М.: Физматгиз, 1963. – Вып. 9. – С. 3–23.

Вначале имелись попытки применения рассмотренных выше мер и для оценки ценности социальной информации. Однако при этом возникают трудности доказательства адекватности таких мер.

Предпринимались также попытки создать *семантическую* теорию информации. Из них наиболее известной является концепция **Р. Карнапа** и **И. Бар-Хиллела**¹, базирующаяся на понятии *логической вероятности* как степени подтверждения гипотезы. В соответствии с этой концепцией наивысшей ценностью обладают гипотезы, подтвержденные достоверным знанием, экспериментом. В этом случае логическая информация приравнивается единице, а семантическая нулю. С уменьшением степени подтверждения гипотезы количество содержащейся в ней семантической информации увеличивается. Следует отметить, что при этом гипотезы формулируются на специальном языке, предложенном в рассматриваемой теории.

Развивая концепцию Карнапа – Бар-Хиллела, **Л. Бриллюэн**² предложил разновидность статистической меры для измерения семантической информации, основанную на измерении и уменьшении неопределенности (энтропии). Опираясь на второе начало термодинамики (принцип Карно), **Л. Бриллюэн** вводит для измерения информации следующую меру снижения энтропии:

$$I_1 = K \ln(P_0/P_1) = K \ln P_0 - R \ln P_1, \quad (2.4)$$

где P_0 – число равновероятных исходов, возникновение которых *априори* считается равновероятным; при этом нет исходной информации о задаче, т. е. априорная информация $I_0 = 0$;

P_1 – число равновероятных исходов при наличии информации об измерениях в аналогичных, родственных ситуациях, и тогда информация, снижающая неопределенность равна I_1 .

Бриллюэн по аналогии с термодинамикой использовал натуральный логарифм, а для того, чтобы выбрать систему единиц ввел коэффициент K , который для приведения к принятым к тому времени битам равен

$$K = 1/\ln 2.$$

А.А. Харкевич³ связал ценность информации с целью деятельности, предложив рассматривать энтропийную меру Шеннона как меру вероятности непопадания в цель, т. е. как меру несоответствия.

Ю.А. Шрейдер предложил построить теорию семантической информации на основе концепции разнообразия, а не концепции снятия неоп-

¹ **Bar-Hillel Y.** Semantic Information / Y Bar-Hillel, R. Carnap // British Journal of the Philosophy of Science, 1953. – V. 4. – № 14.

² **Бриллюэн Л.** Наука и теория информация / Л. Бриллюэн. – М.: Физматлит, 1960. – 392 с.

³ **Харкевич А.А.** О ценности информации / А.А. Харкевич // Проблемы кибернетики: сб. – Вып 4. – М.: Физматгиз, 1960. – С. 33–41.

ределенности, и в частности, на основе учета такого свойства информации как *зависимость получаемой информации от априорной информации*¹. Опираясь на идею **Н. Винера** о том, что для понимания и использования информации получатель должен обладать определенным запасом знаний, т.е. в терминах математической лингвистики и теории языков – *тезаурусом* Θ), **Ю.А. Шрейдер** определяет количество семантической информации, содержащейся в тексте T , как изменение тезауруса $\Theta I(T, \Theta)$.

Такой подход существенно отличается от концепции выбора при статистическом подходе, при котором предполагается, что получаемая информация тем больше, чем меньше априорных сведений содержится в приемнике информации.

Напротив, по Шрейдеру, чем более сложную структуру имеет тезаурус, тем больше существует возможностей для изменений под воздействием одного и того же сообщения. Это хорошо согласуется с законом необходимого разнообразия У.Р. Эшби, согласно которому управляющая (осмысливающая, понимающая, принимающая решения) система должна обладать большим необходимым разнообразием (сложностью) по сравнению с разнообразием поступающей в нее информации от управляемой (понимаемой) системы.

Предлагаемый **Ю.А. Шрейдером** подход согласуется также с наблюдениями за процессами обмена информацией: по мере роста наших знаний об изучаемом объекте растет и количество извлекаемой информации об этом объекте. Динамику насыщения индивидуального тезауруса Ю.А. Шрейдер иллюстрирует некоторой условной зависимостью (рис. 2.3), характер которой зависит от конкретного потребителя информации или накапливающей информации информационная система.

Эта кривая характеризует одну из особенностей информации – *зависимость получаемой информации от априорной информации*.

Для количественной оценки ценности семантической информации предлагались различные меры.

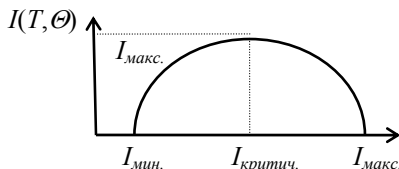


Рис. 2.3. Зависимость информации от априорной

¹ **Шрейдер Ю.А.** О семантических аспектах теории информации / Ю.А. Шрейдер // Информация и кибернетика: сб. – М.: Изд-во Сов. Радио, 1967. – С. 22–28.

Так, в [15] предлагается мера для общечеловеческого тезауруса, под которым понимается совокупность понятий, вырабатываемых человечеством в процессе научного познания природы, общества и мышления, связанных друг с другом. Процесс научного познания представляется как уточнение старых и формирование новых научных понятий, в выявлении и корректировке связей между понятиями, т.е. тезаурус трактуется как определенная структура, которая изменяется под воздействием новой информации в процессе познания и предлагается гипотетический подход, в соответствии с которым информация не одинаково зависит от сложности структуры.

Если под воздействием сообщения внутренняя структура общечеловеческого тезауруса Θ становится менее сложной, чем до этого сообщения, то в таком сообщении содержится больше информации, чем в сообщении, вызывающем усложнение структуры Θ .

Этот вывод согласуется с опытом: известно, что крупные научные открытия в целом упрощают структуру знания – благодаря введению новых, более общих понятий, что и является следствием и проявлением свойства *кумулятивности* информации.

Авторы [15] оговаривают гипотетичность предлагаемого подхода к оценке ценности информации в зависимости от структуры представления общечеловеческого тезауруса, и высказывают идею о необходимости поиска иных мер оценки ценности информации.

Для оценки удовлетворения информационных потребностей в теории научно-технической информации введены меры *релевантности* и *пертинентности* (см. гл. 5).

Затем, когда стали осознавать, что существуют различные виды информации, **Е.К. Войшвилло**¹ предложил различать информацию *восприятия* (знаки) и *информацию-значение*, смысловое содержание для потребителя информации (означаемое знака).

В дальнейшем, как было отмечено выше, информацию стали рассматривать в нескольких аспектах с точки зрения наблюдателя (пользователя): *прагматический* – для достижения целей наблюдателя; *семантический* – с точки зрения смыслового содержания и правильности толкования; *синтактический* (или материально-энергетический, знаковый, технологический) – с точки зрения техники передачи информации. При этом считали, что либо наиболее общим является понятие *прагматической информации*, а *семантический* и *синтактический* аспекты информации имеют подчи-

¹ **Войшвилло Е.К.** Понятие / Е.К. Войшвилло. – М.: МГУ, 1967. – 286 с.; **Войшвилло Е.К.** Символическая логика: Классическая и релевантная: Философско-методологические аспекты: учеб. пособие / Е.К. Войшвилло.. – М.: Высшая школа, 1989. – 149 с.

ненное значение, либо эти аспекты можно считать уровнями существования информации, стали вводить для разных уровней разные оценки (иногда с достаточно экзотическими названиями).

Так, в системе мер оценки экономической информации, изложенной в работах **Ю.И. Черняка**¹ предлагается различать несколько аспектов (уровней) представления и измерения экономической информации (с точки зрения ее полезности для решения задачи, смысла или семантики текста, синтактики знакового отображения, морфологии образования знаков – слов и словосочетаний, и передачи их по каналам связи) и для каждого уровня вводятся свои меры информации в терминах, принимаемых на этом уровне решений (*прагма, сем, знак* или *символ* и т. п.).

Развивалось понятие информации и как философской категории. В частности, **А.Д. Урсул**² стал рассматривать информацию как форму существования материи, подобно массе и энергии.

В теории информационного поля и ее дискретного варианта – информационного подхода к анализу систем **А.А. Денисова** [1, 8, 21] информация рассматривается как структура материи, как категория, парная по отношению к материи, и вводятся понятия чувственной *J* и логической *H* (семантической и прагматической) информации, которые для удобства практических приложений названы информацией восприятия и информационным потенциалом.

Долгое время в предыдущих теориях эти информации рассматривались как отдельно существующие, или как уровни информации, и поэтому единое определение понятия информации являлось предметом дискуссий.

А.А. Денисов показал, что информацией для потребителя является пересечение информации восприятия или чувственной информации и ее потенциала или логической информации, в результате чего формируется единое понятие – информационный смысл, информационная сложность, которая в частных случаях представляет собой декартово произведение) *J* и *H*

$$C = J \cap H \text{ или } C = J \times H.$$

Для конструктивного использования понятий чувственная и логическая информация **А.А. Денисов** ввел соответствующие детерминированные и статистические меры (табл. 2.3).

¹ **Черняк Ю.И.** Информация и управление / Ю.И. Черняк. – М.: Наука, 1974. – 184 с.; **Черняк Ю.И.** Система мер информации как инструмент оптимизации языка управления / Ю.И. Черняк, А.И. Чухнов, Ю.Ю. Финкельштейн // Языки экономического управления и проектирования систем. – М.: Наука, 1973. – С. 24–27.

² **Урсул А.Д.** Информация / А.Д. Урсул. – М.: Наука, 1971. – 295 с.; **Урсул А.Д.** Отражение и информация / А.Д. Урсул. – М.: Мысль, 1973. – 231 с.

Меры информации А.А. Денисова

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; border-radius: 50%; padding: 10px; text-align: center;"> Информация восприятия (чувственная информация) J (элементная база) </div> <div style="font-size: 2em;">C</div> <div style="border: 1px solid black; border-radius: 50%; padding: 10px; text-align: center;"> Логическая информация H (информационный потенциал) (сущность) </div> </div>			
Детерминированный способ измерения	$J_i = A_i / \Delta A_i,$ <p>где A_i – значение измеряемой величины, ΔA_i – «квант», с точностью до которого ЛПР интересуется воспринимаемая информация (единица измерения, разрешающая способность прибора)</p>	$H = \sqrt[\gamma]{\frac{1}{n} \sum_{i=1}^n J_i^\gamma},$ <p>где J_i – результаты измерения A_i; n – объем понятия об охватываемых измерением объектах; γ – параметр усреднения</p>	<p>При $\gamma = 1$ получим среднее арифметическое</p> $H = \sum_{i=1}^n \frac{1}{n} J_i.$ <p>При $\gamma = 0$ получим среднее геометрическое</p> $H = \sqrt[n]{\prod_{i=1}^n J_i}.$ <p>При $\gamma = -1$ – среднее гармоническое</p> $H = n / \sum_{i=1}^n \frac{1}{J_i}$
Вероятностный способ измерения	$J_i = -\log_2 p_i,$ <p>где p_i – вероятность события.</p> <p>В случае использования информации для достижения цели p_i называют вероятностью недостижения цели или степенью нецелесоответствия</p>	$H = \int f(J_i) dJ_i \Rightarrow$ $H = \sum_{i=1}^n q_i J_i = -\sum_{i=1}^n q_i \log p_i,$ <p>где q_i – вероятность использования элемента информации.</p> <p>При $q_i = p_i$ $H = -\sum_{i=1}^n p_i \log p_i.$</p> <p>При равновероятном выборе элемента $p_i = 1/n$ и</p> $H = -\sum_{i=1}^n \frac{1}{n} \log \frac{1}{n} = \log n$ <p>Для прагматической информации</p> $H_{\eta} = -\sum_{i=1}^n q_i \log(1 - p'_i),$ <p>где p'_i – вероятность достижения цели, степень целесообразности</p>	

В концепции *А.А. Денисова* учитывается, что далеко не всегда может быть строго определено вероятностное пространство. В этом случае можно использовать понятие размытой вероятности в смысле Заде. Поэтому предложена информационная мера, учитывающая степень влияния информации как ресурса на реализацию целей.

Эта мера позволяет учесть не только вероятность p_i , влияющая на достижение (или недостижение) цели, но и вероятность q_i того, что данная информация будет использована лицом, принимающим решение:

$$H = - \sum_{i=1}^n q_i \log p_i . \quad (2.5)$$

Мера *К.Шеннона* получается из соотношения (2.5) при $q_i = p_i$ как частный случай.

Существуют и другие меры количества информации, учитывающие процесс ее передачи и изменение вероятности появления сообщения от априорного значения $p(x)$ на входе канала передачи информации к апостериорному значению $p(x/y)$ на выходе канала, связанные с искажением информации в канале, т.е. основанные на условной вероятности ([11], с. 16–21).

2.3. Понятие об информации как ресурсе предприятия (организации)

В течение длительной истории развития цивилизации деятельность, выходящая за пределы материального производства и обслуживания, относилась к непроизводительным затратам. Экономический статус государства определялся имеющимися материальными и энергетическими ресурсами. Однако в настоящее время осознано, что основным ресурсом производственных предприятий и других организаций становится *информация*.

Переход к наукоемкому производству во второй половине XX в. сопровождался возникновением новых видов информационного обмена. Возросла зависимость экономики от источников, объемов и качества информации (научно-технической, экономической, политической и др.), от уровня развития средств доступа к этой информации, что привело к формированию на рубеже 80-х гг. XX столетия принципиально новой экономической категории – *национальные информационные ресурсы*. Интенсивное развитие экономики становится невозможным без эффективного информационного сопровождения. Информация превратилась в стратегический ресурс, существенно возросла роль знаний.

В социально-экономических системах управления в зависимости от формы различают *документальную* и *фактографическую* информацию, в зависимости от приложений – *экономическую*, *социальную* информацию, классифицируемую по различным признакам.

Информационное пространство предприятия (организации) должно включать в себя весь спектр различных видов информации, отображающей состояние и функционирование конкретного предприятия или организации:

- сведения о состоянии производства (или обслуживания) и его обеспечении материальными, финансовыми и иными ресурсами, представляемые в виде данных или фактов (фактографическая информация);
- научно-техническая информация, представляемая в виде текстовых документов (документальная информация);
- нормативно-правовая информация, также представляемая в виде текстов, в которых нужно искать элементы при этом элементы – нормы, статьи, т.е. фрагменты текста;
- нормативно-техническая, нормативно-методическая, маркетинговая, мониторинговая информация, являющаяся документально-фактографической, т.е. представляемой в форме текстов, из которых нужно извлекать конкретные данные о производственных нормативах, состоянии среды, сведения о наличии и ценах товаров на рынках и т.п.

Разными исследователями предлагались различные способы классификации информационного обеспечения.

Так, с точки зрения взаимодействия предприятия (организации) с окружающей средой всю информацию принято делить на *входящую* и *исходящую*. В зависимости от сроков хранения различают *постоянную*, *условно-постоянную* (иногда обновляемую) и *переменную* (регулярно изменяющуюся). Классифицируют информацию и *по уровням управления* (корпоративная, внутризаводская, цеховая, внутрицеховая), *по характеру деятельности* (конструкторско-технологическая, бухгалтерская, учетно-отчетная, плановая и т.п.).

В автоматизированных системах информационное обеспечение делят на *машинное* (в памяти компьютера) и *внемашинное* (на бумажных носителях).

Различные классификации предлагались и использовались в системах управления, как правило, для информации, создаваемой и хранящейся в форме документов (приказов, планов, писем, справочно-табличных форм статистической отчетности и т.п.).

Однако по мере развития автоматизированных средств появилась возможность регистрации и хранения информации в виде отдельных фактов (характеристик предметов, событий, операций и т.п.), т.е. в виде массивов *фактографической* информации, в которых данные могут сортироваться по различным признакам и выводиться в различных формах, удобных для решения той или иной управленческой или проектной задачи.

При создании банков данных или баз данных фактографической информации использовались более разнообразные классификации. Например, в создаваемых для отображения производственного процесса банках данных формировались БАНК, СИОД¹: главный предметный массив (ГП), массив состава изделия (СИ), массив рабочих мест или рабочих центров (РМ, РЦ), массив пооперационных трудовых нормативов (ПТН), т. е. выделялись группы данных, характеризующих предмет, средства, условия труда при производстве конкретных изделий. При создании информационных массивов о кадрах (сотрудниках предприятия) данные классифицировали в соответствии с личным делом работника, выделяя социально-демографическую, производственную, общественную и другую информацию о личности.

Разделение на документальную и фактографическую информацию носит более принципиальный характер, когда речь идет о научно-технической информации (монографиях, статьях, отчетах, патентах, законодательных актах и т. п.). Такая информация формируется человеком всегда в виде текстов, т. е. в форме документальной информации, а тексты (даже относительно структурированные) имеют ряд принципиальных особенностей (синонимия, омонимия, парадоксы), которые затрудняют извлечение из них фактографической информации, необходимой для решения проектных или управленческих задач.

При этом иногда усложняется собственно понятие факта, который должен представлять собой минимальную единицу информации, принципиально значимую для того или иного вида деятельности. В частности, применительно к нормативно-правовой информации фактом следует считать не отдельный термин, а юридически значимую норму, регламентирующую управленческую деятельность, и эта форма факта должна иметь определенную структуру. Иными словами, при фактографическом анализе текста речь идет о расчленении целостного объекта (текста) на минимально значимые части (факты), и здесь следует учитывать рекомендации о расчленении системы на элементы, т. е. применять системный анализ.

Для научно-технической информации также введены и используются различные классификации: по видам источников документальной информации (первичные и вторичные), видам информационных изданий (общегосударственные, отраслевые, внутрифирменные).

В последнее время появился еще один важный вид информации – *программные продукты* (программное обеспечение, сертифицирован-

¹ *Келехсаев А.А.* Системы интеграции и обработки данных СИОД1, СИОД2 / А.А. Келехсаев, А.П. Беляев. – М.: 1977. – 208 с.

ное и предложенное к распространению, в том числе и к продаже). Это программное обеспечение для станков с ЧПУ, автоматических линий и т. п., компьютерные программы для планирования, учета и обслуживания производственных процессов, программная поддержка научных исследований, менеджмента и маркетинга, программные средства автоматизации поиска информации, анализа текстов и т. п.

При этом научно-техническая информация и программные продукты могут быть ориентированы не только на внутреннего потребителя, но и быть предметом экспорта. Для более эффективного удовлетворения производственных потребностей сотрудников предприятия целесообразно разрабатывать классификаторы не только по видам информации, но и по ее содержанию, т. е. по функциям производственного процесса, организационного управления и т. д. В частности, прежде всего, полезно структурировать разнообразную информацию по ее назначению.

Пример структуризации информации для производственной системы приведен на рис. 2.4, на котором информация структурирована по сферам деятельности предприятия: производственная информация, информация для научно-исследовательской деятельности, информация для организационного управления, научно-техническая информация.

Разработка структуры информационного обеспечения предприятий или организаций является важной и сложной задачей, от решения которой во многом зависит эффективность их деятельности.

Эту задачу следует решать с учетом конкретных особенностей предприятия (организации). При этом необходимо разрабатывать многоаспектную классификацию, которая позволит более полно охарактеризовать информационные ресурсы организации (т. е. учесть вид, характер, назначение информации, ее направленность на внутренние потребности или на экспорт и т. д.).

Классификации помогают формировать из фактографической информации документы – формы статистической отчетности, справки для руководителей различных служб системы управления и т. п.

При этом из вышесказанного следует, что информационное обеспечение предприятия – это совокупность данных, языковых средств описания данных, программных средств обработки информации, а также процедур и методов ее организации, хранения, накопления и доступа к ней.

Важно также учесть и многоаспектность проблемы организации сбора, хранения, поиска и представления информации, что можно сделать путем стратифицированного представления (по уровням) информационной инфраструктуры.

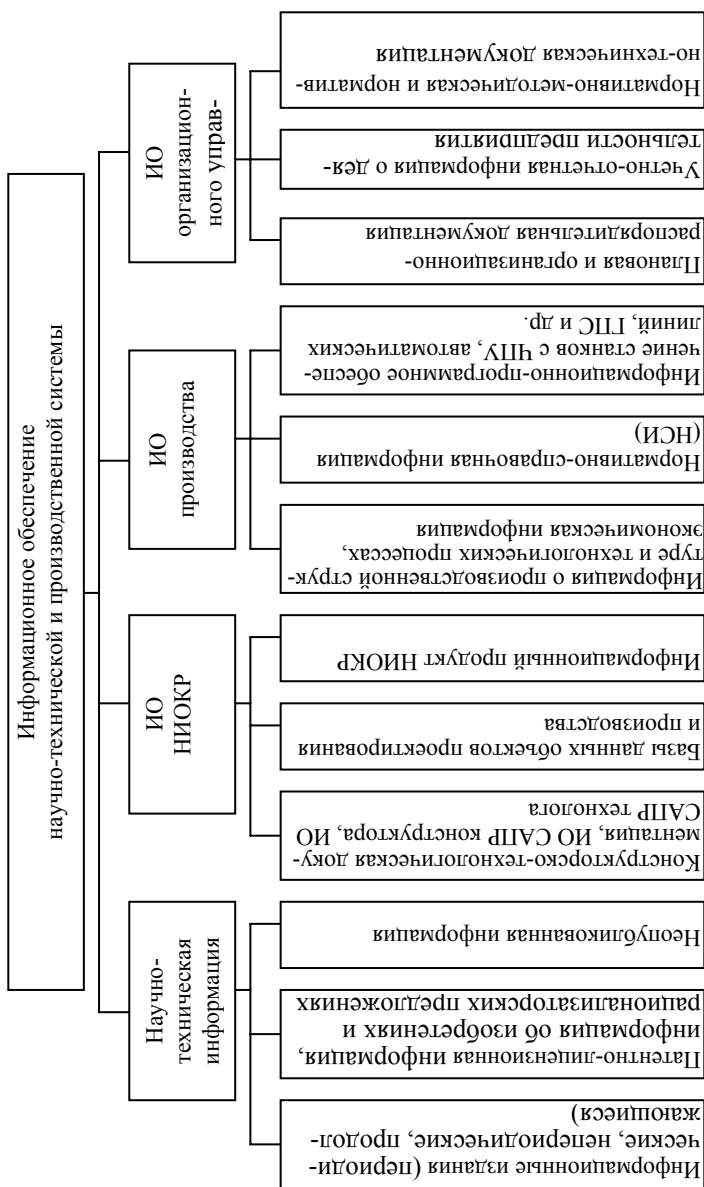


Рис. 2.4. Структура источников информации для производственной системы

Глава 3. ПОНЯТИЕ ОБ ИНФОРМАЦИОННЫХ СИСТЕМАХ И КЛАССИФИКАЦИИ ИС

3.1. Из истории разработки информационных систем

В середине XX в., когда появилась возможность обработки информации с помощью электронно-вычислительных машин (ЭВМ), был распространен термин *системы обработки данных* – СОД (термин был введен в системах радиоуправления ракетами). По мере увеличения памяти ЭВМ основное внимание стали уделять проблемам создания *баз данных* (БД.). В последующем появились термины *автоматизированные информационные системы* (АИС), информационно-поисковые системы (ИПС), информационные системы поддержки принятия решений (СППР).

Начиная с 1960-х гг. в истории развития информационных систем нашей страны относительно независимо сформировались два направления:

- разработка *автоматизированных информационных систем* как первой очереди *автоматизированных систем управления* (АСУ);
- разработка *автоматизированных систем научно-технической информации* – АСНТИ.

Работы по их созданию начались практически одновременно.

Первое направление – разработка АИС и АСУ – было инициировано научно-техническим прогрессом и возникшими в связи с этим проблемами организационного управления.

Зарубежная практика шла по пути разработки отдельных программных процедур для бухгалтерии, учета материальных ценностей и т.п., и основные работы проводились в направлении исследования и совершенствования возможностей вычислительной техники, разработки средств, обеспечивающих наиболее рациональную организацию информационных массивов, удобный для пользователя интерфейс, наращивание памяти компьютера и повышение скорости обработки информации.

В нашей стране задача обеспечения информацией управленческих работников с самого начала поставлена системно. Была разработана классификация АСУ, ориентированная на разные уровни управления предприятиями, включая их структурные подразделения, отраслями и государством в целом. Разрабатываемые отраслевые АСУ (ОАСУ), региональные или территориальные АСУ (РАСУ, ТАСУ) и АСУ предприятий (АСУП) объединялись в Общегосударственную автоматизированную систему (ОГАС) для сбора, хранения и поиска информации.

Для управления разработками автоматизированных систем были подготовлены и изданы соответствующие руководящие методические материалы ¹, в которых отражались теория и основная терминология, обязательная для использования при представлении отчетных материалов по разработке и внедрению АИС и АСУ.

Следует отметить, что АИС создавались как *фактографические информационные системы* с представлением информации пользователям в виде регламентированных форм, в которых фактографическая информация была сгруппирована в соответствии с алгоритмами решаемых на ее основе прикладных задач управления.

Второе направление – разработка АСНТИ – инициировалось научной общественностью как объективная необходимость возрастания роли информации во всех сферах деятельности, и прежде всего, в науке и технике.

Зарубежные работы в этой области первоначально проводились в форме создания разрозненных локальных центров сбора и анализа информации, но в последующем стала наблюдаться тенденция к их объединению, по крайней мере, в рамках одной отрасли.

В нашей стране задача обеспечения научно-исследовательской и управленческой деятельности научно-технической информацией сразу, так же, как и при создании АСУ, была поставлена системно. Были подготовлены и утверждены Техническое задание ² и Единый порядок ³ разработки общегосударственной *Автоматизированной системы научно-технической информации* – АСНТИ, в дальнейшем – Государственной автоматизированной системы научно-технической информации (ГАСНТИ) в стране [15, 31]), в структуре которой предусмотрены общегосударственные, отраслевые и региональные органы НТИ, отделы или бюро научно-технической информации (ОНТИ, БТИ) на предприятиях, в научно-исследовательских институтах и в других организациях.

Для исследования проблем создания систем НТИ, семантического анализа текста, реферирования и перевода с одного языка на другой развивалась теория научно-

¹ **Общепромышленные** руководящие методические материалы по созданию автоматизированных систем управления предприятиями и производственными объединениями (АСУП). – М.: Статистика, 1977. – 284 с.; **Общепромышленные** руководящие методические материалы по созданию многоуровневых интегрированных АСУ производственным объединением (предприятием). – М.: ГКНТ, 1986. – 22 с. И др.

² **Техническое** задание на совершенствование Единой системы научно-технической информации в стране. – М.: ГКНТ СМ СССР, 1972. – 34 с.

³ **Единый** порядок и технические условия разработки и внедрения автоматизированных подсистем обработки, поиска, хранения, выдачи и передачи информации. – М.: ГКНТ при СМ СССР, 1972. – 35 с.

технической информации, базирующаяся на применении математической лингвистики и семиотики. Был введен ряд понятий – *информационно-поисковая система* (ИПС) и ее структура, *информационно-поисковый язык* (ИПЯ), *дескриптор*, *тезаурус*, *грамматика*, *семантика*, понятия *релевантности* и *pertинентности* информационного поиска как характеристики ИПС, на основе которых их можно сравнивать и выбирать, и другие понятия, которые играют важную роль в развитии теории и практики информационного обеспечения предприятий и организаций.

Благодаря возросшим возможностям вычислительной техники, и особенно в результате появления персональных ЭВМ, рассмотренные направления сближаются. По мере роста объемов и усложнения структуры информационных массивов АИС возникает необходимость в заимствовании понятий теории информационного поиска, разработанной ранее для систем НТИ. В свою очередь, в решении проблемы разработки документальных АС НТИ полезны понятия *функциональной части* (ФЧ) и *обеспечивающей части* (ОЧ), введенные в нормативных документах по разработке АИС и АСУ. Кроме того, для управления современными предприятиями и организациями необходим спектр информации разного вида и назначения: от *фактографической* до *документальной*.

Работы по созданию централизованных АСУ и АСНТИ были приостановлены в связи с политическими и экономическими преобразованиями 1990-1991 гг. В течение этого периода вплоть до 1999 г., в основном развивались отечественные или приобретались зарубежные локальные информационные системы и базы данных на предприятиях и в различных фирмах, создавались отдельные страницы документальной и фактографической информации в сети Интернет.

Но затем происходит постепенная адаптация к новым экономическим условиям, и в частности, информационные ресурсы и некоторые принципы организации ГСНТИ сохранились в системе, которую в настоящее время условно называют Российской ГСНТИ (РГСНТИ). Структура этой системы приведена в разделе 5.1.

При введении принципов самокупаемости и самофинансирования в период косыгинских реформ возросла роль еще одного важного вида информации – *нормативно-методической*, регламентирующей деятельность предприятий, для чего стали разрабатывать системы нормативно-методического обеспечения управления (СНМОУ) и их автоматизированный вариант – АСНМОУ. АСНМОУ занимают промежуточное место между фактографическими и документальными информационными системами, поскольку законодательные нормативно-методические документы (законы, постановления, уставы, стандарты, инструкции и т.п.) – тексты, но, как правило, достаточно хорошо структурированные, что облегчает извлечение из них фактографической информации – правовых норм, статей, функций управления и т.п. Поэтому СНМОУ разрабатываются как *документально-фактографические* информационные системы.

С возрастанием роли нормативно-правовой информации при внедрении рыночных принципов управления экономикой создаются автоматизирован-

ные системы *нормативно-правовой документации* (АС НПД) разного рода и назначения (системы «Консультант +», «Кодекс», «Гарант» и др.), которые чаще называют справочно-правовыми системами (СПС).

На новой технической базе с использованием современных технологий развиваются информационные системы (ИС) государственного управления на федеральном и муниципальном уровнях, в частности, ИС государственной статистики; ИС природных ресурсов; государственные системы экономической, финансовой и научно-технической информации, внешнеэкономической деятельности; ИС библиотечной сети РФ и другие ИС специального назначения (см. примеры таких систем в [5]).

3.2. Виды и классификации информационных систем

В настоящее время не существует единой классификации информационных систем. Это связано с тем, что различные направления (СОД, АИС и АСУ, АСНТИ) долгое время развивались относительно независимо, поэтому и классификации в каждом из направлений предлагались также независимые.

Обобщенная классификация, базирующаяся на основных видах информации – *документальной* и *фактографической*, приведена на рис. 3.1.

При классификации АСУ и АИС применяют различные признаки классификации: степень автоматизации (форма участия человека в системе при выполнении функций управления); назначение системы (тип процесса, являющегося объектом автоматизации); степень использования технических средств (ТС) человеком для принятия управленческих решений; структурированность проблем, решаемых системой; степень централизации обработки информации; уровень управления.

По степени автоматизации АСУ и АИС делят на *автоматизированные* и *автоматические*.

В *автоматизированных* системах управления ТС используются для сбора, передачи, обработки информации и выдачи управленческих решений. При этом окончательное решение принадлежит человеку. ТС выполняют весь комплекс заранее предписанных действий. Управляющая информация непосредственно передается исполнительным механизмам, регулируя их работу без участия человека. За человеком остается функция контроля исправности ТС.

По назначению системы (характеру использования) различают такие типы АСУ и АИС:

Автоматизированные системы управления технологическими процессами (АСУТП), *системы организационного или административного управления* (АСОУ), *автоматизированные системы научных исследований* (АСНИ), *системы автоматизированного проектирования* (САПР).

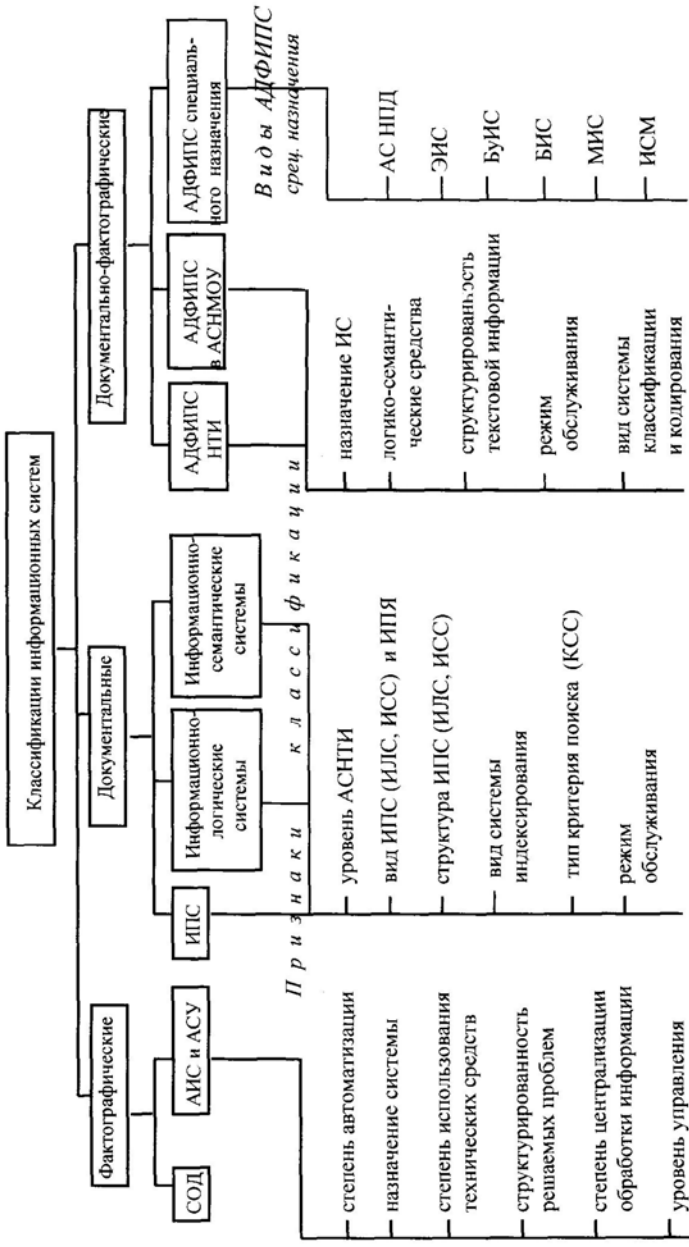


Рис. 3.1. Классификация информационных систем

По степени использования ТС человеком для принятия управленческих решений АСУ и АИС делят на:

информационные, в которых обеспечивается сбор и частичная систематизация первичной информации, и

управляющие, которые кроме этого обеспечивают выработку некоторых управленческих решений, передающихся непосредственно работникам или исполнительным механизмам по заранее заданным программам.

Информационные системы в свою очередь делят на *информационно-справочные*, которые выполняют задачу обеспечения руководства необходимыми справочными данными по запросам, и *информационно-советующие*, в которых кроме сбора, передачи и обработки информации подготавливаются рекомендации, используемые при принятии решений.

Управляющие системы делят на *информационно-управляющие* (например, система управления проектами), *управляющие системы с запрограммированными командами*, в которых решаются задачи регулирования (например, АСУТП), *самоадаптирующиеся* и *самообучающиеся* системы, функционирование которых меняется в зависимости от воздействия внешней среды.

При классификации ИС можно учитывать *структурированность проблем*.

Задачи, решаемые в системе, можно считать *структурированными*, если их решения носят повторяющийся характер (учет материалов, расчет заработной платы и др.). Часть АИС предназначена для решения таких задач. *Плохо структурированные* задачи – это такие, характер решения которых представляется не полностью определенным. Информационные системы, используемые для их решения, делят на системы, ориентированные на обработку данных и формирование специальных управленческих отчетов, и системы, в которых сведения из отчетов используются для принятия решения на основе предлагаемых альтернатив.

По степени централизации обработки информации выделяют *системы, имеющие несколько уровней обработки информации* (характерны для крупных объектов), *системы с централизованной обработкой информации* (характерны для средних объектов), *системы коллективного пользования* (характерны для малых объектов).

По уровню управления различают системы, относящиеся к:

низшему уровню управления (АСУП – для уровня предприятий и организаций, АСОУ, АСУТП и т. д.);

среднему уровню управления (ОАСУ – отраслевые АСУ, РАСУ – республиканские и региональные АСУ или ТАСУ – АСУ территориальных органов и др.);

высшему уровню управления (ОГАС – общегосударственная автоматизированная система).

Возможны и более детальные классификации [7, 10, 11]: по уровню охвата пользователей (индивидуальные, корпоративные, региональные и т.п.); по назначению (управленческие, офисные, научно-исследовательские, редакторские); по характеру функционирования (ИС в режиме реального времени, ИС стратегического планирования); по уровню организации массивов данных и др.

Для систем *научно-технической информации* также разрабатывали различные классификации.

По уровням АСНТИ (*общегосударственные, отраслевые, региональные, ОНТИ и БТИ предприятий и организаций*).

По признаку *виды документальных ИС* информационно-поисковые системы прошли следующие стадии развития: собственно *информационно-поисковые системы* (ИПС), *информационно-логические системы* (ИЛС), *информационно-семантические системы* (ИСС).

По *режимам информационного обслуживания*: по стандартным запросам (СЗ), системы избирательного распространения информации (ИРИ), дифференцированного обслуживания руководителей (ДОР), ретроспективного поиска (РП) по произвольным запросам.

По *видам ИПЯ* различают ИПЯ, и соответственно ИПС, *без грамматики и с грамматикой*. ИПЯ классифицируют и более глубоко – по *парадигматическим отношениям, лексике языка и синтагматическим отношениям*.

Поскольку лексика и синтагматические отношения характеризуют текст, описанный на конкретном ИПЯ, а парадигматические отношения представляют собой внетекстовые смысловые отношения между лексическими единицами ИПЯ, которые устанавливаются на основании потребностей информационного поиска, то и ИПС классифицируют в зависимости от развитости парадигматических отношений ИПЯ.

По *видам критериев поиска* (критериев смыслового соответствия); по тематическому профилю комплектования; формам носителей информации; уровням интеграции лексики и другим, специфическим для ИПС НТИ признакам классификации.

По *видам структур ИПС* различают:

- *ИПС иерархической структуры*, в которых все лексические единицы ИПЯ связаны сильными парадигматическими отношениями (подчинения и соподчинения) и образуют в совокупности иерархическую классификацию.

Иерархические классификации имеют вид древовидного графа или дерева понятий. На практике их представляют в табличной форме записи.

- *ИПС фасетной структуры*, в которых лексические единицы ИПЯ предварительно группируются в фасеты, а иерархические отношения устанавливаются внутри фасетов.

Фасеты, следующие друг за другом в определенной последовательности, образуют фасетную классификацию. Преимущество фасетной структуры по сравнению с иерархической заключается в *многоаспектности*, так как количество фасетов и субфасетов в принципе не ограничивается. Важной особенностью многоаспектной классификации является также то, что последовательность признаков, и соответственно фасетов, может быть произвольной.

- *ИПС неиерархической структуры*, в которых лексические единицы ИПЯ упорядочивают по внешним признакам, например, в алфавитном порядке.

Если иерархические и фасетные классификации строят в основном по заранее заданной схеме, то неиерархические языки получают эмпирически в процессе индексирования вводимых в ИПС документов.

Следует иметь в виду, что признаки, по которым выше были классифицированы системы управления, являются независимыми, вследствие чего каждая система, отнесенная к какому-то классу по одному признаку, может соответствовать нескольким классам по остальным признакам.

Предлагались и другие классификации.

Например, для того чтобы охарактеризовать систему НТИ более полно, разрабатывались многоаспектные классификации. Наиболее развитой из таких классификаций является *фасетная* классификация *А.В. Соколова*¹, в которой признаки классификации определяются *семантическими средствами* ИПС, т.е. по видам информационно-поисковых языков, методам (правилам, алгоритмам) индексирования, методам (правилам, алгоритмам) поиска по запросу.

В последующем представления о видах автоматизированных систем развивались. По мере развития вычислительных средств и технологий сети Интернет, методов и средств информационного поиска возникают новые термины, более точно характеризующие специфику информационных систем специального назначения. В развитие информационно-советующих и информационно-управляющих систем возникли *системы поддержки принятия решений, экспертные системы*, помогающие в принятии решений по проектированию и управлению.

Параллельно с *информационно-логическими* и *информационно-семантическими системами*, усиливающими интеллектуальные возможности информационно-поисковых систем научно-технической информации, развивается направление искусственного интеллекта, базирующееся на кибернетических исследованиях процессов в живых организмах, в том числе функционирования мозга, и моделировании

¹ *Соколов А.В.* Информационно-поисковые системы: учеб. пособие / А.В. Соколов. – М.: Радио и связь, 1981. – 161 с.

этих процессов средствами вычислительной техники, и на этой основе создаются *интеллектуальные информационные системы*.

На новой технической базе с использованием современных технологий развиваются информационные системы государственного управления на федеральном и муниципальном уровнях, системы экономической, финансовой и научно-технической информации, внешнеэкономической деятельности; и другие ИС специального назначения.

В их числе [18]: экономические информационные системы (ЭИС): *бухгалтерские информационные системы* (БуИС), *автоматизированные банковские информационные системы* (АБИС), *ИС фондового рынка*, *ИС в контроллинге*, *ИС бюджетирования*, *маркетинговые информационные системы* (МИС), информационные системы мониторинга (ИСМ); *ИС в кадровом менеджменте*; *ИС в налоговых органах, в страховании и пенсионном обеспечении*; *ИС в туристическом бизнесе*; *ИС в управлении недвижимостью*; *ИС таможенной и внешнеэкономической деятельности* и т.п.; *системы сопровождения жизненного цикла изделий* типа ИПИ–системы (ИПИ – информационная поддержка изделий) и системы управления проектными и инженерными данными предприятия – так называемые PDM–системы (Product Data Management Systems), включая CAD–системы (Computer-Aided Design – компьютерная поддержка проектирования), САМ–системы (Computer-Aided Manufacturing – компьютерная поддержка изготовления), САЕ–системы (Computer-Aided Engineering – поддержка инженерных расчетов) и т.п.; *системы управления взаимоотношениями с клиентами* (CRM – Customer Relationships Management); *ИС документооборота и делопроизводства* (ИСДД).

С развитием систем телекоммуникаций стал активно развиваться информационный рынок, получивший название «пятый рынок». Развиваются информационные технологии (ИТ или IT–технологии) – способы и средства сбора, обработки, хранения и распределения информации.

Для оперативного и интеллектуального анализа хранимых данных используются технология комплексного многомерного анализа данных OLAP (On-Line Analytical Proctssing) и технология интеллектуального анализа информации Data Mining [17, 18 и др.].

На основе новых информационных технологий развиваются *процессо-ориентированные корпоративные информационные системы* – системы, обеспечивающие реализацию процессного подхода к управлению предприятием. Менеджмент бизнес-процессов возник в рамках концепции CPI (Continuous Process Improvement) и его японского аналога TQM (Total Quality Management), согласно которой предполагается управление бизнес-процессом, как единым целым. Этот подход положен в основу стандартов серии ISO 9000, которые разработаны Международной организацией по стандартизации (International Standard Organization – ISO) и определяют базовый набор принципов менеджмента качества.

Для оптимизации и автоматизации внутренней деятельности фирмы применяются ИС планирования потребности в материалах – MRP (Material Requirements Planning); ERP (Enterprise Resources Planning) системы, направленные на усовершенствование таких процессов, как планирование, изготовление, учет и контроль.

. В основе ERP-систем лежит принцип создания единого хранилища (репозитория) данных, содержащего всю корпоративную информацию, включая финансовую информацию, производственные данные, данные по персоналу и т.д. ERP-система – это набор интегрированных приложений, позволяющих создать интегрированную информационную среду (ИИС) для автоматизации планирования, учета, контроля и анализа всех основных бизнес-операций предприятия. В соответствии с современными требованиями APICS, ERP-система должна помимо ядра, реализующего стандарт MRPII, включать ряд модулей: управления логистическими цепочками SCM; усовершенствованного планирования и составления производственных графиков APS (Advanced Planning and Scheduling); CRM; электронной коммерции EC (Electronic Commerce); управления данными об изделии PDM; надстройку Business Intelligence на основе технологий OLAP и др.

В конце 1990-х гг. был разработан также стандарт CSRP (Customer Synchronized Resource Planning), который охватывает взаимодействие предприятия с клиентами: оформление заказа, техническое задание, проектирование и изготовление требуемого товара, поддержку клиентов и др. Если стандарты MRP/MRPII/ERP ориентированы только на внутреннюю организацию предприятия, то в стандарт CSRP включен полный жизненный цикл изделия: от его проектирования с учетом требований заказчика до гарантийного и сервисного обслуживания после продажи.

Специализированные информационные системы, примеры которых приведены на рис. 3.1, отнесены в этой классификации к документально-фактографическим при использовании этого термина в широком смысле, поскольку даже те из этих ИС, которые начинали разрабатываться как фактографические, в последующем, как правило, дополняются нормативно-правовой и другой документальной информацией.

В то же время по мере увеличения числа специализированных ИС, по-видимому, целесообразно выделить их в особый класс специализированных ИС, классифицируемых в соответствии с эволюцией стандартов управления предприятиями, и разработать соответствующую классификацию.

Для целей анализа и создания ИС в сфере экономики удобна используемая на Западе классификация ИС, отличающаяся степенью сложности и участия лиц, принимающих решения¹:

¹ *Матвеев Л.А.* Системы поддержки принятия решений: учеб. пособие / Л.А. Матвеев. – СПб.: СПбГУиЭФ, 1993. – 47 с.

* *Системы электронной обработки данных (СЭОД)* или просто *системы обработки данных (СОД)* –

работают в автоматическом режиме с минимальным участием человека; предназначены для решения хорошо структурированных задач, по которым имеются входные данные, известны алгоритмы решения задач.

* *Информационные системы управления (ИСУ)* –

применяются при худшей структурированности задач; имеют возможность манипулирования данными за счет появления в их составе СУБД; осуществляют поиск и обработку входной информации; выходную информацию представляют в виде специальных управленческих отчетов; все решения принимает человек; используется на уровне стратегического планирования, управленческого и оперативного контроля.

* *Системы поддержки принятия решений (СППР)* –

предназначены для решения в режиме диалога плохо структурированных задач, для которых характерна неполнота входных данных, частичная ясность целей и ограничений; обеспечивают значительное участие в работе системы человека, который может вмешиваться в ход решения, модифицировать входные данные, процедуры обработки, цели и ограничения задачи, выбирать стратегии оценки вариантов решений; включает помимо СУБД СППР базу моделей и систему управления этой базой, а также систему управления диалогом; используется на уровне стратегического планирования, оперативного и управленческого контроля.

* *Экспертные системы (ЭС)* –

основываются на моделировании процесса принятия решения человеком-экспертом с использованием компьютера и разработок в области искусственного интеллекта, на использовании не только данных и информации, но и знаний; но не включают в себя математических моделей, улучшающих принимаемое человеком решение.

* *Гибридные экспертные системы (ГЭС)* –

являются гибридом ЭС и СППР; обеспечивает доступ человека к решению задачи на любой стадии; окончательное решение принимает человек; используются на уровне стратегического планирования и управленческого контроля.

* *Информационные системы мониторинга (ИСМ)* –

предназначены для целей контроля за деятельностью фирмы, обеспечивая вышние звенья управления важной укрупненной информацией; не предназначены для помощи в принятии решений, но полезны для выявления оперативных проблем, а также при анализе разного рода управленческих ситуаций за счет обеспечения текущей и ретроспективной информации.

Вопросы построения и разработки каждого из этих видов ИС могут стать предметом самостоятельного издания. Поэтому в монографии в соответствии с поставленной выше целью выявлены и рассмотрены некоторые общие принципы построения и функционирования основных из выделенных классов ИС – фактографических ИС на примере АИС (гл. 4), документальных ИС на примере АСНТИ (гл. 5), документально-фактографических ИПС на примере АДФИПС для Автоматизированных систем нормативно-методического обеспечения управления (гл. 6).

Глава 4. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ АИС И АСУ

4.1. Принципы управления разработками АИС и АСУ

Особенности построения и организации функционирования АСУ

Системы такой сложности, как АСУ, обладают рядом специфических особенностей, присущих открытым системам с активными элементами, к которым, в частности, относятся неоднозначность использования понятий «цели – средства», «система – подсистема»; трудность прогнозирования (а иногда и принципиальная непредсказуемость) поведения системы при внесении в нее изменений. Для обеспечения адаптивности системы, способности ее к самоорганизации необходимо предусмотреть соответствующие средства, обеспечивающие целеобразование, способность вырабатывать варианты поведения, а при необходимости изменять структуру системы управления и АСУ.

Эти особенности были осознаны с самого начала разработки АСУ и обусловили необходимость привлечения для их объяснения и обеспечения системных представлений, закономерностей функционирования и развития сложных систем (см. в [1, 21, 25]).

Понимая неизбежность и необходимость проявления названных особенностей и обуславливающих их закономерностей, которые действуют в системе независимо от того, учитывают их или нет, и затрудняют управление разработками АИС (как первой очереди АСУ), некоторые специалисты уже на ранних стадиях их создания предлагали создавать системы проектирования и развития АСУ, разрабатывать единые принципы проектирования и терминологию. Были разработаны типовые положения, типовые структуры, порядок разработки и другие методические материалы, объединенные затем в единый документ – *Общепромышленные руководящие методические материалы (ОРММ)*.

С самого начала в конце 1960-х гг. при проектировании АСУ любого уровня (общегосударственного, регионального, предприятия и организаций, их подразделений и т.п.) возникает необходимость разделить сложную систему на части, чтобы распределить работы между исполнителями, занимающимися разработкой и организацией функционирования АСУ.

Столкнувшись со сложностью выбора первоочередных объектов и направлений автоматизации, распределения ресурсов на разработку их

программного обеспечения, в 1970-е гг. для управления разработками АСУ и АИС стали создавать системы проектирования и развития АСУ, разрабатывать единые принципы проектирования и терминологию.

В понимании принципов построения и организации функционирования АСУ большую роль сыграло выделение *функциональной* и *обеспечивающей* частей (ФЧ и ОЧ), что не утрачивает актуальности и в настоящее время при разработке любых автоматизированных информационных и управляющих систем.

Понятие о структуре функциональной и обеспечивающей части

Эти понятия были введены по аналогии с понятиями «цель» и «средства», выделяемыми при постановке любой задачи принятия решения.

Структура ФЧ определяется на основе анализа целей и функций системы управления, для обеспечения деятельности которой создается АСУ, т.е. ФЧ определяет *цели* и основные *функции* АСУ и представляет собой совокупность *подсистем* и *задач*, выбранных для автоматизации.

Структура ОЧ включает различные виды обеспечения (организационное, информационное, техническое, программное, лингвистическое, эргономическое и т.п.), необходимые для реализации подсистем и задач ФЧ АСУ, т.е. ОЧ представляет собой *средства* для достижения целей, реализации задач АСУ.

При этом под структурой ОЧ понимается не просто совокупность средств информационного, технического, алгоритмического, программного и других видов обеспечения, а организационную форму взаимодействия всех видов обеспечения, необходимых для реализации подсистем и задач (входящих в структуру ФЧ) на всех уровнях иерархии структуры ОЧ.

Необходимость введения новых терминов – ФЧ и ОЧ – вместо привычных понятий «цели» и «средства» объясняется особенностями и закономерностями сложных многоуровневых систем.

В таких системах каждый уровень иерархической структуры ведет себя как «двуликий Янус» – как *средство* по отношению к вышестоящему уровню и как *цель* по отношению к нижележащему, т.е. имеет место *закономерность коммуникативности* (см., например, в [1, 21, 25]). Соответственно составляющие любого промежуточного уровня в структуре АСУ можно рассматривать как подсистемы по отношению к вышестоящему уровню, к системе в целом, их объединяющей, и в то же время, взятые сами по себе, они могут рассматриваться как системы. Поэтому иногда при разработке АСУ возникали противоречивые ситуации; отдельные подсистемы «Управление кадрами», «Управление качеством» и т.п. на определенной стадии развития АСУ начинали называть «АСУ-Кадрь», «АСУ-Качество», т.е. считали их как бы самостоятельными системами, в то время, как по отношению к общей системе управления предприятием они продолжали оставаться подсистемами.

Более того, на каждом уровне иерархической структуры ФЧ в силу *закономерности иерархической упорядоченности* всегда следует определять свои цели, функции, задачи, средства. Иными словами, понятия «цель» и «средств-

ва» в иерархической структуре всегда используются неоднозначно, и это должны осознавать разработчики АСУ, не тратя времени на бессмысленные дискуссии по поводу терминов.

Следует также оговорить, что введенные термины ФЧ и ОЧ нельзя однозначно отождествлять с понятиями «цели» и «средства». Они являются более сложными понятиями.

Действительно, иногда говорят об *обеспечивающих подсистемах*, имея в виду определенным образом организованные совокупности информационных, программных, технических средств, используемых для автоматизации не укрупненной функции (для которой используется термин «функциональная подсистема»), а какой-либо вспомогательной функции нижележащих уровней или определенной их совокупности. Например, такого рода обеспечивающими подсистемами были банки данных определенного назначения, типа ИНЭС¹, СИОД², БАНК и т. п., используемые как средства для хранения и предоставления лицам, принимающим решения, информации о состоянии производства. В качестве такого рода подсистем в настоящее время можно рассматривать программные продукты для автоматизации отдельных видов деятельности (бухгалтерского учета, материально-технического снабжения и т. п.).

С другой стороны, когда сдают в эксплуатацию *функциональную подсистему*, то имеют в виду не только совокупность задач и функций, включаемых в нее, но и технические средства, алгоритмы, программы, инструкции по их использованию, т. е. и совокупность средств реализации этой подсистемы.

Таким образом, термины ФЧ и ОЧ являются обобщающими условными понятиями, которые помогают охарактеризовать автоматизированную систему как целое, выделив в ней работы, в большей мере связанные с анализом и описанием целей системы (формирование структуры ФЧ АСУ), и работы, связанные с определением общей структуры средств реализации целей (разработка структуры ОЧ АСУ), что помогает организовать управление разработками АСУ и АИС.

4.2. Принципы организации управления разработками АСУ

Как было показано выше, в понимании принципов построения и организации функционирования АСУ большую роль играет выделение *функциональной и обеспечивающей* частей (охарактеризованных в предыдущем разделе), что не утрачивает актуальности и в настоящее время при разработке любых автоматизированных информационных и управляющих систем.

¹ *Авен О.И.* Что же такое АСУ / О.И. Авен. – М.: Наука, 1981. – 175 с.

² *Келехсаев А.А.* Системы интеграции и обработки данных СИОД1, СИОД2 / А.А. Келехсаев, А.П. Беляев. – М.: 1977. – 208 с.

Соответственно при управлении разработками автоматизированных систем выделяют две основные проблемы:

1. Формирование структуры ФЧ АСУ и выбор на ее основе первоочередных задач автоматизации.

2. Формирование структуры ОЧ АСУ как организационной формы взаимодействия всех видов обеспечения (информационного, технического, организационного и т.д.), необходимых для реализации подсистем и задач, входящих в структуру ФЧ АСУ.

Решение этих двух основных проблем взаимосвязано. С одной стороны, структура ОЧ определяется структурой ФЧ. В то же время выбор структуры ФЧ во многом зависит от имеющихся технических, программных и иных средств, т.е. определяется потенциальными возможностями ОЧ.

Основные проблемы управления разработками АСУ, в свою очередь, делят на задачи, которые часто можно решать параллельно.

Например, первую проблему делили следующим образом:

1.1. Прогнозирование структуры ФЧ АСУ. Разработка прогнозного варианта структуры ФЧ АСУ (на 15–20 лет) и основных направлений развития АСУ (на 5–10 лет).

1.2. Разработка структуры ФЧ АСУ последующей очереди (на 3–5 лет). Эту задачу называли также «Выбор первоочередных подсистем (комплексов задач) автоматизации для последующей очереди АСУ».

1.3. Выбор первоочередных (наиболее значимых) задач в подсистемах АСУ и последовательности их проектирования и внедрения.

1.4. Проектирование подсистем АСУ.

Вторую проблему можно представить следующими подпроблемами:

2.1. Выбор (обоснование) структуры ОЧ АСУ.

2.2. Проектирование отдельных видов обеспечения.

Кроме того, в действующей АСУ важно создавать *организационную структуру АСУ*, которая определяется составом и взаимосвязью отдельных структурных подразделений в условиях эксплуатации АСУ. Поэтому к названным основным проблемам в последующем была добавлена проблема управления разработками АСУ.

Систему управления разработками АСУ можно считать 3-й проблемой и разделить ее на следующие задачи:

3.1. Разработка структуры организационного обеспечения управлением разработками АСУ (на первых этапах эта задача сводится к определению структуры подразделения, разрабатывающего АСУ, а по мере развития – к определению взаимоотношений между подразделениями-разработчиками и подразделениями, использующими результаты разработки в практической деятельности, а также

подразделениями, подготавливающими и контролирующими вводимую информацию). Иногда эту составляющую выделяют в самостоятельную проблему.

3.2. Создание информационной системы для обеспечения проектирования подсистем и задач АСУ.

Структура проблем разработки АСУ приведена на рис. 4.1.

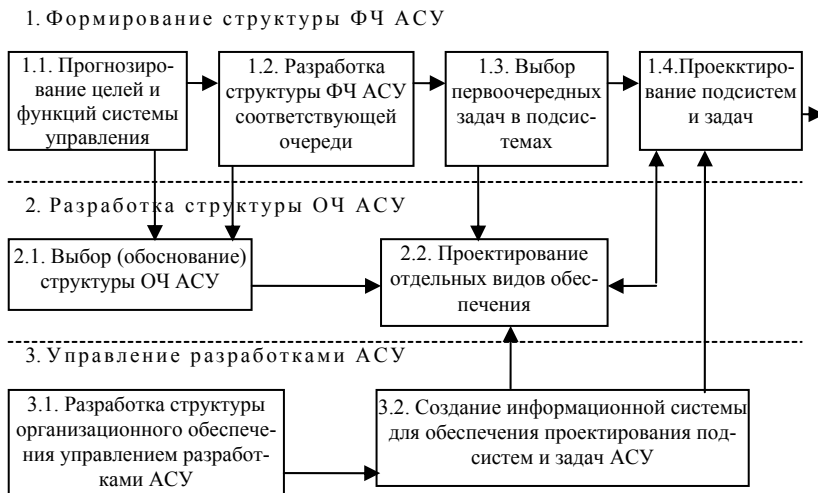


Рис. 4.1. Этапы разработки АСУ

Для управления разработками автоматизированных систем были подготовлены соответствующие руководящие методические материалы, в которых АСУ трактовалась как развивающаяся система и вводилось понятие *очереди*. АСУ первой очереди разрабатывалась как *автоматизированная информационная система* – АИС. В руководящих материалах оговаривался также порядок разработки соответствующей очереди АСУ (АИС) и ввода ее в эксплуатацию (рис. 4.2).

При реализации приведенной на рис. 4.1 структуры ориентировались на методику, приведенную на рис. 4.2. Получаемая структура ФЧ АСУ и выбор состава ОЧ проводились на основе обследования существующей системы управления предприятием и технико-экономического обоснования. (ТЭО) и являлись основой разработки технического задания (ТЗ). После согласования и утверждения ТЭО и ТЗ выполнялся технический проект (ТП) и рабочий проект (РП) и ввод разработанных подсистем и задач эксплуатацию.

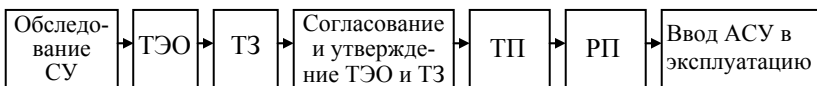


Рис. 4.2. Порядок разработки АСУП

Практика показала, что разработку АСУ невозможно реализовать на основе одного проекта. Было введено понятие очереди АСУ и в процессе реализации предшествующей очереди начиналась разработка структуры ФЧ и ТЗ следующей очереди (рис. 4.3), т.е. применялась одна из важных закономерностей теории систем – *закономерность историчности* (см., например, в [1, 21, 25]).

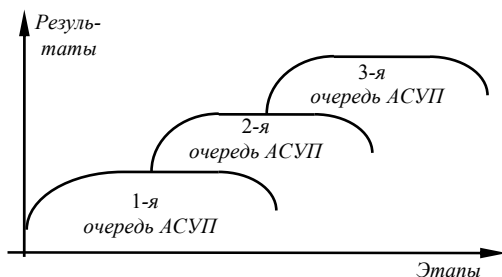


Рис. 4.3. Проявление закономерности историчности при разработке АСУ

При разработке каждой последующей очереди АСУ последовательность этапов, приведенная на рис. 4.2, повторялась. Более того, практика показала, что в период одной очереди при разработке подсистем этапы, приведенные на рис. 4.2, также повторялись применительно к подсистеме.

Были разработаны типовые положения, типовые структуры, порядок разработки и другие методические материалы, объединенные затем в единый документ – *Общотраслевые руководящие методические материалы (ОРММ)*, регламентирующий ход разработки АСУ и ОАСУ.

Четкий порядок, установленный ОРММ, ускорил распространение опыта организации работ по проектированию АСУ, облегчил учет и сравнительный анализ хода работ по созданию АСУ на предприятиях и в отраслях

Однако одновременно ограничил развитие АСУ ряда предприятий и НПО, что неизбежно в силу закона *«необходимого разнообразия»* У.Р. Эшби (см. в [1, 21, 25]): именно за счет ограничения «разнообразия», т.е. типизации, и в конечном счете примитивизации систем, было достигнуто упрощение централизованного управления разработками АСУ.

Негативная роль ОРММ была в последующем осознана, особенно в период перехода предприятий на самокупаемость и хозрасчет. Ти-

повые решения и структуры можно использовать лишь на начальных стадиях создания АСУ, а по мере их развития все больше начинают проявляться индивидуальные особенности конкретных предприятий и объединений и связанная с этим индивидуальность АСУ.

В этих условиях для управления разработками АСУ требуются методические материалы, в которых не только определялось бы, *что* нужно делать в процессе разработки АСУ и диктовались бы готовые типовые проектные решения, а давались бы рекомендации о том, *как* нужно принимать решения по выбору структуры АСУ, средств ее реализации в конкретных условиях, необходимо разрабатывать соответствующие методики и модели для принятия решений по управлению разработками АИС и АСУ с учетом конкретных особенностей предприятий и организаций.

Очевидно, что принципы построения и эффективность АСУ существенно зависят от уровня развития информационных технологий.

С появлением в середине 1980-х гг. персональных ЭВМ происходит корректировка идеи АСУ от ВЦ и централизации управления – к распределенному вычислительному ресурсу и децентрализации управления. При этом уменьшается нагрузка на централизованные вычислительные ресурсы и верхние уровни управления, что позволяет сосредоточить в них решение крупных долгосрочных стратегических задач.

В то же время для обеспечения эффективного управления крупными предприятиями остается актуальной идея создания *интегрированных АСУ*, а для обеспечения информацией по группам основных функций организационного управления предприятий, корпораций – *корпоративных информационных систем* – КИС.

4.3. Виды и принципы разработки структуры функциональной части АСУ

Виды структур функциональной части АСУ

Первоначально основной формой представления структуры ФЧ АСУ была древовидная иерархическая структура. АСУ делилась на подсистемы (или комплексы задач), подсистемы – на группы задач, а последние, в свою очередь, – на отдельные задачи.

В *Общепромышленных руководящих методических материалах по созданию АСУП*¹, утвержденных Постановлением Государственного коми-

¹ *Общепромышленные* руководящие методические материалы по созданию автоматизированных систем управления предприятиями (АСУП)). – Минск, 1971. – 376 с.

тега по науке и технике Совета Министров СССР 10 мая 1971 г. № 145. типовая структура ФЧ первоначально включала 7 подсистем (рис. 4.4).



Рис. 4.4. Типовая структура ФЧ АСУП

Затем типовая структура ФЧ АСУП была дополнена подсистемой управления качеством, а в последующем постепенно расширялась. Подсистемы в первых структурах ФЧ АСКП, а затем – в типовой структуре ФЧ АСУП были ориентированы на основные функциональные подразделения существовавших систем организационного управления, откуда и произошел термин *функциональная подсистема*

Число подсистем в АИС первой очереди конкретных предприятий было разным, но обычно не превышало числа подсистем типовой АСУП (7–8), рекомендованной в Общеотраслевых руководящих методических материалах. Число задач по подсистемам колебалось довольно в широких пределах, в зависимости от принятой детализации задач и конкретных особенностей предприятия.

В 80-90-е гг. XX в. в связи с необходимостью объединения в единую систему разнообразных автоматизированных систем, разрабатывавшихся на предприятии возник и был широко распространен термин «Интегрированные автоматизированные системы управления промышленными предприятиями».

В соответствии с объективными процессами материального производства и управления создавались информационные системы управления следующих основных типов (рис. 4.5):.

- автоматизированные системы научных исследований (АСНИ);
- системы автоматизированного проектирования изделий (САПР) и автоматизированные системы технологической подготовки производства (АС ТПП);
- автоматизированные системы общезаводского управления, ориентированные на автоматизацию функций управления объединением (корпорацией, трестом, концерном и др.) (АСУО) и предприятием (АСУП);

- автоматизированные системы для комбинированного организационного и технологического управления (АСУ ОТ);
- автоматизированные системы управления гибкими производствами (АСУ ГПС), включая АСУ отдельными производствами, цехами, участками, гибкими автоматизированными линиями (ГАЛ), гибкими производственными модулями (ГПМ);
- автоматизированные системы управления технологическими процессами (АСУ ТП);
- автоматизированных систем контроля качества и испытания изделий (АСКИ).

Объединить эти полсистемы помогла структура ФЧ, приведена на рис. 4.5 [33].



Рис. 4.5. Виды производственных информационных систем

Однако по мере развития АСУ число подсистем увеличивалось. При возрастании числа подсистем нарушается важное требование к иерархическим структурам – гипотеза Миллера, согласно которой число составляющих одному узлу должно быть 7 ± 2 .

Трудности управления разработками АСУ при существенном увеличении числа подсистем, т. е. числа составляющих на верхнем уровне иерархической структуры ФЧ, привели к тому, что на некоторых предприятиях стали вводить еще один обобщающий уровень, который был предложен на Волжском автозаводе (история развития структуры ФЧ АСУ ВАЗа приведена на рис. 4.6), и назван *направлениями*.

