

Александров В.В., Андреева Н.А., Кулешов С.В.

Методы построения информационно-логистических систем

Санкт-Петербург
2005

УДК 658.012:004.42
ББК 32.973:26-018.26

Современное состояние программных продуктов, реализующих пользовательские потребности в построении и сопровождении информационных систем, прошли в своем развитии следующие этапы: базы данных, представляющих собой простые отношения между унифицированными данными, представленными табличными структурами; базы знаний, содержащие также правила управления этими отношениями. Современные условия требуют разработки информационных систем, опирающихся на понятийно-терминологический состав предметной области. Реализацию методов организации ассоциативных полей и связей в понятийно-терминологических областях можно видеть на примере тезаурусов предметных областей.

В книге рассматривается тенденции развития информационных систем и становление понятий логистической и семиологической информационных систем, даются примеры семиологических систем логистического информационного сопровождения.

Рассматриваются программные реализации интерфейсной системы динамической визуализации тезаурусов и глоссариев, текстуальных описаний конкретных предметных областей. Приводятся примеры.

Александров В.В., Андреева Н.А., Кулешов С.В.
Методы построения информационно-логистических систем.

Содержание

Содержание	3
Предисловие.....	4
Введение	4
Исторический дискурс	9
Латиница и Кириллица.	15
Текст и смысл.....	16
Понятийная база информационных систем	22
История коммуникативных функций языка.	36
Онтологическая информационная система.....	39
Visual Thesaurus	40
Динамическая визуализация контекста (Визуальный словарь).....	44
Информационные системы.....	46
Электронный документ.....	51
Программные продукты и решаемые задачи.....	53
Базы данных.....	56
Информационно-аналитические системы	58
САПР	59
Экспертные системы.....	62
Логистические системы.....	63
Поисковые системы в Интернет.....	66
CAD и CALS	72
Информационная поддержка жизненного цикла изделий (CALS)	75
ИПИ-технологии. Технологии представления данных	79
Информационные системы в производстве	82
Семиологические информационные системы	84
Информационная безопасность.	86
Глоссарии и семантическое упорядочение информационных потоков.....	90
Основы семиологической теории	94
Литература.	102
Краткий словарь терминов	104
Приложение 1.....	107
Приложение 2.....	107

Предисловие

"Нет ничего кроме текста" ("Il n'y pas de hors-texte")

Ж. Деррида

Современное состояние инфо-телекоммуникационных возможностей привело нас к тем же проблемам, которые на протяжении более трех тысяч лет привели к упорядоченным текстологическим измышлениям и текстовому типу коммуникационного общения.

Проблемы, решение которых привело человечество к языковой форме общения, во всю пытаются применить и в концепции компьютерного инфо-телекоммуникационного общения. Однако, посредник-инструмент в виде компьютера и его программного окружения предъявляет свои собственные как возможности, так и ограничения. Его адекватное восприятие и понимание текстуальных информационных посылок требует своих методов представления информации.

Вместо объяснения пространными фразами, подходящими для человека, для компьютера требуется строить формализованные запросы, вместо переформулировки требования требуется уточнение условий и ограничений и др.

Не понимание принципов, лежащих в основах информационных систем, приводит к парадоксам и общей неудовлетворенности от компьютерных систем в целом.

В связи с этим в книге основное внимание сосредотачивается на проблемах неадекватной **интерпретации текстологического содержания пользователем информационной системы и программной средой компьютера.**

Введение

За несколько тысячелетий человек создал языковую – речевую, текстуальную и визуальную формы информационного общения. Гуманитарно-философское сословие: филологи и лингвисты упражнялось и упражняется в поиске простых составляющих и их связей отражающих семантико-смысловые компоненты инвариантные относительно наций, рас, географии и государств.

Нет никакой мистики в попытках организовать процесс связи букв, слов с семантикой текстового содержания. Мифологическое творчество и различные религиозно-мистические учения, прежде всего, пытались преодолеть вербальные трудности в организации коммуникационных процессов общения. Приведем исторический дискурс:

...Гематрія – самый сложный и трудный приемъ для раскрытія іероглифическаго смысла, но зато онъ является наиболее плодотворнымъ въ своихъ результатахъ. Каждая изъ 22 буквъ еврейскаго алфавита соответствуетъ определенному числу; первыя 9 буквъ соответствуютъ простымъ числамъ, слѣдующія 9 — десяткамъ и т. д. Каждое слово соответствуетъ определенному числу, получающемуся сложениемъ чиселъ, соответствующихъ буквамъ этого слова. Точно такъ-же целые стихи могутъ быть приведены къ одному числу. Общее мистическое правило этой цифровой Каббалы заключается въ томъ, что слова, выраженные однимъ и темъ-же числомъ, имеютъ

внутреннее сродство между собою. Поэтому, получивъ изъ слова число, можно его разложить на новыя слова, соблюдая вышеозначенное правило...

Не обходится этот процесс без мистики... И впрямь, все, что познается, имеет число, ибо невозможно без него ничего ни понять, ни познать, тем более организовать процесс идентификации. Потребность в числовом процессе семантически-смысловой идентификации понималась задолго до появления компьютеров.

Исторически предполагалось решить проблему ...мірозданія — путемъ лингвистическимъ. Вещь состоитъ изъ сущности, формы и имени, причемъ имя не только неотделимый элементъ вещи, оно важное связующее звено между сущностью и формой: именно благодаря ему сущность плюсъ форма образуютъ вещь. Но такъ какъ самое имя есть результатъ известнаго сочетанія буквъ, то, въ конечномъ счете, буквы являются орудіями, при помощи которыхъ созданъ весь реальный міръ, состоящій изъ сущностей и формъ. Созданъ же онъ изъ десяти первоначальныхъ чиселъ (цифръ, «сефиротовъ»), представляющихъ безформенныя сущности. Къ этой теоріи буквъ и цифръ присоединяется теорія контрастовъ, по которой міръ физическій, какъ и міръ моральный, являетъ рядъ тезъ и антитезъ, примиряемыхъ и уравниваемыхъ синтезами... [1]

Проблема «тезъ», «антитезъ», синтеза, а также поиска и реферирования в информационной среде Интернет до сих пор ждет своего решения. На современном уровне реализован лишь примитивный механизм символично-числовой идентификации. Другими словами, **компьютер осуществляет процесс идентификации (пусть даже ее называют ассоциативной, понятийной, интеллектуальной и др.) на основе совпадений заданной последовательности единиц и нулей, как алфавита, где буквы и числа неразделимы.**

И снова мы вынуждены прибегать к лингвистическому подходу и решать проблему недостатков естественных языков с их многозначностью и неточностью. В XVII в. в Европе возникает идея создания совершенно универсального языка символов для описания знаний об окружающем мире. Одним из первых в этой области работал Рене Декарт. По его мнению, универсальный язык, однозначно понимаемый всеми людьми, должен основываться на всеобщей классификации, математике, логике и главным образом философии. Вследствие этого этот язык стали называть “философским языком”. Такой язык, созданный средствами наук, должен быть удобным не только для передачи мысли, но и для самого мышления.

Вслед за Декартом этой темой занимался Готфрид Лейбниц. Он предполагал выразить все содержание человеческого мышления системой символов (“характеров”) – своеобразной “азбукой человеческих мыслей” – и заменить рассуждения вычислениями. Таким образом, язык становился бы алгеброй мысли. Лейбниц считал, что *никогда не кончатся споры и не установится мир в борьбе школ, пока от путаных рассуждений, неясных слов и неопределенных значений мы не перейдем к простым исчислениям и определенным характеристам.* По его мнению, универсальный язык, в том виде как его предполагалось создать, *есть самое возвышенное и самое экономичное употребление человеческого разума с помощью символов и знаков.*

Для этой цели он и предложил «универсальную характеристику» – конгруэнтность $ABC \cong AXC$ – способ переопределения понятий и их идентификаций – основу функционирования современных компьютерных технологий (рис. 1).

Понятие конгруэнтности в современных системах управления базами данных превратилось в простую выборку всех элементов X по условиям совпадения ключа B , где A и C ограниченность словаря (граничные условия множества).

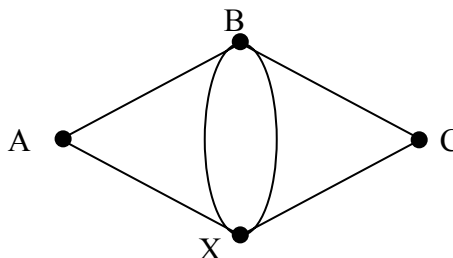


Рис. 1. Принцип идентификации неразличимости Лейбница.

Здесь следует подчеркнуть что слово «число» используется как понятие, обозначающее упорядоченный уникальный идентификатор, на что и адаптирована аппаратно-программная среда компьютеров. Мистика же проистекает из попыток построить ассоциативно-идентифицируемые подмножества слов не переопределяя динамически изменяющийся ассоциативный ряд понятий в развивающихся информационных потоках. Например, первоначально простое понятие англоязычного слова “Tuning” – «тюнинг» (попасть в тональность, тонкая настройка радиотелевидения при получении сигнала) превратилась в рекламный слоган «тюнинг автомобиля», «тюнинг офиса», т.е. приобрело понятийное звучание таких слов, как «модернизация», «настройка», «подгонка» и др.

Это наглядный пример, как по части идентифицируется целое, т.е. реализация голографичности – свойство, универсальность которого пока что воспроизводится и иллюстрируется лишь разумом человека.

Еще Г. И. Гюрджиев [2] как-то завел разговор об универсальном языке. По его мнению, универсальный язык возможен, только люди никогда не изобретут его, поскольку он давно изобретен. Точнее это даже не язык в прямом смысле этого слова, а универсальная схема, которая называлась “энеаграммой” (от гр. Ennea – 9) Эннеаграмма представляет собой круг, на котором размещен ряд натуральных чисел от 1 до 9. Шесть из этих чисел – 1, 2, 4, 5, 7, 8 – связаны друг с другом линиями внутри круга в следующем порядке: 1, 4, 2, 8, 5, 7... Остальные три числа – 3, 6, 9 – образуют треугольник

Гексанема является наиболее характерной деталью эннеаграммы. Она задается циклической дробью $1/7=0,142857...$, о которой знали еще вавилоняне. Это очень занятная дробь. По поводу ее составных частей существует два высказывания Пифагора. Один раз он сказал: “Все – едино”, а другой – “Все есть число 7”. Таким образом, дробь $1/7=0,142857...$ также символизирует “Все”. Следует отметить, что аналогичная циклическая последовательность образуется при делении на 7 любого целого числа (если, конечно, оно не кратно числу 7). При другом числителе начало периода просто сместится (например, $3/7=0,428571...$) – **получается, что часть в какой-то мере тождественна**

целому! Дробь $22/7=3,142857...$ есть одно из приближенных значений числа π , которое употреблял еще Архимед.

Числами Г. И. Гюрджиев кодировал понятия из определенных областей знаний. С помощью такого кода можно описывать устройство какого-либо целостного явления или системы, например, музыкальной гаммы, человеческого организма, планетарной системы. Связь между числами, обозначенная гексанемой, расположение чисел по кругу и т. д. – все это должно отражать взаимоотношения между элементами описываемой системы, причем не только статически, но и в динамике.

Отсюда появляются популярные сегодня рассуждения о коде Библии, коде Леонардо да Винчи, гене человеческого организма и последовательности Фибоначчи как законах пропорционального развития с постоянным отношением 0,618 между целым и его частями.

В настоящее время системная методология стала общепризнанной во всех областях науки. Появились специализированные направления, изучающие те или иные свойства систем, – кибернетика, синергетика и общая теория систем. Кибернетический подход охватывает открытые для информации и безэнтروпийные системы, особое внимание, уделяя изучению процессов управления. В синергетике, в отличие от кибернетики, упор делается на постижении самих принципов построения систем, возникновения самоорганизаций, их развития и самоуничтожения. В исследованиях по общей теории систем, рассматривающей все объекты мира как системные, т.е. являющиеся системами и принадлежащие к какой-либо системе акцент ставится на создании концептуально – алгоритмического программируемого подхода.

Мы подчеркиваем, что **эффективность многофункционального инструмента – компьютера** в ограничении интерпретационных возможностей пользователя. При всем разнообразии природно-биологических форм передвижения: от черепахи до гепарда, научно-технический прогресс доказал энергетическое совершенство искусственно созданных объектов: колеса и двигателя.

Та же самая ситуация наблюдается и при разработке эффективных средств компьютерно-информационных телекоммуникаций.

Конгруэнтность Лейбница – «колесо» информационно-идентифицирующего процесса, а свойство коммуникативности (передача, преобразование и прием информации) – «двигатель» лингвистического (семантико-смыслового) процесса.

Это становится все более актуальной для Интернет и компьютерных технологий. Известный современный философ Ж. Деррида дает следующий тезис «Нет ничего кроме текста» («Il n'y a pas de hors-texte»), где подразумевается предельное расширение понятия текстуальности: вся жизнь становится текстом. А информационные системы – эффективное средство интерфейсного, по-русски – межличностного информационного консенсуса.

Книга – текстуальный способ фиксации и передачи знаний. По мере развития информационных технологий эти функции наиболее эффективно развиваются, в том числе и как средство Интернет-коммуникации.

Рассмотрим, на чем основана информационная коммуникативность.

Можно ли сформулировать некоторый текст, который однозначно понимался бы в разных языках?

В инженерно – технических дисциплинах это строго-формализованный аксиоматический подход, отраженный функционально-аналитическими записями.

Выявление же непротиворечивых логических отношений, реализующих коммуникативный акт из текстуальных описаний, представляет самостоятельную сложную проблему.

В [3] наиболее подробно сформулирована данная проблема.

“Некоторые этические проблемы стали прозрачнее после того, как я продумал некоторые проблемы семантики – только не беспокойтесь из-за того, что говорят, будто выражаемся мы сложно; вероятно, те, кто говорит это, приучены думать чересчур просто по вине средств массовой информации с их «откровениями», по определению предсказуемыми. Пусть же привыкают мыслить более сложно, ибо не просты ни тайное, ни явное.

Суть лингвистической задачи сводилась к следующему: существуют ли «семантические универсалии», то есть элементарные понятия, общие для всего человеческого рода и находящие выражение на любом языке. Проблема эта не так уж легкоразрешима, если учесть, что во многих культурах отсутствуют понятия, кажущиеся нам очевидными: скажем, понятие субстанции, наделенной качествами (например, в выражении «яблоко – красное»), или понятие идентичности ($A = A$). Однако в результате раздумий я заключил, что безусловно имеются понятия общие для всех культур и что все они относятся к положению нашего тела в пространстве”.

В настоящее время одной из основных проблем в области информационных технологий является обучение адекватному описанию, упорядочению и восприятию текстуального отражения информационных составляющих предметной области. В то же время инженерно-техническая мысль стремилась с одной стороны упростить устранение неоднозначностей, возникающих в соответствии со специфическими лингвистическими конструкциями построения текстов человеком, а с другой интенсифицировать процесс накопления и воспроизводства текстуально-ориентированных знаний.

Компьютерно-ориентированный процесс идентификации связан с процессами классификации, ассоциации, анализа и синтеза текстуального описания и базируется на принципе идентификации неразличимости, который был сформулирован Г. Лейбницем следующим образом: "Два объекта считаются неразличимыми, если все их свойства общие". Этот принцип наиболее активно используется при индексном ключевом поиске в компьютерных системах (поиск по однозначному совпадению) в котором в качестве ключа выступают буквы, слова, символы и т.д. Конгруэнтность или “универсальная характеристика” по определению Г. Лейбница, открывала возможность “...человеку рисовать, не имея дара художника...” [4]

Эта возможность заложена в правиле **отождествления X как множества точек, объектов, вещей и их свойств с именем и свойствами известного объекта C** (см. обозначения, введенные ранее). Последовательная реализация этого принципа и привела к созданию Компьютера, Интернета и Мультимедийных программных средств.

Новая информационная технология, казалось бы, дает неограниченные возможности в обучении и интеллектуальном познании. Однако тот, кто активно использует Интернет для поиска научной информации, постоянно сталкивается с

необходимостью “**понятийной идентификации**” – сопоставления искомого с известным, представленным в различных формах и языках.

Современные мультимедийные и Интернет технологии, компьютерные информационные системы не только не влияют на процесс и не облегчают суть установления информационного консенсуса в процессе общения, но, модифицируя сам процесс, включают такого посредника, как компьютер с его специфическими особенностями, в организацию и восприятие информационного процесса.

Научно-образовательные школы являются посредником между индивидуально субъективным восприятием предметных областей и общепризнанными догмами, понятиями, терминами, определениями их текстуального представления.

В настоящий момент все еще продолжает быть актуальным следующий вопрос. Известно, что практика есть критерий истины, отраженной в результативности и эффективной работоспособности системы.

Вопрос состоит в оценке продолжительности положительной динамики законсервированной внутренней программной среды любой информационной системы в быстроменяющейся внешней среде. Ответственность авторов программных продуктов, составляющих фундамент информационных систем, перекладывается на плечи всех потребителей систем. В итоге возникает потребность в корпоративном договоре о правилах, правах и допустимых ограничениях функционирования корпоративных потребителей: телекоммуникационных, компьютерных и информационных систем.

И первое же нажатие на клавишу клавиатуры компьютера вызывает потребность осознавать, что Ваш Алфавит – это не символ Вашего языка, а абстрактные протоколированные (стандартизованные) символы – последовательность 8-битовых значений «1» и «0».

Из этого следует общая для всех информационных систем постоянно разрастающаяся в условиях корпоративности и глобализации проблема совместимости идентификаций: звука в речь или музыку; буквы-символы в слова-понятия и как результат осмысления в действия – принятие решения [5].

Исторический дискурс

Задолго до компьютера и Интернет проблема коммуникативного общения составляла предмет исследования. [6]

*...Так например у грачей долговечных – охрипшие крики,
Или же карканье стаи вороньей, когда, по рассказам,
Просит дождя она и накликает и ветер и бурю.
Стало быть, если различные чувства легко могут вызвать
У бессловесных зверей издавание звуков различных,
То уж тем более роду людей подобало в ту пору
Звуками обозначать все несхожие, разные вещи...*

Хрисипп:

Не одно и то же голос и говор... но голос дело голосовых органов говор – говорных, из которых первое место занимает язык, а затем нос, губы и зубы. А голосовые органы – гортань и движущие ее мускулы и нервы, доставляющие этим последним силу из мозга.

Диоген Вавилонский:

Слово, согласно стоикам членораздельный звук.

Предложение – значащий звук, исходящий из рассудка

Диалект – слово, имеющее и племенную и эллинскую чеканку; или местное слово.

Элементы слова – двадцать четыре буквы. «Буква» говорится в тройном смысле: элемент, начертание элемента и его имя.

Гласные – семь элементов алфавита

Безгласные – шесть

Секст Эмпирик:

Стоики утверждают, что три (вещи) между собой сопряжены – обозначаемое, обозначающее и объект. Из них обозначающее есть звук; обозначаемое – тот предмет, выражаемый звуком, который мы постигаем своим рассудком, как уже заранее существующий, а варвары не воспринимают, хотя и слышат звук.

Сегодня как никогда актуальна, но все еще далека от решения проблема компьютерной идентификации речи.

С текстуальной идентификацией компьютерные системы на данный момент относительно справляются.

Звук и слово отличаются между собою тем, что звуком является и звон, а словом только членораздельное (звучание).

Слово отличается от предложения тем, что предложение всегда значаще, слово (бывает) и не значащим, а предложение – никоим образом.

Высказывание отличается от произнесения: произносятся звуки, высказываются вещи, которые и являются высказываемыми.

*Начинает речь тот, кто манифестирует – объявляет о намерениях (задает ограничения предметной области). То о чем сообщается – это денотат-имя. То, что говорится это значение-вывод – сигнификация. Событие не является ни одним из них: оно говорит не более, чем о нем сообщается или говорится. Тем не менее, событие принадлежит языку и связано с ним настолько, что не существует вне выражающих его предложений. Денотация и манифестация не составляют основу языка, они всего лишь становятся возможными вместе с ним. Что делает язык возможным – так это событие. **Без события все было бы только шумом – невнятным шумом.** [7]*

И снова мы приходим к следующей интерпретации. А – исходная посылка, множество известных манифестанту-пользователю знаний, упакованных и упорядоченных в его компьютерном представлении из которого он и актуализирует конкретный объект – денотат В и получает Х как список из множества эквивалентных объектов из иных информационных источников С, значения которых и сигнифицирует – раскрывает информационное содержание денотата В.

В – раскрывается через Х в том и только в том случае, если информационные источники В текстуально идентифицируемы, полны, достоверны и внутренне непротиворечивы. Проблема С – это проблема адекватности энциклопедических знаний.

Интересна и следующая, обратная, задача. Если В не раскрывается через Х то либо не найден и/или не существуют дополнительные знания о В. Или мы имеем в качестве С

специально организованный информационный мусор. Такие «качели» между произвольным выбором между В и С порождают процесс авто(само)реферирования (self reference). Элемент этой системы частично реализован нами в [8].

Под текстуальной идентификацией мы понимаем фиксированную среду и нормативные акты, которые устанавливают нормы электронного документа.

В рамках информационных систем можно привести следующее утверждение. Пользователь – это манифестация семантической интерпретации предложения, то предметная область – это принцип денотации, а информационная база знаний – принцип сигнификации.

Этимологизирование, как метод истолкования «имени», нисколько не было скомпрометировано новыми постановками вопроса о соотношении между именем и вещью. Если сторонники правильности имен «по природе» могли искать в имени сокровенную сущность предмета, для представителей договорной теории оно являлось документом тех взглядов на вещь, какие существовали у древних «творцов» имени. Платоновский «Кратил» [6] представляет в значительной своей части пародию на излюбленные в то время приемы этимологизирования; приемы эти очень характерны для античного подхода к проблемам языка. Объектом этимологического объяснения являются «имена». Слова, не имеющие непосредственного **вещного значения**, не интересуют этимолога, и даже глаголы попадают в сферу его внимания лишь в какой-либо из именных форм, например в причастии. Имя берется в номинативе как нечто законченное, без различия морфологических элементов; флектируемое окончание рассматривается как реальная, а не формальная лишь часть имени. Этимологический анализ состоит обыкновенно в том, что разбираемый номинатив раскрывается как словосочетание, образованное из двух или нескольких уже наличных в живом языке слов, в их реальных формах без различия морфологических формантов. Часто встречающийся в греческом языке и легко доступный языковому сознанию прием сложения слов рассматривается как основной метод словообразования. Идеальным с этой точки зрения является случай срастания слов (типа русского «умалишенный» – лишенный ума), когда составное слово механически разлагается на части, образующие законное с точки зрения живого синтаксиса словосочетание. Поскольку однако такие случаи чрезвычайно редки, античный этимолог приходил к убеждению, что в живом языке имена уже искажены – бессознательно, от «древности» и долгого употребления, или сознательно, в целях «благозвучия». Изменения, которым подвергается имя, складываются из: 1) вставки, 2) изъятия, 3) замены, 4) перестановки тех или иных букв, – и этимолог учитывает при анализе все эти возможности. Семантическая энергия имени неравномерно распределена по его частям: одни «буквы» являются носительницами смысла, другие – безразличны. В качестве примера Платон приводит названия букв алфавита, – альфа, бэта и т.д., в которых значащими являются лишь начальные буквы (семитическое происхождение названий букв делало их непонятными для греков), а прочая часть слова индифферентна. Необходимым условием правильной этимологии остается конечно требование, чтобы обнаруженное в составе имени словосочетание раскрывало какие-либо свойства или признаки обозначаемого именем предмета – **«вещное знание»**.

Проблема **«вещного знания»** проявляется и в современном интеллектуальном поиске, правда простая замена слов **«вещного знания»** на **логистику знания**, проблему

не решает. Проблема в том, **что и когда** считать **вещным** знанием и **кто** это знание определяет. Отсюда и возникает проблема диахронизма – с развитием профессиональных знаний предметной области быстро изменяется их понятийное окружение, а следовательно содержание словарей и неоднозначность термина интеллектуального поиска., причем по-разному для разных потребителей информационных систем. Требуется постоянно доопределять, что в данный момент является догмой, а что интеллектуальным содержанием.

Мы по-прежнему в положении конкуренции с развивающейся окружающей средой и по-прежнему стоят старые проблемы эффективной корпоративной глобальной коммуникации:

...нужда научила людей помогать друг другу и, собираясь вместе от страха, они начали постепенно друг друга узнавать. Голос их был еще бессмысленным и нечленораздельным, но постепенно они перешли к членораздельным словам и, установив друг с другом символы для каждой вещи, создали понятное для них самих изъяснение относительно всего. А так как такие объединения имели место по всему миру, то язык оказался не у всех равнозвучным, поскольку каждые случайным образом составляли свои слова: отсюда разнообразие в характере языков, а первоначально возникшие объединения положили начало всем племенам.

Языковым проблемам посвящен его (Платона) диалог «Кратил», но задача «Кратила» – показать, что имена не являются орудиями познания вещей и что исследование имен бессильно помочь мысли, стремящейся к познанию «вечносущего». Платон, как это часто бывает в его диалогах, сначала ведет читателя по ложному пути. Он исходит из предположения, что имя является «орудием» познания вещей, и приходит к выводу, что имя должно выражать «идею» вещи с помощью «букв и слогов». Эта мысль конкретизируется затем на целой серии этимологии, в большинстве случаев заключающей в себе момент пародии, и кульминирует в изложении атомистической гипотезы семантики звука; имя оказывается подражанием природе вещи. Дальнейшие рассуждения показывают однако, что все эти выводы ошибочны. «Уподобление» звукового состава имени предмету всегда несовершенно; конкретное имя часто включает в себе «буквы», семантически не только не соответствующие, но даже противоположные природе обозначаемого; тем не менее мы такие имена понимаем, и правильность их – результат «договора». В еще большей степени это относится к именам предметов, лишенных чувственной качественности, например к именам чисел, в отношении которых «уподобление» невозможно. Метод этимологии ненадежен и сам по себе, открывая широкие возможности произвольных и разнообразных толкований, но даже правильный этимологический анализ способен раскрыть лишь мнение творцов имен о вещах, а мнение их может оказаться ложным. Вещи должны познаваться на основании исследования самих вещей, а не на основании исследования имен. Имя, таким образом, не является орудием познания. Платон не задается целью оспаривать атомистическую гипотезу и даже готов ее принять, но для его целей она безразлична. Звук как «подражание вещи» еще более отдален от вечносущего, чем самая чувственная вещь, и имя не может иметь никакой ценности кроме договорной. Т. Бенфей находил в рассуждениях первой части «Кратила» зародыши понятия «философского языка», но Платону эта мысль совершенно чужда. Воспроизвести «идею» в звуке принципиально

немыслимо, а другого соотношения между содержанием и именем Платон не знает. Под конец жизни, излагая в VII письме основные положения своей системы, Платон разъясняет, что знание имени есть наинизшая ступень познания, ниже, чем представление телесного образа вещи, и в резких выражениях формулирует свое согласие с «договорной» теорией.

Аристотель в этом отношении вполне солидарен с Платоном. Он неизменно подчеркивает, что слова семантически лишь по «договору», что они не являются «орудиями» и не заключают в себе ничего «природного». Позднейшие комментаторы, которые принимали пересказанную у Платона атомистическую гипотезу за собственное его учение, потратили много бесплодного труда для того, чтобы ослабить мнимое противоречие между Аристотелем и Платоном. Образчик этой гармонистики читатель найдет в рассуждениях Аммония (около 500 г. н. э.). Специально языковыми проблемами Аристотель не занимался и затрагивал их мимоходом, преимущественно в связи с теми или иными логическими или риторическими исследованиями. Это не случайно, поскольку слово является для Аристотеля в первую очередь звуковым комплексом, а семантическая сторона слова, те душевные переживания, знаком которых слово является, полностью лежат уже вне сферы языка.

Однако те исследования в области логики, которые велись в школе Платона, а затем самим Аристотелем и его учениками, оказали значительное воздействие на языковую теорию. Учение Аристотеля о понятии и суждении опирается во многих своих частях на неоформленные языковые наблюдения, и логический анализ суждения обратно отражается на языковой теории в первых попытках классификации частей речи. Уже у позднего Платона традиционный для греческой философии дуализм истинного «сущего» и ложного «не-сущего», приписывавший истинность и ложность самому объекту, уступил место новому воззрению, которое усматривало момент истинности и ложности в соединении или разделении двух объектов; отсюда развивается учение о субъекте и предикате суждения. Внимание исследователя направляется на глагол. В «Софисте» Платона излагается как новинка, положение, что всякая «речь» состоит по крайней мере из имени и глагола и является их соединением. Если раньше считали предложение сплетением «имен» и «речений», то теперь термин «речение» (ῥημα) специфицируется и начинает означать глагол как противостоящую имени часть «речи», т. е. предложения. Аристотель указывает в качестве характерного признака глагола, что он, помимо своего основного значения, содержит добавочное означение времени. Применение термина «речение» могло быть вызвано и тем, что глагол, взятый в личной форме, всегда соозначает субъект. При этом грамматические понятия еще сливаются с логическими: Аристотель вообще недостаточно разграничивает предмет, мысль и языковое выражение. Лексический состав языка разделяется на две части: слова знаменательные (семантические) и незнаменательные (асемантические). К первым относятся имена и глаголы, вторые играют роль связующих элементов предложения; Аристотель называет их, как и современный ему ректор Теодект, «союзами» (в терминологии нового времени – «частицы» или «служебные части речи»). Понятие имени берется таким образом еще очень широко. Впрочем четкой грани между «именем» и «субъектом», «глаголом» и «предикатом» у Аристотеля вообще нет.

Отождествляя имя с субъектом, Аристотель определяет те именные формы, которые не могут служить субъектом, как отклонение от нормы, «падения» имени, «падежи». Номинатив есть «имя», прочие формы – «падение»; технических обозначений отдельных падежей еще нет. Глаголом является только настоящее время; остальные времена – падения глагола. Не-личные формы глагола, инфинитив и причастие, Аристотель относит, по-видимому, к именам.

Степень оригинальности грамматических построений Аристотеля неясна: проблема классификации частей речи и разработка грамматических категорий вставала в это время и перед риториками (Теодект, Анаксимен) в связи со стилистикой и обучением литературному языку.

Аристотель находится на пороге эллинистической эпохи, того периода в истории греческого общества, когда наступают его застой и загнивание. Для эллинистической науки характерно значительное накопление эмпирического материала при бедности творческими идеями. Языковые исследования не избежали этой общей судьбы. И хотя создание античной грамматической теории по существу, еще впереди, методологические основы ее уже почти окончательно определены предшествующим развитием.

Основной особенностью античной «языковой» теории является то, что она вовсе не ставит своей задачей изучение языка. Термин «языкознание» не античен, не антично и самое понятие. Разумеется, и у греков и у римлян имелось слово «язык» (γλῶττα или διαλέχτος, lingua), но оно всегда оставалось обыденным словом и не достигало уровня научного термина. О «языке» говорили, когда надо было противопоставить один язык (или диалект) другому; но язык не осознавался как объект исследования. Античный теоретик анализирует «речь» (предложение), «слово», «имена», но обходит, как нечто само собой разумеющееся и не проблемное, принадлежность их к сфере языка. Этой сферы с ее внутренней закономерностью для античной теории вовсе не существует. Категории бытия, мышления и языка осознаются античной философией в их единстве, но очень нечетко в их противоположности и почти неизменно сливаются.

В мифологическом мышлении имя принадлежало вещи. Античная наука разорвала эту связь, поместила между именем и вещью мысль, но продолжала соотносить изолированное слово, как полнозначного носителя смысловых отношений, непосредственно с отдельной вещью или мыслью. Слово – «значащий звук», и оно полностью определено своими двумя аспектами – фонетическим и семантическим. Строение слова, его сочетаемость с другими словами отражают строение вещи (или представления) и отношения между вещами в реальном мире (или в человеческом мышлении), зависят только от значения слова и ни в какой мере от строения языка как целого. Язык – не система, а агрегат; стало быть он не создает никакой специфической проблематики.

Теория стремится растворить языковую закономерность в онтологической или логической закономерности, и на долю языка остается очень мало специфического помимо его внешней формы.

Мы подробно привели эти выдержки, чтобы показать, что проблема корпоративного взаимопонимания, по сути, самоорганизующийся процесс в некотором замкнутом окружении и все существовавшие и обсуждающиеся проблемы коммуникативного общения людей лишь переносятся на специфику нового инструмента-

посредника в виде среды электронных и программных средств коммуникации. Приведем самый очевидный пример.

Латиница и Кириллица.

Развитие единого информационного пространства порождает проблему с национальными алфавитами и, в частности, с кириллицей. Традиционно сложившаяся поддержка на системном уровне только латинского алфавита порождает проблемы как с распространением русскоязычного программного обеспечения (ПО) на западный рынок, так и полноценное использование англоязычного ПО в России.

Для понимания проблемы национальных алфавитов вспомним старые печатные машинки, на которых, как правило, имелся только один набор национальных символов (только русский или только английский). Когда необходимо было напечатать символы «чужого» языка часто набирались символы похожие по начертанию: русская “Р” вместо английской “P”, русская “В” вместо английской “B”, цифра “1” вместо английской “I”, не говоря уже о фонетически похожих буквах (“Е”, “Т”, “К” и т.д.). Если для документа, единожды напечатанного на машинке, такой способ является приемлемым, то для документа электронного, с которым производится дальнейшая работа в электронном виде, такой способ уже не применим. Все дело здесь в неоднозначной идентификации конструкций, полученных при такой записи, ведь при поиске “10” (набрано цифрами) и “Ю” (набраны буквами) совершенно разные данные, хотя для человека они выглядят почти одинаково и интерпретируются в зависимости от контекста.

Даже понимая необходимость введения корректной идентификации национальных алфавитов, до сих пор продолжается использование кодировок подобных KOI-8, WIN-1251, которые подменяют одни национальные символы на другие и оказываются неработоспособны при необходимости одновременного использования разных национальных символов.

При этом существуют форматы, не зависящие от национальных особенностей данных. Таким форматом является звуковые форматы (MP3, OGG, PCM и др.). Если они передают речь, то для них не важно, на каком языке выражается говорящий, так как этот формат данных сохраняет физическую величину инвариантную по отношению к национальному языку передаваемого сообщения.

По-другому дело обстоит с форматами, требующими записи непосредственно языковых конструкций – текстовыми и документальными форматами. Здесь кроме необходимости корректного ввода (необходимости поддержки клавиатурой национальных символов и многоязыкового ввода) и корректного вывода (наличие соответствующих национальных шрифтов) требуется еще корректная идентификация национальных текстовых данных.

В качестве примера можно привести проблему сортировки, которая стояла перед программистами больших машин несколько лет назад, и которая в данный момент решена (проблема заключалась в прерывности последовательности национальных символов в кодировке, что требовало искусственного присвоения каждому символу дополнительного идентификатора, по которому производилась сортировка).

Существующие в настоящее время способы решения проблемы однозначности представления данных при использовании национальных алфавитов, в частности

кириллицы, до сих пор являются несовместимыми между собой. Появляющаяся в последнее время единая для всех национальных алфавитов кодировка Unicode охватывает только прикладной уровень программного обеспечения и не позволяет получить «чистого» национального программного обеспечения.

Проблема неоднозначности русского языка, в связи с которой потерпели неудачу попытки создания русскоязычных языков программирования и операционных систем с русскоязычным интерфейсом, а так же возможность многовариантности задания команды или оператора.

Возникает проблема русскоязычных информационных систем. И эта проблема усложняется и с точки зрения отсутствия компьютерной среды рассчитанной на структуру неконвенционального языка. Этот вопрос остается открытым, мы формулируем потребность в создании русскоязычного прототипа динамических структур семантических понятий, организации ассоциативных полей и связей. Основа концепции – иерархические рекурсивно-древовидные структуры как динамика построения понятийных областей были предложены еще в [9].

