

Лекция 6. Функциональные и обеспечивающие подсистемы

1. Элементы функциональных подсистем

На прошлых лекциях мы рисовали схемы, на которых были изображены основные функциональные части – подсистемы – типичных информационно-поисковых систем:

- подсистема анализа, комплектования, и описания (индексирования) документов;
- подсистема хранения документов;
- подсистема каталогов;
- подсистема указателей к каталогам (информационно-поисковых языков);
- подсистема входного пользовательского интерфейса;
- подсистема выходного пользовательского интерфейса.

Такое разбиение на подсистемы не является единственно возможным. Как уже говорилось в лекции 10, границы систем и подсистем никогда не бывают строго определённым, всегда существуют области, которые мы можем рассматривать либо как принадлежащие, либо как не принадлежащие определённой системе в зависимости от того, какие взаимодействия элементов являются существенными в рассматриваемых процессах. Так в числе перечисленных подсистем можно рассматривать как единую подсистему входной и выходной интерфейс. С другой стороны, из подсистемы комплектования обычно выделяют в отдельный блок подсистему индексирования документов по правилам используемых информационно-поисковых языков.

1.1. Блок хранения документов

Каждую подсистему можно дробить при рассмотрении на более мелкие подсистемы, вплоть до элементов, рассматриваемых как неделимые. Пределы делимости также в большой степени условны и зависят от наших потребностей и возможностей анализа. Даже в таком очевидном случае как состав блока хранения документов, элементом которого естественно является **документ**, мы должны рассматривать этот «элемент» как комплекс, состоящий из двух частей: (1) из исходного первичного документа, поступившего из внешней среды, и (2) из метainформации (описания, индексов), приписанной документу блоком комплектования, которая позволяет отыскивать документ в хранилище с помощью применяемых информационно-поисковых языков. Далее документ членится на его структурные части – разделы, абзацы, предложения, слова, буквы... И на буквах не всегда можно остановиться, если рассматривать кодировку символов в компьютерном представлении. И даже на двоичном представлении элементов памяти (в виде состояний «да» и «нет») нельзя

остановиться, если необходимо рассмотреть физическую реализацию этих состояний.

Документы, поступающие из внешней среды в блок комплектования, являются элементами этой внешней среды. Их структура отличается от структуры документов в блоке хранения отсутствием в них данных индексирования, индексной метаинформации. По физической реализации они могут быть как аналогичны документам, хранимым системой, так могут и отличаться. В развитых информационных системах целесообразно накапливать поступающие документы без изменения их формы, структуры и носителя. Однако технология пользования документами может потребовать изменения их формы и носителя. В традиционных книжных хранилищах часто создают резервный фонд оригинальных документов, которые служат преимущественно лишь для создания рабочих копий, реально используемых в остальных технологических процессах. Это позволяет обеспечить сохранность и аутентичность информации, поскольку всегда будет иметься оригинальный документ, с которого можно снять копию взамен испорченной или утерянной, а также предъявить его для доказательства истинности предоставляемых системой сведений. Рабочие копии документов могут создаваться как в традиционной бумажной форме, так и на машиночитаемых носителях или в компьютерных файлах. Большое значение имеет создание архивных фондов документов в виде фотографических микрокопий. В настоящее время именно технология микрокопирования предоставляет наиболее экономные средства долговременного надёжного и компактного хранения данных. Конечно запись на современные машинные носители данных гораздо компактнее микрофотокопий, но отсутствие опыта долговременного хранения не позволяет считать их надёжным средством архивирования. К тому же архивирование на машинных носителях требует сохранения не только носителей данных, но также технических средств и программного обеспечения, требующегося для их чтения.

Документы, поступающие в систему на машинных носителях, целесообразно во всех случаях копировать, чтобы создать в блоке хранения массив рабочих копий в формате, удобном для обработки. Даже поисковые машины Интернета, которые, как мы знаем, не предназначены для хранения информации, выложенной на сайтах, тем не менее сохраняют копии изменённых и исчезнувших страниц и выдают их в ответ на запросы.

Оригиналы документов, если они накапливаются в хранилище, составляют первый контур хранения информации. Вторым контуром составляют копии оригиналов. Третий контур представлен вторичными документами, образуемыми в информационной системе. Первые два контура в некоторых системах могут и отсутствовать, как в тех же поисковых машинах Интернета. Но третий контур всегда имеется, хотя бы на уровне адресов и библиографических описаний. В информационных системах, созданных для делового использования, в третьем контуре бывают представлены и другие

типы вторичных документов: ключевые слова, аннотации, рефераты, рецензии, обзоры.