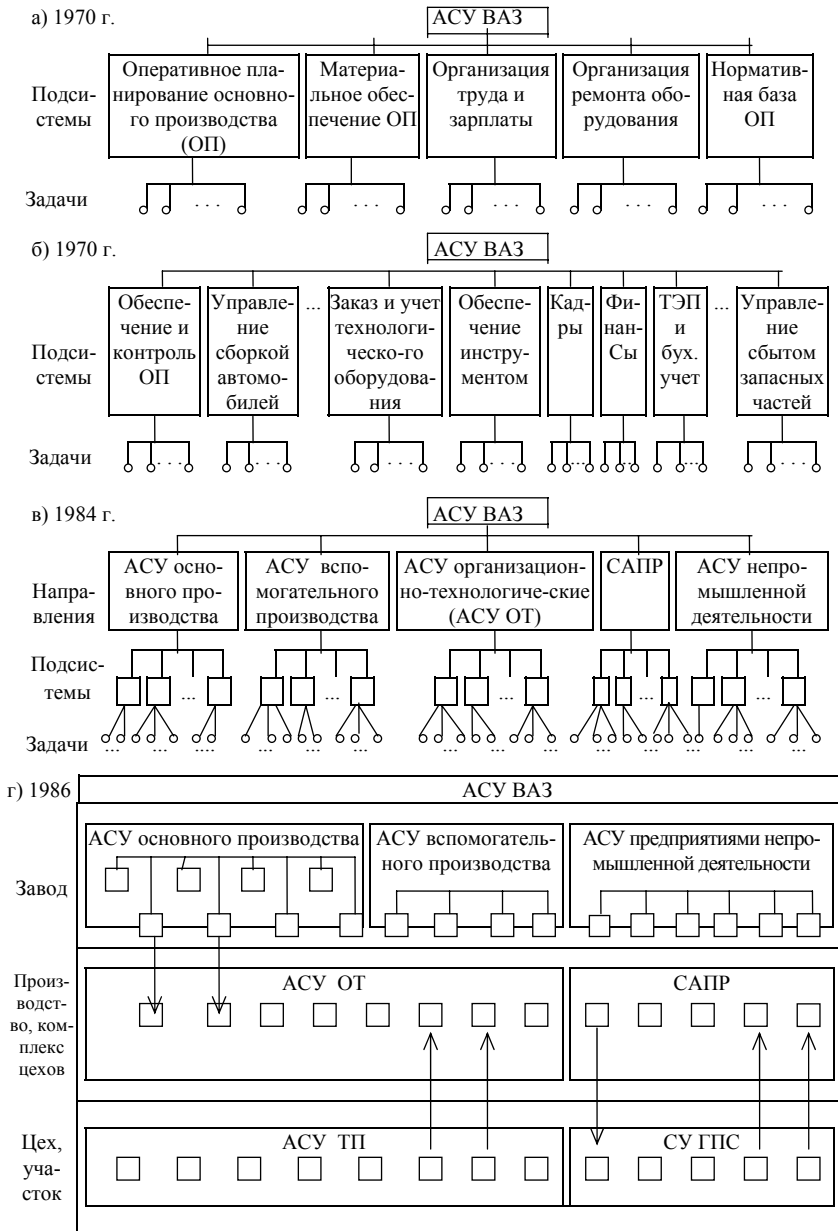


Рис. 4.6. Развитие структуры ФЧ АСУ ВАЗа

Для иллюстрации развития структуры ФЧ АСУ выбран пример ВАЗа, на котором создана наиболее развитая АСУ, являющаяся основой функционирования производства.

Направление объединяло несколько подсистем, связанных между собой. В принципе можно было бы укрупнить подсистемы, однако к тому времени в практике разработок АСУ с понятием подсистемы был связан вполне определенный объем работ, поэтому более удобным оказалось ввести новый термин. Пример направлений, выделенных на ВАЗе в 1984 г., приведен на рис. 4.6, в).

В дальнейшем по мере развития АСУ появилась (впервые на Волжском автозаводе) новая форма автоматизированных систем – *организационно-технологические АСУ (АСУ ОТ)*, которые создавались для повышения степени оперативного управления отдельными участками производства. АСУ ОТ сочетают функции управления технологическими процессами (АСУ ТП) и функции организационного управления производством (АСУП), причем большая часть массивов организационного управления формируется на основе массивов управления техпроцессами (что возможно только при практически полной автоматизации технологических процессов). Подсистемы АСУ ОТ возникали либо на основе АСУ ТП, либо – как развитие подсистем АСУП.

Вначале АСУ ОТ рассматривали как одно из направлений. Но иерархический принцип представления имеет в числе своих специфических особенностей стремление к обособленности выделенных ветвей, т.е. в иерархических структурах сильны вертикальные связи и предельно ослаблены горизонтальные взаимосвязи между направлениями, подсистемами, которые после их выделения в самостоятельные начинают развиваться независимо, как бы соперничая друг с другом. Поэтому, если включить АСУ ОТ в структуру ФЧ АСУ как равноценное направление с АСУ ТП и АСУП, то по мере самостоятельного развития этих направлений возможно дублирование работ в АСУ ОТ и АСУ ТП или в АСУ ОТ и АСУП. Для того, чтобы усилить взаимодействие этих трех направлений, их оказалось удобнее поместить в структуре друг под другом.

В этом случае устанавливаются вертикальные согласующие взаимосвязи между подсистемами и задачами этих направлений, и по мере развития подсистем АСУ ТП в направлении усиления функций организационного управления переводить их на уровень АСУ ОТ, а по мере усиления функций управления технологическими процессами в подсистемах АСУП – на уровень АСУ ОТ.

Аналогичные взаимоотношения целесообразно установить между подсистемами АСУП, САПР и СУ ГПС.

Сформированная в результате *многоуровневая структура* ФЧ АСУ ВАЗа приведена на рис. 4.6, г).

Аналогично была сформирована и структура АСУ объединения АвтоВАЗ, в которой к приведенным на рис. 4.6, *з* уровням добавился еще уровень объединения, а на нижележащей по отношению к нему страте, наряду с АСУ ВАЗа, были помещены АСУ других предприятий, входящих в объединение, структуры ФЧ АСУ которых формировались аналогично АСУ ВАЗа.

Приведенные примеры структур ФЧ АСУ иллюстрировали развитие структуры по мере развития предприятий и их АСУ: от первоначального представления структур ФЧ в виде древовидных иерархических структур до структур ФЧ АСУ в виде *многоуровневой, стратифицированной*.

Для разработки структур ФЧ АСУ конкретных предприятий применялись подходы и методы системного анализа – методики структуризации целей и функций систем управления и методы их оценки, и выбора на этой основе подсистем и задач, подлежащих автоматизации при разработке соответствующей очереди АСУ.

Принципы и примеры разработки структур ФЧ АСУ

Применение методики, базирующейся на двойственном определении системы. А.И.Уёмов, определяя систему через понятия «вещи», «свойства», «отношения», предложил двойственное определение¹, в одном из которых свойства q_i характеризуют элементы a_i , а в другом – свойства q_j характеризуют связи (отношения) r_j (1.1, *e*):

$$S \equiv \left[\begin{array}{l} \{a_i\} \ \& \ \{r_j(q_j)\} \\ \text{def} \quad a_i \in A \quad r_j \in R \quad q_j \in QR \end{array} \right], \quad (4.1)$$

$$S \equiv \left[\begin{array}{l} \{a_i(q_i)\} \ \& \ \{r_j\} \\ \text{def} \quad a_i \in A \quad q_i \in QR \quad r_j \in R \end{array} \right].$$

В работах **Б.Д. Кошарского** (см. ссылки в [1, 21]) было показано, что этим определениям соответствует два способа представления системы управления:

п р о ц е д у р н о е – как множества объектов A , на котором реализуется заранее определенные отношения R с фиксированными свойствами QR ; при этом, если системообразующее отношение определено во времени, то это представление соответствует структуризации системы по циклу управления: планирование, организация, регулирование, учет и т.п. (набор функций цикла управления изменяется по мере развития экономики и зависит от конкретных условий);

ф а к т о р н о е – как множества объектов A , обладающее заранее определенными свойствами QA с фиксированными между ними отношениями R ; при этом могут быть выделены такие составляющие, как основное и вспомогательное производство, основные и оборотные фонды, трудовые ресурсы, материально-техническое обеспечение и другие объекты управления на предприятии (набор их также определяется конкретными условиями).

¹ **Уёмов А.И.** Системный подход и общая теория систем / А.И. Уёмов. – М.: Мысль, 1978. – 272 с.

Б.Д. Кошарский показал, что каждый из этих способов представления системы в отдельности дает неполное описание системы управления, а для выявления системных особенностей конкретного предприятия необходимо один способ описания дополнить другим, двойственным ему, т.е. что только совместное использование процедурного и факторного представлений системы позволяет обеспечить конкретизацию и полноту анализа целей и задач организационного управления.

Отметим, что это утверждение о полноте справедливо лишь в рамках принятой концепции системы.

Такое требование на практике реализуется либо путем формирования и анализа матрицы «цикл управления–объект управления» (табл. 4.1), после оценки которой формируются двойственные структуры и осуществляется выбор из них наилучшей; либо путем параллельного формирования двойственных вариантов структуры (рис. 4.8), с использованием взаимно обратной последовательности признаков структуризации.

Таблица 4.1

Матрица «цикл управления–объект управления»

Объект управления (ОбУ) Производственный признак	Цикл управления (ЦУ) Временной признак					
	Прогнозирование	Перспективное планирование	Организация	Текущее планирование	Оперативное управление	Учет Контроль Анализ
	<i>ПП</i>	<i>ПП</i>	<i>ОРГ</i>	<i>ТП</i>	<i>ОУ</i>	<i>УКА</i>
Научно-исследовательская работа (НИР)	–	–	–	–	–	–
Производство основной продукции (ОП)	+	+	+	+	+	+
Вспомогательно-обслуживающее производство (ВИП)	–	–	–	+	–	+
Транспорт (Т)	–	–	–	–	–	–
Материально-техническое снабжение (МТС)	–	–	–	+	+	+
Трудовые ресурсы (кадры – К)	–	–	–	–	–	+
Сбыт продукции (СбП)	–	–	–	+	–	+
Финансы (Ф)	–	–	–	–	–	–

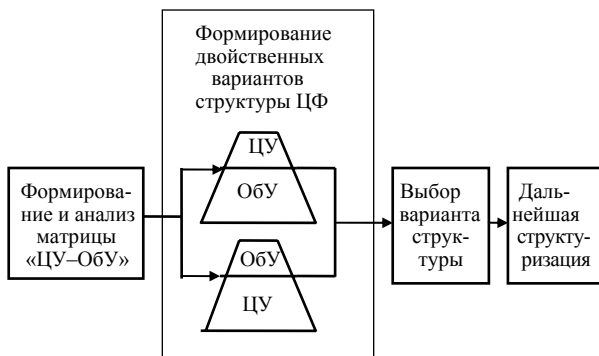


Рис. 4.7. Структура методики, основанной на двойственном определении системы А.И. Уёмова

На рис. 4.8 приведен пример вариантов таких обратных друг другу структур, соответствующих анализу составляющих в табл. 4.1.

В этом примере использовались классификаторы, рекомендованные обычно в справочниках разработчика АСУ, чтобы получить реальную ситуацию, имевшую место при разработке типовой структуры функциональной части (ФЧ) АСУП, которая в течение ряда лет была основой разработки структур ФЧ многих АСУ предприятий.

Полученные структуры анализируют.

Первая структура (рис. 4.8, а) имеет «вырожденные» ветви и не соответствует требованиям равномерности структуризации и др., предъявляемым к структурам целей, рассмотренным, например, в [1, 21, 25]. Вторая структура (рис. 4.8, б) с этой точки зрения лучше, но в ней одна ветвь (первая, связанная с управлением основным производством) «перегружена» по сравнению с остальными. Если разделить ее, выделив в самостоятельные подцели (подсистемы) ТПП, ТЭП и ОУОП (аббревиатуры см. в табл. 4.1), то структура будет удовлетворять требованию равномерности. Эта структура и была положена в основу типовой структуры функциональной части АСУП. В дальнейшем была добавлена подсистема управления качеством продукции (УКачП).

Методика Кошарского–Уёмова нашла широкое применение в различных отраслях при структуризации целей и функций предприятий в процессе разработки структуры функциональной части их автоматизированных систем управления. При этом в ряде случаев подход, положенный в ее основу, оказался столь естественным, что имена первых авторов были забыты, принцип двойственного представления не упоминался, а использовались сразу признаки «цикл управления» и «объект управления».

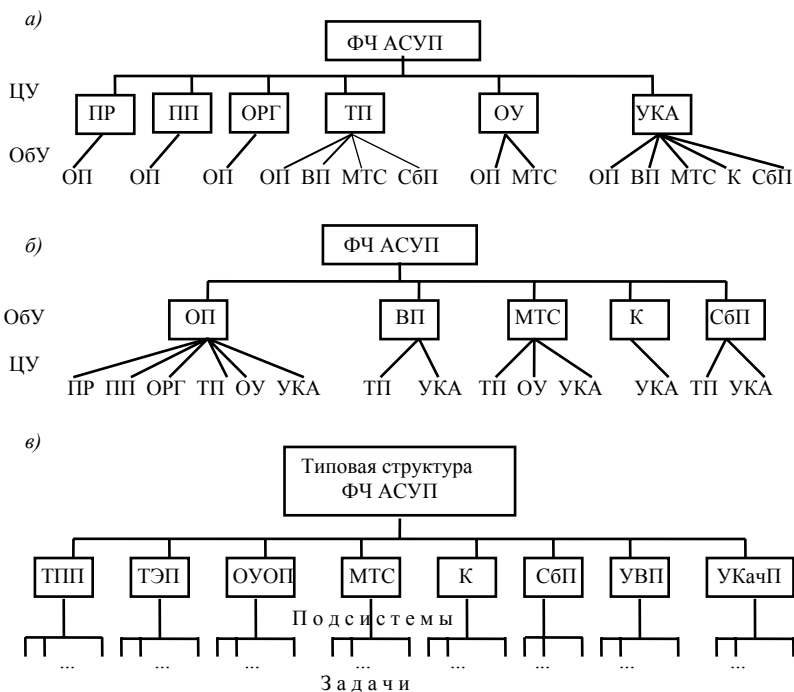


Рис. 4.8. Варианты структур ФЧ АСУП

Методика является удобной для анализа целей и функций действующих предприятий, для которых можно провести обследование существующей системы управления и выявить объекты управления. Однако в ней нет средств для определения новых объектов, новых видов деятельности и функций, связанных с развитием предприятия (внедрением новой техники, технологии и т. п.), что ограничивает применение методики при реконструкции и проектировании новых предприятий.

Применение методики, основанной на концепции системы, учитывающей среду и целеполагание. В основе методики лежит определение системы *В.Н. Сагатовского*¹, в котором учитываются понятия цели Z , среды SR и интервала времени ΔT , периода существования системы, влияющего на процесс целеобразования:

$$S \underset{def}{=} \langle A, R, Z, SR, \Delta T \rangle. \quad (4.2)$$

¹ **Основы** системного подхода и их приложение к разработке территориальных АСУ / Под ред. Ф.И. Перегудова. – Томск: Изд-во ТПИ, 1976.

Структура методики приведена на рис. 4.9.

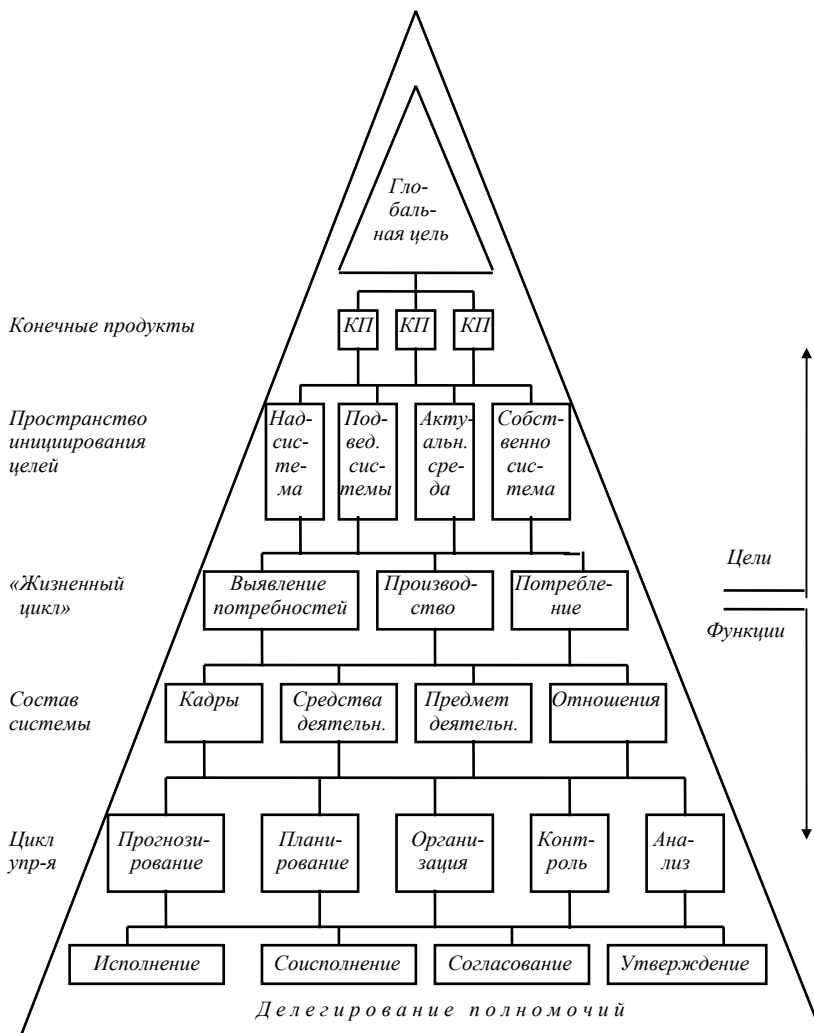


Рис. 4.9. Структура методики, учитывающей взаимодействие системы со средой

На основе этого определения дано обоснование положенной в основу методики философской концепции системы и используемых для ее раскрытия признаков структуризации системных моделей. Разработана и исследована методика группой ученых томских вузов **Ф.И.**

Перегудовым, В.Н. Сагатовским, В.З. Ямпольским, Л.В. Кочневим под руководством **Ф.И. Перегудова**. Поэтому кратко ее иногда называют методикой Ф.И. Перегудова – В.Н. Сагатовского.

Основные этапы методики соответствуют уровням структуризации, Методика излагается в терминах и формулировках ее авторов, но с некоторыми дополнениями, предложенными ими позднее.

Уровень 1. Формирование глобальной цели системы.

Цель либо задается вышестоящей организацией либо воссоздается на основе анализа директивных документов. Цель должна быть ориентирована на конечный продукт, для получения которого существует или создается система.

Конечным продуктом может быть любой результат социальной деятельности: материальная продукция, новый научный результат, научная информация и т. д.

Уровень 2. Декомпозиция по признаку «виды конечного продукта» (ВКП).

Осуществляется в тех случаях, когда система производит разные виды конечного продукта. При наличии большого числа разновидностей продукции классификатор по этому признаку может быть двухуровневым.

Виды конечного продукта зависят от того, для чего строится структура целей. Если речь идет о производстве, то конечным продуктом является выпускаемая продукция, а если структура целей строится для аппарата управления, – то это планы, решения и другие нормативно-методические документы, обеспечивающие выпуск соответствующего вид продукции.

Уровень 3. Декомпозиция по признаку «пространство инициирования целей» (ПИЦ).

Формируются подцели исследуемой системы, иницируемые требованиями и потребностями окружающей среды, которые имеют отношение к производству конечного продукта проектируемой или исследуемой системы.

При этом все системы, с которыми взаимодействует исследуемая в процессе производства конечного продукта, делятся на четыре класса (рис. 4.10): *надсистема* (НС) или *вышестоящие системы* (ВС), формулирующие главные требования к конечному продукту (и потребности в нем); *ниже-*

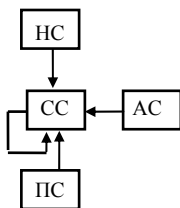


Рис.4.10. ПИЦ

лежащие или *подведомственные системы* (ПС), требования которых выступают в основном в качестве ограничений на свойства конечного продукта или потребностей в организации ремонта и других видов обслуживания материально-технической базы для производства конечного продукта, *существенная* или *актуальная среда* (АС), т.е. системы¹, исследуемая *собственно система* (СС), подцели которой иницируются собственными (внутренними) потребностями, мотивами, программами, посто-

¹ В дальнейшем **В.Г. Колосовым** была предложена идея выделения *дружественной* (поставщики, потребители, аналогичные предприятия, с которыми сотрудничает рассматриваемая система), *конкурентной* (предприятия, выпускающие аналогичную продукцию или поставляющие ее на рынок из других стран или регионов) и *безразличной* среды, которая на данном этапе является нейтральной, но которая со временем может стать дружественной или конкурентной.

янно возникающими в развивающейся системе и также трансформирующимися в требования к конечному продукту.

Отметим, что этот признак структуризации базируется на *закономерности коммуникативности* (см., например, в [1, 21, 25]).

Уровень 4. Декомпозиция по признаку «жизненный цикл».

Определяются различные подэтапы получения конечных продуктов в зависимости от их видов – от формирования или прогнозирования потребностей в продукте до потребления или поставки заказчику (см. примеры на рис. 4.9).

Начиная с этого уровня декомпозиции обычно становится удобнее оперировать не термином «подцель», а термином «функция», и считать, что «*дерево целей*» как бы перерастает в «*дерево функций*».

Уровень 5. Декомпозиция по основным элементам (составу) системы (СС), в результате которой формируются функции, вытекающие из потребностей основных элементов системы, которые объединяются в три основные группы – *кадры К, предмет деятельности ПД и средства деятельности СД*, на практике иногда вызывает затруднения, и его бывает удобно интерпретировать как «объекты деятельности» (основной – ПД, и обеспечивающей – СД, К и др. ресурсы системы).

Уровень 6. Декомпозиция по признаку «*Управленческий цикл*», классификатор которого, предлагаемый авторами методики, приведен на рис. 4.9.

Уровень 7. Декомпозиция по признаку «*Делегирование полномочий*», классификатор по которому также приведен на рис. 4.9.

Примечание. Методика развивалась. В первоначальном ее варианте было пять уровней декомпозиции. Позднее добавилось еще два. В зависимости от приложений менялась последовательность признаков: во многих приложениях удобнее на первое место вынести признак «пространство инициирования целей», помогающий уточнить «виды конечного продукта». По-разному можно раскрывать «жизненный цикл» системы.

Методика нашла наиболее широкое применение из всех рассматриваемых: она использовалась при разработке «дерева целей» управления хозяйством области; при формировании структуры функциональной части территориальной АСУ Томской области; структуры ФЧ отраслевой АСУ (ОАСУ) Минвуза РСФСР; при корректировке организационной структуры Минвуза и т.д. (см. ссылки в [1, 21]).

Наибольшее распространение получили первые три признака структуризации (т.е. признаки формирования собственно структуры целей, и особенно признак «пространство инициирования целей», в основе которого лежит закономерность коммуникативности (разделяющая сложную среду на надсистему, подведомственные системы, актуальную среду и внутреннюю среду – собственно систему, постоянно изменяющуюся в развивающейся системе).

Получаемые с помощью этой методики структуры цели и функций существенно полнее, чем при использовании предыдущей.

Пример структуры приведен на рис. 4.11.



Рис. 4.11. Пример структуры целей и функций, полученных на основе методики, учитывающей взаимодействие системы со средой

Анализ *пространства иницирования целей* заставляет обратить внимание на необходимость определения платежеспособного спроса на продукцию (услуги) предприятия (НС), поддержания в работоспособном состоянии оборудования, организации ремонта, изготовления специнструмента, спецоснастки и на другие функции, обусловленные потребностями ПС, а также на функции координации и кооперирования с предприятиями и организациями АС и функции СС.

Таким образом, благодаря большей полноте определения системы, данного **В.Н. Сагатовским**, по сравнению с двойственным определением системы **А.И. Уёмова** и раскрытию этого определения в хорошо отработанной совокупности признаков структуризации и примеров классификаторов по этим признакам, рассматриваемая методика

обеспечивает большую полноту анализа целей и функций систем управления по сравнению с предыдущей методикой, помогает выявлять новые, ранее не выполнявшиеся на предприятии функции. Методика является хорошим средством анализа целей и функций в условиях изменяющейся среды, развития предприятия (организации), при внедрении в производство и управление различного рода нововведений, техническом перевооружении и реконструкции или проектировании новых предприятий.

В установившемся же режиме функционирования предприятия эта методика может оказаться избыточной. Кроме того, возникают затруднения при определении «видов конечного продукта» четвертой ветви (собственно системы управления), для структуризации которой оказалась удобной методика, основанная на концепции деятельности¹.

Сравнивая методики с точки зрения положенных в их основу концепций, можно дать некоторые рекомендации по их выбору в конкретных условиях.

Так, концепция двойственного определения системы *А.И. Уёмова* ориентирована на описание статики системы, фиксацию уже достигнутых представлений о ней у ЛПР. Она, разумеется, допускает включение новых объектов управления, изменение функций в цикле управления (что и происходит по мере развития методики), однако в ней нет средств, которые помогали бы выявить новые объекты, функции, виды деятельности, такие, как внедрение новой техники, технологии, нововведений в управленческой деятельности. Методику, основанную на концепции системы, учитывающей ее взаимодействие со средой, полезно применять на этапах развития системы, пересмотра производственной и организационной структур, при проектировании новых предприятий. Она, помогает выявить новые виды деятельности, объекты управления.

Существуют и другие методики структуризации целей и систем, рассмотренные в [1, 21, 25], необходимость в использовании которых определяется конкретной проблемной ситуацией.

Необходимость в использовании методики, базирующейся на концепции деятельности¹, возникает в тех случаях, когда исследуемый или создаваемый объект недостаточно изучен, т. е. в случаях постановки новых проблем, структуризации целей развития новых видов деятельности.

Методика *Р. Акоффа* и *Ф. Эмери*² помогает обеспечить полноту выявления подцелей и функций для системы управления районом, городом, для управления непромышленной деятельностью предприятия (культурно-бытовыми, детскими учреждениями и другими учреждениями социальной сферы).

¹ *Применение* системного анализа на разных уровнях правления в высшей школе: Обзор. информация / Под ред. В.Н. Волковой. – М.: НИИВШ, 1977. – 65 с.

² *Акофф Р. О целеустремленных системах* / Р. Акофф, Ф. Эмери. – М.: Сов. радио, 1974. – 272 с.

Таким образом, при выборе и разработке методики структуризации целей и функций системы управления нужно учитывать состояние системы (находится ли она в стабильном состоянии, либо требуется существенный пересмотр ее целей и функций в связи с реконструкцией производства, перестройкой системы управления, изменением принципов организации экономики и т.п., либо система еще мало исследована), характер анализируемого вида деятельности, степень познания объекта (т.е. имеющиеся представления о нем у лиц, формирующих структуру ЦФ), отведенный период времени на проектирование или преобразование системы управления, влияющий на возможность выполнения методики в полном объеме, включая оба этапа, и т.п.

Поэтому целесообразно разрабатывать обобщенную методику формирования и анализа структур целей и функций, которая включала бы несколько методик структуризации, несколько методов оценки структур, предусматривала бы возможность выбора методики получения первоначального варианта (вариантов) структуры и методов оценки компонентов структур ЦФ, наиболее подходящих для соответствующего периода развития предприятия и его системы организационного управления, с учетом характера, объемов и других конкретных особенностей создаваемого предприятия (организации).

Принципы автоматизации формирования структуры целей и функций

Изучение процесса формирования и анализа структур целей и функций управления показало, что это – сложный, итеративный процесс, требующий уточнения признаков структуризации, классификаторов по этим признакам, изменения их последовательности, обсуждения вариантов структуры и внесения изменений в исходные классификаторы. Если структуру параллельно формируют разные специалисты, то даже при использовании одной и той же методики они, как правило, формируют разные варианты структуры, что обусловлено проявлением закономерности эмерджентности (см. [1, 21, 25] на каждом уровне иерархии. При сопоставлении вариантов структуры, согласовании мнений экспертов нужно обеспечить быструю повторяемость формирования новых, уточненных структур, что весьма трудоемко.

Все это обусловило необходимость поиска путей автоматизации формирования и анализа структур целей и функций, которая позволила бы сократить время на получение структуры, не снижая при этом степени полноты.

Анализ пошагового формирования структуры ЦФ с использованием методик структуризации позволяет частично формализовать этот процесс и подготовить возможность автоматизации. Анализ показал также, что по мере добавления признаков структуризации при продвижении с верхнего

уровня на нижний получаемые подцели и функции становятся все более конкретизированными. Это означает, что при помещении рядом составляющих классификаторов смежных – вышестоящего и нижележащего – уровней иерархической структуры и при продвижении по структуре сверху вниз происходит уточнение содержания функций управления, т. е. как бы появляется новый смысл (вследствие осмысления формируемых сочетаний человеком, их воспринимающим и оценивающим).

При сопоставлении таких конкретизированных подцелей (функций) можно ставить вопрос о том, какие из них в условиях конкретного предприятия являются более значимыми, по каким нужно осуществлять централизованное управление, создавать автоматизированные процедуры поддержки принятия решений и т. п.

Эффект появления нового смысла исследуется в теории множеств (при формировании множеств из «пар», «троек», «*n*-ок» элементов исходных множеств), в математической лингвистике и теории языков (при создании тезаурусов).

Например, в соответствии с тезаурусным или лингвистическим представлением рассматриваемой ситуации можно сформировать списки понятий (классификаторы по признакам структуризации) и, помещая рядом составляющие из разных списков (операция *конкатенации*), получить таким образом все возможные конкретизированные функции, которые должны быть в структуре целей и функций при принятой концепции системы и соответствующей методике структуризации.

Организация такой процедуры базируется на математической операции «размещения с повторениями». Программа для ее реализации должна содержать несколько вложенных циклов (по числу списков). Однако число получаемых размещений с повторениями согласно известной теореме комбинаторики определяется как $R = k_1 * k_2 * \dots * k_n$, где k_1, k_2, \dots, k_n – число элементов в исходных списках (классификаторах), и если число списков (признаков структуризации), например, 7, и в каждом классификаторе содержится хотя бы по 7 составляющих, то $R = 7 * 7$, что много не только для человека, но и для ЭВМ, а если элементов и списков больше, то перебор (хотя число размещений с повторениями меньше, чем число сочетаний) становится практически нереализуемым.

Поэтому решение проблемы автоматизации формирования структуры ЦФ оказалось возможным только в диалоговом режиме. Опираясь на основную идею «дерева целей» – отсечение малозначимых (несущественных или несуществующих) элементов «дерева целей» на каждом шаге структуризации – можно существенно сократить перебор и осуществить автоматизацию формирования структуры ЦФ.

На основе этой идеи были определены принципы разработки автоматизированных диалоговых процедур анализа целей и функций, бази-

рующиеся на получении размещений с повторениями из исходных списков классификаторов по используемым признакам структуризации, в результате чего возникает известный в теории множеств эффект появления нового смысла, уточнения подцелей и функций.

Автоматизированная процедура реализуется в диалоговом режиме, что позволяет сократить перебор путем отсеечения малозначимых или не имеющих смысла размещений на каждом шаге структуризации.

Идея организации такой человеко-машинной процедуры иллюстрируется рисунком 4.12.

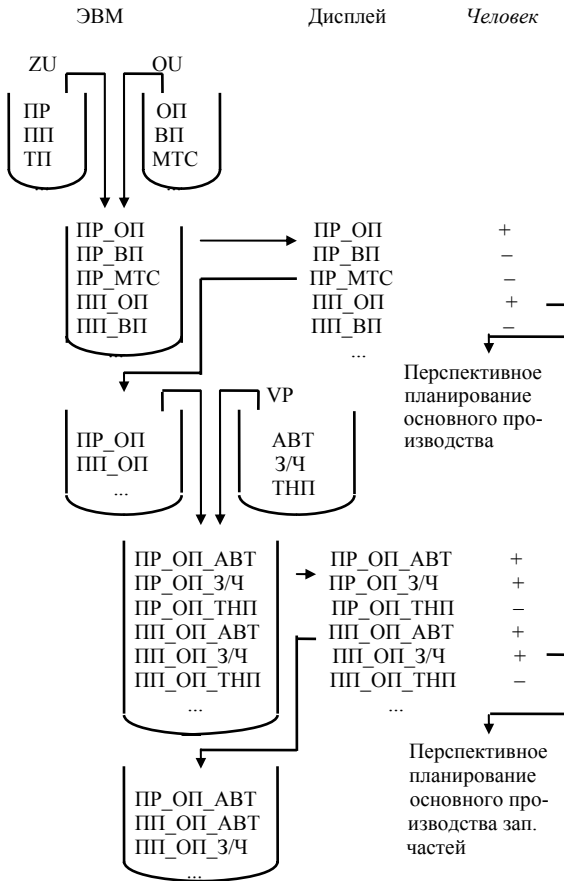


Рис. 4.12. Принципы автоматизации формирования структур ЦФ

На рис. 4.12 приводится пример применения методики, базирующейся на двойственном определении системы А.И. Уёмова, согласно которой на верхних уровнях иерархической структуры применяются признаки структуризации «объект управления (OU)» – «цикл управления (ZU)», а в качестве третьего признака структуризации используется признак «виды продукции (VP)», классификатор которого включает автомобили (АВТ), зап. части (З/Ч), товары народного потребления (ТНП). Участие человека показано ответами по двоичной систем «+» – «-».

Естественно, в классификаторы вводятся только ключевые слова. Поэтому на рисунке приведены в качестве примера формулировки подцелей, получаемые на основе отобранных размещений, составленных из ключевых слов. Глобальная цель названа – «Основные направления (ОН) деятельности предприятия».

Разрабатывая процедуру автоматизации формирования структуры ЦФ, желательно предусмотреть возможность изменения последовательности признаков структуризации при выводе результатов анализа на дисплей или принтер.

Пример двойственных вариантов структуры для двух верхних уровней был приведен на рис. 4.8, а и б).

Появление нового смысла может быть связано не только с операцией конкатенации. Между уровнями могут быть и причинно-следственные отношения, и отношения подчиненности, включения и т. п., что, например, имеет место при использовании других методик структуризации.

После разработки структуры целей и функций проводится оценка составляющих ее структуры для выявления наиболее значимых. При этом следует параллельно использовать *экспертные* и *косвенные количественные оценки*, которые затем обрабатываются совместно (см., например, в [1, 21, 25]).

При экспертной оценке в качестве критериев используется система критериев, аналогичная принятой в методике ПАТТЕРН¹, но с некоторыми модификациями: в ней предлагаются следующие группы критериев – относительной важности (значимости), взаимосвязанности, экономические оценки (последняя – заменяет группу критериев «состояние-срок»).

Идея косвенных количественных оценок предложена на основе анализа иерархических структур с помощью информационного подхода, из результатов которого следует, что структурированность ветвей

¹ Лопухин М.М. ПАТТЕРН – метод планирования и прогнозирования научных работ / М.М. Лопухин. – М.: Сов. радио, 1971. – 160 с.

иерархической структуры определяет придаваемую им фактическую значимость¹.

В качестве косвенных количественных оценок могут быть использованы: число подразделений, выполняющих данную функцию, число документов, подготавливаемых для реализации функции, структурированность ветви (например, число тем, на которые разделена проблема), внимание к соответствующим подцелям и функциям в директивных документах, периодической печати, источниках НТИ и т.п. При выборе косвенных количественных оценок следует учитывать «пространство инициирования целей и факторов», т.е. учитывать требования и потребности надсистемы, отраженные в законодательных актах и директивных документах, аналогичных предприятий актуальной среды, интересы подведомственных подразделений, инициативы структурных единиц собственно системы управления

При обработке результатов оценки применяются не только традиционные методы усреднения, но и выявление *противоречивых мнений* с последующим содержательным анализом этих оценок.

При этом для сопоставления оценок удобно применять графические представления в виде гистограмм.

Для сбора и обработки экспертных и косвенных количественных оценок и для сравнительного информационного анализа структур следует создать группу соответствующих специалистов. В необходимых случаях может быть использованы методы организации сложных экспертиз, базирующиеся на информационном подходе [1, 3, 21].

На основе полученных результатов оценки исходная структура корректируется составляющие, получившие наименьшие оценки значимости по сравнению с другими и не получившие при этом высоких оценок связности с высокозначимыми, либо исключаются из структуры ЦФ, либо опускаются на нижележащие уровни иерархии, и, напротив, составляющие, получившие высокие оценки значимости, могут быть перенесены на более высокие уровни иерархической структуры. При такой корректировке могут возникнуть вырожденные ветви, разные варианты новой структуры. Может возникнуть необходимость изменить логику формирования структуры.

В результате получается структура ФЧ АИС, представляющая собой совокупность подсистем, подлежащих автоматизации и задач, на которые делятся подсистемы. При необходимости могут быть выделены вначале группы задач. А при увеличении числа подсистем подсистемы могут быть объединены в направления, а если возникнет необходимости формируются многоуровневые структуры, подобно тому, как это было сделано на ВАЗе (рис. 4.6).

¹ **Волкова В.Н.** Оценка целевых структур при разработке планов в системе высшей и средней специальной школы: Экспресс-информация / В.Н. Волкова, А.И. Коношенко, А.Ф. Марьенко – М.: НИИВШ, 1980. – 19 с.

4.4. Разработка структуры обеспечивающей части АИС

Разделение АСУ на функциональную и обеспечивающую части, а последней – на информационное обеспечение (ИО), техническое (ТО), организационное (ОргО), программное (ПО) и другие виды обеспечения, позволило привлечь для уточнения соответствующих видов обеспечения специалистов в этих областях.

Такой подход к организации разработок АСУ помог справиться со сложностью системы и ускорить разработку АСУ путем параллельного проведения работ по анализу и выбору структуры отдельных видов обеспечения. Однако, если разрабатывать отдельные проекты ИО, ТО, ОргО и других видов обеспечения, то после разработки этих проектов возникла достаточно сложная задача их согласования, взаимоувязки принятых структур этих видов обеспечения, критериев, учитываемых при их разработке и т. д.

Поэтому на определенном этапе развития работ по созданию АСУ был даже сформулирован специальный принцип – единства ИО, ТО и ОргО как основных видов обеспечения, при разработке структур которых возникали несогласованности – и было рекомендовано проектировать структуру ОЧ АСУ с самого начала как единую с уточнением структур отдельных видов обеспечения в рамках общего проекта.

В качестве примера приводится подход к разработке структуры ОЧ АИС, предложенный в 1970-е годы на основе идеи постепенной формализации модели принятия решений¹.

Задача обоснования структуры ОЧ АСУ была сформулирована следующим образом.

Под структурой ОЧ АСУ понимается сеть информационных служб (Главный информационно-вычислительный центр, локальные вычислительные центры производств, цехов и других подразделений, автоматизированные рабочие места и другие составляющие ОргО) с размещенными в ней массивами хранения информации, документами (ИО), техническими средствами регистрации, хранения, передачи, обработки, представления информации (ТО), программным обеспечением (ПО), методическим обеспечением (инструкциями для пользователей, положениями о подразделениях и т. п.) и другими видами обеспечения.

Под выбором структуры ОЧ – задача наилучшего размещения всех этих компонентов с учетом единой согласованной системы критериев и ограничений, обеспечивающей наиболее эффективную реали-

¹ *Волкова В.Н.* К методике проектирования автоматизированных информационных систем / В.Н. Волкова // Автоматическое управление и вычислительная техника. – М.: Машиностроение, 1975. – Вып. 11. – С. 289–300.

зацию подсистем и задач, включенных в структуру ФЧ АСУ на соответствующем этапе ее развития (т.е. соответствующей очереди АСУ).

Задача в такой постановке на первый взгляд кажется практически нерешаемой. Действительно, представить эту задачу классом *хорошо организованных* систем, т.е. создать математическую модель, в которой взаимосвязи между компонентами структуры ОЧ и целями АСУ (задачами, входящими в структуру ФЧ) были бы описаны в виде аналитических зависимостей, практически невозможно.

Более реально представить эту задачу моделью *плохо организованной* системы, т.е. использовать статистические исследования, отражающие основные характеристики потоков информации, и на этой основе предложить структуру ОргО, информационных массивов, требуемые технические средства, т.е. разработать ориентировочный вариант структуры ОЧ. Однако и этот путь – не лучший с точки зрения затрат времени и доказательства правомерности принятого решения.

Такой подход возможен при корректировке структуры ОЧ, когда есть предварительные исследования информационных потоков. В условиях же создания новой АИС получение статистики на основе «ручного» способа сбора информации может дать неверные результаты, поскольку, во-первых, «ручной» способ обычно реализован по не так, как в автоматизированном варианте, а во-вторых, при автоматизации могут принципиально измениться формы сбора информации.

Наиболее целесообразно отобразить задачу с помощью класса *самоорганизующихся, развивающихся* систем и организовать процесс «выращивания» структуры ОЧ с помощью модели постепенной формализации задачи, изложенной, например, в [1, 4, 21].

Принципиальной особенностью модели постепенной формализации задачи является то, что она ориентирована на развитие представлений исследователя об объекте или процессе принятия решений, на постепенное «выращивание» решения задачи. Поэтому предусматривается не одноразовый выбор методов моделирования, а смена методов по мере развития у ЛПР представлений об объекте и проблемной ситуации в направлении все большей формализации модели принятия решения.

Для формирования и анализа модели постепенной формализации была разработана методика системного анализа, сочетающая методы из групп МАИС и МФПС, которые помогают создать язык моделирования, процедуру оценки вариантов решения и автоматизировать процесс «выращивания» решения.

Возможный вариант смены методов по мере развития модели проиллюстрируем на упрощенном примере моделирования процессов

прохождения информации при выборе структуры обеспечивающей части автоматизированной информационной системы.

Основная идея постепенной формализации иллюстрируется рисунком 4.13. При этом на рис. 4.13 показаны последовательные переходы от методов работы с ЛПР (из группы МАИС) к методам формализованного представления и обратно.

В методике системного анализа для решения данной задачи удобно предусмотреть два основных этапа:

1. Формирование модели, отображающей возможные варианты прохождения информации в АИС (этот этап иллюстрируется рисунками 4.13 и 4.14).

2. Оценка модели и выбор наилучшего варианта пути прохождения информации (рис. 4.15).

Охарактеризуем кратко эти этапы.

Э т а п 1. Формирование модели, отображающей возможные варианты прохождения информации в АИС.

1.1. Отграничение системы от среды («перечисление» элементов системы).

Подэтап может выполняться с применением метода «мозговой атаки», а в реальных условиях – методов типа комиссий, семинаров и других форм коллективного обсуждения, в результате которого определяется некоторый перечень элементов будущей системы. В состав таких комиссий должны входить и разработчики, и будущие пользователи АИС.

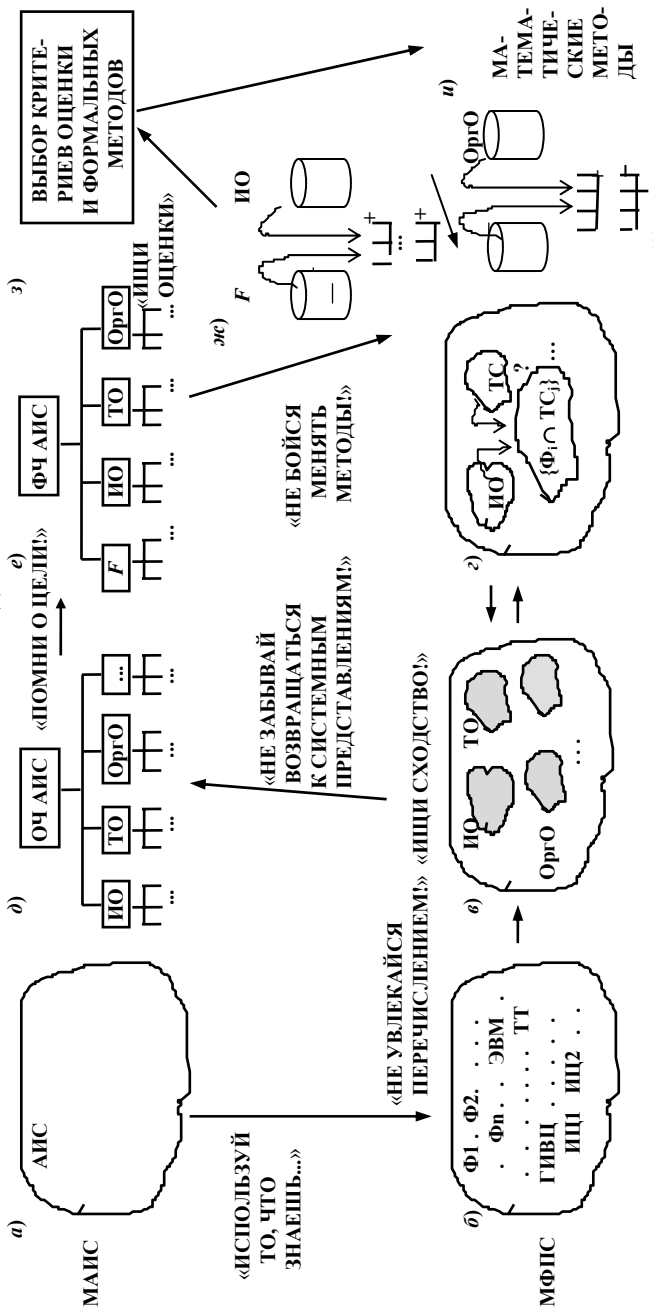
Задачу «перечисления» можно представить на языке теоретико-множественных методов как переход от названия характеристического свойства, отраженного в названии формируемой системы и множества ее элементов, к перечислению элементов, которые отвечают этому свойству и могут быть включены в исходное множество.

На рис. 4.13, б и 4.14, а перечислено для примера небольшое число исходных элементов: *ГИВЦ, ИЦ1, ИЦ2, ...*, *A1, A2, ...* – пункты сбора и обработки информации; *Ф1, Ф2, ...* – формы сбора и представления информации (документы, массивы); *ЭВМ, ТТ* (телетайп), *T* (телефон), ... и т. д. Понятно, что в реальных условиях конкретных видов подобных элементов существенно больше, и они будут названы более конкретно – не *ЭВМ*, а тип *ЭВМ*; аналогично – тип *ТТ*, регистраторов производства (*РП*), наименование или код документов и массивов и т. д.

1.2. Объединение элементов в группы.

Не следует слишком увлекаться подэтапом «перечисления» (хотя на практике это и имело место). Сложную реальную развивающуюся систему невозможно «перечислить» полностью. Следует, набрав некоторое множество элементов, попытаться объединить их в группы, найти меры сходства, «близости» и предложить способ их объединения.

СТРУКТУРНЫЕ ПРЕДСТАВЛЕНИЯ



ТЕОРЕТИКО-МНОЖЕСТВЕННЫЕ ПРЕДСТАВЛЕНИЯ

СТРУКТУРНО-ЛИНГВИСТИЧЕСКИЕ ПРЕДСТАВЛЕНИЯ

Рис. 4.13. Иллюстрация идеи постепенной формализации

Если в качестве метода формализованного отображения совокупности элементов выбраны теоретико-множественные представления, то этот подэтап можно трактовать как образование из элементов исходного множества некоторых подмножеств путем перехода от перечисления сходных по какому-то признаку элементов к названию характеристического свойства этого подмножества. В результате в приводимом примере могут быть образованы подмножества элементов по соответствующим видам обеспечения – *ИО*, *ТО*, *ОргО* (рис. 4.13, в и 4.14, б).

Примечание. Можно было бы предусмотреть еще один подэтап разделения подмножеств на более мелкие.

1.3. Формирование из элементов подмножеств новых множеств, состоящих из «пар», «троек», «п-ок» элементов исходных подмножеств.

В рассматриваемом примере объединяя элементы подмножеств *ИО*, *ТО*, *ОргО* в «пары» и «тройки» можно получить, например: *Ф1_ЭВМ*, *Ф1_ТТ*, *Ф2_ЭВМ* и т.п.; *ЭВМ_ГИВЦ*, *ЭВМ_ИЦ1*, *ЭВМ_А1*, *ТТ_ГИВЦ*, *ТТ_ИВЦ1*, *ТТ_А1* и т.п.; *Ф1_ЭВМ_ГИВЦ*, *Ф1_ТТ_А1* и т.д.

Иногда в задачах моделирования на этом этапе можно получить новый результат, который подсказывает путь дальнейшего анализа.

Но в данном примере, как правило, интерпретация получаемых компонентов затруднена и ввести какое-либо формальное правило сравнения элементов новых множеств для принятия решения о выборе наилучших не удастся. В таких случаях, согласно рассматриваемому подходу, нужно обратиться к системно-структурным представлениям и попытаться поискать дальнейший путь развития модели.

Примечание. Отметим, что отражением данного подэтапа в истории развития проектных работ по созданию АСУ был период, когда формировались матрицы «документы – технические средства», «информационные службы – технические средства» и на основе экспертной оценки элементов этих матриц пытались принимать решения, связанные с формированием структуры ОЧ АИС.

1.4. Содержательный анализ полученных результатов и поиск новых путей развития модели.

Для проведения содержательного анализа следует возвратиться к системным представлениям и использовать один из методов группы МАИС – структуризацию (в данном случае в форме иерархической структуры – рис. 4.13, д и 4.14, в).

Такое представление более удобно для руководителей работ по созданию АСУ, чем теоретико-множественные представления, и помогает им вначале распределить работу между соответствующими специалистами по *ИО*, *ТО* и т.д., а затем найти дальнейший путь развития моде-

ли на основе содержательного анализа сути полученных «пар» и «троек» с точки зрения формулировки решаемой задачи.

В формулировке задачи, помимо упоминания видов обеспечения, говорится, что все эти компоненты нужно разместить так, чтобы обеспечить наиболее эффективную реализацию подсистем и задач, входящих в ФЧ АИС. А поскольку любая задача представляет собой последовательность действий (*функций*) по сбору, хранению и первичной обработке информации, то становится очевидной необходимость внесения в модель нового подмножества «Функции-операции (ФО)», добавление элементов которого к прежним «парам» и «тройкам» позволяет получить новое их осмысление (рис. 4.13, е).

Для того, чтобы глубже понять развитие модели постепенной формализации, проиллюстрируем изложенную идею на конкретном примере.

Предположим, что нужно принять решение о структуре ОЧ АИС отрасли (первой очереди ОАСУ), предприятия которой расположены в разных городах. При этом предварительно рассматривается 2 основных варианта:

1) создание единого Главного информационно-вычислительного центра (ГИВЦ) отрасли и организации централизованного сбора от всех предприятий посредством установленных на них периферийных средств сбора информации (A_1, A_2, \dots, A_k);

2) наряду с ГИВЦ и периферийными средствами сбора на предприятиях, создать региональные ИВЦ (обозначенные на рис. 4.14 $ИЦ_1, ИЦ_2, \dots, ИЦ_n$), которые будут расположены в городах.

Необходимо выбрать вариант и определить вычислительную мощность ГИВЦ и региональных ИВЦ (в случае выбора второго варианта), типы ЭВМ для ГИВЦ и ИВЦ, типы периферийных средств регистрации информации, объемы информационных массивов в ГИВЦ и ИВЦ, формы документов $\Phi_1, \Phi_2, \dots, \Phi_m$ сбора и передачи информации между пунктами, принятыми в соответствующем варианте. При этом следует иметь в виду, что в случае выбора первого варианта возникают проблемы диспетчеризации приема-передачи информации от достаточно многочисленных пунктов первичного сбора информации на предприятиях.

Аналогично может быть поставлена задача для корпорации, фирмы которых расположены в разных городах (как, например, в Объединении Авто-ВАЗ), или для предприятия, крупные производства которых расположены в разных корпусах.

Для ответа на требуемые вопросы и выбора структуры ОЧ АИС необходимо исследовать информационные потоки.

Для простоты на рис. 4.13 и рис. 4.14 показаны только принципиально отличающиеся друг от друга функции – связи C , хранения M (от «*memory*» – «память») и обработки K (от «*компьютер*»). После их добавления получают комбинации, которые ЛПП могут не только сравнивать, но и оценивать.

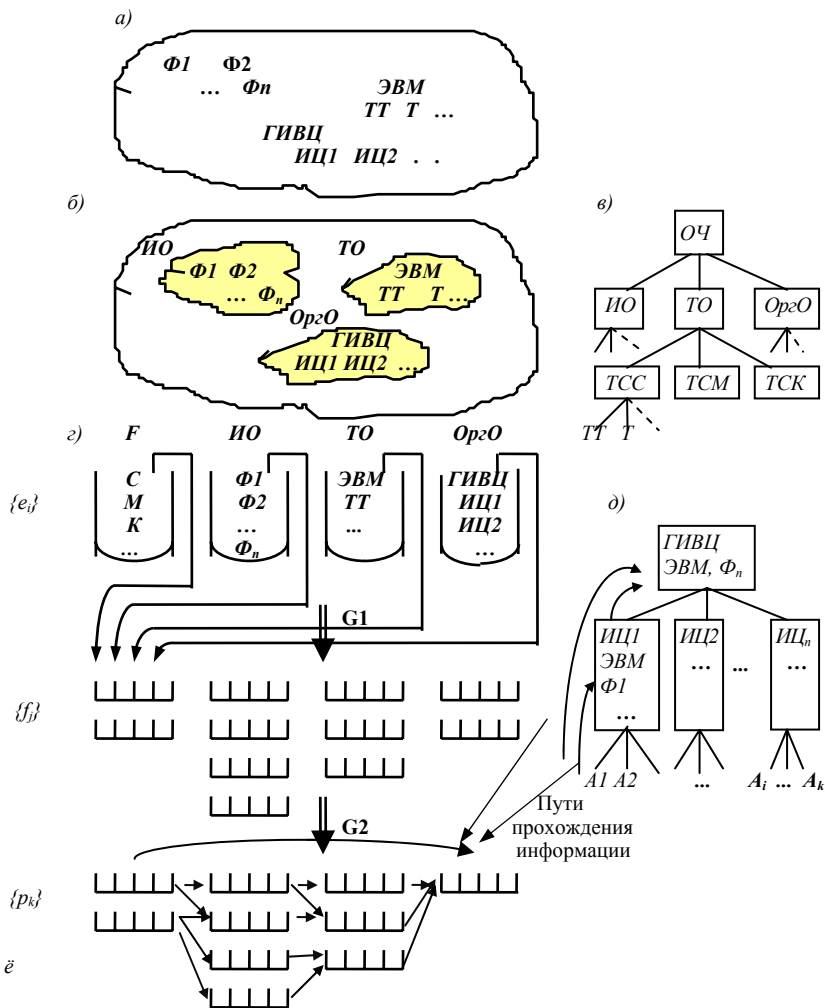


Рис. 4.14. Иллюстрация формирования графо-семиотической модели

Например, комбинации типа $C_{\Phi 1_{TT}}$ и $C_{\Phi 1_T}$ отличаются друг от друга скоростью передачи информации, которые в конкретных условиях можно измерить или вычислить.

Если бы задача не была сформулирована, то при содержательном осмыслении результатов предшествующих этапов нужно было бы

уточнить формулировку задачи, что и помогло бы сделать дальнейший шаг в развитии модели.

Примечание. Отметим, что применение на практике матричных представлений, упомянутых в предыдущем пункте, обратило внимание разработчиков на тот факт, что анализировать названные матрицы практически невозможно, если не задумываться над тем, для решения каких задач предназначены технические средства и документы. Тогда стали формировать матрицы «задачи – методы», «задачи – средства», т. е. вводить фактически новое множество – «Задачи» или «Функции (F)».

1.5. Разработка языка моделирования.

После того, как найдено недостающее подмножество, в принципе можно было бы продолжить дальнейшее формирование модели, пользуясь теоретико-множественными представлениями. Однако, когда осознана необходимость формирования последовательностей функций-операций, конкретизированных путем дополнения их видами обеспечения – конкретизированных функций ($K\Phi$), то целесообразнее выбрать лингвистические представления, которые удобнее для разработки языка моделирования последовательностей $K\Phi$.

Тогда в терминах лингвистических представлений данный этап можно представить следующим образом:

разработка тезауруса языка моделирования;

разработка грамматики (или нескольких грамматик, что зависит от числа уровней модели и различии правил).

Структура тезауруса языка моделирования, приведенная на рис. 4.13, ж и 4.14, з, включает три уровня:

уровень первичных терминов (или слов), которые представлены в виде списков, состоящих из элементов $\{e_{ij}\}$ подмножеств ΦO , $И O$, $Т O$, $OpzO$;

уровень фраз $\{f_{ij}\}$, который в этом конкретном языке можно называть уровнем $K\Phi$, так как абстрактные функции C , M , K , объединяясь с элементами подмножеств $И O$, $Т O$, $OpzO$, конкретизируются применительно к моделируемому процессу;

уровень предложений $\{p_{kj}\}$, отображающий варианты прохождения информации при решении той или иной задачи.

Грамматика языка включает правила двух видов:

G1 – преобразования элементов $\{e_{ij}\}$ первого уровня тезауруса в компоненты $\{f_{ij}\}$ второго уровня, которые имеют характер правил типа «помещения рядом» (конкатенации, сцепления, R_l на рис. 4.15);

G2 – преобразования компонентов $\{f_{ij}\}$ в предложения $\{p_{kj}\}$ (варианты прохождения информации) – правила типа «условного следования за»

(R_{II} на рис. 4.15); правила этого вида исключают из рассмотрения недопустимые варианты следования информации: например, после функции $C1_Ф2_A1-ИЦЦ1_ТТ$ (передача документа $Ф2$ из $Ф1$ в $ИЦЦ1$ с помощью $ТТ$) не может следовать функция $M1_Ф2_ГИВЦ_МН$, так как в результате выполнения предшествующей функции документ $Ф2$ в $ГИВЦ$ не поступил (здесь $МН$ – машинный носитель).

Используя разработанный язык, процедуру формирования модели можно автоматизировать. При этом правила типа G1 и G2 относительно несложно реализуются с помощью языков логического программирования, и в частности языка Турбо-Пролог.

Э т а п 2. Оценка модели и выбор наилучшего варианта пути прохождения информации.

После формирования вариантов следования информации необходимо их оценить. Для этого могут быть приняты также разные варианты – от содержательной оценки путей сбора и первичной обработки информации (нижний уровень рис. 4.13, *жс*) до поиска алгоритмов последовательного преобразования оценок компонентов предшествующих уровней модели в оценки компонентов последующих уровней, что осуществляется путем анализа сформированной графо-семиотической модели. Варианты оценки модели иллюстрируются рисунком 4.15.

2.1. Выбор способа оценки графо-семиотической модели

В рассматриваемом примере можно проводить оценку тремя способами (рис. 4.15):

а) на уровне вариантов прохождения информации $\{p_k\}$, что иногда могут сделать компетентные специалисты путем коллективного обсуждения предложенных им вариантов (если число этих вариантов не очень велико – не более 7 ± 2);

б) на уровне $KФ$ $\{f_{ij}\}$ с последующим преобразованием этих оценок $W\{f_{ij}\}$ в оценки вариантов $W''\{p_k\}$;

в) на уровне элементов $\{e_{ij}\}$ с последующим преобразованием оценок $W\{e_{ij}\}$ в оценки $W'\{f_{ij}\}$, а их – в оценки $W''\{p_k\}$.

При втором способе можно выделить на модели «сферы компетентности», и поручить оценку $KФ$ по сферам соответствующим специалистам; оценки $KФ$ в большинстве случаев также получают экспертно, однако в некоторых случаях они могут быть измерены (например, время передачи информации с помощью $ТТ$ и T в приведенных выше функциях $C_Ф1_ТТ$ и $C_Ф1_T$); этот способ подобен оценке сетевой модели, и при определении алгоритма преобразования оценок $Ф_n$ можно пользоваться опытом сетевого моделирования (для большинства критериев оценки алгоритм преобразования – суммирование, а для критерия надежности передачи или хранения информации, оцениваемых с помощью вероятностей, алгоритм более сложный).

При третьем способе алгоритмы преобразования Φ могут быть найдены путем анализа различных $K\Phi$ с точки зрения влияния на их оценку по тому или иному критерию элементов соответствующего вида. Например, оценка $K\Phi$ передачи информации «С...» по критерию времени t может быть получена на основе выяснения, что в структуре $K\Phi$ «С...» влияет на оценку по t . Если используются технические средства связи, то, зная принципы передачи информации с их помощью, можно определить v_{TC} и зависимости $t = r_{\phi}/v_{TC}$, где r_{ϕ} – объем передаваемой информации (например, измеряемых в числе знаков), т. е. оценка элементов, принадлежащих подмножеству $ИО$; v_{TC} – скорость передачи информации с помощью соответствующего технического средства, т. е. оценка элемента, принадлежащего подмножеству $ТО$. Таким образом, в данном примере на оценки $K\Phi$ «С...» влияют элементы подмножеств $ИО$ и $ТО$, и следует предусмотреть оценку этих элементов в исходных списках элементов. Аналогично можно определить, какие из элементов влияют на оценки $K\Phi$ по стоимости, надежности, срокам внедрения и другим учитываемым критериям оценки.

С учетом сказанного следует сделать вывод о том, что выбор способа оценки модели зависит от результатов этапа 1, т. е. от вида графо-семиотической модели, а алгоритмы преобразования оценок Φ и Φ_{II} определяются на основе анализа этой модели.

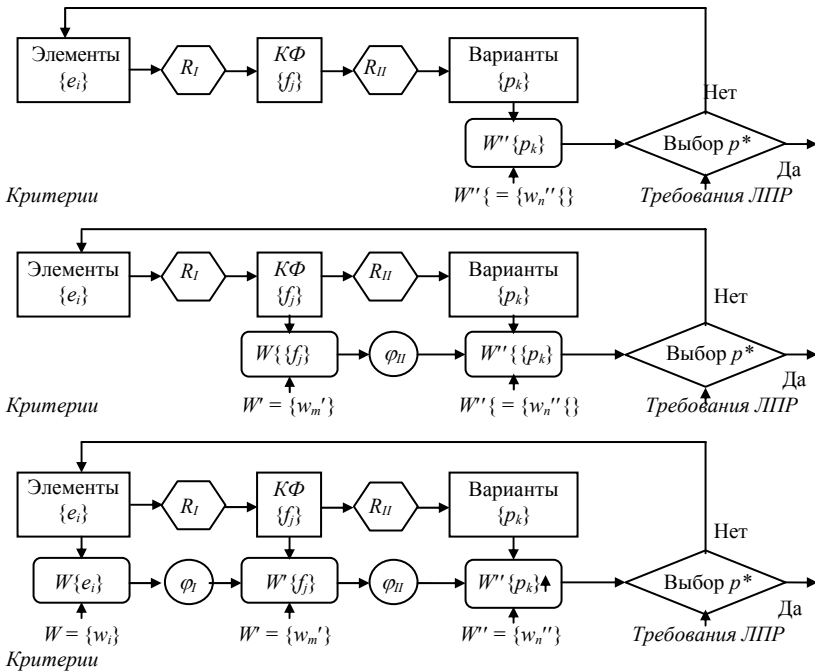


Рис. 4.15. Способы оценки графо-семиотической модели

2.2. Выбор критериев оценки модели.

Выбор критериев оценки зависит от выбранного способа оценки модели.

Например, при первых двух способах оценки (на уровне $\{p_{ij}\}$ и на уровне $\{f_{ij}\}$), могут быть приняты такие оценки, как оперативность (время), достоверность (вероятность сбоя при передаче информации, ошибок при ее обработке и т. п.), трудоемкость, затраты на внедрение, эксплуатационные расходы, сроки внедрения и т. д., а при оценке модели на уровне элементов $\{e_{ij}\}$ – оценки типа r_{ϕ} , v_{TC} и т. п., на основе которых могут быть вычислены оценки $K\Phi$, или оценки трудоемкости, скорости заполнения форм или ввода информации и т. п.

2.3. Оценки модели.

Способ оценки модели на уровне вариантов $\{p_k\}$ – экспертный; на уровне $\{f_{ij}\}$ для экспертного оценивания могут быть выделены сферы компетентности и привлечены соответствующие специалисты, знающие особенности конкретных технических средств и т. п.; и кроме того, наряду с экспертным оцениванием могут быть проведены эксперименты по той или иной $K\Phi$.

Оценки элементов $\{e_{ij}\}$, необходимые для вычисления оценок соответствующих $K\Phi$, могут быть в большинстве случаев получены из справочной литературы или измерены.

2.4. Выбор наилучшего варианта структуры ОЧ.

Основой структуры ОЧ являются выбранные варианты следования информации по задачам и подсистемам ФЧ.

Задачу определения наилучшего варианта можно поставить как многокритериальную. Но можно организовать процесс выбора вариантов путем постепенного ограничения области допустимых решений, вначале исключить все p_k , которые не удовлетворяют граничным значениям учитываемых критериев, затем предложить рассмотреть оставшиеся варианты ЛПР, которые могут либо сразу выбрать из них наиболее предпочтительный, либо ввести весовые коэффициенты критериев, либо исследовать область допустимых решений по Парето.

Можно также добавить новые критерии качественного характера, не включенные в первоначально выбранный перечень критериев из-за невозможности их количественной оценки.

В заключение отметим, что после того, как для какого-то класса задач пройдены все этапы постепенной формализации задачи и найдены основы языка моделирования, можно применять не всю методику, а сразу начинать с подэтапа 1.5.

В то же время в случае, когда нужно поставить задачу для принципиально нового объекта или процесса, полезно при обосновании модели выполнять все подэтапы методики. Это может обосновать

адекватность модели и принципы разработки языка автоматизации моделирования.

В рассматриваемом примере при разработке структуры ОЧ учитываются только задачи сбора, предварительной обработки информации и формирования первичных информационных массивов (что и включалось в основные функции АИС как первой очереди АСУ).

Если предполагается, что об АИС первоначально ничего неизвестно, кроме ее назначения, то модель формируется, как показано выше.

В качестве первого шага системного анализа предлагается принять «ограничение» системы от среды путем «перечисления» ее возможных элементов (рис. 4.13, б).

Затем (рис. 4.13, в и д) для анализа некоторого полученного множества выбраны теоретико-множественные представления, помогающие найти на сформированном пространстве состояний «меры близости» для объединения элементов в группы.

При этом вначале может быть использован эффект получения нового смысла у элементов, сформированных из «пар», «троек», «л-ок» элементов исходных подмножеств, на которые предварительно разделено общее множество элементов АИС).

Далее, когда возможности теоретико-множественных представлений в познании взаимодействия элементов в системе исчерпываются, следует вернуться к системно-структурным представлениям, с помощью которых активизируется использование интуиции и опыта ЛПП, перечень множеств анализируется и при необходимости дополняется (рис. 4.13, е) принципиально важными подмножествами для дальнейшего моделирования (в частности, в рассматриваемой задаче на этом этапе перечень исходных подмножеств *ИО*, *ТО*, *ОргО* дополнен подмножеством функций *F*).

