

Министерство образования и науки Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Д. В. Кудрявцев

**СИСТЕМЫ УПРАВЛЕНИЯ
ЗНАНИЯМИ И ПРИМЕНЕНИЕ
ОНТОЛОГИЙ**

Учебное пособие

Министерство образования и науки Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Д. В. КУДРЯВЦЕВ

**СИСТЕМЫ УПРАВЛЕНИЯ
ЗНАНИЯМИ И ПРИМЕНЕНИЕ
ОНТОЛОГИЙ**

*Рекомендовано Санкт-Петербургским государственным
политехническим университетом в качестве учебного пособия
для студентов высших учебных заведений, обучающихся
по направлению подготовки магистров «Информатика
и вычислительная техника»*

Санкт-Петербург
Издательство Политехнического университета
2010

УДК 004.89 (075.8)

К 88

Рецензенты:

Доктор технических наук, профессор Высшей школы менеджмента
Санкт-Петербургского государственного университета Т.А. Гаврилова

Кандидат технических наук, профессор Санкт-Петербургского
государственного политехнического университета Л.А. Станкевич

Кудрявцев Д.В. **Системы управления знаниями и применение онтологий**: учеб. пособие / Д.В. Кудрявцев. – СПб.: Изд-во Политехн. ун-та, 2010. – 344 с.

Последовательно рассмотрены понятия *знания, управления знаниями, системы управления знаниями*. Представлен обзор современных методологий управления знаниями. Предложено применение онтологий и семантических технологий в системах управления знаниями. Описан процесс управления знаниями в организации. Дано описание задач, методов и программных средств для повышения эффективности процессов создания, накопления, распределения и использования знаний в организации. Рассмотрены вопросы структурирования и представления знаний, работы с метаданными, поиска информации, а также обмена неявными знаниями на предприятии. Систематизированы программные средства для работы со знаниями. Приведены примеры систем управления знаниями, а также указаны направления современных исследований в области представления и управления знаниями. Особое внимание уделено применению онтологий в системах управления знаниями.

Предназначено для студентов вузов, обучающихся по направлению подготовки магистров «Информатика и вычислительная техника». Может быть также использовано при обучении (в системах повышения квалификации, в учреждениях дополнительного профессионального образования и пр.) специалистов в области информационных технологий и управления.

Печатается по решению редакционно-издательского совета
Санкт-Петербургского государственного политехнического университета.

© Кудрявцев Д. В., 2010

© Санкт-Петербургский государственный
политехнический университет, 2010

ISBN 978-5-7422-2982-7

СОДЕРЖАНИЕ

| | |
|--|-----|
| СОКРАЩЕНИЯ, ПРИНЯТЫЕ В ПОСОБИИ..... | 5 |
| ВВЕДЕНИЕ..... | 6 |
| 1. ЗНАНИЕ И УПРАВЛЕНИЕ ЗНАНИЯМИ..... | 10 |
| 1.1. Определение понятия знание..... | 10 |
| 1.2. Источники и виды знаний..... | 13 |
| 1.3. Знания и интеллектуальный капитал..... | 17 |
| 1.4. Определения и назначение управления знаниями..... | 19 |
| 1.5. История возникновения и подходы к управлению знаниями..... | 20 |
| 1.6. Обзор задач управления знаниями..... | 24 |
| 1.7. Междисциплинарность управления знаниями..... | 25 |
| 2. ОБЗОР МЕТОДОЛОГИЙ УПРАВЛЕНИЯ ЗНАНИЯМИ..... | 27 |
| 2.1. Методология Know-Net..... | 27 |
| 2.2. Создание знаний в подходе Нонака и Такеучи..... | 31 |
| 2.3. Методология CommonKADS..... | 35 |
| 2.4. Методология DECOR..... | 40 |
| 2.5. Методология Карла Виига..... | 42 |
| 2.6. Методология On-To-Knowledge..... | 46 |
| 3. СЕМАНТИЧЕСКИЙ ПОДХОД К УПРАВЛЕНИЮ ЗНАНИЯМИ И ОНТОЛОГИИ..... | 50 |
| 3.1. Потребность в общем языке..... | 50 |
| 3.2. Потребность в интеграции информации..... | 55 |
| 3.3. Онтология как общий язык в управлении знаниями..... | 56 |
| 3.4. Примеры онтологий..... | 61 |
| 3.5. Применение онтологий..... | 69 |
| 3.6. Согласование академической и бизнес- терминологий..... | 79 |
| 4. СТРУКТУРА И РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ..... | 80 |
| 4.1. Два уровня деятельности: работа со знаниями и управление..... | 80 |
| 4.2. Уровень знаний и объекты управления..... | 81 |
| 4.3. Процессы преобразования знаний..... | 86 |
| 4.4. Неформальные сети знаний..... | 89 |
| 4.5. Инструменты управления знаниями..... | 92 |
| 4.6. Процесс управления знаниями..... | 94 |
| 4.7. Определение понятия «система управления знаниями»..... | 95 |
| 4.8. Структура проекта разработки и внедрения системы управления знаниями..... | 96 |
| 5. ПРОЦЕСС УПРАВЛЕНИЯ ЗНАНИЯМИ..... | 99 |
| 5.1. Диагностика и анализ..... | 99 |
| 5.2. Проектирование системы управления знаниями..... | 115 |
| 5.3. Планирование внедрения системы управления знаниями..... | 118 |
| 5.4. Контроль и оценка..... | 121 |
| 6. СОЗДАНИЕ ФОРМАЛИЗОВАННЫХ ЗНАНИЙ..... | 129 |
| 6.1. Структурированный ввод информации..... | 130 |
| 6.2. Совместное создание контента и использование вики..... | 132 |
| 6.3. Извлечение знаний..... | 138 |
| 6.4. Распознавание речи..... | 141 |
| 6.5. Специализированные службы по созданию знаний..... | 142 |
| 6.6. Выявление новых знаний в базах данных..... | 143 |