Даже осознавая необходимость поддержки национальных алфавитов иностранные разработчики делают системы в которых заявлена работоспособность с национальными алфавитами, которые без проблем работают с латиницей, но начинают неадекватно работать, когда в качестве системного имени или имени файла используется кириллическое название.

Текст и смысл

Проблема построения информационных систем заключается в монополизации информационных потоков. При этом во всех системах нужно заранее четко определить кто сигнифицирует, т.е. кто определяет границы всех терминов, определяет внутреннюю структуру предметной области, а также определяет интерфейсы взаимодействия с внешней средой.

Вот почему компьютерные информационные системы должны содержать тексты, построенные по принципу тройного сообщения: **символ-знак-идентификатор, форма представления и непосредственно содержание**. Это и есть первый шаг к интеллектуальному интерфейсу информационных систем, поддерживающему поиск совпадений и различий, ориентированного на поиск единой концепции запрограммированного рассеивания и концентрации семантико-смыслового поля.

Сошлемся на труды Деррида: *...разные имена объединяет одно существенное свойство*, которое Деррида именовал «логоцентризмом». Это слово обозначает традиционную для европейской культуры ориентацию на слово как на высший авторитет, и особенно на живое слово – залог присутствия говорящего человека. Раньше, допустим, говорилось: пусть данное явление значит то-то и то-то, а вот на самом деле за ним скрываются Бог, Бытие, Смысл жизни, Природа, Истина – в зависимости от ценностной установки всегда отыскивался некий скрытый сакральный конструкт, обеспечивающий мыслителю «мягкую посадку». Благодаря трудам Деррида философское доверие к подобным центрирующим мифологемам было утрачено: мифологемы оказались при ближайшем рассмотрении лингвистическими фикциями, получающими наполнение в зависимости от контекста. В противовес логоцентрической «метафизике присутствия»

Деррида выдвигал альтернативную систему категорий, главная из которых – письмо как особый дискурс, символический аналог мышления в целом. Его вызывающий тезис «Нет ничего кроме текста» («Il n'y a pas de hors-texte») подразумевал предельное расширение понятия текстуальности: вся жизнь становилась текстом.

В настоящее время одной из основных проблем в области информационных технологий является обучение адекватному описанию, упорядочению и восприятию информационных составляющих предметной области.

Интернет порождает интенсивный обмен информационными потоками, вследствие чего постоянно расширяется используемый словарный состав и меняется тезаурус профессиональных предметных областей. Это, в свою очередь, приводит к терминологическим сложностям при освоении информационных технологий.

Если терминологических соответствий много, невозможно осуществить автоматизированный перевод. Классические подходы к математической лингвистике и структурной лингвистике не решают эти вопросы.

Существует проблема неоднозначности соответствия понятий одного языка понятиям другого: ни одно слово не может быть описано двумя разными выражениями. В языке нет полных синонимов и нет точных соответствий между схожими по значению словами в разных языках. Синонимические понятия одного языка не соответствуют таковым в другом языке. Отсюда вытекает потребность в создании глоссария на том языке, который мы используем в процессе работы с определенными информационными системами.

Умберто Эко в [10] обратил внимание на различие между лингвистической составляющей языка и его коммуникативной функцией.

Семиология рассматривает совокупность информационных процессов как знаковые системы, предполагая, что они таковыми и являются, будучи, таким образом, также феноменами коммуникации. Тем самым она отвечает потребностям самых разнообразных современных научных дисциплин, как раз и пытающихся свести явления самого разного порядка к факту коммуникации. Психология изучает восприятие как факт коммуникации, генетика устанавливает коды наследственной информации, нейрофизиология описывает процесс передачи сигналов с периферии нервных окончаний к коре головного мозга, при этом все эти дисциплины неизбежно обращаются к математической теории **перекодирования**, которая и была создана для того, чтобы объяснить процесс передачи сигнала на уровне автоматов на основе общих положений физико-математических дисциплин. По мере своего развития такие науки, как кибернетика, занимающаяся вопросами контроля и управления автоматическими системами и электронно-вычислительной техникой, сблизилась с биологическими и неврологическими исследованиями. Одновременно коммуникативные модели находят все более широкое применение при изучении жизни общества, при этом на редкость эффективным оказывается сотрудничество структурной лингвистики и теории информации, предоставляющее возможность применять структурные и информационные модели при описании технологий, систем родства, кухни, моды, языка жестов, организации пространства и т. п., и даже эстетика иногда заимствует некоторые понятия теории коммуникации, используя их в своих целях. Наблюдается осязаемая потребность в унификации информационных полей, позволяющих описывать самые разнообразные

явления. Но если любой факт культуры это факт коммуникации и как таковой может быть рассмотрен по тем же параметрам, по которым рассматривается всякий коммуникативный акт, попробуем выделить элементарную коммуникативную структуру там, где имеет место самый, так сказать, примитивный случай коммуникации. Например, можно рассмотреть действие оператора (кассира, диспетчера) – передачу информации от одного простейшего автомата к другому. И вовсе не потому, что сложнейшие коммуникативные феномены, включая **эстетическую информацию, могут быть редуцированы к передаче сигнала от одной машины к другой, но потому что было бы полезно посмотреть, с чем мы имеем дело в самом простом и очевидном случае** коммуникации, построив ее образцовую модель. И только когда нам удастся это сделать, когда мы сможем выделить коммуникативную структуру, лежащую в основе и более сложных случаев коммуникации, только тогда мы будем вправе рассматривать любое **явление технологии или культуры с точки зрения коммуникации.**

То обстоятельство, что все науки, от лингвистики до нейрофизиологии, при описании коммуникативных процессов сводятся к бинарным представлениям, свидетельствует об адекватности описания моделей предметных областей при их компьютерном представлении. Когда мы идентифицируем, какое из двух событий имеет место, мы получаем информацию.

Отношение ряда событий к ряду соответствующих им возможностей – это отношение между арифметической и геометрической прогрессиями. Это означает, что при наличии 64 возможных исходов, когда, например, мы хотим узнать на какую из 64-х клеток шахматной доски пал выбор, мы получаем количество информации, равное шести. Иными словами, чтобы определить, какое из шестидесяти четырех равновероятных событий произошло, нам необходимо последовательно произвести шесть операций выбора из двух.

Мы свели идентификацию к позиционной структуре числа, представленного двоичными символами «0» и «1». Заметим, что эта операция идентификации не требует от числа никаких арифметических и математических свойств (например, аксиом Пеано). Так шахматисты пользуются идентификацией местоположения клеток на плоскости, используя структуру заданной упорядоченности числовой и буквенной нумерации. Этот подход распространен и на туристических картах, заменяя иногда буквенную нумерацию на римские числа. Прогресс же информационных технологий позволяет организовать идентификацию местоположения на оцифрованных картах с привязкой их к геофизическим информационным инвариантам (GPS, Глонас, сотовые терминалы и др.)

Мы приводим эти примеры, чтобы показать, что любые информационные системы содержат специфически упорядоченные структурные идентификаторы. Поэтому первичным элементом любой компьютерной информационной системы является адекватное отображение структуры данных профессиональных предметных областей в структуры данных, поддерживаемые программным окружением – средой компьютера, которая в свою очередь претерпевает изменения, адаптируясь к современным цифровым технологиям.

Когда мы имеем дело с машиной, мы не выходим за рамки кибернетики, а кибернетику интересуют только принципы обработки сигналов.

Но в процессе коммуникации – интерфейсном взаимодействии мы должны говорить не о мире сигнала, но о мире смысла.

С этого момента речь должна идти уже о процессе означивания, ведь в этом случае сигнал – это не просто ряд дискретных единиц, рассчитываемых в битах информации, но, скорее, значащая форма, которую адресат-человек должен наполнить значением.

Теперь нам следует определить семантико-смысловое содержание термина, которое лишь возникает из контекста.

Идя от значения к символу, мы получаем отношения именования (ономастика), когда какие-то смыслы привязываются к какому-то звуковому образу, и напротив, взяв точкой отсчета звучание, мы получаем отношения семасиологического порядка (какой-то звуковой образ получает определенное значение). К тому же, отношения между символом и его значением могут меняться: они могут разрастаться, усложняться, искажаться; символ может оставаться неизменным, тогда как значение может обогащаться или скудеть. Этот безостановочный Динамический процесс и следует интерпретировать, как "смысл".

Означающее – это не вещь (означающее "собака" – это не та собака, которую изучает зоология), и означающее это не ряд звучаний, составляющих имя (звукоряд "собака", изучаемый фонетикой и регистрируемый с помощью электромагнитной ленты). Означающее – это образ этого звукоряда, в то время как означаемое это образ вещи, рождающийся в уме и соотносящийся с другими такими же образами (например, дерево, arbor, tree, baum и т. д.)

Связь означающего и означаемого произвольна, но навязанная языком, который является кодом, она не может быть изменена по усмотрению говорящего. Напротив, именно необходимость подчиниться коду, протоколу, формату помогает уловить разницу между означаемым и понятием, мысленным представлением. Благодаря коду определенное означающее начинает соотноситься с определенным означаемым. И если потом это означающее принимает в голове у говорящего форму понятия или же воплощается в определенных навыках говорения, то это касается таких дисциплин, как психология и статистика. Парадоксальным образом, когда семиология, кажется, вот-вот определит означающее, в тот самый миг она рискует изменить самой себе, превратившись в логику, философию или метафизику. Семиология изучает не мыслительные операции означивания, но только коммуникативные конвенции как феномен культуры (в антропологическом смысле слова).

Не претендуя на исчерпывающее объяснение проблем коммуникации, семиология позволяет деконструировать профессионально предметное текстологическое описание в понятийно терминологическое – корпоративно консенсусное.

Язык – это общественный продукт речевой способности и вместе с тем система необходимых конвенций, принятых в том или ином обществе и обеспечивающих реализацию этой способности говорящими. Язык – это система, стало быть, структура, описываемая отвлеченно и представляющая собой совокупность отношений.

Идея языка как структуры посещала умы многих лингвистов прошлого. "Мы не должны понимать язык как нечто, начинающееся с обозначения разных объектов с помощью слов и составленное из слов. На деле не речь состоит из предшествующих ей слов, но наоборот, слова берут свое начало в речи". "Язык – это система, все части

которой можно и должно рассматривать в их синхронной взаимосвязи. Изменения, которые затрагивают всегда только тот или иной элемент системы, но не всю ее в целом, могут изучаться только вне целостной системы; действительно, любое изменение отражается на всей системе, но изначально оно происходит в какой-то одной точке и никак не связано с совокупностью вытекающих из него следствий, меняющих характер целого. **Это сущностное различие между следованием и сосуществованием, между частными и системными характеристиками не позволяет рассматривать те и другие в рамках одной науки".**

Код, когда мы имеем дело с языком, устанавливается и крепнет в процессе общения, являясь результатом общепринятых навыков говорения; и в тот миг, когда код устанавливается, каждый говорящий начинает неизбежно соотносить одни и те же значки с одними и теми же понятиями, комбинируя их по определенным правилам. Бывает, что код вводится, так сказать, сверху, каким-то авторитетным лицом, и становится обязательным для той или иной группы (например, азбука Морзе), и в этом случае код используется осознанно. В то время как другие коды, и среди них язык, будучи столь же обязательными, используются говорящими неосознанно, последние подчиняются им безотчетно, не ощущая своей зависимости от навязанной им жесткой системы правил.

В последнее время лингвисты обсуждают вопрос о том, как следует понимать этот код, как закрытую или как открытую систему, т. е. речь идет о том, как люди говорят, то ли они повинуются какой-то незыблемой системе правил, установленной некогда раз и навсегда и применяемой ими неосознанно, или же они говорят благодаря врожденной способности к формированию лингвистических последовательностей, обязанных своим происхождением некоторым простейшим комбинаторным принципам, предоставляющим возможность самых разнообразных сочетаний. В таком случае то, что принято называть системами и кодами, окажется не более чем поверхностной структурой, производной от какой-то глубинной структуры, системы правил, которая в отличие от других структур не может быть артикулирована с помощью оппозиций.

В самом деле, в пределах одного кода значение слова зависит от того, есть ли в языке другие слова со сходным или отличным значением. В итальянском языке слово "снег" нагружено разными значениями (белый снег, грязный снег, рыхлый снег, падающий, укрывающий землю, твердый, талый снег), между тем у некоторых групп эскимосов, судя по всему, разные значения слова "снег" закреплены за разными существительными. Итак, перед нами система отношений значений, в которой значение каждого слова устанавливается путем распределения смысловой нагрузки между всеми словами.

Так открывается возможность создания строго структурной семантики, которая, описывая значения, не принимала бы во внимание отношение означающее – означаемое. Можно заметить, что французское слово *arbre* охватывает тот же спектр значений, что и немецкое *Baum*, в то время как французское же слово *bois* соответствует итальянским *legno* и *bosco* одновременно, тогда как на долю *forêt* приходится значение "более густого лесного массива". Напротив, немецкое слово *Holz* значит *legno*, но не *bosco* и подводит значения *bosco* и *foresta* под общее наименование *Wald*. По иному распределяются значения соответствующих слов в датском языке, и не составит труда сравнить его с ситуацией в итальянском.

Французский	Немецкий	Датский	Итальянский
Arbre	Baum	trae	albero
Bois	Holz		legno
	Wald	skov	bosco
forêt			foresta

Если в художественном тексте неоднозначность устраняется за счет контекста, то идентификационные принципы информационных систем не имеют возможности для исправления неоднозначностей. Возникает ситуация «буриданова осла» с равноправным выбором из нескольких направлений.

Разные потребители, употребляя разные понятия, слова – идентификаторы к одному и тому же объекту постепенно приводят всю информационную систему к развалу с точки зрения запрограммированных возможностей.

В таких таблицах, как эта, мы имеем дело не с "идеями", но со значениями, обусловленными собственно языком как системой... Нет никакой нужды знать, что представляет собой означаемое (с онтологической или физической точки зрения): достаточно сказать, сославшись на код, что таким-то означаемым соответствуют такие-то означаемые. Нет ничего особенного в том, что впоследствии эти означаемые повсеместно воспринимаются как "понятия" или даже как "мысли", и то, что они могут быть объяснены правильным словоупотреблением, тоже вполне законно. Но в тот миг, когда семиология устанавливает факт наличия кода, значение уже не есть психологическая, онтологическая или социологическая данность: это факт культуры, который описывается с помощью системы отношений, устанавливаемой кодом и усваиваемой данным сообществом в данное время.

Теперь несколько слов о том, что такое структура. Структура – это модель, выстроенная с помощью некоторых упрощающих операций, которые позволяют рассматривать явление с одной единственной точки зрения. Структура – это способ действия, разрабатываемый для того, чтобы иметь возможность именовать сходным образом разные вещи.

Всякий раз, когда устанавливается общая структура некоего круга явлений, следует задаться вопросом, нет ли у данной структуры своей структуры, нет ли кода у этого кода, такого, который позволил бы охватить большое число явлений, расширить сферу применения кода.

Так, фонологи и лингвисты, описав систему какого-то языка, спрашивают себя, а на что это похоже и нельзя ли сравнить эту систему с системой другого языка, сыскав такой код, который стал бы их общим основанием. А затем, нет ли такого кода, который позволил бы выявить общность системы внутриязыковых отношений с системами родства, а эти, в свою очередь, связать с системой расположения домов в обследуемом селении. Этим успешно занимается структурная антропология.

Итак, переходя от упрощения к упрощению, структуралист грезит о том, чтобы, в конце концов, ему открылся Код Кодов, некий Пра-код, который позволил бы выявить скрытые ритмы (элементарные структуры), управляющие любым поведением, как культурным, так и биологическим. Этот код должен был бы воспроизводить устройство человеческого ума, которое окажется сходным с механизмом органических процессов. И по сути дела, сведение всякого человеческого поведения и всякой органики к коммуникативным процессам, а также сведение всякой коммуникации к двоичному выбору предполагает не что иное, как сведение любого культурного и биологического факта к одному и тому же механизму порождения.

Но это не означает, что к абстракциям структурных моделей приходят путем последовательного упрощения того, что и без того известно, напротив, чаще структурный метод не столько обнаруживает структуру, сколько выстраивает ее, изобретает в качестве гипотезы и теоретической модели и утверждает, что все изучаемые явления должны подчиняться устанавливаемой структурной закономерности.

Найти код – это и значит теоретически постулировать его. Конечно, каждый лингвист, прежде чем установить закономерности языка, изучает языковое поведение во всей его конкретности и многообразии; но он не в силах исчерпать все возможности индивидуального произношения, индивидуальной манеры говорить, индивидуальных намерений говорящего, стало быть, он должен бросить накапливать факты и заняться построением языковой системы.

Код – это модель, являющаяся результатом ряда условных упрощений, производимых ради того, чтобы обеспечить возможность передачи тех или иных сообщений.

Код – это структура, представленная в виде модели, выступающая как основополагающее правило при формировании ряда конкретных сообщений, которые именно благодаря этому и обретают способность быть сообщаемыми. Все коды могут быть сопоставлены между собой на базе общего кода, более простого и всеобъемлющего.

Понятийная база информационных систем

Казалось бы, в наш информационный XXI век общедоступным становится уровень символьных абстракций между именем-идентификатором объекта и его внутренним содержанием. Однако с развитием специализированных систем типа управления предприятием (ERP), логистической поддержки жизненного цикла изделия (CALS), систем управления базами данных (DBMS) их названия привязываются к сущностным характеристикам системы т.е. онтологиям (onthologic – гр. ontos (сущее) + logos (понятие)). На аналитические информационные системы переносятся их примитивные утилитарно-ограниченные функции. Опасность проявляется в неоправданно расширительном толковании полученных результатов в виде заявлений следующего типа «на основе базы знаний, экспертной системы, ассоциативного поиска, самореферирования или автоматического аналитического обзора принято решение о прогнозировании катастрофы и др.». При этом прогностического механизма – методологии и алгоритмов принятия решений на консервативно-упакованных данных – внутри компьютера не

дается. Знания о предыдущем опыте, как правило, распространяются на прогнозируемые события с неограниченным сроком.

Характерно, что именно с развитием информационных технологий (ИТ) все в большей степени встает проблема логистической поддержки. По сути, спонтанное возникновение в Интернет-сообществе различных словарей – попытка на интуитивном уровне преодолеть отсутствие **инфраструктур логистической поддержки**. Трудности возникают и с неадекватным пониманием самого термина логистики.

Приведем англо- и русскоязычную интерпретацию основных понятий, требующихся для дальнейшего изложения:

Антология – anthologia <гр. anthos – цветок + lego – собираю> – сборник различных статей. [11]

Онтология – onthologic <гр. on (ontos) – сущее + logos – понятие, учение> – греч. учение о существе или о сущности, бытии, сути. [12]

Логистика – Logistics – the methods of procuring distributing, maintaining, and replacing materiel and personnel, as in a military operation. [13] – методы доставки (доставания, сводничества), распределения, поддержки и размещения материалов (продуктов технологической деятельности).

Логистика – математическая алгебра, часть тактики, о передвижении войск. [12]

Логистика – полностью отрывает логические формы от его содержания. [11]

Лингвистика – Linguistics – the study of the nature and structure of human speech – учение о природе и структуре человеческой речи. [13]

Лингвистика (фр. Linguistique <лат. lingva – язык>) – языковедение, наука о языке, включающая фонетику, семантику, лексикологию, этимологию, грамматику и орфографию, орфоэпию. [14]

Семантика – Semantics – изучение смысла в языке.

Семиотика – Semiotics – наука о знаках, включая знаки-символы и знаки-сигналы.

Тезаурус – Thesaurus – отобранные слова или понятия, выполняет роль специализированного словаря, заданной предметной области, например медицина, музыка.

Глоссарий – Glossary – понятийно-терминологические определения слов, используемых в тезаурусе конкретной предметной области.

В современных компьютерных информационных системах логистическая поддержка является основным фактором их существования и проявляется в обязательном адекватном начальном владении достаточным базисом знаний.

Подчеркиваем, в компьютерных информационных системах возникает проблема адекватного восприятия различных понятийных знаний, которых не существовало при взаимной коммуникации типа непосредственного общения. Процесс непосредственного общения является процессом обучения и формирования адекватных профессиональных понятий. Барьер в виде компьютерно-информационной системы разрывает этот непосредственный естественный процесс. Каждый пользователь пытается внедрить в информационную систему свою понятийную базу. Необходим некий уровень, который поддерживал бы обновление понятийной профессиональной базы для корпоративного пользователя системы.

Мы в большей степени ориентируемся на терминологически-понятийное соответствие термина и содержания через стандартизованные глоссарии, которые также, в свою очередь, должны иметь механизм согласования между собой и различными предметными областями, что в настоящий момент и акцентируется как проблема логистической поддержки. Недаром логистическая поддержка CALS-технологии переформулирована в аналитическую информационную систему [15].

В данном пособии мы акцентируем внимание на первом шаге – фундаменте любой аналитической информационной системы – выявлении и организации понятийно-терминологического глоссария заданной предметной области на основе динамического визуального интерфейса.

Этот объективный фактор обусловлен необходимостью совмещать знания предметной области: проектирование автомобиля, самолета, разработка электрических схем и др. – со знанием специфики возможностей и ограничений компьютерно-ориентированных программных продуктов, созданных специалистами-программистами на базе своих знаний, понятийно-структурные связи которых, как правило, не совпадают со знаниями специалистов-предметников.

Приведем пример типового активно рекламируемого учебного курса для специалистов в области информационных технологий.

Курс предполагает, что слушатель использует информационные системы предприятия в своей каждодневной работе и качество их работы является важным для него; или слушатель является сотрудником ИТ-отдела предприятия, обеспечивающим работоспособность информационных систем.

Предполагается, что слушатели этого курса должны знать как минимум кириллицу и латиницу.

Курс дает следующие знания:

- понимание важности правильной организации работы в ИТ-подразделении и предоставления ИТ-услуг
- понимание важности системного подхода в управлении ИТ-инфраструктурой как для потребителей, так и для поставщиков ИТ-услуг
- понимание процессно-ориентированный подхода в организации работ и взаимодействии с основными подразделениями предприятия
- знание типовой модели ITIL и умение определять цели процессов, какие виды деятельности они включают в себя, принципы построения и взаимоотношения процессов между собой

Программа курса:

1. Введение

- Место ИТ в современном бизнесе
- Управление на основе процессов
- Особенности процессов управления сервисами
- Библиотека мирового передового опыта ITIL

2. Поддержка услуг

- Service Desk как современный вариант реализации функции Help Desk.
- Процесс Incident Management
- Процесс Problem Management

- Процесс Configuration Management
- Процесс Change Management
- Процесс Release Management

3. Предоставление услуг

- Основные характеристики процессов, входящих в разделы Поддержка и Предоставление услуг
- Процесс IT Service Level Management
- Финансовое управление ИТ услугами (процесс Financial Management for IT Services)
- Процесс Availability Management
- Процесс Capacity Management
- Управление непрерывностью ИТ услуг (процесс IT Service Continuity Management)

Из приведенного содержания курса видно, что основная задача в обучении сводится лишь к натаскиванию последовательности использования функциональных операторов, то есть технологии работы внутри программного продукта или технологии. Как это ориентировано и как это соотносится с конкретными информационными потоками производственного цикла не обсуждается, так как программисты, разрабатывающие программный продукт, редко владеют знаниями конкретных технологий производственного цикла. Например, «краснодеревщик» способен построить любой формат «шкатулок» и «полочек», но их размер и форма возникают лишь из **«вещного значения»** предметов, под которые они создаются. Изучение компьютерно-информационных технологий и их применение в большей степени должно быть привязано к конкретным предметным областям, обладающим собственной структурой информационных описаний. Отсюда и проблема запутанности терминологической базы.

Рассмотрим данную проблему на примере ряда учебных пособий [16-20] основные понятия и определения (рис. 2-5), которые иллюстрируются на разработанном визуальном динамическом интерфейсе [8]. Они выбраны исходя из наличия их оцифрованного электронного представления.

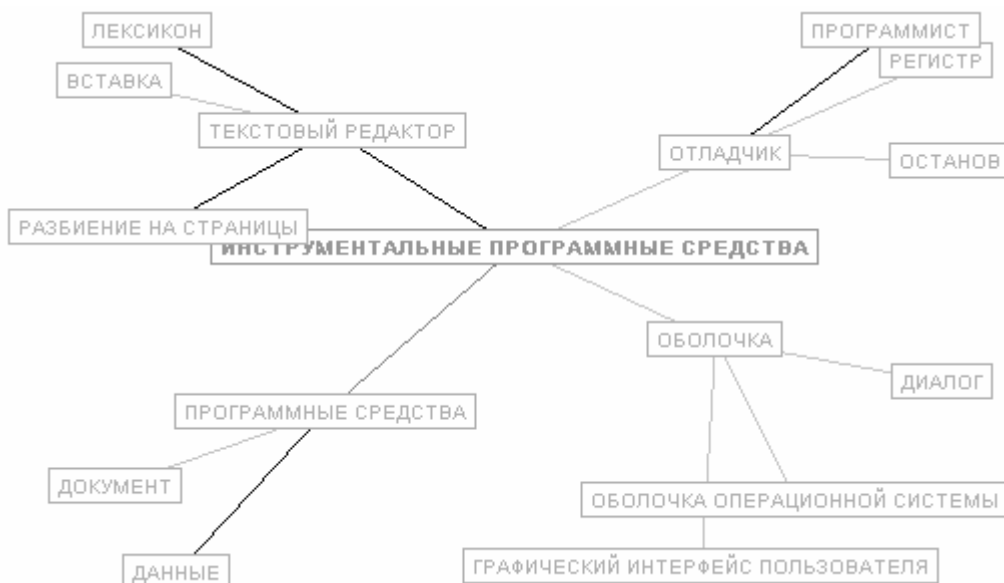


Рис. 2 Инструментальные программные средства



Рис. 3 Информационная система

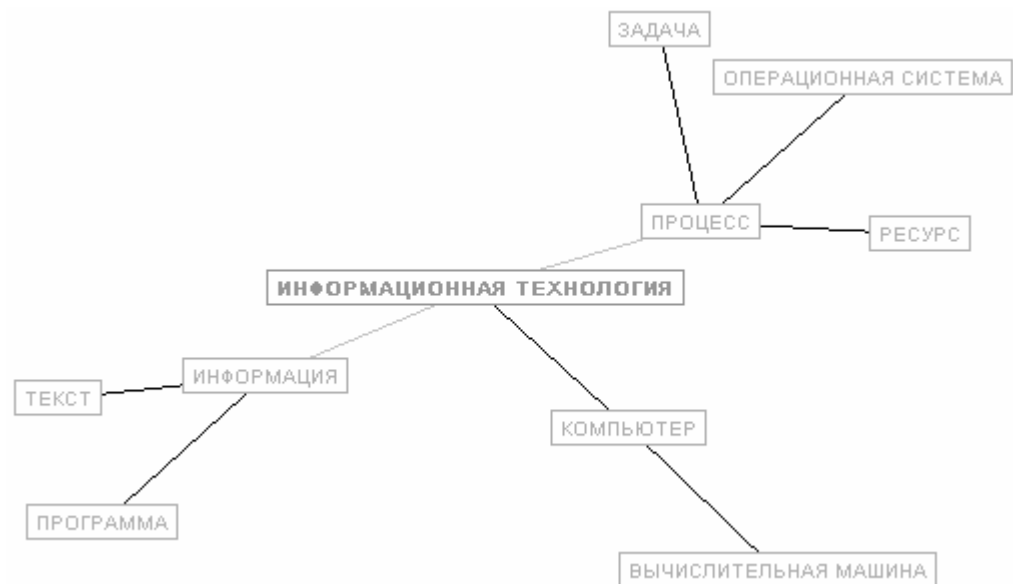


Рис. 4 Информационная технология

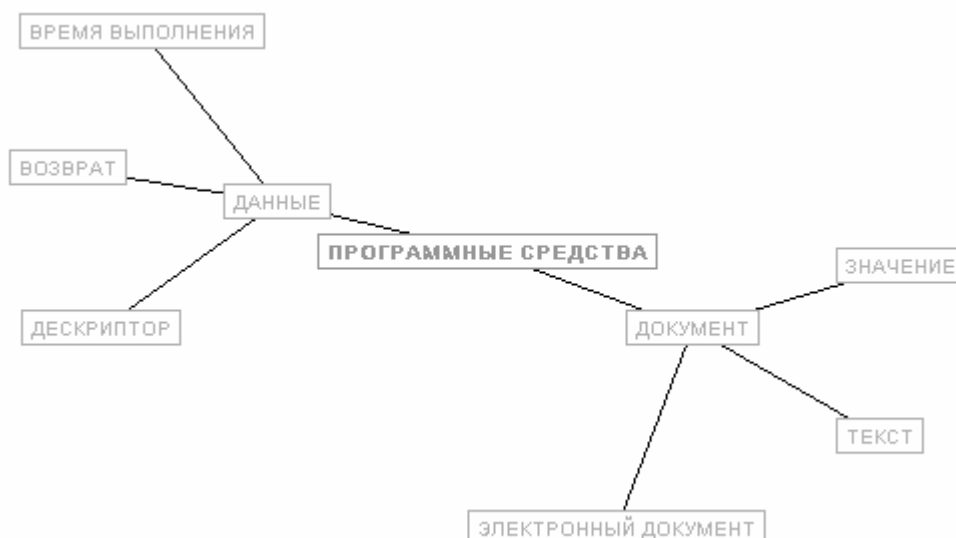


Рис. 5 Программные средства

Совершенно ясно, что учебные курсы должны быть ориентированны на обучение по выбранной специальности, которая и должна быть отражена в структурной совместимости водимых понятий и преемственности определений и развиваемых понятий. К сожалению, следует констатировать, что на приведенных примерах полученные понятийно-структурные связи отражают лишь семантику компьютерных информационных технологий и почти никак не связаны с понятийным окружением предметной области, например, тематике – интеллектуальные системы управления. В настоящий момент это связано в том числе и с тем, что специфика тематики интеллектуальных систем управления недостаточно отражена в учебно-справочном оцифрованном материале.

С точки зрения учебного процесса, любой учебник обязан содержать ключевые слова и понятия, то есть иметь на выходе тезаурус и глоссарий. За время обучения студент набирает определенный понятийный словарь относительно дисциплин, составляющих тематику выбранного курса.

Постоянная модификация информационных систем происходит из развивающейся терминологической неоднозначности между именем и содержанием. Происходит сужение области применимости, так как под разными словами понимают различное сущностное содержание. Происходит разрыв между именем и содержанием объекта. Информационные же системы проектируются под содержание, а работают с именами. **Возникает известный парадокс размножения вербальных сущностей.**

Закон для всех имен, наделенных семантикой, состоит в том, что развитие их понятий опирается на предыдущие имена с расширением семантико-смыслового содержания и потребности расширения информационного поля.

Образ или событие любой природы имеют собственную семантику. Эта собственная семантика универсальна и контекстно-интерпретируема в некотором конечном наборе смыслов или событий. Это утверждение основано на возможности осмысления, хотя возможно и не однозначного, практически любого образа или события

внешнего мира, даже лишенного собственного контекста. Вот почему писатели, фантасты, сказочники всегда опережают научно-технические открытия и инженерные решения.

Например, исследования, связанные со спектральным анализом и фильтрацией в той или иной мере направлены на контекстный поиск в потоке данных. Но это поиск в контексте задачи, во имя которого мы отрешаемся от абстрактного восприятия данных. Что же мешает нам вести исследования “от образа”, сохраняя тем самым собственную семантику потока данных? Что мешает искать “абстрактный фильтр”? Здесь хорошо видны две причины. Во-первых, априорно выбранная физико-математическая модель среды, причем выбор определен прикладными аспектами разработки и технологическими возможностями. Так, например техническая реализация колебательного контура на элементах LRC поддерживает и поддерживала актуальность математической модели на основе Фурье анализа. Транспьютеры и микропроцессоры уже требуют другую математическую среду при моделировании.

Во-вторых, значительные трудности организации системного анализа потоков данных с целью выявления интерпретируемых в различных контекстах структурных образований, их накоплении, хранении, а также организации быстрого поиска на уровне общепринятых структур данных.

Последнее означает, что без базы знания, ориентированной на семантический анализ структур данных, из информационного описания значительное продвижение в области обработки информационных потоков представляется невозможным [10].

Ясно, что базы знания сами по себе без специальных механизмов предварительного анализа не являются работоспособными образованиями. На рис. 6 приведены два основных механизма актуализации знаний: от данных к знаниям и от знаний к фактам – данным.

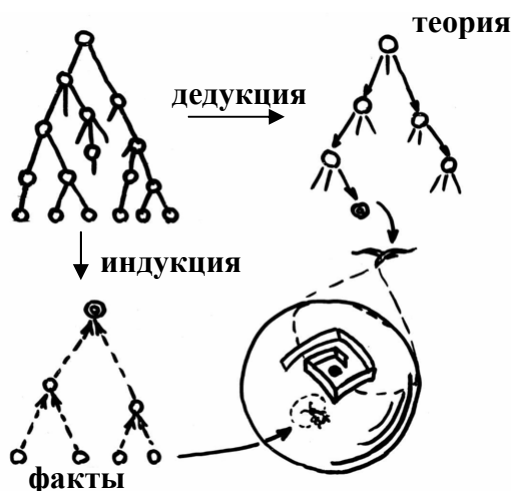


Рис. 6. Представление человеческих знаний. Дедукция и индукция.

Здесь снова возникает проблема соответствия проблемно-ориентированных сред уровню рассмотрения проблемы и их возможностей работы со смыслом. Например, цифровая форма представления сигнала опирается только на абстрактное понятие полосы пропускания и никак не учитывает семантику того, что он несет, – текст, музыку или изображение.

Для уяснения некоторых положений, связанных с понятием смысла, нам придется сейчас отказаться от обобщенного понятия контекста и перейти к рассмотрению трех исходных отношений заключенных в предложениях [7].

Начнем с того, что **событие по самой сути принадлежит языку**. Именно события выражаются или могут быть выражены, высказываются или могут быть высказаны в возможных предложениях. Внутри предложения много отношений. Какие же из них составляют события?

В предложениях выделяют три типа отношений.

Первое принято называть денотацией (обозначением). Процедура денотации заключается в соединении слов с именем конкретного понятия – образа предмета.

Второе отношение называют манифестацией. Это отношение выражает семантико-смысловое отношение к объекту. Следовательно, манифестация представляется как высказывание казуальных причинно-следственных отношений, соответствующих данному предложению.

Третье отношение называют сигнификацией (значением). Оно определяется порядком понятийных импликаций.

Первое и третье отношения уже давно используются для создания проблемно-ориентированных сред. В частности, авторы могут сослаться на монографию [7], где именно в этом ключе был сформирован язык среды, отвечающий и структуре объекта и информационному представлению для компьютерной обработки. Важность второго отношения связана со следующим: первенство манифестации – заявления в отношении не только денотации, но и сигнификации должно быть понято только в порядке контекста, где значения естественным образом содержатся имплицитно от денотации через манифестацию к сигнификации и обратно, рекурсивно воспроизводя сопутствующие понятия, которые и составляет семантику информационного поля.

Отождествление смысла с сигнификацией, ведущее к парадоксу Кэрролла «думай о смысле, слова придут сами», согласно которому значение никогда не играет роли последнего основания, поскольку само неизбежно зависит от всех трех приведенных выше отношений. В этом случае по Расселу: «...все утверждаемое – как наделенное смыслом высказывание – обладает некоторым видом возможности...».

Тогда условие истины можно определять уже не как форму концептуальной возможности, а как некоторый абстрактный объект, то есть не как сигнификацию, а как **смысл** – «голосуй или проиграешь».

Смысл – это четвертое отношение предложения – чистое событие.