Для дальнейшей реализации идеи комбинирования элементов в поисках вариантов решения задачи (в рассматриваемом примере – путей прохождения информации при ее сборе и первичной обработке) могут быть выбраны более удобные и подсказывающие правила формирования вариантов лингвистические представления, являющиеся основой разработки языка моделирования путей прохождения информации.

В рассматриваемом примере использовано сочетание *лингвистических*, *семиотических* и *графических* представлений и разработан язык *графо-семиотического моделирования* (который в первоначальных вариантах использования рассматриваемого подхода иногда носил и другие названия – *структурно-лингвистического*, *сигнатурного* (знакового) моделирования).

Таким образом, с помощью языка моделирования разрабатывается многоуровневая модель (в нашем примере двухуровневая, если считать уровень исходных множеств нулевым (рис. 4.13, ж).

Осмысление этой модели (на уровне МАИС) приводит к преобразованию структуры: первоначально структура ОЧ формировалась как структура-состав, в которой были представлены виды обеспечения и их детализация (рис. 4.13, з и е),

и на той основе осознана необходимость анализа структуры-функционирования, т. е. вариантов структуры информационных потоков (рис. 4.13, з).

Многоуровневую модель, отображающую взаимосвязи между компонентами и целями системы, в обобщенном виде можно представить в виде аналитических зависимостей (4.3, а) и (4.3, б).

Для варианта оценок, приведенного на рис. 4.15, а:

$$\begin{aligned}
 W^{n*}(p_{jn}) &= \text{opt } W^n(p_{jn}); \\
 & p_{jn} \in P_n, \quad P_n \subset S_I \\
 W^n(p_{jn}) &= \varphi^n \{W^{n-1}(p_{j\ n-1})\}; \\
 & p_{j\ n-1} \in P_{n-1}, \quad P_{n-1} \subset S_I \\
 & \dots\dots\dots \\
 W^k(p_{jk}) &= \varphi^k \{W^{k-1}(p_{j\ k-1})\}; \\
 & p_{jk} \in P_k, \quad P_k \subset S_I, \quad p_{j\ k-1} \in P_{k-1}, \quad P_{k-1} \subset S_I \\
 & \dots\dots\dots \\
 W^1(p_{j1}) &= \varphi^1 \{W(e_i)\}; \\
 & p_{j1} \in P_1, \quad P_1 \subset S_I, \quad e_i \in E, \quad E \subset S_I
 \end{aligned}
 \tag{4.3, а}$$

Для варианта, приведенного на рис. 4.15, б:

$$\begin{aligned}
 W^{n*}(p_{jn}) &= \text{opt } W^n(p_{jn}); \\
 & p_{jn} \in P_n, \quad P_n \subset S_I \\
 P_{jn} &= \bigcup_{j_{n-1}=1}^m p_{j\ n-1}; \\
 & p_{j\ n-1} \in P_{n-1}, \quad P_{n-1} \subset S_I \\
 & \dots\dots\dots \\
 P_{jk} &= \bigcup_{j_{k-1}=1}^I p_{j\ k-1}; \\
 & p_{jk} \in P_k, \quad P_k \subset S_I, \quad p_{j\ k-1} \in P_{k-1}, \quad P_{k-1} \subset S_I \\
 & \dots\dots\dots \\
 P_{j1} &= \bigcup_{j=1}^r e_i; \\
 & p_{j1} \in P_1, \quad P_1 \subset S_I, \quad e_i \in E, \quad E \subset S_I.
 \end{aligned}
 \tag{4.3, б}$$

Знаком \bigcup обозначено любое взаимодействие компонент «условное следование за», сложное взаимодействие или просто «помещение рядом»; $W^n(p_{jn})$ – функционал, связывающий критерии оценки выбираемого решения с компонентами p_{jn} , которые зависят от компонентов

предыдущего уровня $p_{j\ n-1}$; в общем случае p_{jk} зависят от компонентов $p_{jk\ i}$; $E, P_1, \dots, P_k, \dots, P_{n-1}, \dots, P_n$ – множества смысловыражающих элементов (тезаурус) задачи; $W(e_i), W^1(p_{i1}), W^k(p_{jk}), W^n(p_{jn})$ – критериальные отображения элементов (компонентов) структурных уровней тезауруса языка моделирования; $\varphi^1, \varphi^k, \varphi^n$ – алгоритмы преобразования критериальных отображений одного структурного уровня в другой.

В результате получается система алгоритмов, обеспечивающая возможность автоматизации, и соответственно повторяемость процесса формирования и анализа модели при изменении наборов первичных элементов и их оценок. Эта система алгоритмов как бы обеспечивает взаимосвязь между компонентами и целями системы (или при моделировании потоков информации по отдельным задачам АИС – между компонентами и этой задачей), т.е. получается в результате формальная, аналитическая модель, только представленная не в виде привычных для такого рода моделей формул или уравнений, а в виде алгоритмов и их программной реализации в памяти ЭВМ.

Однако получить такой сложный алгоритм взаимодействия элементов в системе, позволяющий отобразить конкретную ситуацию и выбрать лучшее решение, практически невозможно без организации направленной постепенной формализации задачи.

На практике после определения видов обеспечения процесс формирования модели для решения данной задачи, приведенный на рис. 4.13, был несколько упрощен.

Так, при решении задачи формирования структуры ОЧ АИЧ первый этап может быть определен тем, что ОЧ АИС разделена на виды обеспечения, т.е. сформирована структура-состав ОЧ (рис. 4.16, а).

Затем для уточнения состава видов обеспечения применены теоретико-множественные представления (рис. 4.16, б), помогающие вначале расширить состав видов обеспечения, найти способы объединения элементов из разных подмножеств (рис. 4.16, в).

В практике разработки АСУ этому этапу соответствовал поиск взаимосвязей между компонентами ОЧ с помощью формирования и экспертной оценки матриц типа «ТО–ОргО» или «ИО–ТО». На основе оценки элементов матриц «ТО–ОргО» размещались технические средства в подразделениях предприятия. В матрицах «ИО–ТО» оценивалось, с помощью каких технических средств следует обрабатывать соответствующие документы, что первоначально при ограниченных мощностях ЭВМ иногда было оправдано.

Однако на основе анализа матриц, отображающих некоторые взаимосвязи между компонентами ОЧ, решить задачу формирования структуры ОЧ не удастся. Анализ матриц, взаимоувязка полученных с их помощью взаимосвязей между элементами – весьма трудоемкая работа, задерживающая разработку АИС. Поэтому согласно рассматриваемому подходу, целесообразно *возвратиться к системно-структурным представлениям*, с помощью которых

активизируется использование интуиции и опыта ЛПР, перечень множеств анализируется и при необходимости дополняется принципиально важными подмножествами для дальнейшего моделирования.

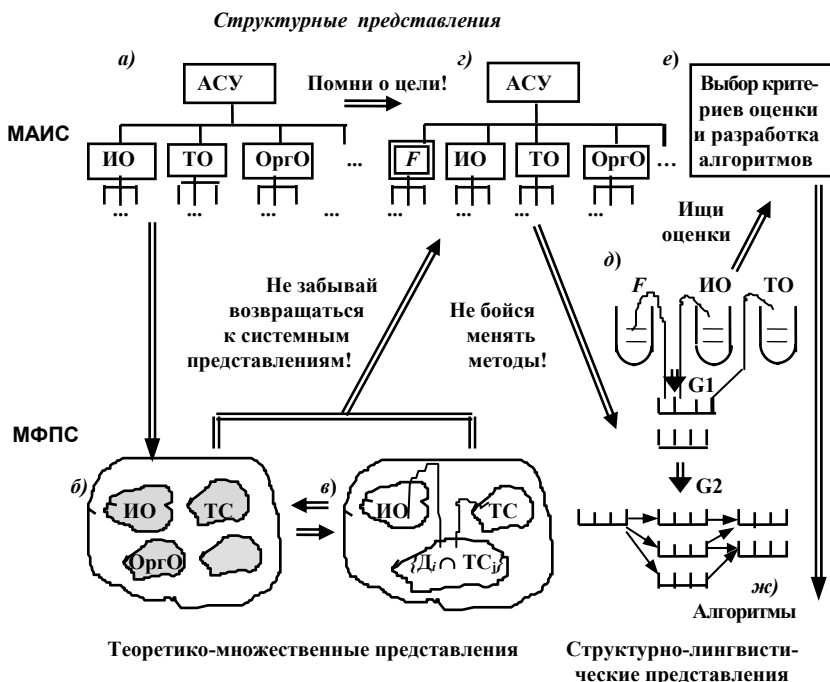


Рис. 4.16. Постепенная формализация модели обоснования структуры ОЧ АИС

В частности, в рассматриваемой задаче перечень исходных подмножеств (ИО, ТО, ОргО соответственно) дополнен подмножеством задач, входящих в структуру ФЧ АСУ, и функций F (рис. 4.16, г), которые выполняются при решении задач сбора и первичной обработке информации, входящими в структуру ФЧ АСУ 1-й очереди.

Для дальнейшей реализации идеи комбинирования элементов в поисках вариантов решения задачи (т.е. путей прохождения информации при ее сборе и первичной обработке) могут быть выбраны более удобные и подсказывающие правила формирования вариантов *лингвистические* представления, являющиеся основой разработки языка моделирования путей прохождения информации.

С помощью языка моделирования можно получить многоуровневую модель (рис. 4.16, д), позволяющую формировать варианты информационных потоков. После формирования вариантов следования информации выбраны критерии и способы оценки графо-семиотической модели (рис. 4.16,е). Приняты разные

варианты – от экспертной оценки вариантов сбора и первичной обработки информации (нижний уровень рис. 4.16, д) до поиска алгоритмов последовательно преобразования оценок компонентов предшествующих уровней модели в оценки компонентов последующих уровней.

Структура полученной методики формирования и анализа информационных потоков приведена на рис. 4.17.



Рис. 4.17. Методика обоснования структуры ОЧ АИС

Этапы 1–7 выполняются по каждой из задач ФЧ АСУ, а затем проводится содержательная интерпретация, обобщение результатов моделирования задач ФЧ АСУ и определение на этой основе состава необходимых и достаточных компонентов структуры ОЧ.

Последовательность этапов может быть иная, чем на рис. 4.16 и 4.17. Например, вначале можно применить целевой подход, и проводить активное обследование, отбор элементов модели с помощью предварительно сформированной структуры функций, на основе которой провести опрос сотрудников исследуемого или аналогичных объектов, получить необходимые сведения в документах и т.д. Или можно обратиться к оценкам раньше, чем это сделано на рис. 4.16, – оценить значимость подцелей и функций, что помогло бы ускорить выбор лучшего варианта и т.д. При выполнении этапа 5 полезно обращаться к этапу 3, а при выполнении этапа 7 – к этапу 5, что показано на рис. 4.17.

Приведенный пример демонстрирует, что постепенная формализация становится своего рода «механизмом» развития системы, «выращивания» модели принятия решения. По мере развития модели происходит смена методов из групп МАИС и МФПС. В результате можно получить разные варианты формализованной модели. В процессе моделирования следует учитывать рекомендации, приведенные на рис. 4.13, 4.14 и 4.16, типа «используй то, что знаешь», «не увлекайся перечислением», «не

забывая возвращаться к системным представлениям», «помни о цели», «не бойся менять методы» и т. п. [4].

При «выращивании» модели можно накапливать информацию об объекте, фиксируя все новые компоненты, связи, правила взаимодействия компонент, и, используя их, получать отображения последовательных состояний развивающейся системы, постепенно создавая все более адекватную модель реального, изучаемого или создаваемого объекта. При этом информация может поступать от различных специалистов и накапливаться во времени по мере ее возникновения в процессе развития объекта и наших представлений о нем.

При постановке задачи для принципиально нового объекта или процесса постепенная формализация позволяет обосновать принципы разработки языка автоматизации моделирования и обобщенную формальную модель с пошаговым доказательством ее адекватности на каждом витке моделирования.

Рассмотренный пример иллюстрирует основную суть подхода – применение идеи постепенной формализации для формирования нового алгоритма, получения новой методики системного анализа в тех случаях, когда имеет место большая начальная неопределенность задачи и отсутствуют аналоги для ее решения.

Такого рода подходы и методики применялись достаточно редко. Обычно структура АСУ развивалась на основе аналогий и возникающих текущих потребностей. В то же время по мере усложнения автоматизированных систем предприятий и организаций усложняется и управление ходом их развития, и идея приведенного подхода может оказаться полезной.

На практике может быть применена упрощенная методика разработки и развития АИС и АСУ организации, идея которой приводятся разделе 4.5.

4.5. Разработка методики для проектирования АИС предприятия

Для достаточно крупных предприятий в соответствии с Общеотраслевыми руководящими методическими материалами по созданию АСУП, был определен порядок разработки АСУ, приведенный на рис. 4.1 и 4.2. Разрабатывались методики формирования структур функциональной и обеспечивающей частей АСУ, подобные приведенным выше.

Однако в последующем практика показала, что принцип единства ИО, ТО и ОргО как основных видов обеспечения, принятый в рассмотренной методике обоснования обеспечивающей части АСУ, важен при разработке первой очереди АСУ, когда принимаются принципиальные решения по выбору технических средств и структуры организационного обеспечения.

Следующие очереди разрабатываются в условиях выбранных технических средств и развитие технической базы осуществляется с учетом

совместимости с существующими, а главной задачей становится формирование структуры *информационного обеспечения*.

Поэтому в методике разработки структуры ФЧ и ОЧ для предприятия (организации) предусматривается два этапа:

1. Разработка (развитие) структуры функциональной части АСУ.

При выполнении этого этапа разрабатывается структура целей и функций организации (подэтап 1.1), проводится оценка составляющих этой структуры (подэтап 1.2.) с использованием *экспертных* и *косвенных количественных* оценок и на этой основе определяется состав подсистем и задач, подлежащих автоматизации в первую очередь, т. е. сформировать структуру ФЧ АИС.

2. Определение состава информационных массивов, необходимых для реализации подсистем и задач ФЧ АИС.

Эту задачу, как правило, можно выполнить, задав предварительно совокупность информационных массивов (с учетом уже созданных), и сопоставив их с подсистемами и задачами ФЧ АИС, определить значимость массивов для реализации этих задач и функции, которые не обеспечены информацией, и дополнить совокупность массивов новыми.

Пример методики приведен на рис. 4.18.

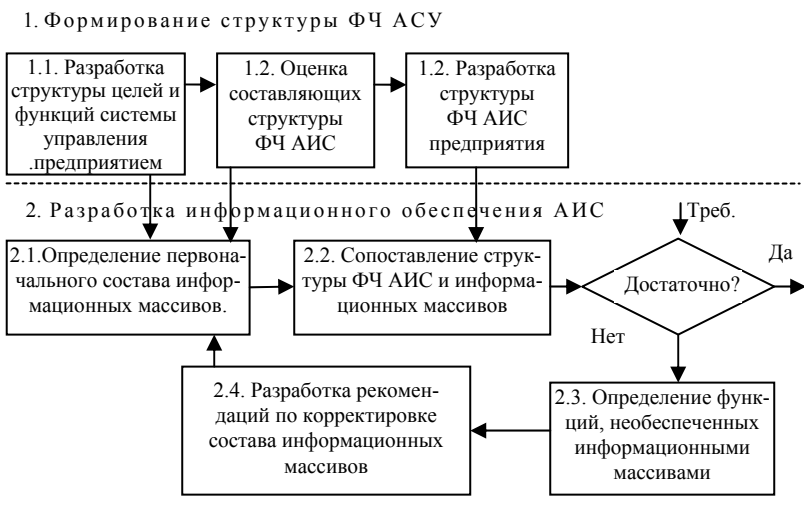


Рис. 4.18. Структура методики проектирования информационной системы

При выполнении подэтапа 1.1. применяются методики структуризации целей систем управления и автоматизированные диалоговые процедуры анализа целей и функций

При выполнении подэтапа 1.2 для оценки структуры целей и функций путем выявления наиболее значимых составляющих предлагается параллельно используются *экспертные оценки* и *косвенные количественные оценки*, которые затем обрабатываются совместно.

При экспертной оценке в качестве критериев используется система критериев, аналогичная принятой в методике ПАТТЕРН¹, но с некоторыми модификациями: в ней предлагаются следующие группы критериев – относительной важности (значимости), взаимосвязанности, экономические оценки (последняя – заменяет группу критериев «состояние-срок»). Идея косвенных количественных оценок предложена² на основе анализа иерархических структур с использованием информационного подхода из результатов которого следует, что структурированность ветвей иерархической структуры определяет придаваемую им фактическую значимость.

В качестве косвенных количественных оценок могут быть использованы: число подразделений, выполняющих данную функцию, число документов, подготавливаемых для реализации функции, частота обращения к выполнению функции и информационным массивам, обеспечивающим ее выполнение, внимание к соответствующим подцелям и функциям в директивных документах, периодической печати, источниках НТИ и т.п. При выборе косвенных количественных оценок следует учитывать «пространство инициирования целей и факторов», т.е. учитывать требования и потребности надсистемы, отраженные в законодательных актах и директивных документах, аналогичных предприятий актуальной среды, интересы подведомственных подразделений, инициативы структурных единиц собственно системы управления

При обработке результатов оценки применяются не только традиционные методы усреднения, но и выявление *противоречивых мнений* с последующим содержательным анализом этих оценок. При этом для сопоставления оценок удобно применять графические представления в виде гистограмм.

При выполнении подэтапа 1.3. на основе полученных результатов оценки исходная структура корректируется составляющие, получившие наименьшие оценки значимости по сравнению с другими и не получившие при этом высоких оценок связности с высокочисленными, либо исключаются из структуры ЦФ, либо опускаются на нижележащие уровни иерархии, и, напротив, составляющие, получившие высокие оценки значимости, могут быть перенесены на более высокие уровни иерархической структуры. При такой корректировке могут возникнуть вырожденные ветви, разные варианты новой структуры целей и функций. В последнем случае следует перейти к подэтапу 1.1 и повторить формирование и оценку структуры ЦФ.

¹ *Лопухин М.М.* ПАТТЕРН – метод планирования и прогнозирования научных работ / М.М. Лопухин. – М.: Сов. радио, 1971. – 160 с.

² *Волкова В.Н.* Оценка целевых структур при разработке планов в системе высшей и средней специальной школы: Экспресс-информация / В.Н. Волкова, А.И. Коношенко, А.Ф. Марьенко – М.: НИИВШ, 1980. – 19 с.

На основе этого анализа формируется структура ФЧ, в которой компоненты первого сверху уровня структуризации первоначально называли подсистемами, следующий – группами задач или задачами. В последующем по мере развития АСУ названия менялись, и при необходимости формировались многоуровневые структуры ФЧ (см. структуру ФЧ АСУ ВАЗа). Но для малых и средних предприятий обычно сохранялось представление структуры ФЧ в виде древовидной иерархической структуры.

При разработке информационного обеспечения (этап 2) в совокупности массивов могут быть массивы, обеспечивающие информацией производственный процесс, материально-техническое снабжение и т.п. массивы фактографической информации, и массивы научно-технической информации. После того, как определен окончательный состав массивов, разрабатывается структура информационного обеспечения, т.е. определяется структура входных, архивных, информационно-поисковых массивов, взаимосвязи между массивами и т.п.

Принципы построения и эффективность АСУ существенно зависят от уровня развития информационных технологий. С появлением в середине 1970-х гг. персональных ЭВМ происходит корректировка идеи АСУ: от ВЦ и централизации управления – к распределенному вычислительному ресурсу и децентрализации управления. Для обеспечения информацией по группам основных функций организационного управления предприятиями в настоящее время приобретают готовые программные продукты. Вариант методики для выбора готовых программных продуктов приводится в следующем разделе. Особенно актуален выбор готовых программных продуктов для предприятий малого и среднего бизнеса, которые не имеют средств для проведения собственных разработок по созданию программного обеспечения для создания необходимых информационных массивов.

4.6. Выбор готовых программных продуктов

Предприятия малого и среднего бизнеса, как правило, не имеют средств для внедрения единой автоматизированной системы, поэтому существует практика выбора имеющихся на рынке программных продуктов, автоматизирующих соответствующие области управления (бухгалтерский учет, материально-техническое обеспечение и т.п.) с последующей интеграцией в единую систему.

При выборе готовой информационной системы (ИС) обычно учитывают технические и стоимостные характеристики: количества рабочих мест (с возможностью одновременного доступа), разделение функций пользователей, скорость функционирования системы, сложность ее администрирования, проработанность механизмов, обеспечивающих права доступа, стоимость ИС и т.п. В то же время важно учиты-

вать такой критерий как полнота автоматизации функций управления и степень использования возможностей приобретаемой готовой ИС.

Поэтому методика должна быть основана на сопоставлении функций системы управления предприятия (или подсистемы АСУ) и функциональных возможностей существующих на рынке программных продуктов и определении *степени автоматизации* функций конкретного предприятия с помощью выбираемого программного средства и *коэффициент использования* этого средства.

Степень автоматизации определяется как отношению количества автоматизированных за счёт конкретного программного продукта функций f_{ia} к общему количеству функций f_{io} организации:

$$A = \frac{\sum f_{ia}}{\sum f_{io}}, \quad (4.4)$$

Степень использования функциональных возможностей автоматизированной системы.

$$k_n = \frac{\sum \varphi_{ia}}{\sum \varphi_{ie}}, \quad (4.5)$$

где φ_{ia} – число функциональных возможностей программного продукта, используемых в кадровой службе конкретной организации; φ_{ie} – общее число функциональных возможностей программного продукта.

Структура методик приведена на рис. 4.19.

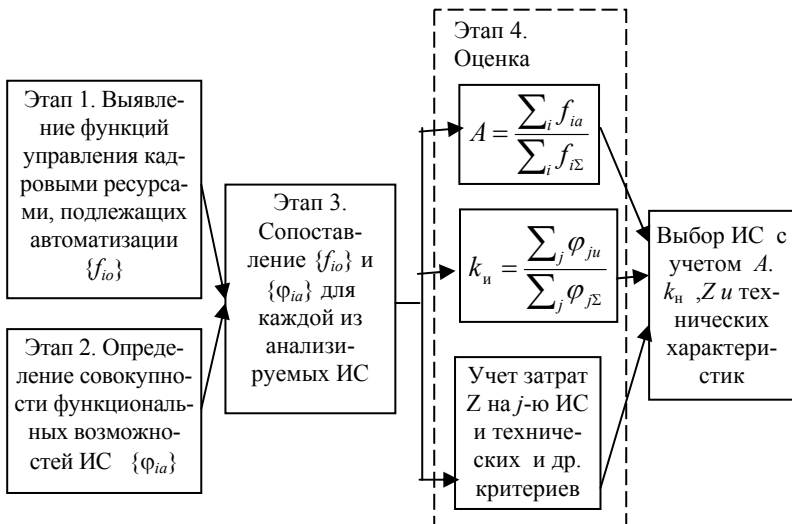


Рис. 4.19. Структура методики выбора ИС

Глава 5. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ ДОКУМЕНТАЛЬНЫХ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

5.1. Виды научно-технической информации и структура Государственной системы НТИ

Научно-техническая информация может быть представлена в *документальной* и *фактографической* форме. Но в любой форме она может быть *семантической* (содержательной и понятной для широкого круга пользователей) или *прагматической* (содержательной и существенной для принятия решений в конкретных условиях).

При этом производственная документация, оперативная информация о состоянии производства или рынка и другие виды фактографической информации, используемые в экономике для принятия текущих решений, не относят к научно-технической информации

Научно-техническую информацию классифицируют по разным признакам: по способу восприятия (визуальная, звуковая, тактильная (осязаемая); по форме представления (текстовая, графическая (изобразительная); по типу, по назначению; по способу распространения; по степени аналитико-синтетической обработки.

Основные виды научно-технической информации приведены в табл. 5.1. [15].

Таблица 5.1. Виды научно-технической информации

Основание деления	Наименование признака	Примечание
1. Назначение информации	Массовая	Предназначенная для всех, а не только для узких специалистов
	Специальная	Предназначенная только для специалистов в конкретной области
2. Тип информации	Документальная	Фиксированная в научных документах
	Фактографическая (точнее концептуально-фактографическая)	Идеи и факты, извлеченные из научных документов

Основание деления	Наименование признака	Примечание
3. Способ распространения информации.	Опубликованная. Неопубликованная (и не публикуемая).	Широко распространяемая посредством тиражного размножения научных документов, прошедших официальную регистрацию.. Не рассчитанная на широкое распространение и/или не прошедшая официальной апробации.
4. Степень аналитико-синтетической переработки.	Первичная. Вторичная.	Непосредственные результаты научно-исследовательской и опытно-конструкторской работы (монографии, статьи, отчеты). Результаты аналитико-синтетической переработки первичной информации, сведения о документах, содержащих научно-техническую информацию.
5. Области получения и/или использования информации в отраслях народного хозяйства и других областях человеческой деятельности.	Название отрасли науки или области применения	Астрономическая, биологическая, географическая, идеологическая, искусствоведческая, историческая, медицинская, промышленная, сельскохозяйственная, строительная, физическая, химическая, экологическая и т. п.

Для сбора, хранения и предоставления научно-технической информации создаются соответствующие хранилища.

За рубежом первоначально формировались *центры анализа информации*, постепенно укрупняемые путем объединения родственных центров, а в нашей стране создавались сразу как *интегральные комплексы*, ставящие своей задачей преодоление дублирования работ по индексированию, хранению и обработке НТИ и объединяемые Государственной системой НТИ.

Центры анализа информации

По мере осознания роли информации как важнейшего ресурса развития общества и исследования ее основных особенностей, в том числе таких, как тиражируемость и многократность использования, старе-

ние (и нередко связанное с этим «загрязнение» (*pollution*) информационного пространства, неаддитивность, некоммуникативность, кумулятивность, зависимость реализуемости и эффективности от степени использования информации и т.п., все больше осознавалась необходимость работ по систематизации, оценке и обобщению информации. Поэтому в 1960-е гг. стали создаваться центры анализа информации.

Термин «*Informations Analysis Center*» (IAC) был предложен Дж. Симпсоном (США) в 1964 г. [19]. В 1967 г. при консультативном Комитете по научно-технической информации президента США была создана Секция по центрам анализа информации (*Panel of IAC*). Эквивалентом этому термину в немецком языке служит *Wissensbewertungsstelle* («центр по оценке знаний»), а во французском *centre d'analyse de l'information*.

В практике создания таких центров сложилось 2 направления:

- центры, занимающиеся сбором, индексированием, хранением, поиском и распространением информации;
- центры, занимающиеся содержательной оценкой и интерпретацией научной информации.

К числу центров первого типа относят: научно-технические библиотеки; издательства, торговые ассоциации или профессиональные организации, занимающиеся изданием и распространением научных книг и журналов; информационные центры, занимающиеся подготовкой библиографических и реферативных изданий; центры, занимающиеся сбором, хранением и распространением документальной информации по собственной инициативе или запросам потребителей.

Центры второго типа являются по сути научными учреждениями и организациями, занимающимися оценкой и обобщением информации. Такие центры вначале создавались по отдельным отраслям фундаментальной науки (биология, физика, химия, науки об окружающей среде) или даже научным направлениям.

Вначале центры создавались независимо, часто дублируя работу друг друга. Например, в США к середине 1960-х гг. было создано более 100 центров разного рода, причем иногда по несколько в одной и той же области знаний (см. табл. в [15, с. 329]).

Затем стала наблюдаться тенденция к их объединению. Так:

* в 1965 г. министерство здравоохранения, образования и социальной политики США объединило 18 IAC в Информационный центр по ресурсам образования (*Educational Resources Informational Center – ERIC*);

* в 1966 г. был создан Комитет по научно-техническим данным (*Comitee on Data for Science and Technology – CODATA*), в который вошли представители международных союзов и ряда ведущих в тот период стран – СССР, США, Великобритании, ФРГ, Канады, Японии и др.;

* в 1970 г. три крупных информационных центра США (*BioScience Information Service of Biological Abstracts* – BIOSIS, *Chemical Abstracts Service* – CAS, *Engineering Index Inc.* – Ei) объединились для изучения проблем дублирования обработок и тех же публикаций и возможностей устранения или хотя бы сокращения такого дублирования путем кооперации и обмена информационными массивами на машинных носителях.

Примерами такого рода центров в нашей стране являются:

* созданный в 1952 г. Всесоюзный институт научной и технической информации (ВИНИТИ);

* созданная в 1965 г. Государственная система стандартных справочных данных (ГСССД), которая должна была обеспечивать не только сбор, обработку и предоставление информации о свойствах веществ и материалов, но и качественную переработку информации, оценку ее достоверности; при этом ГСССД сразу создавалась как разветвленная многоуровневая система с децентрализованным получением и распространением научной информации;

* центральные научно-исследовательские институты технико-экономической информации (ЦНИИТЭИ), которые создавались для ведущих отраслей промышленности (ЦНИИТЭИ Приборостроения, ЦНИИТЭИ Автопрома и т. п.).

Интегральные информационные системы

Для того, чтобы преодолеть проблемы дублирования работ по индексированию, хранению и обработке НИИ, «загрязнения» информационного пространства была предложена [15] идея создания *интегральных информационных систем*.

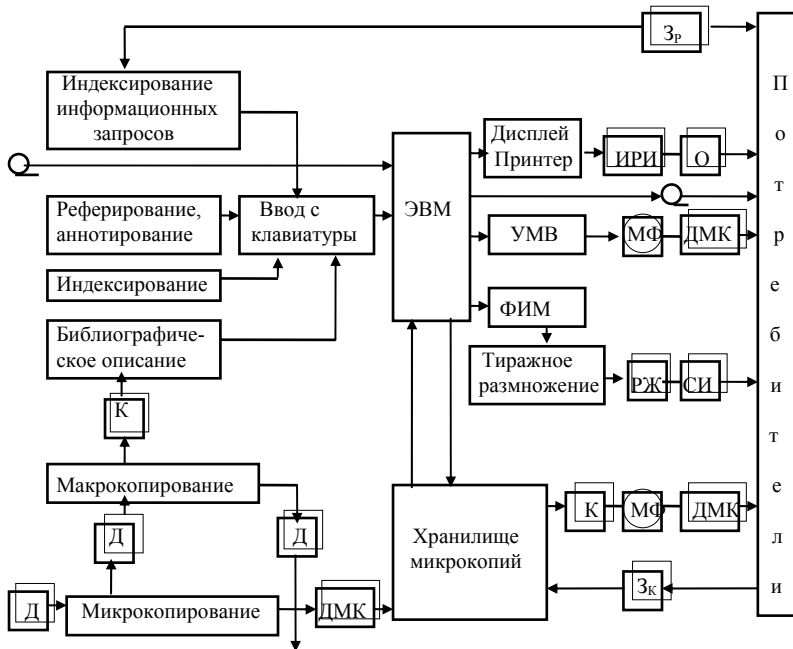
Интегральная информационная система (ИИС) – это совокупность методов и средств, позволяющих при *однократном* описании, индексировании и реферировании научных документов обеспечить *многоаспектную* обработку содержащейся в них информации и *многократное* ее использование для удовлетворения разнообразных информационных потребностей [16, с. 344].

Принцип однократного ввода и многократного использования информации, разумеется, реализовывался и в ряде зарубежных центров анализа информации. Но они, как правило, создавались для отдельных отраслей: по физике (в Американском физическом институте и в Институте инженеров-электриков в Великобритании), по ядерной физике и технике (Международное агентство по атомной энергии, Центр информации и документации Евратома), по медицине, химии и химической технологии и т. д.

В СССР замысел ИИС подразумевал не просто создание многоаспектных информационных банков или баз данных, а получение эффекта *целостности*, т. е. новых *эмерджентных* свойств, которых не было у элементов, объединяемых в ИИС.

Получение эффекта целостности базировалось на создании условий для совместной работы ученых и специалистов как тех, кто обеспечивает подготовку сигнальной информации и реферирование (осуществляемое высококвалифицированными учеными и специалистами, работающими, как правило, на внештатной основе), так и пользователей, которые влияют на возникновение новых системных свойств посредством обратной связи в режиме ИРИ.

Для того, чтобы реализовать этот замысел, была предложена принципиальная схема функционирования ИИС (рис. 5.1), и с 1971 г. в ВИНТИ начала разрабатываться интегральная информационная система АССИ-СТЕНТ (Автоматизированная Справочно-информационная СИСТЕма по Науке и Технике), в функции которой входило оказание помощи ученым и специалистам в решении не только научно-исследовательских, но и производственных, организационных, управленческих задач.



В книгохранилище

Рис. 5.1. Принципиальная схема функционирования интегральной ИИС

Д – документы; ДМК – диамикрокарты; З_р – разовый информационный запрос; З_к – запрос копии документа; ИРИ – избирательное распределение информации; О – ответы на разовые информационные запросы; К – макрокопии документов; МФ – рулонные микрофильмы; РЖ – реферативные журналы; СИ – буллетени сигнальной информации; УМВ – устройство микрофильмирования; ФНБ – фотонаборная машина.

С основными задачами и особенностями этой системы можно познакомиться в [15, с. 361–377], что может быть полезным для развития современных систем информационного обеспечения научных исследований.

Государственная система научно-технической информации

Для более полной реализации замысла целостной интегральной системы в 1970-е гг. в СССР было начато создание *Государственной системы научно-технической информации* (ГСНТИ) и поставлена задача разработки *Единой автоматизированной системы научно-технической информации*, в последующем переименованной в Государственную – ГАСНТИ. Руководил созданием ГСНТИ и ГАСНТИ Государственный комитет Совета Министров СССР по науке и технике.

В настоящее время в стране сохранилась общегосударственная система НТИ, которую условно называют Российской ГСНТИ (РГСНТИ), структура которой приведена на рис. 5.2 [1, 31].

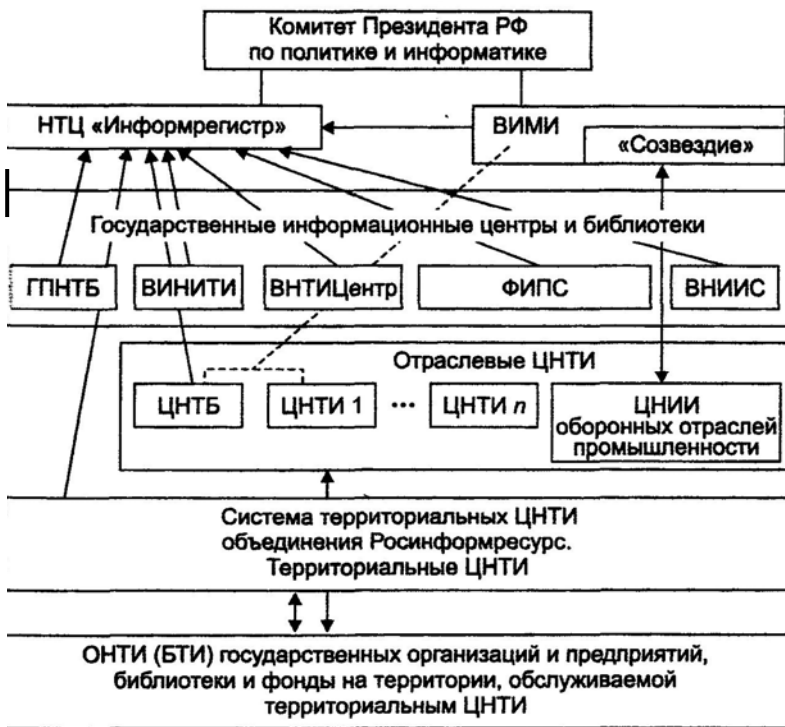


Рис. 5.2. Структура РГСНТИ

Структура РГСНТИ включает [31]:

1. Государственные информационные центры и библиотеки, основными из которых являются:

- Государственная публичная научно-техническая библиотека России (ГПНТБ России) – головной орган научно-технических библиотек, фонды которой составляют отечественные и зарубежные научные издания, неопубликованные переводы, авторефераты диссертаций и др. источники НТИ;

- Всероссийский институт научной и технической информации (ВИНИТИ), фонды которого содержат отечественные и зарубежные периодические издания, сборники трудов, реферативные журналы и ряд др.;

- * Всероссийский научно-технический информационный центр (ВНТИ-Центр), формирующий и сохраняющий документы государственной регистрации НИОКР, отчеты по НИР и ОКР, диссертации, фонды алгоритмов и программ;

- * Федеральный институт промышленной собственности (ФИПС) и Информационно-издательский центр Российского агентства по патентам и товарным знакам, накапливающий фонды отечественной и зарубежной патентной литературы, товарные знаки (бывшее НПО «РОСПАТЕНТ»);

- * Всероссийский научно-исследовательский институт стандартизации (ВНИИС) и Федеральный фонд государственных стандартов, общероссийских классификаторов технико-экономической информации, международных, региональных стандартов, правил, норм и рекомендаций по стандартизации, национальных стандартов зарубежных стран Федерального агентства по техническому регулированию и метрологии (в прошлом – ВНИИ классификации, терминологии и информации по стандартизации и качеству – ВНИИКИ, разрабатывающий и собирающий стандарты и др. нормативно-техническую документацию);

- * НТЦ «Информрегистр» Комитета при Президенте РФ по политике информатизации, занимающийся регистрацией и предоставлением сведений об имеющихся в стране информационных ресурсах на машинных носителях, в том числе библиографических и справочных БД.

2. Система территориальных центров НТИ (ЦНТИ) объединения Росинформресурс Миннауки РФ, включающая 69 центров.

Каждый из территориальных центров НТИ (ЦНТИ) включает научно-технические библиотеки, располагает уникальными фондами опубликованных и неопубликованных информационных изданий, отражающих сведения о научных и технических разработках, материальных и сырьевых ресурсах своих регионов, об организациях и предприятиях региона, его экономике и т.п. ЦНТИ взаимодействует с отделами НТИ (ОНТИ) государственных организаций и предприятий, частных фирм на территории обслуживаемых регионов, имеющих свои библиотеки и фонды документов, отражающие производственную, коммерческую и иную информацию (всего более 100 тыс. организаций).

3. Система НТИ министерств и ведомств России.

Центральные научно-технические библиотеки (ЦНТБ) и отраслевые ЦНТИ оборонных отраслей промышленности, объединяемые автоматизированной информационной системой «Созвездие» с центральным звеном во Всероссийском

научном институте межотраслевой информации (ВИМИ) при Комитете Президента РФ по политике информатизации.

ВИМИ выполняет функции методического и координационного центра для ведущих отраслевых ЦНТИ. Отраслевые ЦНТИ взаимодействуют с информационными подразделениями предприятий и организаций своей отрасли и создают справочно-информационные фонды, содержащие техническую, отраслевую нормативную, технологическую и др. виды документации.

4. Система территориальных ЦНТИ объединения Росинформресурс Миннауки РФ 69 территориальных ЦНТИ.

5. Отделы НТИ (ОНТИ) и бюро технической информации (БТИ) государственных организаций и предприятий, библиотеки и фонды на территории, обслуживаемой территориальным ЦНТИ (более 100 тыс.).

Особое место в в систем ГСНТИ занимаю Российская книжная палата, регистрирующая и хранящая информацию об опубликованных в РФ произведениях печати и государственной библиографии, и Российский государственный архив научно-технической документации Федеральной архивной службы России, содержащий информацию о документах научно-исследовательских, проектных, конструкторских, технологических организаций и предприятий федерального значения, передаваемых на постоянное хранение. Эти общегосударственные хранилища являются относительно независимыми в общей структуре ГСНТИ.

С начала 90-х гг. XX в. в компьютерных сетях началось предоставление возможности широкой аудитории пользователей знакомиться с электронными вариантами выпускаемых в стране компьютерных изданий.

5.2. Информационные потребности и информационное обслуживание

При работе с информацией (особенно с документальной) важное значение приобретает проблема определения потребителя (исследователя), и удовлетворения его потребностей.

Информационной потребностью (ИП) называют осознанную необходимость в знаниях или, другими словами, осознание недостаточности имеющихся знаний.

При выявлении информационных потребностей следует учитывать: *тип информации* (тематика, виды документов; *качество* и *количество* информации (достоверность, полнота, избыточность, надежность информационных источников); *способы представления* информации потребителю (устно, на экране, оригинал, копия, информационное издание, микрофильм, и т. п.); *сроки представления* с учетом минимизации разрыва между появлением, опубликованием информации и доведением ее до потребителя.

Для выявления и обеспечения информационных потребностей необходима система *информационного обслуживания* с соответствующей организацией информационных потоков.

Информационные потребности зависят от различных объективных и субъективных факторов. Различают *субъективные* ИП, которые определяются мнением и требованиями потребителя, его представлениями, знаниями, умением сформулировать запрос и потребности в объективно необходимой информации; и потребности в *объективно необходимой* информации, которая должна быть связана с задачей, проблемой, которую исследует потребитель в соответствии со своими служебными обязанностями, в частности, в системе управления предприятием (организацией).

При выполнении профессиональных обязанностей возникают потребности в текущей и ретроспективной информации, узко- и широко-тематической информации, отраслевой, межотраслевой, правовой, нормативной и др. видов информации.

Для более эффективной организации информационного обслуживания необходимо учитывать принципиальную особенность – субъективизм и неадекватность выражения информационных потребностей. Иными словами, информационная потребность субъекта определяется степенью понимания стоящих перед ним задач и знанием возможностей систем информационного обслуживания.

Изучение запросов потребителей информации является первоочередной задачей научно-информационной деятельности. Задача эта является весьма сложной, поскольку потребители, как правило, не умеют четко сформулировать свои потребности, и они уточняются по мере работы с получаемой информацией.

Существующие методы изучения информационных потребностей делят на две группы:

- *косвенные* или *документальные*, базирующиеся на анализе документальных источников, (карт обратной связи, запросов потребителей информации и т. п., независимо от того, запросили ли специалисты);
- *прямые*, связанные с участием потребителя информации, изучением задачи, проблемы, объекта, конкретной ситуации, которые позволяют выявить потребность в объективно-необходимой информации конкретных потребителей.

Прямые методы обычно реализуются с помощью анкетирования, интервьюирования и других экспертных методов. С их помощью изучают не столько информационные потребности, сколько представление субъекта об этих потребностях. Использование прямых методов позволяет не только обеспечить информационную поддержку членов коллектива, но и определить, какая именно информация нужна каждому для исполнения своих обязанностей. В то же время следует иметь в

виду, что решения, принимаемые на основе экспертных оценок, основаны на субъективных мнениях экспертов и далеко не всегда позволяют определить истинные информационные потребности.

Применение косвенных методов связано с изучением предметной области и функционально-должностных обязанностей специалистов. Здесь широко используются методы анализа систем, базирующихся на анализе структуры целей и функций деятельности потребителя информации, а также методы математического моделирования. В результате формируется информационная модель, отражающая потенциальные информационные потребности и позволяющая организовать информационное обслуживание в процессе управления предметной областью.

Проблема изучения информационных потребностей и запросов потребителей информации может решаться на основе статистических исследований информационных потоков, изучения их закономерностей (см. гл. 7), таких как экспериментально открытый закон рангового распределения *Дж. Ципфа*, теоретически уточняющий его закон *Б. Мандельброта*, закономерности упорядочения параметров документальных информационных потоков *С. Брэдфорда*, *Б. Викери*, закономерность концентрации-рассеяния, исследованная применительно к проблемам определения информационных потребностей *В.И. Горьковой* и т. п.

Перспективным представляется определение запросов потребителей на основе структуризации целей и функций деятельности руководителей (при обслуживании в режиме ДОР) или специалистов (обслуживаемых в режиме ИРИ).

Для решения проблемы обеспечения информационных потребностей в теории научно-информационного поиска предложены различные виды информационного обслуживания: *регламентное обслуживание по стандартным запросам (СЗ)*, *избирательное распределение информации (ИРИ)*, *дифференцированное обслуживание руководителей (ДОР)*, *ретроспективный поиск (РП) по произвольным запросам*.

Доступ к научно-технической информации – вид *информационных услуг, информационного обслуживания*, существующего еще до появления электронной вычислительной техники.

Механизация и автоматизация доступа к НТИ началась в 1970-е гг. При этом стали различать библиотечную информацию, ориентированную на широкий круг читателей, и специальную научно-техническую информацию, которую собирают, формируют справочно-информационные фонды и предоставляют специалистам особые, отделенные от библиотек, подразделения, создаваемые на предприятиях и в организациях – отделы научно-технической информации (ОНТИ) или бюро технической информации (БТИ).

В настоящее время информационное обслуживание (сервис) представляет собой область профессиональной деятельности, обеспечивающая поиск, обработку, оформление и распространение информации.

В процессе развития системы информационных коммуникаций сформировались три вида информационного обслуживания: *документальное, фактографическое и концептографическое.*

Документальное обслуживание общества и различных его институтов, в том числе науки, техники и экономики, традиционно обеспечивалось научно-технической информацией. Сущность его заключается в том, что информационные потребности удовлетворяются путем предоставления документов, сведения из которых потребитель извлекает самостоятельно. Развитие этого вида обслуживания сделало доступным массовому потребителю не только печатные и машиночитаемые документы, но и базы данных различных видов и наименований.

Дальнейшее совершенствование ИТ позволило перейти к *фактографическому* информационному обслуживанию. Его основу составляет предоставление потребителю сведений (данных, фактов) в соответствии с его запросом. Это более сложная форма организации обслуживания. Для формирования и реализации запроса требуется знать характеристики объекта поиска, связи между ними, т.е. необходимо множество сведений, позволяющих сформулировать запрос и понять его.

Различают два вида фактографического обслуживания: информационно-поисковое и информационно-решающее.

В первом случае выполняется лишь поиск информации по заданным характеристикам; во втором – обеспечивается предоставление информации об объекте, получаемой в результате работы алгоритмов преобразования данных.

Необходимость интерпретации, анализа, реферативной обработки полученных сведений (данных), наличие огромных объемов информации привели к возникновению и развитию *концептографического* информационного обслуживания. Оно предполагает предоставление потребителю так называемой ситуативной информации: интерпретации данных, оценки, рекомендации, прогнозы, реферативные обзоры, рецензии, переводы с иностранных языков и т.д. Другими словами, осуществляется предоставление информации, которой в явном виде в документах и базах данных может и не быть. Такая информация может быть получена на основе аналитико-синтетической обработки научно-технической информации или на основе анализа и обработки (на основе моделей принятия решений, моделей прогнозирования и т.п.) фактографической информации. Концептографическое обслуживание можно рассматривать как форму интерпретированного документаль-

ного и фактографического обслуживания, или как разновидность научной деятельности в области информационного анализа.

Существуют различные *виды* документального информационного обслуживания. Основными можно считать следующие:

- *регламентное обслуживание* в форме подготовки и издания обзорных и реферативных информационных, библиографических указателей и т.п., или обслуживание по стандартным запросам;
- оповещение отдельных специалистов (абонентов) о текущих публикациях, представляющих для них потенциальный интерес путем *избирательного* (адресного) *распределение информации (ИРИ)* по «профилям интересов», сформулированных потребителями, или *дифференцированного обслуживания руководителей (ДОР)* по мере поступления информации, необходимой для принятия управленческих решений;
- *ретроспективный поиск*, т.е. отыскание документов, в которых находятся сведения по конкретным запросам, в массиве всех накопленных источников информации, что может осуществляться либо в форме подготовки тематических подборок, либо в оперативном режиме.

Наряду с основными предлагались более полные классификации видов обслуживания. В частности, в [15, с. 247] приводится обобщенная классификация по нескольким признакам (табл. 5.2).

Таблица 5.2

Виды информационного обслуживания

Признак классификации	Вид обслуживания
<i>По источнику инициативы</i>	<ul style="list-style-type: none"> • принудительное • по запросам потребителей
<i>По типам документов</i>	<ul style="list-style-type: none"> • обслуживание опубликованными документами и/или их копиями • обслуживание копиями неопубликованных документов (отчетов по НИР, диссертаций, переводов и т.п.)
<i>По направленности или адресности</i>	<ul style="list-style-type: none"> • информационные издания (много адресов) • избирательное распределение информации (один адрес)
<i>По периодичности или срочности</i>	<ul style="list-style-type: none"> • текущее оповещение • ретроспективный поиск
<i>По способу доведения документов до потребителя</i>	<ul style="list-style-type: none"> • непосредственная передача документов или их копий потребителям • двухступенчатое обслуживание (вначале информационное издание-сигнал, а затем – копии заинтересовавших потребителя документов)
<i>По назначению, цели</i>	<ul style="list-style-type: none"> • ознакомление; • для практического использования

5.3. Понятие о документальном информационном поиске

Теория информационного поиска первоначально сформировалась при исследовании документальных информационно-поисковых систем (ДИПС) научно-технической информации.

Под *информационным поиском* в таких системах понимается некоторая последовательность операций, выполняемых с целью отыскания документов (статей, научно-технических отчетов, описаний к авторским свидетельствам и патентам, книг и т.д.), содержащих определенную информацию (с последующей выдачей самих документов или их копий), или с целью выдачи фактических данных, представляющих собой ответы на заданные вопросы [15, с. 248].

Массив элементов информации, в котором производится информационный поиск, был условно назван *поисковым массивом* [15].

Процесс поиска документов может быть формализованно описан в терминах теории множеств следующим образом: D – некоторое множество документов или библиотека (поисковый массив); Q – множество информационных запросов; R – отношение, свойство, при наличии которого любому $q \in Q$ ставится в соответствие подмножество $D' \subset D$, называемое ответом на информационный запрос.

Очевидно, что прочитать каждый документ информационной базы, чтобы найти необходимый, практически невозможно. Поэтому на протяжении истории развития информационного поиска разрабатывались и совершенствовались различные методы поиска.

Каждому документу, вводимому в поисковый массив, ставится в соответствие *поисковый образ документа* (ПОД), который представляет собой характеристику, отражающую основное смысловое содержание документа (этим ПОД отличается от кода, присваиваемого информационному элементу в фактографических информационных системах).

В виде такой же краткой характеристики – *поискового предписания* или *поискового образа запроса* (ПОЗ) должен быть сформулирован и информационный запрос. Благодаря этому процедура поиска может быть сведена к простому сопоставлению поисковых образов документов с поисковым предписанием.

Однако такое сопоставление допустимо лишь в случае, если ПОД и ПОЗ описаны в терминах единого языка.

Для реализации процедуры описания и документов и запросов с помощью ПОД и ПОЗ разрабатывают *информационно-поисковые языки* (ИПЯ), которые имеют различные возможности.

Простейшим ПОД может являться заглавие документа, переписанное в форме перечня входящих в него слов. При этом необходимо установить некото-

рые правила типа включения в ПОД существительных в единственном числе и именительном падеже, глаголов – в неопределенной форме или в виде отглагольных существительных и т.п. правил, общих для формирования ПОД и ПОЗ.

В более развитых вариантах для уточнения содержания документа можно в качестве ПОД использовать аннотацию или реферат документа, также переписанные в виде перечня входящих в них слов с учетом некоторых правил, общих для формирования ПОД и ПОЗ.

В практических условиях реализации ИПС число слов, включаемых в ПОД и ПОЗ, не должно быть слишком большим, поэтому вводится понятие *ключевые слова*, которые являются наиболее значимыми для отображения содержания документа.

Некоторые идеологи информационного поиска (например, Ч. Мидоу [14]) отождествляли понятия *ключевое слово* и *дескриптор*. Однако в дальнейшем эти понятия стали использоваться в разном смысле.

Для отображения ПОД и ПОЗ разрабатываются *информационно-поисковые языки* (ИПЯ).

Для сопоставления ПОД и ПОЗ применяют различные *критерии поиска* или *критерии смыслового соответствия* (КСС).

Точность поиска определяется точностью отображения документов и запросов с помощью их поисковых образов и степенью совпадения ПОД и ПОЗ. Поэтому возможна неполнота выдачи документов, либо, напротив, «поисковый шум», которые представляют собой своеобразную плату за облегчение процедуры информационного поиска.

Для оценки качества поиска вводят специальные критерии. Обобщенные критерии качества поиска – *релевантность* и *пертинентность* [15, с. 282–287].

Под *релевантностью* понимается соответствие выдачи запросу, т.е. релевантность характеризует качество алгоритма поиска.

Под *пертинентностью* понимается соответствие выдачи потребностям лица (или лиц), для которого (которых) осуществляется поиск информации, т.е. пертинентность характеризует смысловыражающие возможности ИПЯ, точность отображения с его помощью информационных потребностей.

Предлагаются и используются различные конкретные критерии количественной оценки релевантности и пертинентности. Разработка комплекса критериев оценки качества информационного поиска – достаточно сложная проблема: состав и количественные характеристики критериев зависят от конкретного назначения и принципов реализации ИПС.

Информационный поиск осуществляется с помощью информационно-поисковой системы, понятие о которой развивалось. Кратко рассмотрим представления об ИПС и ее структуре.

5.4. Понятие об информационно-поисковой системе и структуре ИПС

Структура и функционирование конкретной ИПС зависят от вида и состава информационных источников, от способов реализации информационного поиска. В то же время есть некоторые общие принципы построения и функционирования ИПС, которые кратко рассматриваются в данной главе.

Анализ определений ИПС. Под *информационно-поисковой системой* первоначально понималась некоторая совокупность или комплекс связанных друг с другом отдельных частей, предназначенных для выявления в каком-либо множестве элементов информации (документов, сведений и т.д.), которые отвечают на информационный запрос, предъявляемый к системе [15].

С учетом приведенного выше описания процесса информационного поиска ИПС (*IPS*) можно определить следующим образом:

$$IPS \underset{\text{def}}{\equiv} \langle D, Q, R, D' \rangle, \quad (5.1)$$

$D' \subset D$

где D – некоторое множество документов или библиотека (поисковый массив); Q – множество информационных запросов; D' – ответ на информационный запрос; R – множество отношений, свойств, при наличии которых любому запросу $q_i \in Q$ ставится в соответствие подмножество D' .

Более полно **А.И. Черный** предложил представить ИПС в виде совокупности четырех основных компонентов [27, с. 18]:

$$IPS \underset{\text{def}}{\equiv} \langle LS, D, TS, N \rangle, \quad (5.2)$$

где LS – *логико-семантический аппарат* (включающий информационно-поисковые языки – один или более, правила индексирования и критерии выдачи); D – *поисковый массив* (т.е. определенное множество снабженных поисковыми образами документов, в котором отыскиваются необходимые); TS – *технические средства* (т.е. какие-то приспособления или устройства, которые необходимы для записи и хранения поисковых образов, для хранения документов и осуществления процесса сопоставления поисковых образов документов с поисковым предписанием или поисковым образом запроса); N – *люди, взаимодействующие с системой* (т.е. те, кто пользуются данной ИПС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска, а также выполняют другие интеллектуальные операции, без которых невозможен информационный поиск).

Затем для обеспечения возможности автоматизации процедуры информационного поиска было предложено [15, 27] в ИПС выделить два уровня рассмотрения – *абстрактный* и *конкретный*.

Абстрактной ИПС была названа совокупность ИПЯ (*retrieval language – RL*), правил индексирования (*IND*) и критерия выдачи или критерия смыслового соответствия (*KSS*):

$$IPS \underset{def}{=} \langle RL, IND, KSS \rangle. \quad (5.3)$$

Конкретной ИПС названа практически реализованная система, включающая массив документов *D*, в котором производится информационный поиск, технические средства *TS* реализации ИПС, а также взаимодействующих с ней людей *N*.

Структура функционирования ИПС в таком понимании приведена на рис. 5.3.

В соответствии с рассмотренным выделением в ИПС *абстрактного* и *конкретного* уровней и с учетом особенностей хранения документальной информации (библиотеки, архивы и т.п. хранилища) процедуру информационного поиска документальной информации было предложено разделить на два контура [27]:

1) семантическое осмысление запроса и выдача адресов (шифров, кодов), соответствующих запросу документов; на рис. 5.3 этот контур показан сплошными линиями;

2) отыскание самих документов (вручную или с помощью специализированных технических средств, если ими оборудовано хранилище); на рис. 5.3 – штриховые линии.

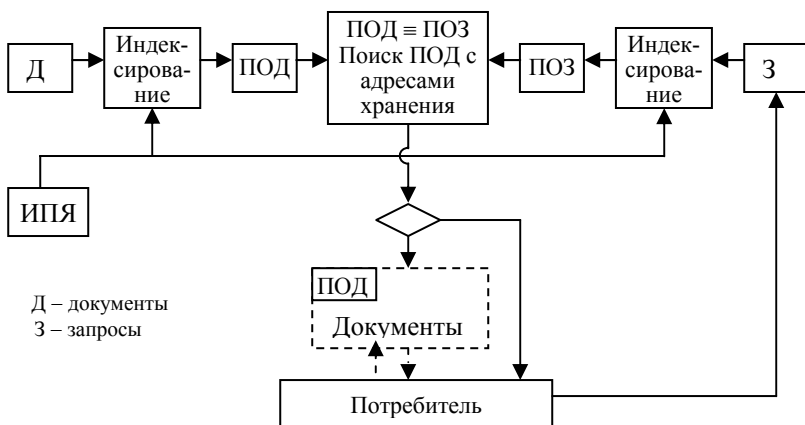


Рис. 5.3. Структура функционирования ИПС

Второй контур связан с разработкой специализированных технических средств хранения больших массивов документов и работой по переоборудованию хранилищ, а собственно проблемы информационного поиска решаются в первом контуре.

С учетом вышерассмотренного первый контур ИПС представляет собой ее логико-семантический аппарат и состоит из трех основных блоков (рис. 5.4):

информационно-поискового языка, системы перевода (индексирования) на этот язык и логики, обеспечивающей поиск, которые, в свою очередь, могут быть детализированы и реализованы разными способами.

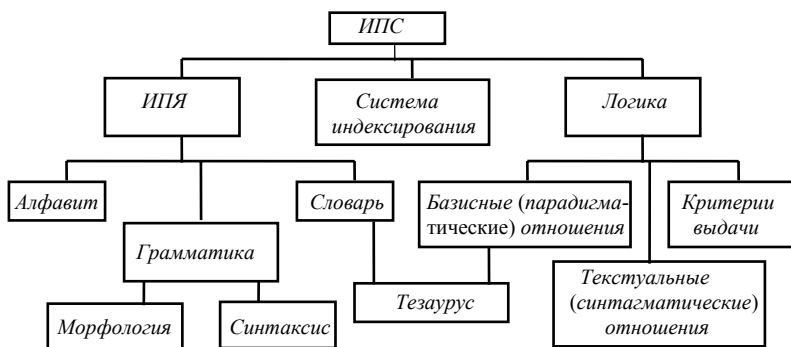


Рис. 5.4. Структура логико-семантического аппарата ИПС

Представление ИПС в виде двух контуров является в настоящее время наиболее распространенным.

В некоторых системах контуры могут быть совмещены.

Напротив, иногда возникает необходимость выделять не два, а большее число контуров, что помогает организовать последовательно углубленный анализ текстов документов. Такие варианты реализуются, например, в документально-фактографических системах нормативно-правовых и нормативно-методических документов.

В символической форме, принятой выше, *абстрактная ИПС (1-й контур)* представляет собой совокупность ИПЯ (RL), правил (системы) индексирования (IND) и логики (LOG), включающей наряду с критериями смыслового соответствия, базисные отношения:

$$IPS \underset{def}{=} \langle RL, IND, LOG \rangle. \quad (5.4)$$

В теории и практике различают ИПС разных видов.

Документальные ИПС (ДИПС) в ответ на вводимые в них информационные запросы выдают оригиналы, копии или адреса хранения документов, содержащих требуемую информацию.

Фактографические ИПС (ФИПС) предназначены для выдачи непосредственно требуемой информации (например, температуры кипения какой-либо жидкости, статистических показателей, содержащихся в соответствующих отчетных документах и т.п.).

При этом существуют фактографические системы двух видов:

1) системы, в которых сразу формируются массивы фактографической информации, параллельно с документальными;

2) системы, в которых массивы фактографической информации формируются на основе массивов документальной информации.

Информационные системы второго вида могут, в свою очередь, формироваться как *документально-фактографические* (ДФИПС и АДФИПС), содержащие массивы двух видов:

- *документальные* и сопряженные с ними массивы *фактографической* информации;

- *информационно-логические* ИПС.

В отличие от документальных, фактографических и документально-фактографических ИПС 1-го вида, которые могут по запросам выдавать только такую информацию, введенную в них ранее, *информационно-логические системы* представляют собой информационные системы более высокого класса: они должны выдавать не только ранее введенную в них информацию, но и производить, если необходимо, логическую переработку этой информации с целью получения новой информации, которая в явном виде не вводилась в ИС.

Информационно-логическую систему (ИЛС или *ILS*) можно определить как совокупность ИПЯ (*RL*), правил перевода с естественного языка на информационный, т.е. правил индексирования (*IND*) и правил логического вывода (*LV*), которая предназначена для алгоритмического получения новой информации (I_n) [15, 27]:

$$IPS \underset{def}{=} \langle RL, IND, LV, I_n \rangle. \quad (5.5)$$

Развивая представления об информационных системах, способных получать новую информацию, **Ю.И. Шемакин** предлагает понятие *информационно-семантической системы* [30, с. 60]:

$$ISS \underset{def}{=} \langle a, St, tp_{iss}, co, t_i \rangle, \quad (5.6)$$

где a – цель; St – структура; $tp_{iss} \in TP$ – подмножество технологических процессов для данной *ISS*; co – условия; t_i – время.

Входящие в определение (5.6) составляющие могут быть детализированы с учетом конкретной реализации ИПС. Особенно важно уточнять состав технологических процессов:

$$tp_{iss} \equiv \langle met, re, SemSI \rangle, \quad (5.7)$$

где *met* – методы; *re* – средства; *SemSI* – семантическая переработка семантической информации.

Для организации проектирования информационных систем **Ю.Ф. Тельнов** [22] предлагает определение, в котором учитываются и цели (запросы) *Q*, и элементы (внешние *E_n* и внутренние *E_i*), и отношения *R* (включая динамические взаимодействия), и функции *F* (процессы, операции), и период времени *T*, и закономерности *Z*, определяющие структуру системы и ее взаимодействие с внешней средой:

$$S \underset{def}{=} \langle \underline{Q}, E_i, E_n, T, F, R, Z \rangle. \quad (5.8)$$

Выбор определения ИПС зависит от конкретного объекта, для которого она разрабатывается, назначения и условий разработки и функционирования информационной системы.

5.5. Информационно-поисковый язык

Понятие об информационно-поисковом языке

Применение естественного языка для отображения ПОД и ПОЗ связано со значительными трудностями, обусловленными наличием в языке синонимов, омонимов и т.п. неоднозначностей использования терминов естественного языка. Поэтому на определенном этапе развития теории и практики создания ИПС вместо естественного языка стали применять искусственные информационно-поисковые языки – ИПЯ.

Существуют различные названия и определения специализированного языка, с помощью которого отражают основное содержание документов, вводимых в ИПС.

Информационно-поисковый язык (retrieval language) – «это специализированный искусственный язык, предназначенный для выражения основного содержания документов или информационных запросов с целью отыскания документов в некотором их множестве» [15, с. 259].

Информационно-поисковый язык (ИПЯ) используется для отображения содержания документов *информационно-поисковой системы* в *поисковом образе документа* – ПОД, и запроса – *поисковом образе запроса* – ПОЗ, или *поисковом предписании*.

Такой язык называли вначале *информационным языком (ИЯ)*, предъявляя к нему требование однозначной записи содержания документа; *языком индексирования (index language)*, определяемым как совокупность или система символов или индексных терминов и правил их использования для выражения предметного содержания документов; *документальным языком (langage documentaire)* и т.п. (подробнее с обзором этих терминов можно познакомиться в [15]).

В окончательном варианте понятийного аппарата теории информационного поиска утвердился термин *информационно-поисковый язык (retrieval language)*.

Обобщая различные представления об информационно-поисковом языке, можно дать следующее определение:

Информационно-поисковый язык (ИПЯ) является формализованной семантической системой, обеспечивающей передачу (запись) содержания документа в объеме, необходимом для целей поиска.

Документ, записанный на этом языке, может быть в принципе и не понят человеком, даже если в записи используются слова естественного языка, поскольку в ИПЯ употребление слов, выражений, отношений между ними стандартизировано определенным образом.

Задачей ИПЯ является перевод содержания документа в *поисковое предписание* или *поисковый образ документа* (при вводе документа в ИПС) и перевод содержания запроса пользователя в *поисковый образ запроса (поисковое предписание)*.

Первые исследователи в качестве составляющих ИПЯ выделяли: *алфавит* (набор буквенных и цифровых символов); *слова*, формируемые из алфавита с помощью морфологических правил – *морфологии*; *словарь* перевода (в котором каждому слову или осмысленной конструкции естественного языка сопоставлено слово или словосочетание ИПЯ); правила, отражающие взаимоотношения между словами документа, которые в конкретных ИПЯ реализуются, например, с помощью *текстуальных* или *контекстуальных* отношений, или с помощью специальных правил грамматики – *синтаксис*.

Словарь может состоять из *ключевых слов (словосочетаний)* или *дескрипторов*. Вначале некоторые авторы (например, **Ч. Мудой** [14]) отождествляли эти понятия и понимали под дескриптором все слова, выбранные для включения в словарь.

Однако в дальнейшем термину *дескриптор* стали придавать более сложный смысл: в отличие от *ключевых слов*, выбираемых предварительно из документов массива, для поиска в котором разрабатывается ИПЯ, под *дескриптором* понимается некоторый (выбранный разра-

ботчиком ИПЯ), обобщающий термин для отображения группы синонимов или слов, которые для целей поиска в конкретной ИПС можно считать синонимами.

Такие слова объединяют в класс *условной эквивалентности*, обобщаемый соответствующим *дескриптором*, и если в тексте документа или запроса встречается слово из данного класса, то его заменяют в ПОД или ПОЗ дескриптором.

Таким образом, *дескриптор* – специальное понятие, введенное и используемое в теории *информационного поиска* [14, 15, 27].

В современных информационно-поисковых языках под *дескриптором* понимают имя класса *условной эквивалентности* [15, 27].

Класс условной эквивалентности формируется из ключевых слов, связанных *парадигматическими отношениями*

Парадигматические (базисные) отношения – один из видов семантических отношений, предложенных в теории *информационного поиска* и применяемых при разработке *информационно-поисковых языков*.

Парадигматические отношения представляют собой внетекстовые смысловые отношения между лексическими единицами ИПЯ, которые устанавливаются на основании потребностей информационного поиска.

Роль парадигматических отношений сводится к следующему. Принципиальной особенностью естественного языка является тот факт, что в нем одни и те же события могут быть описаны в разных терминах. Тогда в *поисковом образе документа* – ПОД, и *поисковом образе запроса* – ПОЗ, могут быть использованы разные слова с сохранением смысла документа и запроса.