1.2. Блок комплектования

Вторичные документы создаются в информационной системе блоком комплектования. Наилучшие результаты даёт составление вторичных документов силами специалистов, владеющих одновременно знаниями в технологии информационного обслуживания и в предметной области, излагаемых в документе сведений. Такую технологию, однако, отнюдь не всегда удаётся обеспечить. В крайнем случае некие вторичные документы могут быть созданы автоматически путём выбора определённых данных из исходного документа. В таком случае роль специалиста выполняет компьютерная программа. Специалисты и программы, анализирующие первичные документы и составляющие вторичные документы, являются элементами подсистемы комплектования.

1.3. Блок каталогов

Первичные документы, отобранные блоком комплектования из внешней среды, поступают в подсистему хранения, а та вторичная информация, которая идентифицирует первичный документ и определяет его место в хранилище (метаинформация), заносится в блок каталогов. Для традиционных изданий идентифицирующую информацию содержит библиографическое описание, для сайтов Интернета – их сетевые адреса (URL). Обычно эта информация дополняется кратким рефератом, аннотацией или некой выжимкой из документа, полученной автоматически. Эти три компоненты – адрес хранения, библиографические данные и аннотация, – называются в целом **каталожной записью** и является элементом подсистемы каталогов. Располагаются каталожные записи согласно системе принятой в данном конкретном каталоге. Тем самым в каталожную запись фактически добавляется четвёртый элемент – адрес (индекс) записи в каталоге. При наличии в системе нескольких каталогов каждая каталожная запись может повторяться в каждом из каталогов с новым индексом. Каталожные записи могут повторяться и в пределах одного каталога в зависимости от того, сколько индексов будет присвоено записи согласно технологии анализа и индексирования документов в блоке комплектования.

1.4. Блок указателей

Система каталожных адресов (индексов) определяется блоком указателя к данному каталогу. Этот блок может состоять из одной программы автоматического формирования адреса, как то происходит в случае алфавитных библиографических каталогов. В случае систематического каталога указатель к нему имеет вид библиографической классификации. Элементом классификации является **классификационная запись**, включающая обычно *индекс* (код) записи и *описание класса* документов, которые должны быть отнесены к заданному индексом месту в

каталоге. Кроме того в классификационные записи могут входить методические указатели индексаторам (используемые при «ручном» индексировании), а также ссылки на другие классы (записи), которые могут быть использованы как при ручном индексировании, так и в процедурах автоматизированного индексирования.

Для каталогов предметного типа указателями могут служить различные словари информационно-поисковых языков, рассмотренные нами ранее. Наиболее традиционным для библиотечной практики является словарь предметных рубрик. Элемент такого словаря – предметная рубрика – состоит часто из одного **заголовка предметной рубрики**, выраженного однословным термином или словосочетанием. В необходимых случаях в запись предметной рубрики вводят указания индексатору. В свою очередь предметная рубрика может включать несколько подзаголовков, представляющих собой рубрики более частных понятий, включаемых в состав понятия, выраженного основным заголовком.

В простейшем случае, характерном для полностью автоматизированных информационных систем, указателем к предметному каталогу служит словарь однословных ключевых слов – **унитермов**, которые и являются элементами подсистемы указателя. Однако наиболее перспективно использовать в качестве указателя **информационно-поисковый тезаурус (ИПТ)**. Элементами ИПТ являются дескрипторы, представленные словарными статьями, состав и структуру которых мы уже рассматривали в лекции 9.

1.5. Блоки пользовательского интерфейса

Как показано на рис. 12.2, указатели предоставляют информацию для соотнесения с каталогами как документов (в блоке комплектования), так и запросов пользователя. Последняя функция осуществляется обслуживающим персоналом системы или программами ввода запросов в блоке пользовательского интерфейса. Алгоритмы отнесения тематики документов и запросов к тому или иному понятию указателя могут быть тождественны и могут выполняться физически одними и теми же работниками, но конкретные программы и инструкции по вводу пользовательских запросов в систему отличны от таковых для ввода документов, что следует хотя бы из различия исходного материала документа и запроса. (Запрос пользователя всегда формируется из его информационной потребности, представленной в психической сфере). Элементами входного пользовательского интерфейса являются персонал, инструкции и компьютерные программы, обеспечивающие перевод информационной потребности пользователя в поисковое предписание для системы.