| | |
|---|-----|
| 7. НАКОПЛЕНИЕ ЗНАНИЙ | 152 |
| 7.1. Структурирование знаний | 152 |
| 7.2. Метаданные | 161 |
| 7.3. Виды информационно-поисковых языков | 166 |
| 7.4. Разработка информационно-поискового языка на предприятии..... | 180 |
| 7.5. Использование онтологий и языков семантического веба..... | 186 |
| 7.6. Индексирование/Аннотирование ресурсов (создание метаданных)..... | 198 |
| 7.7. Референтные модели, справочники, типовые элементы | 204 |
| 8. РАСПРЕДЕЛЕНИЕ ЗНАНИЙ..... | 206 |
| 8.1. Тянущий (поиск) и толкающий (доставка) подходы в распределении знаний... 206 | |
| 8.2. Оценка качества поиска и доставки знаний | 213 |
| 8.3. Применение онтологий для поиска информации..... | 216 |
| 8.4. Формирование пакетов знаний | 222 |
| 8.5. Использование контекста для повышения качества поиска | 224 |
| 8.6. Визуализация знаний на основе онтологий | 228 |
| 8.7. Настройка прав доступа..... | 232 |
| 9. ОБМЕН НЕЯВНЫМИ ЗНАНИЯМИ И ПОДДЕРЖКА СЕТЕЙ ЗНАНИЙ | 233 |
| 9.1. Средства совместной работы | 233 |
| 9.2. Сети знаний индивидуального уровня | 241 |
| 9.3. Сообщества практиков как пример сети знаний | 247 |
| 10. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ | 257 |
| 10.1. Традиционные программные средства для управления знаниями..... | 257 |
| 10.2. Архитектура системы управления знаниями на основе онтологии | 272 |
| 10.3. Программные средства управления знаниями на основе онтологий..... | 274 |
| 10.4. Процесс разработки систем управления знаниями на основе онтологий | 277 |
| 11. ПРАКТИЧЕСКИЕ ПРИМЕРЫ СИСТЕМ УПРАВЛЕНИЯ ЗНАНИЯМИ | 282 |
| 11.1. База знаний компании IBS..... | 282 |
| 11.2. Создание базы знаний проектов по управлению знаниями (на основе онтологий и рассуждений по прецедентам) | 290 |
| 11.3. Технология обработки знаний в области инженерии организации | 296 |
| 11.4. Стандартизация ремонтных операций | 304 |
| 11.5. Навигатор по методологии проектирования организаций – пример использования контекста для персонализации информации..... | 307 |
| 12. НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ В ОБЛАСТИ ОНТОЛОГИЙ И УПРАВЛЕНИЯ ЗНАНИЯМИ | 312 |
| 12.1. Проект ACTIVE | 313 |
| 12.2. Проект NEPOMUK..... | 317 |
| 12.3. Проект OrganiK..... | 325 |
| 12.4. Проект X-Media | 327 |
| 12.5. Проект Mymory..... | 331 |
| 12.6. Проект NEON | 333 |
| 12.7. Проект KnowledgeWeb..... | 334 |
| 12.8. Проект SEKT (Semantically-enabled Knowledge Technologies) | 334 |
| ЗАКЛЮЧЕНИЕ | 336 |
| БИБЛИОГРАФИЧЕСКИЙ СПИСОК..... | 339 |
| ГЛОССАРИЙ..... | 340 |