Кроме того, на практике может оказаться необходимым отыскивать документы, в которых речь идет о более частных понятиях, чем в ПОЗ. Не потерять такие документы может помочь введение парадигматических (базисных) взаимоотношений между дескрипторами ИПЯ.

В широком смысле в состав парадигматических отношений включают отношения синонимии (тождество означаемых при различии означающих), омонимии (тождество означающих при различии означаемых), отношения, основанные на одинаковости основы при различных окончаниях (парадигмы склонения и спряжения).

Однако в более узком смысле при разработке ИПЯ иногда предлагается под *парадигматическими (базисными) отношениями* понимать «*лишь такие отношения между словами (означающими), которые основаны на существовании тех или иных связей между означаемыми*» [15, с. 433].

Разные специалисты предлагают различные способы определения парадигматических связей: по *сходству предметов*, по *принадлежности к одному классу*, *ассоциативные отношения* (ассоциации по *смежности* в пространстве и во времени, по *сходству*, по *контрасту*, отношения *соподчинения*, «*вид-род*», «*причина-следствие*», «*часть-целое*» и т. п.).

При этом допускается произвольное установление отношений в конкретном ИПЯ, с ориентацией на повышение эффективности информационного поиска.

В частности, *Э.С. Бернштейн*, *Д.Г. Лахути* и *В.С. Чернявский*¹ использовали при разработке ИПС «Пусто-Непусто» парадигматические отношения, которые определяют как отношения, существующие между словами поискового языка независимо от контекста, называя именно их базисными отношениями (БО), и задавали их списком (включая в тезаурус). БО увеличивают семантическую силу системы, позволяют формулировать запросы в терминах, отличных от терминов, употребляемых в релевантных документах.

Фиксированные базисные отношения могут быть заданы различными способами: с помощью структуры слов (как в УДК), с помощью системы ссылок, с помощью деревьев дескрипторов и т. п.

Следует иметь в виду, что, стремясь улучшить результаты поиска, можно увеличить «шум», т. е. избыточную выдачу.

В первых работах по теории информационного поиска термин дескриптор использовался иногда [15] как синоним понятия *ключевое слово*

В различных языках эти компоненты ИПЯ используются по-разному. Словарь может иметь достаточно сложную структуру, т. е. представлять собой *тезаурус*, который может включать в себя и алфавит, и слова, и словосочетания, и более сложные конструкции.

Термин *тезаурус* (от греч. «θησαυροζ», «thesauros» – сокровищница, богатство, клад, запас и т. п.) в общем случае характеризует «совокупность научных знаний о явлениях и законах внешнего мира и духовной деятельности людей, накопленную всем человеческим обществом» [15, с. 85]. Этот термин был введен в современную литературу по языкознанию и информатике в 1956 году Кембриджской группой по изучению языков. В то же время термин существовал раньше: в эпоху Возрождения тезаурусами называли энциклопедии. С обзором определений тезауруса и первых тезаурусов можно познакомиться в [15, с. 415–432, 469–505].

¹ *Бернштейн Э.* Вопросы теории поисковых систем / Э. Бернштейн, Д. Лахути, В. Чернявский. – М.: ВНИИЭМ, 1966. – 64 с.

Особую роль в формировании тезауруса играют *базисные (парадигматические)* отношения, которые исторически являются элементом логики ИПС.

В математической лингвистике и семиотике термин тезаурус используется в более узком смысле, для характеристики конкретного языка, его многоуровневой структуры.

Для этих целей удобно пользоваться одним из принятых в лингвистике определений тезауруса как «*множества смысловыражающих элементов языка с заданными смысловыми отношениями*»¹.

Это определение иллюстрирует рис. 5.5 на элементарном примере формирования слов из букв и предложений их слов. Разумеется, в реальных тезаурусах уровни носят иные названия: ключевые слова, дескрипторы, абзацы и иные лингвистические и логические элементы.

При этом между уровнями тезауруса могут существовать различные взаимоотношения – от древовидных иерархических до причинно-следственных.

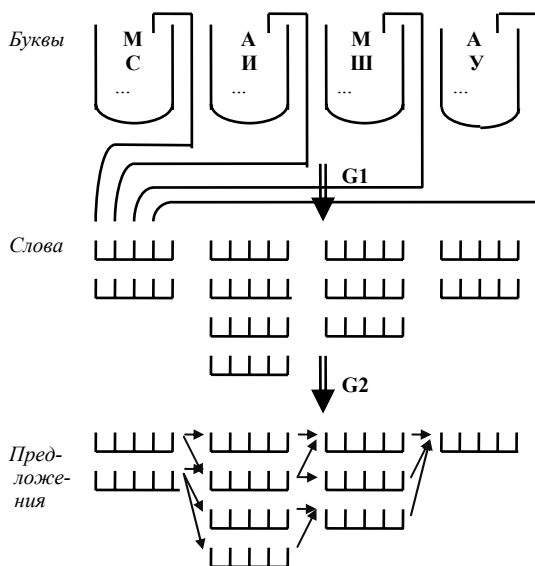


Рис. 5.5. Принципы формирования структуры тезауруса

¹ Шрейдер Ю.А. Информация в структурах с отношениями / Ю.А. Шрейдер // Исследования по математической лингвистике, математической логике и информационным языкам: сб. – М.: Наука, 1972. – С. 147–159.

Таким образом, тезаурус позволяет представить структуру языка в виде уровней (страт) множеств слов, предложений, абзацев и т.п., смысловыражающие элементы каждого из которых формируются из смысловыражающих элементов предшествующих структурных уровней.

Правила формирования смысловыражающих элементов второго, третьего и последующих уровней в тезаурус не входят. Они образуют грамматику информационно-поискового языка (*G1*, *G2* и т.д.). В тезаурусе же определяется только вид и наименование уровня, характер и вид смысловыражающих элементов.

Иногда вместо термина *смысловыражающие элементы* используется термин *синтаксические единицы* тезауруса. Однако, это менее удачный термин, так как при формировании элементов нового множества смысловыражающих элементов каждого последующего уровня (при образовании слов из букв, фраз и предложений из слов и т.д.) у элементов вновь образованного множества появляется новый смысл, т.е. как бы проявляется закономерность целостности, и это хорошо отражает термин «смысловыражающий элемент».

Понятие тезауруса стало в первую очередь использоваться при разработке информационно-поисковых языков, но в последующем его стали применять и при создании других искусственных языков – языков моделирования, автоматизации проектирования.

Тезаурус позволяет охарактеризовать язык с точки зрения уровня обобщения, ввести правила их использования при индексировании информации. В теории научно-технической информации [15, 16, 27, 30 и др.] исследуются различные свойства тезауруса.

Можно говорить о *глубине* тезауруса того или иного языка, характеризуемой числом уровней, о видах *уровней обобщения*, и, пользуясь этими понятиями, сравнивать языки, выбирать более подходящий для рассматриваемой задачи или, охарактеризовав структуру языка, организовать процесс его разработки.

В практике создания информационно-поисковых систем наиболее известен словарь-тезаурус «Тезаурус ASTIA»¹.

В системе SMART² содержится два вида тезаурусов:

- Тезаурус с иерархической структурой понятий.

Дает возможность для любого номера понятия найти их «родителя», «сыновей», «братьев» и множество возможных перекрестных ссылок.

- Словарь синонимов или тезаурус.

Используется для замены значащих слов номерами понятий, каждое из которых представляет класс основ слов, близких по смыслу.

¹ *Vikery B.C.* Thesaurus of ASTIA Descriptors, 2nd Ed //Armed Forces Technical Information Agency. – Arlington, Virginia: Dec. 1962.

² *Сэлтон Г.* Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. –М.: Сов. радио, 1973. – 560 с.

Тезаурусы разрабатываются и в отечественных отраслевых системах научно-технической информации (например, в АСНТИ-геология¹).

Простейшими тезаурусами являются словари *дескрипторов* при толковании дескриптора как имени класса условной эквивалентности, формируемого на основе *парадигматических отношений*.

Термин тезаурус иногда используется в более широком смысле. Например, **Ю.И. Шемакин** тезаурусом называет сложную систему организации в автоматизированных системах управления и обработки информации разных ее видов (научно-технической, управленческой, представляемой в документальной и фактографической форме) [29].

Морфологию и *синтаксис* удобно объединять единым термином – *грамматика*. Тогда говорят, что ИПЯ состоит из *тезауруса* и *грамматики*, а затем рассматривают смысловыражающие элементы (синтаксические единицы) тезауруса и правила грамматики.

Под *грамматикой* (которую иногда называют *синтактикой*, *синтаксисом*, что сужает понятие грамматики, исключая из него *морфологию*) понимаются правила, с помощью которых формируются смысловыражающие элементы языка. Пользуясь этими правилами, можно «порождать» (формировать) грамматически (синтаксически) правильные конструкции или распознавать их грамматическую правильность.

Простейшими правилами грамматики являются синтагматические (текстуальные) отношения.

При создании и использовании искусственных языков для информационно-логических систем применяют понятия математической лингвистики.

В частности, понятия *порождающей* и *распознающей* грамматики.

Под *порождающей* грамматикой понимается совокупность правил, с помощью которых обеспечивается возможность формирования (*порождения*) из первичных элементов (словаря) синтаксически правильных конструкций.

Под *распознающей грамматикой* – правила, с помощью которых обеспечивается возможность распознавания синтаксической правильности предложений, фраз или других фрагментов языка².

На базе математической лингвистики развивается *теория формальных грамматик Н. Хомского*.

¹ *Автоматизированная* система научно-технической информатики – разработка и эксплуатация / К.И. Володин, Л.Л. Гульницкий, И.Ф. Пожарский и др. – М.: Финансы и статистика, 2004. – 192 с.

² *Гладкий А.В.* Элементы математической лингвистики / А.В. Гладкий, И.А. Мельчук. – М.: Наука, 1969. – 88 с.; *Гросс М.* Теория формальных грамматик / М. Гросс, А. Лантен; под ред. А. В. Г л а д к о г о. – М.: Мир, 1971. – 294 с.

Классы формальных грамматик Н. Хомского¹ считаются основой теории формальных языков.

Формальный язык определяется как множество (конечное или бесконечное) предложений (или «цепочек»), каждое из которых имеет конечную длину и построено с помощью некоторых операций (правил) из конечного множества элементов (символов), составляющих алфавит языка.

Формальную грамматику определяют в виде четверки множеств

$$G = \langle V_T, V_N, R, A \rangle, \quad (5.9)$$

где V_T – множество основных или терминальных символов; V_N – множество вспомогательных или нетерминальных символов; R – множество правил вывода, или продужий, которые могут иметь вид

$$\alpha \rightarrow \beta, \quad (5.10)$$

где $\beta \in (V \cup V_N)$, т.е. β – цепочка конечной длины из терминальных и нетерминальных символов множеств V_T и V_N ; $\alpha \in (V_T \cup V_N)V_N(V_T \cup V_N)$, т.е. α является цепочкой из терминальных и нетерминальных символов, содержащей по крайней мере один нетерминальный символ из V_N ; A – множество аксиом (в грамматиках комбинаторного типа, к которым относятся грамматики **Н. Хомского**, A состоит из одного начального символа S , причем $S \subset V_N$).

Учитывая, что в литературе по формальным грамматикам, как правило, не стремятся к содержательной интерпретации получаемых выводов, а рассматривают лишь формальную сторону процессов порождения и распознавания принадлежности цепочек к соответствующему классу грамматик, приведем содержательный пример порождающей грамматики.

Предположим, дано: $V_T = \langle \epsilon_1, \epsilon_2, n, l \rangle,$
 $V_N = \langle S, P \rangle.$

Порождающая грамматика

Распознающая грамматика

$$R = \begin{cases} S \rightarrow SP & (1) \\ S \rightarrow \epsilon_1 S & (2) \\ S \rightarrow \epsilon_2 S & (3) \\ S \rightarrow n & (4) \\ P \rightarrow l & (5) \end{cases} \quad \begin{cases} SP \rightarrow S & (1') \\ \epsilon_1 S \rightarrow S & (2') \\ \epsilon_2 S \rightarrow S & (3') \\ n \rightarrow S & (4') \\ l \rightarrow P & (5') \end{cases} \quad (5.11)$$

Применяя правила R левой части (5.11) в приведенной последовательности, получим

$$S \Rightarrow SP \Rightarrow \epsilon_1 SP \Rightarrow \epsilon_1 \epsilon_2 SP \Rightarrow \epsilon_1 \epsilon_2 n P \Rightarrow \epsilon_1 \epsilon_2 n l.$$

(1) (2) (3) (4) (5)

¹ **Хомский Н.** Три модели для описания языка / Н. Хомский // Кибернетический сборник. – Вып. 2. – М.: Изд-во ИЛ, 1961; **Хомский Н.** Введение в формальный анализ естественных языков / Н. Хомский, Дж. Миллер // Кибернетический сборник: Новая серия. – Вып. 3. – М.: Мир, 1965.

Это – формальная сторона процесса порождения. Для того, чтобы получить интерпретируемое выражение, нужно расшифровать терминальные символы, включенные в V_N , где v_1 – ВСЕ, v_2 – ВОЗРАСТЫ, n – ПОКОРНЫ, l – ЛЮБВИ.

Тогда полученное предложение

« $v_1 v_2 n l$ » – «ВСЕ ВОЗРАСТЫ ПОКОРНЫ ЛЮБВИ».

Если изменить последовательность применения правил, то будут получаться другие предложения. Например, если применить правила в последовательности $(1) \Rightarrow (3) \Rightarrow (2) \Rightarrow (4) \Rightarrow (5)$, то получится «ВОЗРАСТЫ ВСЕ ПОКОРНЫ ЛЮБВИ». Если применить не все правила: например, $(1) \Rightarrow (2) \Rightarrow (4) \Rightarrow (5)$, то получим «ВСЕ ПОКОРНЫ ЛЮБВИ».

Если же попытаться получить предложение, как у А.С. Пушкина – «Любви все возрасты покорны», то, как бы мы не меняли последовательность правил, получить эту фразу не удастся. Нужно изменить первое правило: вместо $S \rightarrow SP$ включить в R правило $S \rightarrow PS$.

Из примера видно, что вид порождаемых цепочек (предложений) зависит от вида правил (исчисления) и от последовательности их применения (алгоритма).

С помощью приведенного примера легко также продемонстрировать тесную связь понятия «грамматически правильный» с языком (грамматикой).

Распознающая грамматика для рассматриваемого примера будет содержать как бы «перевернутые» правила – правая часть (5.11), которые должны применяться в обратной последовательности. Пример представления анализа правильности предложения с помощью правил распознающей грамматики приведен на рис. 5.6.

При распознавании правильности предложения если не оговаривать, что предложение (цепочка) грамматически правильно с точки зрения правил данного формального языка, то можно, пользуясь формальной грамматикой в первоначальном виде, получить вывод, что приведенная фраза Пушкина грамматически неправильна с точки зрения правил грамматики (5.11).

Действительно, с точки зрения правил грамматики для построения делового текста, которым соответствуют правила (5.11), другие поэтические строки часто получали бы формальную оценку «грамматически неправильно». И, напротив, если построить грамматику на основе анализа пушкинского стиля, то в деловом тексте получились бы предложения типа «Я решение свое принял правильно» (подобно фразе «Я памятник себе воздвиг нерукотворный»).

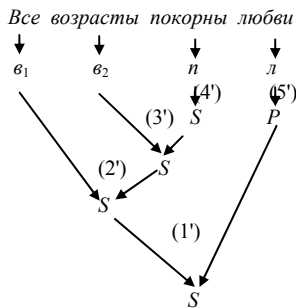


Рис. 5.6

Понятие формальной грамматики используют при создании языка моделирования соответствующего литературного или музыкального произведения – пародий, подражательств или, как иногда принято говорить, произведений соответствующего стиля или класса.

Например, известны работы **Р.Х. Зарипова**¹ по моделированию музыкальных произведений в стиле, или в классе, массовых советских песен, моделирование процесса сочинения стихотворных произведений и т. п.

Подобным же образом можно моделировать порождение деловых писем или других документов, имеющих, как правило, не только формализованный стиль, но и формальную структуру. Аналогично можно создавать языки моделирования структур, языки автоматизации проектирования сложных устройств и систем определенного вида (класса).

Основу таких работ составляют идеи, которые можно пояснить с помощью классов грамматик, впервые предложенных **Н. Хомским**².

Разделение грамматик на классы определяется видом правил вывода R . В зависимости от них можно выделить четыре основных, наиболее часто рассматриваемых класса грамматик (в полной теории формальных грамматик с правилами типа подстановки есть и промежуточные классы) (табл. 5.3).

Таблица 5.3

Основные классы грамматик Н. Хомского

Класс	Характеристика
1-й класс. <i>Неукорачивающие</i> (НУ-грамматики)	На правила вывода накладывается только одно требование, чтобы в левой части правила вывода было всегда меньше символов, чем в правой, т. е. чтобы правила были неукорачивающими, не уменьшали число символов в выводимых цепочках. Иногда эти грамматики называют грамматиками типа ноль (<i>нулевого типа</i>) или <i>алгоритмическими</i> .
2-й класс. <i>Нонтекстные, контекстно-связанные</i>	На правила вывода, помимо требований неукорачиваемости, накладывается ограничение, чтобы на каждом шаге изменялся только один символ в контексте, т. е. чтобы $Z1 B Z2 \rightarrow Z1 W Z2$, где B – один нетерминальный символ, W – непустая цепочка символов, т. е. $W \neq \emptyset$. Иногда применяют термин – <i>грамматика непосредственных составляющих</i> (НС-грамматики).
3-й класс. <i>Контекстно-свободная</i> (КС-грамматика) или <i>бесконтекстная</i>	Кроме неукорачиваемости требуется, чтобы правила имели вид $B \rightarrow \beta$, т. е. α всегда состоит из одного вспомогательного символа.
4-й класс. <i>Автоматные</i> (А-грамматики)	На правила вывода накладывается по сравнению с третьим классом еще одно ограничение, требующее, чтобы в правилах вывода нетерминальный символ всегда стоял справа или слева, т. е. с одной стороны. Если нетерминальный символ стоит слева, т. е. правила имеют вид $A \rightarrow aB$ или $A \rightarrow a$, где $(A, B) \in V_N$, $a \in V_T$, автоматная грамматика является <i>правolineйной</i> ; если нетерминальный символ стоит справа – то автоматную грамматику называют <i>леволинейной</i> .

¹ **Зарипов Р.Х.** Машинный поиск вариантов при моделировании творческого процесса / Р.Х. Зарипов. – М.: Наука, 1983. – 232 с.

² **Хомский Н.** Три модели для описания языка / Н. Хомский // Кибернетический сборник. – Вып. 2. – М.: ИЛ, 1961.

В теории формальных грамматик показано, что имеет место следующее соотношение:

$$A \subseteq KC \subseteq HC \subseteq HY. \quad (5.12)$$

Иногда доказывают, что имеет место строгое вхождение

$$A \subset KC \subset HC \subset HY. \quad (5.12,a)$$

При исследовании разных классов формальных грамматик получены результаты, которые позволяют сделать вывод, что по мере уменьшения числа ограничений, накладываемых на правила вывода, т.е. по мере продвижения в (5.11) слева направо, в языке увеличивается возможность отображения смысла, т.е. возможность выражения с помощью формальных правил семантических особенностей проблемной ситуации. Говорят, что формальная система становится более богатой. Однако при этом в языке растет число алгоритмически неразрешимых проблем, т.е. увеличивается число положений, истинность или ложность которых не может быть доказана в рамках формальной системы языка.

Здесь мы сталкиваемся фактически с проблемой Гёделя, которая в теории формальных языков обсуждается обычно в терминах этой теории. А именно: вводится понятие «*операция определена* (или *не определена*) на множестве языков данного класса»; и считают, что операция определена на множестве языков данного класса, если после применения ее к языкам, входящим в это множество, получается язык, принадлежащий множеству языков этого класса.

Например, если $Y_1 \subset KC$ и $Y_2 \subset KC$, и если $(Y_1 \cup Y_2) \subset KC$, то операция объединения \cup определена на классе KC -языков.

Характеризуя с помощью введенного понятия классы языков, отмечают, что в последовательности (5.11) по мере продвижения слева направо увеличивается число операций, которые не определены на множестве языков данного класса.

Здесь, правда, следует оговорить, что дело обстоит не так прямолинейно. Точнее было бы сказать, что для большого числа операций нет доказательств, что они определены на классах HC -языков и HY -языков, т.е. эти доказательства становятся сложнее или вообще (в силу теоремы Гёделя) нереализуемы средствами теории формальных грамматик.

Приведенное упрощенное представление проблемы помогает обратить внимание тех, кто будет заниматься разработкой языков программирования или программных систем, языков моделирования, автоматизации проектирования, на необходимость учета следующей закономерности: *чем большими смысловыражающими возможностями обладает знаковая система, тем в большей мере растет в ней число алгоритмически неразрешимых проблем* (т.е. тем менее доказательны в ней формальные процедуры).

При выходе в класс произвольных грамматик, в котором не выполняется даже условие неукорачиваемости, доказать допустимость тех или иных формальных преобразований средствами математической лингвистики практически невозможно, и поэтому в поисках новых средств исследователи обратились к семиотическим представлениям. Здесь можно провести как бы формальную границу между лингвистикой и семиотикой.

При создании ИПЯ с тезаурусом и грамматикой важную роль играют понятия *семантики* и *прагматики*.

Под *семантикой* понимается *с о д е р ж а н и е, з н а ч е н и е, с м ы с л* формируемых или распознаваемых конструкций языка; под *прагматикой* – *полезность для данной цели, за д а ч и*.

В естественном языке различить понятия, с помощью которых характеризуются термины «*семантика*» и «*прагматика*», трудно; обычно пояснить различие можно лишь при парном сопоставлении терминов:

<семантика>	:: = <содержание>	<смысл>	<значение>;
<прагматика>	:: = <смысл>	<значение>	<полезность>.

Поэтому принято рассматривать эти понятия на примерах. Поясним различие между семантически и прагматически правильными конструкциями языка на следующих легко запоминающихся примерах.

Традиционно для пояснения синтаксической правильности и семантической бессмыслицы используется предложенный *Л.В. Щербой* пример «*Глокая куздра тцето борзданула бокра и курдычет бокрѣнка*» (в котором просто нет ни одного слова естественного языка, имеющего смысл). Но примеры можно найти и в естественной речи.

Предложение «*Муха лукаво всплеснула зубами*» синтаксически правильное, но не имеет смысла в естественном русском языке в обиходном, широком употреблении, т.е. является с точки зрения пользователей русским языком семантически неправильным (исключим пока гипотетическую ситуацию сказки, в которой муха может быть наделена указанными свойствами).

Другое предложение «*Маленькая девочка собирает цветы на лугу*» – синтаксически и семантически правильное. Однако для директора завода (если это луг, а не заводской газон, и – учтем личный фактор – если эта девочка не его дочь) это предложение не несет никакой информации, т.е. прагматически (с точки зрения целей руководителя) является неправильным. Другое дело, если «Иванов (который в данный момент должен находиться на рабочем месте) собирает цветы на лугу». Тогда это предложение было бы и прагматически правильным.

Возвратимся теперь к примеру с мухой. Приведенное предложение, *семантически неправильное*, может в гипотетической ситуации сказки оказаться *прагматически правильным*, что важно иметь в виду при применении лингвистических представлений.

Пользуясь этими правилами, можно лучше отразить смысл документа или запроса в ПОД и ПОЗ, повышая релевантность поиска.

Виды и классификации ИПЯ

В зависимости от используемых компонентов ИПЯ бывают разных видов.

В качестве первоначально использовавшихся ИПЯ **Ч. Мидоу** [14] выделяет следующие (рис. 5.7).

Иерархические классификации.

Например, Десятичная классификация *Дьюи*¹, классификация библиотеки Конгресса США, Универсальная Десятичная классификация (УДК)², которая представляет собой модификацию системы Дьюи, получившую широкое применение.

Иерархические классификации обеспечивают возможность расширения «вниз», т. е. уточнения описания документа, но они отличаются жесткостью, их достаточно трудно изменять.



Рис. 5.7. Классификация ИПЯ **Ч. Мидоу**

¹ *Dewey V.* Dewey Decimal Classification and Relative Index, Forest Press, Lake Placid Club / V. Dewey. Essex Country, New York, 1959.

² *Сборник* методических материалов по Универсальной Десятичной Классификации. – М.: Гос. Комитет Совета Министров СССР по науке и технике, ГПНТБ, 1969. – 485 с.

- *Язык предметных заголовков*

Подобно иерархической классификации использует фиксированное число предметных классов (часто располагаемых по алфавиту), но для его терминов обычно не используется определенный код. Язык позволяет любому документу приписывать более чем один термин, почти не имеет структуры и средств для выражения взаимоотношений между терминами. Предметные заголовки используются, например, в журналах (рубрики), в классификаторах специальностей вузов и ВАК. Этот язык можно считать языком иерархической классификации, но с ослабленной структурой, что облегчает его разработку, но затрудняет изучение и применение.

Рассмотренные языки иногда называют *предкоординированными* (*pre-coordinate*) системами, поскольку семантические комбинации терминов не определены в словарном составе ИПЯ, а составляются его разработчиками. Такие языки неизбежно имеют пробелы, в них трудно отражать новые предметные области (что легко видеть на примере широко применяемой системы УДК).

Поэтому для более совершенного описания содержания документа стало применяться индексирование ключевыми словами – *координатное индексирование*:

- *Система ключевых слов с фиксированным словарем.*

Такие языки могут применяться для узкоспециализированных ИПС с достаточно формализованной (унифицированной) терминологией (напр., система «УниTERM», разработанная *М. Таубе*¹). Эта система подобна системе предметных заголовков, но в отличие от нее, во-первых, ключевые слова короче предметных заголовков (обычно это единичные слова, иногда – короткие словосочетания), и во-вторых, объем полного словарного состава существенно больше. Словарный состав здесь, как правило, фиксирован, и отсутствуют средства установления связей между словами (синтаксис), но включение набора слов в ПОД или ПОЗ позволяет пользователю как бы угадывать эти связи, что помогает более полно описать исходный документ или запрос.

- *Система ключевых слов со свободным словарем.*

Такие языки позволяют пользователю выбирать для описания документа любые слова (за исключением союзов и предлогов), руководствуясь их ролью в отображении содержания документа. Это позволяет точнее отобразить содержание документа, но может снизить релевантность поиска, поскольку составители ПОЗ не могут предусмотреть точки зрения составителей ПОД.

- *Языки с синтаксисом (грамматикой).*

В качестве простейшего из ИПЯ этого вида *Ч. Мудду* [14] выделяет язык *помеченных дескрипторов* (*tagged descriptors*), с помощью которого отобра-

¹ *Taube M.* Unit Terms in Coordinate Indexing / M. Taube, C.D. Gull, I.S. Wachtel // Amer. Documentations. – 1952. – № 3, 4. – P. 213–218.

жение смысла в ПОД и ПОЗ осуществляется путем присоединения к основному дескриптору (или ключевому слову) уточняющих дескрипторов (ключевых слов), роль которых состоит в том, чтобы либо классифицировать основной дескриптор как имя собственное, характерный признак или действие, либо объединить в одну группу дескрипторы, относящиеся к одному и тому же предмету документа. В современных поисковых системах Интернет применяются операции AND, OR, отображающие логические операции дизъюнкции и конъюнкции), в простейшем варианте – слова заключают в кавычки.

В дальнейшем были разработаны ИПЯ с грамматикой, содержащей более сложные правила.

- *Язык фасетного индексирования.*

В более развитых в синтаксическом отношении вариантах такого языка различные дескрипторы могут изменять значения друг друга. Простейшим примером такого синтаксиса является запись команды ЭВМ, состоящей из собственно оператора и адреса хранения информации. В качестве примера можно также привести уточненное описание товара, включая фасон, цвет, цену и т.п. характеристики товара. Такой синтаксис основан на известном в теории множеств положении: в результате помещения рядом элементов разных множеств возникает эффект появления нового смысла. Языки такого вида позволяют частично устранять омонимию с учетом контекста. Различные роли, которые играют дескрипторы в таких языках, называют *фасетами* ¹.

Для фасет могут быть использованы дескрипторы из одного и того же словаря. Располагаются фасеты в порядке значимости дескрипторов для отображения содержания индексируемого документа. В отличие от иерархических классификаций фасеты можно располагать в произвольном порядке.

- *Язык фраз.*

В качестве ПОД используются индексирующие фразы. В этом случае контекст ключевых слов позволяет частично снять проблемы семантической неоднозначности.

Трудность здесь состоит в выборе фраз, включаемых в язык. Кроме того, отсутствует возможность расширения ПОД.

- *Язык пермутационного индексирования.*

Пермутационный указатель включает контекст каждого слова, содержащегося в фразе и называется указателем ключевых слов, взятых в контексте, или указателем типа KWIC ². Идею такого указателя легче пояснить примером:

Системы индексирования **документов**

Системы **индексирования** документов

Системы индексирования документов

¹ *Vikery B.C.* On Retrieval Systems Theory / B.C. Vikery. – London: Butterworth, 1961, 33 p.

² *Vikery B.C.* General Information Manual, Keyword-in-Context (KWIC) Indexing / B.C. Vikery. New York: IBM Corp., White Plains, 1962.

При этом ПОД образует колонка ключевых слов в центре, расположенных в порядке алфавита. Можно считать этот ПОД фразой, но он эффективнее предыдущего с точки зрения смысловыражающих возможностей. Однако пермутационное индексирование трудоемко и неэффективно экономически.

- *Естественный язык.*

Наиболее точно может отразить семантику текста, однако помимо трудоемкости и экономической неэффективности, возникают проблемы синонимии, омонимии и др. неоднозначности естественного языка, затрудняющие алгоритмизацию поиска.

Уменьшить неоднозначности языка помогает *словарь*.

В истории развития информационного поиска разрабатывались и применялись разнообразные словари: словарь синонимичных пар; словарь с многократными связями (например, двуязычный словарь); классификационная таблица; словарь с определениями на естественном языке; отрицательный словарь (содержащий запрещенные словосочетания); словарь-тезаурус («Тезаурус ASTIA»¹ и др.).

Таким образом, существуют *ИПЯ, использующие ключевые слова; дескрипторные ИПЯ без грамматики и с грамматикой, ИПЯ с отрицательным словарем, ИПЯ с тезаурусом* и т. д.

Существуют и иные классификации ИПЯ.

Так, в [27, с. 31–36] предлагается следующее разделение ИПЯ (рис. 5.8):

1. *Предкоординированные ИПЯ.*

- 1.1. Перечислительные классификации: иерархические, алфавитно-предметные (по Мидоу – язык предметных заголовков).

- 1.2. Фасетные классификации.

2. *Посткоординированные (координатные по Мидоу) ИПЯ.*

- 2.1. Дескрипторные языки (с координацией посредством использования операции логического умножения или пересечения \cap).

- 2.2. Семантические коды, задающие парадигматические отношения структурами лексических единиц (код Перри – Кента², RX-коды языка «Бит»³).

¹ *Vikery B.C.* Thesaurus of ASTIA Descriptors, 2nd Ed. // B.C. Vikery // Armed Forces Technical Information Agency. – Arlington, Virginia: Dec. 1962

² *Perry J.W.* Tools for machine literature searching. Semantic code dictionary. Equipment, Procedures / J.W. Perry and A. Kent. – New York, Interscience Publishers, 1958. – 241 p.

³ *Информационно-поисковая система «БИТ»* / Э.Ф. Скороходько, Л.Э. Пшеничная, П.Н. Кар-Ялайне и др. – Киев: Наукова думка, 1968. – 219 с

2.3. Синтагматические языки с развитой системой средств отображения синтагматических отношений (см., например, язык СИНТОЛ¹).

3. Языки библиографических ссылок.



Рис. 5.8. Классификация ИПЯ *А.И. Черноого*

Предлагались классификации ИПЯ по *типу их словарного состава* (см. ссылки в [18]), *типам языковых единиц, степени их сложности, характеру отношений между этими единицами (виду грамматики), системам индексирования* и т. п.

При этом следует иметь в виду, что на практике конкретный ИПЯ нельзя строго отнести к тому или иному классу, поскольку: во-первых, некоторые ИПС могут работать и в режиме без грамматики, и в режиме с грамматикой (например, системы СИНТОЛ¹, СМАРТ [18, 23]); а, во-вторых, ИПС развиваются, и основой является развитие ИПЯ. Поэтому обычно ИПС и ИПЯ описывают рядом характеристик, с тем чтобы пользователь мог выбрать желаемые.

При выборе ИПЯ необходимо оценивать их эффективность. При оценке эффективности ИПЯ используют различные критерии.

¹ **СИНТОЛ** // Сборник переводов по вопросам информационной теории и практики. – М.: ВИНТИ, 1968. – С. 36–47, 50–52, 66–72, 76–80.

В частности, *Ч Мидоу* предлагает четыре меры: *семантическую силу, многозначность, компактность и стоимость* [14, с. 71–79].

Качественные графические иллюстрации исследований, приводимые в [14], показывают, что оценки семантической силы и многозначности ИПЯ коррелируют друг с другом, вступая в противоречие с компактностью и стоимостью, и что наиболее сложной проблемой является оценка семантической силы ИПЯ.

Проблема оценки ИПЯ, в свою очередь, является составной частью более общей проблемы – оценки *качества информационного поиска*.

5.6. Система индексирования

Процедуру перевода с естественного языка на ИПЯ называют *индексированием*. Результатом такого перевода является ПОД (при вводе документов в ИПС) или ПОЗ (при индексировании запроса пользователя).

Проблема индексирования связана с семантическим анализом текстов документов. Сложность ее связана с тем, что индексирование документов, вводимых в поисковые массивы, и запросов пользователя разнесены во времени.

Для алгоритмизации и автоматизации индексирования необходимо решить проблему выбора для включения в ПОД или ПОЗ наиболее значимых ключевых слов, дескрипторов, фраз (в зависимости от лексических единиц ИПЯ).

Важность можно определить несколькими признаками:

- статистически, т.е. на основе частоты использования термина в документе;
- на основе высказываний автора (его мнения, отраженного в заглавии документа или подзаголовках, выделяемых автором в документе);
- с помощью грамматики, позволяющей отразить взаимосвязи между лексическими единицами, содержащимися в контексте;
- по критериям важности, сформулированным пользователем, для чего при индексировании документов могут быть указаны весовые коэффициенты дескрипторов.

Система индексирования конкретной ИПС в основном определяется возможностями ИПЯ, имеющимися в нем лексическими и синтаксическими средствами. Однако есть и некоторые специфические правила и рекомендации, исследование которых позволило выявить некоторые разновидности систем индексирования.

Типы систем индексирования.

Существуют различные типы систем индексирования:

1. К первому типу относят системы *свободного индексирования*.

При этом способе из индексируемого документа выписываются в ПОД слова или словосочетания, которые, отражают содержание индексируемого документа. Кроме этого, элементами ПОД могут быть слова, отсутствующие в этих документах, но отражающие более точно смысл их текстов с точки зрения целей создания ИПС. Выписанные элементы упорядочиваются в алфавитном порядке. Такой упорядоченный набор слов (словосочетаний) представляет собой ПОД при этом типе индексирования. Аналогично – из текста запроса пользователя формируется ПОЗ.

Такой процесс индексирования является принципиально неалгоритмическим, т. е. неавтоматизируемым.

2. При втором методе, который условно называют методом *полу-свободного индексирования*, из документа выписывают слова и словосочетания вначале так же, как и при свободном индексировании.

Однако выписанные элементы сравнивают затем с фиксированным словарем, не найденные в нем – устраняют, а оставшиеся, упорядочиваемые в алфавитном порядке, представляют собой ПОД (или ПОЗ).

3. Третий способ индексирования основан на *статистическом подходе*.

Выбор слов (выражений) исходного текста, подлежащих включению в ПОД, производится на основе статистического анализа текста, при котором его слова рассматриваются как знаки, не имеющие семантических значений. При этом предлагались различные *статистические критерии*, основанные на сопоставлении относительной частоты употребления слова в документе и относительной частоты употребления слова в представительном массиве документов (т. е. в репрезентативной статистической выборке).

Например, в [15] предлагаются следующие количественные критерии:

$$|F - R| > k; \quad \frac{F}{F + R} < k; \quad \frac{F}{R} > k,$$

где F – относительная частота употребления слова в документе; R – относительная частота употребления слова в представительном массиве документов.

Легко видеть, что в основе приведенных соотношений лежит идея, согласно которой информационная значимость слова определяется расхождением частоты его употребления в данном документе и во всем потоке рассматриваемых документов.

Возможны различные подходы к определению расхождения:

- согласно первому вычисляется расхождение между частотой употребления слов в потоке документов данной тематики (монотематический поток) и частотой встречаемости этого слова в многотемном потоке документов (политематический поток);
- второй принцип основан на вычислении расхождения частоты употребления слова в потоке текстов данной тематики и частоты этого же слова в потоке текстов тематики, далекой от данной («противоположной» тематики).

Статистический способ индексирования может быть алгоритмирован и автоматизирован, и в настоящее время имеются средства автоматизированного статистического анализа текстов.

Однако самостоятельного практического применения в ИПС этот способ не нашел, он используется как вспомогательный в сочетании с семантическим анализом текстов документов.

4. К четвертому типу относят *системы индексирования, контролируемые заданным словарем (тезаурусом)*.

Алгоритм индексирования сводится к тому, что каждое слово текста сравнивается с точностью до основы со словарем, совпавшие слова записываются в ПОД.

В некоторых системах словарь используется как помощник специалисту, занимающемуся индексированием текста.

К таким системам относится, например, УДК. В других – такой словарь является элементом алгоритма индексирования: слово, одновременно встретившееся в тексте и в словаре, записывается в ПОД. В дескрипторных ИПЯ в ПОД (ПОЗ) записываются не само слово текста, а соответствующий ему дескриптор.

Перспективным представляется индексирование документов с использованием специально разработанных иерархических классификаций, отражающих цели поиска и использования документов.

Такие классификаторы могут использоваться как бы в качестве ИПЯ в информационных системах нормативно-методического обеспечения управления: иерархический классификатор, объединяющий нормативно-методические документы, разрабатывается на основе структуры целей (основных направлений) и функций деятельности предприятия.

Иерархический классификатор ИПЯ может быть основой системы избирательного распределения информации: разрабатывается классификатор потребностей категории работников, пользующихся системой ИРИ.

5.7. Логика ИПС. Критерии смыслового соответствия

Как показано на рис. 5.4, под логикой ИПС понимаются *критерии выдачи* или *критерии смыслового соответствия* (обязательный элемент), *базисные* (парадигматические) и *текстуальные* (синтагматические) отношения между словами ИПЯ (базисные и/или текстуальные отношения могут и отсутствовать).

Критерий смыслового соответствия (КСС) или *критерий выдачи* позволяет решать вопрос о выдаче или не выдаче того или иного документа, т.е. являются основой алгоритма поиска.

Существуют следующие виды КСС [15, 27]:

* «На полное вхождение» или «на вхождение».

Условием выдачи документов является полное вхождение ПОЗ в ПОД. Иными словами, документ выдается, если множество дескрипторов, образующее ПОЗ ($M_{\text{ПОЗ}}$), полностью входит (рис. 5.9) в множество дескрипторов, содержащихся в ПОД ($M_{\text{ПОД}}$), или совпадает с $M_{\text{ПОД}}$, т.е.

$$M_{\text{ПОЗ}} \subseteq M_{\text{ПОД}}.$$

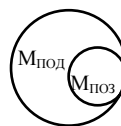


Рис. 5.9

* «На частичное вхождение» ПОЗ в ПОД (пересечение ПОД и ПОЗ).

Документ выдается, если ПОД и ПОЗ совпадают частично, т.е. если часть дескрипторов, содержащихся в $M_{\text{ПОД}}$, совпадает с дескрипторами, входящими в $M_{\text{ПОЗ}}$ (рис. 5.10):

$$M_{\text{ПОЗ}} \cap M_{\text{ПОД}}.$$

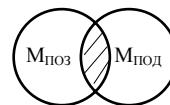


Рис. 5.10

* КСС с учетом базисных отношений.

Документ выдается в том случае, если для каждого дескриптора запроса в ПОД встретился либо сам дескриптор, либо дескриптор, связанный с исходным базисными отношениями.

* КСС с учетом текстуальных и базисных отношений.

Различие с предыдущим заключается в том, что сравнение дескрипторов ПОЗ и ПОД должно осуществляться с точностью до совпадения текстуальных отношений, в которые их прообразы вступают соответственно в запросе и документе.

* КСС с учетом весовых коэффициентов информативных слов или дескрипторов.

Каждому информативному слову в запросе приписывается весовой коэффициент (W_i). Весовые коэффициенты в ПОЗ определяются пользователем и нормируются. Сумма всех весовых коэффициентов в запросе должна быть константой ($\sum W_i = \text{Const}$). Выдача эшелонируется в зависимости от суммы весовых коэффициентов слов запроса, совпавших со словами, употребляемыми

ми в документе. Количество эшелонов выдачи, а также соответствующие каждому из них суммы весовых коэффициентов (порог) определяются разработчиком системы в процессе ее отладки.

** КСС с учетом синтаксических отношений.*

Вводятся правила грамматики и сопоставляются синтагмы, формируемые из дескрипторов (или ключевых слов) с помощью введенных правил.

5.8. Критерии оценки качества поиска и информационно-поисковых систем

В теории информационного поиска предлагаются и используются различные критерии оценки качества информационно-поисковой системы.

Разработка комплекса критериев оценки качества информационного поиска – достаточно сложная проблема: состав и количественные характеристики критериев зависят от конкретного назначения и принципов реализации ИПС.

Оценка или метод оценки – это алгоритмическая процедура, которая любому оцениваемому объекту изданной области ставит в соответствие некоторый другой объект, называемый значением оценки. Полностью алгоритмическую процедуру оценки называют формальной оценкой.

Различают два типа оценок:

* оценки-описания, значения которых характеризуют непосредственно систему безотносительно к другим системам;

* оценки-шкалы, значения которых определяют сравнительные достоинства различных поисковых систем.

От «оценки-описания» требуется, чтобы ее значения позволяли достаточно полно судить о существенных свойствах оцениваемых объектов, например, предсказывать их поведение в тех или иных конкретных условиях. В этом случае «оценка-описание» называется эффективной.

От «оценки-шкалы» требуется, чтобы ее значения упорядочивали множество оцениваемых объектов, например, различных ИПС, не вступая при этом в противоречие с существующими у нас содержательными представлениями о сравнительных достоинствах этих объектов. В этом случае «оценка-шкала» называется здоровой.

Содержательные представления о сравнительных достоинствах систем являются содержательной оценкой. Объективная формальная оценка не должна противоречить содержательной.

Следует иметь в виду, что одна и та же формальная оценка может рассматриваться и как «оценка-шкала», и как «оценка-описание».

Содержательная оценка подразумевает оценку полезности информации для потребителя, для результатов его основной деятельности. При этом оценка эффективности получаемой информации подразумевает оценку ее полезности и затрат на ее получение. Кроме того, для строгой оценки необходимо выделить долю результата, которая получена именно благодаря полученной информации, что крайне затруднено.

С учетом сказанного вместо оценки эффективности поиска, ограничиваются оценкой функциональной эффективности.

Оценки поисковых систем делят на два класса, которые называются *внешними* (или функциональными) и *внутренними* оценками.

Внешние, или *функциональные*, оценки основаны на сравнении результатов работы системы с результатами идеального содержательного поиска, осуществляемого экспертом. В теории *информационного поиска* для этого введены понятия *релевантности* и *пертинентности*.

Под *релевантностью* понимается соответствие выдачи запросу, т.е. релевантность характеризует качество алгоритма поиска. Под *пертинентностью* – соответствие выдачи потребностям лица (или лиц), для которого (которых) осуществляется поиск информации, т. е. пертинентность характеризует смысловыражающие возможности ИПЯ, точность отображения с его помощью информационных потребностей.

В настоящее время иногда термин релевантность используют в более широком смысле (напр., [13]), и различают релевантность первого рода (формальную релевантность), которая соответствует термину, первоначально введенному в теории информационного поиска [15], и релевантность второго рода, соответствующую понятию пертинентности.

Для оценки релевантности используют такие критерии как *полнота*, *точность* поиска, *потери*, *шум*, которые могут быть представлены в виде различных соотношений.

В качестве критериев оценки качества информационного поиска в [13] вводится также понятие коэффициента корреляции поиска:

$$r = \frac{ad - bc}{\sqrt{(a+c)(b+d)(d+c)(a+b)}}, \quad -1 \leq r \leq 1.$$

Кроме того, к числу показателей функциональной эффективности в [13] относят *оперативность* поиска; *специфичность поиска* – отношение числа невыданных нерелевантных документов (d) к общему числу нерелевантных документов ($d + b$), где b – число выданных нерелевантных документов:

$$C = \frac{d}{d+b}; \quad 0 \leq C \leq 1.$$

На практике при оценке ИПС с большими массивами информации точные измерения числа релевантных и нерелевантных документов в общем массиве или в массивах выданных документов затруднено. Поэтому могут использоваться энтропийные показатели

Энтропийные меры могут быть получены на основе исследования выборки из информационного массива, т.е. могут использоваться вероятностные меры неопределенности исходного массива, p_0 , массива выданных p_1 , и массива невыданных p_2 документов, вычисленные на их основе H_0, H_b, H_{nb} и соответствующие меры W .

В частности, в [13] предлагается интегральный энтропийный показатель как мера упорядоченности поискового массива документов, являющаяся результатом процесса поиска по заданному запросу:

$$W = \frac{H_0 - H_n}{H_0},$$

где H_0 – допоисковая (априорная) энтропия, H_n – послепоисковая (апостериорная) энтропия.

При этом для измерения априорной и апостериорной энтропии в [13] предлагается использовать меры концентрации релевантных документов в общем массиве, в массиве выданных документов и в массиве невыданных документов, в относительных единицах:

$$P_0 = \frac{a+z}{a+g+z+d}; \quad P_1 = \frac{a}{a+g}; \quad P_2 = \frac{z}{z+d},$$

где a – число релевантных документов, выданных в результате поиска; z – число релевантных, не выданных в результате поиска; g – число нерелевантных документов, выданных в результате поиска; d – число нерелевантных документов, невыданных в результате поиска.

Энтропийная мера может быть представлена и в логарифмической форме.

В [13], например, для оценки энтропии используются натуральная логарифмическая мера (т.е. неопределенность измеряется в неперах):

неопределенность исходного массива

$$H_0 = [P_0 \ln P_0 + (1 - P_0) \ln(1 - P_0)];$$

неопределенность массива выданных документов:

$$H_b = [P_1 \ln P_1 + (1 - P_1) \ln(1 - P_1)];$$

неопределенность массива невыданных документов

$$H_{nb} = [P_2 \ln P_2 + (1 - P_2) \ln(1 - P_2)].$$

При формировании подобных мер могут быть использованы и иные логарифмические шкалы: двоичные логарифмы (биты), восьмеричные логарифмы (байты), десятичные логарифмы.

Внутренние оценки основываются на таких структурных качествах системы как *сложность, степень близости к человеческой логике или естественному языку, степень алгоритмичности*, на оценке компонентов ИПС, и в частности информационно-поискового языка (ИПЯ) и т.п.

Например, **Ч. Мидоу** [14] предлагает оценивать качество информационно-поискового языка, используя следующие критерии: *семантическая сила (expressiveness), многозначность и компактность* языка, *стоимость выбора термина*.

Семантическая сила – это способность языка идентифицировать объект, различать мелкие особенности объектов, описывать объект с разной степенью детализации.

Подразумеваются потенциальные возможности ИПЯ, а не умение им пользоваться. Самой большой семантической силой обладает естественный язык.

Многозначность означает, что слово или синтаксическая единица тезауруса имеет более, чем одно значение (омографы), или, напротив, что некоторое значение может иметь более одного символического представления в словарном составе ИПЯ (синонимия). Кроме того, одинаково звучащие слова могут иметь различные значения (полисемия или омонимия).

Синонимия и омография могут существовать и в синтаксических единицах, состоящих из нескольких слов.

Компактность характеризуют физический размер или длину терминов словаря или поисковых образов, составленных их числа терминов, необходимых для отображения смысла документов и запросов.

Стоимость характеризует цену процесса принятия решения по выбору терминов (ключевых слов, дескрипторов или иных синтаксических единиц) для отображения смысла документа или запроса.

В общую стоимость входят: *стоимость обучения пользованию языка, стоимость составления и совершенствования словаря, затраты, связанные с устранением ошибок, допущенных при выборе терминов, затраты времени на индексирование документов и составление ПОЗ*.

Предлагаемые **Ч. Мидоу** оценки не являются независимыми и взаимоисключающими.

ИПЯ может быть семантически сильным, но многозначным. Компактность слов в словарном составе языка не определяет стоимости, т.е. затрат времени, труда на выбор терминов.

ИПЯ характеризуют также *словарным составом* и наличием *грамматики*. При наличии тезауруса ИПЯ можно охарактеризовать глубиной *тезауруса*, т.е. количеством уровней, видов смысловыражающих элементов или синтаксических единиц тезауруса. Характеристики ИПЯ являются *внутренними* оценками информационно-поисковой системы, влияющими на оценку качества информационного поиска, по критерию pertinентности.

Количественные оценки релевантности

Формальные оценки релевантности можно получить только для релевантности первого рода, т.е. для релевантности в исходном ее понимании в теории информационного поиска.

Для введения критерия релевантности следует задать процедуру определения меры семантической близости поискового образа документа поисковому образу запроса и некоторое пороговое значение этой меры. Если мера превышает пороговое значение, то документ релевантен запросу.

ПОД и ПОЗ представляют собой множества ключевых слов или дескрипторов в зависимости от вида информационно-поискового языка. Для их сопоставления используют критерии смыслового соответствия, которые определяют на основе совпадения ключевых слов (дескрипторов) в ПОД и ПОЗ.

ПОД и ПОЗ можно представить в виде четких и нечетких множеств. Со способами введения мер релевантности на основе нечетких множеств можно познакомиться в [13]. При введении мер релевантности обычно используют представление ПОД и ПОЗ в виде множеств.

Для четких множеств вводят нормированную меру релевантности

$$0 \leq \mu_R(a, b) \leq 1.$$

Критерий смыслового соответствия можно представить в виде:

$$r_a = \langle \mu_R(a, b) \geq \alpha \rangle,$$

где $\mu_R(a, b)$ – функция вычисления меры релевантности (или просто мера релевантности); α – пороговое значение релевантности, такое, что

$$r_a = \begin{cases} 1 & \text{при } \mu(a, b) \geq \alpha \text{ документ выдается} \\ 0 & \text{при } \mu_r(a, b) < \alpha \text{ документ не выдается} \end{cases}$$

Изменяя пороговое значение α , можно организовать выдачу различных совокупностей документов, которую в теории информационного поиска называют *эшелонированной выдачей*. Каждый эшелон соответствует определенной мере семантической близости совокупности документов запросу.

Очевидно, что чем больше пороговое значение α , тем более жесткие условия налагаются на смысловую близость документа запросу. В нормированных мерах при $\alpha = 1$ для выдачи документа требуется полное совпадение ПОД и ПОЗ.

Оценку релевантности можно характеризовать *полнотой* выдачи (или *потерями*), т.е. числом невыданных релевантных документов), и *точностью*, (или *шумом*), т.е. числом или процентом «лишних» документов, которые выданы в результате поиска, но не являются релевантными.

Например, в [27] оценки полноты R и точности и T вводятся следующим образом:

$$R_i = \frac{a_i}{n_i}, \quad T_i = \frac{a_i}{m_i}, \quad 0 \leq R_i \leq 1, \quad 0 \leq T_i \leq 1,$$

где a_i – число релевантных документов, формально выданных системой на i -й запрос; m_i – число всех формально выданных на i -запрос системой документов; n_i – число всех релевантных документов, соответствующих запросу.

При этом рекомендуется полноту и точность определять на основе нескольких поисков N по запросу, и определять их средние значения и суммарные относительные оценки:

$$R = \frac{\sum_{i=1}^N \frac{a_i}{n_i}}{N}, \quad T = \frac{\sum_{i=1}^N \frac{a_i}{m_i}}{N} \quad \text{– средние относительные оценки;}$$

$$R = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N n_i}, \quad T = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N m_i} \quad \text{– суммарные относительные оценки,}$$

где N – число поисков.

Величины $(1 - T)$ и $(1 - R)$ называются соответственно шумом и потерями.

В [15, с. 306] предлагается наглядная матрица для определения полноты и потерь, точности и шума (табл. 5.4).

Таблица 5.4

		Релевантны	Нерелевантны		
		A_1	A_2		
Выдано	B_1	a	b	$A+b$	
Не выдано	B_2	c	d	$C+d$	
		$a+c$	$b+d$	$a+b+c+d$	

Полноту поиска измеряют отношением числа выданных релевантных документов (a) к общему числу релевантных документов массива ($a+c$):

$$R = \frac{a}{a+c}, \quad 0 \leq R \leq 1.$$

Точность поиска T – отношение числа выданных релевантных документов (a) к числу общему выданных документов ($a+b$)

$$T = \frac{a}{a+b}, \quad 0 \leq T \leq 1.$$

Соответственно потери L и шум S можно представить следующим образом:

$$L = \frac{c}{a+c} \quad \text{и} \quad S = \frac{b}{a+b}, \quad 0 \leq L \leq 1, \quad 0 \leq S \leq 1.$$

В [27] предлагается, проведя серию экспериментов n по определению полноты и точности поиска, определить среднюю полноту и среднюю точность:

$$P_{cp} = \frac{1}{n} \sum_{i=1}^m \frac{a_i}{a_i + b_i}; \quad T_{cp} = \frac{1}{n} \sum_{i=1}^m \frac{a_i}{a_i + c_i}.$$

Используются и иные способы усреднения (см., напр., [15, 27]).

Например, в связи с оценкой системы СМАРТ Сэлтон¹ ввел нормированную полноту R_N и нормированную точность P_N

$$R_N = \frac{1}{N} \sum_{i=1}^N \frac{n_i}{n},$$

где N – число документов в массиве; n – число всех релевантных документов в массиве; n_i – число релевантных документов, выданных до i -го ранга включительно,

$$P_N = \frac{1}{N} \sum_{i=1}^N \frac{n_i}{i},$$

где i – номер ранга.

В [15] предлагается также, пользуясь табл. 5.4, ввести (с определенной степенью приближения) показатели в терминах теории вероятностей:

$$\frac{a}{a+c} = P_{B1}(A_1) \quad \text{– условная вероятность выдачи релевантных документов;}$$

$$\frac{b}{b+d} = P_{B2}(A_1) \quad \text{– условная вероятность выдачи нерелевантных документов;}$$

¹ Сэлтон Г. Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. – М.: Сов. радио, 1973. – 560 с.

$\frac{c}{a+c} = P_{B1}(A_2)$ – условная вероятность невыдачи релевантных документов;

$\frac{d}{b+d} = P_{B2}(A_2)$ – условная вероятность невыдачи нерелевантных документов.

$$P_{B1}(A) + P_{B1}(A_2) = 1$$

$$P_{B2}(A) + P_{B2}(A_2) = 1$$

Предлагались и другие способы оценки релевантности (см. обзор в [15])

Определение полноты системы связано с определением содержательной выдачи на каждый запрос.

Существует несколько способов определения этой выдачи:

* Сплошной просмотр всего экспериментального массива.

Достоинством этого способа является надежность, недостатком – трудоемкость.

* Метод документа-источника («метод Клевердона»).

Состоит в том, что по некоторым документам массива, выбранным более или менее случайно, составляются запросы с таким расчетом, чтобы каждый документ-источник был релевантен составленному по нему запросу.

* Метод контрольных документов.

По запросу, полученному по произвольно выбранному документу-источнику, проводится содержательный поиск путем сплошного просмотра массива, начиная, например, с документа-источника, до нахождения первого релевантного документа, который объявляется контрольным. Значение полноты для системы считается теперь как доля запросов, по которым система выдала контрольный документ в общем количестве запросов.

* Метод объединения формальных выдач.

Применяется при сравнении нескольких поисковых систем («оценка-шкала»). Он состоит в том, что по каждому запросу эксперт просматривает только те документы, которые выдавались хотя бы одной из этих поисковых систем. Содержательной выдачей считается совокупность обнаруженных релевантных документов, и относительно нее определяется полнота, которая отличается от истинной полноты каждой из рассматриваемых систем.

Вопрос о представительности массива документов и массива запросов, выбранных для определения формальных оценок, в общем виде не решен. Считается, что более или менее устойчивые оценки (колебания не превышают 5%) можно получить на массиве в 4000 документов, массив запросов при этом должен быть порядка нескольких сотен.

Таким образом, релевантность (формальная) характеризует свойства средств логико-семантического аппарата информационно-поисковой системы и зависит от возможности отображения ПОД и ПОЗ с помощью информационно-поискового языка, принятых в ИПС алгоритмов поиска и системы индексирования.

5.9. Разработка и отладка ИПС

При разработке ИПС обычно выполняются следующие этапы:

1) анализ совокупности документов, представляющих соответствующую научно-техническую область, для которой создается ИПС; выбор ключевых слов (КС), характеризующих содержание каждого документа; КС выбирают из заголовков (или из аннотаций) документов, включаемых в информационный массив (информационную базу) ИПС. Ключевые слова – существительных или отглагольных существительных, взятые в единственном числе и именительном падеже, прилагательные, глаголы в неопределенной форме);

2) формирование из ключевых слов словаря (первой стадии разработки информационно-поискового языка); ИПЯ является основой для формирования ПОД, а в последующем ПОЗ.

3) выбор системы индексирования документов; обычно вначале выбирается индексирование по алфавиту;

4) формирование ПОД каждого документа с использованием словаря; первоначально при выборе системы индексирования по алфавиту слова в ПОД располагаются по алфавиту; к дальнейшему в процессе отладки системы может быть выбран иной способ индексирования.

Примечание. Возможно свободное индексирование, без словаря (система Унитерм); тогда будет отсутствовать п. 2.

При формировании информационного массива в структуру записи (сложного термина – при использовании языка логического программирования¹) включают номер документа, наименование документа, фактографические сведения (дата его издания, место издания, автор и т.п., в зависимости от характера информации), фасета из КС ПОЗ. ПОЗ может описываться отдельно, тогда в него помимо КС включается номер документа.

5) формируется ПОЗ или поисковое предписание; при применении ИПЯ – с его использованием;

¹ *Анастасиади П.Г.* Турбо-Пролог и его применение для создания информационно-поисковых систем: учеб. пособие / П.Г. Анастасиади, В.Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2007. – 56 с

б) проводится отладка ИПС; в процессе которой оценивается релевантность выдачи и исследуется несколько способов повышения релевантности:

изменение алгоритма поиска и/или КСС;

изменение способа индексирования (статистический, с учетом значимости КС для отображения смысла документа или запроса, т.е. с учетом весовых коэффициентов КС в ПОД и ПОЗ);

развитие ИПЯ: внесение изменений в словарь путем введения классов условной эквивалентности, дескрипторов или иной способ изменения словаря; введение грамматики.

Изложенная методика разработки и отладки ИПС представляется в виде структурной схемы типа приведенной на рис. 5.11.

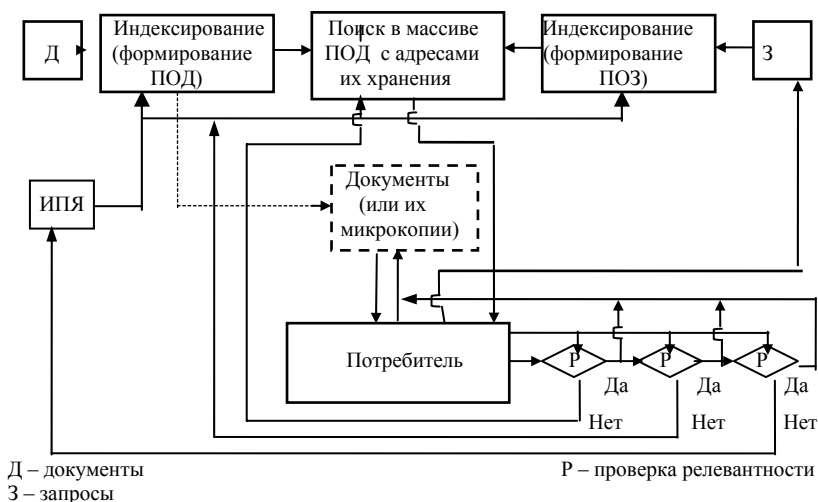


Рис. 5.11. Схема отладки ИПС

С применением рассмотренной методики студенты выполняют лабораторную работу с использованием языка логического программирования Турбо-Пролог¹.

¹ *Анастасиади П.Г.* Турбо-Пролог и его применение для создания информационно-поисковых систем: учеб. пособие / П.Г. Анастасиади, В.Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2007. – 56 с.; Доорс Дж. Пролог – язык программирования будущего / Дж. Доорс, А.Р. Рейблейн, С. Вадера / Пер. с англ. А.Н. Волкова. – М.: Финансы и статистика, 1990. – 144 с. (Prolog – Progrzmming for Tomorrow by J. Doors, A.R. Reblein & S. Vadera. – Signa Press *Winstow).

Глава 6. ПРИНЦИПЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ ДОКУМЕНТАЛЬНО-ФАКТОГРАФИЧЕСКИХ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ

6.1. Понятие о документально-фактографических информационно-поисковых системах

Документально-фактографическая информационно-поисковая система (ДФИПС) представляет собой ИПС, предназначенную для выдачи непосредственно требуемой информации (например, температуры кипения конкретной жидкости, статистических показателей и иных конкретных сведений, содержащихся в соответствующих отчетных документах, публикациях и т.п. документальных информационных массивах, и др.).

При этом существуют ДФИПС двух видов:

- * *документальные* и сопряженные с ними массивы *фактографической* информации, которые формируются параллельно;

- * *информационно-логические системы*, в которых фактографическая информация извлекается из документальной в процессе поиска.

Первоначально разрабатывались документально-фактографические системы, в которых массивы фактографической информации формировались параллельно с массивами документальной информации.

Наиболее известными примерами таких систем являются ДФИПС «Фтор» и «Спектр» [15, 16].

ДФИПС «Фтор». Была разработана в ВИНТИ в 1963 г. и предназначена для оперативного оповещения химиков о новых публикациях по фтороорганическим соединениям (документальная ИПС), а также для быстрого поиска сведений о химических соединениях и реакциях (фактографический режим работы ИПС).

Для реализации фактографического режима формируются фактографические массивы, в которых записываются:

- * реакция или соединение, краткие сведения о них, содержащиеся в публикации;

- * дополнительные сведения (механизм и кинетика реакции, спектральные данные, физические и физико-химические свойства и т.д.);

- * выходные данные публикации-источника.

Для описания документов разработан специальный ИПЯ, предназначенный для исчерпывающего описания реакций и свойств фтороорганических соединений.

В АДФИПС «Фтор» предусмотрена возможность работы в следующих режимах обслуживания: ИРИ, ретроспективный поиск; издание пермутационных библиографических указателей по тематике ИПС; изготовление копий публикаций-источников по заказам подписчиков.

АДФИПС «Спектр». Разработана в 1970 г. группой сотрудников Новосибирского отделения органической химии и Вычислительного центра Сибирского отделения Академии наук СССР. Первоначально эта АДФИПС предназначалась для идентификации органических соединений по их инфракрасным спектрам.

Для этого разработан специальный вариант кодирования инфракрасного спектра, поскольку АДФИПС предназначалась не только для опознания веществ по их инфракрасным спектрам, но и для выявления различных корреляционных связей. В АДФИПС предусмотрен режим стандартных запросов. При этом возможен поиск одновременно по 4-м запросам, а в последующем была создана подсистема АДФИПС для электронной спектроскопии, в которой возможен поиск одновременно по 8 запросам.

Документально-фактографические ИПС особого типа составляют основу *Автоматизированных систем нормативно-методического обеспечения управления* (АСНМОУ) предприятием (организацией) и Информационных систем нормативно-правовой документов.

6.2. Автоматизированные системы нормативно-методического обеспечения управления предприятиями и организациями

Автоматизированная система нормативно-методического обеспечения управления (АСНМОУ) предприятием (организацией) должна содержать нормативно-правовые, нормативно-методические, нормативно-технические и организационно-распорядительные документы (НПД, НМД, НТД и ОРД), которые обеспечивают реализацию принятых проектных и управленческих решений в процессе функционирования предприятия (организации) [28].

Пример структуры СНМОУ и АСНМОУ, соответствующей основным функциям, которые должна выполнять СНМОУ, приведен на рис. 6.1.