Обслуживающий персонал и/или терминальные устройства системы, взаимодействующие с пользователем, выполняют двойную функцию: ввод в систему запросов и выдачу из системы результатов поиска информации. Эти две функции могут выполняться физически различными работниками и

устройствами, но как правило они совмещены в материале одного организационно-технологического комплекса. Тем не менее процессы (и осуществляющие их программы), необходимые для перевода найденной информации в доступную для восприятия пользователем форму, имеют мало общего с процессами ввода запросов. Поэтому вполне оправданным является выделение независимого блока выходного пользовательского интерфейса, основными элементами которого являются именно программы, инструкции и технические средства (например, печатающие устройства) для выдачи документов пользователю.

Следует отметить, что выявленные в ходе нашего обзора основные элементы блоков не исчерпывают состав информационной системы. В каждом блоке кроме основных всегда представлены «вспомогательные» компоненты, образующие связи основных элементов и выполняющие другие особые функции. Например, в составе ИПТ кроме дескрипторов имеются аскрипторы – термины, не включаемые в поисковые предписания, а заменяемые дескрипторами. Все элементы блоков реализованы в каких-либо материальных объектах, которые в свою очередь являются компонентами системы. В систему входят также чисто технические средства поддержки работоспособности, например, элементы электропитания, помещения или ремонтный персонал.

2. Обеспечивающие подсистемы

Те же самые элементы системы, которые мы в предыдущих параграфах выявили в составе отдельных блоков информационной системы, могут быть систематизированы иным образом. Блоки схемы 11.1 показывают те группы элементов, которые оказываются наиболее тесно связанными при взаимодействиях, обусловленных процессами обычного функционирования системы. Эти блоки потому и называются **функциональными подсистемами**.

Особые воздействия на систему выявляют особые группировки элементов на подсистемы, спаянные особыми взаимодействиями. Важнейшим «особым» воздействием на систему является процесс её создания и поддержания работоспособности. Задача этого процесса – обеспечить систему необходимыми ресурсами. Поскольку информационная система является материальным образованием в социальной сфере, естественным образом выделяются три категории обеспечивающих процессов и соответственно три группировки элементов на подсистемы – **обеспечивающие подсистемы**:

1. процессы обеспéчения материальными ресурсами формируют подсистему «**материально-техническое обеспéчение**»,
2. процессы обеспéчения информационными ресурсами формируют подсистему «**информационное обеспéчение**»,
3. процессы обеспéчения социальными ресурсами формируют подсистему «**организационное обеспéчение**».

2.1. Техническое и организационное обеспечение

Материально-техническое и организационное обеспечение иногда рассматривают нерасчленённо как одну подсистему, представляющую собой субстрат, на котором действует информационная компонента системы. Тем не менее технические средства и другие материальные ресурсы (помещения, здания, энергоснабжение, коммунальные услуги), составляющие элементы материально-технического обеспечения, создаются и поддерживаются совсем другими методами чем те социальные структуры (предприятия, организации или другие коллективы), представляющие собой организационное обеспечение информационной системы.

Сердцевиной материально-технического обеспечения является та компьютерная техника, которая хранит запасённую в системе информацию и выполняет алгоритмы информационных процессов. Она чаще всего и подразумевается под термином «**техническое обеспечение**». Сюда также относятся каналы передачи данных и другие средства связи компьютеров, в частности средства связи общего назначения (телефонные, телевизионные и радиовещательные каналы). В большинстве случаев в качестве технического обеспечения информационных систем применяются стандартные компоненты универсальных вычислительных машин независимо от типов и масштабов систем (так называемое **универсальное** техническое обеспечение). Однако в ряде случаев, для реализации особо мощных информационных систем целесообразно применять специализированные вычислительные комплексы – так называемые **машины баз данных**. Специфическая архитектура этих устройств позволяет существенно повысить эффективность обработки терабайтных массивов информации за счёт применения многопроцессорных параллельных процедур.

В противоположность техническому обеспечению в подсистему **организационного обеспечения** входят не материальные компоненты, а в первую очередь «дематериализованные» социальные отношения людей, объединяющие их в коллектив, обеспечивающий работу информационной системы. Эти социальные отношения в некоторой степени подкрепляются общеправовыми и профессиональными нормативными документами, но главным образом реализованы в психической сфере персонала информационной системы, где мы выделяем следующие категории специалистов: администраторы системы (определяющие методы поддержания и развития системы), операторы (осуществляющие текущие процессы), пользовательские посредники (осуществляющие связь с клиентами). В сфере организационного обеспечения лежит **ответственность** за все процессы функционирования системы: обслуживание пользователей, эксплуатацию задач, подготовку данных, техническое обслуживание устройств.