Поскольку смысл – это отнюдь не только лишь один из двух терминов дуальности, противопоставляющей имя и свойства объекта. Обычно такого рода проблемы рассматриваются через серии парадоксов, из которых мы здесь упомянем парадокс неопределенного размножения.

Когда мы что-либо обозначаем, мы исходим из того, что смысл понят, что он уже налицо. Мы не переходим от звуков к образам и от образов к смыслу: мы с самого начала помещены в смысл. Смысл предписывает возможные обозначения и условия. Более того, смысл также является в свою очередь объектом следующего предложения. Тогда если принять предложение за некое имя, то ясно, что каждое имя, обозначающее объект, само может стать объектом нового имени, обозначающего его смысл: **n1** отсылает к **n2**, которое

обозначает смысл **n1**; **n2** отсылает к **n3** и так далее. Для каждого из своих имен язык должен содержать некоторое имя для смысла этого имени. Такое бесконечное размножение вербальных сущностей известно как парадокс Фреге.

В сжатой формулировке он звучит так: пусть дано множество *C* всех множеств, не содержащих самих себя в качестве своего элемента; тогда если *C* не принадлежит *C*, то по определению *C*, *C* принадлежит *C*; если же *C* принадлежит *C*, то по определению *C*, *C* не принадлежит *C*. Известно, что в теоретико-множественном варианте с этим парадоксом работали практически все выдающиеся ученые, ибо он относится к самым началам теории множеств.

В варианте смысл-объект этот парадокс, который мы будем называть все-таки парадоксом Фреге, раскрывает “место поиска смысла”, указывает момент возникновения того, что позже стали называть контекстной смысловой зависимостью. Парадокс Фреге используется нами как идейная основа развивающихся структур связи данных в проблемно-ориентированных средах, где эти связи, в отличие от произвольных сред, имеют выраженный характер через ассоциативно организованную идентификацию и допускают проведение **событийной, то есть смысловой, структуризации**.

Именно этим человек отличается от других животных, это и принято считать абстрактным мышлением, и это по мере возможностей пытаются искусственно воспроизвести, опираясь на компьютерно-информационные технологии. Абстрактное мышление опосредуется емкой знаковой системой, вне которой обобщающие абстракции не возможны – не будучи закреплены знаком, они быстро “расшатываются” и стираются. С момента же наречения, связывания свойств с именем, понятие обретает устойчивость и дееспособность (становится возможным применять его в логических операциях). Аспект динамики структурности шире того, что принято выводить под термином самоорганизация, поскольку язык также охватывает всякие процессы разупорядочения, разрушения, в том числе не укладывающиеся в русло естественной эволюции, и стало быть, предполагает широкий анализ деструктивных проявлений homo sapiens – будь то в аспекте психопатогенеза личности, (суб)культуры или в плане прогрессирующего вытеснения человеком других животных видов, подрыва баланса фито-зоомасс.

Для характеристики состояний структуризации: упорядоченности и разупорядоченности, используются такие понятия как: форма, структура, система, сложность, и им оппозиитные: бесформенность, бесструктурность, деформация неделимость и другие. Рассмотрим их кратко.

Форма (в русском языке встречается также греческий корень морфэ) и структура – весьма близкие термины. Форма – исходное в истории философии понятие, ведущее к анализу структуры объекта. С одной стороны, оно означает наружный вид вещи, внешние очертания, с другой – строение, внутренняя организация содержания.

Наиболее развитую концепцию формы построил в свое время Аристотель. Он заметил, что всю реальность можно свести к последовательным переходам от материи к форме и обратно, причем форма выступает «сутью бытия», его «первой сущностью». Именно она дарует определенность и действительность вещам в противоположность материи – неопределенной, бесформенной лишь возможности существования.

Структура, и система используются для характеристики упорядоченного целого – как совокупности элементов и их связей. При этом понятие структура (по лат. строение)

практически означает «совокупность устойчивых связей объекта», а в понятии система акцент делается на качестве образования быть целостным (само слово по-гречески означает целое, составленное из частей), это понятие более полно охватывает разнохарактерные связи и взаимодействия целого и включает элементы.

Структура выражает то, что остается устойчивым, (относительно) неизменным при всех преобразованиях системы. Другими словами, это инвариант (-ный аспект) системы. Структура выступает интегрирующим фактором системы и детерминирует ее качество.

Термин «система» используется для именования практически любого образования, проявляющего признаки упорядочения, и включает «жесткий», инвариантный и «флюидный», вариантный аспекты упорядочения. Возьмем, например, определение системы, данное одним из основоположников общей теории систем Л. Берталанди, как «комплекса взаимодействующих элементов» (к исходным относится и такое: система есть ограниченное множество взаимодействующих элементов). Под эти определения подпадает как одноклеточный или многоклеточный высокоразвитый организм, так и, пожалуй, немного спрессованное содержимое мусорной корзины.

Итак, структура предполагает «жесткий», высококогерентный тип упорядочения, структура – это всегда порядок. Начиная работать с некоторой предметной областью, как правило, заранее не известен объем и весь словарь информационных требований и ограничений для инструментальных средств, которые потребуются в процессе работы. На пример, желание коллекционировать альбомы по живописи должно быть согласовано, по крайней мере, по размеру с книжными полками, купив полки для обычных книг, можно попасть в ситуацию, когда ни один альбом (как правило, имеющий большой формат) не поместится на полку.

В подобной ситуации и оказываются специалисты, работающие в сфере информационных технологий, выбирающие инструментальные средства для работы с информационными потоками конкретной предметной области, зачастую не имея опыта работы в этой области и отсутствием знаний о динамике ее информационного развития.

Сегодня проблема стоит уже не в создании информационных систем, а в адаптации существующих программных средств к потребностям пользователя.

На данный момент создано огромное количество информационных систем общего назначения для различных задач обработки данных, чтобы создавать наукоемкий информационный продукт: СУБД, CAD-системы, CALS-системы. Пусть читателя не пугают разные предназначения перечисленных систем – все они по своей сути являются информационными системами, в большей или меньшей мере адаптированными под некоторый класс задач, суть которых – перелопачивать мусор потока данных в информационно-значимый продукт.

Как правило, все компьютерные информационные системы содержат блок под названием СУБД и лишь окружающий интерфейс адаптирован под конкретного пользователя и технологию, создающую интегрированную информационную среду. Конкретные производственные системы, работая с инфраструктурой внутренних связей ее области, по сути носит служебный внутрикомпьютерный характер стандартных протоколов и лишь форма входных и выходных документов (отчетов), их понятийный словарь – глоссарий обязан быть приведен к терминологии технологии производства и его

обслуживающего персонала. Структурная иерархия технологического процесса должна отражаться в структуре информационного содержания, обмена.

Компьютерные информационные системы, по сути, ориентированы на классы задач, такие как универсальные оболочки САД (работа с чертежной документацией), САЕ (инженерные расчеты) и т.д.

Появление англоязычного термина “реинженеринг” – это попытка адаптировать технологическую структуру производства к компьютерным информационным возможностям.

Однако проектировщики этих систем – программисты левополушарные с мышлением вербального свойства. И естественно, их продукты с трудом воспринимаются правополушарными людьми, т.е. специалистами, имеющими дело с интуитивно-образным, объемным моделированием: дизайнерами, архитекторами и др.

Из различия ролей: программиста – разработчика системы, пользователя – специалиста в предметной области, системного интегратора и проистекают различия как в тезаурусах, используемых для работы с информационной системой, так и в потребностях, которым система должна удовлетворять (рис. 7).

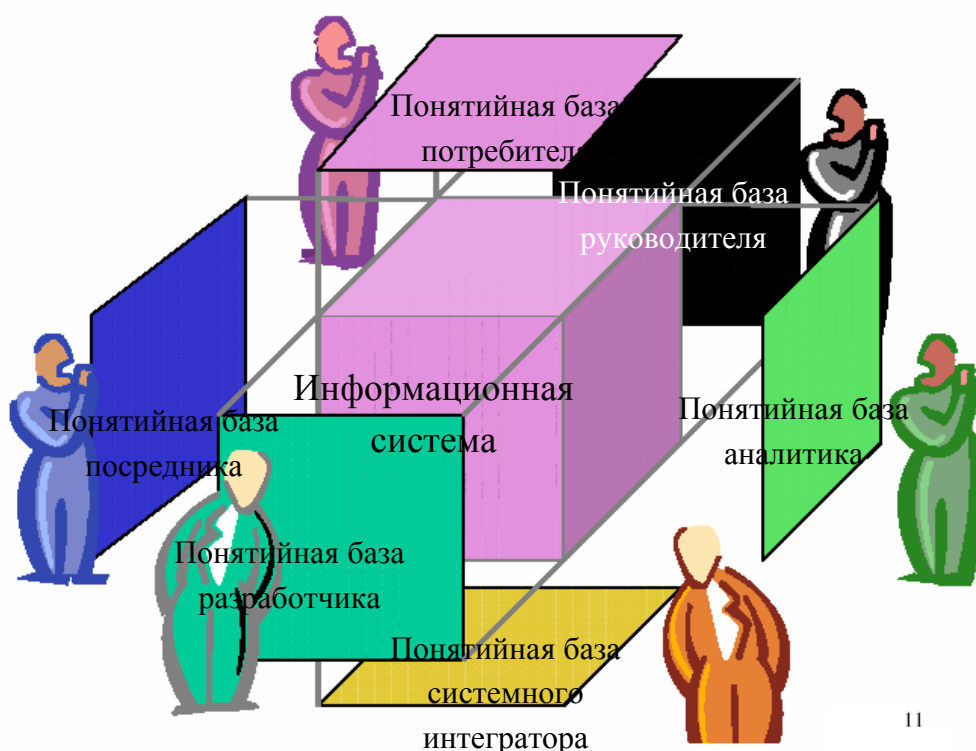


Рис. 7. Различие понятийных баз пользователей информационной системы.

Различные пользователи-потребители информационных систем используют различные термины, но все вместе они размывают предметную область. Любая же существующая программная система жестко постулирует технологический терминологический базис (тезаурус системы) – не только их словарь, но и понятийные

связи. Поэтому, даже если пользователи имеют общий тезаурус, все равно возникают разночтения по полученным смысловым результатам.

Подобные проблемы кроме естественного запутывания потребителей, затрудняют процесс внедрения некоторых программных средств, например предиктивного ввода (автоматическое дописывание слова из словаря по введенным начальным символам):

...Канадская компания WordLogic выпустила новую программу Predictive Keyboard, которая позволяет определять слово, вводимое пользователем, уже по первым буквам и автоматически дописывает его до конца. В сущности, Predictive Keyboard, анализируя первые нажатые клавиши, пытается угадать то, что пользователь хочет напечатать, и предлагает возможные варианты. По мере того как пользователь продолжает вводить следующие буквы, число вариантов сокращается, и в специальном окошке появляется наиболее подходящее слово...

Чтобы для пользователя – специалиста в некоторой прикладной предметной области (например, конструктора двигателей автомобиля) использование информационной системы помогало выполнению основной профессиональной деятельности, а не нагружало необходимостью дополнительного изучения целого ряда компьютерных дисциплин, необходимы следующие условия. Интерфейс системы должен полностью соответствовать профессиональному глоссарию, то есть оперировать теми словами, которые использовал бы специалист, не прибегая к помощи компьютера, а не изобилловать подсказками типа «введите уникальный идентификатор изделия» в случае необходимости простого указания названия детали. Последовательность операций для получения «виртуального» результата с помощью компьютера должна соответствовать последовательности операций, которые бы пришлось проделать без «виртуального помощника».

Соблюдение этих принципов позволяет перейти от лозунга, который соответствовал недавней эпохе «Сделаем каждого слесаря программистом, а потом научим их работать по специальности» к лозунгу современных систем «Дадим каждому такую систему, в которой бы он мог работать и без знания внутренних «потрохов» компьютера».

Книга под названием «Руководство пользователя по системе XXX» или «Краткий курс по YYY», а уж тем более «ZZZ для чайника» никогда не смогут научить правильно мыслить, а лишь привязывают человека к конкретной версии конкретной программы, давая объяснения типа «в случае А нажмите В» или «чтобы сделать С, выполните D».

Возможность переходить от одной программной среды к другой не читая руководств пользователя, кажущаяся некоторым «высшей степенью хакерского искусства» заключается всего лишь в понимании процессов лежащих в основе любой информационной системы и мышление на более высоком уровне абстракции и обобщения, чем традиционно предоставленный интерфейс пользователя.

Наличие у системы толстой книжки-руководства признак не наличия у компании производителя хорошей службы поддержки, а плохого уровня интерфейса пользователя. Наличие толстой книги – руководства и специализированных курсов свидетельствует лишь о том, что требуется дополнительно объяснять основы функционирования программы, вследствие не правильно организованного интерфейса.

Информационная система с хорошим интерфейсом не должна требовать книги для начала работы с ней. В свою очередь, хорошая книга (прочтение которой должно повышать эффективность работы пользователя с системой) должна рассказывать концепцию системы на самом верхнем уровне общности, не опускаясь до объяснения примеров, в какой момент какую кнопку нужно нажимать.

Интерфейсные возможности компьютера все в большей степени приближаются к вокабулярным, вокальным и зрительным способностям человека и механистический принцип функционирования Робота – не навреди следует заменить на интеллектуальную функцию компьютера помощи. Например, семантическая неоднозначность восприятия команды авиадиспетчера «направо» или «налево», зеркального отображения направления полета приводит самолет к катастрофе, несмотря на наличие автомата аварийного сближения. Возникает проблема веры и доверия человеку и способности современных информационных технологий корректировать и устранять как ошибки, так и злой умысел.

Развитие информационных компьютерно-ориентированных систем привело к терминологическим сложностям и проблеме адекватного под задачу выбора программных сред и автоматизированных систем типа CAD, CAE, CALS, PDM и др. Более подробно анализа путаницы терминологических понятий великолепно рассмотрена в [15.]. Там же рассматривается и новый подход к решению проблемы – создания ALS, проекта «Автоматической логистической поддержки», которая предназначена взять на себя функцию информационной поддержки процессов компьютерного проектирования, производства и сопровождения изделий на протяжении всего жизненного цикла.

Отметим, что понятие «логистики» в рассматриваемом контексте англоязычного документа отличается от принятого в настоящее время в русскоязычной компьютерно-ориентированном информационном сообществе.

Стоит ли после этого удивляться заявлению ведущего специалиста британской фирмы «Кинетик»: «В последний раз я слышал о CALS восемь лет тому назад», сделанному на встрече в Лондоне. Ведь, как специалист по логистическим системам, он мог понимать под CALS именно подобные системы, восемь лет назад попавшие в невидимую часть «CALS-айсберга».

Вполне понятным становится и то, что после 1994 г. успешно развивавшаяся программа CALS вдруг куда-то исчезла, оставив после себя лишь серию отдельных работ типа «Computer Aided Logistics Support», рассыпанных по различным программам и контрактам Министерство обороны США с промышленными фирмами. (Этот факт и был воспринят в среде русских специалистов как изменение первоначальной направленности программы CALS). Наиболее весомая часть этих работ оказалась в рамках программы JSF (Joint Strike Fighter), начатой все в том же «поворотном» 1994 г. Именно в этой программе сосредоточились все работы по созданию Министерством обороны автоматизированных систем логистического обеспечения эксплуатации одного из видов военной техники, конкретно, тактических боевых самолетов.

В период 1994-2001 гг. только прямые расходы на проведение в рамках программы JSF работ по логистике составили больше 150 млн. А в 2001 г. боровшиеся за получение основного контракта фирмы «Боинг» и «Локхид Мартин» со своими проектами боевого самолета JSF представили в конкурсную комиссию и предложения по

автоматизированной системе логистического обеспечения (рис. 8). Она получила обозначение ALS (Autonomic Logistics Support).

По аналогии с этим и «Life-Cycle» в развернутом определении CALS-технологии следует понимать не как весь жизненный цикл изделия, а как стадию его эксплуатации. Ведь стадии разработки и производства уже учтены термином «acquisition». А это приводит к необходимости уточнить еще одно слово в рассматриваемом определении CALS-технологий – «continuous». Да, его можно перевести как «непрерывный», имея в виду постоянство выполнения каких-либо функций по времени реализации процесса в целом (в данном случае по времени реализации всего жизненного цикла изделия). Но американцы явно имели в виду другое [15]:

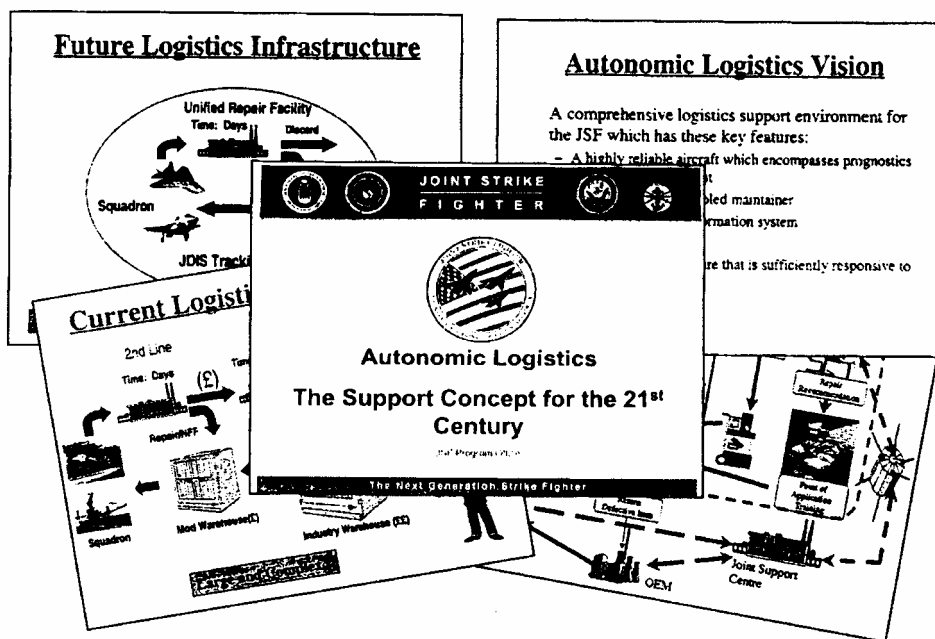


Рис. 8. Автоматизированная система логистического обеспечения ALS пока еще на бумаге, но это только «пока».

Однако попытки создания следующего уровня информационной системы, которые бы помогали разобраться с предыдущими версиями программных систем, приводят к известной проблеме о «Вавилонской башне».

...Человек мечтает о том, чтобы построить каменную башню “высотой до небес”, о том, чтобы создать на земле идеальную жизнь. Люди верят в интеллектуальные методы, в технические средства, в формальные утверждения. Неизбежно наступает момент, когда люди перестают понимать друг друга. Каждый человек воздвигает свою Вавилонскую башню: его стремления, жизненные цели, достижения – все это Вавилонская башня. **Вся** жизнь человека, накопление богатств, приобретение власти или **знаний** – все это постройка Вавилонской башни, ибо должно закончиться катастрофой. Неизбежна потребность перехода на новый план бытия под которым мы и понимаем научно-технический прогресс... [21]

Миф о «Вавилонской башне» или «Вечном двигателе» – устойчивое желание человека к «глобализации» приводит к мифологеме нереализуемости проекта в целом в заданной трактовке без ограничений.

Ограничения на языковые формы общения хорошо рассмотрены у фантаста Роберта Шекли («Потолкуем малость»), которые мы сформируем в следующем виде. **Скорость изменения – диахронизм языка должен быть ниже темпа овладения им. Вследствие этого естественный язык и содержит механизм демпфирования – ограничения скорости изменения словарного состава за счет своих структурных построений и иерархизации понятий в виде построения тезаурусов и глоссариев.**

Любая информационная система терминологически очень быстро переживает сама себя, так как **инкапсулирует** в себе словарный запас разработчика и не развивается вместе с развитием естественного языка.

Естественный язык – пример открытой развивающейся системы, исторически и эмпирически доказавшей свою эффективность в части информационно-коммуникативного общения и реализующей функции аутентификации, идентификации и самореферирования.

Но как далеки современные авторы статей, компьютерных программных продуктов и автоматизированных информационных систем от исходных понятий употребляемых слов, которые и появились как потребность в расширительном толковании, интерпретации накопленных сведений и знаний.

К сожалению, коммерческая потребность заставляет производителей коммерческих программных продуктов придумывать все новые и новые красивые мифологизированные названия.

История коммуникативных функций языка.

Развитие языка, как средства общения и представления знания, основано на коммуникативной потребности одновременно в единообразном и расширительном толковании и интерпретации накопленных сведений и знаний.

Так в IV в. до н.э Аристотель писал:

...Всякое слово может быть или общеупотребительным, или глоссой, или метафорой, или украшением, или вновь сочиненным, или растяженным, или сокращенным, или измененным.

Общеупотребительным я называю такое, которым пользуются все, а глоссой – которым пользуются немногие. Ясно, что глоссой и общеупотребительным может быть одно и то же слово, но не у одних и тех же людей.

Метафора – перенесение слова с измененным значением из рода в вид, из вида в род, или из вида в вид, или в форме пропорции. Я говорю о перенесении из рода в вид, например «а корабль мой вот стоит», так как стоять на якоре – это особый вид понятия «стоять»...

...Слова семантически лишь по «договору», они не являются «орудиями» и не заключают в себе ничего «природного»...

Поскольку имя непосредственно принадлежит вещи, для архаического мышления нет надобности относить имена к какой-либо сфере, отличной от сферы бытия вещей. Язык как целое есть лишь совокупность имен, которая может быть противопоставлена совокупности имен чужого языка, но не включает в себе ничего специфически-

языкового, несвойственного самим вещам, и не рождает никаких проблем, кроме вопроса об отношении отдельных имен к отдельным предметам, «правильности» наименования. Различие в именах есть различие вещей, сходство и близость вещей обнаруживается в сходстве и близости этимологически сопоставляемых имен, как они были установлены ономотетом, одним или многими. Таков итог размышлений над языком ко времени разложения мифологического мирозерцания и то идейное наследие, которое греческая философия получила в вопросах языка от предшествующих стадий истории мысли.

И хотя греческая мифология не приписывает ни одному из своих персонажей подобных функций, самое представление об «установителе имен», «ономотете», засвидетельствовано философами. Древнепифагорейское изречение гласит: *что мудрее всего? Число, а на втором месте тот, кто положил вещам имена...*[6]

Следующий отрывок из книги «Язык в современном мире» характеризует состояние вопроса коммуникативной функции языка с позиций лингвистов на 1960 г., ниже приводятся основные необходимые нам цитаты.

По наблюдениям Эдварда Сафира, «мысленный мир – это миниатюра, которую человек носит с собой, она является мерой и пониманием того, как он оценивает то, на что способен в этом мире». Этой теории, в дальнейшем, предавалось большое значение, она много изучалась знаменитым американским философом Бенжамином Ли Ворфом (1897-1941), который глубоко изучил хопи, шоуни, нутку, и другие языки аборигенов. Так называемые гипотезы Ворфа, о том, что «мировоззрение человека предопределяется той языковой средой, в которой он был воспитан», возможно, были преувеличены его многочисленными сторонниками, но и сейчас очень редкие филологи будут отрицать эти свойства.

В то же время, те же филологи все еще захвачены идеей о том, что то, или иное значение может быть «структурировано». Эта идея в течение какого-то времени обсуждалась участниками восьмого конгресса лингвистов, проходившего в г. Осло в 1957 г., спустя 60 лет после представления Майклом Брилом своего труда о значениях и публикации «эссе о семантике, наука о значении». Является ли это структурой значений, существующей в самом языке, или это – рамка, надетая на язык «снаружи»? Как только мы отбросим все формы, уйдем от соответствий фонетики и морфологии через полные соответствия синтаксиса и стилистики к этой сфере соответствий, которые мы называем семантикой. Мы находимся на том же уровне, что и Джейр Эймос Каменски (или Комениус) был в семнадцатом веке, когда он попытался разделить все знания на разделы, передвигаясь от идей и вещей к формам, вместо (как в привычном словаре или энциклопедии) от формы к идее и мысли. Очень давно, греки пытались сделать нечто подобное, их попытки явились миру в виде Ономастикона во втором веке нашей эры. Оригинальный текст этой первой попытки утрачен, но из вторичных источников мы узнаем, что это исходило от богов к человеку, человеческому телу, роду, науке, искусству, охоте, пище, производству, праву, правлению и утвари. Недавно, мечта Комениуса о завершенной «картине мира мыслей», частично материализовалась Питером Марком Рогетом и его последователями, чей практический «тезаурус английских слов и фраз» сейчас широко используется в исправленном виде в Великобритании и в дополненном виде под редакцией Сильвестра Моусона в США. Этот мир идей первую очередь упорядочен в шесть широких классов абстрактных взаимоотношений, пространство,

материя, интеллект, воля и любовь. Они, в свою очередь, разделены и подразделены. Но как читатель может управлять своим путем через неизведанное море абстракций? «Компас» предлагается в виде алфавитного указателя, который неизбежно занимает значимую часть книги. Современный Рогет имеет много общего с предметными указателями библиотечных каталогов, в особенности с теми, которые основаны на десятичной классификации Деви, предложенном Мелвином Деви в 1873 году, его дополнили издатели британской национальной библиографии. По его схеме, знания подразделяются на десять основных секторов, каждый из которых определяется цифрами от 0 до 9: современная философия, религия, социология, филология, наука, полезное искусство, прикладное искусство, литература и история. Каждый из этих секторов в дальнейшем подразделяется на десять подсекторов, и так далее. Филология или лингвистика, на своем месте между социальными и физическими науками.

Анализ естественного языка (ЕЯ), автоматическое понимание ЕЯ имеют давнюю историю с множеством подходов, которые очень часто противоречили друг другу, хотя критерий пригодности здесь только один – решение задачи понимания ЕЯ. К слову сказать, проблема машинного понимания ЕЯ не решена до сих пор. Почему? Искусственный интеллект – слишком общее и туманное понятие. То же справедливо и для обработки ЕЯ как одной из важных составляющих мира искусственного интеллекта. Задачи понимания ЕЯ как таковой нет. Есть множество довольно частных задач – и многие из них уже имеют приемлемые решения. Возьмем, к примеру, машинный перевод. Анекдотическое качество существующих систем перевода иллюстрирует ту идею, что орешек не по зубам, по крайней мере, пока. Но если выделить из проблемы наиболее актуальные частные задачи, и решать каждую из них в отдельности, можно получить приемлемое качество за приемлемую цену.

Современный компьютерный инфо-коммуникационный интерфейс, является надстройкой над миром Тьюринга, программная реализация которого опирается на греко-латинский язык, в котором первична грамматика на которой строится текст, а затем возникает смысл.

Все естественные языки проявили удивительную способность к понятийно-ассоциативной идентификации. Изменения в окружающей среде – развитие компьютерных информационных технологий и телекоммуникационных сетей – привело к резким изменениям коммуникативных функций. Это вызвало не просто резкое изменение тезауруса, но и определенные требования к языку, конструкция которого либо адекватна современным коммуникативным возможностям либо нет.

Современный компьютер основан на понятии машины Тьюринга, которая эквивалентна и практически реализует понятия: вычислимых функций и алгоритма [22], а также предложенной Г. Лейбницем конгруэнтности (сопоставления, взаимосвязи) между объектами – неизвестному объекту приписываются свойства известного. Принцип функционирования компьютера одна из основных причин, почему операционная система и интерфейсная поддержка пользователя в своей основе опирается на греко-латинскую модель языка с его строгой грамматической конструкцией. Когда грамматика как бы алгоритм, формирующий текст, а текст инициализирует смысл. В то время как русский язык принципиально не **конвенциален**, т.е. первичен смысл, отображенный в тексте, а затем уже грамматический разбор.

...Русский язык, несмотря на свою схожесть с европейскими языками по существу обладает не просто иной языковой структурой, но даже во многом противоположной структуре европейских языков. Это выражается в том, что русский язык в своем существе принципиально не конвенционален. Его суть проявляет себя в трактовке языкового закона. Русский язык всегда предполагает лишь неграмматический способ его освоения. Обучение ему реализовывалось изнутри самого языка в качестве системы отказов от его книжных форм и системы подмен книжных форм русского языка на книжные формы... [23].

Чтобы соединить два слова в русском, нужны значительные усилия. Тут и согласование по роду, числу и падежу, и глагольное управление. Весь труд уходит на синтаксис и морфологию, а не на семантику.

Приведем следующие цитаты из Дж. Оруэлла «Английский язык» [24]

...Английский язык это обширнейший вокабуляр и простота грамматического строя. Для нужд интернационального общения английский может быть сведен к простейшему «птичьему» языку в диапазоне от «бейсик инглиш» до «бичламара» на котором изъясняются в южной части Тихого океана. Таким образом, он соответствует функции инструмента общения народов разных стран...

Поэтому легко было «бейсик инглиш» превратить в язык программирования «Бэйсик». Зато обширнейший вокабуляр все еще не доступен компьютеру для реализации качественного пользовательского речевого интерфейса, т.е. анализ и синтез речи максимально приближенный к возможностям человека все еще ждет своего решения.

Онтологическая информационная система

Формирование информационной системы проходит через следующие этапы (рис. 9). Первая фаза любого добросовестного работника, специалиста-профессионала, аналитика и ученого для заданной темы выбрать уже опубликованные основополагающие тексты, т.е. осуществить антологию. Антология – сборник тематически-ориентированных текстов и является платформой составления тезаурусов и глоссариев. Результат же информационного анализа и исследования, представленный как краткий реферат, и составляет суть понятия онтология.

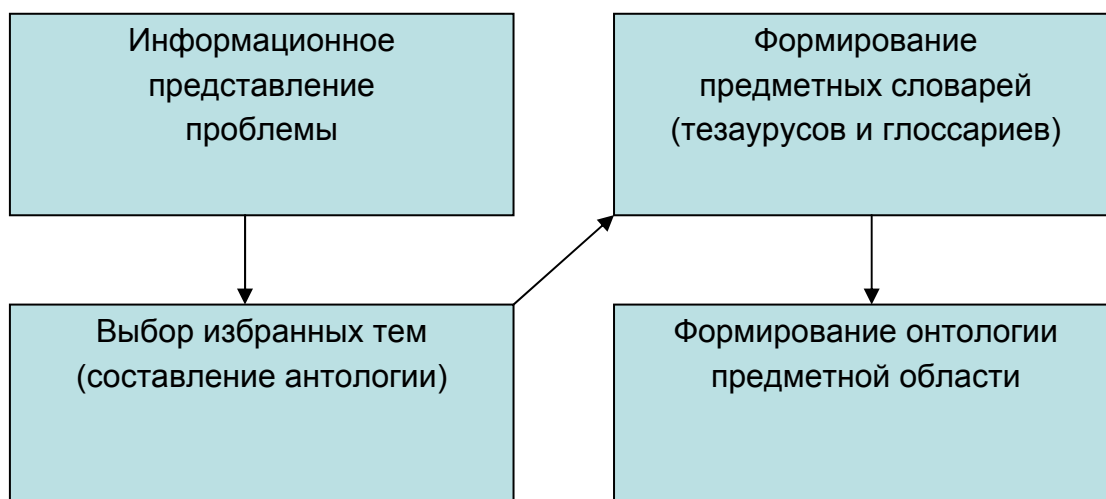


Рис. 9. Этапы создания онтологической информационной системы

Структуры данных определяют применимость информационной системы к различным задачам. Структуры данных, как правило, определяются при анализе задачи исходя из структуры информационных потоков. В худшем случае такая система будет применима только к той задаче, под которую она разрабатывалась и именно в тех условиях, которые были при постановке задачи. В лучшем случае, если рассмотрен более общий случай возможных условий применимости, систему возможно применить под класс задач, сходных с исходной (например, автоматизированное рабочее место продавца супермаркета).

Другая крайность, связанная с созданием системы в «слишком общем случае», когда разработчик создает систему с элементами, обладающими очень высоким уровнем абстракции. При этом такие элементы при задании их параметров могут быть адаптированы под конкретную задачу. Недостаток таких систем заключается в излишней сложности настройки таких систем, когда требуется определить все параметры этой системы в соответствии с задачей.

Применимость системы во многом определяется базовыми типами данных, которые представляют содержащуюся в системе информацию. Очевидно, что информационная система, работающая с данными векторной графики (например, некоторая геоинформационная система) будет не эффективна при использовании ее в качестве электронной библиотеки художественных текстов, хотя при очень большом желании и соответствующих затратах ее можно заставить работать и в таком режиме. Менее очевидно такое сравнение для систем, работающих с различными разновидностями текстовых и числовых табличных данных.

Табличные данные и многомерные данные в компьютерных информационных системах на основе баз данных – это одномерная цепочка табличных отношений. В рамках реляционной алгебры (а эти рамки автоматически возникают, как только в составе системы появляется реляционные базы данных) любые многомерные данные есть объединение элементов отношений, выявленных и разбитых при декомпозиции задачи. В связи с этим можно отметить, что жаргонное употребление «многовекторное направление» или «многовекторный поиск» не имеет никакого смысла.

Проследить различие в применимости разных систем к разным задачам можно на примере различия платформ Wintel (Windows + Intel) и Mac (Apple Macintosh). В связи с тем, что компьютеры Apple изначально разрабатывались с учетом традиций, принятых в издательском деле (например, разрешение экрана соответствует пропорциям стандартного бумажного листа) компьютеры Mac стали стандартом *de facto* в типографиях и дизайнерском деле. При этом принципиально схемотехнически и логически платформы практически равноценны, и лишь наличие совокупности мелких условий становится определяющим в выборе предпочтения в использовании платформы в конкретной задаче.

Visual Thesaurus

В настоящее время осуществляется прямые попытки воспроизводить «бумажные» технологии представления юридической, нормативно-правовой, финансовой и прочей документации, а также книг, учебников и энциклопедий в электронной форме, что вызывает значительные трудности как непосредственно при создании для компьютера

программного продукта, так и для его восприятия пользователем. Понимание этого факта уже привело к появлению в Интернет программных продуктов, адаптированных под пользователя и приближающегося к реализации структурной схемы на рис. 10.

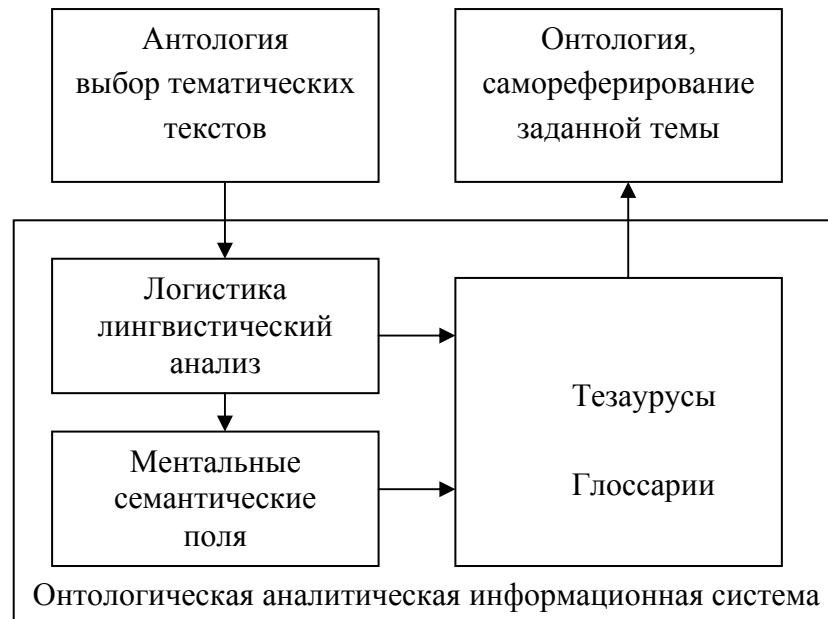


Рис. 10. Структура онтологической аналитической информационной системы.