СОКРАЩЕНИЯ, ПРИНЯТЫЕ В ПОСОБИИ

БД – база данных,

БЗ – база знаний,

ИАД – интеллектуальный анализ данных,

ИПТ – информационно-поисковый тезаурус,

ИПЯ – информационно-поисковый язык,

ИР – интеллектуальные ресурсы,

ИТ – информационные технологии,

КП – корпоративная память,

КС – классификационная система,

НИР – научно-исследовательская работа,

НИОКР – научно-исследовательские и опытно-конструкторские работы,

СУЗ – система управления знаниями,

ССПК – социальный семантический персональный компьютер,

УЗ – управление знаниями,

OWL – Web Ontology Language, язык веб-онтологий,

RDF – Resource Description Framework, система описания ресурсов,

RDFS – RDF-схема.

ВВЕДЕНИЕ

В постиндустриальном обществе знания являются главным источником устойчивых конкурентных преимуществ предприятий, а управление знаниями (УЗ) – основным механизмом их обеспечения.

Основными задачами управления знаниями являются:

- выявление критически важных знаний или интеллектуальных ресурсов предприятия, необходимых для существования предприятия и/или реализации его стратегии,
- определение проблем и возможностей, связанных с этими знаниями,
- улучшение процессов создания, накопления, распределения и использования критически важных знаний,
- стимуляция обмена знаниями в неформальных социальных сетях,
- оценка интеллектуального капитала и деятельности в области знаний.

Одним из инструментов для решения указанных задач выступают информационные технологии, позволяющие лучше искать нужные знания, получать персонализированные рассылки и уведомления, проводить автоматизированный анализ имеющейся информации, а также поддерживающие групповую работу и коммуникацию сотрудников. Комплекс информационных технологий для поддержки процессов преобразования знаний можно называть системой управления знаниями (СУЗ), а точнее ее программной реализацией.

Любая система автоматизации затрагивает проблемы хранения корпоративных знаний, но только системы управления знаниями ориентированы на это в явном виде, тем самым способствуя сохранению этого ценнейшего ресурса, а не растворяя его в алгоритмах, программах, документации, технологических процессах.

В учебном пособии сделан акцент на разработке СУЗ на основе онтологий, которые представляют собой концептуальную модель предметной области в виде системы понятий, их свойств и отношений, понятную как людям, так компьютерам. Онтологии формируют единый язык для коммуникации, а также позволяют интегрировать разнородные интеллектуальные ресурсы предприятия в рамках единой системы. В

пособии предложена структура комплексной СУЗ, которая интегрирует основные подходы в данной области и позволяет четко разделить саму комплексную СУЗ и ее информационно-технологическую составляющую. Представлены уникальные переводы основных методологий управления знаниями, знакомство с которыми крайне полезно для разработки и внедрения СУЗ, как студенту, так и специалисту. Пособие обобщает широкий диапазон современных материалов (большая часть источников – 2000-е годы), а также интегрирует практические методы и инструменты с современными исследовательскими разработками в области.

Необходимо отметить, что в пособии детально рассмотрены только вопросы, связанные с применением информационных технологий в УЗ. Однако разработка комплексной СУЗ требует также согласованных мероприятий в области организационного управления, работы с персоналом, технической инфраструктуры и права. Поэтому для специалистов, ответственных за разработку комплексной СУЗ, необходимо дополнительно более детально ознакомиться с соответствующими областями или привлекать соответствующих специалистов. Кроме того, в пособии не рассмотрен вопрос *обучения*, которое, по мнению автора, является частным случаем процесса преобразования знаний, а также имеет много своей специфики и многовековых традиций.

Краткое описание разделов пособия и рекомендации по их изучению:

Раздел 1 дает общее представление о рассматриваемой предметной области и определения базовых понятий.

В разделе 2 дан обзор существующих методологий управления знаниями, на основе которых разработан интегрированный подход (*раздел 4*), изложенный в пособии.

Центральными и приоритетными разделами пособия являются 3 и 4.

В разделе 3 показывается роль онтологий и семантических технологий в СУЗ.

В разделе 4 приводятся основные компоненты СУЗ и дается их обобщенное описание, которое детализируется в последующих главах. Также в данном разделе описывается структура проекта по разработке и внедрению СУЗ. Изучение данного раздела необходимо для понимания позиционирования разделов с 5 по 10, которое позволит читателю

определить состав и порядок их изучения.

Раздел 5 описывает процесс управления знаниями. Элементы данного процесса также составляют проект по разработке СУЗ.