В качестве методической основы создания АСНМОУ используют идею стратифицированного представления процедуры поиска информации с углублением на каждой страте анализа документов, содержащихся в АСНМОУ, путем структуризации их текстов (рис. 6.2):

на верхней страте – поиск документов по функциям управления;

на второй сверху – поиск разделов документов в соответствии с запросом пользователя, взаимоотношений между разделами связанных друг с другом документов;

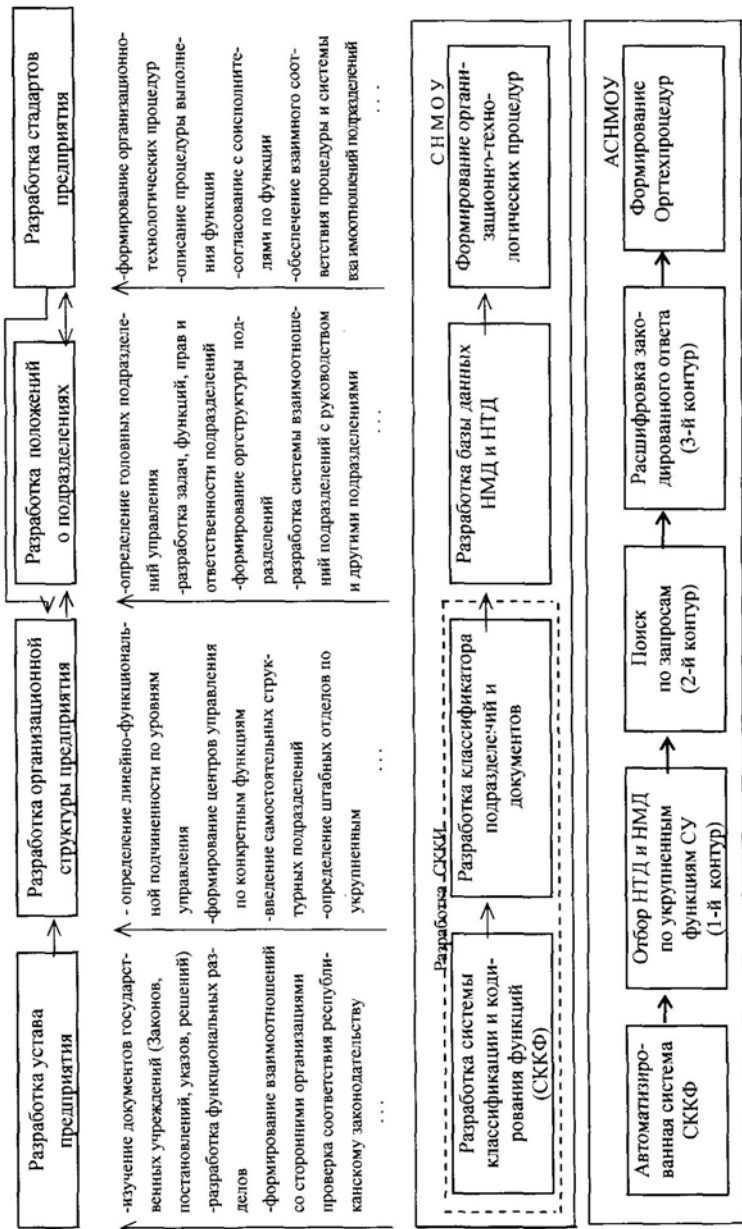


Рис. 6.1. Функции и структура СНМОУ и АСНМОУ

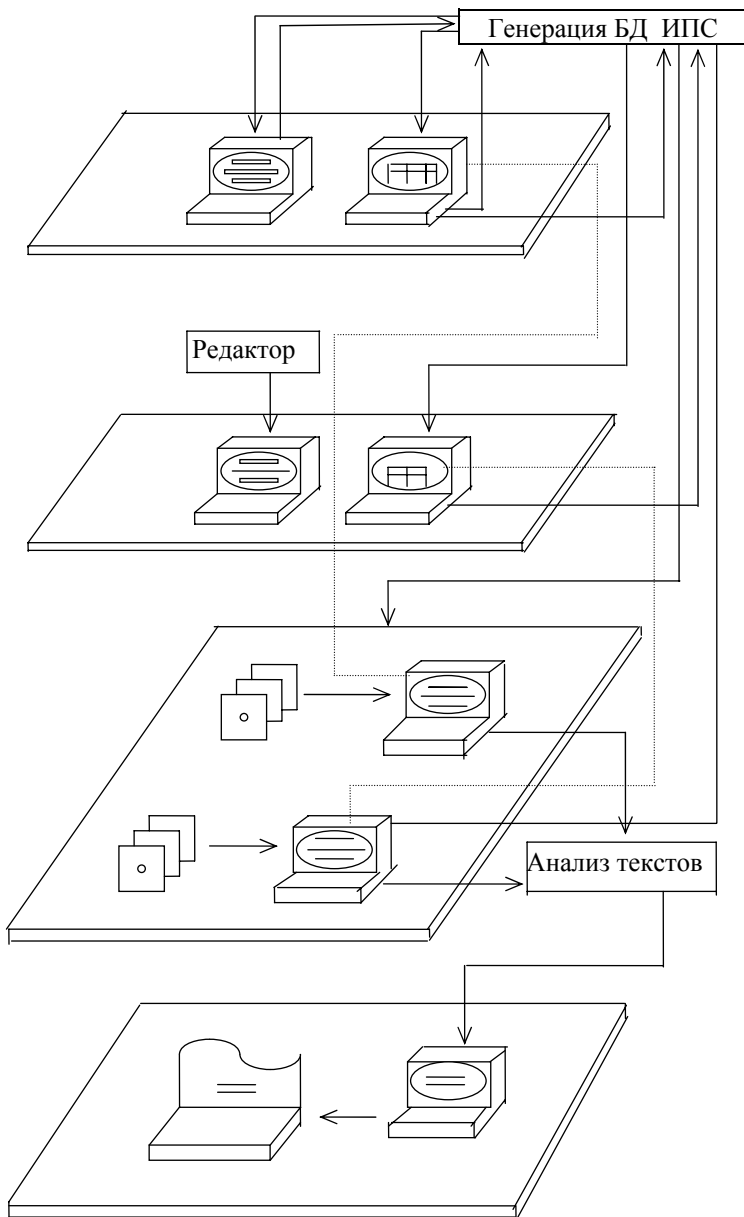


Рис. 6.2. Принципы стратифицированного построения АСНМОУ

на третьей страте – вывод текстов на дисплей или принтер (полного текста или его разделов);

на четвертой страте (которая реализуется не для всех документов) – аналитико-синтетическая обработка (АСО) текстов документов (что, например, требуется при поиске информации в текстах законов и других НПА).

Виды баз данных определяются спецификой конкретного предприятия (организации).

Например, базы данных можно выделить таким образом: создать базу данных НПА (общегосударственных, региональных), базу данных об НМД, НТД, и ОРД органов отраслевого управления и предприятия. Для АСНМОУ предприятия может оказаться целесообразным создать отдельные базы СП, должностных инструкций, положений о подразделениях и т.п. Базы данных для крупных предприятий и организаций имеют значительный объем. Для организации эффективных процедур поиска и корректировки информации, их нужно структурировать. Исследование особенностей БД АСНМОУ и работы с ними показало, что выбрать наиболее целесообразную жесткую структуру БД практически невозможно, так как, с одной стороны, подразделения предприятия принимают участие в выполнении нескольких укрупненных функций управления, а, с другой стороны, одна и та же функция выполняется несколькими подразделениями.

Кроме того, одна и та же функция регламентируется в документах разного вида – и в положениях о подразделениях, и в СП, и в ОРД и т.п. При этом одинаковые или сходные с различной степенью детализации и с несколько модифицированными формулировками функции в разных документах закодированы по-разному, в соответствии с группированием и индексированием функций, принимаемым при разработке документа его авторами.

Используя терминологию теории научно-технической информации, удобно говорить не о создании баз данных для различных страт, а о создании двух-, трех- или более контурных АДФИПС.

В частности, при создании БД нормативно-правовых документов исследовалось два варианта АДФИПС: двухконтурный – с поиском документов в *первом* контуре и извлечения из них разделов (которые содержат сведения о правовых нормах, представляющих собой фактографическую информацию для пользователя) во *втором*; и трехконтурный с поиском в контурах последовательно в первом – документов, во втором – разделов, и в третьем – статей (содержащих еще более конкретную фактографическую информацию о правовых нормах).

Аналогично могут быть организованы БД нормативно-методических документов.

Например, при создании АДФИПС положений о подразделениях предприятия или БД НМД, НТД и ОРД могут быть образованы следующие контуры: 1) отбор положений или других НМД, НТД, ОРД, соответствующих запросу (т. е. структуризация под запрос – например, по укрупненным функциям, по группам подразделений); 2) поиск в базе отобранных документов по запросам пользователя необходимых ему сведений о функциях, сроках и способах их выполнения и т. п. фактографической информации (поиск может осуществляться по признакам, предусмотренным при разработке АДФИПС); 3) вывод фрагментов документов, отобранных в соответствии с запросами во втором контуре, на дисплей или принтер в удобной для пользователя форме (например, разделов положений, содержащих необходимые сведения о функциях, сроках их выполнения исполнителях, в виде таблицы, и т. п.).

Пример трехконтурной АДФИПС нормативно-методических документов АСНМОУ (положений о подразделениях оргструктуры предприятия), приведен на рис. 6.3 [28].

Структуры БД в этом примере имеют следующий вид (в приводимом примере принята фасетная классификация).

Структура первого контура БД положений о подразделениях:

$$nmo = \langle NI, N2, F1, F2, F3 \rangle,$$

где NI – номер отдела в существующей СНМОУ; $N2$ – наименование отдела; $F1, F2, F3$ – индексы укрупненной функции в структуре целей и функций предприятия, выполнение которой регламентируются в положении.

Структура БД раздела «Взаимосвязь» положений о подразделениях предприятия (второго контура АДФИПС), составляющими которой являются функции положения:

$$nmo = \langle F1, F2, F3, NI, DI, FJ, DK, NP \rangle,$$

где $F1, F2, F3$ – индексы укрупненной функции в структуре целей и функций предприятия, выполнение которой регламентирует подготавливаемый документ; NI – номер отдела-изготовителя документа в существующей СНМОУ; DI – наименование вида исходного документа (или конкретного документа), для подготовки которого выполняется функция; FJ – индекс функции внутри положения; DK – индекс документа, получаемого в результате выполнения функции; NP – номер отдела, в который передается документ для дальнейшей обработки (согласования). Поиск фактографической информации в этом контуре может осуществляться по любому параметру.

Третий контур – вывод соответствующих отобранным кодам текстов.

В приведенном примере АДФИПС предусмотрен также блок формирования и анализа организационно-технологических процедур (ОТП) подготовки и реализации управленческих решений (подробнее характеризуем, например, в [1, 21, 28]).

Соответственно – для БД нормативно-правовых документов.

Структура БД НПД первого контура:

$$npd = \langle N, T, D, M, F1, F2, F3, IF \rangle,$$

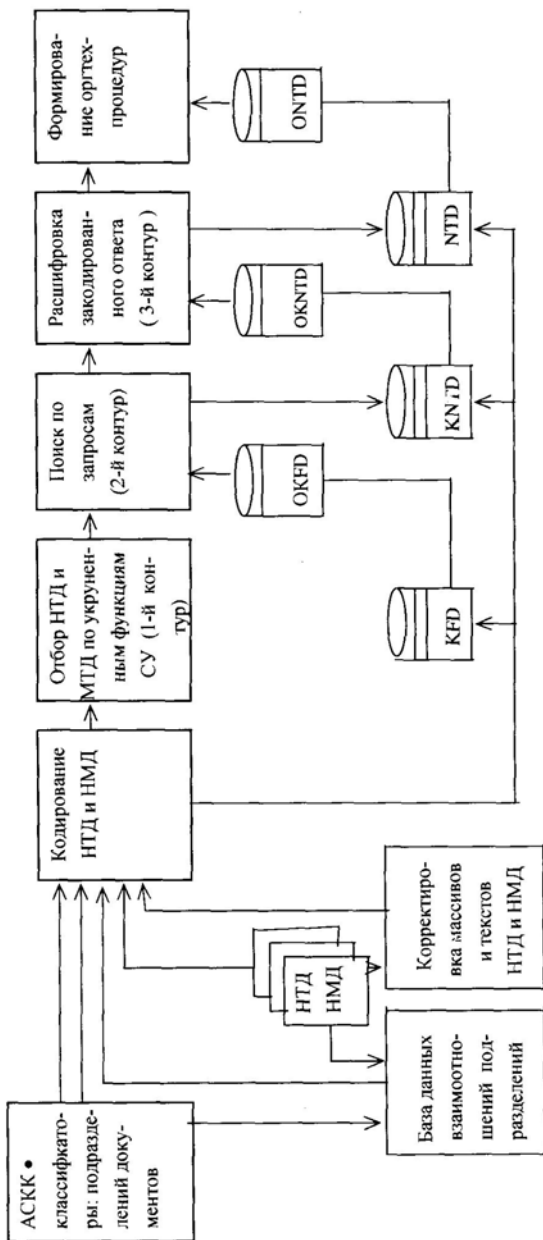


Рис. 6.3. Структура трехконтурной АДПИС

где N – номер документа в БД НПД; T – тип документа (закон, постановление и т. п.); D – наименование документа; M – место опубликования или хранения документа; $F1, F2, F3$ – индекс укрупненной функции, регламентируемой документом; IF – имя файла, в котором хранится текст документа (или фрагменты текста).

Структура БД НПД второго контура:

$$npd = \langle N, D, R, F1, F2, F3, F4, F5, F6, IFR \rangle,$$

где N – номер документа в БД НПД; D – наименование документа (закона, постановления, и т. п.); R – наименование раздела документа; $F1, F2, F3, F4, F5, F6$ – индекс функции, регламентируемой разделом; IFR – имя файла, в котором хранится раздел текста документа.

При необходимости может быть сформирован контур для статей НПД или контур аналитико-синтетической обработки информации.

Для ускорения разработки АСНМОУ для конкретных предприятий и организаций разрабатываются системы ее генерации АДФИПС, с примерами которых можно познакомиться в [1, 18].

Приведенные примеры ДФИПС относятся к классу ИПС, в которых фактографический поиск понимается как процесс отыскания уже готовых данных и фактов, извлекаемые из текстов в процессе подготовки и индексирования входных данных и вводятся в массивы ИПС. В этом случае теория и методика разработки таких ИПС практически не отличается от чисто документальных ИПС, и нередко ДИПС развились в ДФИПС.

В последующем стали понимать документально-фактографический поиск как процесс автоматизации извлечения и логического преобразования фактографических данных с целью получения новой информации. Такие ИС являются одним из направлений развития информационно-логических систем. Принципы построения и разработки таких ИЛС существенно отличаются от рассмотренных. Необходимо проводить анализ системы понятий и методов рассуждений, принятых в соответствующей отрасли знаний, разрабатывать более сложные ИПЯ с правилами логического вывода. ИС такого рода базируются на применении методов дискретной математики и искусственного интеллекта.

Глава 7. ЗАКОНОМЕРНОСТИ ИНФОРМЕТРИИ И СТРОЕНИЯ ДОКУМЕНТАЛЬНЫХ ИНФОРМАЦИОННЫХ ПОТОКОВ

7.1. Понятие об информетрии

Термин *информетрия*¹ был введен в начале 80-е гг. XX в. по аналогии с наукометрией, библиометрией для краткой характеристики количественных методов исследования научно-технической информации. Наиболее полно этот термин был раскрыт **В.И. Горьковой** [5].

Закономерности информетрии определяют распределения информации в документальных информационных потоках (ДИП). количественные и качественные параметры организации частотных словарей, использования слов в текстах документов. ДИП формируют официальные, периодические и продолжающиеся издания и др. опубликованные и неопубликованные документы *научно-технической информации*

Первые результаты исследований лингвистических закономерностей естественного языка были получены **Дж. Эсту (J.B. Estoup, 1916 г.)**, **А. Лоткой (A.J. Lotka, 1926 г.)**.

Характеристику качественных свойств частотных словарей определил в 1916 г. Дж. Эсту, который обнаружил, что частота использования слов в тексте обратно пропорциональна его номеру в частотном словаре.

Наиболее исследованы следующие закономерности информетрии.

7.2. Закон Ципфа

George **G.K. Zipf (Дж. Ципф)**, или в некоторых современных переводах – **Г. Зипф**) в начале 30-х гг. XX в. на основе статистических исследований получил следующую закономерность³.

Пусть есть текст длиной N слов и словарь объемом m слов с указанием частоты появления слова в тексте. Слова в словаре расположе-

¹ **Bonitz M.** Scientometrie, Bibliometrie, Informetrie / M. Bonitz // Zbl. Bibliotheksw. – 1982. – Vol. 86. – № 1. – S. 19–24.

³ **Zipf G.K.** Human behavior and principle of least effort / G.K. Zipf // Cambridge (Mass.): Addison – Wesley, 1949. – Vol. XI. – 573 p. **Zipf G.K.** Selected studies of the principle of relative frequency in language. / G.K. Zipf // Cambridge (Mass.): Harvard Univ. Press., 1932. – 51 p.

ны в порядке убывания их по частоте и пронумерованы от 1 до m . Ранг, равный 1 присваивается слову, частота появления которого наибольшая; ранг, равный m , – наименее употребляемому слову. Тогда

$$p_{r_i} = \frac{f_{r_i}}{N},$$

где p_{r_i} – относительная частота появления слова в тексте;

f_{r_i} – абсолютная частота появления слова r_i ранга в тексте определенной длины;

N – число слов в тексте; r_i – ранг слова, где $1 \leq i \leq m$.

Если умножить вероятность или относительную частоту обнаружения слова в тексте на ранг r_i слова, то получим:

$$p_{r_i} r_i = f_{r_i} \frac{r_i}{N} = k,$$

где k – константа; $1 \leq r_i \leq m$.

Если преобразовать формулу, то получим $p_{r_i} = \frac{k}{r_i}$, т.е. функцию типа $y = k/x$, график которой – равнобочная гиперболола.

Таким образом, на основе анализа полученных зависимостей Ципф предложил эмпирическую формулу, устанавливающую связь между частотой появления слов в тексте и его рангом в словаре:

$$p_{r_i} = k r_i^{-1},$$

где k – эмпирически определяемая константа, изменяющаяся для разных текстов.

При этом $1 \leq r_i \leq m$;

p_1 – частота наиболее употребляемого слова;

p_m – частота наименее употребляемого слова;

$p_{r_i} = \varphi(r_i)$ – «гиперболическая лестница», поскольку ранговое распределение имеет ступенчатый характер (ряд слов появляется с одинаковой частотой), но при аппроксимации можно считать распределение Ципфа гиперболой (рис. 7.1).

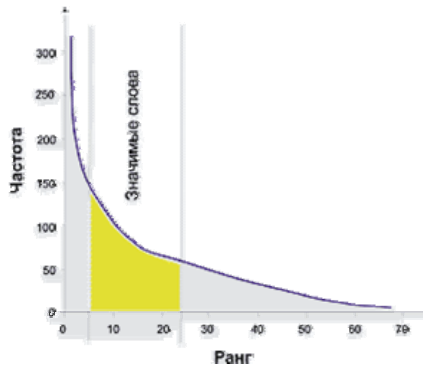


Рис. 7.1

Значение константы в разных языках различно, но внутри одной языковой группы остается неизменно, какой бы текст мы ни взяли. Так, имеются исследования, показывающие, что, например, для английских текстов константа Ципфа равна приблизительно 0,1; а для русского языка – примерно 0,06–0,7.

Поэтому Ципф приводил также запись этой закономерности в форме

$$p_{r_i} = 0,1r_i^{-1}, \quad 1 < r_i < m, \quad ,$$

где $k = 0,1$ (для естественных языков).

Опираясь на экспериментальные данные, собранные в результате статистического исследования многих текстов на различных языках, Ципф обнаружил также, что распределение слов естественного языка подчиняется единому простому закону, который он назвал «принципом наименьшего усилия»: выражая мысли с помощью языка, мы подвергаемся действию двух противоположных сил – силе унификации и силе диверсификации, проявляющихся, с одной стороны, в необходимости быть понятыми, а с другой – желанием выразить мысль покороче.

Ципф установил, что частота и количество слов, входящих в текст с этой частотой, связаны между собой. Если построить зависимость количества слов в данной частоте от частоты вхождения слова, то получится кривая аналогичная рис. 7.1, которая будет сохранять свои параметры для всех без исключения созданных человеком текстов с некоторыми отклонениями для разных естественных языков (рис. 7.2).

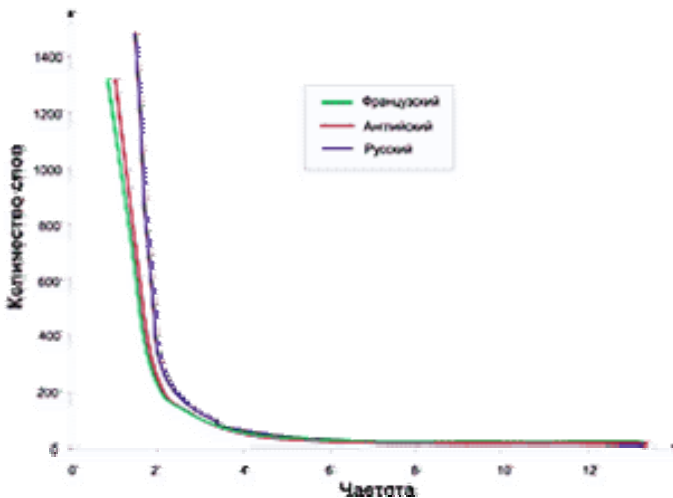


Рис. 7.2

Эту закономерность считают иногда *вторым законом Ципфа*.

Исследования показали, что наиболее значимые слова лежат в средней части гиперболы (см. рис. 7.1). Слова, которые попадают слишком часто, в основном оказываются предлогами, местоимениями, в английском – артиклями и т.п. Редко встречающиеся слова тоже в большинстве случаев не имеют решающего смыслового значения.

От того, как будет выставлен диапазон значимых слов, зависят свойства информационно-поисковой системы.

Если поставить широко – нужные термины потонут в море вспомогательных слов; если установить узкий диапазон – можно потерять смысловые термины. В каждой поисковой системе эта проблема решается по-своему, с учетом общего объема текста, специальных словарей и т.п.

Таким образом, закономерности Ципфа отражает некоторое общее свойство, присущее разным языкам. Это свойство заключается в том, что в каждом тексте на любом естественном языке имеется некоторое количество наиболее употребимых слов. Причем число этих слов значительно меньше общего числа слов, используемых в тексте.

Законы Ципфа универсальны. В принципе, они применимы не только к текстам.

В аналогичную форму выливается, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности узлов в сети Интернет – тоже отвечают законам Ципфа.

Закономерность Ципфа проявляется и при исследовании документальных информационных потоков (ДИП). В этом случае закон Ципфа представляют через абсолютную частоту появления слов:

$$f_i = Cr_i^{-1},$$

где f_i – абсолютная частота появления слова в текстах документального потока; r_i – ранг слова в ранговом распределении; C – частота появления слова 1-го ранга, которая для данного ДИП можно считать эмпирической константой.

7.3. Закон Мандельброта

Benoit Mandelbrot (Б. Мандельброт) в 1954 г.¹ предложил теоретическое обоснование эмпирически открытого закона Ципфа. Представляя слова в виде последовательности букв, разделенных пустыми промежутками и присваивая знакам определенную «стоимость» (затраты усилий,

¹ *Мандельброт Б.* Теория информации и психологическая теория частот слов / Б. Мандельброт // Математические методы в социальных науках. – М.: Прогресс, 1973. – С. 326–337.

времени), Мандельброт показывает, что словам можно приписать априорные вероятности, так чтобы их общая «стоимость» в среднем была минимальной и при этом количество информации оставалось инвариантным. На основе этих представлений математическим путем Мандельброт показал, что результирующее соотношение между частотой слова и его рангом соответствует эмпирическому закону Ципфа с небольшой поправкой:

$$p_{r_i} = k r_i^{-\gamma},$$

где p_{r_i} – относительная частота появления слова в тексте;

r_i – ранг слова; k – эмпирическая постоянная; γ – величина, близкая к единице, но изменяющаяся в зависимости от свойств текста.

Коэффициент γ характеризует определенные свойства языка – степень его формализованности, при этом с уменьшением γ степень формализованности языка уменьшается.

7.4. Закон Брэдфорда

Открыт английским химиком и библиографом *С. Брэдфордом* в 1948 г.¹ на основе обнаружения общих принципов распределения публикаций по изданиям в разных областях (на примере изданий по геофизике и химии).

Основной смысл закономерности состоит в следующем: если научные журналы расположить в порядке убывания числа статей по конкретной проблеме, то журналы можно разбить на три зоны таким образом, чтобы количество статей в каждой зоне по заданной теме было одинаковым.

При этом в *первую зону*, названной Брэдфордом зоной *ядра*, входят профильные журналы, непосредственно посвященные рассматриваемой тематике. Количество журналов в зоне ядра невелико. *Вторую зону* образуют журналы, частично посвященные заданной области, причем их число существенно возрастает по сравнению с числом журналов в ядре. *Третья зона* – самая большая по количеству изданий – объединяет журналы, количество которых весьма далека от рассматриваемой предметной области.

При равном числе публикаций в каждой зоне число источников (в данном примере – наименований журналов) резко возрастает при переходе от одной зоны к другой. Брэдфорд установил, что число на-

¹ *Bradford S.C.* Documentation / S.C. Bradford. – London: Grosby Lockwool, 1948. – 156 p.; *Brookes B.C.* Theory of Bradford law / B.C. Brookes // J. Doc. – 1977. – Vol. 33. – № 3. – P. 180–209.

именований журналов в третьей зоне примерно во столько раз больше, чем во второй зоне, во сколько раз число наименований журналов во второй зоне больше, чем в ядре:

$$\frac{P_3}{P_2} \approx \frac{P_2}{P_1} \approx a \quad \text{и} \quad \frac{P_3}{P_4} \approx a^2, \quad \text{т.е.} \quad P_3 : P_2 : P_1 \approx 1 : a : a^2 : a^3,$$

где

$P_1 : P_2 : P_3$ – число наименований журналов в 1-й, 2-й и 3-й зонах соответственно.

7.5. Закон Викери

Б. Викери¹ уточнил модель **С. Брэдфорда**. Он выяснил, что журналы, проранжированные в порядке уменьшения в них статей по конкретному вопросу, можно разбить не на три зоны, а на любое число зон. Основной смысл закономерности Викери (рис. 7.3): если периодические издания расположить в порядке уменьшения в них количества статей по конкретному запросу, то в полученном списке можно выделить ряд зон, каждая из которых содержит одинаковое количество статей.

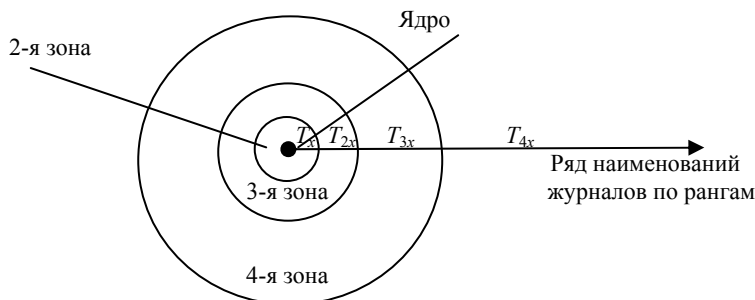


Рис. 7.3

При этом число журналов в первой зоне и нарастающее их число в последующих зонах соотносятся следующим образом:

$$T_x : T_{2x} : T_{3x} : T_{4x} : \dots = 1 : a : a^2 : a^3 : a^4 : \dots,$$

где x – количество статей в каждой зоне; T_x , – количество журналов, содержащих x статей; T_{2x} , T_{3x} , T_{4x} , ... – количество журналов, содержащих $2x$, $3x$, $4x$ и т.д. статей соответственно.

Часто этот закон называют законом Брэдфорда в толковании Викиери.

¹ *Vickery B.C.* Bradford's law of scattering // *J. Doc.* – 1948. – Vol. 4. P. 198 – 203.

7.6. Закономерность концентрации–рассеяния В.И. Горьковой

В работах **В.И. Горьковой**¹ было показано, что количественные закономерности строения ДИП могут быть выражены как посредством количественных параметров ранговых распределений (закономерность Ципфа), так и посредством параметров упорядоченных потоков элементов ДИП (закономерность Брэдфорда–Викери), т. е. что эти закономерности отображают одно феноменологическое явление.

Смысловый анализ эмпирических данных строения подсистем ДИП, проведенный в работах **В.И. Горьковой** и ее учеников¹, позволил зафиксировать наличие специфических функциональных свойств различных частотных зон ранговых распределений и упорядоченных потоков. Научные журналы, наиболее продуктивные по числу научных статей для соответствующего тематического раздела, сосредоточены в зоне рангового распределения, где частоты появления наименований элементов наибольшие, т. е. в ядре рангового распределения.

Структурное *подмножество ядра* рангового распределения состоит из *профильных научных журналов*, которые составляют ~10% от всех наименований журналов, содержащих публикации по данному тематическому разделу; при этом в научных журналах, входящих в ядро рангового распределения, содержится 50–60% всех публикаций по данной тематике.

Тематическое содержание профильных журналов ядра рангового распределения определяет тематические признаки данного упорядоченного потока.

¹ **Горькова В.И.** Информетрия: Количественные методы в научно-технической информации / В.И. Горькова // Итоги науки и техники. – Сер. Информатика. – Т. 10. – М.: ВИНТИ, 1988. – 328 с.; **Горькова В.И.** Анализ документальных информационных потоков и изучение запросов потребителей информации: Лекции / В.И. Горькова, Т.И. Гусева. – М.: ИПКИР, 1974. – 60 с.; **Горькова В.И.** Совершенствование системы информационного обеспечения на основе статистического анализа информационных потребностей специалистов / В.И. Горькова, Б.В. Петренко. – Минск: БелНИИИТИ, 1973. – 543 с.; **Горькова В.И.** Ранговое распределение на множествах научно-технической информации / В.И. Горькова // НТИ. – 1969. – Сер. 2 – № 7. – С. 5–11.; **Горькова В.И.** Статистические оценки статистических совокупностей документальных информационных потоков / В.И. Горькова // НТИ. – 1971. – Сер. 2. – № 12. – С 14–20.; **Горькова В.И.** Закономерности распределения публикаций в периодических и продолжающихся изданиях по электротехнике и энергетике / В.И. Горькова, С.П. Меллион // НТИ. – 1968. – Сер. 2. – № 11 – С. 3–7.; **Горькова В.И.** Частотное распределение множества ключевых слов / В.И. Горькова, К.И. Нумычева // НТИ. – 1970. – Сер. 2. – № 6. – С. 3–8.; **Горькова В.И.** Системное исследование документального информационного потока / В.И. Горькова // Системные исследования: Ежегодник, 1979. – М.: Наука, 1980. – С. 240–266.

Для ядерного подмножества терминов рангового распределения характерно вхождение в него заглавных понятий, тематическое содержание которых определяет классификатор данного тематического раздела.

В зону ядра рангового распределения авторов публикаций входят имена исследователей, которым принадлежат основополагающие положения тематического раздела ¹.

На основе исследования этих специфических особенностей ядра **В.И. Горьковой** был сделан вывод о том, что явление образования ядра рангового распределения системы ДИП можно трактовать как свойство системы ДИП концентрировать в ядре рангового распределения логическую информацию, определяющую основные понятия предметной области, объектов и методов исследования соответствующей отрасли науки (техники), научной дисциплины.

Явление образования ядра Горькова рассматривает как организацию высокого уровня, когда между элементами подмножества ядра существуют взаимосвязи, играющие более важную роль, чем связи между элементами других подмножеств. В качестве одного из важных параметров взаимосвязанности для формирования ядра В.И. Горьковой предложены меры «совстречаемости» понятий.

Элементы ядра выступают в роли «организаторов» подсистемы ДИП в конкретной отрасли науки или научной дисциплине, т.е. выполняют функции организующих отношений системы ДИП. Наименования элементов зоны ядра ДИП определяют условия включения элементов всех других частотных зон ранговых распределений в данную упорядоченную совокупность.

Одновременно с явлением образования ядра рангового распределения элементов ДИП по тематическому разделу наблюдается рассеяние элементов.

Так, если ядро содержит ~10% от всех наименований элементов ДИП, что составляет 50–60 % упорядоченного потока подсистем ДИП, то в зоне рассеяния при наименовании элементов ~90% сосредоточено всего 40–50 % упорядоченного потока документов.

Зона рассеяния рангового распределения определяет структуру связи подсистем ДИП различных отраслей науки (техники) или научных дисциплин, взаимно обогащающих друг друга семантической информацией.

Рассмотренные функциональные свойства двух зон рангового распределения и упорядоченного потока огрубленно трактуют функциональные свой-

¹ **Горькова В.И.** Частотное распределение множества ключевых слов / В.И. Горькова, К.И. Нумычева // НТИ. – 1970. – Сер. 2. – № 6. – С. 3–8.

ства ранговых распределения. В зоне рассеяния можно выделить третье структурное подмножество элементов с малой частотой появления, обладающее специфическими свойствами. В дальнейшем, видимо, возможна и более детальная структуризации зоны рассеяния.

Рассмотренное феноменологическое явление концентрации и рассеяния элементов ДИП было названо *закономерностью концентрации–рассеяния*.

Действие закономерности концентрации–рассеяния основано на том, что упорядоченное множество элементов ДИП имеет два структурно-функциональных свойства.

Во-первых, концентрировать подмножества ядерных элементов, функция которых состоит в том, чтобы:

- а) сосредотачивать логическую информацию, определяющую основные предметы, объекты и методы исследований в отрасли науки;
- б) выполнять роль «организующих отношений».

Во-вторых, рассеивать подмножество элементов, функция которых состоит в том, чтобы:

- а) определять структурные связи подсистем ДИП различных отраслей науки и научных дисциплин, обогащающих друг друга семантической информацией;
- б) обеспечивать определенную целостность всей подсистемы ДИП.

Закономерности концентрации–рассеяния позволяет использовать параметры аппроксимирующих функций, описывающих ранговые распределения и упорядоченные потоки в качестве параметров, определяющих функциональные свойства упорядоченных совокупностей элементов ДИП.

Следует еще раз обратить внимание на тот факт, что численные меры упорядоченности строения ДИП, определяемые статистическими методами только тогда могут использоваться в практике научно-информационной деятельности, когда они сочетаются с методами исследования функциональных свойств упорядоченных совокупностей ДИП. С примерами статистических оценок параметров закономерностей ДИП можно познакомиться в работах В.И. Горьковой¹.

¹ **Горькова В.И.** Информетрия: Количественные методы в научно-технической информации / В.И. Горькова // Итоги науки и техники. – Сер. Информатика. – Т. 10. – М.: ВИНТИ, 1988. – 328 с.; **Горькова В.И.** Анализ документальных информационных потоков и изучение запросов потребителей информации: Лекции / В.И. Горькова, Т.И. Гусева. – М.: ИПКИР, 1974. – 60 с.

7.7. Перспективы развития информетрии

На основе идей законов Ципфа–Мандельброта и Брэдфорда–Викери, закономерности концентрации–рассеяния, сформулированной В.И. Горьковой, развиваются методики автоматизации индексирования и анализа текстов, введения весовых коэффициентов терминов (см. ссылки в <http://www.medialingvo.ru>).

Вводятся меры веса ключевых слов.

Так, в работах *Спарка Джонса* экспериментально показано, что если N число документов и n число документов, в которых встречается данный индексный термин (ключевое слово), то вычисление его веса по формуле:

$$W = \log \frac{N}{n} + 1$$

приводит к более эффективным результатам поиска, чем без использования оценки значимости индексного термина, т. е. определенное значение имеет не только частота применения слова в конкретном документе, но и число документов, в которых это слово встречается.

Вводятся логарифмические меры.

Например, чтобы избавиться от лишних слов и в тоже время поднять рейтинг значимых слов, вводят инверсную частоту термина

$$i_x = \log (N / n_i),$$

где N – количество документов в базе данных; n_i – количество документов с термином i .

А затем каждому термину присваивают весовой коэффициент, отражающий его значимость в форме:

$$j = j_x / i_x,$$

где j – вес термина i в документе; j_x – частота термина i в документе; i_x – инверсная частота термина.

В новом смысле используется термин «ядро».

В 1995 г. на симпозиуме в Дублине предложена интересная и полезная для совершенствования информационного поиска идея «Дублинского ядра» (Dublin Core)¹, основанная на формировании метаданных, зафиксированных в спецификации определенного стандарта, и представлении k -го документа множеством пар $D_k = \{N_{ik}, V_{ik}\}$, где N_{ik} – имя i -го элемента метаданных Дублинского ядра в описании содержания k -го документа; V_{ik} – значение этого элемента метаданных. Аналогично описывается запрос.

Перспективным представляется использование для формирования «Дублинского ядра» закономерности концентрации–рассеяния.

¹ *Коголовский М.Р.* Перспективные технологии информационных систем / М.Р. Коголовский. – М.: ДМК Пресс; М.: Компания АйТи, 2003. – 288 с.

Возрастает интерес и к способам оценки текстов. Например, к работам *Г. Луна*², в которых предложения текста оцениваются в соответствии с параметром:

$$V_{np} = \frac{N_{zc}^2}{N_c},$$

где V_{np} – значимость предложения; N_{zc} – число значимых слов в предложении; а N_c – полное число слов в предложении.

Используя этот критерий, из любого документа можно отобрать некоторое число предложений. Понятно, что они не будут составлять членораздельного текста. Нужно учитывать также, что значимые слова должны браться из тематического тезауруса или отбираться экспертом. По той причине методика может лишь помочь человеку, а не заменить его (во всяком случае, на современном этапе развития вычислительной техники).

Закономерности организации ДИП, введения количественных мер терминов, предложений и др. компонентов текста полезно использовать на всех этапах создания информационно-поисковых систем: при комплектовании информационных фондов, создании информационно-поисковых языков и логико-семантического аппарата ИПС, при организации справочно-информационного обслуживания в библиотеках и отделах научно-технической информации, при создании и совершенствовании классификационных систем, выявлении тенденций роста и старения ДИП, при аналитико-синтетической обработке текстовой информации.

В последнее время на основе идеи закономерности концентрации–рассеяния разрабатываются методы выявления информационного ядра предметной области при построении информационной системы для реорганизации бизнес-процессов, при создании виртуальных предприятий.

² **Luhn H.P.** Automatic creation of literature abstracts BM / H.P. Luhn // Journal of Research and Development, 2, 1958, p. 159–165.

8. ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ТЕОРИИ ИНФОРМАЦИОННЫХ СИСТЕМ

8.1. Концепция архитектуры предприятия и ее применение для развития теории информационных систем

Предпосылки возникновения понятия «архитектура предприятия»

В настоящее время все больше внимания уделяют новому направлению, которое стали называть *архитектурой предприятия*.

Это направление изначально предназначалось для решения двух основных проблем в сфере информационных технологий (ИТ), выявленных двадцать лет назад.

Первая проблема заключается в постоянном увеличении сложности ИТ-систем, что приводило к увеличению затрат на их построение.

Вторая вызвана тем, что со временем получить реальную отдачу от ИТ-систем становится труднее, т.е. несмотря на всевозрастающую стоимость ИТ-систем, организациям с большим трудом удавалось поддерживать их соответствии требованиям бизнеса.

Очевидно, эти проблемы взаимосвязаны. Чем сложнее система, тем труднее ею управлять. Чем лучше удастся справиться со сложностью системы, тем выше вероятность получения от нее реальной выгоды.

В итоге – высокие затраты, низкая эффективность. Эти проблемы, впервые выявленные 20 лет назад, постепенно достигли критической точки. Стоимость и сложность ИТ-систем выросли экспоненциально, а реальная польза от них резко снизилась.

Для создания простой однопользовательской информационной системы понятие «архитектора» не нужно. Но в процессе создания корпоративной, распределенной системы представление об архитектуре все более усложняется.

Принципы построения архитектуры предприятия, которые двадцать лет назад представлялись странными и далекими от действительности, в настоящее время становятся все более востребованными.

За прошедший период было разработано множество методологий построения архитектуры предприятия.

История развития архитектур предприятия приведена в следующем подразделе и в Приложении 1.

Краткая история развития концепции архитектуры предприятия

Понятие «архитектура предприятия» появилось в 1987 г. в статье *Дж. А. Захмана* «Структура архитектуры информационных систем», опубликованной в журнале «IBM Systems Journal»¹.

Видение Захмана заключалось в том, что для обеспечения высокой ценности и гибкости бизнеса необходим целостный подход к архитектуре систем, в рамках которого каждая существенная проблема рассматривается с разных точек зрения. Такой подход к созданию архитектуры систем представляет собой то, что Захман изначально называл *архитектурной структурой* информационных систем, а впоследствии – *структурой архитектуры предприятия*.

В последующем Захман внес вклад в разработку архитектуры ИС Министерства обороны США. Эта попытка была предпринята в 1994 г., а сама концепция получила название «Базовая архитектура технического обеспечения для управления информацией» (Technical Architecture Framework for Information Management – TAFIM).

Преимущества, обеспечиваемые такими архитектурами предприятий, как TAFIM, были отмечены Конгрессом США. Конгресс в 1996 г. принял закон, известный как акт Клингера – Коэна², который определил реформу управления информационными технологиями.

В соответствии с этим законом всем федеральным агентствам было предписано принять меры по повышению эффективности инвестиций в ИТ. Для надзора за выполнением закона был сформирован совет директоров по информационным технологиям, в который вошли директора по информационным технологиям из всех основных правительственных органов.

С этого периода началось активное развитие идеи архитектуры предприятий и организаций.

В апреле 1998 г. Совет директоров по информационным технологиям начал работу над первым крупным проектом – структурой архитектуры федеральной организации (Federal Enterprise Architecture Framework – FEAF). Версия 1.1 данной структуры была выпущена в сентябре 1999 г.

В 1998 г., четыре года спустя после разработки TAFIM и два года спустя после оформления этой методологии в виде акта Клингера – Коэна, методология TAFIM была официально отменена Министерст-

¹ *Захман Дж.А.* Структура архитектуры информационных систем / Дж. А. Захман // IBM Systems Journal. – 1987. – Том 26. – № 3.

² Акт Клингера – Коэна от 1996 г. (PL 107–347) (см. THOMAS [Библиотека Конгресса]).

вом обороны. Все наработки по TAFIM были преобразованы в открытую группу, а затем – в новый стандарт, известный в настоящее время под названием TOGAF (The Open Group Architectural Framework).

Через некоторое время полномочия совета директоров по информационным технологиям по архитектуре федеральной организации были переданы Административно-бюджетному управлению.

В 2002 г. Административно-бюджетное управление переработало методологию FEAF и переименовало ее в архитектуру федеральной организации (Federal Enterprise Architecture – FEA).

Несмотря на весьма продуктивную деятельность федерального правительства по разработке архитектуры предприятия прогресс практически отсутствовал. В 2004 г., восемь лет спустя после принятия акта Клингера – Коэна, предписывающего использовать эффективные процессы планирования ИТ, Бюджетно-контрольное управление представило отчет о внедрении идеи архитектуры.

В 2003 г. Министерством Обороны США была разработана DoDAF (Department of Defence Architecture Framework), а консорциумом The Open Group – методология TOGAF (The Open Group Architecture Framework – структура архитектуры The Open Group)¹.

С января 2005 г. Бюджетно-контрольное управление серьезно наказало ряд правительственных агентств за невыполнение предписаний по внедрению и использованию архитектуры предприятия. В их числе ФБР, Министерство обороны, Министерство национальной безопасности и НАСА.

В 2005 г., примерно в то же время, когда Административно-бюджетное управление стало доминировать в области разработки архитектуры предприятия в государственном секторе экономики, доминировать в частном секторе экономики стала другая организация – группа Gartner.

Методология Gartner не является ни таксономией (как модель Захмана), ни процессом (как TOGAF), ни полной методологией (как FEA). Эта методология представляет собой набор практических рекомендаций по построению архитектуры предприятия от одной из наиболее известных в мире исследовательских и консалтинговых ИТ-организаций – компании Gartner.

Компания Gartner считает, что архитектура предприятия должна начинаться с того, что организация собирается достичь, а не с текущего положения дел, т.е. базироваться на системно-целевом подходе. После того как в организации сформировано единое представление о будущем, определяется влияние этого представления на архитектуру

¹ www.opengroup.org.

бизнеса, технологическую архитектуру, информационную архитектуру и архитектуру решений.

На основе этого подхода в 2006 г сформулированы рекомендации, касающиеся разработки архитектуры в виде последовательности шагов и задач участников, которые, однако, не детализированы до уровня моделей процесса разработки архитектуры, а затем разработана Методика описания архитектуры Gartner, которая представляет собой трехмерный куб: горизонтальные слои – бизнес-архитектура; вертикальные домены – информационная архитектура (приложения, данные, интеграция, доступ); вертикальные элементы – техническая архитектура (инфраструктура, системное управление, безопасность).

К 2005 г. компания Gartner стала одной из наиболее влиятельных организаций, занимающихся консалтингом на уровне директоров по информационным технологиям. Однако в области разработки архитектуры предприятия наиболее известной ИТ- и консалтинговой компанией была не Gartner, а Meta Group.

В 2005 г. Компания Meta Group предложила Методику META Group, отличительной особенностью которой является более детальное и формализованное описание именно *процесса* разработки архитектуры и всех его составляющих. META Group отмечает, что архитектура реализуется на практике через процесс управления ИТ- программами и проектами.

Компания Gartner пыталась создать рекомендации по разработке архитектуры предприятия, однако ей не удалось превзойти Meta Group. В 2005 г. компания Gartner решила, что, раз ей не удастся конкурировать с Meta Group, можно поступить иначе: компания Gartner приобрела компанию Meta Group.

После приобретения Meta Group компания Gartner/Meta потратила год на то, чтобы разобраться, какой вклад внесла каждая из компаний в методологию разработки архитектуры предприятия и как наилучшим образом объединить свои заметно различающиеся подходы.

В конце концов, был применен достаточно простой алгоритм: если компания Meta Group одобряла какой-либо пункт, он включался в методологию, если не одобряла, соответствующий пункт исключался. Так, компании Gartner нравились архитектурные *структуры*, компании Meta Group нравился архитектурный *процесс*. Поэтому с учетом принятого принципа *структуры* были исключены, а *процессы* включены.

История возникновения основных видов архитектур приведена на рис. 6.1. Краткая характеристика архитектур – в табл. 8.1. Более подробно основные современные методологии архитектуры предприятия рассматриваются в Приложении 1.

Таблица 8.1

Краткая характеристика концепций архитектуры предприятия

Дата	Автор Название	Основная идея
1987	Модель Дж.А. Захмана.	Модель предприятия представляется в виде набора согласованных описаний, которые соотносятся с ячейками формализованной матрицы (рис. П1.1 в Приложении 1). В матрице учтены существенные для архитектуры аспекты предприятия. Заполнение матрицы происходит сверху вниз.
1994	Модель Министерства обороны США TAFIM	Концепция «Базовая архитектура технического обеспечения для управления информацией» – TAFIM (<i>Technical Architecture Framework for Information Management</i>).
1992	Стивен Спивак. Модель ЕАР (Enterprise Architecture Planning – Планирование архитектуры предприятия) [6, 20].	Сегментный подход к разработке архитектуры, взгляд на предприятие с точки зрения его бизнес-функций и требований в области информации. Включает 7 шагов, определяющих архитектуру и соответствующий план ее реализации (рис. П1.3 в Приложении 1). Модель основана на упрощенной матрице Захмана 1987 г. Суть процесса ЕАР состоит в определении верхних строк этой матрицы (при этом включена концепция технических средств), Методика ЕАР обеспечивает взгляд на предприятие с точки зрения его бизнес-функций и требований в области информации.
1994	В. Iyer, R. Gottlieb. Модель 4-доменной архитектуры (Four Domains architecture, FDA)	Предлагается условно разбить ячейки модели Захмана на 2 компонента – архитектуру описания (Architecture-in-Design) и архитектуру исполнения (Architecture-in-Operation). При этом первая компонента описывает ход, средства и артефакты процесса разработки архитектуры предприятия, в то время как вторая предназначена для описания непосредственно бизнес-процессов и реализации ИТ-систем.
1995	Филипп Кручтен. Модель «4+1» (точнее 2The 4+1) View Model of Architecture	Предлагается использовать 5 представлений (рис. П1.4 [6]). Четырьмя основными представлениями в этой методике являются следующие. Логическое представление – отвечает на вопрос: что система должна выполнять в терминах конечных пользователей? Для иллюстрирования могут применяться диаграммы классов (в нотации языка UML).

Продолжение таблицы 8.1

Дата	Автор Название	Основная идея
		<p>Процессное представление. Учитывает нефункциональные требования к системе, включая производительность и доступность. Описывает вопросы параллельного исполнения и синхронизации процессов.</p> <p>Физическое представление. Описывает размещение программных компонент системы на аппаратных платформах и аспекты, связанные с физическим расположением системы.</p> <p>Представление уровня разработки. Описывает фактическую организацию модулей системы, разделение ее на подсистемы, которые могут разрабатываться независимо.</p> <p>Архитектура системы во многом определяется сценариям, которые объединяют все представления вместе.</p>
1996	Английская консалтинговая компания Systems Advisers Ltd. Стратегическая модель архитектуры SAM (Strategic Architecture Model)	SAM – это надстройка над моделью архитектуры предприятия Захмана. Она предоставляет общие структуры для определения архитектуры и механизмы, позволяющие организовать и анализировать информацию об архитектуре. SAM использует нотацию «сфер интересов» для представления целостного набора фактов о предприятии и «отношений», которые связывают эти факты в «полезные» группы (рис. П1.5).
1996–1997	Евгений Зиндер. «3D –модель» предприятия	Введена ось времени, где располагаются интервалы осуществления различных проектов и стадий развития ИС и всего предприятия. В качестве других осей выступает матрица Захмана (см. рис. П1.6) [20]: 1) ось уровня проектирования и использования ИС; 2) ось раздела обеспечения и аспекта работы ИС; 3) ось времени, в котором развивается предприятие и его ИС.
1999	Ассоциация CIMOSA, рабочая группа IFIP-IFAC. Обобщенная референсная архитектура и методология предприятия GERAM	<p>Методология GERAM (Generalised Enterprise Reference Architecture and Methodology) включена в качестве приложения в действующий базовый стандарт по архитектуре предприятия – ISO 15704:2000 «Requirements for enterprise-reference architectures and methodologies».</p> <p>Схема GERAM предусматривала (см. рис. П1.7 [6, 20]).</p>

Дата	Автор Название	Основная идея
1999	Совет директоров по информационным технологиям основных правительственных органов(CIO Council). Структура архитектуры федеральной организации FEAF. Версия 1.1 ISO 15704.	FEAF (Federal Enterprise Architecture Framework) состоит из 8 компонент (рис. П1.8 [6]). Методология описания архитектуры, ориентированная на государственные ведомства Подход к описанию архитектуры – со стороны бизнес-процессов. В документе содержался ряд инновационных идей, например, идея «сегментированных архитектур», т.е. рассмотрение в архитектурном аспекте сегментированных подмножеств крупного предприятия.
2002 – 2006	Административно-бюджетное управление. Архитектура федеральной организации FEA – Federal Enterprise Architecture.	Переработка методологии FEAF в FEA. FEA можно рассматривать и как методологию создания архитектуры предприятия, и как результат применения этой процедуры к конкретной организации – Правительству США. Схема сегментов федерального правительства приведена на рис. П1.14 [20]. Из рисунка видно, что многие сегменты используются во многих агентствах и все или почти все эти сегменты можно использовать повторно.
Де-кабрь 2003	Министерство Обороны США. Рамочная архитектура Министерства Обороны или DoDAF – Department of Defence Architecture Framework.	DoDAF содержит правила, руководства и документы, которые должны использоваться при разработке и описании архитектуры различных систем, используемых военными ведомствами США. Структура основных представлений, используемых в DoDAF, приведена на рис. П1.13 [6]. Архитектуру описывают 3 представления: <ul style="list-style-type: none"> ● операционное; ● системное; ● представление технических параметров. Каждое из них используется для отражения различных архитектурных характеристик и атрибутов. Имеются пересечения – некоторые из атрибутов как бы объединяют два различных представления, что обеспечивает целостность, единство и единообразие в описании архитектуры.

Продолжение таблицы 8.1

Дата	Автор Название	Основная идея
2003	<p>Консорциум The Open Group .</p> <p>TOGAF – The Open Group Architectural Framework (Структура архитектуры The Open Group).</p>	<p>Модель TOGAF применяется для описания интеграционных компонент, использующихся для поддержки широкого спектра корпоративных приложений. Как архитектурный <i>процесс</i> модель TOGAF дополняет модель Захмана. Захман показывает, как следует классифицировать артефакты. Модель TOGAF описывает процесс создания артефактов.</p> <p>В состав модели TOGAF входят две основные компоненты – Методика ADM (Architecture Development Method), определяющая процесс разработки архитектуры, и Базовая Архитектура (Foundation Architecture), которая дополняется соответствующей базой данных ресурсов, включающей описания архитектурных принципов, примеров реализации, а также специализированный язык ADML. Общая структура TOGAF показана на рис. П1.9 и П1.10 в Приложении 1.</p>
2005	<p>Компания Gartner. Модель Gartner</p>	<p>Компания Gartner считает, что архитектура предприятия должна начинаться с того, что организация собирается достичь, а не с текущего положения дел. После того как в организации будет сформировано единое представление о будущем, можно рассматривать влияние этого представления на архитектуру бизнеса, технологическую архитектуру, информационную архитектуру и архитектуру решений.</p> <p>В данном подходе сформулированы рекомендации, касающиеся разработки архитектуры в виде последовательности шагов и задач участников, которые, однако, не детализированы до уровня моделей процесса разработки архитектуры.</p> <p>Методика описания архитектуры Gartner представляет собой как бы трехмерный куб (см. рис. П1.15 в Приложении 1)</p>
2005	<p>Компания Meta Group. Методика META Group</p>	<p>Отличительной особенностью методики META является более детальное и формализованное описание именно процесса разработки архитектуры и всех его составляющих (рис. П1.16).</p> <p>META Group отмечает, что архитектура реализуется на практике через процесс управления ИТ- программами и проектами.</p>

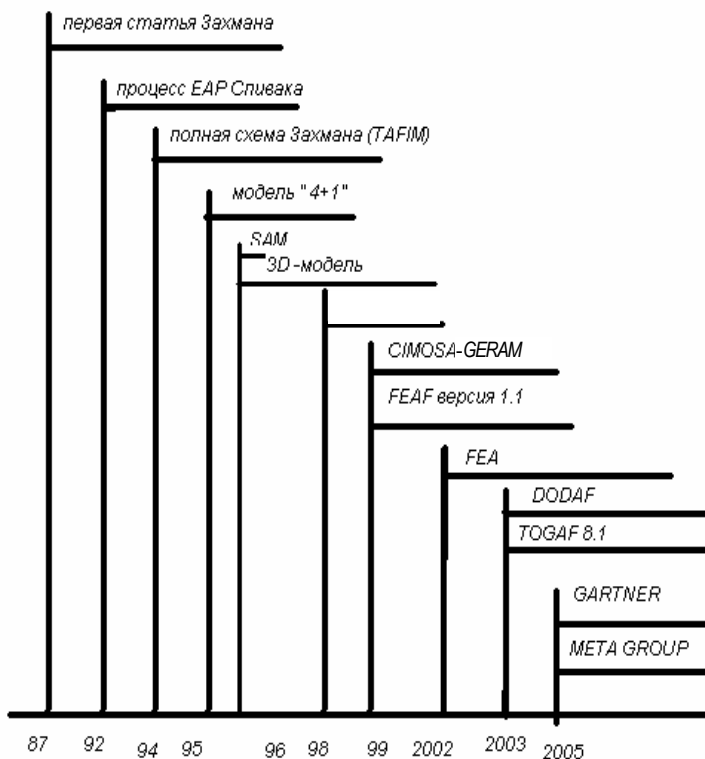


Рис. 8.1. Диаграмма развития методов построения Архитектуры предприятия

Развитие представлений об архитектуре предприятия

Термин «архитектура» означает множество близких по смыслу понятий, но применяющихся в различных приложениях. Известных формальных определений архитектуры существует несколько сотен¹.

В англоязычной литературе – методиках, статьях, стандартах – термину «архитектура предприятия» соответствует термин *enterprise architecture*.

Дж. А. Захман в 1987 г. в статье «Структура архитектуры информационных систем» вначале называл свою концепцию *архитек-*

¹ См., например, сайт Института Проектирования Программного Обеспечения Карнеги-Меллона (SEI – Carnegie Mellon Software Engineering Institute) www.sei.cmu.edu/technology/architecture.

турной структурой информационных систем, а впоследствии – *структурой архитектуры предприятия*.

В публикациях, в основном в Интернете, содержится большое количество рекомендаций, материалы аналитических компаний, теоретические и практические наработки в области архитектуры ИТ таких организаций, как Совет Директоров по Информационным Технологичам госорганизаций США (CIO Council) и ряда других ¹.

Поскольку Захман предложил термин «архитектура» применительно к информационной системе, то первоначально было распространено определение в программистском стиле: *«Архитектура системы состоит из нескольких компонент, внешних свойств и интерфейсов, связей и накладываемых ограничений, а также архитектуры этих внутренних компонент»*.

Однако вскоре **Е. Зингер** предложил более широкое определение: *«Архитектура предприятия – это вовсе не архитектура информационных систем или технологий, данное понятие охватывает и устройство бизнеса (деятельности по государственному управлению, если это министерство), и базовые технологии (например, станки, банкоматы, технологические процессы и т. д.), и работников всех видов и рангов, и информационные технологии»* ².

«Архитектурный взгляд» на системы (как ИТ-системы, так и бизнес-системы) определен в стандарте ANSI/IEEE 1471-2000 как *фундаментальная организация системы, состоящая из совокупности компонент, их связей между собой и внешней средой, и принципы, которыми руководствуются при их создании и развитии*.

Формальное определение в стандарте IEEE1471 ³. Института инженеров-электриков и электронщиков для определения архитектуры предлагает метамодель. Этот стандарт определяет такие абстрактные элементы архитектуры, как представления, системы, среды, обоснования, заинтересованные стороны и т.д. В соответствии с этим представлением *«система обладает некоторой архитектурой, которая может быть определенным образом описана с различных точек зре-*

¹ Enterprise Architecture – Sergey Orlik (Microsoft Platforma 2011). www.slideshare.net/sorlik/enterprise-architecture-serge

² **Зингер Е.** «3D-предприятие» – модель трансформирующейся системы / Евгений Зингер // http://www.iteam.ru/publications/it/section_53/article_1272.

³ **Стандарт** IEEE 1471-2000. Рекомендации IEEE по архитектурному описанию преимущественно программных систем. Подробную информацию об этом стандарте описания архитектуры можно получить по адресу [http://www.enterprise-architecture.info/Images/Documents/IEEE 201471-2000.pdf](http://www.enterprise-architecture.info/Images/Documents/IEEE%201471-2000.pdf).

ния в зависимости от интереса тех людей (участников), которые рассматривают архитектуру системы. Каждой точке зрения на архитектуру системы соответствует определенное представление, основу которого составляет некоторый набор моделей».

Однако этот стандарт не определяет структуру собственно архитектуры предприятия. Например, говорится о том, что необходимо иметь различные представления архитектуры, но при этом не указывается, какие это должны быть представления.

Применительно к архитектуре предприятия формальное описание впервые было сформулировано в стандарте ISO 15704, который был предложен рабочей группой IFAC/IFIP (International Federation of Automatic Control / International Federation for Information Processing). Идея состояла в том, чтобы разработать максимально общую, так называемую эталонную (reference) модель архитектуры предприятия, которая охватывала бы дополнительно процесс развития предприятия во времени как проект, а также учитывала бы роль человеческого фактора. Тогда архитектуры отдельных подсистем, в том числе ИТ-системы предприятия, могут быть разработаны как специфические уточнения такой общей модели.

Разработанная как приложение к данному стандарту, такая общая эталонная модель архитектуры получила название GERAM (Generalized Enterprise Reference Architecture and Methodology – Обобщенная референсная архитектура и методология предприятия). Фактически GERAM представляет собой абстрактное описание архитектуры общего уровня, которая может быть использована для «привязки» и сравнения между собой различных практических моделей архитектур. В частности, в ее рамках определяются такие понятия, как «продукт», «жизненный цикл», «роль персонала в системе», «моделирование процессов» применительно к задачам описания функционирования предприятия ¹.

В Национальном стандарте Российской Федерации ² архитектура определяется как *«описание (модель) основного устройства (структуры) и связей, частей системы (физического или концептуального объекта или сущности)»*.

При этом предлагается различать два типа архитектур, имеющих отношение к интеграции предприятия, а именно:

а) системные архитектуры (называемые иногда архитектурами «типа 1»), действие которых распространяется на проектирование

¹ Базовый стандарт «Архитектура Предприятия – ISO 15704:2000 «Requirements for enterprise-reference architectures and methodologies».

² Национальный стандарт Российской Федерации «Промышленные автоматизированные системы. Требования к стандартным архитектурам и методологиям предприятия. Industrial automation systems. Requirements for enterprise-reference architectures and methodologies» (Дата введения – 2010-01-01)

системы, например, на компьютеризированную, являющуюся частью системы интеграции предприятия;

b) стандартные проекты предприятия (называемые иногда архитектурами «типа 2»), действие которых распространяется на организацию разработки и выполнения проекта, например, интеграцию предприятия или другую программу развития предприятия.

В некоторых работах выделяют такие виды, как *системная* архитектура (архитектура систем – System Architecture) и *программная* архитектура (архитектура программного обеспечения – Software Architecture).

При этом, хотя методика описания и проектирования архитектуры отдельных прикладных систем имеет много общего с подходами к описанию архитектуры предприятия в целом, тем не менее, архитектура программных систем является отдельной областью знаний, которой посвящено большое количество соответствующих публикаций.

Под «программной архитектурой» в зависимости от контекста может пониматься как архитектура взаимодействия приложений в рамках информационной системы предприятия (т. е. *архитектура приложений*), так и *архитектура программных модулей* или *архитектура взаимодействия различных классов в рамках одного приложения*.

Каждая из архитектур, в свою очередь, может рассматриваться с тем или иным уровнем детализации. Так, для *программной* архитектуры выделяют следующие уровни описания архитектуры:

- **концептуальная архитектура** определяет компоненты системы и их назначения, обычно в неформальном виде. Это представление часто используется для обсуждения с нетехническими специалистами, такими как руководство, бизнес-менеджеры и конечные пользователи функциональных характеристик системы (что система должна уметь делать, в основном, с точки зрения конечного пользователя);

- **логическая архитектура** выделяет, прежде всего, вопросы взаимодействия компонент системы, интерфейсы и используемые протоколы. Это представление позволяет эффективно организовать параллельную разработку;

- **физическая реализация**, которая описывает привязку к конкретным узлам размещения, типам оборудования, характеристикам окружения, таким как, например, используемые операционные системы и т.п.

Интересный пример анализа различных аспектов деятельности программного архитектора предлагается в стандарте IEEE 1471-2000 ¹.

В соответствии с определениями компании Gartner ² архитектура – это:

- *общий план или концепция, используемая для создания системы, такой как здание или информационная система, или «абстрактное описание системы, ее структуры, компонентов и их взаимосвязей»;*

¹ <http://www.bredemeyer.com>.

² Структура архитектуры предприятия Gartner: развитие, 2005 г. / Грета А Джеймс, Роберт А. Хэндлер, Энн Лапкин и Николас Галл. – 25 октября 2005 г. – Код Gartner: G00130855.

• *семейство руководящих принципов, концепций, правил, шаблонов, интерфейсов и стандартов, используемых при построении совокупности информационных технологий предприятия.*

Первое определение ориентировано на описание существующих и будущих систем, второе – на процесс их построения.

В.И. Галактионов¹ предлагает рассматривать архитектуру предприятия в двух аспектах:

- статическом – в некоторый фиксированный момент времени;
- динамическом – как процесс перехода (миграции) от текущего состояния к некоторому желаемому состоянию в будущем.

Рассматриваемая в статике архитектура предприятия состоит из следующих элементов:

- миссия и стратегия, стратегические цели и задачи;
- бизнес-архитектура;
- системная архитектура.

Рассматриваемая в динамике архитектура предприятия — это логически связанный цельный план действий и скоординированных проектов, необходимых для преобразования сложившейся архитектуры организации к состоянию, определенному как долгосрочная цель, базирующийся на текущих и планируемых бизнес-целях и бизнес-процессах организации.

Таким образом, архитектура предприятия в общем случае описывается следующими последовательно зависимыми разделами:

Проектирование системной архитектуры предполагает разделение системы на **наиболее крупные составные части** и принятие **конструктивных решений**, которые после их принятия *с трудом поддаются изменению*. Если впоследствии оказывается, что нечто изменить легче, чем казалось вначале, это «нечто» легко исключается из «архитектурной» категории.

Статический срез системной архитектуры на определённый момент времени включает:

- архитектуру приложений – функциональный и компонентный состав информационной системы;
 - архитектуру данных – способы взаимодействия систем и хранения данных;
 - архитектуру оборудования – используемые технические средства/решения.
- Другими аспектами системной архитектуры являются:
- способы и планы миграции от текущего состояния архитектуры к целевому;
 - способы передачи реализаций между средами;
 - стоимость решения, включая капитальные и операционные расходы.

При этом В.И. Галактионов считает, что существует более одного способа описания архитектуры. Степень важности каждого из этих способов. меняется в продолжении жизненного цикла.

¹ **Галактионов В.И.** Системная архитектура и ее место в архитектуре предприятия / В.И. Галактионов // Директор информационной службы, 2002. – № 05.

А.В. Данилин и А.И. Слюсаренко [6] проводят анализ определений архитектуры и предлагают в зависимости от контекста определять архитектуру системы как [6]:

1) многоаспектное описание или план задуманной или развиваемой системы на уровне ее компонентов, детализированное в достаточной мере для руководства ее воплощением, а также принципы и руководящие материалы, определяющие руководство конструированием и развитием системы во времени;

2) структура существующей системы как совокупность ее компонентов и их взаимосвязей.

В варианте второго пункта этого определения архитектура у предприятий была и всегда.

Опираясь на методологию Gartner, в соответствии с которой подход к формулировке архитектуры должен основываться на анализе бизнес-процессов и поддерживающих их приложений, авторы [6] делают вывод о том, что *«архитектура предприятия является одним из инструментов организационных изменений всего предприятия в целом с использованием ИТ, и особенно той части организации, которая отвечает за информационные технологии»*, и что *«представление об архитектуре предприятия имеет свои корни в дисциплине, которая получила название “системное мышление”»*.

И дают более подробное разъяснение этой идеи следующим образом.

«Отличительной характеристикой решений, принимаемых в отношении архитектуры, является то, что эти решения должны приниматься с учетом широкой, или системной, перспективы. Любое решение, которое может быть принято локально (например, в рамках подсистемы), не является архитектурным для системы в целом. Это позволяет делать различие между детальным проектированием и принятием решений по поводу практической реализации системы, с одной стороны, и архитектурными решениями – с другой. Первые решения имеют локальное влияние, а вторые – систематическое. Поэтому для проектных решений нужна соответствующая более широкая перспектива, позволяющая учесть системное влияние решений более высокого уровня, что дает возможность достичь желаемого уровня компромиссов и соглашений между составными частями для обеспечения должного уровня качества системы в целом».

При этом в [6] отмечается, что организации испытывают постоянные трудности в синхронизации целей и задач бизнеса и процессов развития своих информационных систем. Существует как бы «облако неопределенности» между определением организацией и обеспечивающей ее ИТ-инфраструктурой своих целей и задач.

Архитектура информационных технологий и архитектура предприятия в целом по мнению авторов [6] является основным механизмом

интерпретации и реализации целей организации через адекватные ИТ-инфраструктуру. Это достигается через создание определенного количества взаимосвязанных архитектурных представлений, которые с использованием различных методик описания архитектуры разбивают архитектуру предприятия на различное количество моделей и определений, относящихся к таким областям, как *бизнес, информация, прикладные системы, технологическая инфраструктура*.

Бизнес-модели описывают стратегию организации, структуры управления, требования, ограничения и правила, а также основные бизнес-процессы, включая взаимосвязи и зависимости между ними. Т.е. *бизнес-архитектура* описывает на уровне предприятия в целом то, как реализуются основные функции организации, включая организационные и функциональные структуры, роли и ответственности, расположение, время, типы файлов и баз данных и других информационных хранилищ.