Одним из примеров реализации системы является Visual Thesaurus [25]. Ее возможность динамического развития семантических понятий основана на системе WordNet – ментальной модели лексикона человека для английского языка. Лексикон (греч. *lexikon* – словарь) – запас слов, выражений.

Заметим, что эта несложная программная адаптация интерфейса в динамическую структуру визуальных понятий в корне меняет и значительно интенсифицирует процесс обучения, освоения и применения данного продукта. Однако, обратим особое внимание, что этот эффект достигается, если заранее тезаурусы предметных областей и глоссарии используют не энциклопедические словари, а сами построены на новой информационной платформе. Для англоязычной аудитории это WordNet, электронный тезаурус, отражающий все возможные толкования слов английского языка, и показывающий взаимосвязи между ними, который является одним из наиболее авторитетных и широко используемых стандартов для построения лексико-семантических баз данных.. Методология с которой разрабатывался и создавался WordNet была предложена в 1985 г. Дж. Миллером и его коллегами из Лаборатории когнитологии Принстонского Университета (США) на основе модели ментального лексикона человека. Описание структуры WordNet имеется в [26].

Популярность и широкое распространение WordNet обусловлены, прежде всего, его существенными содержательными и структурными характеристиками. Принстонский WordNet и все последующие варианты для других языков направлены на отображение состава и структуры лексической системы языка в целом, а не отдельных тематических областей. Например, настоящая версия WordNet охватывает общеупотребительную лексику современного английского языка – более 120 тысяч слов. Словарь состоит из 4 отдельных блоков для основных знаменательных частей речи: существительных,

глаголов, прилагательных и наречий. Базовой структурной единицей Принстонского WordNet является синонимический ряд (синсет), объединяющий слова со схожим значением. Предполагается, что каждый синсет репрезентирует в словаре некоторое лексикализованное понятие данного языка. Для удобства использования словаря человеком каждый синсет дополнен дефиницией и примерами употребления слов в контексте. Синсеты в WordNet связаны между собой такими семантическими отношениями, как гипонимия (родовидовое), меронимия (часть-целое), лексический вывод (каузация, пресуппозиция) и др.; среди них особую роль играет гипонимия: она позволяет организовывать синсеты в иерархические структуры (деревья). Лексика каждой части речи представлена в виде набора деревьев (леса). Для разных частей речи родовидовые отношения могут иметь дополнительные характеристики и различаться областью охвата, например, только некоторые группы прилагательных в WordNet связаны гипонимическими отношениями.

Аналогом термина «синсет» в отечественных источниках можно считать такие понятия, как класс эквивалентности или семантический класс. Наглядно, понятие синсета может быть представлено лексической матрицей, рис. 11.

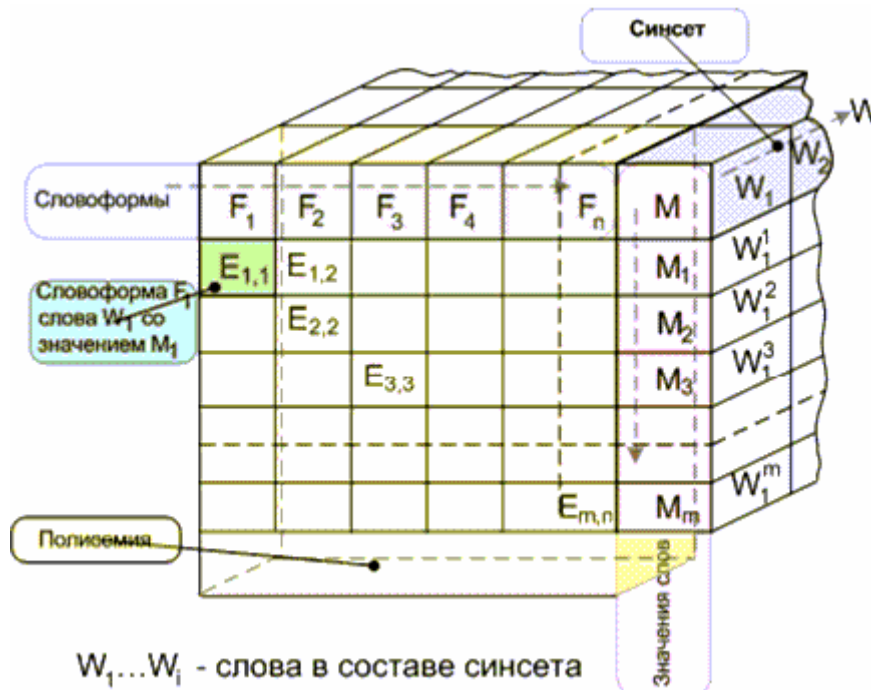


Рис. 11. Структура WordNet – модель семантических понятий

В виде синсетов (synset) представлены группы существительных, глаголов, прилагательных и наречий, между которыми определяются различные отношения. Именованные связи между словами синсетов позволяют выделять антонимы для заданного слова. Каждый синсет содержит список синонимов или синонимичных словосочетаний и указатели, описывающие отношения между ним и другими синсетами. Эти указатели позволяют устанавливать различные отношения: иерархические, соответствующие иерархиям род/вид, часть/целое; не иерархические, в частности определяющие атрибуты для значений некоторых синсетов, например, вес {большой, маленький}. Слова в синсетах логически сгруппированы так, что могут быть

взаимозаменяемыми в некотором контексте. Можно выделить три определенных уровня классификации, рис. 12.



Рис. 12. Уровни классификации в WordNet

Для каждого синсета однозначно определяется его положение в грамматической и лексической классификации. Слово или словосочетание может появляться более чем в одном синсете и более чем в одной синтаксической категории (части речи). Слова, для которых существует омонимия и полисемия одновременно включаются в несколько синсетов и могут быть причислены к различным синтаксическим и лексическим классам.

Стандартная процедура построения WordNet предполагает обращение к частотным словарям языка для выбора употребительной лексики, с которой обычно начинается построение тезауруса. Наполнение структур осуществляется следующим образом: сначала исследуется ядро лексики русского языка – наиболее частотные слова с наиболее общим значением, затем полученные иерархические структуры расширяются за счет менее употребительной лексики. Исследование семантических связей лексикализованных понятий осуществляется в рамках лексико-семантических групп, набор и состав которых определяется составом ядерной части тезауруса, при этом используются методы дефиниционного, контекстного и деривационного анализа. Каждый синсет сопровождается словесным описанием его значения. Это позволяет для слова составить графический портрет семантических соответствий и, следовательно, выделить контекстное содержание текста (рис. 13).

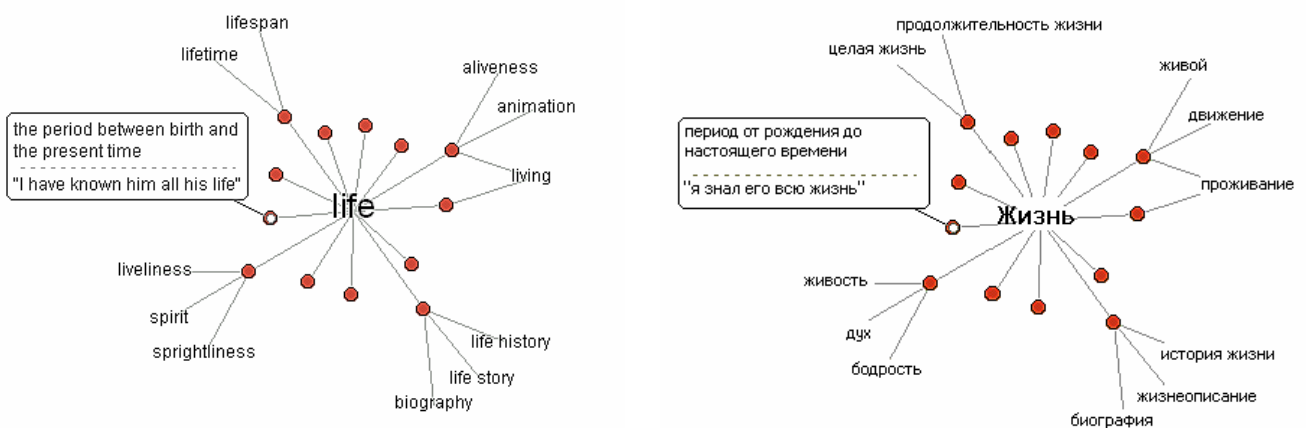


Рис. 13. Понятийное развитие слова и создание контекстного окружения.

Подобная организация ментальных семантических полей и ориентирована на реализацию таких интеллектуальных функций как ассоциативный поиск, самореферирование, идентификация и аутентификация информационных источников.

Динамическая визуализация контекста (Визуальный словарь)

Необходимость разработки русскоязычного прототипа WordNet – ментальной модели лексикона человека для построения предметно-ориентированных семантических полей как основы конструирования русскоязычных тезаурусов и глоссариев проявилась в реализации ряда проектов.

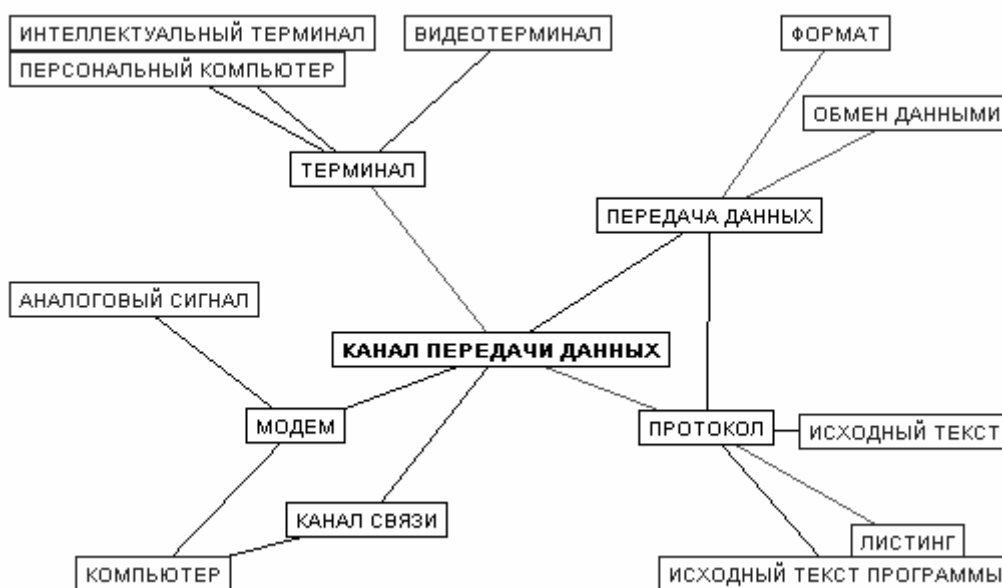


Рис.14. Контекстное окружение в системе «Визуальный словарь» на примере предметной области «информационные технологии».

Одним из таких проектов является визуальная понятийно-терминологическая информационная система «Визуальный словарь» [8].

В этой системе предлагается методология построения понятийного окружения понятий предметной области путем формирования связей между зависимыми терминами. Дополнение информационной системы интерактивным визуальным интерфейсом позволяет создать интерактивную среду для быстрого ознакомления с предметной областью.

Глоссарием в рамках семиологической информационной системы является множество терминов, являющееся базисом предметной области в котором для каждого элемента построены его связи с другими терминами предметной области. Связями назовем отношение на множестве терминов, дающее связь между определяемыми и определяющими словами.

Для создания глоссария используется антология – сборник тематически-ориентированных текстов.

Ниже приводится сокращенное описание алгоритма автоматизированного формирования глоссария заданной предметной области.

1. Формирование общего словаря из антологии для создания множества всех слов во всех формах.

2. Формирование списка терминов – анализ антологии (текстов по предметной области) с целью выделения терминологического базиса для предметной области. Предварительно сформированное множество терминов уточняется на стадии статистической обработки. Отдельным проходом алгоритма выделяются термины, состоящие из нескольких слов.

3. Морфологическая обработка словаря и приведение слов в нормальную форму на основании общего словаря с помощью модуля морфологического анализа.

4. Формирование семантической сети – многопроходный алгоритм формирующий множество связей между элементами списка терминов на основании контента и структуры текстов антологии. Реализация процедуры принятия решения включения новой связи, использует механизм сравнения слов в начальных формах, причем приоритет отдается терминам, состоящим из нескольких слов.

5. Определение ранговых распределений для списка терминов – формирование индексов частоты использования терминов в обрабатываемой предметной области и частоты использования связей между терминами.

6. Реструктуризация семантической сети на основе ранговых распределений терминов, словаря стоп-слов; уточнение состава базиса предметной области. Также на этом этапе проводится формирование иерархии связей на основе опциональных параметров, задающих количество уровней иерархии и ограничение количества элементов на одном уровне, а также стратегию фильтрации при ограничениях. Ограничения влияют только на визуализацию и позволяют более наглядно представить термины, являющиеся центрами большого количества связей.

7. Формирование визуального представления системы путем заполнения интерфейсно-независимых xml-ориентированных шаблонов для последующего формирования конечного информационного ресурса.

Для функционирования поисковых возможностей также формируется индексный файл.

Экспериментальная версия система подготовки данных представляет собой интегрированный программный комплекс, обеспечивающий выполнение всех этапов алгоритма.

В результате работы система формирует следующие объекты:

- множество всех терминов, используемых в данной предметной области, ранжированных по степени информативности с толкованием их значения (в множестве исключается наличие стоп-слов и служебных слов языка),
- семантическую модель предметной области, описывающую отношения между понятиями,
- список терминов, которые не удалось включить в семантическую модель, т.е. связать с другими из-за недостаточности соответствующей информации в тексте.

Визуальное представление, формируемое семантическими системами типа «визуального словаря» или «Visual Thesaurus» являются аналогом иероглифической записи, которая позволяет воспринимать содержимое текста не последовательно, а

одномоментно. Это позволяет воспринимать структуру связей предметной области в комплексе, притом именно в том, который соответствует связям, сформированным специалистом или разработчиком системы, а не формировать его самостоятельно при прочтении груды технической документации.

Кроме наглядного представления связей предложенный подход позволяет оценивать качество текстов антологии.

Об числовых характеристиках текста – рейтинговых распределениях, способных отнести его к определенной категории (художественный текст, технический текст, текст на формализованном языке, простой перечень слов) говорилось в [9].

Самым очевидным критерием является «разваливание» семантического окружения на большое количество независимых областей. Это означает слабую связность внутри текста, «перескакивание» с одних терминов на другие, отсутствие общей темы, зашумленность терминологического базиса. Другими показателями качества являются скорость развития – величина, показывающая скорость введения новых терминов и понятий в тексте. При слишком малой скорости развития текст кажется скучным, при слишком высокой – непонятным. Эмпирически установлено [9], что оптимальные значения лежат вблизи скорости роста равной 1,618.

Конечно, автоматическая оценка качества текстов не является строгой и может давать ошибки, но ее использование статистически оправдано, например, при упорядочивании результатов в поиске при равной релевантности нескольких документов. В этом случае раньше в списке найденного встретятся те документы, показатель качества которых выше.

Информационные системы.

Применение современных информационных технологий является одним из главных инструментов повышения эффективности промышленного производства. Анализ мирового опыта показывает, что в основе создания и сопровождения конкурентных высокотехнологичных изделий лежит стратегия использования непрерывной информационной поддержки жизненного цикла изделия. Применение указанных технологий позволяет существенно сократить стоимость наукоемкой продукции и затраты на ее разработку и производство, уменьшить сроки вывода новых изделий на рынок, значительно повысить качество производимой продукции при одновременном повышении удобства эксплуатации.

Реализация концепции непрерывной поддержки жизненного цикла изделия предполагает создание единой информационной среды, охватывающей всю информацию об изделии, созданную на этапах изучения рынка, составления технического задания, проектирования, технической подготовки производства, выпуска продукции, упаковки и хранения, реализации, установки и ввода в эксплуатацию, технической поддержки и обслуживания и утилизации, управление информационными потоками в ходе различных этапов и исполняемыми в рамках этих потоков заданиями.

Информационная среда – это совокупность технических и программных средств хранения, обработки и передачи данных, а также организационно-методические меры реализации процессов информатизации (рис. 15).

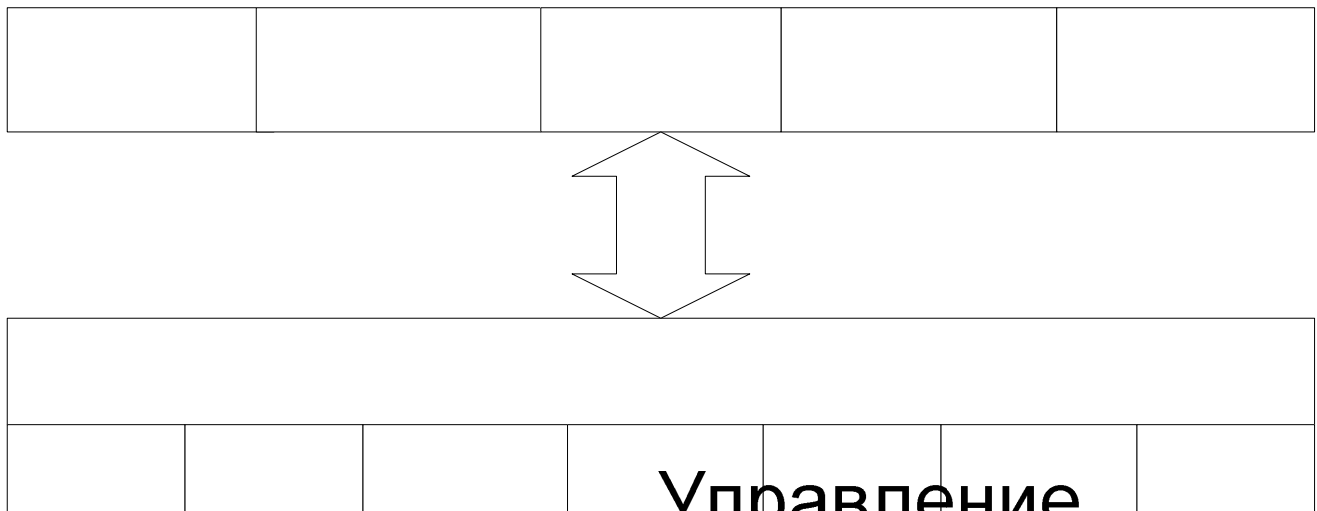


Рис. 15. Информационная среда.

В основе комплексного решения такой задачи лежит управление всеми документами предприятия на всех этапах их жизненного цикла независимо от их предназначения, формата, способа хранения и использования. Крупнейшие мировые промышленные корпорации строят такие решения на платформах, обеспечивающих полный спектр потребностей по управлению информационными потоками крупных предприятий, входящих в структуры холдингов и взаимодействующих в процессе своей деятельности с предприятиями-поставщиками, заказчиками и партнерами. На основе единой платформы строятся хранилища организационно-распорядительных, нормативно-справочных, проектно-конструкторских, технологических, эксплуатационных, финансовых и прочих документов, создается распределенная корпоративная система делопроизводства, автоматизируются основные и вспомогательные бизнес-процессы внутри предприятия и процессы его взаимодействия с партнерами, поставщиками и клиентами, процессы публикации информации на сайтах в Интернет и корпоративных информационных порталах. При этом обеспечивается:

- Возможность автоматизации на единой платформе полного спектра задач, связанных с управлением неструктурированными документами.
- Поддержка распределенности и высокая масштабируемость системы, проверенная мировым опытом использования в крупнейших корпорациях.
- Открытость системы — возможность интеграции с ERP, CRM, CAD/CAM-системами и интеграция с типовым пользовательским интерфейсом.
- Возможность функционирования в различном аппаратном и программном окружении – поддержка различных операционных систем.
- Безопасность, обеспечиваемая как многоуровневой защитой доступа к документам, так и возможностью подключения внешних средств криптографии и электронной подписи российской разработки.
- Мощные административные возможности для централизованного управления распределенной системой.
- Широкий набор средств разработки приложений.

Документы

Медиа
данные

Основным преимуществом построения информационной системы непрерывной поддержки жизненного цикла экспортно-ориентированной наукоемкой продукции является соответствие требованиям международных отраслевых стандартов АТА, АЕСМА и стандартов CALS в части подготовки эксплуатационной документации в виде Интерактивных Электронных Технических Руководств.

Interactive Electronic Technical Manual — Интерактивное Электронное Техническое Руководство (ИТЭР) — это интерактивная интеллектуальная среда доступа к большим объемам графической и текстовой технической информации, которая представляет собой полное электронное техническое руководство, связанное с текстом, рисунками, фотографиями и видео. Стандарты определяют требования к структуре ИТЭР, номенклатуре составляющих документов, стандартной системе нумерации разделов и подразделов ИТЭР, кодированию атрибутов документов, а также определяют структуру и типы документов.

В настоящее время для изделий, поставляемых на экспорт, ИТЭР создается по требованию иностранных заказчиков путем перевода в электронный вид существующей бумажной документации. Такой подход отличается высокой трудоемкостью разработки и, в особенности, поддержки изменений в ИТЭР в течение всего жизненного цикла изделия. Дополнительные трудности возникают при необходимости создания и поддержки многоязычных версий ИТЭР.

Под информационной системой понимается многокомпонентная система, направленная на удовлетворение потребностей потребителей информации и работающая с электронными документами. Составные части информационных систем:

- Программно-аппаратное обеспечение
 - программные комплексы
 - аппаратные средства вычислительной техники и связи
- Информационное обеспечение
 - словари; тезаурусы и классификаторы
 - схемы и их описания; эксплуатационная и сопроводительная документация
- Методическое обеспечение
 - формальные модели деловых процессов
 - инструкции и методики
- Организационное обеспечение
 - положения, уставы; должностные инструкции

Классификация систем по масштабу применения

- *“Коробочные” системы.* Основные заказчики — малые предприятия. Представляют собой некое законченное решение, разработанное под заказ, либо выполняющее типовой набор задач, который может подойти многим предприятиям.
 - *Best, Галактика, БОСС*
- *Системы групповой работы.* Основные заказчики — средние предприятия. Более гибкие механизмы настройки и среда для разработки.
 - *PC Docs, Lotus Notes, 1C, MySQL, Allaire JRun*

- *Системы корпоративного уровня.* Данный класс систем ориентирован на крупные предприятия, где будут работать тысячи пользователей, и будут обрабатываться большие объемы информации. Используются корпоративные стандарты в области хранения и представления информации, а также в области настройки логики и интерфейсов.
 - *SAP R3, PeopleSoft, Documentum, Interwoven, BEA*

Требования к системам корпоративного уровня

- *Производительность.* Системы обрабатывают огромные объемы информации, с которыми могут работать десятки тысяч человек. Очень важно, чтобы с повышением числа пользователей и объема информации, производительность оставалась на прежнем уровне.
- *Распределенность.* Большинство крупных организаций имеют большую степень территориальной распределенности. Необходимо, чтобы была поддержка возможности обмена и синхронизации информации между удаленными системами. При этом каждая из систем может иметь различные показатели производительности и объемы обрабатываемой информации.
- *Открытость.* Система должна иметь инструменты для возможности интеграции с другими системами. Это одно из важнейших требований к такому классу систем, без которого не представляется возможным создание правильно функционирующей модели документооборота.
- *Гибкость.* От того, насколько тонко можно настраивать систему, зависит тот объем задач, который можно решить при помощи нее. Очень часто ценность системы определяется именно этим фактором.
- *Модульность.* Во многих случаях заказчику нужен лишь узкий спектр функций, выполняемых системой. В данном случае, очевидно, что система должна состоять из отдельных модулей, которые могут функционировать как вместе, так и раздельно.
- *Многоплатформенность.* Система должна обеспечивать свою работу на большинстве используемых в крупных предприятиях платформах, при этом быть платформенно-независимой, т.е. работая под одной платформой, уметь взаимодействовать с системой, работающей под другой.
- *Защищенность.* Система должна быть защищенной от вторжений извне с целью получить какую-либо информацию. Одновременно с этим система должна обеспечивать разделение данных между пользователями в соответствии с их полномочиями на просмотр или модификацию определенных видов данных.

Рассмотрим подробнее, что понимается под информационной поддержкой жизненного цикла изделия.

Обычная процедура движения продукта от производителя к потребителю, а также последующее сопровождение, обслуживание являются составными компонентами жизненного цикла изделия. Традиционно поддержка жизненного цикла включает в себя

первоначальное обучение и последующие консультации по вопросам функционирования продукта.

Тенденции систем поддержки жизненного цикла изделий движутся в направлении исключения этапа обучения и заменой его информационной системой. Это значит, что работа потребителя с продуктом может начинаться непосредственно после его приобретения, путем согласования понятийно-терминологического пространства потребителя с понятийно-терминологическим пространством производителя (рис. 16).

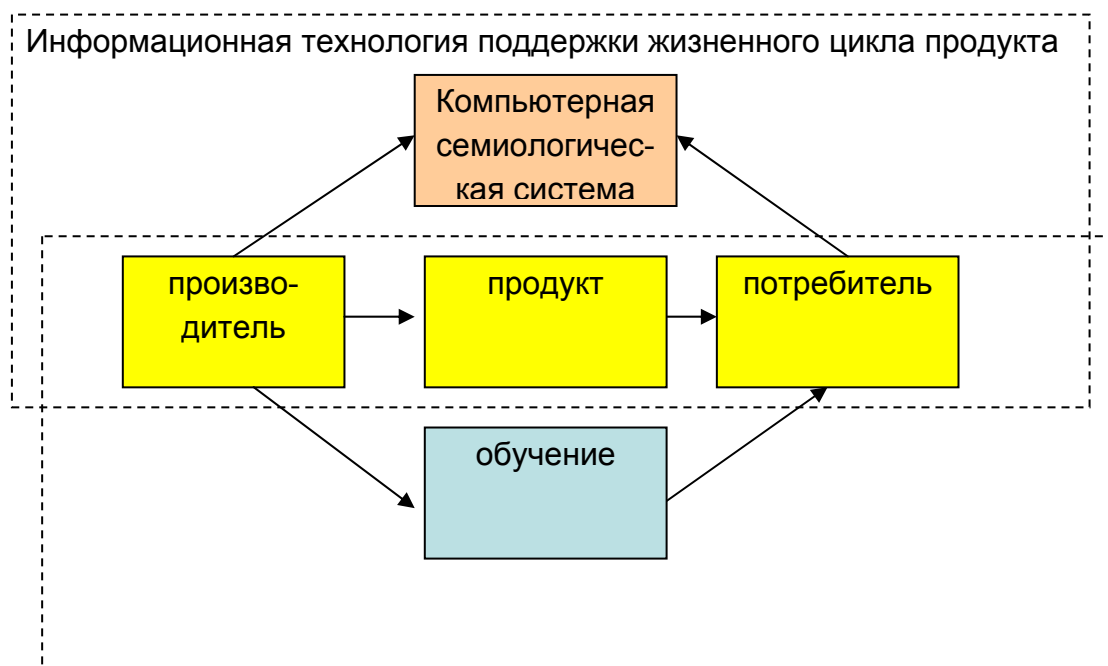


Рис. 16. Схема поддержки жизненного цикла продукта

Существующие информационные системы в основном были направлены на изменение понятийно-терминологического пространства потребителя для работы в новых условиях. Новое поколение информационных систем допускает также и обратный процесс — использование понятийно-терминологического пространства потребителя для осуществления его коммуникативной функции с продуктом, т. е. адаптацию продукта под конкретного потребителя. При этом, как правило, информационная система является компонентом продукта.

Другим видом таких систем является их неявное присутствие в компонентах продукта, имеющего интерактивный интерфейс (это в большей мере касается высокотехнологичных и программных продуктов).

Технологическая база изделия, понятийная область программного продукта, и понятийное поле потребителя при этом разделяются, что позволяет создавать универсальные системы (оболочки), позволяющие работать с различным информационным содержанием, и, соответственно, в разных предметных областях.

Существующие информационные системы не дают непосредственного «контакта» между понятийными полями производителя и потребителя через семиологическую модель предметной области. В основном это определяется «наследием» баз знаний и экспертных систем — когда знания о предметной области воплощаются в реализации физического уровня, как правило, зависящего от платформы. Разработчик (производитель) и

потребитель должны быть связаны семиологической моделью предметной области через интерфейсы, соответственно разработчика и пользователя (рис. 17).

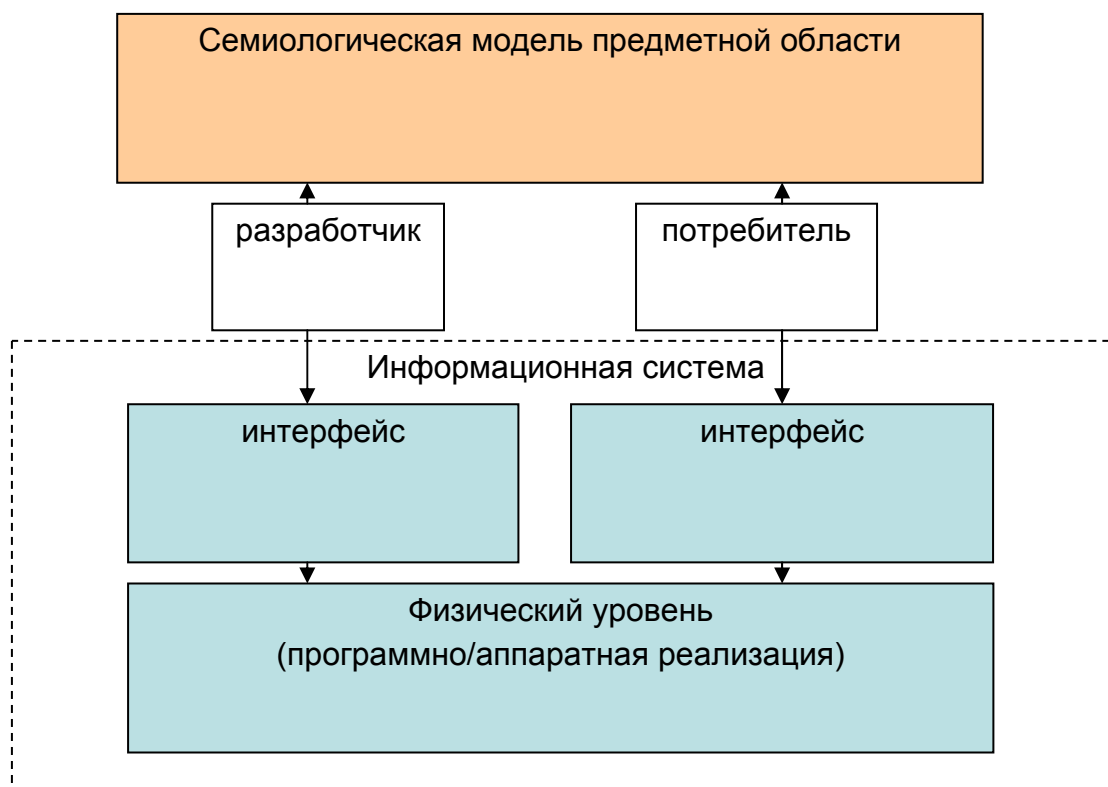


Рис. 17. Структура взаимодействия

Электронный документ

Основой всех информационных систем является понятие документа. Требования конкретного формата данных, представляющего документ проистекают из выбранного программного продукта.

Подчеркнем, что понятие *документ* используется как унифицированный формат (в рамках данного корпоративного стандарта) электронного представления данных – упорядоченный текст, графика, видео, чертеж и др.

Формат электронного документа создается на основе деятельности отделов предприятия и исследования информационных потоков. Здесь могут возникать проблемы совместимости между допустимыми протоколами и форматами программного обеспечения информационной системы и семантикой предметной области.

Электронный технический документ

Электронный технический документ (ЭТД) – оформленная надлежащим образом в установленном порядке и зафиксированная на машинном носителе техническая информация, которая может быть представлена в форме, пригодной для ее восприятия человеком. ЭТД логически состоит из двух частей: содержательной и реквизитной.

Содержательная часть представляет собой информацию непосредственно об изделии и/или способах и средствах поддержки ЖЦ изделия.

Реквизитная часть – аутентификационные и идентификационные данные ЭТД, включающие набор обязательных информационных атрибутов и аутентификационные признаки (одну или несколько электронных цифровых подписей).

ЭТД может существовать как официальный документ только в рамках автоматизированной системы (АС), в которой он обрабатывается, приобретает статус, регистрируется и/или хранится.

ЭТД может иметь две формы представления: внутреннюю и внешнюю. Внутренней формой является фиксация ЭТД на машинном носителе или в памяти ЭВМ. Внешней формой является представление документа в доступном для визуального обозрения и пригодном для восприятия человеком виде, например в произведение на экране дисплея. Ее варианты могут иметь различия с точки зрения визуального восприятия, обусловленные техническими характеристиками используемых для их получения устройств вывода, а также быть отображен в унифицированной, так и в другой форме, в зависимости от требований к содержанию документа в соответствии с решаемой задачей.

Необходимо заметить, что стандартная процедура электронного документооборота, внедряемая на большинстве предприятий должна дополняться системами, обеспечивающими поддержку понятийной базы этого документооборота, что обеспечит возможность работы с системой человека, имеющего терминологическую базу отличную от стандарта принятого для вводимого документооборота. Это необходимо для комфортной работы также как уже ставшая стандартной интеграция информационных систем с офисными приложениями для создания привычного окружения в смысле пользовательского интерфейса.

Типовые функции обработки документов

- Ввод бумажных документов
 - В формат образа документа
 - В формат электронного документа (Структурированный или Неструктурированный)
- Создание (разработка) документов
- Регистрация документа
 - В момент создания
 - При поступлении в систему извне
- Атрибутирование документов
- Хранение документов
- Редактирование документов
- Поиск документов
- Координация коллективной работы

Виды субъектов при обработке документов

- Собственник информационных ресурсов – субъект, в полном объеме реализующий полномочия владения, пользования, распоряжения указанными объектами

- Владелец информационных ресурсов – субъект, осуществляющий владение и пользование указанными объектами и реализующий полномочия распоряжения в пределах, установленных законами
- Пользователь (потребитель) информации – субъект, обращающийся к информационной системе или посреднику за получением необходимой ему информации и пользующийся ею

Программные продукты и решаемые задачи

Развитие информационных систем, прошедшее через базы данных (получение данных через отчеты и запросы), OLAP, Data Mining, можно рассмотреть по «вопросам», на которые они способны дать ответ.

Информационная система	Назначение	Вопрос, на который система может давать ответ
Базы данных (отчеты и запросы)	Получение данных (оперативный анализ информации)	Что имеется?
OLAP	Проверка взаимосвязи данных относительно технологических ограничений	Что будет если ..., то...?
Data Mining	Выявление знаний о развитии процесса	Что необходимо узнать?

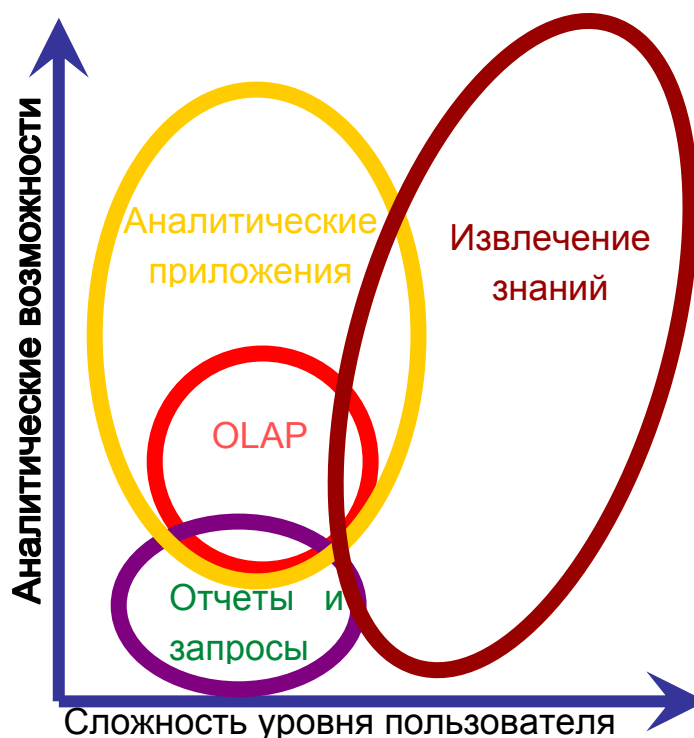


Рис. 18 Соотношение методов построения информационных систем.