2.2. Информационное обеспечение

К информационному обеспечению в широком смысле слова относится вся информация, записанная на элементах технического обеспечения и в документах организационного обеспечения. Однако в практике работы с информационными системами принято выделять различные слои информации, образующиеся в результате различных процессов и играющие в информационной системе различную роль.

Прежде всего к информационному обеспечению относится **информация первичных и вторичных документов**, находящаяся в блоке хранения документов. Так что этот слой информационного обеспечения совпадает с блоком хранения, и рассматривать их отдельно нет смысла. Компьютерная часть блока хранения состоит из одной или нескольких баз данных и, может быть, отдельных файлов памяти. В свою очередь базы данных представляют собой совокупности физических записей, разделённых на поля, придающие заранее определённый смысл записываемым в полях символам и тем самым определяющие применимые к данным операции обработки.

Кроме блока хранения документов в составе информационного обеспечения выделяется **информационно-технологическое** обеспечение. Оно зафиксировано, главным образом, в проектной и эксплуатационной документации информационной системы и включает следующие сведения:

- общие принципы организации фонда документов и средств их поиска,
- видовые, тематические, объёмные и временные характеристики информационных потоков (документов и запросов),
- виды носителей и описание форматов представления данных во входных и выходных потоках,
- методы хранения первоисточников и вторичных документов, состав и логическая структура создаваемых машинных файлов и документальных массивов,
- методы (алгоритмы) контроля, корректировки и защиты информации,
- средства и способы передачи данных и обмена информацией с внешними пользователями и системами,
- общее описание процессов сбора, обработки, хранения, поиска и выдачи информации по каждой из задач информационной системы,
- перечень технологических инструкций, регламентирующих порядок осуществления обработки информации и обслуживания пользователей.

2.3. Лингвистическое обеспечение

Большой слой информации в системе относится к лингвистическому обеспечению. Часто эту подсистему трактуют как содержащую только информационно-поисковые языки, т. е. средства описания и поиска документов в хранилище (см. лекции 7 – 9). При широком подходе сюда включают все средства представления информации и средства интерпретации данных. Тогда лингвистическое обеспечение охватывает объекты трёх типов: (1) форматы данных, (2) словари данных и (3) команды оперирования с данными. Во многих из этих элементов лингвистическое обеспечение непосредственно соприкасается с программным обеспечением, и чётко провести границу между ними зачастую бывает невозможно.

Форматы данных – это соглашения об интерпретации определённых частей файлов для представления информации различного типа и об агрегировании простых данных в логически связанные комплексы. Разные системы программирования предусматривают разные наборы типов данных, обладающих соответствующими форматами. Для примера можно назвать такие популярные типы данных как

- «логическое значение», занимающее в памяти один бит,
- «целое число», занимающее строку битов длиной в 2, 4 или 8 байтов,
- «действительное число», представляемое строкой байтов, где часть битов выражает показатель степени числа 10, а другая часть битов – мантисса, т. е. десятичная дробь, а общее значение есть произведение двух этих частей,
- «символьное значение», занимающее ровно один байт, смысл которого определяется таблицей используемой кодировки символов (например, популярная кодировка – ASCII)
- «бинарный объект» - неопределённая последовательность байтов, ограниченная стандартными разделительными символами,
- «текстовое поле» - бинарный объект, в котором разделители выделяют структурные элементы текста (слова, предложения, параграфы)
- и др.

Среди форматов выделяются **коммуникативные форматы**, предназначенные для обмена данными между системами. Они задаются официально утверждёнными нормативными документами, что гарантирует взаимную осмысленность данных, выданных разными системами на машинных носителях.

В библиотечной практике широкое применение находят форматы серии MARC, разрабатываемые в рамках Международной библиотечной ассоциации для обмена данными электронных каталогов и организации межбиблиотечного обмена документальной информацией. Для обмена между

центрами научно-технической информации предназначен отечественный формат МЕКОФ, заданный государственным стандартом ГОСТ 7.19. Оба эти формата соответствуют международному стандарту ИСО 2709, но отличаются друг от друга, хотя цель их применения – общая. В настоящее время всё большее значение принимает обмен информацией по сети Интернет, где стандартным представлением данных является язык HTML (HyperText Marc-Up Language = язык гипертекстовой разметки) или варианты его дальнейшего развития (XML и др.). Эти языки стандартизуются в рамках международной «Ассоциации 3W», которая объединяет ведущих разработчиков и провайдеров Интернета.