Архитектура информации определяет ключевые активы, связанные со структурированной и неструктурированной информацией, требующейся для бизнеса, включая расположение, время, типы файлов и баз данных и других информационных хранилищ.

Архитектура прикладных систем описывает те системы, которые и обеспечивают необходимый функционал для реализации логики бизнес-процессов организации.

С точки зрения *технологической архитектуры*, важные модели включают описание ИТ-сервисов, которые требуются для реализации перечисленных выше трех других областей архитектуры.

При этом *логические модели* ИТ-сервисов построены в абстрактной, технологически независимой форме и оставляют свободу для оптимального выбора конкретных технологий. Но, в конце концов, архитектура предприятия завершается *физическими моделями*, которые определяются технологиями, аппаратными и программными платформами, выбранными для реализации ИТ-сервисов.

Иерархическое построение архитектуры позволяет облегчить ее восприятие человеком.

В определениях можно выделить, по крайней мере, три различных аспекта:

- иерархия архитектур различных организационных систем;
- соотношения между объективной реальностью и субъективным восприятием;
- соотношения между общесистемной архитектурой и частными архитектурами.

Можно говорить об *архитектуре предприятия в целом, архитектуре уровня отдельных проектов или семейства продуктов, об архитектуре отдельной прикладной системы*.

При этом все виды архитектура должны использовать сходные средства описания и представления результатов, опираться на методы декомпозиции (структуризации) сложных систем, а архитектура представляет собой некоторую модель реальной системы, которая динамически изменяется, сохраняя соответствие оригиналу.

Профессионалы в области информационных технологий понимают под архитектурой ИТ достаточно большой спектр понятий – структурированное семейство технических руководств, включая концепции, принципы, правила, шаблоны и интерфейсы, а также взаимосвязи между ними, которые используются при создании новых информационных систем и развитии существующих систем. В отличие от них, профессионалы в области бизнеса не рассматривают этот вопрос как вопрос исключительно технологий. Наоборот, они разговаривают в терминах бизнес-моделей, бизнес-процессов и иногда – бизнес-архитектуры.

В процессе эволюции понятия «Архитектура предприятия» этот термин означал все более комплексный и всеобъемлющий подход к описанию и практике использования информационных технологий по обеспечению основной деятельности организации и, как результат, получение все более широкого спектра преимуществ.

В ранних работах ИТ-архитектура понималась в основном как *технологическая* архитектура или архитектура, определяющая *инфраструктуру* информационной системы. Работы по описанию архитектуры были сосредоточены на формировании технологических стандартов и принципов, включая проведение инвентаризации различных технологий, используемых в организации. Такой подход позволяет добиться определенных частных выгод, связанных прежде всего с уменьшением стоимости закупок и эксплуатации информационных систем, уменьшением затрат на разработку приложений и обучение персонала. Однако он был ограниченным, так как не подразумевал ориентацию на решение бизнес-задач.

Следующей ступенью явилось понятие *корпоративной информационно-технологической архитектуры* масштаба предприятия (EWITA – *Enterprise-wide information technology architecture*). Это означало, что усилия по описанию архитектуры предприятия должны включать в себя описание архитектуры информации и архитектуры прикладных систем, а не только технологический уровень. Основное направление работ при этом состояло в совместном использовании общих данных, исключении дублирования бизнес-функций, координации управления пользователями, ресурсами, информационной безопасностью за счет улучшений в управлении портфелем прикладных систем. Корпоративная информационно-технологическая архитектура масштаба предприятия описывает то, как компоненты информационной

системы связаны между собой; точно так же бизнес-архитектура описывает то, как элементы бизнеса связаны между собой.

Такой подход обеспечивает более эффективное взаимодействие различных структурных подразделений организации, совместный доступ к информации различных подразделений и внешних организаций (клиентов, партнеров, поставщиков); уменьшение дублирования с точки зрения параллельной реализации близких по функционалу прикладных систем для различных бизнес-подразделений; решение проблем, которые затрагивают интересы нескольких подразделений, например, интеграция и взаимодействие информационных систем.

Важным логическим шагом для эффективного описания существующих в организации процессов и планируемых изменений явилось введение понятия *архитектуры предприятия (Enterprise Architecture)*, которая объединяет *корпоративную ИТ-архитектуру* масштаба предприятия с *бизнес-архитектурой* и позволяет обеспечить достижение стратегических целей предприятия.

Преимуществами такого включения бизнес-архитектуры в контекст рассмотрения целостной архитектуры предприятия являются большая способность организации к изменениям или динамичность (*agility*) и синхронизация возможностей информационных технологий с бизнес-стратегией, обеспечение вариативности бизнес-стратегии за счет возможности изменений в обеспечивающих процессах и технологических решениях; лучшие перспективы, с точки зрения использования возможностей информационных технологий по формированию бизнес-стратегии.

Таким образом, концепция архитектуры предприятия явилась результатом поиска некоторого целостного подхода, который обеспечил бы взгляд на организацию как единую систему.

Введены понятия *архитектуры предприятия, архитектуры приложений, архитектуры данных, технологической архитектуры*, что позволяет кратко охарактеризовать различные многомерные структуры различных страт ИС – страты данных, программных приложений, клиент-серверов, технологической страты, ИС предприятия в целом, т.е. *архитектуры предприятия*.

Архитектура предприятия представляет собой концепцию, помогающую организациям изучить собственную структуру и принципы работы. Она представляет собой комплексный набор моделей, описывающих структуру и функции предприятия. Важные сферы применения этих моделей – систематическое планирование и создание архитектуры ИТ, а также принятие стратегических решений.

Отдельные модели в архитектуре предприятия упорядочены логически, что позволяет получить высокодетализированное описание предприятия, включая цели и задачи; процессы и их организация; системы и данные; используемые технологии.

При этом понятия *структуры* и *архитектуры* используются неоднозначно:

- либо понятие «структура» характеризует строение, расположение, порядок, т.е. конфигурацию ИС – «структура архитектуры» у Захмана;

- либо существует точка зрения ¹, в соответствии с которой под *архитектурой* понимается описание системы с точки зрения конечных пользователей и интерфейсов взаимодействия с внешней средой, т.е. как внешний взгляд на ИС, а *структуру* ИС описывают в виде взаимодействующих между собой подсистем, т.е. внутреннюю конфигурацию ИС. При этом каждая подсистема может быть разделена на составные части в иерархии, вплоть до модулей прикладных программ, принимаемых за неделимые элементы. Таким образом, понятие структуры представляется в форме иерархии (древовидной или стратифицированной), включающей несколько уровней разбиения, и полученные структурные единицы, в свою очередь, могут быть представлены в виде архитектурного описания по отношению к внешним структурным единицам.

8.2. Концепция ситуационных центров

В теории разработки АСУ предусматривалось, что АИС являются первой очередью АСУ. В дальнейшем предпринимались попытки разработки вначале *информационно-справочных* и *информационно-советующих систем* (ИСС).

Затем появился термин «*системы поддержки принятия решений* (СППР)».

В то же время практика развития информационных систем для предприятий и организаций в основном пока ограничивается разработкой или приобретением программных продуктов, обеспечивающих учетно-отчетную деятельность систем организационного управления предприятиями.

Эти ИС хорошо обеспечивают отчетность предприятий перед вышестоящими органами отраслевого и территориального управления и разнообразными контролирующими организациями.

Если разработаны требования к объему, времени выпуска продукции и др. количественным характеристикам производственных предприятий или обслуживающих организаций, то АИС позволяют определить рассогласование реальных результатов с желаемыми, и представить эту информацию лицам, принимающим решение.

¹ **Бойченко А.В.** Основы открытых информационных систем / А.В. Бойченко, В.К. Кондратьев, Е.Н. Филинов. – М.: Изд. Центр АНО «ЕОАИ», 2004. – 126 с.

Первоначально поставленная задача создания именно АСУ, была реализована в основном на уровне технологических процессов – АСУ ТП, которые разрабатывались для ряда технологических процессов, для непрерывных производств, иногда – для управления конвейером.

Например, на Волжском автозаводе (ВАЗ) поставка комплектующих на конвейер сборки автомобиля обеспечивалась автоматически на основе регистрации комплектующих на выходе автомобиля с конвейера и обработки этих данных информационным центром.

А для системы организационного управления даже промышленными предприятиями реализованы были только АИС, обеспечивающие информацией лиц, принимающих решения, а процессы собственно принятия решений остаются за человеком.

В то же время для ряда прикладных проблем требуются новые подходы к информационно-аналитическому обеспечению управленческой деятельности.

Определенным шагом на пути к повышению степени автоматизации управления предприятиями и организациями являются системы бизнес-аналитики¹, которые на основе различных технологий обработки данных обеспечивают кластеризацию систем, группирование и классификацию информации, что в ряде ситуаций является действенной помощью при принятии решений.

В настоящее время развивается форма информационно-аналитических систем, называемых *ситуационными центрами* (СЦ), которые должны представлять собой инструмент поддержки управленческой деятельности.

Концепция ситуационного центра была предложена английским кибернетиком **Стаффордом Биром** в 1970-е гг. Первый ситуационный центр для первых лиц государства был создан под руководством Бира в конце 1970-х гг. В настоящее время их число составляет несколько сотен и продолжает возрастать.

В нашей стране активное развитие ситуационных центров началось в 1990-е гг.

Современный ситуационный центр определяют как организационно-техническую систему, реализующую функции подготовки и поддержки принятия управленческих решений, которая позволяет наиболее полно и оперативно представлять информацию о сложившейся ситуации органам управления, прогнозировать возможные сценарии ее развития, оперативно подготавливать возможные альтернативные варианты управленческих решений и оценивать их последствия.

¹ См., напр., **Паклин Н.Б.** Бизнес-аналитика: от данных к знаниям (+ CD)/ Н.Б. Паклин, В.И. Орешков. – СПб.: Питер, 2009. – 624 с.

Создание и распространение ситуационных центров было вызвано высокой динамичностью, сложностью, многоаспектностью, увеличением степени неопределенности задач управления, решаемых в органах государственной власти и крупных корпорациях. Создание таких систем помогает лицам, принимающим решение, в переработке больших объемов информации, в подготовке сценариев принятия решений.

Функционирование ситуационных центров основано на следующих принципах [9]:

- непрерывный мониторинг и моделирование протекающих процессов, прогнозирование сценариев развития ситуаций;
- визуализация управленческих ситуаций и причинно-следственных связей анализируемых событий;
- организация коллективной выработки решений с использованием информационных ресурсов, интеллектуальных информационных технологий и средств отображения информации;
- обеспечение оперативного синтеза альтернативных решений.

Таким образом, ситуационные центры интегрируют в своем составе системы поддержки принятия управленческих решений (Decision Support System) и презентационные центры (Presentation Center).

Современные ситуационные центры позволяют решать такие задачи, как:

- обеспечение информационной поддержки руководителей;
- осуществление непосредственного доступа руководителей к территориально удаленным информационным ресурсам как структурных подразделений, так и других организаций;
- согласование и обеспечение целостности функционирования информационно-коммуникационных систем;
- осуществление доступа к информации организаций, взаимодействующих при принятии решения;
- сокращение временных и финансовых затрат, вызванных несовместимостью информационно-телекоммуникационных систем, дублированием подготовки данных, их противоречивостью, затруднениями с доступом, выборкой и передачей информации;
- интеграция информационных систем структурных подразделений в единое информационное пространство.

По характеру поддерживаемых управленческих задач выделены три класса ситуационных центров (табл. 8.2)¹.

¹ Выступление *Данчула А.Н.* / Роль ситуационных центров в системе поддержки принятия решений государственного управления: круглый стол 29.11.2010.

Классы ситуационных центров

Класс СЦ	Назначение	Примеры
<p>1. Оперативно-диспетчерские центры, решающие в реальном времени задачи оперативного управления сложными организационно-технологическими процессами с многочисленными информационными потоками</p>	<p>Изначально в центрах этого класса основное внимание уделялось презентационной компоненте, с развитием средств вычислительной техники в них не только стали появляться современные средства отображения информации коллективного пользования, но и все большую роль стало играть аналитическое информационно-программное обеспечение принятия решений.</p>	<ul style="list-style-type: none"> • Центр управления космическими полетами. • Центры управления перевозками. • И т.п.
<p>2. Ситуационные центры анализа и управления кризисными ситуациями.</p>	<p>Основное назначение подобных центров заключается в предотвращении кризиса за счет своевременного предоставления лицам, принимающим решения (ЛПР), исчерпывающей информации по текущему состоянию контролируемых объектов и прогнозов возможных сценариев развития событий.</p> <p>В случае же, если кризиса избежать не удалось, такие ситуационные центры становятся, по сути, оперативными штабами по управлению процессами локализации (ликвидации) последствий кризиса</p>	<p>Центр управления кризисными ситуациями Министерства по чрезвычайным ситуациям, созданный на основе решений, отработанных при ликвидации последствий Чернобыльской катастрофы в 1986 г. Этот центр обеспечивает визуализацию текущего и прогнозируемого состояния анализируемой ситуации, отображая имеющиеся силы и средства, предлагаемые рекомендации. На основе этой информации принимаются решения, которые в режиме реального времени доводятся до исполнительных структур и подразделений.</p> <p>Ситуационные центры этого класса имеются в Минатоме, Минприроде, других ведомствах и организациях.</p>

Класс СЦ	Назначение	Примеры
3. Информационно-аналитические ситуационные центры	Ориентированы на принятие не только оперативных, но и стратегических решений. К ним относятся многие из созданных в последнее время ситуационных центров для самых разнообразных применений. К этому классу можно отнести и широко известную «ситуационную комнату» президента США. В различных источниках указывается, что Президента США обслуживают от 4 до 5 ситуационных центров.	<ul style="list-style-type: none"> • Ситуационный центр в Совете Безопасности (СБ) при Президенте РФ. СБ позволяет осуществлять мониторинг, моделирование последствий, анализ событий, которые происходят в экономике, социальной сфере, в области национальной безопасности, помогая тем самым выработать решения (создан в 1994 г.) • Ситуационный центр в резиденции Президента РФ (введен в строй в феврале 1996 г.) • Ситуационные центры полномочных представителей Президента РФ в федеральных округах Пусковые комплексы развернуты в 2004 г.) • Информационно-аналитические ситуационные центры в ряде министерств и ведомств Российской Федерации, в администрациях глав свыше десяти субъектов Федерации, наиболее развитым из которых является СЦ ОАО «Российские железные дороги» – Ц РЖД (сдан в постоянную эксплуатацию в 2002 г.)

В настоящее время активно развиваются ситуационные центры третьего класса. Ситуационный центр представляет собой помещение, предназначенное для оперативного принятия управленческих решений, а также для контроля разного рода объектов и событий. Ситуационные центры используются федеральными и региональными органами власти, органами местного самоуправления, крупными промышленными предприятиями, образовательными учебными учреждениями.

Ситуационный центр оснащается новейшими средствами коммуникаций: оборудованием для интерактивного представления информации, системой конференцсвязи видеоконференцсвязи. Средства, используемые в ситуационных центрах, дают возможность наглядного

отображения огромных объемов информации, на основе которой принимаются решения и доводятся до исполнителей.

С конкретными примерами истории создания и развития ситуационных центров можно познакомиться, например, в [9, 18].

Существуют и другие классификации СЦ: по составу систем ситуационного моделирования (СЦ наблюдения, аналитические, полнофункциональные); по масштабу (степени сложности) решаемой задачи; по размещению; (стационарные, мобильные, виртуальные); по степени детерминированности решаемых задач; по целевой направленности (контроля, управления, кризисных ситуаций, обучения); по способу отображения информации (коллективные, индивидуальные); по универсальности (специальные, настраиваемые).

В связи с необходимостью ознакомления государственных служащих с современными управленческими и информационно-аналитическими технологиями, реализуемыми в ситуационных центрах, в Российской академии государственной службы (РАГС) при Президенте РФ в 2004 г. создан учебно-исследовательский ситуационный центр. В СЦ РАГС¹ предусматривается проведение занятий в форме: лекций-демонстраций, лекций-дискуссий, кейс-стадий, деловых игр. Поддержка практических занятий в СЦ РАГС реализуется в виде учебно-аналитических задач (УАЗ) по решению управленческих проблем. На практических занятиях обеспечивается автоматизированная поддержка коллективного доступа к информационным ресурсам, обмена данными между автоматизированными рабочими местами участников занятий и визуализация результатов их действий на видеостене.

Занятия, проводимые в СЦ РАГС, могут быть посвящены всем или некоторым из следующих укрупненных этапов решения проблемы:

- неформальное обсуждение аналитического доклада;
- формализация проблемной ситуации и формулировка проблемы;
- групповое моделирование проблемной ситуации;
- принятие коллективных решений.

Основное содержание кейс-стади – групповое моделирование проблемной ситуации; это комплекс действий по установлению возможных тенденций развития проблемной ситуации; возможных последствий ее развития; экспертному прогнозированию возможного развития проблемной ситуации в условиях применения различных

¹ *Данчул А.Н.* Первая очередь ситуационного центра РАГС: Концепция построения и состав // Информационно-аналитические средства поддержки принятия решений и ситуационные центры: Материалы научно-практической конференции, состоявшейся в РАГС 28 – 29 марта 2005 / Под общ. ред. А. Н. Данчула. – М.: изд-во РАГС, 2006. – С. 19–27.

стратегических решений. При использовании видеоконференций возможна организация микрогрупп удаленных участников.

В деловых играх групповое моделирование целесообразно дополнять процедурами принятия коллективных решений. Между микрогруппами распределяется несколько различных функциональных ролей участников коллективных действий.

Информационное обеспечение (ИО) СЦ РАГС строится по четырехуровневой схеме:

- 1) локальное ИО учебно-аналитических задач;
- 2) информационное обеспечение СЦ РАГС, включающее адаптированные базы данных статистической информации об экономике России и ее регионов;
- 3) информационное обеспечение общеакадемического уровня, включающее правовые базы данных и ресурсы, доступные с помощью информационно-обучающего портала «Государственное управление и местное самоуправление»;
- 4) информационные ресурсы Интернет.

При выборе технологии реализации информационных процессов (ИП) в СЦ РАГС учитывались их следующие основные характеристики:

- коллективность – участие в ИП группы взаимодействующих в интересах достижения общей цели пользователей;
- интеллектность – вследствие нечеткости исходной постановки и слабой формализуемости решаемой проблемы доля интеллектуальных действий, выполняемых человеком или в редких случаях программами искусственного интеллекта, соизмерима с долей рутинных действий или превышает ее;
- интерактивность – высокая чередуемость действий, выполняемых человеком и программно-техническими средствами, обуславливающая диалоговый характер реализуемых ИП;
- уникальность – отсутствие типовых полномасштабных технологий реализации ИП решения проблем;
- большой объем и разнообразие форм представления используемой информации, в том числе – значительная доля и существенная важность визуальной и аудиоинформации;
- сеансовость – возможность разбиения ИП решения проблемы на сеансы, разделенные определенными или неопределенными промежутками времени;
- высокая доля вспомогательных действий по подготовке сеанса: разработка или выбор сценария сеанса, сбор, отбор и подготовка данных, выбор и подготовка (настройка) необходимой программно-технической среды.

Ориентированность СЦ на решение уникальных, а не типовых задач означает, что он должен рассматриваться как система с развивающимися (за счет включения новых задач) функциями, а создание новых задач как штатный процесс функционирования. Уникальность также влечет достаточно длительный и ресурсоемкий цикл подготовки к решению задачи. В связи с этим в составе программного обеспечения СЦ РАГС весомую часть составляет инструментальное программное обеспечение, предназначенное для автоматизации разработки этих задач.

Большой объем и разнообразие форм представления информации обеспечивается в СЦ РАГС прежде всего за счет использования видеостены, позволяющей одновременно воспроизводить визуальную и с различных АРМ и видеисточника. Для полного использования преимуществ, предоставляемых видеостеной, важно чтобы эти формы поддерживались и специальным программным обеспечением.

Решение основных прикладных проблем информационно-аналитической поддержки управленческой деятельности в ситуационных центрах может быть достигнуто в ходе выполнения исследований по следующим направлениям.

1. Разработка теории (методов) создания информационно-аналитических систем как перманентно развивающихся систем решения уникальных слабо-формализованных задач.

2. Разработка методов и средств интеграции разнообразных программных средств и баз данных в условиях постоянного обновления их перечня.

3. Разработка методов и средств поддержки различных классов коллективной (в том числе территориально-распределенной) деятельности.

4. Разработка методов визуализации (создания информационных образов) решаемых задач, организация хранения и поиска визуальной информации.

Развитию идей и форм реализации ситуационных центров способствуют регулярно проводимые на базе РАГС конференции¹ и семинары по проблемам создания ситуационных центров, заседания за «круглым столом» на базе Аналитического центра при Правительстве РФ.

Эти конференции и заседания и экспертные обследования ситуационных центров позволили выявить основные проблемы, которые

¹ *Ситуационные* центры и перспективные информационно-аналитические технологии поддержки принятия решений: Матер. науч.-практич. конф. РАГС, 7-9 апреля 2008 г. – М.; Изд-во РАГС, 209. – 418 с.; *Ситуационные* центры 2010. Современные информационно-аналитические технологии поддержки принятия решений: Матер. науч.-практич. конференции / Под общ. ред. д-ра техн. наук, профессора А. Н. Данчула. . РАГС, 27-28 апреля 2010. – М.; Изд-во РАГС, 2011. – 420 с.

необходимо решить для развития ситуационных центров, основными из которых являются следующие ¹:

- СЦ не встроены в цепочку принятия решения;
- нет понимания роли СЦ как со стороны руководителей, так и со стороны обслуживающего персонала;
- не в полной мере используются имеющиеся ресурсы
- основные функции большинства СЦ – пока еще только презентации и учебные мероприятия;
- на уровне органов управления территориями и в надведомственных структурах ЛПР, как правило, не заинтересованы в развитии СЦ;
- кадры не готовы;
- слаба методическая подготовка;
- основными источниками информации являются Интернет и Росстат;
- редко используется многомерное представление и многофакторный анализ;
- документальное обеспечение сводится к должностным инструкциям.

На Федеральном уровне СЦ функционируют более эффективно, чем на региональном

В то же время анализ показывает, что, развития требуют исследовательские, методические, образовательные, организационные проблемы, проблемы нормативно-методического обеспечения управления функционированием СЦ. Необходимо разрабатывать типовые сценарии, типовую методологию функционирования СЦ. Необходима также пропаганда полезности СЦ для организации коллективного принятия решений. Заимствуя зарубежный опыт, необходимо учитывать, что в России в отличие от многих зарубежных стран, не ситуационная психология, а когнитивное мышление.

Таким образом, создание и практическое использование ситуационных центров как важного средства информационно-аналитических технологий поддержки принятия решений требует развития теоретических и методических основ процессов принятия решений на базе СЦ.

Утверждение Стратегии национальной безопасности и Основ стратегического планирования в Российской Федерации как инструментов Государственного управления дополнительно к прежним функция ситуационных центров добавила задачу создания системы ситуационных центров, ориентированных главным образом на стратегическое планирование и согласование стратегических целей общества

¹ *Роль* ситуационных центров в системе поддержки принятия решений государственного управления. 26 ноября 2010: Круглый стол. – М.: Аналитический центр при Правительстве РФ. – С. 7-8.

и стратегических целей государства в лице Правительства РФ. Создание системы ситуационных центров является компонентом более крупной задачи обеспечения вертикальных и горизонтальных связей в системе стратегического планирования в РФ, основанной на принципах долгосрочного и среднесрочного программно-целевого управления, решением задачи мониторинга и прогнозирования чрезвычайных ситуаций, создания системы независимой экспертной оценки программно-проектной деятельности органов государственной власти¹.

В следующих разделах рассматриваются подходы к разработке информационных систем организаций, основанные на применении теории систем и системного анализа.

8.3. Подходы к анализу и проектированию систем и их применение для развития теории проектирования информационных систем

На протяжении всей истории развития теории систем предлагались и применялись различные подходы к представлению (отображению), анализу и проектированию систем.

Традиционный подход, применяющийся в математических исследованиях определить элементы (переменные, константы) и связать их соответствующим соотношением (формулой, уравнением, системой уравнений), отображающим принцип взаимодействия элементов.

Когда задачи усложнились, и такое соотношение не удавалось сразу найти, то предлагалось формировать «пространство состояний» элементов и вводить «меры близости» между элементами этого пространства. Такой подход вначале пытались применить для исследования сложных систем.

Предлагалось обследовать систему, выявить все элементы и связи между ними. Этот подход называли иногда «перечислением» системы. При обследовании применялись разные способы: 1) *архивный* (изучение документов и архивов предприятия); 2) *опросный* или *анкетный* (опрос сотрудников, в том числе с помощью специально разработанных вопросников – анкет).

Однако первые же попытки применить такой подход к исследованию систем управления предприятиями и организациями показали, что «перечислить» сложную систему практически невозможно. В истории разработки автоматизированных систем управления был такой случай. Разработчики написали несколько десятков томов обследования системы, а так и не смогли присту-

¹ *Ситуационные* центры: нереализованный потенциал / В.И. Ерохин, М.В. Садофьев, Т.В. Еферина, И.М. Алексеева // Стратегическое управление: от идеи до результата. – М.: Аналитический центр при Правительстве РФ, 2011. С. 84–103.

пить к созданию АСУ, поскольку не могли гарантировать полноты описания. Руководитель разработки вынужден был уволиться, и в последствии стал изучать системный подход и популяризировать его.

Учитывая трудности «перечисления» системы, с самого начала возникновения системных теорий исследователи искали подходы к ее анализу и созданию.

Основными подходами к анализу систем, предлагавшимися разными исследователями, являются следующие:

- в начальный период становления теории систем развивался бихевиористский подход, основанный на исследовании поведения (*behaviour* – поведение) систем; однако этот подход весьма трудоемок и не всегда реализуем;
- американский ученый **М. Месарович**¹ предложил подходы, которые назвал *целенаправленным* и *терминальным* (от *терм* – элементарная частица, интересующая исследователя);
- польский ученый **Р. Куликовски**² предложил называть аналогичные подходы *декомпозицией* и *композицией* системы;
- швейцарский астроном, венгр по происхождению **Ф. Цвикки**³ предложил и развил *морфологический подход*, который помогает искать полезные объединения элементов путем их комбинаций;
- американская корпорация **RAND**⁴ предложила подход к созданию сложных программ и проектов, названный «*деревом целей*»;
- в практике проектирования сложных технических комплексов возникли термины *язык моделирования*, *язык автоматизации проектирования*, применяющиеся для отображения взаимосвязей между компонентами проекта; при разработке языков моделирования применяют *математическую логику* и *математическую лингвистику*, в которой есть удобный термин для описания структуры языка – *тезаурус* (см. гл. 5), и подход называют иногда *лингвистическим* или *тезаурным*;
- при исследовании и формировании структур были предложены следующие подходы: путем поиска связей между элементами; или, напротив, путем устранения лишних связей (см. в [1, 21]).

С учетом рассмотренных подходов в настоящее время на основе обобщения предшествующего опыта сформировалось два основных подхода к отображению систем, первоначально предложенных для формирования структур целей⁵.

¹ **Месарович М.** Общая теория систем и ее математические основы / М. Месарович // Исследования по общей теории систем: сб. переводов / Под ред. В. Н. Садовского и Э. Г. Юдина. М.: Прогресс, 1969. – 520 с.

² **Куликовски Р.** Оптимальные и адаптивные процессы в системе автоматического регулирования / Р. Куликовски. – М.: Наука, 1967.

³ **F.** Morphological astronomy. – Berlin: Springer-Verlag, 1957. – 299 p.

⁴ **Лопухин М.М.** ПАТТЕРН – метод планирования и прогнозирования научных работ / М.М. Лопухин. – М.: Сов. радио, 1971. 160 с.

⁵ **Теория** систем и методы системного анализа в управлении и связи / В.Н. Волкова, В.А. Воронков, А.А. Денисов и др. М.: Радио и связь, 1983. 248 с.

а) «сверху» – методы *структуризации* или *декомпозиции*, *целевой* или *целенаправленный* подход;

б) «снизу» – подход, который называют *морфологическим* (в широком смысле), *лингвистическим*, *тезаурусным*, *терминальным*, методом «языка» системы. С помощью этого подхода определяется «пространство состояний» системы и реализуется поиск взаимосвязей (мер близости) между элементами.

Подход «снизу» можно реализовать, применяя не только комбинаторные приемы (морфологический и т.п.), но и бехвиористский подход, вариант которого при автоматизации моделирования поведения объектов в настоящее время иногда называют *процессным*.

Подходы «сверху» и «снизу» называют также *аксиологическим* и *каузальным*, соответственно ¹.

Аксиологическое представление системы – отображение системы в терминах *целей* и целевых функционалов. Этот термин используют в тех случаях, когда необходимо выбрать подход к отображению системы на начальном этапе моделирования и противопоставить это отображение описанию системы в терминах «перечисления» элементов системы и их непосредственного влияния друг на друга, т.е. *каузального представления*.

Каузальное представление системы – описание системы в терминах влияния одних переменных на другие, без употребления понятий *цели* и *средств* достижения целей. Этот термин происходит от понятия «cause» – причина, т.е. подразумевает причинно-следственные отношения. При каузальном представлении будущее состояние системы определяется предыдущими состояниями и воздействиями среды. Такое представление является развитием отображения системы в виде «пространства состояний», характерного для большинства математических методов моделирования. Применяют каузальное представление в случае предварительного описания системы, когда *цель* сразу не может быть сформулирована, и для отображения системы или проблемной ситуации не может быть применено *аксиологическое представление*.

На практике обычно эти подходы сочетают

В 1970–1980-е гг. при проектировании организационных структур были предложены три подхода к решению этой проблемы ².

• *Нормативно-функциональный* подход направлен на унификацию организационных форм управления в рамках отрасли. Разработка и внедрение типовых организационных структур явилось первым шагом на пути внедрения принципов их научно обоснованного построения. Однако ориентация на типовую номенклатуру функций управления и структур-

¹ *Математика* и кибернетика в экономике: словарь-справочник / Отв. ред. Н.П. Федоренко – М.: Экономика, 1975. – 700 с.

² *Мильнер Б.З.* Системный подход к организации управления / Б.З. Мильнер, Л.И. Евенко, В.С. Рапопорт. – М.: Экономика, 1983. – 224 с.

ных управленческих подразделений не позволяет учесть особенности конкретных предприятий и условия их деятельности.

- *Функционально-технологический* подход основан на рационализации потоков информации и технологии ее обработки, на формировании и анализе организационно-технологических процедур подготовки и реализации управленческих решений. Этот подход обеспечивает возможность достаточно полно учесть особенности конкретного предприятия (организации), отличается гибкостью и универсальностью. Вместе с тем, он характеризуется высокой трудоемкостью, использованием стабильной номенклатуры сложившихся функций управления, подчинением оргструктуры схеме документооборота.

- *Системно-целевой* подход заключается в построении структуры целей, определении на ее основе функций управления и их организационном оформлении. Преимущества этого подхода заключаются в возможности учитывать особенности объекта управления и условия его деятельности, изменять и расширять состав функций, проектировать разнообразные организационно-правовые формы предприятий. Трудности в использовании подхода связаны с проблемой перехода от совокупности целей и функций к составу и подчиненности структурных звеньев, обеспечивающих их реализацию.

Применительно к исследованию и разработке ИС эти подходы можно интерпретировать следующим образом.

Обобщающий подход «сверху», называемый целевым, целенаправленным, системно-целевым, основан на структуризации или декомпозиции системы *в пространстве*. Этот подход позволяет расчленить исходную большую неопределенность на более обозримые и выбрать методы их анализа и проектирования, сохраняя целостность представления об исследуемой системе или решаемой проблеме на основе иерархической структуры (древовидной, стратифицированной). Подход применялся при разработке АИС и АСУ для крупных предприятий, при реструктуризации систем организационного управления.

Подход «снизу», основанный на анализе пространства состояний, поиске «мер близости» между компонентами с помощью различных, в том числе статистических методов, морфологического моделирования (подход «снизу») отличается большой трудоемкостью. В настоящее время для анализа пространства состояний в экономике разработаны методы бизнес-аналитики Data-Mining (см., напр., [17]), реализуемые с помощью соответствующих программных продуктов (Deductor и т. п.).

В настоящее время для проектировании информационных систем широкое применение нашел подход, основанный на анализе бизнес-процессов, кратко называемый *процессным*.

Процессный подход (который можно считать развитием *функционально-технологического подхода*) основан на структуризации *во времени*, на представлении процессов в форме графов.

При несомненной привлекательности применение функционально-технологического подхода долгое время было практически нереализуемым из-за большой трудоемкости, отсутствия правил и средств автоматизации формирования графов, отображающих процессы в системах.

В 1990-е годы была разработана методология SADT – (Structured Analysis and Design – структурный анализ и проектирование (предложена *Дугласом Россом*¹), представляющая собой совокупность методов, правил и процедур, предназначенных для построения функциональной модели объекта какой-либо предметной области. На ее основе разработаны и стали широко применяться функционально-ориентированные и объектно-ориентированные CASE- и RAD-технологии. Компьютерная реализация методологии SADT получила название IDEF (Icam Definition). Основными структурными моделями являются модели процессов IDEF0 и IDEF3, модель данных IDEF1X [26]. Созданы стандарты IDEF и DFD, ориентированные на анализ процессов (в том числе бизнес-процессов), что позволило создать теорию проектирования экономических информационных систем [22, 24]. Для реализации моделей применяются автоматизированные средства – BPWin, ARIS, язык UML (Unified Modeling Language – Унифицированный язык моделирования)².

Популярность CASE-методологии и RAD-технологий базируется на разработке принципов и автоматизации формирования процессов, на развитии методов их формирования (на основе анализа «жизненного цикла» производства, обслуживания или других процессов, причинно-следственных связей и т.п.), что и обеспечило развитие *процессного подхода*, преимущества которого заключаются в возможности учитывать особенности конкретного объекта и условий его деятельности.

На основе анализа методик проектирования документальных, фактографических и документально-фактографических информационных систем, приведенных в гл.гл. 4–6, с точки зрения рассмотренных подходов можно сделать следующие выводы.

1. Проектирование АИС и АСУ разных уровней в 1970-е гг. осуществлялась на основе *системно-целевого подхода*, т.е. на основе *структуризации целей и функций* системы управления (*целевой* или *целенаправ-*

¹ **Ross D.** Applications and extension of SADT // IEEE. Computer. – April, 1995.

² **Фаулер М.** UML в кратком изложении: Применение стандартного языка объектного моделирования / Пер. с англ./ М. Фаулер, К.Скотт. – М.: Мир, 1996. – 112 с.

ленный подход, подход «сверху»), формирования на ее основе структуры *функциональной части* АСУ (АИС) и определения средств для ее реализации – *обеспечивающей части*, т.е. информационного, технического, программного, организационного и других видов обеспечения.

Для управления разработками АСУ были подготовлены и изданы соответствующие общепромышленные руководящие методические материалы, в которых отражались теория и основная терминология, обязательная для использования при представлении отчетных материалов по разработке и внедрению АИС и АСУ.

2. При разработке ГАСНТИ также вначале был применен *системно-целевой подход*, разработана классификация АС НТИ для разных уровней управления – общегосударственном, отраслевом, региональном, предприятий и организаций, и соответствующие нормативные документы. Однако в настоящее время ГАСНТИ развивается в основном только на общегосударственном уровне.

3. Разработка документальных информационно-поисковых систем обычно проводится в виде следующих этапов: а) анализ совокупности документов, представляющих соответствующую научно-техническую область, для которой создается ИПС; б) выбор ключевых слов, характеризующих содержание каждого документа; в) формирование из ключевых слов словаря (первой стадии разработки информационно-поискового языка); г) индексирование документов с использованием словаря; д) поиск; е) проверка релевантности выдачи) отладка ИПС.

Постепенно ИПЯ развиваются, вначале на основе формирования дескрипторного словаря, затем – тезауруса, а при необходимости – и разработки грамматики. Существующие ДИПС (в том числе поисковых систем Интернет) редко развиты до уровня наличия в них многоуровневого тезауруса и грамматики. В них обычно используются простейшие правила грамматики – конъюнкция и дизъюнкция (названные упрощенно «AND», «OR» и т.п.), а термин «тезаурус» используется в упрощенном варианте, в лучшем случае – это двухуровневый словарь.

С использованием терминологии рассмотренных подходов, можно считать, что разработка ДИПС начинается с анализа «пространства состояния», т.е. квалифицировать подход как *терминальный, лингвистический, тезаурусный*, метод «языка» системы, в упрощенном названии – как подход «снизу».

По мере развития ДИПС при разработке систем, реализующих режимы избирательного распределения информации и дифференцированного обслуживания руководителей, применяют *системно-целевой подход*, т.е. начинают разработку с исследования потребностей научных подразделений и руководителей.

4. Для разработки экономических информационных систем создана теория проектирования, основанная на применении *процессного подхода*, анализе бизнес-процессов, представляемых специальными графами.

Идею применения графического представления и анализа информационных потоков (*функционально-технологический подход*) при проектировании информационных систем начинали применять в 1970-е гг. ¹. Был разработан язык моделирования информационных потоков, реализованы средствами языка логического программирования РЕФАЛ ². Однако его применение долгое время было практически нереализуемым из-за большой трудоемкости и отсутствия стандартных автоматизированных процедур (проведенные исследования с применением языка РЕФАЛ были экспериментальными).

Разработка CASE- и RAD-технологий, стандартов IDEFO и DFD, ориентированных на анализ процессов (в том числе бизнес-процессов и соответствующих им информационных потоков) и средств их автоматизации позволила создать широко применяющуюся в настоящее время теорию проектирования информационных систем, основанную на *процессном подходе* [22, 24 и др.].

5. При автоматизации управленческой деятельности предприятий малого и среднего бизнеса, которые, как правило, не имеют средств для внедрения единой автоматизированной системы, существует практика выбора имеющихся на рынке программных продуктов, автоматизирующих соответствующие области управления (бухгалтерский учет, материально-техническое обеспечение и т.п.) с последующей их интеграцией в единую систему. Для выбора этих программных продуктов разрабатывают соответствующие методики (пример приведен в разделе 4.6), основанные на применении *системно-целевого подхода*, т.е. разработке структуры целей и функций организации и выбора программных продуктов, функциональные возможности которых обеспечивают реализацию функций и задач организации.

¹ **Волкова В.Н.** Некоторые вопросы автоматизации проектирования АИС: автореф. дис. на соискание ученой степени канд. техн. наук по специальности 05.13.01 – Техническая кибернетика и теория информации / В.Н. Волкова. – М., 1973. – 24 с.; **Волкова В.Н.** К методике проектирования автоматизированных информационных систем / В.Н. Волкова // Автоматизация управления и вычислительная техника: кн. – Вып. 11. – М.: Машиностроение, 1975. – С. 189–300.

² **Волкова В.Н.** Автоматизация проектных работ при создании автоматизированных информационных систем / В.Н. Волкова // Типизация и автоматизация процессов проектирования АСУ: матер. Всесоюз. Семинара. – Душанбе: Изд-во Дониш, 17–19 апреля 1972. – С. 79–87; **Волкова В.Н.** Исследование возможностей сокращения перебора при семиотическом моделировании проектных задач АСУ / В.Н. Волкова, А.Г. Красовский // АСУ и синтез оптимальных систем: сб. тезис. докл. науч.-техн. конф. – М.: НТОРЭС, 1974. – С. 18–20.

8.4. Анализ определений системы и их применение для развития теории информационных систем

Существуют различные определения системы (см. обзоры в [1, 21, 25]).

В первых определениях система рассматривалась как совокупность только элементов a_i и связей r_j между ними ((Л. фон Берталанфи и др.):

$$S \stackrel{\text{def}}{=} \langle A, R \rangle, \text{ где } A = \{a_i\}, R = \{r_j\}$$

$$S \stackrel{\text{def}}{=} \langle \{a_i\}, \{r_j\} \rangle, \quad a_i \in A \quad r_j \in R \quad (8.1)$$

$$S \stackrel{\text{def}}{=} \langle \{a_i\} \& \{r_j\} \rangle, \quad \{a_i\} \& \{r_j\}$$

Затем – с учетом их свойств Q (А. Холл, А.И. Уёмов и др.):

$$S \stackrel{\text{def}}{=} \langle A, R, Q \rangle. \quad (8.2)$$

Затем в определения стал вводиться *системообразующий критерий* (М.Г. Гаазе-Рапопорт), а позднее – и *цель* Z в явном виде:

$$S \stackrel{\text{def}}{=} \langle A, R, Z \rangle, \quad (8.3)$$

где Z – цель, совокупность или структура целей.

В определении **В. Н. Сагатовского** уточняются условия целеобразования – *среда* SR , *интервал времени* ΔT , т.е. период, в рамках которого будет существовать система и ее цели: система – «конечное множество функциональных элементов и отношений между ними, выделенное из среды в соответствии с определенной целью в рамках определенного временного интервала»¹:

$$S \stackrel{\text{def}}{=} \langle A, R, Z, SR, \Delta T \rangle. \quad (8.3a)$$

При этом понятие цели использовалось в перечислении наряду с элементами, связями, свойствами, т.е. рассмотрение системы все же начиналось с элементов и связей, а не с формулирования целей (цель предполагалась заданной).

В последующем в определении предлагается учитывать наблюдателя N :

$$S \stackrel{\text{def}}{=} \langle A, R, Z, N \rangle. \quad (8.4)$$

На необходимость учета взаимодействия между изучаемой системой и исследователем или проектировщиком первоначально указал

¹ **Основы** системного подхода и их применение при разработке территориальных АСУ / Под ред. Ф. И. Перегудова. – Томск: ТГУ, 1976. – с. 13–14.

У.Р. Эшби¹ и в качестве обобщающего названия лиц, исследующих или проектирующих систему, ввел термин «наблюдатель».

Первое определение, в которое в явном виде включен наблюдатель, дал **Ю. И. Черняк**: «Система есть отражение в сознании субъекта (исследователя, наблюдателя) свойств объектов и их отношений в решении задачи исследования, познания»²:

$$S \underset{\text{def}}{\equiv} \langle A, Q_A, R, Z, N \rangle. \quad (8.4, a)$$

Ю.И. Черняк стал учитывать и язык наблюдателя L_N : «Система – отображение на языке наблюдателя объектов, отношений и их свойств в решении задачи исследования, познания»:

$$S \underset{\text{def}}{\equiv} \langle A, Q_A, R, Z, N, L_N \rangle. \quad (8.4, б)$$

Существовали и определения, предлагаемые для информационно-поисковых систем научно-технической информации.

Первоначально определения ИПС отображали процесс информационного поиска в множестве документов D ; с помощью множества информационных запросов Q и отношений R , при наличии которых любому запросу $q_i \in Q$ ставится в соответствие подмножество D' ответов на информационный запрос [15, 27].

$$IPS \underset{\text{def}}{\equiv} \langle D, Q, R, D' \rangle, \quad (8.5)$$

$D' \subset D$

Затем **А.И. Черный** [27] дал определение, отображающее принципы построения и функционирования ИПС

$$IPS \underset{\text{def}}{\equiv} \langle LS, D, TS, N \rangle, \quad (8.6)$$

$$LS \underset{\text{def}}{\equiv} \langle RL, IND, KSS \rangle. \quad (8.7)$$

где LS – *логико-семантический аппарат* (т.е. информационно-поисковый язык, правила индексирования и критерии выдачи);

D – *поисковый массив* (т.е. определенное множество снабженных поисковыми образами документов, в котором отыскиваются необходимые);

TS – *технические средства* (т.е. приспособления или устройства, которые необходимы для записи и хранения поисковых образов, для хранения документов и осуществления процесса сопоставления поиско-

¹ **Эшби У.Р.** Введение в кибернетику / У.Р. Эшби. – М.: Ин. лит., 1959. – 432 с.

² **Черняк Ю.И.** Системный анализ в управлении экономикой / Ю.И. Черняк. – М.: Экономика, 1975. – 191 с.

вых образов документов с поисковым предписанием или поисковым образом запроса);

N – люди, взаимодействующие с системой (т.е. те, кто пользуются данной ИПС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска, а также выполняют другие интеллектуальные операции, без которых невозможен информационный поиск).

В последующем **Ю.И. Шемакин** [30] дал определение информационно-семантической системы. Развивая представления об информационных системах, способных получать новую информацию, **Ю.И. Шемакин** [30, с. 60] предлагает понятие *информационно-семантической системы*, в котором учитываются понятия цели, структуры, технологических процессов, методов, средств и условий поиска:

$$ISS \underset{def}{=} \langle a, St, tp_{iss}, co, t_i \rangle, \quad (8.8)$$

где *a* – цель; *St* – структура; $tp_{iss} \in TP$ – подмножество технологических процессов для данной *ISS*; *co* – условия; *t_i* – время.

Входящие в определение (8.8) составляющие могут быть детализованы с учетом конкретной реализации ИПС. Особенно важно уточнять состав технологических процессов:

$$tp_{iss} \underset{def}{=} \langle met, re, SemSI \rangle,$$

где *met* – методы; *re* – средства; *SemSI* – семантическая переработка семантической информации.

Последнее определение (8.8) позволяет лучше учесть конкретные условия и особенности средств реализации при проектировании современных ИПС.

Для организации проектирования информационных систем **Ю.Ф. Тельнов** [22] предлагает определение, в котором учитываются и цели, и элементы, разделяя их на внешние и внутренние, и отношения, и функции (процессы, операции), и период времени *T*, и закономерности *Z*, определяющие структуру системы и ее взаимодействие с внешней средой:

$$S \underset{def}{=} \langle G, E, E_n, T, F, R, Z \rangle. \quad (8.9)$$

где *G* – цели; *E_n* – внешние элементы; *E*, – внутренние элементы; *R* – отношения, включая динамические взаимодействия; *F* – функции (процессы, операции); *T* – период существования системы; *Z*, – закономерности, определяющие структуру системы и ее взаимодействие с внешней средой.

Анализ определений системы показывает, что первые определения опирались на подход к исследованию и проектированию системы, бази-

рующийся на отображении пространства состояний (элементов, связей, их свойств) и поиске мер близости на этом пространстве. Этот подход в теории систем **М. Месарович** называет терминальным, **Ю.И. Черняк** – лингвистическим или методом «языка» системы; для краткости в теории систем принят упрощенный термин – подход к исследованию или проектированию системы от элементов, т. е. как бы «снизу».

Наиболее предпочтительным при создании информационных систем для социально-экономических объектов представляется аксиологический подход – от целей, потребностей (т. е. «сверху»).

По аналогии с определением, которое было дано в [30] для информационно-семантических систем, предложено более общее определение, в котором реализуется системно-целевой подход¹:

$$S \underset{def}{\equiv} \langle Z, STR, TECH, COND, N \rangle, \quad (8.10)$$

где $Z = \{z\}$ – совокупность или структура целей; $STR = \{STR_{пр}, STR_{орг}, STR_{ИТ}, \dots\}$ – совокупность структур, реализующих цели ($STR_{пр}$ – производственная, $STR_{орг}$ – организационная, $STR_{ИТ}$ – ИТ-структура т. п.); $TECH = \{meth, means, alg, \dots\}$ – совокупность технологий (методы – *meth*, средства – *means*, алгоритмы – *alg* и т. п.), реализующих систему; $COND = \{\varphi_{ex}, \varphi_{in}\}$ – условия существования системы, т. е. факторы, влияющие на ее создание и функционирование (φ_{ex} – внешние, φ_{in} – внутренние); N – «наблюдатели» (по **У.Р. Эшби** и **Ю.И. Ченняку**), т. е. лица, принимающие и исполняющие решения, осуществляющие структуризацию целей, корректировку организационной и производственной структуры, осуществляющие выбор методов и средств моделирования и т. п.

Обобщенное определение ИПС. С учетом анализа развития определений системы и более подробного анализа определений информационно-поисковых систем, приведенных в гл. 5, можно дать определение ИС, в котором учитывались бы и *элементы* разного рода, и *цели* (как правило, многоуровневая структура функциональной части АСУ или структура информационных потребностей), и *структуры* (структура информационного, программного и др. видов обеспечения в АСУ, структура ИПС), и *отношения* между компонентами, и *технологии* в широком смысле (методы, технические средства, алгоритмы, информационные технологии обработки информации), и *среда* (внешние и внутренние факторы), и *временной интервал*, и *логико-семантический аппарат*, и *люди*, взаимодействующих с системой (проектировщики,

¹ **Волкова В.Н.** Развитие определения системы / В.Н. Волкова // Матер. Международной научно-практической конференции «Системный анализ в проектировании и управлении»: сб. – СПб.: Изд-во СПбГПУ, 2001. – С. 12–14.

пользователи и обслуживающий персонал): При этом с учетом большого числа факторов, можно вначале включить в систему укрупненные множества, а затем раскрыть их:

$$S \underset{def}{=} \langle Z, STR, LS, TECH, COND, SR, \Delta T, N \rangle, \quad (8.11)$$

где Z – цели, которые в АИС могут интерпретироваться как структура функциональной части АСУ, а в документальных и документально-фактографических ИС – как потребности;

STR – структуры информационных массивов, т.е. баз данных в фактографических системах; поисковых массивов документов D (определенное множество снабженных поисковыми образами документов) в документальных ИПС, хранилищ информации разного рода (выбор термина диктуется объемами информационных массивов и конкретными условиями);

$LS \equiv \langle RL, IND, KSS \rangle$ – логико-семантический аппарат, включающий информационно-поисковые языки RL , систему индексирования IND и критерии выдачи (или критерии смыслового соответствия KSS);

$TECH$ – технологии в широком смысле, включая TS – технические средства (т.е. приспособления или устройства, которые необходимы для сбора, регистрации, хранения, обработки и представления информации), $meth$ – методы сбора, хранения, обработки информации, включая алгоритмы alg , программные процедуры или пакеты прикладных программ PPP , информационные технологии IT и т.п.;

$COND$ – условия, т.е. внешние φ_{ext} и внутренние φ_{int} факторы, влияющие на создание и функционирование ИС; для их анализа полезно использовать признак «пространство инициирования целей», т.е. выявлять факторы надсистемы и актуальной среды (φ_{ext}), подведомственной и собственно системы (φ_{int});

ΔT – временной интервал создания и функционирования («жизни») информационной системы,

N – люди, взаимодействующие с системой, т.е. те, кто заказывают ИС, проектируют ИС, пользуются данной ИС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска, а также выполняют другие интеллектуальные операции, без которых невозможен информационный поиск.

Для обобщающего названия всех, кто имеет отношение к разработке и использованию систем, **У.Р. Эшби** в начале развития кибернетики и теории систем ввел термин «наблюдатель»; а в настоящее время часто используют термин «стейкхолдеры». Буквальный перевод термина «stakeholders» – «держатели ставок», т.е. игроки в тотализатор на ипподроме. При переводе на русский язык предлагались варианты «акционеры», «заинтересованные стороны». Но каждый

из них неточен: «акционеры» – неполный состав лиц, взаимодействующих с системой, в их числе могут быть и «незаинтересованные». Поэтому был охранен термин без перевода, и его часто применяют при исследовании или проектировании именно информационных систем.

Определение (8.11) может интерпретироваться с учетом назначения, вида информационной системы, условий ее разработки.

Запись определений в формализованном виде помогает сохранять целостное представление о концепции, положенной в основу определения.

Компоненты, входящие в определения, конечно, можно трактовать по-разному. Можно и для элементарной базы данных назвать цель или назначение ее создания, перечислить запросы, отвечающие какой-то совокупности потребностей, определить структуру БД, методы и средства ее реализации, условия использования.

Однако важно понимать, что, разрабатывая информационную систему для предприятий, организаций, следует исходить из их целей, т. е. сформировать структуру ФЧ АИС на основе анализа целей и функций организации и определения наиболее значимых подсистем для автоматизации. А, следовательно, нужно начинать с формулировки концепции проекта системы, формирования определения системы, соответствующего этой концепции, анализа целей и роли всех видов информационных ресурсов для реализации целей, а уж затем выбирать технологии (методы, алгоритмы, средства, в том числе готовые программные продукты) с учетом их соответствия поставленным целям.

Приведенное определение помогает теоретически обосновать концепцию многоуровневой информационной системы, рассматриваемую в следующем параграфе.

8.5. Концепция многоуровневой информационной системы

По мере развития представлений об автоматизации информационного обеспечения любых организаций становится очевидным, что для функционирования и управления организацией нужен широкий спектр информационного обеспечения, включая разнородную информацию от фактографической до документальной

Основной особенностью автоматизации информационного обеспечения процессов принятия решений является необходимость интеграции между собой большого количества независимых друг от друга и территориально распределенных информационных массивов.

В этих условиях задача информационного обеспечения процессов принятия решения представляет собой задачу формирования алгоритмов вызова информационных массивов, хранящихся и актуализируе-

мых в различных распределенных и независимых базах данных и информационных систем.

Решение задачи в такой постановке потребовало разработки новой концепции строения и функционирования автоматизированной информационной системы.

Для регулирования распределения финансовых и других ресурсов на разработку и поддержку функционирования разнородных баз данных и информационных систем и алгоритмов сбора информации для обеспечения потребностей лиц, принимающих решения, или других пользователей, необходимо объединить независимые базы данных и алгоритмы в единую систему.

Появилось понятие интегрированной информационной системы. Предлагались различные способы представления структуры таких систем.

В связи с осознанием огромного влияния информационных ресурсов на социальное и экономическое развитие всего мира, отдельных стран и регионов появилось понятие *информационной инфраструктуры* (ИИ). Термин был введен в зарубежных и отечественных публикациях в конце 1990-х гг.¹

Термин «*инфраструктура*» (лат. *infra* – под, внизу, *structure* – строение, постройка) появился вначале в строительном деле и означал основание, фундамент, нижнее строение.

В дальнейшем понятие было расширено, и в переносном смысле стало означать совокупность внешних по отношению к рассматриваемому объекту сооружений, в том числе не только в строительстве, но и по отношению к производству, к организациям непромышленной сферы.

В нашей стране на необходимость введения понятие *информационной инфраструктуры* (ИИ) впервые было обращено внимание в статье **Ф. Широкова** и **В. Дрожжинова**².

Попытки разработки концепции информационной инфраструктуры организации на примере вуза были инициированы в 1990-е гг. студентами **А.А. Ионовым** и **И.В. Релиной** – членами Молодежного академического общества при Научном совете по проблемам высшей школы, подготовки и аттестации научно-педагогических кадров Санкт-Петербургского Отделения Международной академии наук высшей школы³.

¹ **Clinton W.J., Gore A.** Technology for America's Economic Growth, a New Direction to Build//Executive Office of the President. Washington, DC, 1993, 39 pp.

² **Широков Ф.** Три ступени Альберта Гора / Ф. Широков, В. Дрожжинов // Компьютер-Пресс, 1994. – № 9. – С.43–46. – № 10. – С. 87–95.

³ **Волкова В.Н.** Информационная инфраструктура как средство управления учебно-научным процессом / В.Н. Волкова, А.А. Ионов, И.В. Релина // Сб. тезисов докладов III Международной научно-технической. конференции: Высокие интеллектуальные технологии образования и науки. – СПб.: СПбГТУ, 1996. – С. 185–186

Для управления развитием ИИ ее предлагалось условно представить в виде 4-х страт (рис. 8.2): *пользовательской, функциональной, информационной, коммуникационной* (информационная супермагистраль).

<p><i>Страта 1. Пользовательская</i> Собственно пользователи, включая формирование их потребностей в информации и правила взаимодействия с информационной системой</p>
<p><i>Страта 2. Функциональный слой</i> Совершенство услуг, предоставляемых пользователям различными поставщиками информации: доступ к базам данных и библиотекам, стандартам и т. д.</p>
<p><i>Страта 3. Информационный слой</i> Информация в телекоммуникационных сетях, базах данных и информационных хранилищах библиотек, отделов научно-технической информации и т. п. Информационные ресурсы: справочники, сайты, статистическая информация, нормативно-правовая информация и пр.</p>
<p><i>Страта 4. Информационная супермагистраль</i> Информационные телекоммуникации. Информационные технические средства</p>

Рис. 8.2.

Для обоснования концепции создания интегрированной информационной системы современного вуза были проанализированы определения систем и подходы к их разработке, предложенные в теории систем.

Определение (8.11), отражающее системно-целевой подход к проектированию систем, может интерпретироваться с учетом назначения, вида информационной системы, условий ее разработки.

Например, стратифицированную структуру, приведенную на рис. 8.3, учитывая сложившиеся концепции и термины в теории разработки АСУ и систем научно-технической информации, можно обосновать следующим образом.

Цели Z в определении (8.11) реализуются в форме *функциональной страты*.

Назначение *функциональной* страты (страта 2) – обеспечить удобный доступ к информационным хранилищам – *информационной* страте (страта 3).

Для этого нужно структурировать направления деятельности, цели Z организации, определить взаимосвязи между направлениями деятельности, целями (подцелями) и составляющими информационной страты, как показано на рис. 8.3.

Одной из форм реализации доступа к информации является структура функциональной части АСУ, предназначенной для обеспечения информацией сферы организационного управления.

В библиотеках подобную функцию выполняет предметный указатель.



Рис. 8.3. Многоуровневая структура информационной системы

Информационная страта реализует структуру информационного обеспечения – *STR* в определении (8.11). Она объединяет хранилища информации разного рода, создаваемые в организации в разных формах и на различных носителях.

Нижняя страта, названная *коммуникационной*, реализует компоненту *TECH*, т.е. включает техническое, алгоритмической, программное обеспечение информационной системы, в том числе взаимодействие в сети Интранет, если она создана в организации, и взаимодействия между локальными сетями или отдельными компьютерами, технические средства сбора, регистрации, хранения и обработки информации.

Компоненту *N* отражает *пользовательская* страта. Эта страта должна обеспечить доступ к информации не только руководителям, но и всем сотрудникам организации. Для ее реализации необходимо ставить задачи определения информационных потребностей пользователей.

В системах научно-технической информации ставится задача определения информационных потребностей пользователей и создания систем избирательно-го распределения информации (ИРИ) и дифференцированного обслуживания руководителей (ДОР).

В АСУ для обеспечения фактографической информацией имеются работы по созданию автоматизированных рабочих мест (АРМ) для специалистов соответствующей квалификации, должностных лиц.

Для реализации пользовательской страты необходимо установить взаимосвязи соответствующих групп пользователей.

В научно-исследовательской организации – это научные сотрудники, руководители проектов, руководители организации и т.п., на производственных предприятиях – руководители отделов организационной структуры, в медицинском учреждении – врачи, медицинские сестры, администрация; в вузе – это, например, преподаватели, студенты, сотрудники лабораторий и организационных подразделений аппарата управления; и т.п.

Необходимо определить взаимосвязи компонентов пользовательской страты с существующими направлениями и структурами функциональной страты, обеспечить доступ к этим структурам, в том числе в форме рекламы, информационных писем (например, о конференциях, проводимых мероприятиях, формах доступа к информационным источникам, к средствам издания и размножения информации и т.п.).

В идеале желательно создать соответствующую информационно-поисковую систему для информирования пользователей о существующих функциональных направлениях и хранилищах по запросам.

Пользовательская страта в организациях, как правило, пока не создана, даже для руководящих работников организации. Она складывается в процессе опыта принятия решений.

Стратифицированное представление, приведенное на рис. 8.3, помогает уточнить назначение и взаимодействия различных аспектов реализации компонентов информационной системы, входящих в определение системы (8.11).

В связи с неоднозначным использованием в публикациях термина «инфраструктура», более целесообразно трактовать структуру рис. 8.3 как многоуровневую (стратифицированную) интегрированную информационную систему (МИИС).

Концепция стратифицированной информационной системы в конкретных условиях требует уточнения и развития.

Например, если учесть другие составляющие компоненты технологии *TECH*, входящей в определение (8.1), то между функциональной и информационной, и между пользовательской и функциональной стратами возможно введение дополнительных страт, содержащих средства (алгоритмы, программные продукты) для обеспечения возможности более полного исследования взаимосвязей между компонентами этих страт.

Если учесть логико-семантический аппарат *LS*, то практическая реализация рассмотренной концепции осуществляется путем создания информационно-поисковых систем, работающих в режимах избирательного распределения информации и ретроспективного поиска по произвольным запросам, на базе ресурсов информационной страты, сети Интернет и локальных вычислительных сетей.

Создание МИИС требует разработки и применения соответствующих методов и автоматизированных процедур для определения состава компонентов каждой из страт, оценки влияния того или иного проекта в сфере МИИС на развитие организации.

Стратифицированное представление помогает решать проблему управления проектами и программами развития ИИ на основе их оценки с целью распределения финансовых, материальных и кадровых ресурсов, для чего разрабатывается методика, учитывающая степень влияния проектов на цели организации.

Для исследования взаимосвязей между стратами необходимо использовать методы и модели системного анализа: методики структуризации целей и функций, методы организации сложных экспертиз.

В частности, для анализа и формирования *функциональной страты* целесообразно использовать методики структуризации целей и функций, методы анализа информационных потребностей, разрабатываемые в теории информационного поиска. Оценку значимости элементов каждой из страт можно проводить с применением методов организации сложных экспертиз [1, 3, 21] (метода парных сравнений в модификации *T. Saatu*, метода решающих мат-

риц *Г.С. Поспелова*, информационного подхода *А.А. Денисова*), учитывать разнородные критерии – технические, экономические, социальные.

Пример моделей для исследования взаимосвязей между стратами многоуровневой информационной системы приведен в Приложении 2.

Таким образом, предлагаемая концепция развития интегрированной информационной системы основана на стратифицированном представлении ее структуры, в условиях использования территориально распределенных и взаимно независимых информационных баз и временного объединения их на основе разработки заранее исследованных и формализованных алгоритмов и программных процедур, обеспечивающих подбор необходимой информации для удовлетворения запросов потребителей.

При реализации предлагаемой концепции информационная система представляет собой систему организационных структур, обеспечивающих функционирование и развитие информационного пространства вуза и средств информационного взаимодействия, т.е. включает совокупность информационных центров, банков данных и знаний, систем связи; обеспечивает доступ потребителей к информационным ресурсам на основе разработанных алгоритмов и нормативного обеспечения, регламентирующего доступ к информации.

Автоматизация информационного обеспечения процессов принятия управленческих решений и информационных потребностей пользователей при обращении ЛПР или другого лица к информационной системе должно инициировать процесс, с помощью которого обеспечивается поиск необходимой пользователю информации на основе использования территориально распределенных информационных массивы.

Интересно отметить, что многоуровневой структуры информационной системы, полученной на основе определения системы, могут быть поставлены в соответствие приведенные на рис. П1.9 компоненты, определяемые концепцией архитектуры предприятия типа TOGAF. Если эти компоненты поместить друг под другом – от компоненты «Пользователи» до компоненты «Технологическая архитектура», то получится многоуровневая структура, приведенная на рис. 8.4.

В этой структуре функциональная страта формируется на основе анализа бизнес-процессов, а между функциональной стратой (архитектурой бизнеса) и информационной стратой (архитектурой данных) помещена страта программных приложений, которую можно интерпретировать как определенный логико-семантический аппарат *LS*.

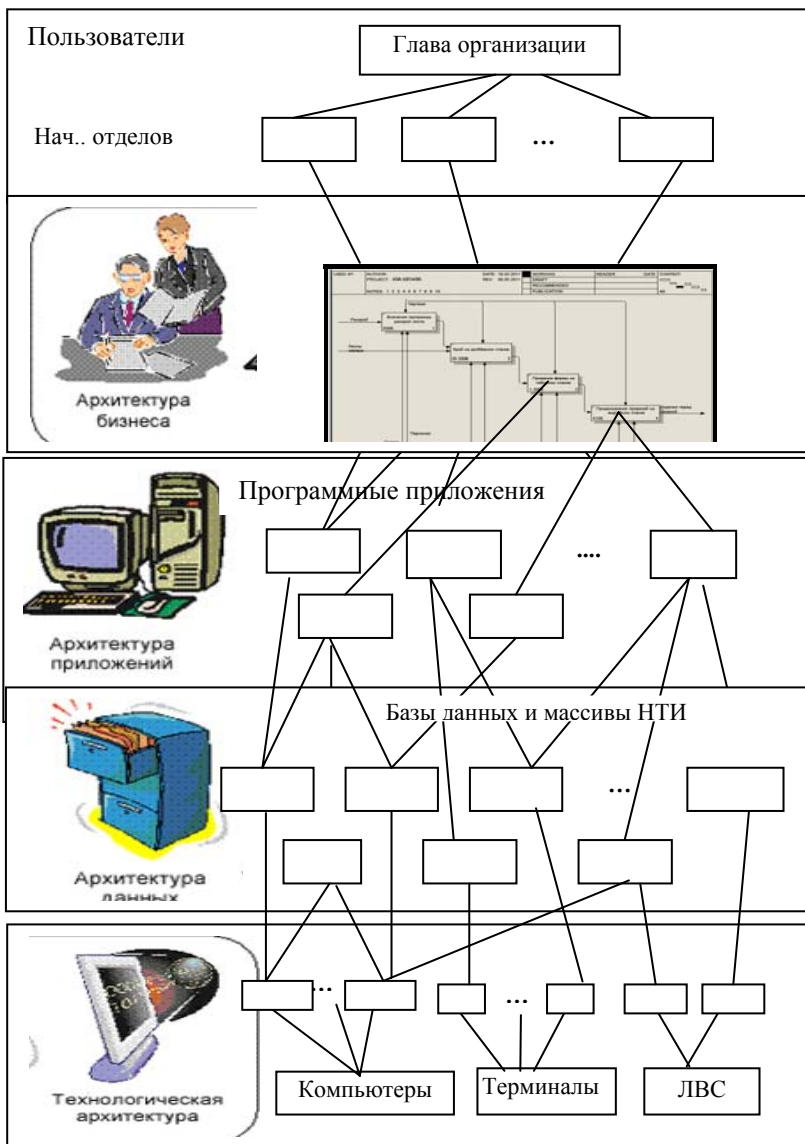


Рис. 8.4. Многоуровневая структура ИС, основанная на идее TOGAF

В рассмотренных примерах реализации концепции системы, лежащей в основе определения системы (8.11), не использовалась пока компонента *COND* – условия функционирования системы с учетом внутренних факторов φ_{int} и внешних φ_{ext} факторов среды, в которой функционирует информационная система.

В то же время эту компоненту необходимо учитывать как на этапе проектирования, так и в процессе функционирования ИС, что и подсказывает определение (8.11).

В частности, в процессе функционирования ИС при поиске информации в ответ на запросы пользователей необходимо учитывать, что информация может содержаться не только в массивах собственно ИС, но и во внешних базах данных, содержащихся, например, в сети Internet, локальных вычислительных сетях, в источниках научно-технической информации, доступных по межбиблиотечному абонементу (МБА) и т.п.