Средства обеспечения автоматизированных информационных систем и их технологий – программные, технические, лингвистические, правовые, организационные средства (программы для электронных вычислительных машин; средства вычислительной техники и связи; словари; тезаурусы и классификаторы; инструкции и методики; положения, уставы; должностные инструкции; схемы и их описания; другая эксплуатационная и сопроводительная документация), используемые при проектировании информационных систем и обеспечивающие их эксплуатацию.



Рис. 19 Жизненный цикл информации.

Проведем классификацию программных продуктов, которые входят в состав информационных систем:

- По структуре хранимой информации
 - Фактографические (структурированные)
 - Документографические (полнотекстовые)
- По характеру обработки данных
 - Динамические
 - Статические
 - Аналитические
- По видам доступа к информации
 - Локальные
 - Территориально-распределенные
 - Общего доступа
- По классу решаемых задач
 - Обработка данных
 - Поисковые
 - Управляющие
 - Аналитические

По специализации программные продукты делятся на:

- Специализированные системы
 - Решает один класс задач
 - Требуется применение нескольких продуктов для решения конкретной задачи
 - Меньшее число установок
 - Хорошо масштабируются
- Комплексные системы
 - Решает несколько задач
 - Более легкие в установке и настройке
 - Больше число установок
 - Меньшая гибкость продукта
 - Обычно более узкие пределы масштабирования

Научно-техническая реализация компьютерной техники последних поколений идет все более быстрыми темпами, совершенствуются и развиваются соответствующие программные продукты. Рассмотрим развитие наиболее известных (наиболее используемых) информационных систем.

Системы управления данными

- Являются фундаментом функционирования любой информационной системы.
- Примеры:
 - ORACLE
 - MS SQL
 - MySQL
 - Verity
 - Excalibur
 - другие

Системы управления информацией и знаниями

- Автоматизация управления различных направлений производственной деятельности предприятия.
- Системы документооборота, финансовые системы, ERP, системы управления содержанием и знаниями, системы управления проектами, системы анализа и принятия решений:
 - SAP, 1C, Галактика, PeopleSoft, Documentum, другие.
- Основные функции таких систем:
 - выполнение библиотечных функций, связанных с областью применения;
 - формирование и поддержка предметной информационной среды;
 - передача информации и ее предметная интерпретация;
 - автоматизация производственных процессов;
 - взаимодействие с другими программными системами и пользовательскими приложениями на уровне передачи информации.

Системы представления содержания

- Web ориентированные серверные компоненты, которые позволяют динамически представлять информацию для конечных пользователей в предметно-ориентированном формате.
- Сервера приложений, Web сервера.
 - IIS, ATG Dynamo, IBM WebSphere, BEA, TomCat, Allaire Jrun
- Представление информации в составе и виде, необходимом пользователю.

Системы создания и обработки содержания (контента)

- Пользовательские приложения, ориентированные на создание и обработку предметного содержания различных форматов, которое будет являться частью производственных процессов предприятия.
- Примеры:
 - MS Office, AutoCAD, Компас, Visio, и другие.

Базы данных.

СУБД являются одним из основных и самых важных компонентов информационных систем.

Базы данных – это совокупность сведений (о реальных объектах, процессах, событиях или явлениях), относящихся к определенной теме или задаче, организованная таким образом, чтобы обеспечить удобное представление этой совокупности, как в целом, так и любой ее части. Реляционная база данных представляет собой множество взаимосвязанных таблиц, каждая из которых содержит информацию об объектах определенного типа. Каждая строка таблицы включает данные об одном объекте (например, клиенте, автомобиле, документе), а столбцы таблицы содержат различные характеристики этих объектов – атрибуты (например, наименования и адреса клиентов, марки и цены автомобилей). Строки таблицы называются записями; все записи имеют одинаковую структуру – они состоят из полей, в которых хранятся атрибуты объекта. Каждое поле записи содержит одну характеристику объекта и имеет строго определенный тип данных (например, текстовая строка, число, дата). Все записи имеют одни и те же поля, только в них содержатся разные значения атрибутов.

Для работы с данными используются системы управления базами данных (СУБД). Основные функции СУБД – это определение данных (описание структуры баз данных), обработка данных и управление данными.

Прежде чем заносить данные в таблицы, нужно определить структуру этих таблиц. Под этим понимается не только описание наименований и типов полей, но и ряд других характеристик (например, формат, критерии проверки вводимых данных). Кроме описания структуры таблиц, обычно задаются связи между таблицами. Связи в реляционных базах данных определяются по совпадению значений полей в разных таблицах. Например, клиенты и заказы связаны отношением "один-ко-многим", т. к. одной записи в таблице, содержащей сведения о клиентах, может соответствовать несколько записей в таблице заказов этих клиентов. Если же рассмотреть отношение между преподавателями и курсами лекций, которые они читают, это будет отношение

"многие-ко-многим", т. к. один преподаватель может читать несколько курсов, но и один курс может читаться несколькими преподавателями. И последний тип связей между таблицами – это отношение "один-к-одному". Такой тип отношений встречается гораздо реже. Как правило, это бывает в двух случаях: запись имеет большое количество полей, и тогда данные об одном типе объектов разносятся по двум связанным таблицам, или нужно определить дополнительные атрибуты для некоторого количества записей в таблице, тогда создается отдельная таблица для этих дополнительных атрибутов, которая связывается отношением "один-к-одному" с основной таблицей.

Любая СУБД позволяет выполнять четыре простейшие операции с данными:

- добавлять в таблицу одну или несколько записей;
- удалять из таблицы одну или несколько записей;
- обновлять значения некоторых полей в одной или нескольких записях;
- находить одну или несколько записей, удовлетворяющих заданному условию.

Для выполнения этих операций используется механизм запросов. Результатом выполнения запросов является либо отобранное по определенным критериям множество записей, либо изменения в таблицах. Запросы к базе формируются на специально созданном для этого языке, который так и называется язык структурированных запросов (SQL – Structured Query Language).

И последняя функция СУБД – это управление данными. Под управлением данными обычно понимают защиту данных от несанкционированного доступа, поддержку многопользовательского режима работы с данными и обеспечение целостности и согласованности данных.

Проблема, возникающая при использовании баз данных, следует из метода их применения – необходимо формализовать предметную область в терминах реляционной алгебры, например с помощью ER-диаграмм. Но необходимость этой формализации уже требует от человека определенного набора знаний, никак не связанных с предметной областью, для которой требуется создание информационной системы, содержащей СУБД.

Надстройкой над СУБД является система управления документами (СУД). Система управления документами предоставляет следующие возможности:

- Создание (разработка) документов
 - Интеграция с приложениями
- Атрибутирование документов
- Хранение документов
 - Многоуровневое хранение документов
 - Определение жизненного цикла документов
- Редактирование документов
 - Интеграция с приложениями
- Поиск документов
 - Атрибутивный поиск документов
 - Полнотекстовой поиск документов

Информационно-аналитические системы

В последнее время особую актуальность обрели различные компьютерные системы автоматизации и анализа информации. Подобные системы внедряются на производственных предприятиях, в организациях по продаже товаров, в сфере потребительских услуг и производстве нематериальной продукции. Информационно-аналитическая система (ИАС) – это компьютерная система, позволяющая получать информацию, создавать ее и производить ее обработку и анализ. Система должна предоставлять возможность сбора и обработки оперативной информации в режиме реального времени с удобным интерфейсом. Статистическая и аналитическая информация должна предоставляться в соответствии с любыми возникающими запросами с возможностью дальнейшей детализации. Построение и внедрение ИАС можно условно разбить на два этапа.

На первом этапе создается подсистема сбора и редактирования информации в режиме реального времени. Данная подсистема должна предоставлять пользователю удобный интуитивно понятный интерфейс. Данные в ней должны отображаться таким образом, чтобы наглядно отображалось состояние различных этапов бизнес-процесса.

На втором этапе строится подсистема статистической и аналитической обработки данных. Пользователь должен иметь возможность получить необходимый отчет, использовать его результаты в качестве входных параметров для других отчетов, произвести анализ имеющихся данных за любой доступный промежуток времени, построить на основании этого анализа прогноз и после сверить результаты прогноза с реальными данными. Оптимальным решением будет, если система будет построена по технологии "клиент-сервер". Среди распространенных в настоящее время (и представляющихся перспективными в будущем) архитектур в качестве основы для подсистемы оперативной обработки информации разумнее всего использовать какую-либо промышленную базу данных, поддерживающую язык SQL (например, Microsoft SQL Server, Oracle, Sybase). Для подсистемы анализа данных лучше всего подойдет технология OLAP-систем (многомерная модель данных, например, Oracle OLAP, Microsoft OLAP Services, IBM DB2 OLAP Server). Наиболее известные на рынке информационно-аналитические системы можно условно разделить на два больших класса: финансово-управленческие и производственные системы.

Финансово-управленческие системы предназначены для ведения учета по одному или нескольким направлениям: бухгалтерия, сбыт, склады, учет кадров и т.д. Практически любое предприятие, которому необходимо управление финансовыми потоками и автоматизация учетных функций, может воспользоваться системами этой группы (например, 1С, Парус). Системы этого класса универсальны по многим критериям, хотя разработчики часто предлагают и решения отраслевых проблем, например, особые способы начисления налогов или управление персоналом с учетом местной специфики региона. Благодаря универсальности таких систем, процесс их внедрения занимает небольшой промежуток времени. Поскольку такие системы, фактически, являются готовым решением, их можно использовать, просто купив "коробочный" вариант и самостоятельно установив его на компьютере. Системы (особенно российские разработки) зачастую построены по модульной схеме и имеют средства программирования, что

позволяет подстроить исходную систему под конкретную задачу. Архитектура таких систем, как правило, представляет собой выделенный файл-сервер с подключенными к нему клиентами в сети Novell Netware или Microsoft Windows NT. В качестве языка программирования, обычно, используются Clipper, FoxPro, dBase или C++. Системы, построенные по такой технологии, создают высокую загрузку локальной сети и при накоплении большого количества информации становятся нестабильными в работе.

Производственные системы предназначены, в первую очередь, для управления и планирования производственного процесса. Функции учета выполняют в них вспомогательную роль, и порой невозможно выделить модуль бухгалтерского учета, так как информация в бухгалтерию поступает автоматически из других модулей (например, PRMS, SAP/R3). Для подобных систем характерна архитектура "клиент-сервер", что позволяет эффективно распределять нагрузку между сервером баз данных и клиентом и уменьшать загрузженность локальной сети. Однако и здесь существуют свои минусы. Стоимость производственных систем очень высока (покупка и внедрение системы может достигать до \$1 млн. долларов), период их внедрения может занимать от полугода до нескольких лет, кроме того, их внедрение сопряжено с большими трудностями, возникающими при адаптации к системе сотрудников предприятия. Для настройки и поддержания такой системы необходимо привлекать высококвалифицированного специалиста компании-производителя, что также очень дорого.

САПР

CAD-системы (computer-aided design – компьютерная поддержка проектирования) предназначены для решения конструкторских задач и оформления конструкторской документации (более привычно они именуются системами автоматизированного проектирования – САПР). Как правило, в современные CAD-системы входят модули моделирования трехмерной объемной конструкции (детали) и оформления чертежей и текстовой конструкторской документации (спецификаций, ведомостей и т. д.). Ведущие трехмерные CAD-системы позволяют реализовать идею сквозного цикла подготовки и производства сложных промышленных изделий.

В свою очередь, САМ-системы (computer-aided manufacturing – компьютерная поддержка изготовления) предназначены для проектирования обработки изделий на станках с числовым программным управлением (ЧПУ) и выдачи программ для этих станков (фрезерных, сверлильных, эрозионных, пробивных, токарных, шлифовальных и др.). САМ-системы еще называют системами технологической подготовки производства. В настоящее время они являются практически единственным способом для изготовления сложнопрофильных деталей и сокращения цикла их производства. В САМ-системах используется трехмерная модель детали, созданная в CAD-системе.

САЕ-системы – (computer-aided engineering – поддержка инженерных расчетов) представляют собой обширный класс систем, каждая из которых позволяет решать определенную расчетную задачу (группу задач), начиная от расчетов на прочность, анализа и моделирования тепловых процессов до расчетов гидравлических систем и машин, расчетов процессов литья. В САЕ-системах также используется трехмерная модель изделия, созданная в CAD-системе. САЕ-системы еще называют системами инженерного анализа.

Системы автоматического проектирования САПР (CAD), представляя собой подсистему обработки содержания, являются компонентами информационных систем, позволяющими работать с конструкторской документацией.

Рассмотрим наиболее популярные системы автоматического проектирования:

AutoCAD – самая популярная в мире система автоматизированного проектирования и выпуска рабочей конструкторской и проектной документации. С его помощью создаются двумерные и трехмерные проекты различной степени сложности в области архитектуры и строительства, машиностроения, генплана, геодезии и т.д. Формат хранения данных AutoCAD де-факто признан международным стандартом хранения и передачи проектной документации.

AutoCAD является платформой, на которой построено множество специализированных программ, имеющих общий формат хранения данных. Высокопрофессиональные приложения от авторизованных разработчиков Autodesk дополняются утилитами и программами, которые (с помощью встроенных языков программирования) создают сами пользователи. На сегодня AutoCAD охватывает весь спектр инженерных задач: создание трехмерных моделей, разработка и оформление чертежей, выполнение различного рода расчетов, инженерный анализ, формирование фотореалистичных изображений готовой продукции. Также существуют пакеты AutoCAD для решения более узких конкретных задач:

- двумерное проектирование, оформления конструкторской и проектной документации (AutoCAD LT)
- машиностроительное проектирование, средства проектирования деталей машин, деталей типа «тело вращения», библиотеку стандартных деталей по восемнадцати международным стандартам (AutoCAD Mechanical)
- строительный пакет (Autodesk Architectural Desktop)
- эскизное и концептуальное архитектурное проектирование, подготовка презентационных материалов (Autodesk Architectural Studio)
- проектирования внутренних инженерных коммуникаций (Autodesk Building Systems)
- проектирование наземных сооружений (Autodesk Civil Design)
- визуализация, пространственный анализ, доступ к различным картографическим и проектным данным удаленных источников в сетях LAN и Intranet/Интернет (Autodesk Envision)
- трехмерное твердотельное проектирование (Autodesk Inventor)
- выполнение проектов крупных машиностроительных сборок (Autodesk Inventor Series)
- земельное проектирование (Autodesk Land Desktop)
- инструментальная геоинформационная система (Autodesk Map)
- трехмерное проектирование (Autodesk Mechanical Desktop)
- визуализация выполнения пространственных запросов, поиска данных, ГИС-анализа, а также для представления данных в сетях (Autodesk OnSite Desktop)
- редактирование растровых данных (Autodesk Raster Design)

- трехмерное моделирование, визуализация материалов в области дизайна интерьеров и экстерьеров, архитектуры и строительства (Autodesk VIZ)

Combustion – программный пакет, предоставляющий широкие возможности рисования, анимации, визуализации и наложения спецэффектов в единой среде. Интуитивный интерфейс, передовая технология управления параметрами объектов и трехмерная рабочая среда позволяют за минимальное время создавать клипы высочайшего качества. Поддержка платформ Windows NT и Macintosh обеспечивает беспрепятственный обмен данными с коллегами и заказчиками.

P-CAD – система проектирования радиоэлектронной аппаратуры (РЭА) P-CAD является интегрированным набором специализированных программных пакетов и имеет иерархическую модульную структуру. P-CAD ориентирован на эффективное проектирование принципиальных схем (ПС), программируемых логических матриц (ПЛМ) и топологии печатных плат (ПП). Программные средства P-CAD'a позволяют автоматизировать процесс проектирования ПП: начать с конструирования ПС, выполнить автоматическое или интерактивное размещение радиоэлектронных компонентов (РЭК) на ПП и автотрассировку соединений, провести проверку соответствия правилам проектирования, получить конструкторскую документацию и подготовить информацию для производства плат на технологическом оборудовании. То есть набор взаимосвязанных пакетов обеспечивает сквозное проектирование РЭА. При этом используются уже существующие в P-CAD'e библиотеки РЭК или пользователь может их расширить, дополнив новыми элементами.

Система P-CAD позволяет выполнять следующие проектные операции:

- создание условных графических обозначений элементов принципиальной электрической схемы (УГО) и их физических образов (конструктивов);
- графический ввод чертежа принципиальной электрической схемы и конструктивов проектируемого устройства;
- математическое моделирование цифровых электронных устройств, в том числе программируемых логических матриц;
- одно- и двустороннее размещение разногабаритных элементов с планарными и многослойными контактными площадками на поле ПП с печатными и навесными (вырубными) шинами питания в интерактивном и автоматическом режимах;
- ручную и автоматическую трассировку печатных проводников произвольной ширины в интерактивном режиме (число слоев 1...32);
- размещение межслойных переходов;
- автоматизированный контроль результатов проектирования ПП на соответствие принципиальной электрической схеме и конструкторско-технологическим ограничениям;
- автоматическую коррекцию электрической принципиальной схемы по результатам размещения элементов на ПП (после эквивалентной перестановки компонентов или их выводов);
- полуавтоматическую корректировку разработанной ПП по изменениям, внесенным в принципиальную электрическую схему;

- выпуск конструкторской документации (чертеж принципиальной схемы, детализовочный и сборочный чертежи) и технологической информации (фотошаблоны и файлы данных для сверления отверстий с помощью станков с ЧПУ) на проектируемую ПП.

3ds max – наиболее популярное программное обеспечение для 3D-моделирования, анимации и визуализации. Включает высокопроизводительные инструменты, необходимые для создания зрелищных кинофильмов и телевизионных заставок, современных компьютерных игр и презентационных материалов. 3ds max располагает целым набором новых функциональных возможностей. Аниматорам, работающим в кинопроизводстве, предлагается новый уровень реализма визуализации и динамики объектов в сочетании с упрощением рабочего процесса.

Character Studio – приложение к 3ds max, предназначенное для решения задач персонажной анимации. Незаменимый инструмент разработчиков компьютерных игр, а также художников-аниматоров, работающих в области телерекламы. Технология работы с пакетом предполагает монтаж, использование систем захвата движения и традиционные средства построения анимации по ключевым кадрам – в сочетании с новейшим инструментарием для деформации "кожи" и анимации "по шагам".

Экспертные системы

Экспертные системы (ЭС) – это системы искусственного интеллекта, использующие знания из сравнительно узкой предметной области для решения возникающих в ней задач, причем так, как это делал бы эксперт-человек, то есть в процессе диалога с заинтересованным лицом, поставляющим необходимые сведения по конкретному вопросу. Экспертные системы, в отличие от других ориентированных на практику программных комплексов искусственного интеллекта, имеют массовое применение, так как каждая экспертная система позволяет решить неограниченное число сложных для человека задач из соответствующей предметной области. Экспертная система может быть построена для любой предметной области, о чем говорит следующий перечень уже решаемых экспертными системами классов задач: интерпретация, диагностика, распознавание, планирование, проектирование, управление, наблюдение, исправление.

Двумя самыми известными экспертными программами в искусственном интеллекте являются *DENDRAL* и *MACSYMA*. Система *DENDRAL* получает масс-спектрограмму некоторого вещества и определяет его молекулярную структуру. Система *MACSYMA* помогает ученым при решении математических задач. Обе системы представляют собой крупные программы, создававшиеся на протяжении нескольких лет.

Наиболее широко встречающиеся области деятельности, где используются экспертные системы: медицина, вычислительная техника, военное дело, микроэлектроника, радиоэлектроника, юриспруденция, экономика, экология, геология (поиск полезных ископаемых), математика.

Примеры широко известных и эффективно используемых (или использованных в свое время) экспертных систем:

DENDRAL – ЭС для распознавания структуры сложных органических молекул по результатам их спектрального анализа (считается первой в мире ЭС);

MOLGEN – ЭС для выработки гипотез о структуре ДНК на основе экспериментов с ферментами;

MYCIN – ЭС диагностики кишечных заболеваний;

PUFF – ЭС диагностики легочных заболеваний;

MACSYMA – ЭС для символьных преобразований алгебраических выражений;

PROSPECTOR – ЭС для консультаций при поиске залежей полезных ископаемых;

ЭСПЛАН – ЭС для планирования производства на Бакинском нефтеперерабатывающем заводе;

YES/MVS – ЭС для управления многозадачной операционной системой MVS больших ЭВМ корпорации IBM;

DART – ЭС для диагностики больших НМД корпорации IBM;

XCON – ЭС для конфигурирования (проектирования) вычислительных комплексов VAX-11 в корпорации DEC в соответствии с заказом покупателя;

МОДИС – ЭС диагностики различных форм гипертонии.

Логистические системы

Схему любой ЛС можно представить в виде блок-схемы с обратной связью (рис. 20).

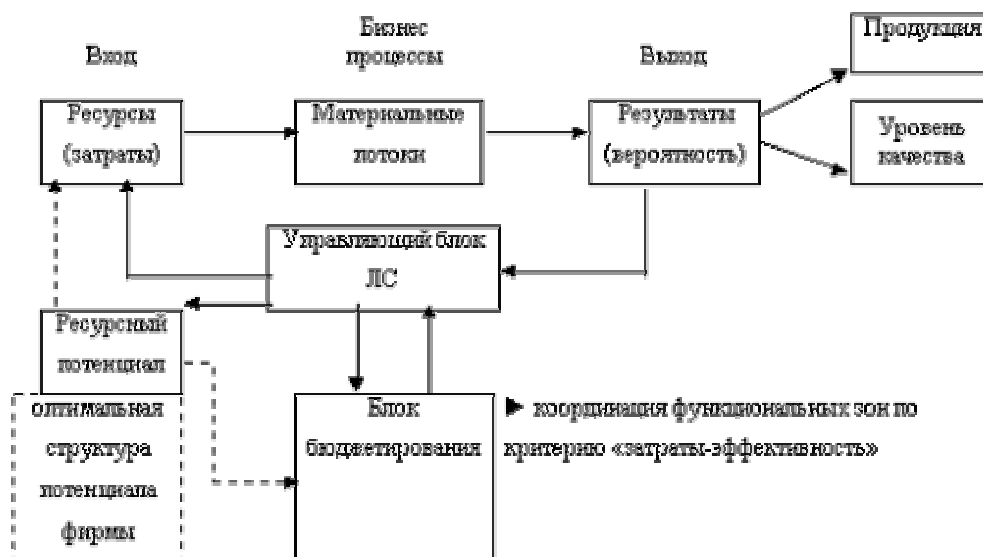


Рис. 20. Структура логистической системы.

Логистические системы делят на макро- и микрологистические.

Макрологистическая система – это крупная система управления материальными потоками, охватывающая предприятия и организации промышленности, посреднические, торговые и транспортные организации различных ведомств, расположенных в разных регионах страны или в разных странах. Макрологистическая система представляет собой определенную инфраструктуру экономики региона, страны или группы стран.

При формировании макрологистической системы, охватывающей разные страны, необходимо преодолеть трудности, связанные с правовыми и экономическими особенностями международных экономических отношений, с неодинаковыми условиями поставки товаров, различиями в транспортном законодательстве стран, а также ряд других барьеров.

Формирование макрологистических систем в межгосударственных программах требует создания единого экономического пространства, единого рынка без внутренних границ, таможенных препятствий транспортировке товаров, капиталов, информации, трудовых ресурсов.

Микрологистические системы являются подсистемами, структурными составляющими макрологистических систем. К ним относят различные производственные и торговые предприятия, территориально-производственные комплексы. Микрологистические системы представляют собой класс внутрипроизводственных логистических систем, в состав которых входят технологически связанные производства, объединенные единой инфраструктурой.

В рамках макрологистики связи между отдельными микрологистическими системами устанавливаются на базе товарноденежных отношений. Внутри микрологистической системы также функционируют подсистемы. Однако основа их взаимодействия бестоварная. Это отдельные подразделения внутри фирмы, объединения, либо другой хозяйственной системы, работающие на единый экономический результат.

На уровне макрологистики выделяют три вида логистических систем.

- Логистические системы с прямыми связями. В этих логистических системах материальный поток проходит непосредственно от производителя продукции к ее потребителю, минуя посредников.
- Эшелонированные логистические системы. В таких системах на пути материального потока есть хотя бы один посредник.
- Гибкие логистические системы. Здесь движение материального потока от производителя продукции к ее потребителю может осуществляться как напрямую, так и через посредников.

Объектом логистическими системами, как известно, является сквозной материальный поток, тем не менее, на отдельных участках управление им имеет известную специфику. В соответствии с этой спецификой выполняют пять функциональных областей логистики, которые в свою очередь осуществляют управление различными логистическими системами. В управление системами входят следующие структуры: закупочная, производственная, распределительная, транспортная и информационная. Укажем специфику каждой структуры и ее место в общей системе логистики.

1. В процессе обеспечения предприятия сырьем и материалами решаются задачи закупочной логистики. На этом этапе изучаются и выбираются поставщики, заключаются договоры и контролируется их исполнение, принимаются меры в случае нарушения условий поставки. Любое производственное предприятие имеет службу, которая осуществляет перечисленные функции. Логистический подход к управлению материальными потоками требует, чтобы деятельность этой службы, связанная с формированием параметров сквозного материального потока, не была обособленной, а подчинялась стратегии управления сквозным материальным потоком. В то же время задачи, решаемые в процессе доведения материального потока от складов готовой продукции поставщика до цехов предприятия потребителя, имеют известную специфику. На практике границы деятельности, составляющей основное содержание закупочной

логистики, определяются условиями договора с поставщиками и составом функций службы снабжения внутри предприятия.

2. В процессе управления материальным потоком внутри предприятия, создающего материальные блага или оказывающего материальные услуги, в основном решаются задачи производственной логистики. Специфика этой структуры управления заключается в том, что основной объем работ по проведению потока выполняется в пределах территории одного предприятия. Участники логистического процесса при этом, как правило, не вступают в товароденежные отношения. Поток идет не в результате заключенных договоров, а в результате решений, принимаемых системой управления предприятием.

Сфера производственной логистики тесно соприкасается со сферами закупок материалов и распределения готовой продукции. Однако основной круг задач в этой области – управление материальными потоками в процессе осуществления именно производства.

3. При управлении материальными потоками в процессе реализации готовой продукции решаются задачи распределительной логистики. Это обширный круг задач, решением которых занимаются как производственные предприятия, так и предприятия, осуществляющие торгово-посредническую деятельность. К решению этих задач имеют отношение властные структуры, так как от организации распределения существенно зависит состояние экономики региона. Например, в случае неудовлетворительной организации системы распределения продовольственных товаров в регионе положение местной власти будет нестабильным.

Реализация функции распределения на производственном предприятии иначе называется сбытом продукции. В сферу внимания этой структуры управления материальный поток попадает, еще находясь в производственных цехах. Это означает, что вопросы тары и упаковки, размера изготавливаемой партии и времени, к которому эта партия должна быть изготовлена, а также много других вопросов, существенных для процесса реализации, начинают решаться на более ранних стадиях управления материальным потоком.

4. При управлении материальными потоками на транспортных участках решаются специфические задачи транспортной логистики. Совокупный объем транспортной работы, выполняемой в процессе доведения материального потока от первичного источника сырья до конечного потребителя, можно разделить на две большие группы (примерно равные):

- работа, выполняемая транспортом, принадлежащим специальным транспортным организациям (транспорт общего пользования);
- работа, выполняемая собственным транспортом всех остальных (нетранспортных) предприятий.

Также как и другие функциональные области логистики, транспортная логистика четко очерченных границ не имеет. Методы транспортной логистики применяются при организации любых перевозок. Однако приоритетным объектом изучения и управления в этом разделе является материальный поток, имеющий место в процессе перевозок транспортом общего пользования.

5. Информационная логистика. Результаты движения материальных потоков находятся в прямой связи с рациональностью организации движения информационных

потоков. В последние десятилетия именно возможность эффективного управления мощными информационными потоками позволила ставить и решать задачу сквозного управления потоками материальными. Высокая значимость информационной составляющей в логистических процессах стала причиной выделения специального раздела логистики – информационной логистики. Объект исследования здесь – информационные системы, обеспечивающие управление материальными потоками, используемая микропроцессорная техника, информационные технологии и другие вопросы, связанные с организацией информационных потоков (сопряженных с материальными).

Информационная логистика тесно связана с остальными структурами логистических систем. Этот раздел рассматривает организацию информационных потоков внутри предприятия, а также обмен информацией между различными участниками логистических процессов, находящимися на значительных расстояниях друг от друга (например, с помощью средств спутниковой связи).

Поисковые системы в Интернет

Классифицируем виды поиска документов:

- Поиск документов по атрибутам
 - Формальные признаки документов: дата создания, наименование документа, аннотация документа
- Поиск документа по содержанию
 - Поиск по точному запросу
 - Поиск по нечеткому запросу
 - Поиск с использованием тезаурусов

Сеть Интернет, как и всякий объект эволюции, имеет дело с организующим и дезорганизующим началами. Организующее начало проявляется обыденно – появление в океанах информации островов, на берегах которых можно быстро найти нужную вам информацию. Это и грамотно сделанный тематический авторский сайт, и удобный портал, или просто страничка с упорядоченными ссылками на ресурсы по некоторой теме. Второе, противоположное начало, способствует хаосу в Сети, исчезают страницы и сайты, повисают в пустоте ссылки. Устаревший контент заполняет Сеть. Совершенно невозможно разобраться в плохо спроектированных и неудачно исполненных сайтах.

Обратимся к поисковым машинам. Казалось бы, вот борцы против хаоса. Как ориентироваться в океане информации без них? Да, верно – именно поисковые машины позволяют найти нужную информацию, иногда они даже дублируют контент Web. Но, увы, хаос побеждает и их. Противостоять хаосу в одиночку, пытаться "спасти Вселенную" от него не удалось и не удастся никому из традиционных глобальных поисковых машин. Почему?

Вспомним, как работает обычная поисковая машина. С помощью роботов она "обходит" Интернет, в частности, по ссылкам и индексирует текстовый контент страниц, часто с учетом задаваемой автором простой метаданных. Заметьте, одна поисковая машина пытается проиндексировать всю Сеть. Идем дальше. Как осуществляется собственно поиск по запросу пользователя? Совпадение слов запроса и слов страницы. Об анализе смысла запроса и соответствия его ресурсу Сети обычно речи не идет – отсюда

так велика доля мусора на, казалось бы, хорошие запросы. И еще одно: кто отвечает за качество поиска?

Вот и живут в индексах поисковых машин ссылки на несуществующие уже ресурсы, а разбираться со всем этим приходится потребителям информации - пользователям – чаще как посетителям, иногда как клиентам и покупателям. Что получается в результате? Поисковые машины индексируют лишь небольшую часть содержимого Сети, причем делают это с опозданием, а релевантность поиска иногда анекдотична. И получается, что поисковые машины, отчаянно борющиеся с хаосом, сами становятся его проводниками.

В Интернет существует большое количество поисковых систем. По самым скромным оценкам, их более восьми тысяч, считая классические поисковые машины, общие и специализированные каталоги, а также метапоисковые Web-узлы (которые посылают запросы сразу на несколько поисковых серверов). В дополнение к этому существует ряд альтернативных средств поиска, способных вам пригодиться, включая утилиты, которые, работая совместно с браузером, добывают информацию из Web, и так называемые "экспертные узлы", где с вашими запросами работают живые люди.

Поисковые машины и каталоги

При всем изобилии методов поиска в Интернет наиболее распространенными средствами нахождения информации по-прежнему остаются поисковые машины и каталоги. Каждый из этих инструментов имеет определенные преимущества, а основная разница между ними заключается в участии/неучастии человека.

Поисковые машины запускают в Web программных "пауков" (spiders), которые путешествуют со страницы на страницу и на каждой индексируют ее полный текст. Каталоги же формируются людьми-редакторами, которые прочитывают страницы, отсеивают неподходящие и классифицируют узлы по темам.

Но из-за того, что каталоги создаются вручную, они охватывают намного меньше ресурсов, чем поисковые машины. В Web сейчас, по самым скромным оценкам, насчитываются миллиарды страниц (причем их число ежедневно увеличивается). Большинство поисковых машин не подошли сколько-нибудь близко к тому, чтобы проиндексировать всю Сеть. Исключением является Google, который претендует именно на эту цифру – миллиард страниц, частично или полностью охваченных его индексами. Самый большой каталог – Open Directory Project – на этом фоне кажется крошечным: в него занесено лишь около 2 млн. страниц.

Гибридные поисковые серверы

Поскольку у каталогов свои достоинства, а у поисковых машин – свои, большинство крупных поисковых серверов в том или ином виде реализуют оба метода. Поисковую машину AltaVista, например, дополняет каталог, основанный на индексах с серверов LookSmart и Open Directory Project. Когда же вы что-то ищете на LookSmart или Yahoo, сервер сначала выдает результаты из своего каталога, а потом переадресует ваш запрос поисковой машине.

Разумеется, услуги многих серверов далеко не исчерпываются каталогом и поисковой машиной. Excite, Go.com, Lycos, Yahoo и другие превратились в порталы, на

которых есть различная информация, и электронная почта, и много других сервисов. Короче, они делают все возможное, чтобы подольше задержать посетителя на своих страницах: это повышает вероятность того, что он прочтет рекламу, за счет которой и существует сервер. Но хотя многие материалы, предлагаемые порталами, полезны или забавны, в результате сервер настолько загромождается второстепенными функциями, что средства поиска начинают казаться необязательным дополнением.

Некоторые поисковые серверы возвращаются к основам. Первым это движение начал в прошлом году Google: на его элегантной главной странице нет почти ничего, кроме логотипа сервера, поля, куда вводится запрос, и пары кнопок. Несколько позже от AltaVista отпочковался сервер Raging Search, который предоставляет доступ к поисковой машине AltaVista без сутолоки соответствующего узла. Утверждается, что поиск на двух узлах может дать различные результаты, но во всех наших тестах списки ссылок оказались полностью идентичными. Как бы то ни было, существование чисто поисковых узлов должно порадовать завсегдатаев Web, которым мешает отвлекающий внимание сервис порталов.

Каталоги

В 1994 г., когда начинался бурный рост "Всемирной паутины", выбор средств поиска в Сети был весьма ограниченным. Одним из первых стал сервис Yahoo. Этот сервер и по сей день остается краеугольным камнем исследования Web, но как каталог он столкнулся сейчас с жесткой конкуренцией со стороны Open Directory Project и LookSmart.

Все три каталога весьма полезны, но с учетом всех обстоятельств предпочтение стоит отдать Open Directory Project. Проект Open Directory Project, инициированный компанией Netscape, реализуется усилиями редакторов-добровольцев со всего мира, которых насчитывается более 24 тыс. и которые проиндексировали около 2 млн. узлов, расклассифицировав их по более чем 200 тыс. категорий. Любой поисковый сервер может получить лицензию Open Directory Project и использовать его базу данных при обработке запросов, и на многих это сделано: AltaVista, HotBot, Lycos, MetaCrawler и около сотни других серверов ныряют туда за ссылками.

Можно было бы ожидать, что, коль скоро каталог Open Directory Project создается силами добровольцев, качество результатов будет колебаться. Но в результате мы получаем хорошо организованные списки относящихся к теме страниц с четкими описаниями каждой ссылки. А узел Open Directory Project производит такое же впечатление, как Google: это "чистый поиск" без отвлекающих моментов типа ссылок на магазины.