О **словарях**, составляющих основу информационно-поисковых языков, мы много говорили в лекциях 7, 8 и 9. Но в лингвистическое обеспечение входят также и другие словари, которые используются совместно с форматами данных для представления теми или иными битовыми последовательностями определённых понятий. Эти словари не обязательно должны быть такими сложными как известная нам универсальная десятичная классификация. В простейшем случае они могут содержать всего два элемента, как например кодификатор обозначений пола для системы персональных данных. Пол может быть обозначен в системе известными буквами – М и Ж, но не исключены и другие обозначения¹, например – всего лишь одним битом, принимающим значения *плюс* или *минус*.

Обращаясь к информационно-поисковым языкам, следует помнить, что они включают не только словарные компоненты, но и правила, которые обеспечивают процедуры применения словарей. Запуск этих процедур осуществляется специальными управляющими кодами, которые вместе с алгоритмами процедур, соответствующих этим командам, составляют так называемые **языки описания данных (ЯОД)** и **языки манипулирования данными (ЯМД)**. Первые из них (ЯОД) содержат команды записи определённых значений в информационные файлы системы в соответствии с теми или иными характеристиками документов или запросов. ЯМД содержат команды двух видов – команды поиска информации и команды отображения информации для пользователя. Совокупность команд поиска информации называются также **языком запросов**. В настоящее время фактическим стандартом является язык запросов SQL (Structured Query Language = язык структурированных запросов).

2.4. Программное обеспечение

Программное обеспечение (ПО, software) – это вся совокупность команд и процедур, заставляющая действовать техническое обеспечение при реализации автоматизированных функций системы. В составе программного обеспечения можно выделить три слоя команд, которые разрабатываются разными лицами и на разном уровне программирования.

¹ В российских православных монастырях встречается обозначение туалетов буквами «Б» и «С», что означает «братья» и «сестры».

Наиболее глубокий слой – это **операционные системы (ОС)**. Они составляют универсальную основу программного обеспечения. Их задача – управление данными и заданиями независимо от содержания и назначения обрабатываемых данных. Операционная система также осуществляет связь с человеком-оператором, который даёт задания и определяет данные для обработки. Операционная система современных персональных компьютеров – Windows разных последовательно совершенствующихся модификаций. Известны и другие операционные системы: Unix, OS/2, MS DOS. Разработка операционных систем тесно связана с конструированием архитектуры компьютеров, и компьютеры в редких случаях эксплуатируются без установки какой-либо из имеющихся в широкой продаже операционной системы.

Создавая информационную систему, дополнительно к ОС компьютер часто снабжают имеющимися на рынке специальными комплексами программ, обладающими всеми основными операциями, необходимыми для формирования баз данных и поиска в них информации. Это позволяет оперативно ввести в компьютер документы, сформировать критерии обработки и поиска их и тем самым получить действующую информационную систему с минимальными затратами средств и времени. Такие программные комплексы называют **оболочками** информационных систем или **системами программирования** информационных систем.

На следующем уровне программирования находятся программы, специально созданные разработчиками информационной системы для выполнения специфических функций, уникальных для имеющегося фактического материала и задания от заказчиков разработки.

2.5. Нормативное обеспечение

Говоря об организационном обеспечении, мы уже упомянули существование нормативных документов, фиксирующих отношения членов персонала информационной системы. Они образуют подсистему **правового обеспечения**. В учебнике Н. З. Емельяновой и др.² приводится следующий перечень документов, необходимых для юридического обоснования функционирования систем:

- Положение о службах, обеспечивающих функционирование информационной системы
- Должностные инструкции
- Нормативные акты о порядке создания и использования информации
- Нормативные акты, регламентирующие процесс обработки информации
- Нормативные акты на использование вычислительной техники.

² Н. З. Емельянова, Т. Л. Партыка, И. И. Попов. Основы построения автоматизированных информационных систем. – М.: Форум, Инфра-М, 2005. – С. 84.

К этому комплексу присоединяются инструктивные материалы по выполнению всего комплекса процедур сбора, обработки и выдачи документов, основой которого являются государственные стандарты (ГОСТы) группы СИБИД (Система стандартов по информации, библиотечному и издательскому делу).

Все эти документы нормируют юридические и технологические отношения в процессе функционирования информационной системы и составляют её подсистему **нормативного обеспечения**.

В процессах создания и эксплуатации информационных систем следует оказывать должное внимание всем подсистемам (функциональным и обеспечивающим) поскольку отказ каждой из перечисленных подсистем приводит к отказам информационной системы в целом. Нормативные документы, разработанные по опыту создания автоматизированных информационно-поисковых систем в нашей стране требуют в проектной документации подробно описывать каждый из перечисленных здесь блоков и обеспечений.