При углубленной детализации страт могут быть использованы методики структуризации, базирующиеся на других концепциях системы [1, 21, 25 и др.], в том числе рассмотренных в гл. 4.

ОСНОВНЫЕ КОНЦЕПЦИИ АРХИТЕКТУРЫ ПРЕДПРИЯТИЯ

П1. Архитектура Д. Захмана и ее модификации

П1.1. Исходная структура Д. Захмана для архитектуры предприятий

Понятие «архитектура предприятия» предложил в 1987 г. *Дж. А. Захман*², который связал это понятие с необходимостью для обеспечения высокой ценности и гибкости бизнеса целостного подхода, рассмотрения каждой проблемы с разных точек зрения. Такой подход к созданию архитектуры систем представляет собой то, что Захман изначально называл *архитектурной структурой* информационных систем, а впоследствии – *структурой архитектуры предприятия*.

Захман изначально объяснил ИТ-таксономию на примере строительной отрасли.

Структура Д. Захмана для строительной отрасли представляет собой таксономию для упорядочения архитектурных артефактов (артефакт – конкретный документ, отчет, аналитический отчет, модель или любой другой компонент архитектурного описания)), в которой учитываются лица, которым адресован артефакт (например, владелец бизнеса и строитель), и конкретная проблема (например, проблема с данными и функциональностью), которую необходимо устранить.

В строительной отрасли архитектурные артефакты неявным образом организованы в двумерную структуру. Одним измерением являются различные «игроки». В случае здания такими игроками являются владелец (тот, кто оплачивает проект), строитель (тот, кто координирует процесс постройки) и специалист по планированию (тот, кто обеспечивает соблюдение строительных норм и правил).

Архитектор здания подготавливает для каждого из «игроков» различные артефакты. Каждому «игроку» необходима полная информация, однако понятие полноты для каждого «игрока» свое. Владелец заинтересован в полном описании функциональности и эстетики здания. Строитель заинтересован в полном описании строительных материалов и процесса постройки. Владельца не интересуют гвозди в стенах. Строителя не интересует, виден ли из окон спальни восход солнца.

В исходной статье Захман говорит следующее:

«...каждое архитектурное представление отличается от других по существу, а не только уровнем детализации»².

¹ Подготовлено студентами М.А. Пашоликовым и М.В. Столяровой в 2011 г.

² *Захман Дж. А.* Структура архитектуры информационных систем / Дж. А. Захман // IBM Systems Journal. – 1987. – Том 26. – № 3.

В первой статье и последующей работе в 1992 г.¹ Захман предложил шесть описательных аспектов (данные, функция, сеть, люди, время и мотивация) и шесть «игроков» (планировщик, владелец, проектировщик, строитель, субподрядчик и предприятие). Эти два измерения можно представить в виде таблицы, как показано на рис. П1.1.

Первая строка соответствует уровню планирования бизнеса в целом (**бизнес-модель**). Отражает интересы высшего руководства. На этом уровне вводятся достаточно общие основные понятия, определяющие бизнес – продукты и услуги, клиенты, расположение объектов бизнеса, а также формулируется бизнес-стратегия.

Вторая строка (**концептуальная модель**) предназначена для определения в терминах бизнеса структуры организации, ключевых и вспомогательных бизнес-процессов.

Третий уровень (**логическая модель**). Здесь бизнес-процессы описываются в терминах информационных систем, включая различные типы данных, правила их преобразования и обработки для выполнения определенных бизнес-функций.

На четвертом уровне – **технологической или физической модели** – осуществляется выбор реляционной СУБД, или средств работы с неструктурированными данными.

Пятый уровень соответствует **детальной реализации** системы, включая конкретные модели оборудования, топологию сети, средства разработки, готовый программный код.

Шестой уровень **описывает работающую систему** (инструкции для работы с системой, фактические базы данных).

Вторым измерением классификации архитектурных артефактов являются описательные аспекты: *кто, что, где, когда, как и почему*.

Второе измерение не зависит от первого. И строитель, и владелец должны знать *что-то*, но это *что-то* для владельца отличается от *что-то* для строителя. Конкретное *что-то* зависит от того, кто задает вопрос.

С точки зрения владельца бизнеса, «данные» – это бизнес-объекты. Эти данные могут включать сведения как о самих объектах, например о клиентах и продукции, так и об отношениях между этими объектами, например о демографических группах клиентов и складских запасах. В разговоре о данных с владельцем бизнеса следует использовать именно этот язык.

¹ *Захман Дж. А.* Расширение и формализация структуры для архитектуры информационных систем / Дж. А. Захман, Дж. Ф. Сова // IBM Systems Journal, 1992. – Том 31. – № 3.

С точки зрения разработчика базы данных, «данные» – это не бизнес-объекты, а строки и столбцы, объединенные в таблицы и связанные друг с другом с помощью математических операций соединения и проекции. В разговоре о данных с разработчиком баз данных следует говорить не о демографических группах клиентов, а о реляционных таблицах в третьей нормальной форме.

Ни одна из этих точек зрения не является более предпочтительной или детализированной. Все они критически важны для целостного понимания архитектуры системы. Захман говорит:

*«Мы испытываем трудности при обсуждении друг с другом архитектуры информационных систем, поскольку вместо единой архитектуры используется набор архитектурных представлений. При таком положении дел неправы оба участника обсуждения. Архитектуры различны. Они взаимно дополняют друг друга. Существуют веские причины выделить ресурсы на разработку каждого архитектурного представления. Если какие-либо архитектурные представления не разработаны, организация подвергается риску»*¹.

Д. Захман так описал свою работу: *«Структура [архитектуры предприятия] по отношению к предприятиям представляет собой просто логическую структуру для классификации и упорядочения описательных представлений предприятия, существенно важных для управления предприятием, а также для разработки корпоративных систем»*².

Многие сторонники структуры Захмана рассматривают ее как междисциплинарную, распространяющуюся далеко за пределы ИТ.

Например, в одной популярной книге, посвященной методологии Захмана, говорится следующее: *«...рано или поздно вы обнаружите, что структура Захмана присутствует во всем, чем вы занимаетесь, а не только в ИТ-проектах. Когда вы поймете структуру Захмана, вы станете более эффективным во всем. Во всем, не больше и не меньше. Это утверждение абсолютно серьезно»*³.

Сам Джон Захман сказал в своем интервью следующее:

*«...схема структуры использовалась несколько тысяч лет, и я уверен, будет использоваться еще несколько тысяч лет. Меняется только наше понимание структуры и принципы ее использования в проектировании предприятий и производстве»*⁴.

¹ Захман Дж. А. Структура архитектуры информационных систем / Дж. А. Захман // IBM Systems Journal, 1987. – Том 26. – № 3.

² Захман Дж. А. Расширение и формализация структуры для архитектуры информационных систем / Дж. А. Захман, Дж. Ф. Сова // IBM Systems Journal, 1992. – Том 31. – № 3.

³ Архитектура предприятия на основе структуры Захмана / Кэрл О'Рурк, Нил Фишман и Уоррен Селкоу // Технология курса. – Бостон, штат Массачусетс, 2003.

⁴ Интервью с Джоном Захманом / Роджер Сешнс, главный редактор журнала Перспективы Международной ассоциации архитекторов программного обеспечения. – Вып. 6.

Как было показано, структура Захмана состоит из шести функциональных аспектов, каждый из которых рассматривается с точки зрения основного «игрока». Структура Захмана в ее современном виде представлена на рис. П1.1.

Как видно из рис. П1.1, таблица Захмана состоит из 36 ячеек – по одной для каждого сочетания точки зрения «игрока» (например, владельца бизнеса) и описательного аспекта (например, данных). При перемещении по таблице по горизонтали (например, слева направо) мы получаем различные описания системы – с точки зрения одного и того же «игрока». При перемещении по вертикали (например, сверху вниз) мы рассматриваем один аспект, но изменяем «игрока», с точки зрения которого рассматривается этот аспект.

Таблица Захмана может помочь организации в разработке архитектуры тремя способами.

Во-первых, в таксономии Захмана каждый архитектурный артефакт должен находиться в одной и только в одной ячейке. Местоположение конкретного артефакта не должно быть неопределенным. Если неясно, в какой ячейке должен находиться артефакт, скорее всего, проблема заключается в самом артефакте. По мере накопления артефактов в процессе разработки можно воспользоваться таблицей Захмана, чтобы определить назначение каждого артефакта.

Во-вторых, в таксономии Захмана архитектура считается полной только в том случае, если заполнены все ячейки. Ячейка считается заполненной, если в ней находятся артефакты, полностью определяющие систему для конкретного «игрока» в конкретном описательном аспекте.

Если все ячейки заполнены артефактами, это дает достаточно информации для полного описания системы с точки зрения каждого «игрока» (современный термин – заинтересованные лица, стейкхолдеры) и под любым возможным углом (в любом описательном аспекте).

В-третьих, в таблице Захмана ячейки в столбцах должны быть связаны друг с другом.

Хотя руководство организации представляет данные иначе, чем администратор баз данных, между этими двумя точками зрения должна быть связь. Кто-то должен изучить бизнес-требования руководства и продемонстрировать, что структура базы данных соответствует этим требованиям.

Итак, существует пять способов использования таблицы Захмана при разработке архитектуры организации. Таблица Захмана позволяет:

- Убедиться в том, что точка зрения каждого заинтересованного лица была рассмотрена в каждом описательном аспекте.
- Улучшить артефакты архитектуры путем точной подгонки описательных аспектов под конкретную аудиторию.
- Убедиться в том, что все бизнес-требования руководства сводятся к технической реализации.

- Убедить руководство в том, что технические специалисты не планируют реализовывать бесполезные функции.

- Убедить руководство в том, что сотрудники бизнес-подразделения учитывают при планировании интересы ИТ-специалистов.

Таким образом, модель предприятия представляется в виде набора согласованных описаний, которые соотносятся с ячейками формализованной матрицы, в которой учтены существенные для архитектуры аспекты предприятия. Заполнение матрицы происходит сверху вниз.

Перспективы (строки в таблице) соответствуют различному уровню управления предприятием, если речь идет об архитектуре предприятия. По сути, на каждом из этих уровней рассматриваются одни и те же категории вопросов, только с различным уровнем абстракции и детализации. Представления соотносятся с категориями специалистов, например, менеджер бизнес-процесса, разработчик и субподрядчик). По столбцам матрицы разнесены основные аспекты деятельности (объекты – «что», действия – «как», местоположения – «где», люди – «кто», время – «когда» и мотивы – «почему»). Формализованная структура этой матрицы приведена на рис. П1.2.

Такая модель описания в целом полезна для идентификации возможных ограничений (например, указание руководства компании о выборе тех или иных средств работы).

Недостаток модели: отсутствие рассмотрения системы в динамике. Модель не содержит «временных срезов».

На матрице Захмана основывается специализированный продукт – Popkin Software Architect.

В последующем Захман внес вклад в разработку архитектуры предприятия Министерством обороны США. Эта попытка была предпринята в 1994 г.

Концепция Д. Захмана получила название «Базовая архитектура технического обеспечения для управления информацией» (*Technical Architecture Framework for Information Management – TAFIM*)¹. Преимущества, обеспечиваемые такими архитектурами предприятий, как TAFIM (например, приведение в соответствие технических проектов и потребностей бизнеса), были отмечены Конгрессом США, который Конгресс в 1996 г. принял закон, известный как акт Клингера – Коэна² и определивший реформу управления информационными технологиями.

¹ *Захман Дж. А.* Расширение и формализация структуры для архитектуры информационных систем / Дж. А. Захман, Дж. Ф. Сова // IBM Systems Journal, 1992. – Том 31. – № 3.

² Акт Клингера – Коэна от 1996 г. (PL 107-347) (см. THOMAS [Библиотека Конгресса]).

		Данные	Функции	Дислокация, сеть	Люди	Время	Мотивация	
		ЧТО	КАК	ГДЕ	КТО	КОГДА	ПОЧЕМУ	
ИТ-менеджеры и разработчики	Планировщик	Список важных понятий и объектов	Список основных бизнес-процессов	Территориальное расположение	Ключевые организации	Важнейшие события	Бизнес-цели и стратегии	Сфера действия (контекст)
	Владелец, менеджер	Концептуальная модель данных	Модель бизнес-процессов	Схема логистики	Модель потока работ (workflow)	Мастер-план реализации	Бизнес-план	Модель предприятия
	Конструктор, архитектор	Логическая модель данных	Архитектура приложений	Модель распределенной архитектуры	Архитектура интерфейса пользователя	Структура процессов	Роли и модели бизнес-правил	Модель системы
	Проектировщик	Физическая модель данных	Системный проект	Технологич. архитектура	Архитектура презентации	Структуры управления	Описания бизнес-правил	Технологическая (физическая) модель
	Разработчик	Описание структуры данных	Программный код	Сетевая архитектура	Архитектура безопасности	Определение временных привязок	Реализация бизнес-логики	Детали реализации
			Данные	Работающие программы	Сеть	Реальные люди, организации	Бизнес-события	Работающие бизнес-стратегии
		Данные	Функции, Процессы	Сеть, расположение систем	Люди, организации	Время, расписание	Мотивация	

Рис. П1.2. Формализованное представление схемы архитектуры предприятия Захмана 1992 г.

В соответствии с этим законом всем федеральным агентствам было предписано принять меры по повышению эффективности инвестиций в ИТ.

Для надзора за выполнением закона был сформирован Совет директоров по информационным технологиям ³, в который вошли директора по информационным технологиям из всех основных правительственных органов.

В 1998 г., четыре года спустя после разработки TAFIM и два года спустя после оформления этой методологии в виде акта Клингера – Коэна, методология TAFIM была официально отменена Министерством обороны. Все наработки по TAFIM были преобразованы в открытую группу, а затем – в новый стандарт, известный в настоящее время под названием TOGAF (The Open Group Architectural Framework).

На базе идеи архитектуры Дж. Захмана был разработан ряд концепций, модифицирующих и развивающих ее.

³ Совет директоров по информационным технологиям A04. Структура архитектуры федеральной организации, версия 1.1. Сентябрь 1999 г.

П1.2. Модель ЕАР

Модель ЕАР (Enterprise Architecture Planning – Планирование архитектуры предприятия) предложена в 1992 году *Стивеном Спиваком*.

Модель основана на упрощенной матрице Захмана 1987 года. Суть процесса ЕАР состоит в определении верхних строк этой матрицы (при этом включена концепция технических средств), т.е. это перспективы, соответствующие представлениям об архитектуре бизнес-руководителей: «планировщика» и «владельца».

Модель включает 7 шагов, определяющих архитектуру и соответствующий план ее реализации (рис. П1.3) [6].



Рис. П1.3. Методика ЕАР планирования Архитектуры предприятия

Методика ЕАР основана на сегментном подходе к разработке архитектуры, обеспечивает взгляд на предприятие с точки зрения его бизнес-функций и требований в области информации. Это инструмент планирования, а не детального проектирования архитектуры. Результаты планирования используются в качестве основы для интегрированной разработки прикладных систем и технологий, которые обеспечивают потребности бизнеса. Отличительными характеристиками этого подхода к планированию архитектуры являются следующие:

- в основе – потребности бизнеса, а не технологические факторы;
- основное внимание сосредоточено более на данных и потребностях в информации, чем на процессах;
- ответственность за процесс в большей степени несут представители бизнес-подразделений, чем специалисты по ИТ.

Привязка ЕАР к стабильности бизнес-модели, а также к подходу, «управляемому данными», ограничивает применение ЕАР.

Подходом Спивака пользовались такие организации как Министерство энергетики США, Штаб Военно-воздушных сил США.

П1.3. Модель 4-доменной архитектуры FDA)

В 1994 г *B. Iyer, R. Gottlieb* разработали модифицированную модель Д. Захмана – модель 4-доменной архитектуры (Four Domains architecture, FDA), в которой предлагается условно разбить ячейки модели Захмана на 2 компоненты – архитектуру описания (Architecture-in-Design) и архитектуру исполнения (Architecture-in-Operation). При этом первая компонента описывает ход, средства и артефакты процесса разработки архитектуры предприятия, в то время как вторая предназначена для описания непосредственно бизнес-процессов и реализации ИТ-систем.

П2. Модель «4+1»

В 1995 году *Филипп Кручен* предложил модель «4+1» (точнее The 4+1 View Model of Architecture») [6].

В этой модели предлагается использовать 5 представлений (рис. П1.4 [6]).

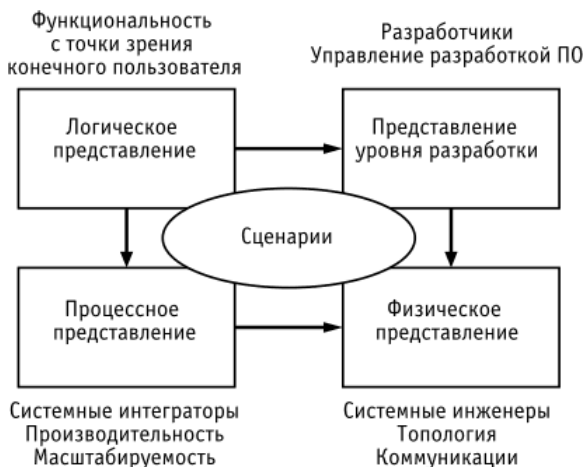


Рис. П1.4. Модель «4+1»

Четырьмя основными представлениями являются следующие:

- **Логическое представление.** Отвечает на вопрос: что система должна выполнять в терминах конечных пользователей? Для иллюстрирования могут применяться диаграммы классов (в нотации языка UML).

- **Процессное представление.** Учитывает нефункциональные требования к системе, включая производительность и доступность. Описывает вопросы параллельного исполнения и синхронизации процессов.

- **Физическое представление.** Описывает размещение программных компонент системы на аппаратных платформах и аспекты, связанные с физическим расположением системы.

- **Представление уровня разработки.** Описывает фактическую организацию модулей системы, разделение ее на подсистемы, которые могут разрабатываться независимо.

Архитектура системы во многом определяется *сценариями*, они объединяют все представления вместе.

Сценарии использования описываются как последовательность взаимодействия объектов и процессов. Они отражают наиболее важные требования, которым должна удовлетворять система. Это представление в каком-то смысле является избыточным и пересекается с четырьмя предыдущими, но оно важно по следующим причинам:

Сценарии использования позволяют идентифицировать элементы архитектуры, которые требуются для эффективно работающей системы. С помощью сценариев можно выполнять проверку и иллюстрацию того, что архитектура является работоспособной и полной.

П.3. Модель архитектуры SAM

В 1996 г. английская консалтинговая компания Systems Advisers Ltd. разработала стратегическую модель архитектуры SAM (Strategic Architecture Model) [6].

SAM – это надстройка над моделью архитектуры предприятия Захмана. Она предоставляет общие структуры для определения архитектуры и механизмы, позволяющие организовать и анализировать информацию об архитектуре.

SAM использует нотацию «сфер интересов» для представления целостного набора фактов о предприятии и «отношений», которые связывают эти факты в «полезные» группы.

«Сферы интересов» SAM позволяют систематизировать всю информацию, имеющую отношение к определенному предмету. Сфера может заполняться в направлении «снизу–вверх» путем сбора относящейся к предметной области информации, а на более высоких уровнях эта информация будет обобщаться. Либо же заполнение может идти в направлении «сверху–вниз» с постепенной детализацией.

Типичные сферы интересов SAM представлены на рис. П1.5 [6].

Иерархическая структура – напоминает ящик с файлами. Минимальный объем информации, относящейся к какой-либо сфере, называется элементом.

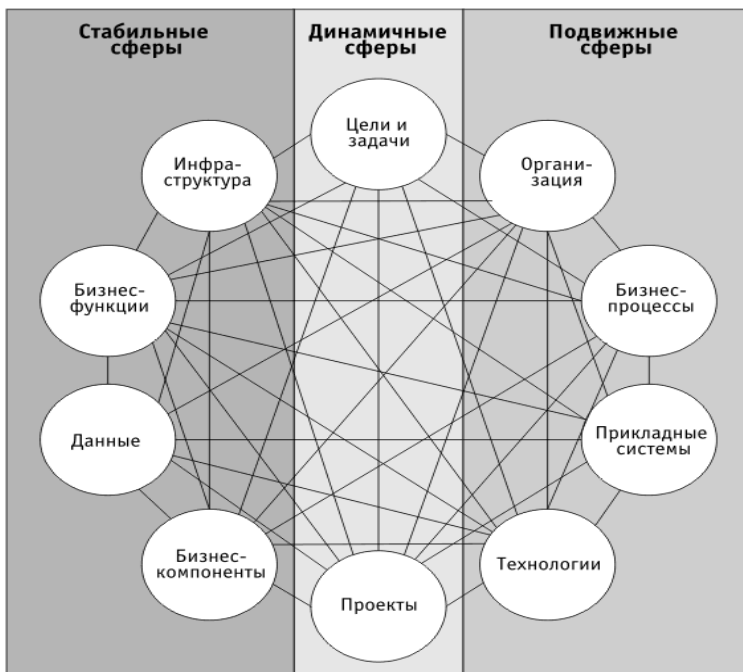


Рис. П1.5. Типичные сферы интересов SAM

Примерами элементов различных сфер являются:

- конкретные бизнес-процессы, например, «Прием заказа» в сфере «Бизнес-процессы»;
- конкретные информационные объекты, такие как «Клиент» в сфере «Данные».

В модели SAM выделены следующие категории сфер:

• **Стабильные.** Эти сферы представляют фундаментальные структуры: бизнес-функции, данные, бизнес-компоненты и инфраструктуру.

• **Подвижные.** Эти сферы описывают то, что предприятие делает или может делать с точки зрения бизнеса, в том числе для того чтобы обеспечить отличия от конкурентов и динамичность в своей деятельности. Сферы, которые относятся к этому разделу – представляют собой области, которые организация может изменить достаточно быстро.

• **Динамические.** Эти сферы задают направления бизнеса, рабочие программы, управление изменениями. Они описывают основные области, в которых работает предприятие, и усилия, которые требуются

для движения в сторону достижения целей и задач посредством связанных между собой проектов.

Предложенная в SAM классификация позволяет понимать, какая часть архитектуры конкретного предприятия носит достаточно стабильный характер, а какая требует постоянных изменений.

П.4. «3D –модель» предприятия

В 1996–1997 гг. *Евгений Зиндер* предложил «3D –модель» предприятия¹, в которой введена ось времени, где располагаются интервалы осуществления различных проектов и стадий развития ИС и всего предприятия. В качестве других осей выступает матрица Захмана (рис. П1.6):

- 1) ось уровня проектирования и использования ИС;
- 2) ось раздела обеспечения и аспекта работы ИС;
- 3) ось времени, в котором развивается предприятие и его ИС.



Рис. П1.6. Обобщенная схема «3D-предприятие»

Стратегический и детальный анализ могут рассматриваться и как разные стадии, что демонстрирует принцип адаптации схемы к жизненным циклам разных типов. Характер расположения архитектурных компонентов ИС в этом третьем измерении отличается большим раз-

¹ *Зиндер Е.* «3D-предприятие» – модель трансформирующейся системы / Евгений Зиндер // http://www.iteam.ru/publications/it/section_53/article_1272.

нообразием, поскольку в реальной жизни многие процессы трансформации предприятия идут параллельно и итерационно.

Создаваемые в ячейках частные модели должны быть согласованы в своих взаимосвязях. Особенность 3D-предприятия в том, что эти взаимосвязи определяются не только для какого-то одного момента, но и на концах всех отрезков оси времени, которым приписаны рассматриваемые проекты, стадии и работы.

Применение схемы можно разбить на следующие шаги:

Первым шагом является общее обсуждение 3D-схемы руководителями предприятия и его подразделений. Оно имеет своей целью достижение общего понимания всех типов сущностей этой схемы как компонентов и представлений системы в процессах их жизни – процессах создания и последующих трансформаций.

Вторым шагом является разделение работ по построению самой общей модели 3D-предприятия между руководителями. Координатором всех работ может быть заместитель генерального директора – директор по развитию.

Третьим шагом является привлечение специалистов и руководителей среднего звена с закреплением за ними конкретных «участков», то есть областей моделирования, на уровнях 3D-модели со второго и ниже.

Четвертым шагом является первоначальное и неформальное описание тех частных моделей, которые актуальны для погружения в общую модель «как есть».

Пятым шагом является рассмотрение совокупности частных моделей с их взаимосвязями.

П.5. Методология GERAM

В 1999 г. рабочая группа IFIP-IFAC Ассоциации CIMOSA разработала Обобщенную референсную архитектуру и методологию предприятия – Generalised Enterprise Reference Architecture and Methodology (GERAM) ¹.

Схема GERAM предусматривала (рис. П1.7) [6]:

- четыре группы аспектов архитектуры предприятия, названных представлениями (Views) – типы моделей («функции», «данные», «ресур-

¹ Методология GERAM включена в качестве приложения в действующий базовый стандарт «Архитектура Предприятия – ISO 15704:2000 «Requirements for enterprise-reference architectures and methodologies» и Национальный стандарт Российской Федерации «Промышленные автоматизированные системы. Требования к стандартным архитектурам и методологиям предприятия. Industrial automation systems. Requirements for enterprise-reference architectures and methodologies. Дата введения - 2010-01-01.

сы», «организация»), назначения (может быть ассоциировано со столбцом «ЗАЧЕМ» Захмана), реализации и «физические представления» (аппаратура, ПО) и возможность определять дополнительные аспекты;

- описание всех аспектов или какой-то их части на каждой из семи или восьми фаз формирования архитектуры и функционирования предприятия;

- конкретизацию модели архитектуры на трех уровнях – обобщенном, уровне частичных моделей и конкретных моделей.

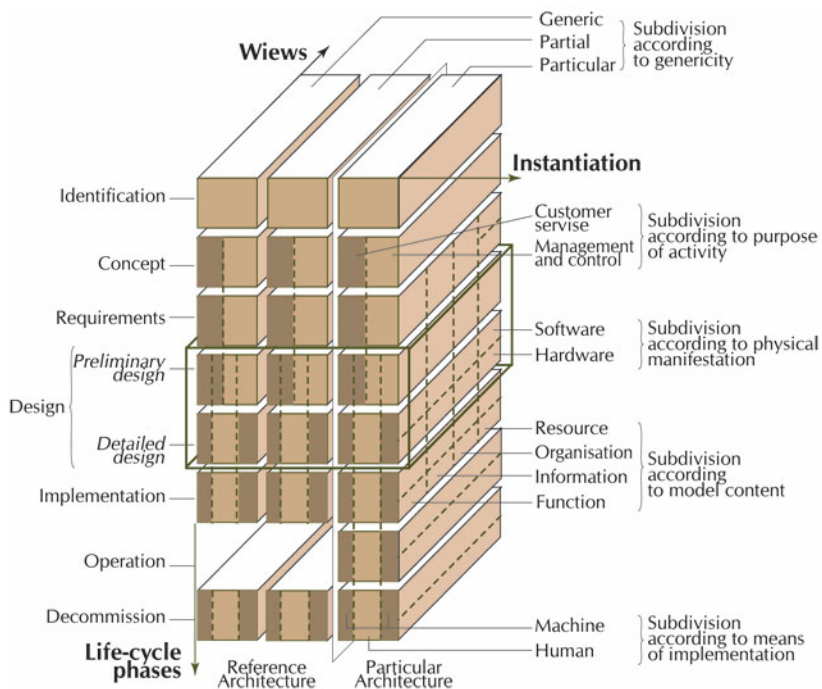


Рис. П1.7. Обобщенная схема GERAM

Особенность GERAM состоит в том, что ось фаз развития архитектуры в принципе может отражать ход времени и не связана жестко с представлениями разных участников процесса.

Недостатки схемы таковы:

- ось фаз формирования архитектуры недостаточно явно соответствует спиральной форме истории жизни предприятия в реальном времени его развития;

- структура схемы и ее описание, достаточно сложны для восприятия, в результате чего GERAM, оставаясь частью стандарта ISO 15704:2000, в широкой практике не получила заметного распространения¹.

Тем не менее, GERAM содержит множество полезных методических элементов и до сих пор учитывается в новых стандартах (например, ISO 19439:2006 «Enterprise integration – Framework for enterprise modelling»).

П.6. FEAF

В апреле 1998 г. совет директоров по информационным технологиям начал работу над первым крупным проектом, структурой архитектуры федеральной организации FEAF – Federal Enterprise Architecture Framework.

В 1999 г. Совет директоров по информационным технологиям основных правительственных органов(CIO Council) предложил структуру архитектуры федеральной организации (FEAF). Версия 1.1 ISO 15704, которая была выпущена в сентябре 1999 г. В этом документе содержался ряд инновационных идей, например идея «сегментированных архитектур» – т.е рассмотрение в архитектурном аспекте сегментированных подмножеств крупного предприятия.

Методология описания архитектуры, ориентированная на государственные ведомства
Подход к описанию архитектуры – со стороны бизнес-процессов. В документе содержался ряд инновационных идей, например идея «сегментированных архитектур» – то есть рассмотрение в архитектурном аспекте сегментированных подмножеств крупного предприятия.

FEAF состоит из 8 компонент (рис. П1.8) [6]:

1. Двигатели архитектуры. Отражают два типа внешних стимулов или источников изменения архитектуры: бизнес-стимулы и технические стимулы. В качестве бизнес-стимула могут выступать новое законодательство, новые инициативы Президента. В роли технических двигателей – новое и улучшенное программное обеспечение, аппаратные средства.

2. Стратегическое направление. Определяет разработку целевой архитектуры. Оно включает в себя видение, сжатое и стратегическое описание поставленной цели развития архитектуры на предстоящие 5 лет, принципы руководства развитием этой архитектуры, а также цели и объекты для управления процессом развития архитектуры.

¹ Стандарт IEEE 1471-2000. Рекомендации IEEE по архитектурному описанию преимущественно программных систем.

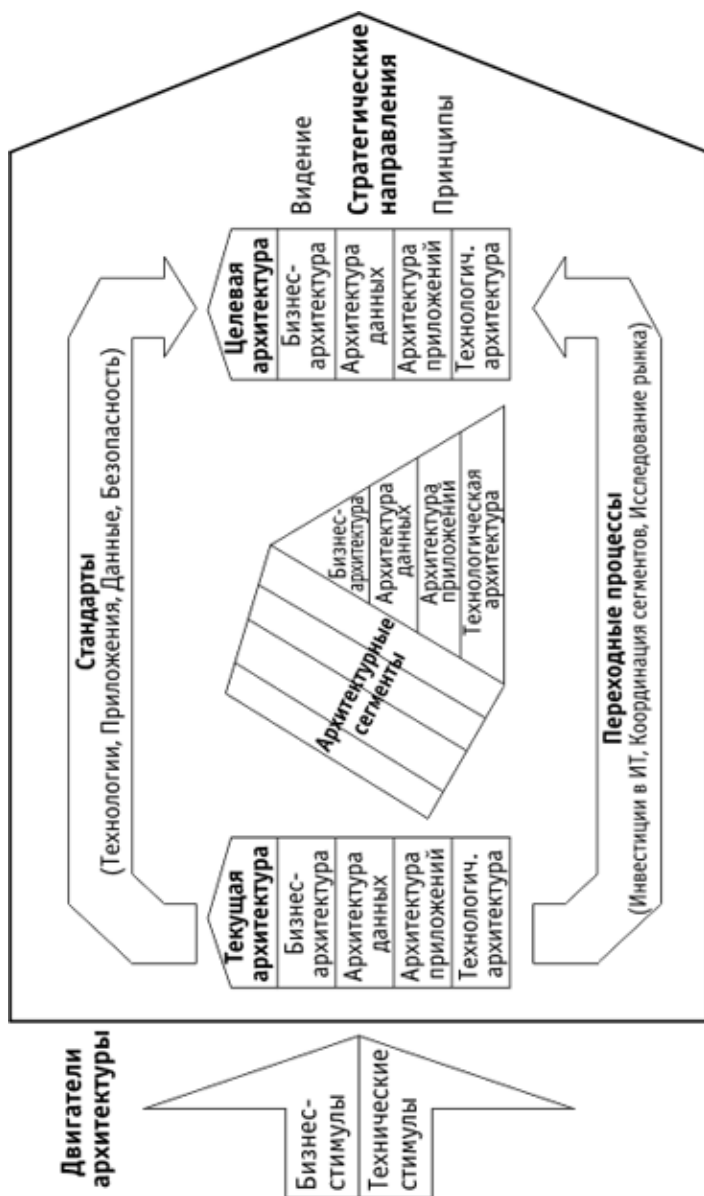


Рис. П1.8. Методология Федеральной Архитектуры и ее компоненты

3. Текущая архитектура. Определяет архитектуру «как есть» и состоит из двух частей:

- Текущая бизнес-архитектура – определяет сегодняшние потребности с точки зрения основной деятельности. Отвечает на вопрос о том, какие имеются в распоряжении функции, процессы, ресурсы.

- Текущая архитектура информационных технологий (т.е. архитектура данных, приложений и технологическая архитектура) – отображает текущее состояние возможностей технологий по обеспечению деятельности организаций.

4. Целевая архитектура. Определяет архитектуру «как должно быть». Она дает представление о будущих возможностях и технологиях, которые явятся результатом соответствующих изменений.

5. Переходные процессы. Поддерживают процесс перехода от текущей архитектуры к целевой архитектуре.

Примеры переходных процессов:

- планирование и принятие решения по инвестициям и ИТ;
- исследование рынка;
- управление архитектурой – координация усилий по сопровождению и управлению архитектурой.

6. Архитектурные сегменты. Отражают разбиение общей архитектуры на отдельные, существенные области деятельности, например:

- общие административные системы;
- области федеральных программ, таких как внешняя торговля или предоставление грантов; электронная торговля для проведения небольших закупок.

7. Архитектурные модели. Архитектурные модели задают бизнес-архитектуру и архитектуру информационных технологий.

При этом рассматриваются:

- Бизнес-модели. Это модели, которые отражают появление бизнес-потребностей, инициированных бизнес-двигателями. Моделирование предполагает создание общего набора определений, диаграмм, а также, возможно, использование автоматизированных инструментальных средств, которые облегчают понимание бизнес-функций, применяемой информации, процессов и продуктов.

- Модели технической среды. Включают модели данных, модели прикладных систем и технологические модели, которые требуются для того, чтобы поддержать реализацию бизнес-потребностей.

8. Стандарты. Включают стандарты, руководящие принципы, а также передовой опыт.

Таким образом, в методике выделены следующие представления:

- бизнес-архитектура (функциональная архитектура деятельности правительства);
- архитектура информации (данных);

- архитектура приложений;
- архитектура инфраструктуры (технологическая или системная архитектура).

Методология FEAF рассматривается в качестве ориентира и многими европейскими странами, и Евросоюзом в целом.

На определенном этапе развития работ по ИТ-архитектурам полномочия Совета директоров по информационным технологиям по архитектуре федеральной организации были переданы Административно-бюджетному управлению, которое в 2002 г. переработало методологию FEAF и переименовало ее в архитектуру федеральной организации (Federal Enterprise Architecture – FEA), информация о которой появилась в 2006 г. FEA является в настоящее время наиболее полной методологией и рассматривается после методологий, идеи которых возникли в 2003 г.

П.7 Методология TOGAF

Методология TOGAF (аббревиатура от The Open Group Architecture Framework – структура архитектуры The Open Group) разработана в 2003 г консорциумом The Open Group [www.open/group.org].

Представление архитектуры предприятия в методологии TOGAF показано на рис. П1.9.



Рис. П1.9. Архитектура предприятия TOGAF

Как показано на рисунке, в модели TOGAF архитектура предприятия подразделяется на четыре категории:

1. Архитектура бизнеса – описывает процессы, используемые для достижения бизнес-целей.
2. Архитектура приложений – описывает структуру конкретных приложений и их взаимодействие друг с другом.
3. Архитектура данных – описывает структуру корпоративных хранилищ данных и процедуры доступа к ним.

4. Технологическая архитектура – описывает инфраструктуру оборудования и программного обеспечения, в которой запускаются и взаимодействуют приложения.

Модель TOGAF позиционируется как *структура*, однако наиболее важным ее компонентом является методика разработки архитектуры (ADM), которую можно квалифицировать как *процесс*. С учетом того, что методика разработки архитектуры является наиболее значимой составляющей модели TOGAF, рассмотрим TOGAF в целом как архитектурный процесс, а не как архитектурную структуру (как позиционирует TOGAF консорциум The Open Group) или методологию (как позиционируется методика разработки архитектуры).

Модель TOGAF применяется для описания интеграционных компонент, использующихся для поддержки широкого спектра корпоративных приложений. Как архитектурный *процесс* модель TOGAF дополняет модель Захмана. Захман показывает, как следует классифицировать артефакты. Модель TOGAF описывает процесс создания артефактов.

В состав модели TOGAF входят две основные компоненты – **Методика ADM** (Architecture Development Method), определяющая процесс разработки архитектуры, и **Базовая Архитектура** (Foundation Architecture), которая дополняется соответствующей базой данных ресурсов, включающей описания архитектурных принципов, примеров реализации, а также специализированный язык ADML.

Общая структура TOGAF показана на рис. П1.10.



Рис. П1.10. Структура TOGAF

На рис. П1.11 показано соотношение между континуумом предприятия и методикой разработки архитектуры (ADM).

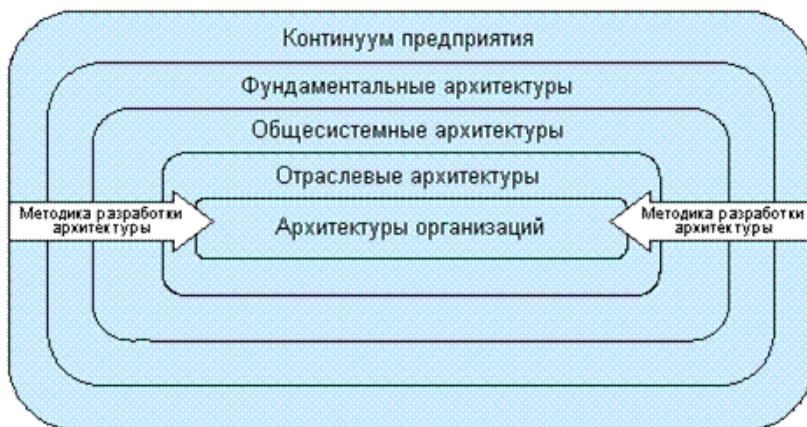


Рис. П1.11. Континуум предприятия в модели TOGAF

Процесс создания конкретной архитектуры предприятия, рассматривается как переход от общей архитектуры к специализированной. Фундаментальная \Rightarrow Общесистемная \Rightarrow Отраслевая \Rightarrow Архитектура организации.

В модели TOGAF мир архитектуры предприятия рассматривается как континуум архитектур, от максимально обобщенных до максимально специализированных. Этот континуум называется континуумом предприятия. Процесс создания конкретной архитектуры предприятия рассматривается как переход от общей архитектуры к специализированной. Методика разработки архитектуры в модели TOGAF представляет собой процесс осуществления такого перехода.

Фундаментальная архитектура включает:

- набор наиболее общих служб и функций, объединенных в Техническую Эталонную Модель (Technical Reference Model – TRM); Техническая эталонная модель является рекомендуемым описанием общей ИТ-архитектуры. Сюда входит система общих служб, включающая такие службы, как Обмен и Преобразование данных, Управление данными, Поддержка интернационализации, Службы Каталогов и т. п.;
- набор элементарных архитектурных элементов, которые используются как «строительные блоки» при построении конкретных решений (архитектура безопасности, сетевая архитектура и т. п.);
- база данных стандартов (Standards Information Base).

Эти принципы построения архитектуры теоретически могут использоваться практически любой ИТ-организацией в мире.

Следующий уровень специализации в модели TOGAF называется **общесистемными архитектурами**. Эти принципы прослеживаются во многих – возможно, не во всех – типах предприятий.

Уровень **Отраслевой Архитектуры** добавляет специфичные для каждой индустрии модели данных, приложения, стандарты, бизнес-правила, а также, при необходимости, процедуры взаимодействия различных отраслевых систем между собой. Эти принципы характерны для предприятий, занятых в одной сфере деятельности.

Самый высокий уровень специализации в модели TOGAF называется **архитектурами организаций**. Формируется архитектура ИТ-систем конкретного предприятия, учитывающая все его особенности.

Для всех используемых в архитектуре служб, наряду с функциональным назначением, необходимо определить и уровень качества реализации, то есть такие характеристики как управляемость, гибкость, гарантированность, удобство использования и т.п.

TOGAF распространяется свободно и может быть использована бесплатно любой организацией для разработки внутренних проектов. Лицензируется только коммерческое использование.

В модели TOGAF на уровне фундаментальных архитектур определяется ряд баз знаний. Основные из них: техническая эталонная модель TRM и информационная база стандартов SIB. TRM является рекомендуемым описанием общей ИТ-архитектуры. SIB представляет собой набор стандартов и псевдостандартов, которые консорциум The Open Group рекомендует использовать при построении ИТ-архитектуры.

В TOGAF техническая эталонная модель и информационная база стандартов рекомендуются к использованию, но не являются обязательными. По мнению авторов, и техническая эталонная модель, и информационная база стандартов не лишены недостатков по следующей причине: они направлены на обеспечение переносимости приложений в ущерб их способности к взаимодействию и автономности. Я считаю, что такое представление технических архитектур устарело.

Модель TOGAF в значительной степени сводится к методике разработки архитектуры. Сотрудники организации будут работать с континуумом предприятия, информационной базой стандартов и технической эталонной моделью (а также с рядом других возможностей TOGAF). Однако для каждодневного создания архитектуры предприятия в основном будет использоваться методика разработки архитектуры, высокоуровневое представление которой показано на рис.П1.12.

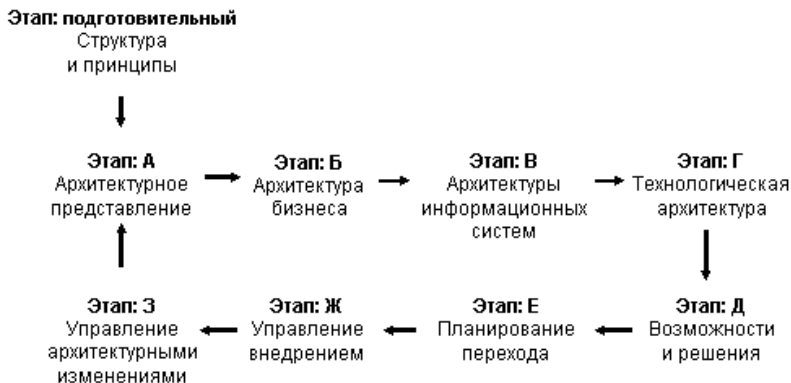


Рис. П1.12. Методика разработки архитектуры ADM в модели TOGAF

В соответствии с **Методикой ADM**, процесс разработки архитектуры включает следующие фазы [20]:

- Подготовка: уточнение модели под особенности организации, определение принципов реализации проекта.
- Фаза А: определение границ проекта, разработка общего представления (Vision) архитектуры; утверждение плана работ и подхода руководством.
- Фаза В: разработка бизнес-архитектуры предприятия.
- Фаза С: разработка архитектуры данных и архитектуры приложений.
- Фаза D: разработка технологической архитектуры.
- Фаза Е: проверка возможности реализации предложенных решений.
- Фаза F: планирование перехода к новой системе.
- Фаза G: формирование системы управления преобразованиями.
- Фаза H: управление изменением архитектуры.

Каждая фаза, в свою очередь разбивается на подпроцессы (этапы), отдельные работы.

Как показано на рис. П1.12, методика разработки архитектуры в модели TOGAF состоит из восьми этапов, которые циклически повторяются.

Перед тем как компания сможет приступить к использованию методики разработки архитектуры, ей необходимо изучить модель TOGAF. Изучить модель TOGAF можно двумя способами.

Во-первых, компания может изучить модель TOGAF самостоятельно. Компания может загрузить документацию по TOGAF [www.opengroup.org/togaf] – в ней достаточно подробно описаны все возможности TOGAF, в том числе методика разработки архитектуры.

Можно также приобрести книги по TOGAF¹. О модели TOGAF доступно больше информации (как бесплатной, так и за умеренную цену), чем обо всех остальных методологиях построения архитектуры вместе взятых.

Во-вторых, компания может обратиться к экспертам по TOGAF. На рынке работает множество консультантов по TOGAF, обладающих сертификатами Open Group [www.opengroup.org/togaf/cert].

В некоторых организациях осознание необходимости создания архитектуры предприятия дается с большим трудом. Такая ситуация возникает, когда инициатива исходит от ИТ-подразделения, и особенно в случае продолжительного противостояния бизнес- и ИТ-подразделений организации.

По завершении подготовительного этапа можно перейти к этапу А. Этот этап начинается, по крайней мере в теории, с Запроса на разработку архитектуры от какого-либо подразделения компании.

После получения «Запроса на разработку архитектуры» (или аналогичного документа) консультант по TOGAF переходит к этапу А. На этом этапе предстоит убедиться в том, что проект надлежащим образом поддерживается компанией, определить область действия проекта, выявить ограничения, документально оформить бизнес-требования и разработать высокоуровневые определения как для базовой (начальной), так и для целевой (желаемой) архитектуры.

Эти базовые и целевые определения включают высокоуровневые определения для всех архитектур, составляющих архитектуру предприятия, приведенных на рис. П1.9, а именно для архитектуры бизнеса, технологической архитектуры, архитектуры данных и архитектуры приложений.

Основным документом, создаваемым на этапе А, является «Постановление о разработке архитектуры», которое должно быть утверждено всеми заинтересованными лицами. После этого можно переходить к следующему этапу. В конце этого этапа формируется архитектурное представление для первой итерации цикла разработки архитектуры.

Архитектурное представление, созданное на этапе А, является основой для этапа Б. Основная цель на этапе Б заключается в создании детализированной базовой и целевой архитектуры бизнеса и всестороннем анализе различий между этими архитектурами.

Этап Б довольно сложен: он включает бизнес-моделирование, подробный анализ бизнеса и документирование технических требований. Для успешной реализации этапа Б необходимо участие многих

¹ Руководство по ИТ-архитектуре предприятия / Перкс Кол и Тони Беверидж. – Нью-Йорк, штат Нью-Йорк: Springer, 2003. ISBN: 0-387-95132-6.

заинтересованных лиц. Основным результатом является подробное описание базовых и желаемых бизнес-целей, а также описание различий между двумя архитектурами бизнеса.

На этапе В выполняются все те же действия, что и на этапе Б, только для архитектуры информационных систем. В модели TOGAF определено девять этапов, каждый из которых разбит на несколько подэтапов:

1. Разработка описания базовой архитектуры данных.
2. Просмотр и проверка принципов, эталонных моделей, точек зрения и средств.
3. Создание архитектурных моделей, в том числе логических моделей данных, моделей управления данными и моделей отношений, в которых бизнес-функции сопоставляются операциям над данными (создание, чтение, обновление и удаление).
4. Выбор компонентов архитектуры данных.
5. Формальный анализ контрольных точек модели архитектуры и ее компонентов вместе с заинтересованными лицами.
6. Анализ качественных критериев (например, производительности, надежности, безопасности и целостности).
7. Завершение архитектуры данных.
8. Анализ контрольных точек и последствий.
9. Анализ различий.

Наиболее важным результатом этого этапа является информация о целях и архитектура приложений.

На этапе Г формируется технологическая архитектура – инфраструктура, необходимая для поддержки новой архитектуры.

На этапе Д выполняется оценка различных возможностей реализации, определяются основные проекты по внедрению, которые необходимо выполнить, а также связанные с каждым проектом возможности для бизнеса. Стандартом TOGAF на этапе Д рекомендуется «сосредоточиться на проектах, которые дадут результаты в краткосрочной перспективе и позволят реализовать долгосрочные проекты».

Этап Е тесно связан с этапом Д. На этом этапе консультант по TOGAF работает с руководством компании, чтобы упорядочить по приоритетам проекты, выбранные на этапе Д, с учетом не только затрат и выхода (определенных на этапе Д), но и факторов риска.

На этапе Ж на основе упорядоченного по приоритетам списка проектов создаются спецификации архитектуры для проектов реализации. Эти спецификации включают критерии приемки и списки рисков и проблем.

Последним этапом является этап З. На этом этапе консультант по TOGAF корректирует процесс управления архитектурными измене-

ниями с учетом новых артефактов, созданных на последней итерации, и новых данных.

После этого цикл повторяется. Одной из целей первого цикла является передача информации.

Многие результаты процесса TOGAF можно получить как с помощью консультанта, так и на основе собственно спецификации TOGAF. Методология TOGAF весьма гибкая, а детали реализации архитектурных артефактов могут быть различны. В одной из книг по TOGAF говорится:

«Методология TOGAF – это не только и не столько создаваемые документы; фактически она в меньшей степени ориентирована на шаблоны документов, а в большей – на то, что мы получаем на входе и на выходе»¹.

Спецификация TOGAF также позволяет гибко работать с этапами. В самой спецификации говорится следующее:

Перед применением методики разработки архитектуры необходимо проверить компоненты на применимость, а затем связать их с конкретными обстоятельствами отдельного предприятия. Это позволяет создать методику разработки архитектуры для конкретного предприятия².

Модель TOGAF позволяет выполнять этапы частично, пропускать их, объединять, изменять порядок и вносить изменения в соответствии с конкретными требованиями. Поэтому два сертифицированных консультанта по TOGAF могут разработать два совершенно различных процесса – даже при работе с одной и той же организацией.

Модель TOGAF обладает еще большей гибкостью в отношении созданной архитектуры. Окончательная архитектура может с одинаковым успехом быть хорошей, плохой или неопределенного качества. В TOGAF описывается, как создать архитектуру предприятия, но не описывается, как создать хорошую архитектуру. Качество конечного продукта зависит от опыта персонала компании и консультанта по TOGAF.

П.6. DoDAF

В декабре 2003 была разработана рамочная архитектура Министерства Обороны или DoDAF (Department of Defence Architecture Framework)

DoDAF содержит правила, руководства и документы, которые должны использоваться при разработке и описании архитектуры различных систем,

¹ Руководство по ИТ-архитектуре предприятия / Перкс Кол и Тони Беверидж. – Нью-Йорк, штат Нью-Йорк: Springer, 2003. ISBN: 0-387-95132-6

² TOGAF, версия 8.1.1.

используемых военными ведомствами США. Структура основных представлений, используемых в DoDAF приведена на рис. П1.13 [6].

Архитектуру описывают 3 представления:

- операционное,
- системное,
- представление технических параметров.

Каждое из них используется для отражения различных архитектурных характеристик и атрибутов. Имеются пересечения - некоторые из атрибутов как бы объединяют два различных представления, что обеспечивает целостность, единство и единообразие в описании архитектуры.

Операционное представление включает описание задач и активностей, операционных элементов и информационных потоков, которые требуются для выполнения миссии министерства. Миссия включает как военные операции, так и традиционные бизнес-процессы. Это представление содержит описания (часто графические), которые включают узлы выполнения операций и элементы, назначенные задачи и активности, и информационные потоки между узлами. Оно определяет тип данных в информационном обмене, частоту обмена, характер информационного обмена.



Рис. П1.13. Три основных представления, используемые в DoDAF

Системное представление включает текстовые и графические описания систем и связей между ними.

Представление технических стандартов определяет минимальный набор правил, которыми руководствуются при создании, обеспечении взаимодействия и взаимосвязей между частями систем и элементами (технические руководства, технические стандарты, соглашения по их реализации, опции при использовании стандартов, правила и критерии).

Аспекты, имеющие отношение ко всем трем представлениям архитектуры (общие аспекты) отражаются с помощью общих для всех представлений продуктов (описаний – текстовых и графических). Такие продукты предоставляют информацию, относящуюся к архитектуре в целом, которая не подпадает под область ответственности отдельного представления. Они задают:

- масштаб;
- охват (включает предметную область и временные рамки для архитектуры);
- контекст архитектуры (тактика; связанные с ней цели и видение; концепция операций; сценарии; условия внешней среды).

Важнейшим элементом и сильной стороной методики являются так называемые архитектурные продукты. Это графические, текстовые, табличные описания, которые создаются в процессе описания архитектуры и которые фиксируют характеристики, имеющие отношение к процессу. Будучи частью описания архитектуры, все продукты, в том числе и графические, должны содержать пояснительный текст. Например, для графических продуктов должны быть указаны все используемые сокращения и пояснения по поводу их смысла. Имеются подробнейшие описания этих архитектурных продуктов, а также рекомендации, связанные с тем, как эти различные продукты должны быть связаны между собой при создании интегрированного описания архитектуры.

П.8. Архитектура федеральной организации (FEA)

На определенном этапе развития работ по ИТ-архитектурам полномочия Совета директоров по информационным технологиям по архитектуре федеральной организации были переданы Административно-бюджетному управлению. В 2002 г. Административно-бюджетное управление переработало методологию FEAF и переименовало ее в архитектуру федеральной организации (Federal Enterprise Architecture – FEA).

FEA можно рассматривать и как методологию создания архитектуры предприятия, и как результат применения этой процедуры к конкретной организации – Правительству США, и как методологию коммерческой организации.

Архитектура федеральной организации ФЕА – это последняя попытка федерального правительства привести бесчисленное множество агентств к единой и повсеместно используемой архитектуре. ФЕА по-прежнему находится в самой ранней стадии развития: большинство компонентов этой методологии стали доступны только в 2006 г. Тем не менее, за этой методологией стоит длительная традиция и множество неудачных попыток, из которых были извлечены полезные уроки.

ФЕА является наиболее полной методологией из всех упомянутых. Она включает и всеобъемлющую таксономию, как в методологии Захмана, и архитектурный процесс, как в модели TOGAF.

Методологии ФЕА включает следующие пункты:

- Набор эталонных моделей, описывающих различные точки зрения на архитектуру предприятия (большинство авторов описывают ФЕА как набор из пяти эталонных моделей: модель бизнеса, модель обслуживания, модель компонентов, технологическая модель и модель данных).
- Процесс создания архитектуры предприятия.
- Процесс перехода от старой парадигмы (до создания архитектуры предприятия) к новой (после создания архитектуры предприятия).
- Таксономия для классификации активов, которые попадают в область действия архитектуры предприятия.
- Методика, позволяющая оценить успешность использования архитектуры предприятия для повышения ценности бизнеса.

Архитектура состоит из отдельных *сегментов*. Сегмент представляет собой один из основных аспектов организации, например трудовые ресурсы.

Сегменты подразделяются на два типа: *базовые* и *служебные*.

Базовый сегмент – это ключевой аспект деятельности организации. Например, для Министерства здравоохранения базовым сегментом является *здоровье*.

Служебный сегмент – это сегмент, который является фундаментальным если не для всех, то для большинства организаций. Например, *управление финансами* является служебным сегментом, обязательным для всех федеральных агентств.

Другим типом активов в архитектуре предприятия являются *службы предприятия*. Например, управление безопасностью – это служба, единообразно реализованная по всей организации.

Схема сегментов федерального правительства приведена на рис. П1.14.

Из рисунка видно, что многие сегменты используются во многих агентствах и все или почти все эти сегменты можно использовать повторно.

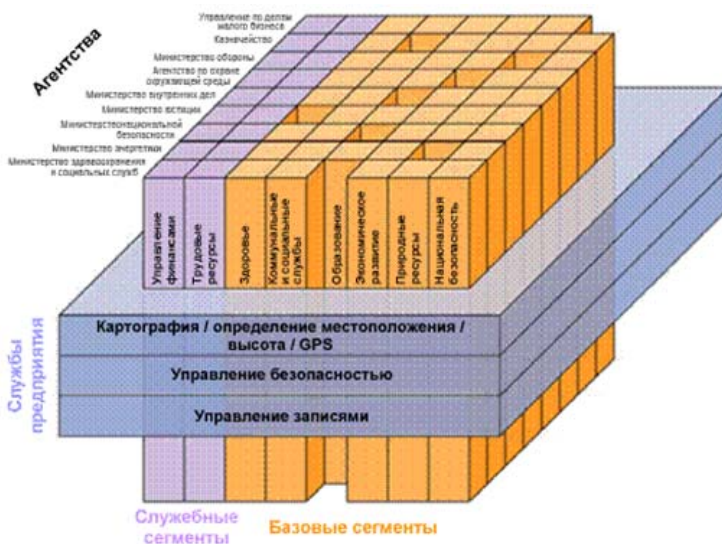


Рис. П1.14. Схема сегментов федерального правительства

Административно-бюджетное управление присваивает каждому агентству рейтинг успеха на основе оценок по каждой категории и совокупному показателю следующим образом: например, зеленый – агентство показало хорошие результаты в области *завершенности* (имеет развитую архитектуру предприятия).

Эта структура применима и за пределами государственного сектора экономики. Рейтинги по категориям можно эффективно адаптировать ко многим предприятиям для оценки степени готовности их архитектуры.

FEA действительно включает эти пять моделей, однако представляет собой нечто намного большее, чем просто набор эталонных моделей. Исчерпывающее описание методологии FEA должно включать следующие пункты:

- Точка зрения, с которой будут рассматриваться архитектуры предприятия (модель сегмента, которая вкратце будет рассмотрена ниже).
- Набор эталонных моделей, описывающих различные точки зрения на архитектуру предприятия (пять перечисленных выше моделей).
- Процесс создания архитектуры предприятия.
- Процесс перехода от старой парадигмы (до создания архитектуры предприятия) к новой (после создания архитектуры предприятия).
- Таксономия для классификации активов, которые попадают в область действия архитектуры предприятия.

- Методика, позволяющая оценить успешность использования архитектуры предприятия для повышения ценности бизнеса

Очевидно, что FEA представляет собой нечто большее, чем набор моделей. Эта методология включает в себя все необходимое для построения архитектуры предприятия даже для самой сложной организации в мире – Правительства США. По заявлению управления по реализации программы FEA (FEAPMO), методология FEA в целом обеспечивает:

«...общий язык и структуру для описания и анализа инвестиций в ИТ, повышает эффективность совместной работы и позволяет преобразовать федеральное правительство в организацию, ориентированную на граждан, направленную на достижение высоких результатов и отвечающую требованиям рынка в соответствии с программой президента США»¹.

Хотя на первый взгляд разве что вмешательство высших сил позволит «преобразовать федеральное правительство в организацию, ориентированную на граждан, направленную на достижение высоких результатов и отвечающую требованиям рынка», есть надежда, что методология FEA может помочь нашей многострадальной корпорации MedAMore справиться с гораздо менее глобальными проблемами. Так что же может предложить методология FEA при ближайшем рассмотрении?

С точки зрения FEA, архитектура предприятия также состоит из отдельных сегментов. Эта идея была впервые изложена в FEAF². Сегмент представляет собой один из основных аспектов бизнеса, например трудовые ресурсы. Сегменты подразделяются на два типа: базовые и служебные.

Базовый сегмент представляет собой ключевой аспект деятельности предприятия в границах политико-административного деления.

Служебный сегмент – это сегмент, который является фундаментальным если не для всех, то для большинства организаций данного типа.

Другим типом активов в архитектуре предприятия являются службы предприятия. Служба предприятия – это четко определенная функция в границах политико-административного деления.

Различие между службами предприятия и сегментами, особенно служебными сегментами, неочевидно. И службы, и сегменты охватывают все предприятие. Различие заключается в том, что область действия служебных сегментов распространяется только на одну политиче-

¹ «Документация по эталонным моделям FEA, версия 2.1», декабрь 2006 г., опубликовано FEAPMO, Административно-бюджетное управление.

² Практическое руководство по FEA совета директоров по информационным технологиям, версия 1.0, февраль 2001 г.

скую организацию. Область действия служб предприятия распространяется на все предприятие.

Например, и в Министерстве здравоохранения и социальных служб, и в Агентстве по охране окружающей среды федерального правительства США используется служебный сегмент трудовые ресурсы. Однако трудовые ресурсы для Министерства здравоохранения и социальных служб отличаются от трудовых ресурсов для Агентства по охране окружающей среды.

И в Министерстве здравоохранения и социальных служб, и в Агентстве по охране окружающей среды используется такая служба предприятия, как управление безопасностью. При этом учетные записи для безопасного доступа, используемые службой управления безопасностью, одинаковы для обоих агентств. Эффективное управление учетными записями для безопасного доступа обеспечивается только в том случае, если оно осуществляется на уровне предприятия.

Возникает соблазн приравнять сегменты или службы к службам, используемым в сервис-ориентированных архитектурах. Такой подход небезупречен по двум причинам. Во-первых, службы предприятия, служебные и базовые сегменты намного шире по охвату, чем службы в сервис-ориентированных архитектурах. Во-вторых, сегменты являются организационной единицей для архитектуры предприятия, а службы – организационной единицей для технической реализации. Что касается организационных единиц для архитектуры предприятия, они охватывают не только технологическую архитектуру, но и архитектуры бизнеса и данных.

Последнее примечание относительно сегментов. Хотя сегменты функционируют на политическом уровне (т.е. на уровне агентств), они определяются на уровне предприятия (то есть на уровне правительства). Службы предприятия, естественно, функционируют и определяются на уровне предприятия.

Тот факт, что сегменты определяются глобально, упрощает их повторное использование в границах политико-административного деления. Можно спланировать использование сегментов в границах политико-административного деления предприятия, а затем воспользоваться этим планом для поиска возможностей повторного использования разработанной архитектуры.

Эталонные модели FEA

Все пять эталонных моделей FEA предназначены для формирования единого языка. Цель – упростить взаимодействие, сотрудничество и совместную работу, минуя границы политико-административного деления. Согласно заявлению управления по реализации программы FEA (FEARMO) Методология FEA включает в себя набор взаимосвязанных «эталонных моделей», предназначенных для упрощения анали-

за деятельности агентств и выявления дублирующих инвестиций, несоответствий и возможностей для совместной работы внутри агентств и между ними. Совместно эталонные модели образуют структуру для единообразного описания важных элементов методологии FEA¹.

Эталонная модель бизнеса (BRM) дает бизнес-представление различных функций федерального правительства. Например, в этой модели определяется стандартная функция бизнеса использование водных ресурсов, которая, в свою очередь, является функцией природных ресурсов, являющейся критически важной для более широкой области обслуживания населения¹.

Эталонная модель компонентов (CRM) дает ИТ-представление систем, поддерживающих бизнес. Например, в эталонной модели компонентов определяется аналитическая система, упомянутая выше в гипотетическом описании взаимодействия между Налоговым управлением и Управлением правительственной печати¹.

В Технической эталонной модели (TRM) определяются различные технологии и стандарты, используемые при построении ИТ-систем. Например, в этой модели HTTP определяется как протокол, который является подмножеством служебного транспорта, который, в свою очередь, является подмножеством служебного доступа и доставки¹.

В эталонной модели данных (DRM) определяются стандартные способы описания данных. Например, сущность в этой модели определяется как нечто, обладающее атрибутами и участвующее в отношениях².

В эталонной модели производительности (PRM) определяются стандартные способы описания полезности, обеспечиваемой архитектурами предприятий. Например, качество в этой модели определяется как область измерений, которая, в свою очередь, определяется как «степень соответствия технологии требованиям к функциональности или возможностям»³.

Процесс FEA

Процесс FEA в основном направлен на создание архитектуры сегмента для подмножества общей архитектуры предприятия (в случае с FEA предприятием является федеральное правительство, а подмножеством — правительственное агентство). Описание процесса приве-

¹ «Практическое руководство по FEA». – Декабрь 2006 г. – Опубликовано FEAPMO, Административно-бюджетное управление.

² «Эталонная модель данных, версия 2.0». – Ноябрь 2005 г. – Опубликовано FEAPMO, Административно-бюджетное управление.

³ «Документация по эталонным моделям FEA, версия 2.1». – Декабрь 2006 г. – Опубликовано FEAPMO, Административно-бюджетное управление.

дено в «Практическом руководстве по FEA»¹. Сегменты предприятия в рамках методологии FEA были рассмотрены выше. Общий процесс разработки архитектуры сегмента (на самом верхнем уровне) выглядит следующим образом.

Этап 1. Анализ архитектуры: формирование простого и лаконичного представления сегмента с привязкой к плану организации.

Этап 2. Архитектурное определение: задание желаемого состояния сегмента, документация целевых показателей производительности, рассмотрение альтернатив и разработка архитектуры предприятия для сегмента, в том числе архитектуры бизнеса, архитектуры данных, архитектуры служб и технологической архитектуры.

Этап 3. Стратегия инвестиций и финансирования: рассмотрение способов финансирования проекта.

Этап 4. План управления программой и реализация проектов: создание плана управления проектом и его реализации, включающего контрольные точки и показатели производительности для оценки успешности проекта.

Оценка успешности в методологии FEA

Структура FEA для оценки успеха в использовании архитектуры предприятия описана в документе «Структура оценки архитектуры предприятия по программе FEA 2.1»². Уровень готовности федеральных агентств оценивается по трем категориям:

1. Завершенность архитектуры – уровень готовности собственно архитектуры.
2. Использование архитектуры – эффективность использования агентством архитектуры при принятии решений.
3. Результаты использования архитектуры – преимущества, достигнутые благодаря использованию архитектуры.

П.8. Методология Gartner

До сих пор речь шла о методологиях, объединенных под знаменем архитектуры предприятия.

Несмотря на весьма продуктивную деятельность федерального правительства по разработке архитектуры предприятия прогресс практически отсутствовал. В 2004 г., восемь лет спустя после принятия акта Клиндера – Коэна, предписывающего использовать эффективные

¹ «Практическое руководство по FEA». – Декабрь 2006 г. – Опубликовано FEAPMO, Административно-бюджетное управление.

² Структура оценки архитектуры предприятия по программе FEA 2.1. – Декабрь 2006 г.

процессы планирования ИТ, Бюджетно-контрольное управление представило отчет о внедрении идеи архитектуры. С января 2005 г. Бюджетно-контрольное управление серьезно наказало ряд правительственных агентств за невыполнение предписаний по внедрению и использованию архитектуры предприятия.

В качестве примера можно привести ФБР¹, Министерство обороны, Министерство национальной безопасности² и НАСА³.

В 2005 г., примерно в то же время, когда Административно-бюджетное управление стало доминировать в области разработки архитектуры предприятия в государственном секторе экономики, доминировать в частном секторе экономики стала другая организация – группа Gartner.

Методология Gartner не является ни таксономией (как модель Захмана), ни процессом (как TOGAF), ни полной методологией (как FEА). Эта методология представляет собой набор практических рекомендаций по построению архитектуры предприятия от одной из наиболее известных в мире исследовательских и консалтинговых ИТ-организаций – компании Gartner.