Какой каталог ни выбрать, у всех есть одно преимущество перед поисковыми машинами: их можно систематически просматривать, пользуясь иерархической системой меню. Например, вы щелкаете на LookSmart на категории Sports (спорт) и получаете список тем, таких как Baseball (бейсбол), Olympics (Олимпийские игры) или Motor Sports (авто- и мотоспорт). Щелкнув в этом списке на пункте Motor Sports, вы сможете сузить поиск, выбрав подкатегорию, например, Formula One Racing (гонки "Формула-1") или NASCAR. Таким путем вы найдете нужные сведения быстрее, чем при обычном поиске по ключевым словам.

Метапоисковые серверы

Такие серверы не ведут собственных индексов, а каждый запрос передают другим поисковым машинам и каталогам. Большинство из них объединяют полученные результаты, убирают повторы и представляют вам обработанный список, содержащий больше адресов, чем при поиске на какой-то одной машине.

На заре развития Web, когда ни на одной поисковой машине число проиндексированных страниц не превышало нескольких миллионов, метапоиск был очень важен. Но сейчас AltaVista, Google и другие машины могут похвастаться индексами на сотни миллионов страниц, так что каждая из них даже в одиночку позволит вам "прочесать" значительную часть Web. Тем не менее, с лучшими метапоисковыми серверами стоит познакомиться; попробуйте обратиться на какой-нибудь из них, когда ваша любимая поисковая машина в очередной раз ничем вам не сможет помочь. К сожалению, их результатам часто недостает релевантности Google – вероятно, из-за того, что логично объединить ссылки из разных источников очень сложно.

В настоящее время в Web работает множество метапоисковых серверов. Среди наиболее выдающихся – Dogpile, Mamma.com, MetaCrawler, ProFusion, Search.com и TheBigHub.com. Каждый из них черпает информацию из своего набора поисковых машин. MetaCrawler, например, обращается к 13 машинам, включая AltaVista, Excite, Google и LookSmart, а TheBigHub.com довольствуется всего восемью. Необходимо понимать, что чем больше машин занимается запросом, тем дольше он обрабатывается: метапоисковый сервер должен каждой машине его передать, от каждой получить ответ и свести все результаты в один список. Ограничение числа запрашиваемых машин помогает метапоисковому серверу держать время ожидания ответа в разумных пределах.

Популярность метапоиска вдохновила владельцев ряда традиционных поисковых машин на реализацию сходных функций. К примеру, Ask Jeeves наряду с результатами поиска по его собственной базе данных приводит ссылки с About.com, AltaVista, Excite и WebCrawler.

Специализированные поисковые машины

Тематические поисковые системы, от TaxTopic.com до James T. Kirk Search Engine охватывают практически все мыслимые предметы. Есть большое количество специализированных поисковых серверов, посвященных астрономии, политике, рыбной ловле и т. д. Лучшие из них демонстрируют такую компетентность в рамках своей темы, какую редко можно встретить у поисковых машин общего назначения, и почти во всех наших тестах они помогли нам найти то, что требовалось. Вдобавок благодаря узкой специализации подобных систем результаты отличаются высокой релевантностью: по запросу bacon (бекон) на сервере Recipes-For-All.com вы наверняка получите рецепты блюд с беконом, а не фильмы с Кевином Беконем или эссе сэра Френсиса Бекона. Как узнать, существует ли поисковая система, например, по коккер-спаниелям? Очень просто: справьтесь об этом в путеводителе по тематическим "искалкам". Например – InvisibleWeb.com и Search Engine Guide. Специализированные поисковые системы редко являются полностью автоматическими (примером таковой может служить Ditto.com, помогающая находить в Web изображения): большинство фактически представляют собой

вручную составленные каталоги узлов, относящихся к данной теме. Поэтому тематический поиск на них дает лучшие результаты, чем узкоспециализированные запросы. Например, задав на сервере All Magic Guide запрос Houdini (Гудини), вы получите 15 полезных ссылок на страницы, посвященные легендарному иллюзионисту, а по запросу harry houdini straitjacket escape photo (фотография Гарри Гудини, выпутывающегося из смиренной рубашки) не найдете ничего.

Многие специализированные серверы представляют собой плод трудов одного энтузиаста, и в силу этого там может наступить застой, если владелец потеряет интерес к теме или у него станет меньше свободного времени. Посмотрите на дату последнего обновления на основной странице, и если оно было давно, двигайтесь дальше.

Экспертные узлы

Разработчики поисковых машин долгое время (и без особого успеха) пытаются научить свои автоматизированные системы мыслить по-человечески. А почему бы не сделать наоборот? Такова идея экспертных узлов – информационных центров, поддерживаемых знаниями и суждениями живых людей.

Порядок действий почти на всех экспертных узлах одинаковый: по систематическому каталогу в стиле Yahoo вы разыскиваете нужную категорию, затем вводите свой вопрос и ждете ответов от посетителей узла ("экспертов"), утверждающих, что они разбираются в данном предмете. Чаще всего результатом бывает весьма квалифицированный и индивидуальный совет, какой не способна дать ни одна поисковая машина. Когда мы, к примеру, спросили о покупке велосипеда для поездок на работу, эксперт на AskMe поделился списком своих любимых страниц, посвященных велосипеду в городе, а велосипедист-ветеран на Abuzz выдал памятку, которую следовало взять с собой в велосипедный магазин. Ответ на таких узлах дается не сразу, но очень часто вы получите его в течение часа или двух.

Описание работы механизмов поиска

Что такое поисковая система в Интернет?

Поисковые системы обычно состоят из трех компонент:

- агент (паук или краулер), который перемещается по Сети и собирает информацию;
- база данных, которая содержит всю информацию, собираемую пауками;
- поисковый механизм, который люди используют как интерфейс для взаимодействия с базой данных.

Средства поиска и структурирования, иногда называемые поисковыми механизмами, используются для того, чтобы помочь людям найти информацию, в которой они нуждаются. Средства поиска типа агентов, пауков, краулеров и роботов используются для сбора информации о документах, находящихся в Сети Интернет. Это специальные программы, которые занимаются поиском страниц в Сети, извлекают гипертекстовые ссылки на этих страницах и автоматически индексируют информацию, которую они находят для построения базы данных. Каждый поисковый механизм имеет собственный набор правил, определяющих, как собирать документы. Некоторые следуют за каждой ссылкой на каждой найденной странице и затем, в свою очередь, исследуют каждую

ссылку на каждой из новых страниц, и так далее. Некоторые игнорируют ссылки, которые ведут к графическим и звуковым файлам, файлам мультимедиа; другие игнорируют ссылки к ресурсам типа баз данных WAIS; другие проинструктированы, что нужно просматривать прежде всего наиболее популярные страницы.

- Агенты – самые "интеллектуальные" из поисковых средств. Они могут делать больше, чем просто искать: они могут выполнять даже транзакции от Вашего имени. Уже сейчас они могут искать сайты специфической тематики и возвращать списки сайтов, отсортированных по их посещаемости. Агенты могут обрабатывать содержание документов, находить и индексировать другие виды ресурсов, не только страницы. Они могут также быть запрограммированы для извлечения информации из уже существующих баз данных. Независимо от информации, которую агенты индексируют, они передают ее обратно базе данных поискового механизма.
- Общий поиск информации в Сети осуществляют программы, известные как пауки. Пауки сообщают о содержании найденного документа, индексируют его и извлекают итоговую информацию. Также они просматривают заголовки, некоторые ссылки и посылают проиндексированную информацию базе данных поискового механизма.
- Краулеры просматривают заголовки и возвращают только первую ссылку.
- Роботы могут быть запрограммированы так, чтобы переходить по различным ссылкам различной глубины вложенности, выполнять индексацию и даже проверять ссылки в документе. Из-за их природы они могут застревать в циклах, поэтому, проходя по ссылкам, им нужны значительные ресурсы Сети. Однако, имеются методы, предназначенные для того, чтобы запретить роботам поиск по сайтам, владельцы которых не желают, чтобы они были проиндексированы.

Агенты извлекают и индексируют различные виды информации. Некоторые, например, индексируют каждое отдельное слово во встречающемся документе, в то время как другие индексируют только наиболее важных 100 слов в каждом, индексируют размер документа и число слов в нем, название, заголовки и подзаголовки и так далее. Вид построенного индекса определяет, какой поиск может быть сделан поисковым механизмом и как полученная информация будет интерпретирована.

Агенты могут также перемещаться по Интернет и находить информацию, после чего помещать ее в базу данных поискового механизма. Администраторы поисковых систем могут определить, какие сайты или типы сайтов агенты должны посетить и проиндексировать. Проиндексированная информация отсылается базе данных поискового механизма так же, как было описано выше.

Люди могут помещать информацию прямо в индекс, заполняя особую форму для того раздела, в который они хотели бы поместить свою информацию. Эти данные передаются базе данных. Когда кто-либо хочет найти информацию, доступную в Интернет, он посещает страницу поисковой системы и заполняет форму, детализирующую информацию, которая ему необходима. Здесь могут использоваться ключевые слова, даты и другие критерии. Критерии в форме поиска должны

соответствовать критериям, используемым агентами при индексации информации, которую они нашли при перемещении по Сети.

База данных отыскивает предмет запроса, основанный на информации, указанной в заполненной форме, и выводит соответствующие документы, подготовленные базой данных. Чтобы определить порядок, в котором список документов будет показан, база данных применяет алгоритм ранжирования. В идеальном случае, документы, наиболее релевантные пользовательскому запросу будут помещены первыми в списке. Различные поисковые системы используют различные алгоритмы ранжирования, однако, наиболее популярные в настоящее время принципы определения релевантности следующие:

1. Количество слов запроса в текстовом содержимом документа (т.е. в html-коде).
2. Тэги, в которых эти слова располагаются.
3. Местоположение искомых слов в документе.
4. Удельный вес слов, относительно которых определяется релевантность, в общем количестве слов документа.

База данных выводит ранжированный подобным образом список документов с HTML и возвращает его человеку, сделавшему запрос. Различные поисковые механизмы также выбирают различные способы показа полученного списка – некоторые показывают только ссылки; другие выводят ссылки с первыми несколькими предложениями, содержащимися в документе или заголовок документа вместе с ссылкой.

Когда Вы щелкаете на ссылке к одному из документов, который вас интересует, этот документ запрашивается у того сервера, на котором он находится.

CAD и CALS

Традиционный подход, сложившийся в первоначальный период внедрения вычислительной техники в производственные процессы, состоял в том, что с ее помощью решались отдельные, частные задачи, относившиеся к различным стадиям ЖЦ изделий. Исторически первыми здесь были задачи, позволяющие автоматизировать отдельные учетно-управленческие функции в рамках так называемой автоматизированной системы управления производством (АСУП). Почти одновременно с ними появились автоматизированные системы управления технологическими процессами (АСУТП). Затем стали разрабатываться и внедряться системы автоматизированного проектирования (САПР), которые позволяли использовать средства вычислительной техники в процессах конструкторской и технологической подготовки производства (в зарубежной технической литературе они известны под аббревиатурами CAE, CAD, CAM).

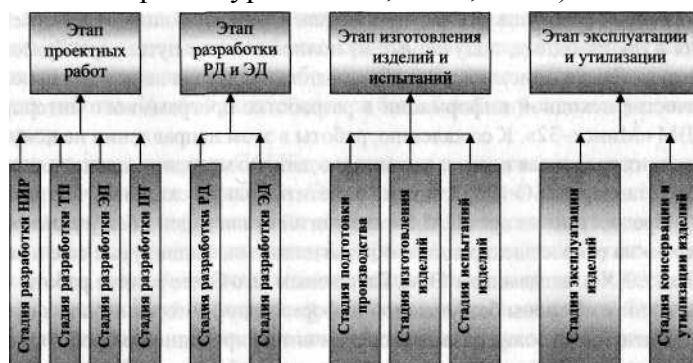


Рис. 21. Жизненный цикл изделия.

По мере увеличения числа таких систем на предприятиях возникла проблема передачи информации между ними, что естественным образом привело к идее информационной интеграции.

Еще в 1976 г. в методике Госстандарта СССР было указано, что в условиях автоматизации проектирования необходима единая система описания изделий. Методика была предназначена для выработки единого системного подхода к формализованному представлению конструкторской и технологической информации и формированию общего информационного языка. В ее основу была положена модель, структура которой состояла из совокупности элементов изделия и их отношений. Полная модель, содержащая всю информацию об изделии, представляла собой совокупность частных моделей, описывающих отдельные части изделия и производные, получаемые из полной модели путем преобразования параметров. Языки описания и преобразования предназначались для использования в качестве исходной информации в разработке программного интерпретатора на ЭВМ «Минск-32». К сожалению, работы в этом направлении не получили развития, хотя сравнивая идеи, заложенные в данную методику, с концепцией современного стандарта ISO 10303, можно отметить больше сходства, чем различий.

Предвестником века CALS-технологий стали идеи «безбумажной информатики» на основе электронного обмена данными, выдвинутые еще в самом начале 70-х гг. XX в. академиком В.М. Глушковым (его более ранние работы были собраны в книге «Основы безбумажной информатики», которая вышла в свет в 1982 г.).

Следующая волна развития систем интегрированной разработки и изготовления сложных наукоемких изделий в России наблюдалась в начале 80-х гг. XX в., когда в оборонных отраслях, а также в подразделениях среднего машиностроения осознали необходимость организации комплексного технологического процесса проектирования и производства.

Именно в то время родились такие системы, как КАПРИ (комплексное автоматизированное проектирование, разработка и изготовление).

Технологии с каждым днем усложняются, требуют все более квалифицированных сотрудников. Опыт и знания сотрудников предприятия становятся интеллектуальным капиталом и конкурентным преимуществом. Нельзя позволить бесследно исчезнуть опыту квалифицированных специалистов с многолетним стажем – опыт каждого из них бесценен. Старые методы передачи знаний, требующие не менее трех – пяти лет работы молодого специалиста в опытном коллективе, становятся непозволительной роскошью.



Рис. 22. Структура PLM.

В ответ на требования рынка в 90-х годах XX века была разработана концепция PLM, дополнившая три другие базовые концепции -ERP, CRM и SCM. К этому времени многие корпорации прошли большой путь автоматизации своей жизнедеятельности и достигли значительных успехов в отдельных направлениях. Но, вместе с тем, стало очевидно, что просто сумма отдельных информационных технологий уже не дает конкурентных преимуществ, не позволяет повышать эффективность работы, а самое главное – не позволяет удовлетворить требования заказчиков и потребителей. В первую очередь это коснулось корпораций выпускающих высокотехнологичную продукцию оборонного и стратегического назначения. Оборонная, авиа-космическая, судостроительная, приборостроительная, энергомашиностроительная и даже автомобильная, как и ряд других отраслей промышленности обнаружили, что удовлетворить требованиям CALS -технологий (или их идеологических эквивалентов в других отраслях), непрерывно не отслеживая эволюцию информации связанной с изделием, очень сложно и дорого.

Концепция PLM не просто охватывает функциональные области, покрываемые решениями CRM, SCM, ERP, но требует глубокой интеграции этих систем с процессом проектирования через управление составом изделия, поскольку именно на этой основе производится выбор поставщиков, управление себестоимостью, совместное планирование поставок комплектующих и сборочного производства с целью обеспечения сроков поставки при минимизации складских запасов, а также управление обслуживанием с учетом данных по реальной эксплуатации каждого изделия.



Рис. 23. Взаимодействие участников жизненного цикла изделия.

PLM не является догмой – это идеология, а не набор программных продуктов или бизнес-процессов. Для каждого предприятия существует свое оптимальное воплощение PLM. Вместе с тем эти воплощения для взаимодействующих предприятий должны включать в себя не только системы отдельных предприятий, но, в первую очередь, – систему управления бизнес-процессами взаимодействия участников на едином информационном пространстве, доступ к которому обуславливается потребностями участников при обязательном соблюдении требований безопасности.

Информационная поддержка жизненного цикла изделий (CALS)

Современная эпоха развития интеграции производственных данных во всем мире проходит под эгидой CALS-технологий – новой концепции развития производственной и коммерческой информатики [27].

Термин CALS появился в 1985 г. в оборонном комплексе США как аббревиатура интегрированной системы информационной поддержки процессов заказа, поставки, обслуживания, эксплуатации и ремонта средств вооружений и военной техники. Речь шла о стандартизации электронного представления и обмена технической и коммерческой информацией, позволяющей упорядочить и ускорить соответствующие процессы в федеральных структурах и вооруженных силах и сократить затраты, связанные с этим сложным информационным взаимодействием.

По сути, CALS – это протокол цифровой передачи данных, обеспечивающий стандартные механизмы их доставки и текущего инжиниринга для проектирования сложных технических объектов. При этом в качестве форматов данных в CALS используют специальные стандарты, например, IGES и STEP. В CALS входят также стандарты электронного обмена данными, электронной технической документации и руководства для совершенствования процессов.

Цель CALS формулируется достаточно просто: производитель обязан поставлять, предположим ВМС США, боевой корабль в комплекте не с эшелонной бумажной эксплуатационно-конструкторской документацией (и вагонами томов изменений к ней), а с актуальной трехмерной электронной моделью. На основе этой модели должна существовать возможность получения всех необходимых в процессе эксплуатации данных как о самом корабле, его конструктивных и тактико-технических характеристиках, так и

обо всех предписанных регламентом профилактических работах. И такой подход должен быть к любому сложному техническому объекту.

За прошедшие годы понятие CALS существенно расширилось и перестало быть прерогативой военного комплекса. Оказалось, что задачи совместного использования электронной информации и обмена ею в части данных о составе и структуре изделий, геометрических моделей, чертежей, технических руководств, описаний процессов, данных, касающихся материально-технического обеспечения, технологии информационной поддержки процессов эксплуатации сложной техники, не менее актуальны и в других отраслях, связанных с наукоемкой машинно-технической продукцией.

В России в последнее время устоялась следующая русскоязычная интерпретация термина CALS – информационная поддержка жизненного цикла изделий (ИПИ). Однако чаще всего этот русскоязычный термин используется, когда речь идет о средствах реализации (методах, технологиях, стандартах и т. д.) концепции и стратегии CALS. Подчеркивая, в общем случае, идентичность этих терминов, авторы пишут "CALS", когда говорят о концептуальных и стратегических вопросах информационной поддержки ЖЦ изделий, а "ИПИ" используется, если речь идет о средствах реализации этой информационной поддержки (ИПИ-методы, ИПИ-технологии, ИПИ-стандарты и т. д.).

В рамках международного комитета по стандартизации (ISO) были разработаны несколько десятков стандартов, закрепляющих накопленный в мире опыт ведения производственной деятельности с использованием электронного обмена данными.

В настоящее время работа многих крупных корпораций, разрабатывающих и производящих наукоемкую продукцию (авиакосмическая и автомобильная промышленности, судостроение) базируется на этих стандартах. Фактически понятие CALS получило новое звучание – сегодня это концепция организации и интегрированной информационной поддержки ЖЦ изделия, основанная на безбумажном обмене данными и стандартизации представления данных на каждом этапе ЖЦ.



Рис. 24. Концептуальная модель CALS.

Концепция и стандарты CALS определяют набор правил и регламентов, в соответствии с которыми строится взаимодействие субъектов в процессах проектирования, производства, испытаний, эксплуатации, сервиса и т. д.

В отличие от бумажного и простейших форм электронного документооборота, основанного на электронных образах бумажных документов, в рамках CALS речь идет об использовании интегрированных информационных моделей (баз данных) продукции и процессов. Здесь на практике реализуются соответствующие информационные технологии (CALS-технологий) и нормативные базы (стандарты).

Целью применения CALS как концепции организации и информационной поддержки бизнес-деятельности является повышение эффективности процессов разработки, производства, послепродажного сервиса, эксплуатации изделий за счет:

- ускорения процессов исследования и разработки продукции;
- сокращения издержек при производстве и эксплуатации продукции;
- придания изделию новых свойств и повышения уровня сервиса в процессах его эксплуатации и технического обслуживания.

Таким образом, CALS необходимо рассматривать как инструмент повышения эффективности бизнеса, конкурентоспособности и привлекательности продукции.

Применение CALS активно развивается, прежде всего, в разработке и производстве сложной наукоемкой продукции, создаваемой интегрированными промышленными структурами, включающими в себя НИИ, КБ, основных подрядчиков, субподрядчиков, поставщиков готовой продукции, потребителей, предприятия технического обслуживания, ремонта и утилизации. Развитие концепции CALS обусловило появление новой организационной формы выполнения крупных проектов – "виртуального предприятия" –

объединения на контрактной основе фирм, участвующих в процессах поддержки ЖЦ изделия и действующих на основе общей системы стандартов информационного взаимодействия. В рамках "виртуальных предприятий" реализуются совместные проекты по разработке, производству, сбыту и обеспечению сервисного обслуживания различных видов наукоемких товаров.

CALS-технологии позволяют эффективно, в едином ключе решать проблемы обеспечения качества продукции, поскольку электронное описание процессов разработки, производства, монтажа и т.д. полностью соответствует международным стандартам ISO 9000.

В настоящее время целый ряд отечественных предприятий в рамках международного сотрудничества, в частности при продаже сложных наукоемких изделий, и лицензий на их производство, уже столкнулись с требованиями соблюдения стандартов CALS в электронной форме технической документации, а также в средствах компьютерной информационной поддержки процессов технического обслуживания, материально-технического обеспечения, заказа запасных частей и ремонта.

Похожие проблемы возникают при электронном взаимодействии и совместном использовании конструкторской, производственной и коммерческой информации в электронной форме. Таким образом, практическое применение CALS-технологий является чрезвычайно актуальной задачей.

Новизна концепции CALS заключается в следующем:

1. Широта охвата и системность подхода (речь идет не только о производстве или проектировании, но и о поддержке всех процессов в жизненном цикле – от замысла до утилизации продукта).
2. При расширении использования компьютерных технологий на повестку дня выходят проблемы информационной интеграции автоматизированных систем.
3. Интеграция достигается путем стандартизации представления информации (или результатов) в процессах проектирования, материально-технического снабжения, производства, ремонта, послепродажного сервиса и т. д. Это обеспечивает оперативную передачу функций одного подрядчика другому, который, в свою очередь, может воспользоваться результатами уже проделанной работы. Такая возможность особенно важна для изделий, имеющих длительный ЖЦ, когда необходимо поддерживать преемственность информационной поддержки продукции независимо от складывающейся рыночной или политической ситуации.
4. Эффективный бизнес в данный момент имеет явную тенденцию к географической распределенности. В случае изменения состава участников – смены поставщиков или исполнителей обеспечиваются преемственность и сохранность уже полученных результатов (моделей, расчетов, документации, баз данных).

Основные преимущества применения CALS:

- сокращение времени выхода изделия на рынок (сокращение временных издержек);
- сокращение стоимости ЖЦ (сокращение материальных издержек);
- повышение качества изделия.

Главными проблемами, мешающими эффективному управлению информацией об изделии, являются огромное количество информации ("информационный хаос") и коммуникационные барьеры между участниками ЖЦ изделия.

Пути их решения заложены в стратегии CALS. Стратегией CALS является создание единого информационного пространства (ЕИП) для всех участников ЖЦ изделия, включая потребителя.

Преодоление информационного хаоса и коммуникационных барьеров между участниками ЖЦ изделия приведет к улучшению взаимодействия между ними и повышению эффективности процессов ЖЦ. Результатом станет снижение временных и материальных издержек и возрастание удовлетворения потребностей заказчика, а это, в свою очередь, неизбежно повысит конкурентоспособность изделия. В основе ЕИП лежит использование открытых архитектур, международных стандартов, совместных хранилищ данных и апробированных программно-технических средств. ЕИП обеспечивает совместную работу проектных организаций, производственных предприятий, поставщиков, организаций сервиса и конечного потребителя на всех стадиях ЖЦ.

Стратегия CALS предусматривает двухэтапный переход к ЕИП:

1. Автоматизация отдельных процессов (или этапов) ЖЦ изделия и представление данных на них в электронном виде в соответствии с требованиями ЕИП. Предполагается, что на этом этапе обмен данными между исходными системами осуществляется отдельными файлами (электронными документами) на магнитных носителях либо по сетям.

2. Интеграция автоматизированных процессов и относящихся к ним данных, уже представленных в электронном виде, в рамках ЕИП. Здесь взаимодействие осуществляется с помощью программных средств в режиме реального времени, параллельная работа исполнителей организуется через единую компьютерную среду.

ИПИ-технологии представляют собой набор методов реализации стратегии CALS (то есть набор методов создания ЕИП) и достижения целей, заложенных в концепцию CALS. Всего можно выделить три группы ИПИ-технологий:

- Технологии представления данных об изделии в электронном виде (первый этап создания ЕИП).
- Технологии интеграции данных об изделии в рамках ЕИП (второй этап создания ЕИП).
- Технологии реинжиниринга бизнес-процессов. Они применяются для изменения структуры процессов ЖЦ. Эти технологии достаточно хорошо известны и активно используются в реальных проектах. В данной работе авторы на них останавливаться не будут.

ИПИ-технологии. Технологии представления данных

Технологии представления данных являются набором методов, языков и моделей для представления в электронном виде информации об изделии, относящейся к отдельным процессам его ЖЦ.

Предметная область – это неструктурированное множество объектов в совокупности со значениями их свойств и существующие на этом множестве отношения.

Информационная модель представляет собой множество понятий (сущностей) в совокупности со значениями их свойств (атрибутов) и заданные на этом множестве отношения.

При автоматизации каждого этапа ЖЦ изделия (или его части) создают частную информационную модель предметной области, для чего:

- формируют множество понятий, отображающих объекты предметной области, необходимые для решения поставленной задачи;
- формируют множество атрибутов понятий, отображающих свойства объектов предметной области, необходимых для решения поставленной задачи;
- устанавливают отношения между понятиями, соответствующие отношениям между объектами предметной области.

Множество понятий и присущих им атрибутов образуют модель частной задачи, а множество отношений между понятиями – логическую основу процедур и алгоритмов обработки данных.

Как правило, модель на каждом этапе служит только для обмена информацией об изделии, то есть является источником первоначальной информации для всех прикладных систем, использующихся на данном этапе, и собирает все результаты их работы. При этом преобразование исходной информации в результирующую на каждом этапе происходит только в прикладных системах.

Основная трудность технологий представления данных состоит в том, что в реальной практике используется слишком большой набор методов, языков и моделей представления информации об изделии.

Технологии интеграции данных представляют собой набор методов для интеграции автоматизированных процессов ЖЦ и относящихся к ним данных, представленных в электронном виде, в рамках ЕИП.

Интегрированная модель ЖЦ изделия должна включать в себя всю информацию о моделях частных задач для каждого этапа ЖЦ и об их взаимосвязях.

Интеграция нужна для создания структуры, объединяющей частные модели, соединения множеств понятий, атрибутов и отношений, а также определения отношений между множествами понятий частных моделей.

Понятия ЕИП и информационной интеграции за время развития концепции CALS менялись.

Цель интеграции на первом этапе – формирование на основе информационных стандартов совокупности электронных документов, аналогичных бумажным, и передача их участникам ЖЦ изделия. Основные понятия данного этапа: электронный документ и электронный документооборот. Главные вопросы, подлежащие решению: идентификация документов и их авторов, взаимно однозначное восприятие документов человеком и компьютерными системами с минимальным количеством преобразований, защита информации от потери и несанкционированного доступа на этапах хранения, передачи и использования.

Второй этап интеграции – создание взаимосвязанной совокупности процессов, составляющих ЖЦ изделия. В результате образуется виртуальное предприятие.

Информационные процессы являются процессами ЖЦ изделия, создающими или использующими данные о нем. Примером служит формальная процедура изменения изделия. Совокупность информационных процессов представляет собой документооборот, происходящий в течение ЖЦ изделия. Естественно, документооборот, управляемый PDM-системой, называется электронным документооборотом. При решении глобальной задачи ИПИ-технологий – повышения эффективности управления информацией об изделии – роль PDM-технологии состоит в том, чтобы сделать информационные процессы максимально прозрачными и управляемыми. Основным методом, применяемым для этого, является повышение доступности данных для всех участников ЖЦ изделия, что требует интеграции данных об изделии в логически единую информационную модель.

Существует много задач, решаемых с помощью PDM-технологии, среди которых можно выделить наиболее распространенные:

- создание ЕИП для всех подразделений предприятия; автоматизация управления конфигурацией изделия;
- построение системы качества продукции на предприятии согласно международным стандартам серии ISO 9000;
- создание электронного архива чертежей и прочей технической документации (наиболее простой способ применения PDM-технологии).

Для реализации PDM-технологии существуют специализированные программные средства, называемые PDM-системами (системы управления данными об изделии). Они представляют собой новое поколение компьютерных средств управления всеми связанными с изделием данными и информационными процессами ЖЦ. В отличие от АСУП, контролирующей информацию о ресурсах предприятия, PDM-системы направлены именно на управление информацией о продукте. У нас в стране PDM-системы часто называют системами управления проектом. И действительно, PDM-система фактически предназначена для работы над проектом по разработке, производству и продвижению на рынок наукоемкого промышленного изделия.

Среди задач PDM-системы можно выделить две основные:

- PDM-система как рабочая среда пользователя.
- PDM-система как средство интеграции данных на протяжении всего ЖЦ изделия.

PDM-система должна выступать в качестве рабочей среды любого сотрудника предприятия, который в процессе своей работы постоянно находится в ней, а система, в свою очередь, обеспечивает абсолютно все его потребности, начиная от просмотра спецификации узла и кончая изменением твердотельной модели детали или утверждением измененной детали начальником. При необходимости PDM-система должна пользоваться помощью других систем для обработки данных (например, САПР), самостоятельно определяя, какое именно внешнее приложение необходимо запустить для обработки той или иной информации.

Пользователями PDM-системы выступают все сотрудники предприятий-участников ЖЦ изделия (конструкторы, технологи, работники технического архива), а также служащие, работающие в других предметных областях: сбыт, маркетинг, снабжение, финансы, сервис, эксплуатация и т. п. Главной задачей PDM-системы как

рабочей среды пользователя является предоставление соответствующему сотруднику нужной ему информации в нужное время в удобной форме (в соответствии с правами доступа).

Что же должна делать идеальная рабочая среда? Ниже приведен классический перечень функций PDM-системы:

- Управление хранением данных и документов. Все данные и документы в PDM-системе хранятся в специальной подсистеме (хранилище данных), которая обеспечивает их целостность, организует доступ к ним в соответствии с правами доступа и позволяет осуществлять их поиск.
- Данные документы являются электронными, то есть обладают электронной подписью.
- Управление процессами. PDM-система отслеживает все операции пользователей с данными, в том числе следит за версиями. Кроме того, PDM-система управляет потоком работ (например, в процессе проектирования изделия) и занимается протоколированием действий пользователей и изменений данных.
- Управление составом изделия. PDM-система содержит информацию о составе изделия, его вариантах. Важной особенностью является наличие нескольких представлений состава изделия для различных предметных областей (конструкторский состав, технологический состав, маркетинговый состав и т. д.).
- Классификация. PDM-система позволяет производить распределение изделий и документов в соответствии с различными классификаторами.
- Это может быть использовано при поиске продукта с нужными характеристиками с целью его повторного использования или для автоматизации присваивания обозначений компонентам изделия.

Информационные системы в производстве

Сегодня создано огромное количество информационных систем общего назначения для различных задач обработки данных, их применение является одним из главных инструментов повышения эффективности промышленного производства.

Однако, часто необоснованно трафаретно применяется докомпьютерная производственная классификация подразделений: снабжение, сбыт, финансы, производство, вместо создания новой структуры. При этом информационная система должна обеспечивать сквозную поддержку функционирования производственного процесса.

Проект внедрения информационных систем позволяет автоматизировать основные управленческие процессы, в том числе и производственный сегмент – один из самых сложных участков для автоматизации, охватывая все производственные процессы и больше других подходит для сложного производства.

Рассмотрим для примера что дает внедрение EPR системы на производственном предприятии на примере IFS Applications.

Особенностью этой системы является автоматизация четырех ключевых управленческих областей компании (снабжение, сбыт, финансы и производство) в рамках единой программно-аппаратной платформы.

Она необходима заводу для проведения автоматизированной загрузки и актуализация данных о составе изделия, нормах расхода материалов, технологическом маршруте, временных нормативах на работу оборудования и т.д.

Внедрение решения «Производство» начинается с автоматизации подсистемы конструкторско-технологической подготовки производства, импорта нормативной базы из существовавшей ранее системы с последующей актуализацией. Технологически обоснованные нормы – основа принятия рациональных управленческих решений. Теперь на отработанной нормативной базе о техмаршрутах и спецификациях в соответствии с методологией МРП на предприятии рассчитывается и актуализируется график обеспечения потребностей в материалах, деталях и сборочных единицах. Он позволяет определять, когда и в каком количестве потребуются покупные материалы и комплектующие, и когда необходимо запускать работу в цехах для того, чтобы произвести готовую продукцию к заданному сроку, а также проверить достаточность для выполнения графика существующих производственных мощностей. Корпоративная информационная система также снабжает менеджмент завода действенным механизмом контроля прямых производственных затрат, позволяющим локализовать источники отклонений фактических затрат от нормативных.

При внедрении модуля «Снабжение» на предприятии появляется возможность снижать издержки в области закупок благодаря проведению анализа предложений поставщиков с учетом цен и минимальных партий поставки, а также путем своевременного выполнения графика закупок. Благодаря модулю «Сбыт» можно отследить заказ клиента от начала его планирования до отгрузки потребителю. Работа в системе ускоряет и упрощает получение информации о складских остатках продукции, заказах и отгрузках.

При этом важным требованием является необходимость настройки системы для возможности учета особенностей российского законодательства.

Что касается потребностей клиентов, то еще лет десять назад можно было сказать, что та или иная система российского производства не имеет необходимых возможностей. Сейчас работать на корпоративном рынке и не обладать тем или иным функционалом нельзя. В части охвата предметных областей все основные продукты практически равны. Если у разработчика нет готового решения, о котором заказчики только задумываются, то на этом рынке он долго не продержится.

В данном случае вопрос несколько шире: речь идет об адаптивности системы. Часто проект ведется следующим образом: принятое руководством решение не учитывает и интересов в плане удобства каждодневной работы рядовых пользователей. Усложняется подготовка отчетов, выкидываются привычные и совершенно необходимые для конкретных исполнителей вещи.

Необходимо при внедрении системы, ориентироваться, в том числе и на исполнителей, стремиться обеспечить комфорт на каждом рабочем месте. Если сотрудника компании-заказчика обяжут вводить данные параллельно основному учету еще в одну систему, а в начале работы он не обнаружит в системе привычных мелочей, на порядок

облегчавших работу (например, невозможно определить, в какое дело подшит документ), то для служащего автоматизация станет шоком.

Вторая составляющая «бесшокового» внедрения заключается в том, что в процессе выполнения проектов осуществляется постепенный переход от использования старого программного обеспечения к новой системе. Данные автоматически попадают и в старую, и в новую базы, что дает возможность избежать двойного ввода, сохранив при этом работоспособность обеих систем и прежний порядок деятельности.

Как правило, большинство крупных компаний идет на такие временные затраты, сохраняя как работоспособность предприятия, так и лояльность персонала, и не увеличивая при этом затраты на оплату труда за сверхурочную работу специалистов. Для предприятий, где реализуются короткие проекты, такие проблемы нехарактерны: при сроке проекта в 6-9 месяцев можно и потерпеть двойной ввод на определенных участках.