Компания Gartner считает, что архитектура предприятия должна начинаться с того, что организация собирается достичь, а не с текущего положения дел. После того как в организации будет сформировано единое представление о будущем, можно рассматривать влияние этого представления на архитектуру бизнеса, технологическую архитектуру, информационную архитектуру и архитектуру решений.

В 2006 году сформулированы рекомендации, касающиеся разработки архитектуры в виде последовательности шагов и задач участников, которые, однако, не детализированы до уровня моделей процесса разработки архитектуры.

Методика описания архитектуры Gartner представляет собой как бы трехмерный куб, состоящий из следующих элементов (см. рис.

¹ Информационные технологии. ФБР начало разрабатывать архитектуру предприятия, однако многое еще предстоит сделать // GAO-05-363. – 9 сентября 2005 г.

² Модернизация бизнес-систем Министерства обороны США. Необходимо устранить давно существующие недостатки в разработке архитектуры предприятия // GAO-05-702. – 22 июля 2005 г. Министерство национальной безопасности США: Заявление для подкомитетов Конгресса / Рэндольф К. Хайт, директор, проблемы ИТ-архитектуры и ИТ-систем // Бюджетно-контрольное управление. – 29 марта 2006 г.

³ Модернизация бизнеса. Есть прогресс по внедрению рекомендаций Бюджетно-контрольного управления в отношении интегрированной программы управления финансами НАСА // GAO-05-799R. – 9 сентября 2005 г.

П1.15) [20]: горизонтальные слои: бизнес-архитектура (это четыре связанных, взаимозависимых и усложняющихся уровней; вертикальные домены – информационная архитектура (Приложения, Данные, Интеграция, Доступ); вертикальные элементы технической архитектуры (Инфраструктура, Системное управление, Безопасность).



Рис. П1.15. Уровни модели архитектуры Gartner

При этом описанные выше слои бизнес-архитектуры пересекаются со всеми элементами информационной и технической архитектур. В этой схеме верхние два уровня ориентированы на совместное обсуждение с бизнес-руководителями и ИТ-специалистами и в какой-то степени соответствуют тому, что называют бизнес-архитектурой, а нижние два уровня входят во внутреннюю компетенцию ИТ-службы:

- верхний уровень «Среды бизнес-взаимодействия» описывает новую модель «виртуального» бизнеса, а также все, что связано с кооперацией предприятий и бизнесом.
- второй уровень «Стили бизнес-процессов» описывает, как организация выполняет свои ключевые функции, т.е. включает в себя бизнес-процессы предприятия, такие как обработка заказа, мониторинг производственных процессов, анализ использования критически важных ресурсов, совместная работа с информацией;
- следующий уровень «Шаблоны» описывает модели и алгоритмы, которые могут широко использоваться для решения различных задач на предприятии. Примерами шаблонов является использование хранилища данных.
- нижний уровень «Строительные блоки» соответствует технологической архитектуре и включает в себя операционные системы, серверы, базы данных, сами данные и пр.

Этот подход является адекватным с точки зрения того, что он раскрывает руководству механизм влияния решений в области ведения бизнеса на решения в области использования ИТ на предприятии.

Подход компании Gartner к архитектуре предприятия основан на авторитете компании. Вы полагаетесь на то, что в этой компании работают высококвалифицированные специалисты, в ней уже налажена эффективная совместная работа и накоплен большой опыт.

Клиенты Gartner могут найти в ее библиотеке исследовательские материалы, посвященные архитектуре предприятия. Например, такими документами являются «Процесс создания архитектуры предприятия Gartner: развитие, 2005 г.»¹ и «Структура архитектуры предприятия Gartner: развитие, 2005 г.»². Однако эти документы содержат мало описательной информации и, в любом случае, датированы концом 2005 г. Компания Gartner утверждает, что эти рекомендации носят вневременной характер, и продолжает их наращивать по мере необходимости.

Компания Gartner уверена, что архитектура предприятия призвана объединить три группы профессионалов: владельцев бизнеса, ИТ-специалистов и специалистов по внедрению технологий. Если вам удалось объединить эти группы и сформировать у них единое представление о факторах, влияющих на ценность бизнеса, вы победили, если нет – проиграли. Успех оценивается чисто прагматически, например по доходности бизнеса, а не по количеству отмеченных элементов в матрице процесса.

Компания Gartner рекомендует начать работу с написания прогноза о стратегическом направлении развития организации и бизнес-факторах, на которые необходимо реагировать. Прогноз должен быть написан простым языком (соблюдать стандарты документации необязательно), без использования аббревиатур, специальной терминологии и технических рассуждений. Текст должен быть всем понятен и направлен на формирование у всех единого представления.

Большинство организаций сталкиваются с необходимостью внесения в бизнес-процессы существенных изменений. Процесс формирования представления об архитектуре предприятия дает сотрудникам организации шанс собраться вместе, отвлечься от повседневной текучки и убедиться в том, что все понимают природу, область действия и последствия ожидаемых изменений.

После того как в организации будет сформировано единое представление о будущем, можно рассмотреть влияние этого представления на архитектуру бизнеса, технологическую архитектуру, информа-

¹ Процесс создания архитектуры предприятия Gartner: развитие, 2005 г. / Скотт Р. Биттлер и Грег Крейцман. – 21 октября 2005 г. – Код Gartner: G00130849.

² Структура архитектуры предприятия Gartner: развитие, 2005 г. / Грета А Джеймс, Роберт А. Хэндлер, Энн Лапкин и Николас Галл. – 25 октября 2005 г. – Код Gartner: G00130855.

ционную архитектуру и архитектуру решений. Общее представление о будущем определяет изменения, которые необходимо внести во все перечисленные выше архитектуры, приоритеты этих изменений и привязку этих изменений к ценности бизнеса.

Архитектура предприятия, согласно представлению Gartner, связана со стратегией, а не с технической реализацией. Она направлена на достижение цели. Два самых важных вопроса, которыми задается компания Gartner, – это куда организация стремится и как она туда попадет. Любое действие, не связанное напрямую с этими вопросами, считается неуместным. Аналитики Gartner любят употреблять следующую фразу: «Ровно столько архитектуры, сколько необходимо, и точно в срок».

П. 10. Методика META Group

В 2005 году Компания Meta Group. Предложила Методику META Group, отличительной особенностью которой является более детальное и формализованное описание именно процесса разработки архитектуры и всех его составляющих.

META Group отмечает, что архитектура реализуется на практике через процесс управления ИТ-программами и проектами.

Объединяющим для всех доменов архитектуры META Group является процесс формулировки бизнес-требований к ИТ-архитектуре, что оформляется в виде двух документов: «Видения общих требований» и «Принципы концептуальной архитектуры».

Организация процесса разработки архитектуры и создание начальной версии архитектуры предприятия, согласно META Group, состоит в прохождении следующих этапов (рис. П1.16).

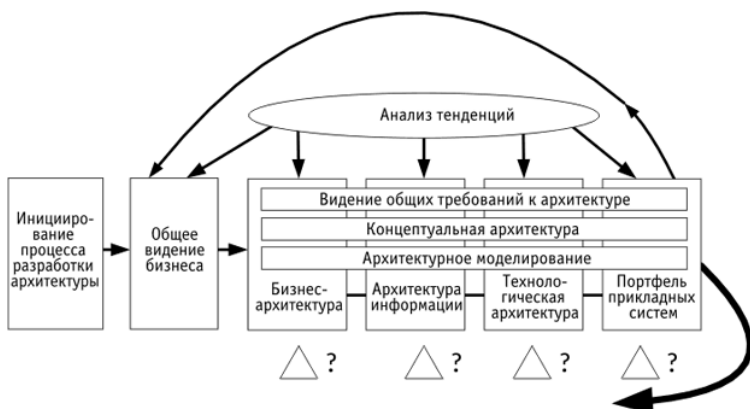


Рис. П1.16. Аналитическая работа и компоненты Архитектуры предприятия META Group

На **этапе 1** разрабатывается Видение общих требований. Разработка Видения общих требований включает в себя:

- анализ тенденций развития внешней для предприятия среды, включая технологические тенденции;
- бизнес-стратегии и основные движущие силы с точки зрения бизнеса;
- требования к информационным системам со стороны бизнеса;
- требования к технологической архитектуре, которая обеспечивает адекватные возможности для информационных систем с точки зрения потребностей бизнеса.

Данная методика предлагает формализованные шаблоны, обеспечивающие разработку Видения общих требований и Концептуальной архитектуры.

Рекомендация относительно Видения общих требований. Главное – это совместное участие представителей бизнес-подразделений и ИТ в выработке общего понимания набора требований, согласованных со стратегическим направлением развития компании. Размер этого документа может быть 10–15 страниц. Основное содержание документа может состоять из четких утверждений, которые тематически связаны между собой, например, как представлено в табл. П1.1.

Таблица П1.1

Пример общих требований

Тенденция	Бизнес-стратегия предприятия	Требования к ИС	Требование к Архитектуре
Задержки в предоставлении услуги затрагивают 20% клиентов	Процесс обслуживания, уменьшающий ожидание клиента, приведет к увеличению доли рынка	Информация о заказах, независимо от канала и места их получения, должна немедленно передаваться в производство	ИТ-инфраструктура должна обеспечивать управляемый доступ и своевременную передачу информации, чтобы обеспечить операционную эффективность

Результатом первого этапа работ могут быть четыре документа:

- список ключевых технологических тенденций;
- список бизнес-стратегий;
- список требований к информационным системам;
- список требований к технологической архитектуре.

Этап 2 состоит в разработке Концептуальной архитектуры, которая определяет логически связанный набор принципов, обеспечиваю-

щий общее руководство для развития информационных систем предприятия и технологической инфраструктуры. Разработка наиболее приоритетных доменов архитектуры, анализ на несоответствие между текущим и желаемым состоянием архитектуры. Каждый домен технологической архитектуры включает описание принципов, технологий, стандартов, продуктов, конфигураций, которые являются многократно используемыми строительными блоками при построении ИТ-систем.

Этап 3 состоит в разработке плана реализации, обеспечивающего переход в сторону желаемого состояния архитектуры.

В полном описании методики META Group приводятся также следующие аспекты:

- практическая реализация архитектуры через процесс управления корпоративными ИТ-программами и проектами;
- вопросы управления и контроля архитектурного процесса;
- оценка зрелости архитектуры;
- анализ технологических тенденций и планирование;
- управление портфелем ИТ-активов и проектов.

Сравнение основных рассмотренных методов построения Архитектуры предприятия приведено в табл. П1.2.

Таблица П1.2

Сравнение методов построения Архитектуры предприятия

Модель	Модель Захмана	3D - модель	FEAF	TOGAF	Методики Gartner	Методики META Group
Характеристика						
Иерархический подход, возможность связи с бизнес-стратегией	+	+	+		+	+
Поддержка различных уровней абстракции	+	+	+	+	+	+
Формальный язык и система обозначений				+		
Описание процесса разработки архитектуры			+	+	+	+
Рекомендации по управлению архитектурой			+	+	+	+

В настоящее время в основном используется одна из четырех методологий:

- Структура Захмана для архитектуры предприятий – эта модель еще не утвердила окончательно свое название.
- TOGAF (The Open Group Architectural Framework) – эта модель называется тем, чем фактически она и является.
- Архитектура федеральной организации – эту модель можно рассматривать по созданию архитектуры предприятия.
- Методология Gartner – эту модель можно описать как набор рекомендаций по созданию архитектуры предприятия.

К 2005 г. компания Gartner стала одной из наиболее влиятельных организаций, занимающихся консалтингом на уровне директоров по информационным технологиям. Однако в области разработки архитектуры предприятия наиболее известной ИТ- и консалтинговой компанией была не Gartner, а Meta Group.

Компания Gartner пыталась создать рекомендации по разработке архитектуры предприятия, однако ей не удалось превзойти Meta Group. В 2005 г. компания Gartner решила, что, раз ей не удастся конкурировать с Meta Group, можно поступить иначе: компания Gartner приобрела компанию Meta Group.

После приобретения Meta Group компания Gartner/Meta потратила год на то, чтобы разобраться, какой вклад внесла каждая из компаний в методологию разработки архитектуры предприятия. Две компании обсуждали, как наилучшим образом объединить свои заметно различающиеся подходы.

В конце концов, был применен достаточно простой алгоритм: если компания Meta Group одобряла какой-либо пункт, он включался в методологию, если не одобряла, соответствующий пункт исключался.

Компании Gartner нравились архитектурные структуры. Компании Meta Group нравился архитектурный процесс. Таким образом, структуры были исключены, а процессы включены.

Один из наиболее важных выводов, которые можно сделать из проведенного в Приложении 1 анализа, заключается в том, что ни одна из методологий не является полной. У каждой из них есть свои достоинства и недостатки.

Этот вывод совпадает с тезисом, сформулированным Мега групп *«в индустрии ИТ нет одного, единственно правильного стандарта на определение архитектуры ИТ, поэтому общие соглашения внутри организации важнее теоретической точности»*.

При выборе вида архитектуры следует руководствоваться рядом критериев.

В частности:

- сложность ИТ-систем, затрудняющая управление ими;
- высокая стоимость текущего сопровождения ИТ-системы, не позволяющие организации своевременно и эффективно реагировать на нынешнюю и будущую конъюнктуру рынка;
- наличие критически важных данных, которые могут оказаться устаревшими либо неверными;
- взаимное недоверие между подразделениями организации, ответственными за ведение бизнеса и технологические аспекты.

Наиболее значимым общим критерием является снижение сложности ИТ-систем и затрат на их сопровождение при увеличении ценности бизнеса и росте эффективности – или, выражаясь более простыми словами, повышение конкурентоспособности организации в условиях все более жесткой конкуренции.

При выборе наиболее подходящей методологии иногда используют следующий подход:

Пройдитесь по строкам (критериям) табл. П1.2 и удалите те критерии, которые не важны для вашей организации.

Добавьте строки (критерии), которые для вас важны, и оцените каждую из методологий по этим критериям.

После этого вы получите хорошее представление о достоинствах и недостатках каждой методологии с точки зрения потребностей вашего предприятия. Если определится однозначный победитель, считайте, что вам повезло. Найдите консультанта, специализирующегося на выбранной вами методологии, и приступайте к работе.

Для многих предприятий ни одна из отдельных методологий не будет полным решением. Таким организациям предлагается использовать подход, при котором создается собственная методология построения архитектуры предприятия, состоящая из наиболее полезных для организации компонентов других методологий или собственных концепций.

Для этого потребуется консультант другого типа – специалист, одинаково хорошо владеющий разными методологиями и специализирующийся на создании эффективных методологий с учетом потребностей и реалий конкретного предприятия.

При выборе методологии в качестве критериев можно руководствоваться некоторыми преимуществами, которые можно получить при успешном внедрении архитектуры предприятия:

1. Более эффективное использование информационных технологий, повышающее приспособляемость бизнеса.
2. Более тесное сотрудничество бизнес- и ИТ-подразделений.
3. Большая ориентированность на цели организации.
4. Высокий моральный дух, поскольку больше сотрудников теперь видят прямую связь между их трудом и успехом организации.
5. Сокращение количества отказов ИТ-систем.
6. Упрощение существующих ИТ-систем.
7. Более высокая адаптируемость новых ИТ-систем.
8. Более тесная связь между ИТ-показателями и бизнес-требованиями.

Очевидно, что организация достигнет большего успеха, чем неэффективная организация. Это верно независимо от того, как оценивается успех: по материальным показателям (например, по доходности или рентабельности), или нематериальным (например, по степени удовлетворенности клиентов или текучести кадров).

Следует иметь в виду, что для большинства организаций однозначного выбора готовой методологии разработки архитектуры не будет. В такой ситуации рекомендуется применить подход, который можно назвать смешанной методологией, основанной на выборе из каждой методологии необходимых разделов и изменении их в соответствии с конкретными потребностями своей организации. Однако даже смешанная методология будет работать только в том случае, если организация готова к изменениям. Такое решение должно быть принято на высшем уровне.

Приложение 2¹

Пример создания многоуровневой информационной системы для обеспечения информационных потребностей студентов

Концепция многоуровневой структуры информационной системы (МСИС) вуза реализована на примере МСИС для обеспечения информационных потребностей конкретной группы пользователей – студентов.

Многоуровневая структура ИС, разработанная для вуза на основе концепции, изложенной в разделе 8.4, представлена на рис. П2.1 ².

Определение потребностей студентов проведено с учетом концепций теории потребностей: иерархии (пирамиды) потребностей **А. Маслоу**, включающей пять групп потребностей (физиологические потребности *physiological needs*; потребности в безопасности, защищенности *need for safety*, потребность в хорошем отношении *need for group*, потребности в уважении, одобрении *respect*); теории мотивации **Ф. Герцберга**, учитывающей влияние материальных и нематериальных факторов на мотивацию человека; концепции **Ф.М. Достоевского**, который делил множество интересов и потребностей людей по усложнению их содержания на три группы (потребности в материальных благах, необходимых для поддержания жизни, потребности познания, всемирного объединения людей); концепции **Г. Гегеля**, который выделял 4 группы (физические потребности, потребности права, законов, религиозные потребности, потребности познания).

Для упорядочения и обеспечения полноты выявления информационных потребностей применены методики системного анализа, рассмотренные в гл. 4: методика структуризации, основанная на учете взаимодействия системы со средой (раздел 4.4) и методика, основанная на концепции системы, стремящейся к идеалу (раздел 4.6).

Структуризация потребностей, полученная на основе комбинирования признаков структуризации названных методик, приведена в верхней части рис. П2.2. На этом рисунке приведены также взаимосвязи между стратами, на основе которых для обеспечения информационных потребности студентов в информационной системе должен

¹ Приложение подготовлено аспиранткой Ю.А. Голуб.

² **Волкова В.Н.** Автоматизированные информационные системы в высшей школе: История и перспективы / В.Н. Волкова, Ю.А. Голуб. – СПб.: Изд-во Политехн. ун-та, 2011. – 112 с.

быть инициирован запуск определенного информационного процесса, в рамках которого используются территориально распределенные информационные массивы – базы данных, массивы научно-технической, нормативно-правовой информации и др.

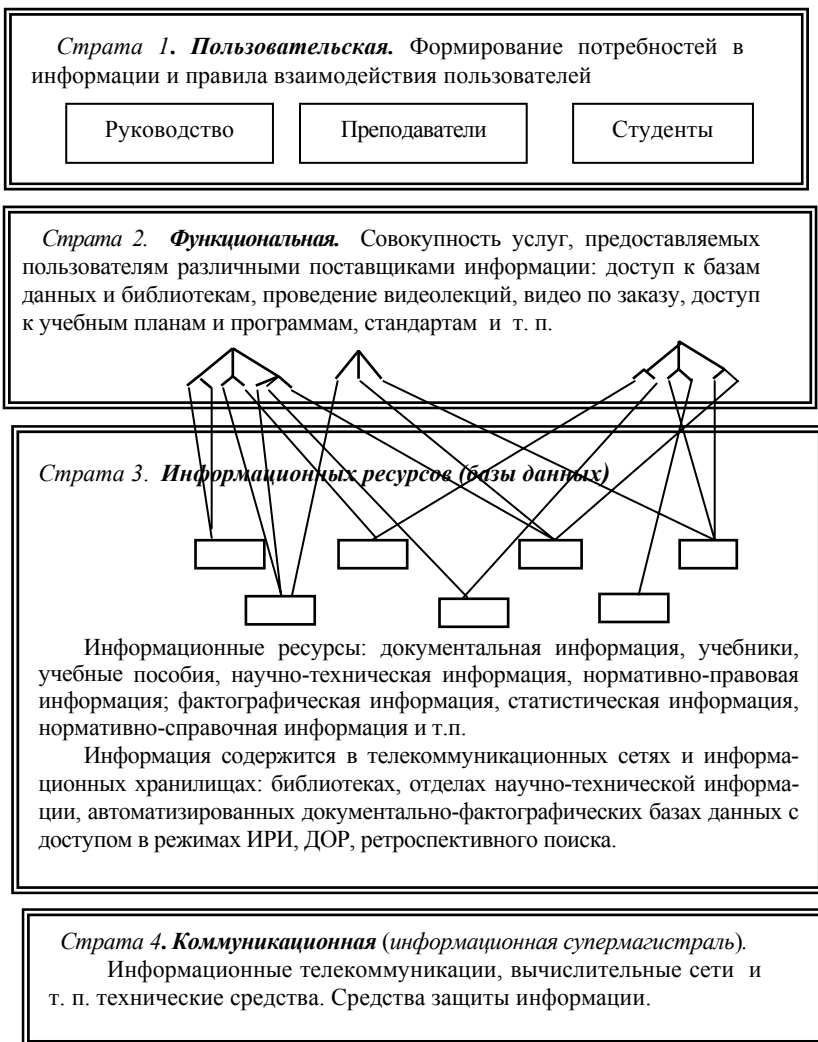


Рис. П2.1. Многоуровневая структура информационной системы

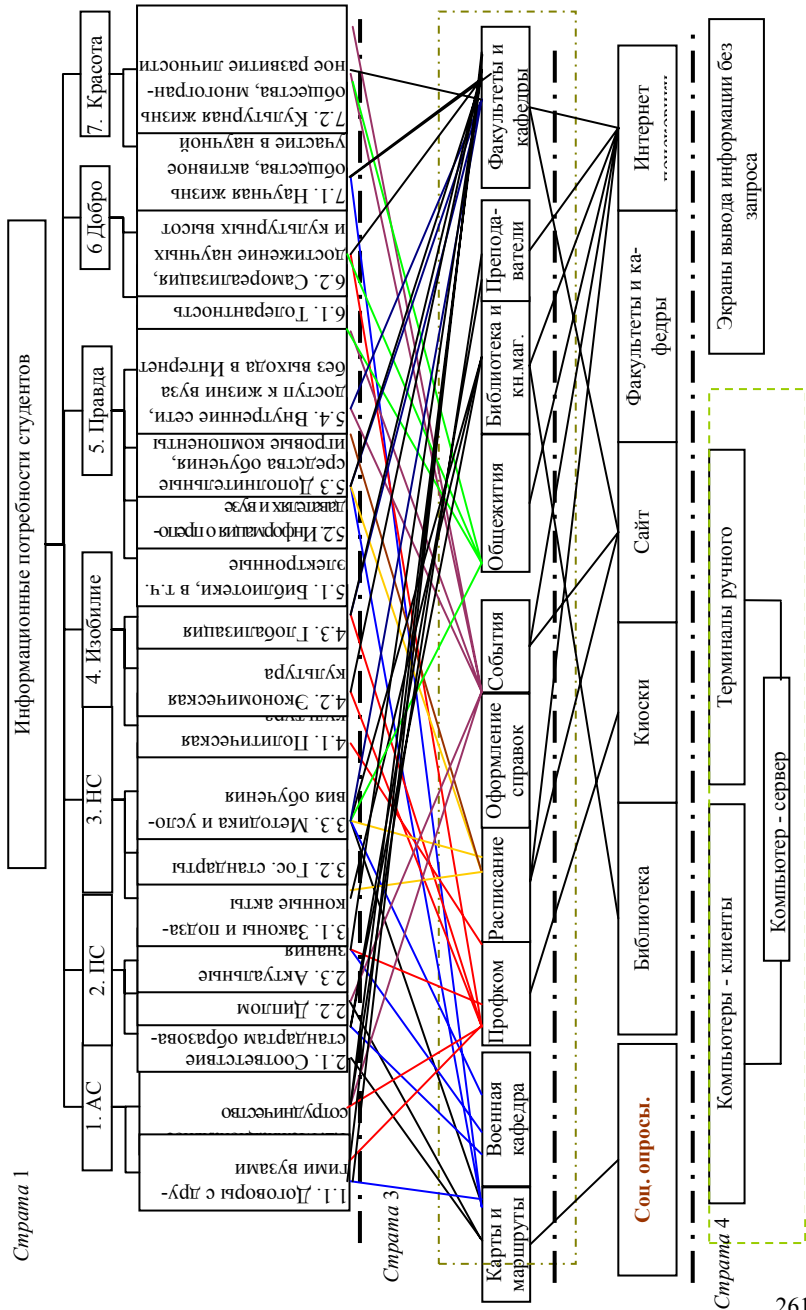


Рис. П2.2. Многоуровневая структура ИС для обеспечения потребностей студентов

Стратифицированное представление, приведенное на рис. П2.1, помогает решать проблему оценки значимости и выбора для реализации компонентов страт для того, чтобы процессы сбора информации при ответе на запрос выполнялись в обозримые сроки. Для оценки компонентов страт разработана методика, основанная на применении методов организации сложных экспертиз.

Методы организации сложных экспертиз – методы и модели, повышающие объективность получения оценок путем расчленения большой первоначальной неопределенности проблемы, предлагаемой эксперту для оценки, на более мелкие, лучше поддающиеся осмыслению.

Четырехстратную структуру, представленную на рис. П2.2, можно сократить до трехстратной (рис. П2.3), пропустив *функциональную страту*, которая практически отсутствует в существующих в настоящее время информационных системах вуза.

Для организации сложной экспертизы применим вначале метод решающих матриц (МРМ), предложенный *Г.С. Поспеловым* как средство стратифицированного расчленения проблемы с большой неопределенностью на подпроблемы и пошагового получения оценок.

С помощью метода решающих матриц можно оценить относительные информационные потребности, затем – компонентов информационной страты их обеспечивающие, а затем – компоненты нижней страты, т.е. технические и программные средства, используемые для создания информационной страты.

Тогда в модели могут быть реализованы следующие страты (сверху вниз): информационные потребности с относительными оценками $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_{na}$; информационные массивы – $\beta_1, \beta_2, \dots, \beta_i, \dots, \beta_{n\beta}$; составляющие (компоненты) информационной супермагистральной – $\gamma_1, \gamma_2, \dots, \gamma_k, \dots, \gamma_{ny}$. (рис. П2.3).

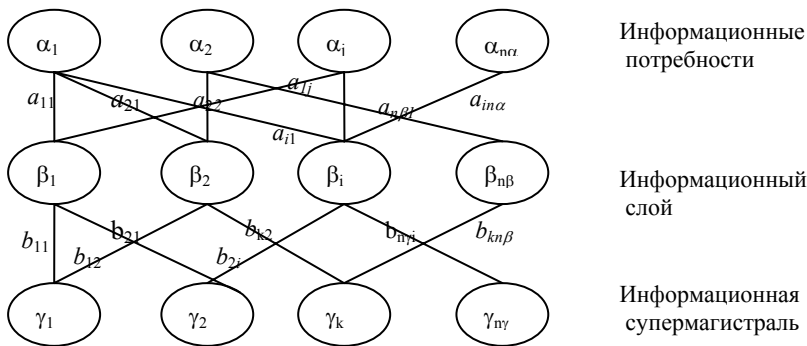


Рис П2.3. Модель решающих матриц для анализа взаимосвязей в МСИС

Каждая строка решающей матрицы характеризует относительную значимость компонентов нижележащего уровня для реализации вышестоящего.

Вначале оценивается относительная значимость α_j компонентов верхнего уровня – информационных потребностей.

Оценка проводится методом нормирования. Проверяется выполнение условия нормирования $\sum_{j=1}^{n\alpha} \alpha_j = 1$.

Затем оценивается a_{ij} – относительная значимость компонентами информационного слоя (баз данных, массивов научно-технической информации и т.п.) для реализации информационных потребностей пользователей.

Проверяется выполнение условия нормирования для каждой j :

$$\sum_i a_{ij} = 1.$$

На основе α_j и a_{ij} вычисляется β_i – относительная значимость компонент информационной страты

$$\beta_i = \sum_j a_{ij} \alpha_j.$$

Проверяется выполнение условия нормирования $\sum_{i=1}^{n\beta} \beta_i = 1$.

Далее оценивается b_{ki} – относительная значимость компонентов информационной супермагистрали (т.е. информационных и программных средств) для реализации информационной страты.

Проверяется выполнение условия нормирования для каждой i

$$\sum_k b_{ki} = 1.$$

Вычисляется γ_k – относительные веса компонент информационной супермагистрали для реализации информационной страты:

$$\gamma_k = \sum_i b_{ki} \beta_i.$$

Проверяется выполнение условия нормирования $\sum_{k=1}^{n\gamma} \gamma_k = 1$.

Алгоритм реализации метода решающих матриц приведен на рис. П2.4.

При оценке технических и программных средств возможно формирование двух отдельных матриц и оценка значимости их компонентов для реализации информационной страты.

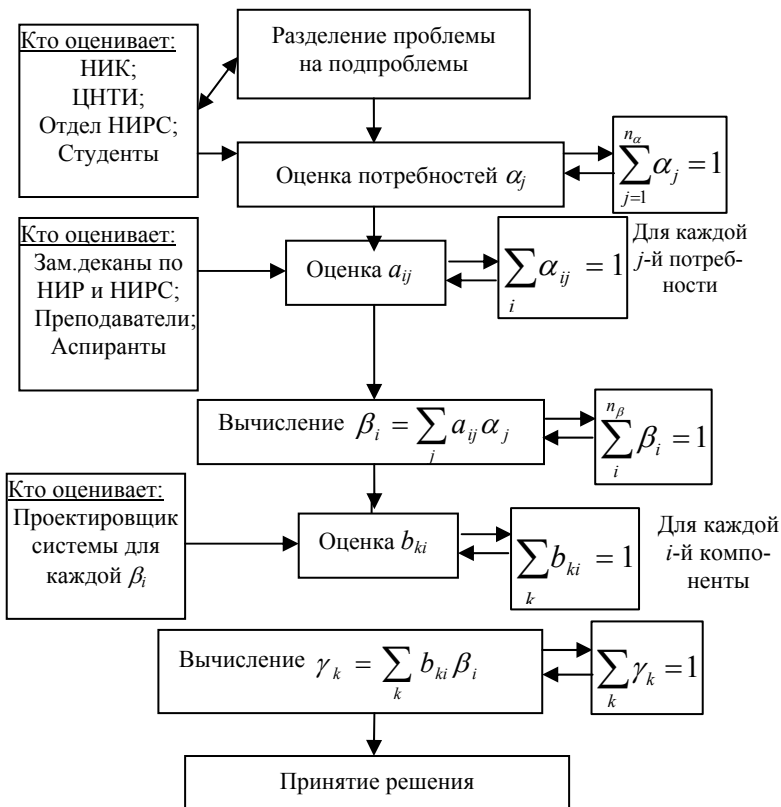


Рис. П2.4. Алгоритм реализации модели, приведенной на рис. П2.3

Еще более объективный анализ можно получить с помощью подхода к разработке модели организации сложных экспертиз, базирующегося на использовании методов структуризации и информационных оценок степени целесообразности исследуемых компонентов, что реализуется на основе многоуровневой оценки влияния компонентов каждой ниже лежащей страты на реализацию компонентов вышестоящей страты.

В основу построения модели положена идея модели решающих матриц, а оценки между уровнями уточняются на основе информационного подхода *А.А. Денисова* [1, 3, 8, 21], позволяющего приводить разнородные критерии (количественные и качественные) к единым информационным единицам, что помогает сопоставлять эти критерии и получать оценки потребности в различных видах услуг, используемых для проведения сравнительного анализа.

В качестве информационной меры, характеризующей значимость компонентов нижележащего уровня на вышестоящий принята предложенная *А.А. Денисовым* [1, 3, 8, 21], мера потенциала H , которая позволяет учесть одновременно два критерия p' и q .

$$H_i = -q_i \log(1 - p_i'),$$

где p_i' – степень влияния нижестоящего уровня на вышестоящий (изменяется в пределах от 0,7 до 0,99); q_i – вероятность выбора и реализации компонентов нижестоящего уровня с точки зрения лиц, принимающих решение при их выборе на вышестоящем уровне. При этом выполняется условие нормирования $\sum_{j=1}^{na} q_j = 1$.

Тогда в модели, сохраняя страты (сверху вниз): информационные потребности с относительными оценками $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_{na}$; информационные массивы – с оценками $\beta_1, \beta_2, \dots, \beta_b, \dots, \beta_{nb}$; составляющие (компоненты) информационной супермагистральной – с оценками $\gamma_1, \gamma_2, \dots, \gamma_k, \dots, \gamma_{ny}$, вводятся оценок между p' и q между стратами (рис. П2.5).

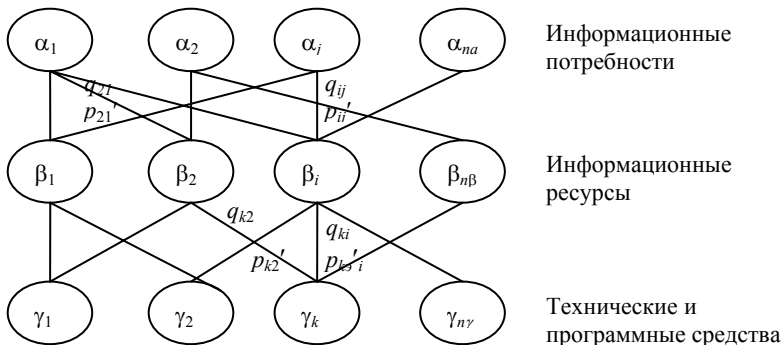


Рис. П2.5. Модель для анализа взаимосвязей в МСИС, основанная на информационных оценках

Оценка p' производится единичными экспертами, для которых определяются сферы компетентности.

Получение обобщенной оценки проводится в соответствии с алгоритмом, приведенным на рис. П2.6, что позволяет получить уточненные оценки каждой компоненты.

Компоненты верхней страты могут быть оценены либо методом нормирования, либо для верхней страты информационная мера степени влияния информационных потребностей может быть оценена как степень влияния p' потребностей пользователей на реализацию целей, т.е. без учета q $H_{\alpha_j} = -\log(1 - p_{\alpha_j}')$.

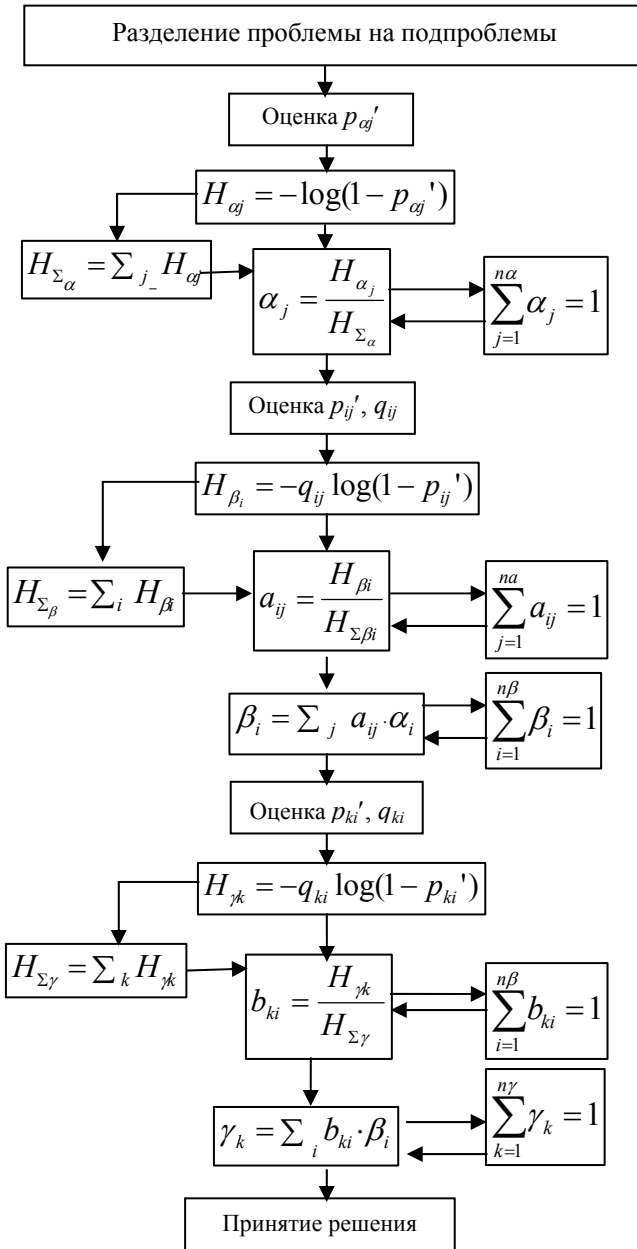


Рис. П2.6 Алгоритм реализации модели рис. П2.5

Для получения значения α_j необходимо выполнить процедуру нормирования, т. е.

$$H_{\Sigma\alpha} = \sum_{j=1}^{n\alpha} H_{\alpha_j}; \quad \alpha_j = \frac{H_{\alpha_j}}{H_{\Sigma\alpha}}.$$

На рис. П2.5. и П2.6. приняты следующие обозначения:

$$\alpha_j = \frac{H_{\alpha_j}}{H_{\Sigma\alpha}} \quad - \quad \text{относительная значимость компонентов верхнего уровня – информационных потребностей.}$$

H_β – информационная мера степени влияния информационного слоя (информационных баз) на реализацию информационных потребностей. $H_{\beta_i} = -q_{ij} \log(1 - p_{ij}')$;

где p_β – степень влияния компонентов информационной страты (информационных баз) на реализацию информационных потребностей;

q_{ij} – вероятность выбора и реализации компонентов информационной страты с точки зрения лиц, принимающих решение при их выборе на вышестоящем уровне.

β_i – относительная значимость компонентов информационной страты $\beta_i = \sum_{j=1}^{n\beta} a_{ij} \alpha_j$;

H_γ – информационная мера степени влияния компонентов информационной супермагистрали (технических средств) на реализацию информационных баз. $H_{\gamma_k} = -q_{ki} \log(1 - p_{ki}')$;

где p_{ki}' – степень влияния компонентов информационной супермагистрали (технических средств) на реализацию информационных баз;

q_{ki} – относительная значимость компонентов информационной супермагистрали на реализацию информационных баз.

$$b_{ki} = \frac{H_{\gamma_k}}{H_{\Sigma\gamma_k}} \quad - \quad \text{относительная значимость компонентов супермагистрали для реализации компонентов информационной страты;}$$

γ_k – относительные веса компонент информационной супермагистрали. $\gamma_k = \sum_{i=1}^{n\gamma} b_{ki} \cdot \beta_i$.

Рассматриваемую модель можно представить в обобщенном формализованном виде.

$$\beta_i = \sum_j a_{ij} \cdot \alpha_j = \sum_j \left(\frac{H_{\beta_i}}{H_{\Sigma\beta_i}} \cdot \frac{H_{\alpha_j}}{H_{\Sigma\alpha}} \right);$$

$$\gamma_k = \sum_i b_{ki} \cdot \beta_i = \sum_i \frac{H_{\gamma_k}}{H_{\Sigma\gamma_k}} \left[\sum_j \left(\frac{H_{\beta_i}}{H_{\Sigma\beta_i}} \cdot \frac{H_{\alpha_j}}{H_{\Sigma\alpha}} \right) \right];$$

$$\gamma_k = \sum_i \left[\frac{q_{ki} \log(1 - p_{ki}')}{\sum_i q_{ki} \log(1 - p_{ki}')} \times \sum_j \left(\frac{q_{ij} \log(1 - p_{ij}')}{\sum_i q_{ij} \log(1 - p_{ij}')} \cdot \frac{\log(1 - p_{\alpha_j}')}{\sum_j \log(1 - p_{\alpha_j}')} \right) \right].$$

Для реализация рассмотренных методов организации сложных экспертиз разработаны автоматизированные диалоговые процедуры [3].

Экспериментальные исследования рассмотренных моделей проведены на примере разработки пилотного проекта информационной системы для обеспечения информационных потребностей студентов. Оценивались компоненты стратифицированной структуры, приведенной на рис. П2.2.

Согласно приведенным на рис. П2.2 потребностям студентов и полученных оценок в информационно-поисковой системе «Политех» предусмотрены следующие функции и информационные массивы:

- географическое расположение всех корпусов СПбГПУ (Политех);
- сведения о факультетах и кафедрах, с указанием географического положения;
 - данные о преподавателях;
 - единый центр расписания с указанием аудиторий и преподавателей;
 - сведения о конкурсах, конференциях;
 - сведения о культурных мероприятиях;
 - база данных об участии студентов в НИРС и итогах НИРС;
 - справочная система для получения информации о профкоме, военной кафедре, порядке оформления обходного листа, документов для налоговой инспекции и т.п.;
 - справочная система по общежитиям (следует реализовать отдельным модулем)
 - справочная библиотечная система (уже реализовано и функционирует отдельным приложением).

Разработка ведется на платформе Java (J2EE). Система представляет собой веб-приложение с многопользовательским доступом; имеет веб-интерфейс, т.е. запускается при помощи Интернет-обозревателя (Internet Explorer, Opera и др.). Реализована на основе сервера приложений Apache Tomcat и базы данных MySQL. Реализация в виде приложения с «тонким клиентом», т.е. для доступа к системе пользователь не нуждается ни в каких клиентских приложениях или предустановках помимо интернет-обозревателя. Система поддерживает многопользовательский режим.

Система предоставляет доступ к следующим категориям информации через поиск и переходы по иерархическим ссылкам (рис. П2.7): 1. Факультеты (адрес, контактные телефоны, декан, список кафедр и др.). 2. Кафедры (адрес, контакты и т.п., список преподавателей. *Возможно: дисциплины, читаемые преподавателями кафедр*). 3. Преподаватели (ФИО, контактная информация, научные интересы и т.п.; *В перспективе, индивидуальные расписания преподавателей*). 4. События/мероприятия (конференции, спорт, гостевые лекции, конкурсы, концерты и др.). 5. Расписание занятий. Составляется для каждой группы, информация хранится в базе данных; с возможностью редактирования (при наличии прав доступа) непосредственно через основной интерфейс. 6. *В перспективе – лента новостей*. 7. Карты (расположение корпусов, схемы проездов и др.). 8. Информация о военной кафедре. 9. Информация о профкоме. 10. Оформление справок (информация о том, где и в какое время, можно получить ту или иную справку или выполнить любую другую формальную процедуру). 11. Общежитие (основная информация, правила проживания, контакты). 12. Библиотека (контакты). 13. Книжные магазины (адреса, контакты).

Система предоставляет возможность добавлять, редактировать и удалять информацию по всем основным информационным категориям через основной интерфейс (через интернет-обозреватель).

В системе осуществлен метод контроля доступа на основе ролевого разграничение доступа (англ. *Role Based Access Control, RBAC*) – развитие политики дискреционного доступа, при этом права доступа субъектов системы на объекты группируются с учетом специфики их применения, образуя роли администратор, менеджер, наблюдатель).

Стратифицированное представление, приведенное на рис. П2.2, помогает решать проблему управления проектами и программами развития МСИС (в том числе предлагаемыми студентами) на основе их оценки с целью распределения финансовых, материальных и кадровых ресурсов.

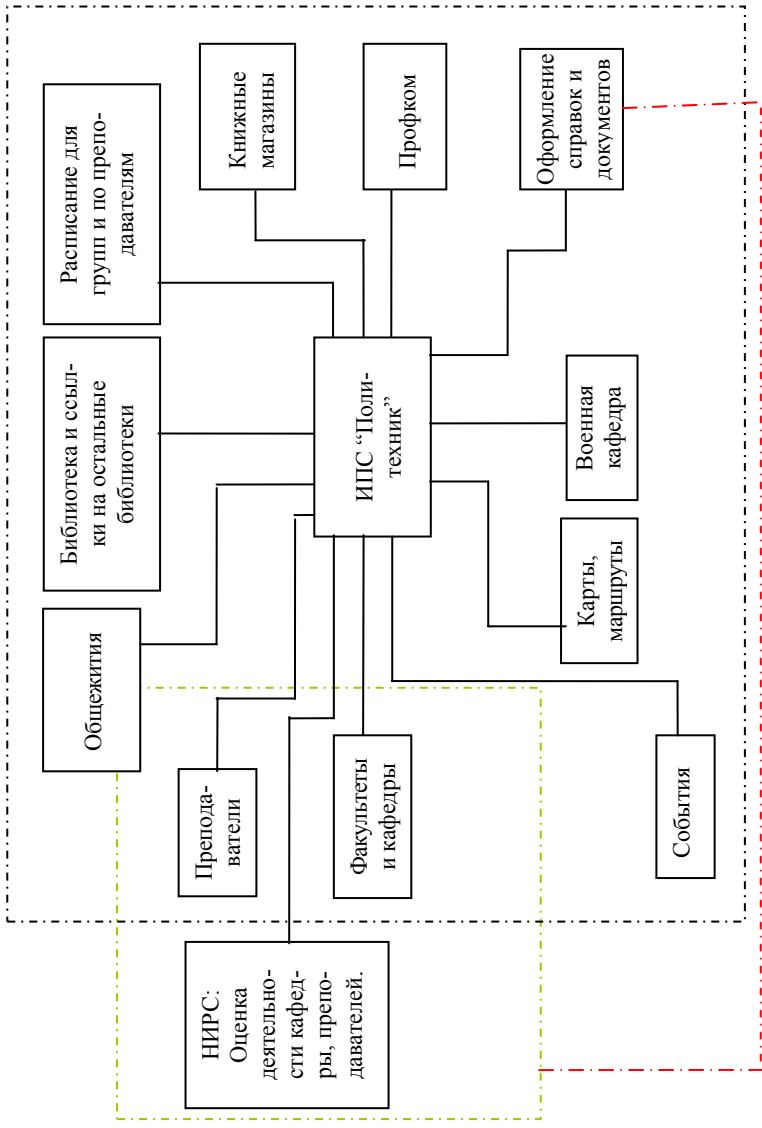


Рис. П2.7

ЗАКЛЮЧЕНИЕ

Как было сказано во введении становление любой науки проходит 4 стадии. Интерпретируем состояние реализации этих стадий для становления теории информационных систем как научного направления.

1. Определение места теории информационных систем среди других научных направлений. Систематизация и классификация ИС.

На основе приведенных в гл.гл. 1 и 3 сведений о структуризации информатики, определении места в этой теории и классификации информационных систем можно сделать вывод, что первую стадию теория информационных систем как наука прошла.

В то же время представляется, что принятую на рис. 3.1 классификацию ИС можно в свою очередь структурировать более детально. Развития и постоянной корректировки требует классификация ИС специального назначения, поскольку в настоящее время информационные технологии активно развиваются и на их основе создаются новые программные продукты и платформы.

2 Углубленное изучение информационных систем, их элементов, структуры и принципов функционирования.

В структуре информатики выделены информационные элементы, информационные процессы и *информационные системы*.

Виды, структуры, возможности **информационных элементов** (включая формы регистрации и хранения информации, технические средства сбора, передачи, хранения, обработки, представления информации), в настоящее время достаточно хорошо изучены и представлены в литературе.

Рассматриваются во многих монографиях и учебных пособиях, изучаются в различных дисциплинах – «Вычислительные системы, сети и телекоммуникации», «Операционные системы», «Базы данных», «Программная инженерия», «Алгоритмизация и программирование» и т. п.

Для исследования **информационных процессов** активизировалось применение предложенного в 1970-е гг. функционально-технологического (процессного) подхода, реализуемого на основе предложенной в 1990-е гг. методологии SADT и семейства IDEF-технологий, которые служат полезным инструментом для анализа процессов в информационных системах.

Изучаются специальными дисциплинами – «Информационные системы и технологии», «Проектирование информационных систем» «Моделирование бизнес-процессов».

Кроме того, информационные процессы как процессы возникновения и распространения научно-технической информации исследуются на основе закономерностей информетрии, что рассматривается в гл. 7.

В данной монографии рассмотрены основные сведения об **информационных системах**: их виды и особенности построения и проектирования, теоретические основы разработки АИС и АСУ предприятия (организации), включая принципы формирования структуры функцио-

нальной и обеспечивающей частей, принципы выбора готовых программных продуктов (гл. 4); приведены основные сведения о документальных информационно-поисковых системах (ДИПС), включая понятие об информационном поиске и информационно-поисковой системе, информационно-поисковом языке, системе индексирования, логике ИПС, критериях смыслового соответствия, критериях оценки ИПС (гл. 5); дано представление о документально-фактографических информационно-поисковых системах (ДФИПС), о системе нормативно-методического обеспечения управления (СНМОУ) предприятия (организации) и ее автоматизированном варианте (АСНМОУ) как информационной системе типа ДФИПС (гл. 6).

3. Формирование научной методологии и методик исследования информационных систем и их компонентов, использование методов точных наук, формальных методов и моделей.

К настоящему времени в информатике и теории информационных систем получен ряд методологических и методических результатов:

- в 1970-е гг. для управления разработками АСУ и ГАСНТИ подготовлены Общеотраслевые методические материалы, регламентирующие разработку АИС и АСУ, Единый порядок и технические условия разработки и внедрения автоматизированных подсистем обработки, поиска, хранения, выдачи и передачи информации, Техническое задание на совершенствование Единой системы научно-технической информации в стране, которые можно считать стандартами того периода;

- в 1990-е гг. разработана методология SADT – (Structured Analysis and Design Technique) – структурный анализ и проектирование, представляющая собой совокупность методов, правил и процедур, предназначенных для построения функциональной модели объекта какой-либо предметной области. Компьютерная реализация методологии SADT получила название IDEF (Icam DEFinition или Integrated DEFinition). Для реализации моделей применяются автоматизированные средства – BPWin, ARIS, язык UML (Unified Modeling Language – Унифицированный язык моделирования);

- для управления ИТ-проектами корпоративных информационных систем предприятий разработаны разнообразные многочисленные стандарты (см. их перечень в [19, с 338–340]);

- разработаны подходы к проектированию систем и методики проектирования и выбора информационных систем для конкретных организаций (гл. 5 и анализ подходов в разделе 8.3); при разработке ИС применяют расчетные модели оценки их эффективности;

- при создании автоматизированных систем научно-технической информации разработана теория информационного поиска и документальных информационно-поисковых систем (гл. 5);

- при создании систем нормативно-методического обеспечения управления разработан подход, принципы создания многоконтурных автоматизированных систем нормативно-методического обеспечения управления (гл. 6);

- исследованы *закономерности информетрии* (гл. 7), базирующиеся на применении статистических методов и помогающие в исследовании информационных потоков и совершенствовании процесса информационного поиска;

- при разработке структуры обеспечивающей части АИС применен подход постепенной формализации модели, основанный на применении методов дискретной математики для доказательства адекватности формализованных модели;

- разработана концепция многоуровневой информационной системы (раздел 8.5);

- для исследования взаимосвязей между стратами многоуровневой информационной системы предложено использовать формальные модели, основанные на применении методов организации сложных экспертиз, базирующихся на идее метода решающих матриц *Г.С. Поспелова* и информационном подходе к анализу систем *А.А. Денисова* (Приложение 2).

Таким образом, в теории информационных систем применяются статистические методы, некоторые расчетные модели, методы дискретной математики, модели организации сложных экспертиз, основанные на формализованных методах.

В то же время формальные математические методы и оптимизационные модели в современной информатике и теории информационных систем применяются недостаточно, т. е. эту стадию развития теории информационных систем еще нельзя считать завершенной.

4. Развитие теоретического компонента – выработка понятий, категорий, методологических установок, создание теоретических концепций, формирование категориального аппарата.

В современной литературе существует множество определений информатики и информационных систем. Это происходит оттого, что данная область знаний относительно новая и соответствующий понятийный аппарат не совсем устоялся.

В процессе разработке АСУ и АИС сформировалась определенная терминология (гл. 4). При создании автоматизированных систем научно-технической информации разработана теория информационного поиска, теория документальных и документально-фактографических информационно-поисковых систем, введен ряд специальных понятий (ПОД и ПОЗ, парадигматические и синтагматические отношения, критерии смыслового соответствия, критерии оценки – релевантность, пертинентность и т.д. (гл.гл. 5 и 6). При разработке специализированных информационных систем применяется терминология, удобная для конкретных приложений.

Таким образом, к настоящему времени сложились комплексы понятий, категорий, методологических установок, теоретических концепций, ориентированных на разработку и эксплуатацию информационных систем разного вида и назначения. Подготовлены словари-справочник [18], учебники и учебные пособия [7, 10–12, 22, 33 и др.].

В то же время обобщающие понятия категориального аппарата теории систем, обеспечивающие взаимопонимание специалистов различных областей знаний, в независимом складывающихся теориях информационных систем разных видов трактуются неодинаково.

Например, понятие «цель».

Применительно к информационным системам *в технических* приложениях трактуется как требования к проектируемой системе.

В *социально-экономических организациях* цель трактуется и как *назначение, миссия* (для внешней среды), и как *цели и функции* организации, и как *потребности* ее сотрудников. Для того, чтобы индивидуальные потребности были подчинены целям и задачам организации, в теории автоматизированных систем управления введено понятие *функциональной части АСУ*, что позволило проще решить проблему целей и потребностей. Структура ФЧ АСУ формируется на основе анализа целей и функций организации и включает подсистемы и задачи, подлежащие автоматизации в соответствующий период планирования работ по созданию АСУ (для удобства вводилось понятие первой, второй и т. п. очереди АСУ).

В теории *научно-технической информации, документальных ИПС* вместо понятия «цель» вводится понятие «*информационные потребности*», которое первоначально было довольно жестко «привязано» к индивидуальным потребностям, но в последующем стало исследоваться применительно к потребностям организаций.

Различные названия присваивают неделимым **элементам** или **компонентам** системы и **способам их отображения**:

в АИС – *базы данных и кодирование*; в документальных ИПС – *информационные массивы и индексирование* с помощью информационно-поискового языка в форме ПОД и ПОЗ.

По разному называют и исследуют **информационные процессы**: *документальные информационные потоки* в документальных системах НТИ, *бизнес-процессы* в АИС и АСУ, *организационно-технологические процедуры* подготовки и реализации управленческих решений в АСНМОУ.

Неоднозначно используется и понятие «**структура**».

Если в системах научно-технической информации сохраняется понятие *структуры* ИПС, то для объединения разнообразных баз данных, используемых для управления предприятием, введено понятие *архитектуры предприятия, архитектуры предложений, архитектуры данных, технологической архитектуры*, что позволяет кратко охарактеризовать структуры различных страт многоуровневой ИС – страты данных, программных приложений, клиент-серверов, технологической страты, ИС предприятия в целом (архитектура предприятия). При этом понятия *структуры* и *архитектуры* используются

неоднозначно: либо понятие «структура» – конфигурацию ИС («структура архитектуры» у Захмана), либо существуют точка зрения, в соответствии с которой под архитектурой понимается описание системы с точки зрения конечных пользователей и интерфейсов взаимодействия с внешней средой, т.е. как внешний взгляд на ИС, а структуру ИС описывают в виде взаимодействующих между собой подсистем, т.е. как внутреннюю конфигурацию ИС. При этом к структура представляется в форме древовидной или стратифицированной иерархии, и полученные структурные единицы, в свою очередь, могут быть представлены в виде архитектурного описания по отношению к внешним структурным единицам. Благодаря возросшим возможностям вычислительной техники, и особенно в результате появления персональных ЭВМ, первоначально развивающиеся самостоятельно *теории построения и функционирования АИС* и *теории поиска в документальных ИПС* сближаются, происходит взаимное проникновение понятий, принципов и теоретических положений.

В частности, по мере роста объемов и усложнения структуры АИС возникает необходимость в заимствовании понятий теории информационного поиска, разработанной ранее для систем НТИ (в частности, понятия *тезауруса*, *многоконтурности* ИПС и т.п.).

В свою очередь, в решении проблемы разработки ИПЯ с высокими смысловыражающими возможностями может помочь структуризация, которая исходно была основой классификации и кодирования информации в АИС и АСУ, а при разработке систем НТИ могут оказаться полезными понятия *функциональной* и *обеспечивающей* частей, введенные в нормативных документах по разработке АИС и АСУ.

Для развития теории информационных систем как науки в гл. 8 предложены концепции, базирующиеся на исследовании *закономерностей информетрии* и их применения; в гл. 8 проведен анализ концепции *архитектуры предприятия*, концепции *ситуационных центров*, возможностей применения для развития теории информационных систем подходов, принципов и моделей *теории систем* и *системного анализа*, и в частности, проведен анализ подходов к проектированию систем и возможности их реализации при создании информационных систем; анализ определений системы, предложено обобщенное определение информационной системы, реализующее системно-целевой подход, и на этой основе – *концепция многоуровневой информационной системы*.

Важным направлением развития теории информационных систем является развитие *методов семантического анализа информации*, которые исследуются в самостоятельно развивающихся направлениях теории информационных систем – теориях *интеллектуальных систем* и *анализа текстов*. Представляется также перспективным продолжить исследование семантики текстовой информации, начатые в 1960–1970-е гг. на базе математической лингвистики и приостановленные развитием средств сети Интернет, обеспечивающих поиск избыточной информации, но не решающих проблему помощи пользователю в ее семантическом анализе.

СПИСОК ЛИТЕРАТУРЫ

1. **Волкова В.Н.** Теория систем и системный анализ: учебник / В.Н. Волкова, А.А. Денисов. – М.: Изд-во Юрайт, 2012. – 579 с.
2. **Волкова В.Н.** Теоретические основы информатики: учеб. пособие / В.Н. Волкова, А.В. Логинова. – СПб.: Изд-во Политехн. ун-та, 2011. – 160 с.
3. **Волкова В.Н.** Методы организации сложных экспертиз: учеб. пособие / В.Н. Волкова, А.А. Денисов. – СПб.: Изд-во Политехн. ун-та, 2010. – 128 с.
4. **Волкова В.Н.** Постепенная формализация моделей принятия решений / В.Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2006. – 120 с.
5. **Горькова В.И.** Информетрия: Количественные методы в научно-технической информации / В.И. Горькова // Итоги науки и техники. Сер. Информатика. Т. 10. – М.: ВИНТИ, 1988. – 328 с.
6. **Данилин А.** Архитектура и Стратегия. «Инь» и «Янь» информационных технологий / А. Данилин, А. Слюсаренко. – М.: Интернет-университет информационных технологий – ИНТУИТ.ру, 2005 // razum.ru/komp/59613-danilin-a-slyusarenko-a.
7. **Данильченко И.А.** Автоматизированные системы управления предприятием: учебник для инженерных специальностей вузов / И.А. Данильченко, В.А. Мясников, В.Н. Четвериков. – М.: Машиностроение, 1984. – 360 с.
8. **Денисов А.А.** Современные проблемы системного анализа: учебник. / А.А. Денисов. – 3-е изд., перераб. и доп. – СПб.: Изд-во Политехн. ун-та, 2008. – 304 с.
9. **Информационно-аналитические средства поддержки принятия решений и ситуационные центры /** Н.И. Ильин, Н.Н. Демидов, А.Н. Данчул и др.; под ред. А. Н. Данчула. – М.: Изд-во РАГС, 2006. – 326 с.
10. **Информационные системы:** учеб. пособие / Под ред. В.Н. Волковой и Б.И. Кузина. – СПб.: Изд-во СПбГТУ, 1998. – 213 с.
11. **Козлов В.Н.** Теория информационных систем: учеб. пособие / В.Н. Козлов, В.Ф. Кисоржевский, Б.И. Морозов. – СПб.: Изд-во СПбГПУ, 2003 – 210 с.
12. **Макарова Н.В.** Информатика: учебник для вузов / Н.В. Макарова, В.Б. Волков. – СПб.: Питер, 2011. – 576 с.
13. **Максимович Г.Ю.** Информационные системы / Г.Ю. Максимович, А.Г. Романенко, О.Ф. Самойлок. – М.: Изд-во Рос. экон. акад., 1999. – 198 с.
14. **Мидоу Ч.** Анализ информационно-поисковых систем / Ч. Мидоу. – М.: Мир, 1970. – 368 с.
15. **Михайлов А.И.** Основы информатики / А.И. Михайлов, А.И. Черный, Р.С. Гиляревский. – М.: Наука, 1968. – 756 с.
16. **Михайлов А.И.** Научные коммуникации и информатика / А.И. Михайлов, А.И. Черный, Р.С. Гиляревский. – М.: Наука, 1976. – 435 с.
17. **Паклин Н.** Бизнес-аналитика: от данных к знаниям / Н. Паклин, В. Орешков. – СПб.: Питер, 2009. – 624 с.
18. **Прикладная информатика:** Справочник: учеб. пособие / Под ред. В. Н. Волковой и В.Н. Юрьева. – М.: Финансы и статистика, ИНФРА-М, 2008. – 768 с.

19. **Сатунина А.Е.** Управление проектами корпоративной информационной системы предприятия: учеб. пособие / А.Е. Сатунина, Л.А. Сысоева. – М.: Финансы и статистика, ИНФРА–М, 2009. – 352 с.
20. **Сеинс Р.** Сравнение четырех ведущих методологий построения Архитектуры / Роджер Сеинс // <http://msdn.microsoft.com/ru-ru/library/ee914379.aspx> Май 2007.
21. **Системный анализ** и принятие решений / Словарь-справочник: учеб. пособие / Под ред. В.Н. Волковой и В.Н. Козлова. – М.: Высшая школа, 2004. – 616 с.
22. **Смирнова Г.Н.** Проектирование экономических информационных систем: учебник / Г.Н. Смирнова, А.А. Сорокин, Ю.Ф. Тельнов / Под ред. Ю.Ф. Тельнова. – М.: Финансы и статистика, 2005. – 512 с.
23. **Сэлтон Г.** Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. – М.: Сов. Радио, 1973. – 560 с.
24. **Тельнов Ю.Ф.** Реинжиниринг бизнес-процессов / Ю.Ф. Тельнов. – М.: Финансы и статистика, 2004. – 320 с.
25. **Теория систем и системный анализ** в управлении организациями: Справочник: учеб. пособие / Под ред. В.Н. Волковой и А.А. Емельянова. – М.: Финансы и статистика, ИНФРА-М, 2009. – 848 с.
26. **Черемных С.В.** Структурный анализ систем: IDEF-технологии / С.В. Черемных, И.О. Семенов, В.С. Ручкин. – М.: Финансы и статистика, 2003. – 208 с.
27. **Черный А.И.** Введение в теорию информационного поиска / А.И. Черный. – М.: Наука, 1975. – 298 с.
28. **Чудесова Г.П.** Преобразование организационной структуры при изменении формы собственности предприятия / Г.П. Чудесова. – СПб.: Изд-во СПбГТУ, 1995. – 190 с.
29. **Шемакин Ю.И.** Тезаурус в автоматизированных системах управления и обработка информации / Ю.И. Шемакин. – М.: Воениздат, 1974. – 172 с.
30. **Шемакин Ю.И.** Компьютерная семантика / Ю.И. Шемакин, А.А. Романов. – М.: Научно-образовательный центр «Школа Китайгородской», 1995. – 343 с.
31. **Шрайберг Я.Л.** Автоматизированные библиотечно-информационные системы России: состояние, выбор, внедрение, развитие / Я.Л. Шрайберг, В.С. Воройский. – М.: Изд-во «Либерия», 1996. – 271 с.
32. **Шрейдер Ю.А.** Логика знаковых систем / Ю.А. Шрейдер. – М.: Знание, 1974. – 43 с.
33. **Юрьев В.Н.** Информационные системы в экономике: учебник / В.Н. Юрьев, В.Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2006. – 538 с.
34. **Юсупов Р.М.** Научно-методологические основы информатизации / Р.М. Юсупов, В.П. Заболотский. – СПб.: Наука, 2002. – 455 с.

СОДЕРЖАНИЕ

Введение	3
Глава 1. Понятие об информатике как научном направлении	5
1.1. История развития термина «информатика»	5
1.2. Информатика – наука об информационных элементах, информационных процессах и информационных системах	12
Глава 2. Понятие об информации и информационных ресурсах предприятий и организаций	21
2.1. Информация: понятие и основные свойства	21
2.2. Подходы к измерению и оценке ценности информации	33
2.3. Понятие об информации как ресурсе предприятия (организации)	40
Глава 3. Понятие об информационных системах и классификации ИС. ...	45
3.1. Из истории разработки информационных систем	45
3.2. Виды и классификации информационных систем	48
Глава 4. Теоретические основы построения и проектирования АИС и АСУ	56
4.1. Принципы построения АИС и АСУ	56
4.2. Принципы организации управления разработками АИС и АСУ	58
4.3. Виды и принципы разработки структуры функциональной части АСУ	62
4.4. Разработка структуры обеспечивающей части АИС	81
4.5. Разработка методики для проектирования АИС предприятия	97
4.6. Выбор для предприятия готовых программных продуктов	100
Глава 5. Теоретические основы построения и проектирования документальных информационно-поисковых систем научно-технической информации	102
5.1. Виды научно-технической информации и структура Государственной системы НТИ	102
5.2. Информационные потребности и информационное обслуживание	109
5.3. Понятие об информационном поиске	114
5.4. Информационно-поисковая система. Структура ИПС	116
5.5. Информационно-поисковый язык	120
5.6. Система индексирования	137
5.7. Логика ИПС. Критерии смыслового соответствия	140
5.8. Критерии оценки ИПС	141
5.9. Разработка и отладка ИПС.	149

Глава 6. Принципы построения и проектирования документально-фактографических информационно-поисковых систем	151
6.1. Понятие о документально-фактографических ИПС	151
6.2. Автоматизированные системы нормативно-методического обеспечения управления предприятиями и организациями.....	152
Глава 7. Закономерности информетрии и их применение	159
7.1. Понятие об информетрии	159
7.2. Закон Ципфа	159
7.3. Закон Манделъброта	162
7.4. Закон Брэдфорда	164
7.5. Закон Викери	164
7.6. Закономерность концентрации – рассеяния В.И. Горьковой	165
7.7. Перспективы развития информатрии	168
Глава 8. Перспективные направления развития теории информационных систем	170
8.1. Концепция архитектуры предприятия и ее применение для развития теории информационных систем	170
8.2. Концепция ситуационных центров	187
8.3. Подходы к анализу и проектированию систем и их применение для развития теории проектирования информационных систем	196
8.4. Анализ определений системы и их применение для развития теории информационных систем	203
8.5. Концепция многоуровневой информационной системы	208
Приложение 1. Основные концепции архитектуры предприятия	217
Приложение 2. Пример создания многоуровневой информационной системы для обеспечения информационных потребностей студентов	260
Заключение.....	274
Список литературы.....	276

Научное издание

Волкова Виолетта Николаевна

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ СИСТЕМ

Налоговая льгота – Общероссийский классификатор продукции
ОК 005-93, т. 2; 95 3004 – научная и производственная литература

Подписано в печать 25.04.2012. Формат 60х84/16. Печать цифровая.
Усл. печ. л. 17,5. Тираж 100. Заказ 9093б.

Отпечатано с готового оригинал-макета, предоставленного автором,
в типографии Издательства Политехнического университета.
195251, Санкт-Петербург, Политехническая ул., 29.
Тел.: (812) 550-40-14
Тел./факс: (812) 297-57-76