Последняя составляющая «бесшокового» внедрения – это проведение работ в целом по проекту и по каждому из участков в соответствии с технологией, основанной на российских и международных стандартах создания информационных систем и управления проектами. Можно было бы не акцентировать внимание на данном аспекте, однако существует достаточно много примеров, когда попытка ввода систем в эксплуатацию без периода тестирования (по принципу «настроили – запустили») заканчивается тем, что большая часть настроек переделывается, а учет приходится нагонять в авральном порядке, вернувшись к использованию старого программного обеспечения. Поэтому «бесшоковое» внедрение – это не только отсутствие двойного ввода при параллельном учете, но и обеспечение комфортного режима и удобства работы – как в процессе внедрения новой информационной системы, так и при ее эксплуатации.

Семиологические информационные системы

Приведенный выше обзор информационных систем и технологий наглядно демонстрирует, что наблюдается разрыв между узкоспециализированными высокоэффективными программными средствами и постоянно усложняемым интерфейсом **корпоративного** общения с информационными системами. По сути, это приводит к необходимости создания следующего уровня так называемых семиологических информационных систем, где во главу угла ставятся принципы адекватного восприятия коммуникаций как согласованного процесса между программируемой средой и интерпретационным текстологическим содержанием, адекватно воспринимаемого на каждом уровне пользователем корпоративной системы.

Для коммуникативного акта не важна грамматика и прагматика. Онтология предметной области проявляется не в словах, а в связях между понятиями. Словом в любой онтологии могут служить любые идентификаторы и способ их связи определяются только структурой языка. Люди, изучающие иностранные языки, приходят к выводу: бесполезно добиваться ответа на вопрос «почему так, а не иначе?». Ответ гласит: «потому, что так принято в данном языке и, возможно, только в нем».

Кроме того, необходимо отметить, что в обыденном языке, безусловно, нет четкого соответствия между знаком (словом) и обозначаемым.

1. Слово имеет, как правило, несколько значений. Даже узкоспециализированные термины переходят из одной отрасли знания в другую, меняя при этом свое значение; в других случаях углубление знаний об объекте, класс которого называется термином, приводит к уточнению и обогащению его значения. Атом для античных ученых и для современных ядерщиков – разные слова (разные значения при одной неизменной форме). Многозначность называют полисемия.

2. Внешне на полисемию похоже явление омонимии: два или более слова с совершенно разными истинными значениями по форме оказались идентичными: топить печку и топить лодку.

3. Разные формы могут служить вместилищем одних и тех же или очень близких значений (синонимия): идти в ногу – маршировать; пурга – вьюга – буря – метель.

4. Одна и та же форма может вмещать прямо противоположные значения (энантиосемия): просмотрели означает "видели, увидели" и "не увидели".

Получается, что значение не прикреплено к определенной форме (или наоборот: одна форма не прикреплена к определенному значению). Можно назвать мудрецом человека, допустившего глупость, мастером – неумеху, шедевром – бездарное изделие.

Ирония понимается с помощью ситуации. Иронической интонации мы не услышим, но поймем верно.

Мы понимаем слово не как таковое, а слово в ситуации, вместе с ней, через нее или в контексте, заменяющим ситуацию.

Семиологический подход исходит из следующих тезисов.

1. Человек или компьютер или любая другая коммуникационная система (Интернет и др.) используя феномен коммуникации должны учитывать:

- язык предшествует и учреждает акт коммуникации;
- не источник и потребитель использует тот или иной язык, а язык лишь выражает семантико-смысловое содержание информационного сообщения.

2. Потенциал информационной защиты состоит в недостаточности познавательных способностей агента как субъекта источника и потребителя, что превращает коммуникацию в чередование того «что знаем», с тем «чего не знаем». Стало быть, структуры разных языков и исторически сложившиеся коды, протоколы, форматы существуют как трансляторы информационного содержания, но не структура языка как такового, не некая Пра-Система, не Код Кодов.

Требуется некая метасистема, постоянно корректирующая и осуществляющая адаптацию новых терминов и понятий, и первые шаги в этом направлении и связаны с построением таких систем CALS-технологии, системы логистики.

В качестве примера можно привести подход [28] к построению подобных систем (рис.25, 26). В данном подходе под пространством посредников «Agent Space» понимается информационная модель предметной области, которая реализуется и поддерживается автономной системой коммуникаций. Форма преподнесения знаний должна соответствовать уровню понимания потребителя информационных сообщений. Используемые английские термины «Autonomous Negotiating Teams» подчеркивают автономность адекватного понимания информационного сообщения (агентом-потребителем). Адаптеры выполняют функцию согласования форматов, протоколов,

«языков» физического уровня коммуникаций, в том числе и различных версий программных продуктов.

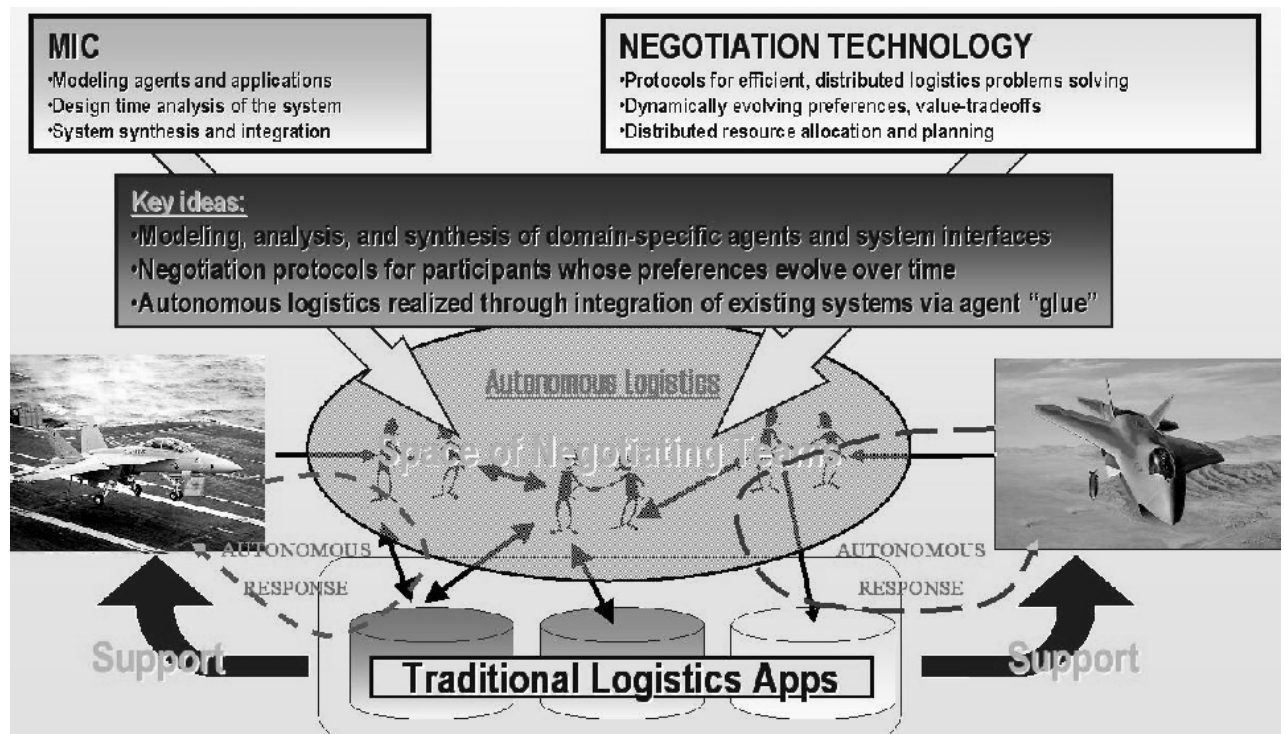


Рис. 25. Общая структура системы MICANTS
(Model Integrated Computing and Autonomous Negotiating Teams)

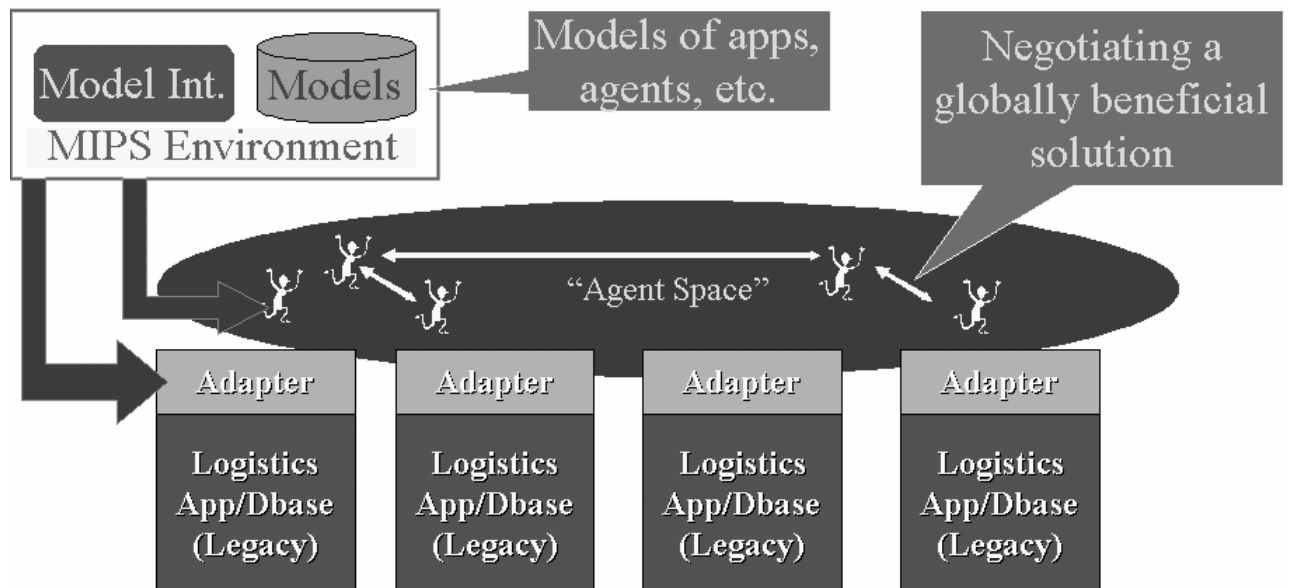


Рис. 26. Концепция системы MICANTS

Информационная безопасность.

Предлагаемый подход семиологической неопределенности, являясь аналогом физического принципа Гейзенберга, выделяет два направления информационной защиты: традиционное **физическое** (физический канал, криптография и др.) и **языковой** принцип соответствия понятийно-предметной ориентации, т.е. семантико-смысловое кодирование.

Мы специально оставили английские термины на рис. 25, 26, так как проблема выбора языка функционирования специфических информационных компьютерных систем все в большей степени заостряется в связи с глобализацией экономических и политических процессов. Так в Известиях от 05.04.2005 приводится интервью с главой военной миссии НАТО в Москве Питером Вильямсом: «Военные, участвующие в совместной операции, должны говорить на **одном языке**»:

*...В Генштабе есть понимание, что одним из главных аспектов этого должны быть **овладение языком, совместимость оборудования и военной техники**. С точки зрения НАТО мы рассматриваем проект создания Самарской бригады как основу для будущей подготовки других российских частей, которые, возможно, будут принимать участие в совместных с нами боевых операциях...*

...Пока одна из самых сложных проблем— это лингвистическая совместимость. Солдаты и офицеры, участвующие в совместной операции, должны говорить на одном языке. Штабные офицеры должны знать общую терминологию при выработке решений и планов боевого применения. Как мне показалось, командование бригады, Генштаб и Минобороны России это очень хорошо осознают, и это серьезный шаг вперед по сравнению с тем, что было у нас на Балканах.

Тогда только участие США позволяло решить проблему лингвистической совместимости. В Самаре безусловно хорошие солдаты, но они должны уметь разговаривать со своими коллегами...

...Например, еще 10 лет назад французские офицеры не владели английским языком. А сегодня операцией в Афганистане руководит французский генерал.

У нас существуют программы дистанционной помощи в овладении языком. Но для их применения нужны компьютеры и интернет в каждом воинском подразделении. В миротворческой бригаде пока такой техники нет. Поэтому мы пока не рассматриваем возможности перевода наших обучающих программ на русский язык. Солдаты и офицеры очень молоды, но у них правильное командование. Мы видели, что в часть вкладываются серьезные ресурсы, что под Самарой создается хорошая база, на которой потом будет строиться дальнейшее здание. Возможно, скоро дело дойдет и до языка, во всяком случае, и в Минобороны, и в Генштабе есть понимание того, что это необходимо...

Проблема информационной безопасности сегодня все еще преподносится в большей степени как инженерно-техническая, структурно-организационная и технологическая дисциплина. Самые совершенные научные и технологические разработки касаются программ криптографии, стеганографии, защиты от несанкционированного пользования.

Эта информационная защита не имеет ничего общего с защитой информационного общества от общественно опасных, разрушающих основы личности, культуры и духовности, гипнотизирующих информационных потоков. Вспомним известные факты погружения в транс и неадекватное поведение при воздействии спецэффектов аудиовизуального ряда. Пока это экспериментальные и случайно обнаруженные эффекты. Теория нейролингвистического программирования, разработанная задолго до компьютерной эры, позволяет управлять процессом личностного изменения. Ограничения в распространении были связаны с отсутствием возможности эффективного тиражирования, с появлением Интернет-сообществ это ограничение уже не действует.

Другой пример – рецепт изготовления яда или атомной бомбы легко найти на сайтах Интернет-сообществ. Прогресс электронных технологий в получении и распространении информации и знаний вызывает потребность в создании концепции безопасности информационного общества, аналогичной концепции безопасности распространения ядерного вооружения. Уже раздаются призывы к запрещению информационных войн и созданию международных комитетов по контролю за информационным зомбированием. Примеры попыток ограничения распространения различных сект и учений даже в отсутствии е-технологий очень часто были и неудачными и приводили к огромным жертвам. Новый виток информационных технологий требует не свободы, а построения независимых, несвязанных друг с другом общественных и властных структур контроля, цензуры и ответственности. От диспута об ответственности писателя, журналиста к статье об ответственности за информационное провоцирование.

Семиологический подход расширяет область информационной безопасности с традиционных методов, использующих лишь кодирование физических носителей информационного содержания, на методы восприятия (инициализации) семантико-смыслового содержания. При этом информация понимается как возможность целенаправленного управления (выбора) на основе полученного сообщения и ставшего означающим (семантико-смысловым компонентом), когда оно получает истолкование на основе того или иного лексикода и, следовательно, зависит от адресата.

Эта информационно-смысловая компонента, инициализируемая сообщением, отличается от количественно исчисляемой информации источника, имеющей дело с физическим каналом, и называется семиологической информацией. Она не исчисляется с помощью количественных методов, но определима через ряд значений, которые могут возникнуть под воздействием разных кодов.

В мире знаков семиология раскрывает мир концептуального мышления, выражающегося в уже устоявшихся способах текстуального общения. В этом случае следует называть ее системой, которая объясняет себя путем последовательного разворачивания все новых и новых конвенциональных систем. Поэтому язык, описывающий другой язык (метаязык), определяется семиологией как иерархией метаязыков. Следовательно, исходные означаемые – метки Боно [29] или денотативные значения, устанавливаются кодом, а созначения (ассоциативные синонимы) зависят от вторичных кодов, или лексикодов, присущих не всем, а только какой-то части носителей языка; и так вплоть до возникновения жаргона, неожиданных метонимий и др. В результате адресат должен сам справляться с контекстом, т.е. раскодировкой информационного сообщения и выделением смысловой компоненты. Вербальное мышление человека интуитивно справляется с возникающими сложностями коммуникационного процесса. Однако интенсивное развитие инфотелекоммуникационных сетей, компьютерных технологий обработки информационных (аудио, видео) потоков, а также глобальный информационный обмен посредством Интернет-технологий и вызывает необходимость создания семиологических компьютерных технологий.

Нет никакой необходимости знать, что представляет собой означаемое (с онтологической или физической точки зрения), достаточно сказать, сославшись на код (MP3, PAL, SECAM, MPEG), что таким означающим соответствуют такие означаемые.

Впоследствии эти означаемые повсеместно воспринимаются как "понятия" или даже как "мысли", и то, что они могут быть объяснены правильным словоупотреблением, тоже вполне законно. Но в тот миг, когда семиология устанавливает факт наличия кода, значение уже не есть психологическая, онтологическая или социологическая данность: это факт культуры, который описывается с помощью системы отношений, устанавливаемый кодом и усваиваемой данным сообществом в данное время. Это и есть основное свойство семиологического подхода с точки зрения сокрытия информации.

Заметим, что гносеологическое понятие информации связано со свойствами ее структурированности, подмеченное еще Г. Лейбницем и сформулированное через понятие конгруэнтности – принципа идентификации неразличимости как универсальной характеристики. Подчеркнем, что физическая и семиологическая информация не вполне одно и то же, но та и другая несут смысл понятия «информации».

Устаревшее наивное понимание проблемы информационной безопасности как защиты каналов с помощью методов криптографии уже не является достаточным. Новый уровень защиты информации можно обеспечить за счет использования понятийных соотношений.

Технология поддержки жизненного цикла изделия (CALS – Computer Aided Logistics Support), включая в себя этап разработки, позволяет разрабатывать и системы информационной безопасности. Но даже простое использование CALS при создании любых продуктов дает базу для систем информационной безопасности нового поколения в виде принципа семантического несоответствия. Применение закрытых систем поддержки жизненного цикла изделия включает в себя использование непротиворечивого онтологического словаря предметной области и использование парадигмы мышления, оптимальной для данной конкретной предметной области.

Таким образом, терминологическое обеспечение проблемы моделирования некоторой предметной области состоит в переупорядочении идентификаторов понятий относительно этой проблемы.

Следует решить такие вопросы, которые требуют уточнения множества понятий, которые в разных науках трактуются по-разному, либо отсутствуют вообще. Возникает сложнейшая проблема создания терминологического обеспечения, в которой образуется два основных направления: использования традиционной терминологии на стыках наук и ее адаптация на множество порожденных понятий, и создания в связи с этим порождением новой терминологии. Кроме терминологических нововведений, вносимых в каждую из наук, оказываются необходимы средства для интерпретации порожденных на этих стыках отношений.

Идеальным способом защиты информации является различие парадигм мышления защищаемой группы и внешней, по отношению к ней, среды. Тогда для группы людей, работающих над единым проектом, в едином нормативном поле при отсутствии пересечений нормативных полей этой группы и внешней среды обеспечивается эффективная защита.

Традиционное понимание информационной безопасности как задачи технической, юридической и других подсистем является экстенсивным подходом к проблеме. Например, криптография, т.е. алгоритмы перекодировки символьной информации,

потенциально приводит к возможности раскрытия, а несоответствие понятийной базы принципиально не позволяет реконструировать исходное сообщение.

Формирование системы, обеспечивающей информационную безопасность объекта, требует обычно решения ряда задач, связанных с формализованной информацией – информацией взаимодействия в форме документов или обменных сигналов технических систем. В этих случаях вполне применимы методы математической теории информации и удастся сформировать весьма точные значения параметров, характеризующих защищенность системы на уровне информации взаимодействия. Однако для полной оценки защищенности эти параметры приходится сопоставлять с оценками семантического соответствия.

Работа в некотором нормативном поле предполагает развитие конкретного сценария при возникновении определенной ситуации или подаче некоторой команды. Примером может служить команда на запуск ракеты, которая влечет за собой конкретную последовательность действий по проверке санкционирования, организации режима подготовки и непосредственно запуска. Человек, не находящийся внутри этого нормативного поля, не имеет представления о точном смысле этой команды, т.е. о ее реализации. По аналогичному пути идут различные группы людей, создавая сленг, жаргон, пытаясь при этом скрыть смысл при передаче информации, используя собственный словарь. Закрывается смысловая сторона сообщения, а не ее сигнальные атрибуты.

Принципиально это позволяет говорить о стеганографии не как о скрытой передаче, а как о возможности в одном потоке данных передавать несколько, по-разному семантически наполненных, информационных потоков. При этом один поток выделяется с помощью общеупотребительного словаря, а другие потоки восстанавливаются с помощью специальных словарей.

Даже в случае перехвата сообщения воспользоваться информацией будет невозможно, т.к. работа происходит в другом формализме. Человек, работающий в другом семантическом поле, не сможет правильно интерпретировать полученные сведения. Попытка активного несанкционированного участия в информационном обмене, в таком случае, может выявляться по неадекватным ответам или, в общем случае, по противоречиям, возникающим в словаре, если изначально выполняется условие работы в непротиворечивом словаре, которая может осуществляться простой формальной проверкой.

Глоссарии и семантическое упорядочение информационных потоков

Ключевым моментом создания глоссария, от которого впоследствии зависит качество конечного продукта, является выбор метода отбора наиболее информативных и характерных для исследуемой предметной области терминов и определений, а также метода формирования глоссария.

Основной проблемой создания глоссария является выбор антологии, т.е. формирование набора документов, с помощью которого будет сформировано множество терминов и дано их толкование.

Многие авторы предлагают для выбора базовых терминов и задания связей между этими терминами использовать экспертные методы. Следует сказать, что привлечение экспертов в данной задаче необходимо, но полностью полагаться на них нецелесообразно в связи с возможностью появления следующих ошибок: различные понятия у различных людей ассоциируются с различными объектами, зависимость терминологии от сферы деятельности эксперта, и др.

Формулировка задачи для абстрактной системы, решающей задачу автоматического построения глоссария некоторой предметной области как ментальной модели лексикона человека, может быть дана в следующем виде. Имеется несколько видов текстов: указание, вопрос, повествование, а также их комбинации. Для задачи создания глоссариев наибольший интерес представляют повествовательные тексты, как текстовые данные, несущие описание новых понятий и устанавливающие связи между новыми и уже существующими понятиями. Система анализа текстов в процессе работы создает математическую модель анализируемых текстовых материалов, которая расширяется и уточняется при поступлении новых текстовых данных. В результате работы система формирует следующие объекты:

- список всех терминов, используемых в данной предметной области, ранжированных по степени информативности с толкованием их значения (в списке не должно оказаться стоп-слов и служебных слов языка),
- семантическую модель предметной области, описывающую отношения между понятиями в некотором формализме (например, в виде графа, понятия в котором являются вершинами, а отношения синонимии, антонимии и гипонимии являются дугами),
- список терминов, которые не удалось включить в семантическую модель, т.е. связать с другими из-за недостаточности соответствующей информации в тексте,
- реферат текста, который, в пределе, является одним предложением, выражающим смысл текста.

Естественные языки проявили удивительную способность к понятийно-ассоциативной идентификации. Появление компьютерных информационных технологий привело к резким изменениям коммуникативных функций, что вызвало не только резкое изменение тезауруса, но и определенные требования к языку, конструкция которого либо адекватна современным коммуникативным возможностям либо нет.

Человек–программист, реализуя интерфейс общения компьютера с пользователем, невольно переносит на компьютер ограничения свойственные человеку. Компьютер, имея другие принципы представления информации, оказывается ограниченным с одной стороны своими техническими возможностями, а с другой искусственными ограничениями, созданными человеком. Это сильно ограничивает возможности компьютера по сравнению с человеком.

Сравнение математических и лингвистических методов автоматизированного анализа текстов, допустимых при формировании глоссария, дается в [30]

Внедрение математических методов в обработку текста проходит на фоне отсталости собственно лингвистической составляющей алгоритмов, что не позволяет достичь высокого качества работы прикладных систем. Ставший устойчивым уклон в

область статистических методов анализа привел к тому, что компьютерная лингвистика на время оказалась оставлена в стороне. Господствующего мнения, что лингвистические алгоритмы являются ненадежными, слабо масштабируемыми и чересчур медленными для решения реальных задач, до недавнего времени придерживались многие авторы. Действительно, в известных системах анализа русскоязычного текста из лингвистического обеспечения используется не более, чем словарь морфологического анализа, приводящий слова к нормальной форме. Алгоритмы синтаксического анализа промышленно реализованы только в автоматических переводчиках и до сих пор вызывают множество нареканий за недостаточную точность.

Сложность реализации высокоточного анализатора связана с наличием тесной связи между синтаксисом и семантикой, присутствием в текстах русского языка большого количества синтаксически омонимичных конструкций, не допускающих однозначной интерпретации без привлечения знаний о семантической сочетаемости слов.

Модель языка, призванная объединить синтаксическую и семантическую составляющие, известна под названием толково-комбинаторного словаря и призвана описать ограничения на сочетаемость лексических единиц. Однако, помимо колоссального объема необходимого ручного труда, разработку подобного словаря усложняет отсутствие достаточно полной и устоявшейся системы классификации типов синтагматических отношений (например, базиса лексических функций), а также парадигматической классификации лексики (тезауруса).

Указанные проблемы привели к установившемуся мнению о нецелесообразности введения модуля синтаксического разбора в системы автоматического анализа текста. Однако оказалось, что, несмотря на ограниченную точность синтаксических анализаторов, их использование способно заметно повысить качество таких систем в случае комбинирования с известными статистическими методами, не выходя за рамки стандартных ограничений на вычислительные ресурсы.

Дело в том, что отдельные ошибки синтаксического анализатора поглощаются в дальнейшем при фильтрации статистическим анализатором, поскольку общее число "нормальных" языковых конструкций в тексте существенно превышает число "нестандартных". Некоторые ошибки могут оказаться заметными при анализе текстов небольшого объема, однако стоит учесть, что статистическому анализу такие тексты вообще не поддаются.

В процессе синтаксического разбора происходит построение дерева синтаксических зависимостей между словами во фразе. В случае удачного разбора предложение сворачивается в полносвязное дерево с единственной корневой вершиной.

Поскольку одна словоформа может соответствовать нескольким грамматическим формам слова, в том числе формам различных слов, в ходе анализа необходимо производить свертку предложения для всех возможных вариантов грамматических форм. Те грамматические формы, которые обеспечивают максимальную свертку дерева (минимальное число висячих вершин), следует считать наиболее достоверными.

Как показала практика, статистические методы анализа текста, на которых до настоящего времени были сконцентрированы усилия разработчиков интеллектуальных систем, достигли своего естественного предела. Дальнейшее развитие приведенных выше методов не позволит заметно повысить качество подобных систем.

Представляется необходимым использовать многоуровневое абстрагирование понятий, которые и приводит с одной стороны к компрессированию сообщений, а с другой – расширяет понятийный аппарат. Впервые этот подход предложен в [9], где, используя аналитические записи рекурсивных структур, показана возможность формализации на основе древовидных структур изменения уровня абстрагирования, т.е. изменение понятий в зависимости от контекста.

Существующие на сегодняшний день глоссарии обладают следующими недостатками:

- зашумленностью терминологического ряда несущественными для рассматриваемого предмета терминами и определениями и, как следствие, избыточностью, затрудняющей работу (некоторые глоссарии носят откровенно спекулятивный характер);
- нечеткостью и разнообразием одних и тех же определений, что влечет за собой постоянные разночтения и недоразумения.

Понимание проблемы разработки глоссариев достигла уже уровня Совета Безопасности Российской Федерации, где эта проблема рассматривалась и на секции информационной безопасности. Без глоссариев не реализовать потребность унификации множества понятий и установление связи терминологии с правовой лексикой, без чего не возникает ни право, ни ответственность, ни структура управленческих решений. Можно совершенно точно сказать, что область информационной безопасности не является чисто научно-технической, а тесно связана с юриспруденцией и правом. При этом термины, закрепленные в правовых актах, должны использоваться максимально строго.

На сегодняшний день в области информационной безопасности существуют несколько десятков глоссариев. Однако это как раз тот случай, когда количество не переходит в качество. Помимо того, что по используемым в них понятийным определениям имеется множество разночтений, вызывает много вопросов сам набор терминов, представленный в этих глоссариях. Существует также достаточное количество словарей, большая часть из которых не дает ни обоснованных определений, ни полного обзора терминов, входящих в данную область. К числу наиболее полных из них следует отнести глоссарии, разработанные Минюстом России и Гостехкомиссией России.

Вместе с тем, терминология, приведенная в данных глоссариях, ориентирована в основном на использование в органах государственного управления и возможности ее применения при проведении научных исследований весьма ограничены.

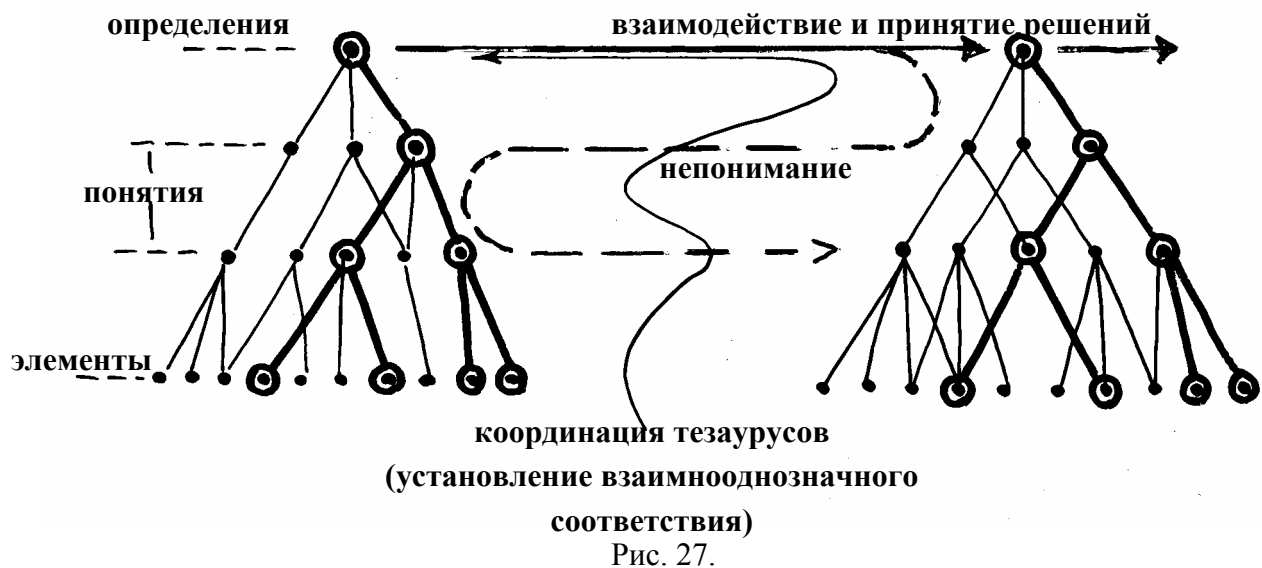
Интеграция России в мировое информационное пространство делает актуальным проведение исследований совместно с зарубежными научными организациями, работающими в области информационной безопасности, так как ни одна страна мира в условиях глобализации развития общества не может в одиночку обеспечить безопасности национальной информационной инфраструктуры. Это делает актуальным создание терминологического аппарата с учетом терминологии, сложившейся за рубежом.

Использование общепринятой терминологии при проведении научных исследований в области информационной безопасности стала настоятельной необходимостью.

В связи с бурным развитием области знаний об информационной безопасности, глоссарий должен быть оперативно и динамично обновляемым. В идеале глоссарий

должен дополняться и уточняться после выхода любого текста в данной предметной области.

Понимание проблемы информационной безопасности с учетом изложенных особенностей приводит к необходимости создания динамических развивающихся систем обеспечения информационной безопасности. Подобные системы должны обеспечивать долговременный мониторинг среды с одновременным обновлением собственных параметров (актуализацией антологии) и принятием активных действий для обеспечения безопасности в соответствии с этими параметрами.



Основы семиологической теории

А для низкой жизни были числа,
Как домашний, подъяремный скот,
Потому что все оттенки смысла
Умное число передает.
Н. Гумилев.

Коммуникация – установление однозначных связей между любым знаком и символом. В силу специфики возможностей человека с его речевой, текстуальной и визуальной формами восприятия, имеются разные формы коммуникативного общения. Интерфейсные возможности компьютера, с одной стороны, развивают коммуникативные возможности, а с другой – жестко ограничиваются по эффективности. Сегодня наиболее предпочитаемый способ интерфейсного общения между человеком и компьютером, а также между компьютерами – это текстуальная форма.

Введение понятия семиологии связано с тем, что для эффективного коммуникационного общения нет необходимости использовать все функциональные свойства лингвистики, особенно те, которые лишь усиливают интуитивно-эмоциональные свойства восприятия.

С точки зрения семиологии мы подчеркиваем, что однотипность текстуальной формы коммуникативного общения не предполагает требования полной лингвистической совместимости между человеком и компьютером.

Практика компьютерного коммуникативного общения показала: прагматика, грамматика и лексика, как составные части лингвистики, не расширяют эффективность функционирования информационных систем. Естественный язык, особенно в области инженерной и прикладной терминологии, требует постоянных усилий лишь для поддержания непротиворечивости своих определений и понятий.

Понимание важности поддержания и постоянного контроля за однозначно понимаемой терминологией уже реализуется и в учебном процессе. В 2000 и в 2004 годах школьники из Финляндии оставили позади себя ровесников из 42 стран. Феномен финского образования основан не на экзаменах и объемах заученных дат и понятий, а на умении пользоваться специально организованными текстами слов, словарями и справочниками, постоянно обновляемыми и предоставляемыми с помощью ресурсов сети Интернет. Непонимание этой концепции представления знаний привело к тому, что по международной оценки PISA каждый второй школьник России **читает лучше, чем понимает**, т.е. сильно «плавает» в функциональной грамоте: не способен оформить квитанции, заказать товар по Интернет, заполнить таможенную декларацию, составить резюме при поступлении на работу [31].

У. Эко вводит понятие семиологии как основы коммуникативного общения между компьютерными информационными системами. Это поиски вербально-лингвистического механизма для выявления и установления связей между различными словарями, синонимами и семантикой внутри заданного информационного поля.

В рамках семиологии основное внимание уделяется однозначно понимаемой терминологии: однозначности, подобия, ассоциации.

Необходимость введения понятия семиологии связано с особой ролью компьютерного коммуникативного общения. Очень многие функции лексики, прагматики и грамматики, изучаемые под понятием «лингвистика», обычно ориентируются на специфические свойства вербального мышления человека. И, следовательно, текстуальный интерфейс при своей компьютерной реализации требует введения специфических операций, которые условно можно называть семиологическими, включающими в себя такие понятия, как: антология, онтология, идентификация и ассоциация.

Длительное время существовало подспудно мнение о необходимости распространения всех функций человеческого общения на интерфейсную поддержку компьютерных информационных систем. Это связано с тем, что развитие интеллекта человека, по выражению Ницше, для выживания человечества необходимы следующие функции:

*...Тот, кто, например, не всегда умел **обнаружить одинаковое**, когда это касалось пищи или хищных зверей, т.е. тот, кто **слишком медленно** переваривал полученные новые сведения, **сопоставляя** их с уже **накопленным опытом**, тот, кто был слишком осторожен в этом сопоставлении, – тот имел **меньше шансов** на выживание, **чем тот**, кто, заметив малейшее сходство, **угадывал в нем одинаковость**....*

В принципе, формализация компьютерных информационных систем опирается на полный перебор объектов и их свойств – по их текстуальному представлению, что требует экспоненциального роста количества операций сравнений. Любое же сокращение количества операций связано с теми или иными ограничениями, вводимыми на решаемую в данный момент задачу в терминах данной информационной системы.

Следует понимать, что с этой целью заранее создается схема **данных** и форматы электронных документов. По мере роста объема данных и изменения их типов, а также коррекции информационных задач, постоянно возникает необходимость в обновляемой логистике – данных и области их ограничения.

Неудачное, но модное сочетание «логистика знаний» в принципе предназначено для выявления специфических характеристик при построении той или иной инструментальной информационной системы, например, инженерное представление чертежной документации в AutoCAD или пространственный дизайн в среде 3D MAX. Универсальность и специализированность системы должна отражать уровень и специфику применимости той или иной инструментальной системы, выступая дополнительным ограничением в информационном поле. Так, инструментальная среда для электронного правительства не эффективна для производственных систем локального масштаба. Приведем следующий пример. Попытка использовать каналы Санкт-Петербурга как транспортные артерии не выдерживает никакой критики, так как их **логистика** никак не соотносится с **топографией** расположения государственных и массовых учреждений, предприятий, музеев, ресторанов и т.д.

По сути, логистика в информационных системах – это выбор границ представления проблемно-ориентированных знаний, то есть формирование баз данных как множеств A и B в принципе идентификации неразличимости по Лейбницу.

Кратко повторим рассмотренные выше положения семиологической теории, рассмотрев их в составе единой теории:

1. Принцип идентификации неразличимости Лейбница.

Г. Лейбниц предполагал выразить все содержание человеческого знания системой символов (“характеров”) – своеобразной “азбукой человеческих мыслей” – чтобы заменить рассуждения вычислениями. Таким образом, язык становился бы алгеброй мысли. Лейбниц считал, что *«никогда не кончаются споры и не установится мир в борьбе школ, пока от путанных рассуждений, неясных слов и неопределенных значений мы не перейдем к простым исчислениям и определенным характеристикам»*. По его мнению, универсальный язык, в том виде как его предполагалось создать, есть *“самое возвышенное и самое экономичное употребление человеческого разума с помощью символов и знаков”*.

С этой целью он и предложил «универсальную характеристику» – конгруэнтность $ABC \sim AXC$ способ идентификации и переопределения понятий – основу функционирования современных компьютерных технологий.

На рис. 1 приведен принцип идентификации неразличимости Лейбница, согласно которому известные объекты из множеств A и C и их свойства связаны следующим отношением ABC . Операция конгруэнтности $ABC \sim AXC$ позволяет интерпретировать X не как отдельный конкретный объект, а как множество объектов из A и C , для которых справедливо заданное отношение B .

В [32] как пример показано, что наиболее простая структура восприятия и воспроизводства развивающихся понятий, которые создаются и воспринимаются с раннего возраста, иллюстрируются на примере рекурсивных сказок.

Рекурсивные сказки – повествование, в которых приходится использовать предшествующие элементы данной последовательности.

Такие сказки могут иметь структуру, которую можно описать с помощью формул: $a + b + c + \dots$ и $a + (a + b) + (a + b + c) \dots$. Не требуется большой наблюдательности, чтобы заметить, что вторая формула является разновидностью первой. Рекурсивные сказки моделируются с помощью самоподобных рекурсивных структур, лишь в последнее время ставших предметом пристального интереса математиков и обнаруживших большое количество полезных свойств и применений в самых различных областях, имеющих дело со сложными развивающимися объектами. Геометрически самоподобные рекурсивные структуры представляются в виде графов – деревьев зависимостей [9].

Из рассмотренных свойств самоподобного развития следуют и найденные эмпирические рейтинговые распределения слов в текстах (законы Ципфа, Мандельброта), что отражает общий концептуальный подход к информационному общению, сказать или написать как можно меньше слов, знаков, символов и быть понятым. Тойнби [33] этот принцип назвал принципом этерификации.

Мы подчеркиваем, что определение конгруэнтности устанавливает однозначное перечисление объектов из A и C по заданному отношению. В этом случае объект X является размножающейся сущностью, в заданных границах знаний, представленных множествами A и B .

Понятие конгруэнтности в современных системах управления базами данных превратилось в простую выборку всех элементов X по условиям совпадения ключа B , где A и C ограниченность словаря (граничные условия множества).

Здесь следует подчеркнуть что слово «число» используется как понятие, обозначающее упорядоченный уникальный идентификатор, на что и адаптирована аппаратно-программная среда компьютеров.

2. Размножение сущностей Фреге

С точки зрения лингвистических основ определений и переопределений мы приводим подход из книги Ж. Делеза [7], состоящий в том, что *“...должно существовать нечто безусловное, способное обеспечить реальный генезис денотации и других отношений предложения. Тогда условие истины можно было бы определять уже не как форму концептуальной возможности, а как некоторую идеальную материю, то есть не как сигнификацию, а как смысл...”*

Тогда если принять предложение за некое имя, то ясно, что каждое имя, обозначающее объект, само может стать объектом нового имени, обозначающего его смысл: $n1$ отсылает к $n2$, которое обозначает смысл $n1$; $n2$ отсылает к $n3$ и так далее. Для каждого из своих имен язык должен содержать некоторое имя как идентификатор следующего уровня понятийного смысла. Такое бесконечное размножение вербальных сущностей известно как парадокс Фреге.

Во всяком развитии, будь то рост живого организма, постепенная эволюция технической системы или возникновение новых биологических видов можно различить

три составляющие: 1) наследственность – предопределенность предысторией развивающейся системы, 2) влияние условий среды, 3) чистое творчество. Для языковых конструкций наследственностью является ограничения языка – алфавит, синтаксис, языковой базис, влияние условий среды. Творчеством является множество понятий, порождаемых в заданных условиях и допустимых в данной конструкции для обозначения требуемой ситуации.

В варианте смысл-объект этот парадокс, который мы будем называть все-таки парадоксом Фреге, раскрывает “место поиска смысла”, указывает момент возникновения того, что позже стали называть контекстной смысловой зависимостью. Парадокс Фреге используется нами как идейная основа развивающихся структур связи данных в проблемно-ориентированных средах, где эти связи, в отличие от произвольных сред, имеют выраженный характер через ассоциативно организованную идентификацию и допускают проведение событийной, то есть смысловой, структуризации.

3. Принцип этерификации Тойнби

Опыт, накопленный в результате попыток формализации многих предметных областей в области их терминологии, а также построение естественно-языковых интерфейсов показал, что возникающие при этом закономерности типа закона Ципфа удачно совпадают с предпочтениями человека по ограничению размножающихся сущностей. Другими словами, существуют механизмы, основанные на параметрах типа частоты использования некоторого термина в тексте или его постоянного использования совместно с другим термином, позволяющие производить переупорядочение связей и выделение иерархических уровней.

Характеристики параметров предполагают возможность их изменения при развитии предметных областей. Это означает перераспределение отношений и создание новых семантических уровней при введении новых понятий или изменении смысла уже имеющихся терминов.

Впервые этерификация как принцип, который обеспечивает коммуникационный прогресс, сформулирован Дж. Тойнби [33] *...В истории письменности наблюдается не только соответствие между развитием техники письма и упрощением формы, но и эти две тенденции фактически тождественны друг другу, поскольку вся техническая проблема, которую должно решить письмо как фиксатор, посредник человеческой речи, – это отчетливая репрезентация широчайшей сферы человеческого языка с максимальной экономией визуальных символов, т.е. этерификация – есть закон прогрессирующего упрощения....*

Отметим, что наиболее эксплуатируется и понятен принцип построения иерархии понятий через словари, энциклопедии и глоссарии к информационно-аналитическому реферированию. Возникают проблемы распознавания нового или узнавания известного.

Распознавание есть интерпретация смысла, отражаемого в семантике входных сигналов как предварительная стадия понимания. Одномоментное индуктивно-интуитивное восприятие визуальной информации конфликтует с последовательным предикатным описанием и дедуктивным выводом.

Принцип этерификации позволяет выявлять области действия предикативных отношений внутри информационных полей, чтобы не допускать бесконечное разрастание сущностей.

Сформированное информационное поле позволяет ограничивать семантику используемых терминов, что устраняет неоднозначное исполнение требований, команд и распоряжений на уровне исполнителей. В примере с военной системой это означает ограничение полномочий исполнителей всех уровней рамками выполняемой задачи и выдачу четких распоряжений вида «взять изделие X и положить его в Y», которые исключают многозначные трактовки исполнителем.

Неоднозначность и сложность предикатного описания поясним следующей выдержкой из [7]:

...Зеленеть указывает на сингулярность-событие, в окрестности которого задается дерево, а грешить – на сингулярность-событие, в окрестности которого образуется Адам. Но быть зеленым или быть грешным – теперь это уже аналитические предикаты образовавшихся субъектов, а именно, дерева и Адама...

В отличие от этого «прагматического» значения мы имеем в виду упорядочение, происходящее из основного содержания объекта. Например, из значения слова «шесть» следует, что данному объекту предшествует объект с именем «три» (для линейных систем), а из самого значения слова «отец» следует последующее положение объекта «сын» (для иерархических систем). Разумеется, можно было бы придумать примеры конкретных ситуаций, где эти элементы должны быть переставлены, но в этом случае мы будем использовать слова лишь в качестве временных меток, не используя вложенный в них смысл. Провести такого рода упорядочение (обычно частичное) можно и не над числами. Трудность состоит в том, что такие аудиовизуальные среды, как источники информации, адекватно воспринимаются участниками общения только за счет либо формирования *естественных образно-зрительных идентификаторов*, от луны, солнца, пейзажей к пиктограммам, иероглифам и т.д., либо общей психофизической реакции на звукоголосовое восприятие.

В зависимости от семантики информационного поля предикат «зеленеть» приведет нас либо на зеленую лужайку, либо к понятию «грех» и «Адама».

Подчеркнем, что семиологический подход расширяет область применения информационных систем до восприятия (инициализации) семантико-смыслового содержания. При этом информация понимается как возможность целенаправленного управления (выбора) на основе полученного сообщения и **ставшего означающим** (семантико-смысловым компонентом), когда оно получает истолкование на основе того или иного лексикода и, следовательно, зависит **от адресата**.

Эта информационно-смысловая компонента, инициализируемая сообщением, отличается от количественно исчисляемой информации источника, имеющей дело с физическим каналом, и называется семиологической информацией. Она не исчисляется с помощью количественных методов, но определима через ряд ассоциативных определений и значений, которые могут возникнуть под воздействием разных кодов информационных полей.

4. Упорядочение Кантора

Основой семиологической теории являются объекты, связи между объектами и описания объектов. Необходимым свойством объектов должна быть возможность идентификации объектов, то есть возможность отличить один объект от остальных. Связи задают отношения между элементами, как между уровнями, так и внутри уровня. Семиологический подход связан с выявлением вербально-лингвистического установления и механизма связей как внутри уровня, так и между уровнями. Какие операции с объектами как с идентификаторами составляют предметное выявление семантики и смысловых понятий, и составляет основу семиологической теории. Существенно при этом, что каждый выводимый объект конечен, точнее, содержит в себе конечное количество информации. Можно сказать, что речь всегда идет об объектах, которые могут быть закодированы (без потери «информации», что бы ни значило это слово) конечной последовательностью символов (букв), каждый из которых принадлежит заранее указанному множеству (алфавиту). Существенно еще, что каждая буква сама по себе не несет никакой другой «информации», кроме возможности отличить ее от других букв алфавита, и при сохранении этой возможности буква может быть как угодно заменена и переобозначена.

В варианте смысл-объект этот парадокс, который мы будем называть все-таки парадоксом Фреге, раскрывает “место поиска смысла”, указывает момент возникновения того, что позже стали называть контекстной смысловой зависимостью. Парадокс Фреге используется нами как идейная основа развивающихся структур связи данных в проблемно-ориентированных средах, где эти связи, в отличие от произвольных сред, имеют выраженный характер через ассоциативно организованную идентификацию и допускают проведение событийной, то есть смысловой, структуризации.

...Обобщающие родовые понятия последовательно развиваются по цепочке, т.е. вытекают друг из друга...” – в этом открытии Гегеля закреплялось основополагающее понятие Развитие: “...Мы инстинктивно придаем более глубокий смысл и большую значимость Развитию, а не тому, что есть в настоящий момент...”

Важную часть семиологической теории базируется на исследованиях алгебраистов С. Ю. Маслова и Ю.И. Манина, которые предложили гипотезу финитности, т.е. возможность адекватного описания состояния развивающейся системы, содержащего лишь конечную информацию. Реально мы будем говорить о состояниях, задаваемых как конечные множества (возможно, с повторениями) слов в конечном алфавите.

На практике часто встречаются системы, которые можно описать следующим образом: имеется некоторое количество исходных объектов и некоторое количество правил построения новых объектов из исходных и уже построенных. Или так: имеется начальная позиция (начальное состояние) и заданы «правила игры» (правила перехода от состояний к состояниям).

Суть работы информационных систем оказывается процесс перехода от одного исчисления (описания) к другому.

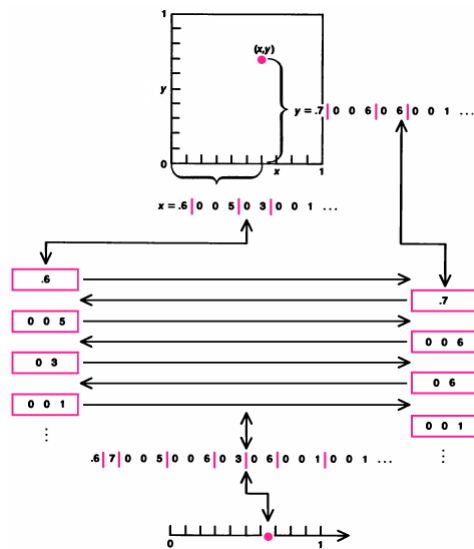


Рис. 29. Взаимнооднозначное соответствие между точками плоскости и точками прямой и способ его построения, предложенный Г. Кантором.

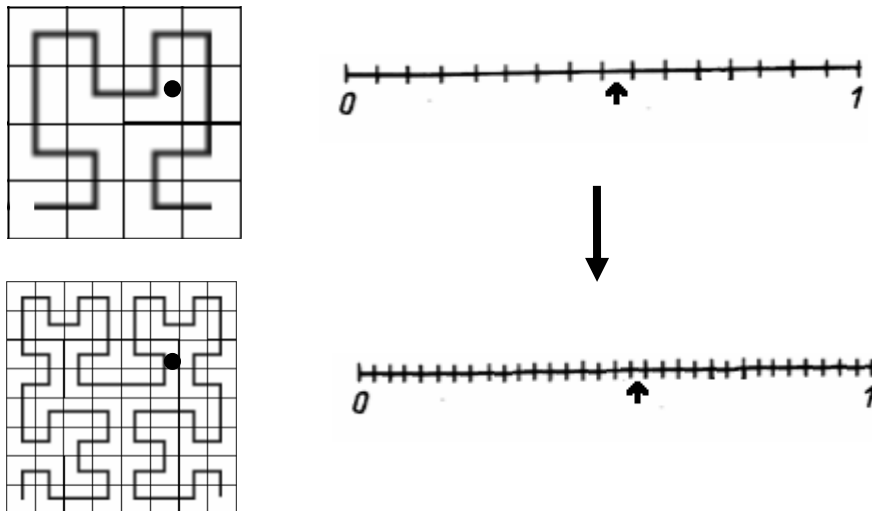


Рис. 30. Взаимнооднозначное соответствие между точками плоскости и точками прямой и способ его построения с помощью заполняющей пространство кривой (ЗПК). На каждом шаге происходит последовательное уточнение положения точки.

Информационные системы рассматривают соответствия между объектами и их свойствами. Математические же модели этих систем используют понятие точек (пример на рис. 29, 30).

Рассмотрим пример применения семиологической системы в области лингвистики.

Для любого алфавита A множество слов в A перечислимо тогда и только тогда, когда перечислимо множество кодов этих слов. Для задания любого перечислимого множества достаточно, по существу, алфавита $\{0, 1\}$.

Так, реализация двухбуквенного алфавита в компьютере есть двоичные числа. Двоичные коды формируют символы, образующие нижний уровень иерархии. На следующем уровне символы формируют слова, а слова формируют предложения. При этом формальное представление понятийных элементов предложения сохраняется в виде двоичных последовательностей.

Семиологический подход обязан формировать как прямой механизм формирования элементов (из формального битового алфавита понятийных конструкций), так и обратный механизм декомпозиции, позволяющий работать на уровне идентификаторов объектов. Такой механизм должен включать операции идентификации, упорядочения и сравнения элементов. В данном случае можно рассмотреть примеры связей, формируемых между словами, являющимися синонимами в данном языке, или между словами, создающими понятийные окружения других слов, а также ассоциативные цепи.

Семиология и лингвистика

Тенденция построения информационных коммуникативных систем проделала путь от баз данных до логистических и CALS-систем. Этот путь связан с попытками организовать взаимодействие человека с компьютерными системами различного рода управления, навигации, контроля ситуации, мониторинга среды на привычной основе естественно-языкового общения.

Однако, несмотря на достигнутые успехи, все еще происходит подмена таких понятий, как аналитическое реферирование и ассоциативный поиск информации человеком, поисковыми серверами компьютерных систем. Есть принципиально вычислимые и невычислимые конфигурации лингвистических конструкций. Именно понимание этого факта с точки зрения построения эффективных коммуникационных функций компьютера и привело к новому классу информационных систем – семиологическим системам, как приближение к недостижимому пока лингвистико-вербальному механизму мышлению человека.

Семиологическая теория, включающая вышеперечисленные компоненты, не противоречит основам функционирования компьютерных систем. Частичное присутствие некоторых компонентов дает различные виды информационных систем, реализация же всех компонентов дает семиологическую информационную систему. Важно отметить, что информационные системы, а в том числе семиологическая информационная система, не дает знаний человеку. И все же, приведенные положения семиологического подхода позволяют автоматизировать процесс построения референтных аналитических систем. Главной, хотя и не единственной их целью, является переупорядочение данных и представление их в удобной для восприятия человека форме. Интерпретирование же данных и накопление знаний целиком остается прерогативой человеческого мышления.

Литература.

1. Папюс. Каббала. — СПб., 1910 — 360 с.
2. Еремеев В. Е. Чертеж антропокосмоса. — АСМ, Москва, 1993.
3. У. Эко. Пять эссе на темы этики. — СПб.: "Симпозиум", 2003, 158 с.
4. Г. Лейбниц. Новые опыты о человеческом разумении автора системы предустановленной гармонии. — Сочинения в 4-х томах, т.2, М.: Мысль, 1983, 686 с.
5. Александров В.В. Интеллект и компьютер. — СПб.: Издательство «Анатолия»., 2004. — 285 с.

6. Античные теории языка и стиля. Под общей редакцией О. М. Фрейденберг. — ОГИЗ, Государственное социально-экономическое издательство. Москва-Ленинград, 1936 -343 с.
7. Делез Ж. Логика смысла. — М., Издательский Центр «Академия», 1995.
8. Визуальный словарь. On-line версия. — <http://www.vslovar.org.ru>
9. В. В. Александров, А. В. Арсентьева. Структурный анализ диалога. — Л.: ЛНИВЦ АН СССР, 1983.
10. У. Эко. Отсутствующая структура. Введение в семиологию. — СПб.: "Симпозиум", 2004.
11. Яндекс - энциклопедия. — <http://encycl.yandex.ru/>
12. Даль В. Толковый словарь живого великорусского языка. — М.: ГИИНС, 1955.
13. Webster's II. Berkley ed., 1984
14. Словарь иностранных слов. — М.: ГИИНС, 1949.
15. В. Володин. CALS: технология или система. — Вестник воздушного флота. № 03-04, 2003 г., с. 52-55.
16. А.Е. Городецкий, В.В. Дубаренко, И.Л.Тарасова, А.В. Шереверов. Программные средства интеллектуальных систем. СПб, издательство СПбГТУ, 2000 г.
17. Александров В.В., Булкин Г.А., Поляков А.О. Автоматизированная обработка информации на языке предикатов. — М.: Наука, 1982, 105 с.
18. А.О. Поляков, Ю.М. Смирнов, А.А.Турчак. Информодинамические основы организации управления предприятиями и холдинговыми компаниями. СПб, издательство СПбГПУ, 2002 г.
19. Д.В. Гаскаров. Интеллектуальные информационные системы. Москва, "Высшая школа", 2003 г.
20. А.В. Алексеев, А.Н. Борисов, Э.Р.Вилюмс, Н.Н. Слядзь, С.А.Фомин. Интеллектуальные системы принятия решений. Рига, издательство "Зинатие", 1997 г.
21. Успенский П. Д. Новая модель вселенной. — С.-Пб., 1994.
22. Колмогоров А.Н. Теория информатизации и теория алгоритмов. — М.: Наука, 1987 - 304 с.
23. XXI век. Еще раз о русском воззрении. — <http://piramyd.express.ru/disput/anonsens/anonsens.htm>
24. Оруэлл Дж. "1984" и эссе разных лет. — М.: Прогресс, 1989.
25. Visual Thesaurus. On-line версия. — <http://www.visualthesaurus.com>
26. Яблонский С., Сухоногов А. Проект русской версии WordNet. — Петербургский ун-т путей сообщения. — <http://www.pgups.ru/WebWN/wordnet.uix>
27. Информационный портал RPM-технологий — <http://www.rpm-novation.com/>
28. Model Integrated Computing and Autonomous Negotiating Teams for Autonomic Logistics — <http://www.isis.vanderbilt.edu/Projects/micants/Tech/Briefings/pi/9911/>
29. Александров В. В., Арсентьева А. В. Информация и развивающиеся структуры. — Л.: ЛНИВЦ АН СССР, 1984.
30. Ермаков А. Е., Плешко В. В. Синтаксический разбор в системах статистического анализа текста. — Информационные технологии. — № 7, 2002.
31. «Почему Россия нефинляндия» Новая Газета. № 36, 23.05-25.05.2005

32. Крестов А. А. Рекурсивные сказки. — Воронежский университет.
33. Тойнби А. Дж. Постигание истории. — Пер. с англ./Сост. Огурцов А.П.; Вступ. ст. Уколовой В. И.; Закл. ст. Рашковского Е. Б. — М.: Прогресс. 1996 — 608 с.
34. А. Ф. Колчин, М. В. Овсянников, А. Ф. Стрекалов, С. В. Сумароков. Управление жизненным циклом продукции. — М.: Анахарсис, 2002. — 304 с.
35. Электронное правительство: рекомендации по внедрению в Российской Федерации. Под ред. В. И. Дрожжинова, Е. З. Зиндера. — М.: Эко-Трендз, 2004. — 352 с.
36. В. В. Александров, С. В. Кулешов. Метоглоссарии — основа аналитических информационных систем. — Проблемы транспорта № 10 — СПб., 2004, с. 248.
37. Лексическая база знаний английского языка WordNet — <http://www.cogsci.princeton.edu/~wn>

Краткий словарь терминов

ASP (Application Service Providers) предоставление услуг доступа к приложениям, расположенным на удаленном от потребителя сервере, через Интернет или частную сеть.

B2B (Business-to-Business) схема "бизнес — бизнес", включает в себя все уровни электронного взаимодействия на уровне компаний с использованием специальных технологий и стандартов электронного обмена данными.

B2C (Business-to-Consumer) схема "бизнес — потребитель": розничная продажа товаров и услуг разным лицам через Интернет.

B2G/B2A (Business-Administration) схема "бизнес — администрация (правительство)". Схема подразумевает взаимодействие компании и госадминистрации страны, на территории которой она работает.

C2G/C2A (Business-Government) схема "потребитель — администрация (правительство)". Взаимодействие правительства и потребителя.

CAD (Computer-Aided Design) — проектирование с помощью ЭВМ, применение информационной технологии к элементам процесса проектирования производимых, собираемых и конструируемых продуктов как в области черчения (для создания, изменения, хранения и вывода инженерных и прочих технических чертежей), так и в области моделирования (для генерирования и использования цельных трехмерных моделей).

CAE (Computer-Aided Engineering) разработка с помощью ЭВМ, применение информационной технологии к элементам процесса проектирования и разработки. Включает в себя все типы функциональных систем, напр. анализ теплообмена, структурный, электромагнитный, воздушный и звуковой анализ

CALS (Continuous Acquisition and Life-cycle Support) — непрерывное развитие и информационная поддержка жизненного цикла изделий. Это глобальная стратегия, направленная на более тесную интеграцию предприятий с помощью упрощения бизнес-процессов и применения стандартов и технологий к разработке, управлению, обмену и использованию деловой и технической информации (ИПИ).

CAM (Computer-Aided Manufacturing) – производство с помощью ЭВМ, применение информационной технологии к контролю и управлению процессами производства, обычно ограничиваемое контролем таких машинных инструментов, как токарные и фрезерные станки, когда инструмент контролируется непосредственно компьютером

CIM (Computer-Integrated Manufacturing) – интегрированное автоматизированное производство. Применение информационной технологии к управлению законченными системами или подсистемами на производстве, для которого характерна интеграция многих отдельных приложений, таких как CAD, CAM.

CM (Corporate Memory) – корпоративная память, подразумевает идентификацию и анализ доступных и необходимых знаний, а также вытекающее из такого анализа планирование и управление действиями по созданию интеллектуального капитала, обеспечивающего достижение целей организации

CRM (Customer Relations Management) – системы поддержания отношений с клиентами

DBMS (Data Base Management System) – системы управления базами данных

E-business – электронный бизнес. Это преобразование основных бизнес процессов компании путем внедрения Интернет-технологий, нацеленное на повышение эффективности деятельности

E-commerce – электронная коммерция. Это любые формы бизнес-деятельности, использующие возможности информационных сетей, – маркетинг, продажи товаров, аренда приложений, предоставление услуг или информатизации.

EDM (Electronic Document Management) – системы управления электронным архивом.

ERP (Enterprise Resource Planning) – системы управления ресурсами масштаба предприятия

MIS (Management Information Systems) – информационная система для менеджеров.

OLAP (On-line Analytical Processing) – системы многомерного анализа массивов данных информационной системы

P2P (Peer-to-Peer или Partner-to-Partner) – схема "равный-равный". Бизнес отношения между партнерами, находящимися в равном положении в Интернет

PDM (Product Data Management) – системы управления данными об изделиях (системы управления проектом)

PLM (Product Lifecycle Management) – идеология непрерывного управления данными изделия на протяжении всего жизненного цикла.

PMS (Project Management System) – системы управления проектами

SCP (Supply Chain Planning System) – системы управления «логистическими цепочками» – (планирование и оптимизация схем поставок для улучшения показателей предприятия, а также контроль за происхождением всех его компонент (составных частей) и их «пути движения» (реализуется в PDM).

Work Flow System – системы управления деловыми процессами.

Антология – anthologia <гр. anthos – цветок + lego – собираю> – сборник различных статей.

Глоссарий – это понятийно-терминологические определения слов, используемых в тезаурусе конкретной предметной области.

Данные – документированная информация, организованная таким образом, что обеспечивается возможность быстрого поиска документов.

Документ – используется как унифицированный формат (в рамках данного корпоративного стандарта) электронного представления данных – упорядоченный текст, графика, видео, чертеж и др.

Документированная информация (документ) – зафиксированная на материальном носителе информация с реквизитами, позволяющими ее идентифицировать.

ЕИП – единое информационное пространство.

Инкапсуляция – термин объектно-ориентированного программирования, означающий структурирование программы на модули особого вида, называемые классами и объединяющие данные и процедуры их обработки, причем внутренние данные класса не могут быть обработаны иначе, кроме как предусмотренными для этого процедурами. Реализация класса как бы заключена в капсулу и скрыта, откуда и происходит название И.

Информационная система – организационно упорядоченная совокупность документов (массив документов и информационных технологий, в том числе и с использованием средств вычислительной техники и связи).

Информационные процессы – процессы сбора, обработки, накопления, хранения, поиска и распространения информации.

Информация – сведения о лицах, предметах фактах, событиях, явлениях и процессах независимо от формы их представления.

КИП – компактное интеллектуальное производство

КИС – корпоративная информационная система

Лингвистика – Linguistics – the study of the nature and structure of human speech – учение о природе и структуре человеческой речи.

Лингвистика (фр. Linguistique <лат. lingua – язык>) – языковедение, наука о языке, включающая фонетику, семантику, лексикологию, этимологию, грамматику и орфографию, орфоэпию.

Логистика – Logistics – the methods of procuring distributing, maintaining, and replacing materiel and personnel, as in a military operation. – методы доставки (доставания, сводничества), распределения, поддержки и размещения материалов (продуктов технологической деятельности).

Онтология – onthologic <гр. on (ontos) – сущее + logos – понятие, учение> – греч. учение о существе или о сущности, бытии, сути.

САПР – система автоматизации и проектирования

Семантика – Semantics – изучение смысла в языке.

Семиотика – Semiotics – наука о знаках, включая знаки–символы и знаки–сигналы.

Референтная группа – на мнение которой ориентируется большинство.

Референтность – основное свойство глоссария, наиболее употребляемое определение, понятие. Глоссарий основывается на выявлении референтных понятий.

Тезаурус – отобранные слова или понятия, выполняет роль специализированного словаря, заданной предметной области, например медицина, музыка.

ТП – технологический процесс

ТПП – техническая подготовка производства

Приложение 1.

Табл. 1. Основные типы информационных систем и их возможности

Система	Функции			
	Поиск по ключу	«Интеллектуальный» поиск	Иерархическое представление структуры	Интерфейс на естественном языке
DBMS (СУБД)	+	±	-	-
Поисковые системы	+	+	±	±
CAD	+	-	+	-
CALS	+	+	?	±
ERP	+	?	+	-
Системы поддержки принятия решений	+	±	+	±
Экспертные системы	±	±	±	±
Онтологические системы	+	+	+	±

Обозначения:

+ функция реализована в необходимом объеме

- функция не реализована

± функция заявлена, но реализована лишь частично

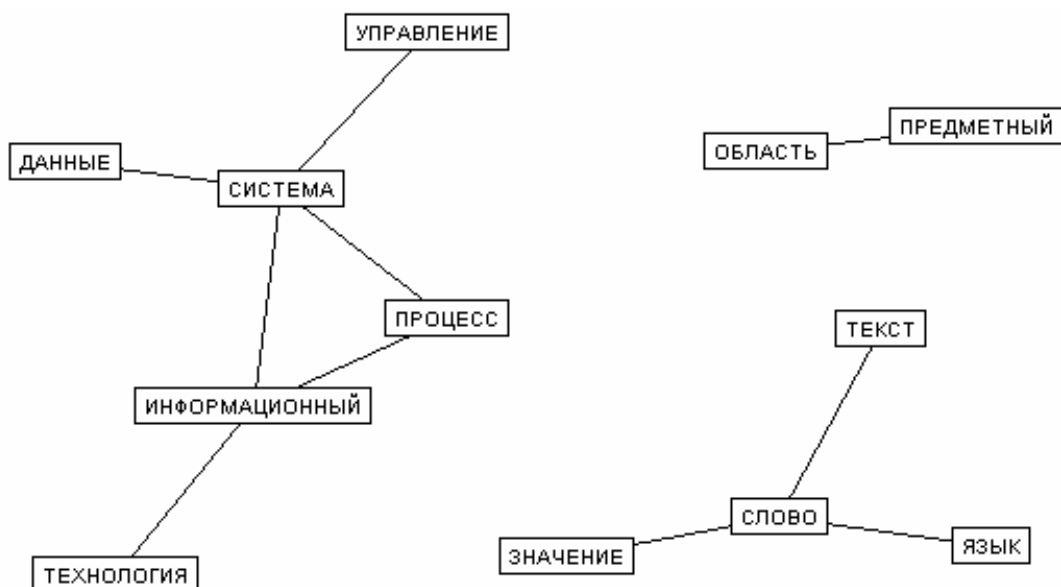
? сравнение не применимо к системе, либо возможность сильно зависит от реализации и версии системы.

Табл. 2. Основные типы информационных систем по типам данных

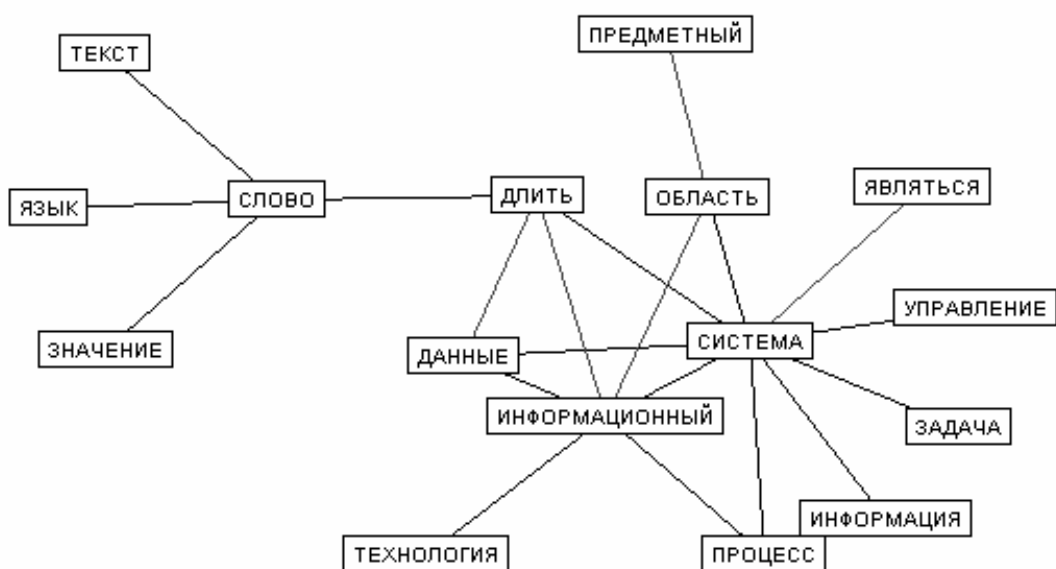
Типы информационных систем				
	СУБД	CAD	Поисковые системы	Онтологические информационные системы
Типы поддерживаемых данных				Неструктурированные данные
			Данные на естественном языке	
		Иерархические данные		
	Табличные данные			

Приложение 2.

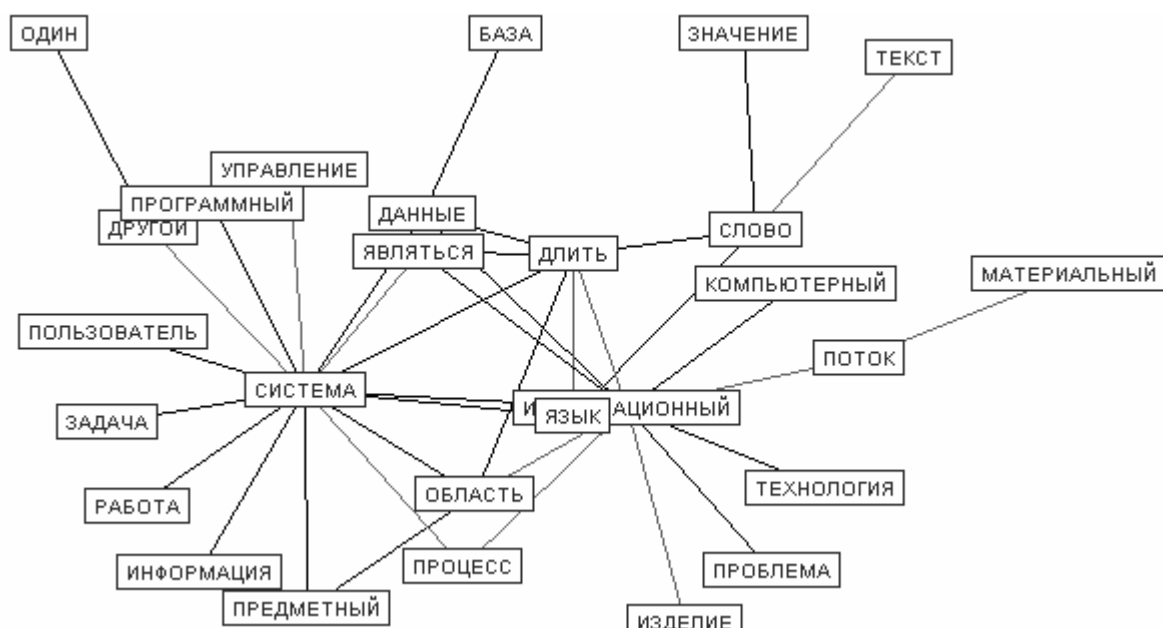
Рассмотрим применение рассмотренного в данной книге метода (реализованного в экспериментальной системе VISGLOSS) после обработки текста самой этой книги.



10 фундаментальных понятийных связей терминов



20 фундаментальных понятийных связей терминов



40 фундаментальных понятийных связей терминов