Разделы 6, 7, 8 и 9 описывают современные подходы, модели, методы и инструменты создания, накопления, распределения и обмена знаниями, а *раздел 10* систематизирует ИТ-инструменты УЗ. Данные разделы можно рассматривать в качестве источника справочной информации для выбора и обоснования необходимых элементов СУЗ.

Разделы 11 и 12 являются, довольно, самостоятельными, а их содержание следует из названия. *Раздел 11* позволяет лучше понять практическую ценность УЗ, а *раздел 12* будет, в первую очередь, интересен читателям, интересующимся научной стороной применения онтологий в СУЗ.

Пособие направлено на формирование следующих профессиональных компетенций:

- знание методов проектирования и производства программного продукта, принципы построения, структуры и приемы работы с инструментальными средствами, поддерживающими создание программного обеспечения;
- навыки использования основных моделей информационных технологий и способов их применения для решения задач в предметных областях;
- навыки выбора, проектирования, реализации, оценки качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях.

В результате освоения дисциплины обучающийся должен:

Знать:

- модели, методы и техники описания информационных ресурсов (формирования метаданных), структурирования, представления, поиска, доставки и визуализации знаний,
- модели, методы и техники поддержки совместной работы,
- рынок программных продуктов для ИТ-поддержки УЗ;
- функциональные возможности и основные характеристики средств ИТ-поддержки УЗ;

- модели и методы проектирования СУЗ;
- области и способы практического применения систем, основанных на знаниях;

Уметь:

- выявлять ключевые знания организации;
- определять проблемы и возможности, связанные с ключевыми знаниями;
- проводить анализ требований и разрабатывать архитектуру программной реализации СУЗ;
- проводить сравнительный анализ программных средств для УЗ;
- обосновывать потребность во внедрении средств ИТ-поддержки УЗ.

Владеть:

- Навыками диагностики и анализа знаний организации;
- Навыками выявления проблем и возможностей, связанных со знаниями;
- Методами анализа требований к программной реализации СУЗ.

Автор выражает благодарность своему научному руководителю Гавриловой Т. А. за полученные знания, помощь в определении целей и критику, а также Григорьеву Л. Ю. за знакомство автора с понятием «управление знаниями», за помощь в понимании места УЗ в управлении предприятием и за опыт практического применения семантических технологий для задач оргпроектирования. Автор также благодарен Горелику С. Л. за культивацию высоких требований к качеству результатов работы и самостоятельного мышления; Гордукаловой Г. Ф. за знакомство автора с подходами и методами информатики, а также за моральную поддержку; Шаламову А. В. за помощь в подготовке материала по оценке СУЗ; коллегам за интересные идеи, новые точки зрения и плодотворные обсуждения; студентам за конструктивные советы и критику; СПбГПУ за импульс к написанию пособия. И, конечно, хочется сказать «спасибо» близким за поддержку и понимание.

1. ЗНАНИЕ И УПРАВЛЕНИЕ ЗНАНИЯМИ

«Большинство ресурсов не являются специфичными: капитал, оборудование, материальные ценности – не могут быть использованы для того, чтобы отличить, один бизнес от другого по своей глубинной сути. Самый главный ресурс, отличающий бизнес и дающий решающие конкурентные преимущества, – это используемые в нем специфические производственные и управленческие знания».

Питер Друкер «Задачи менеджмента в XXI веке»

1.1. Определение понятия знание

Существует множество определений понятия «знание». Прежде чем дать рабочее определение данного понятия, рассмотрим существующие распространенные полезные определения.

Знание – проверенный практикой и достоверный логикой результат познания действительности, отраженный в сознании человека в виде представлений, понятий, суждений и теорий.

Компоненты такого определения также требуют пояснения:

Проверка практикой осуществляется путем наблюдений и экспериментов, где наблюдение – преднамеренное, планомерное восприятие, осуществляемое с целью выявить существенные свойства и отношения объекта познания; эксперимент – метод исследования, с помощью которого объект или воспроизводится искусственно, или ставится в определенные условия, отвечающие целям исследования.

Удостоверенность логикой подразумевает подчиненность правилам, принципам и законам, по которым мысль движется к истине, от одной истины к другой, более глубокой. Логические законы – обобщенное отражение объективных отношений вещей на основе практики.

Основными способами отражения действительности в сознании являются:

Представления – это образы тех предметов, которые когда-то воздействовали на органы чувств человека и потом восстанавливаются по сохранившимся в мозгу связям.

Понятие (в философии) – это мысль, в которой отражаются общие существенные свойства и связи предметов и явлений.

Суждение – это такая форма мысли, в которой посредством связи понятий утверждается (или отрицается) что-либо о чем-либо.

Теория – это система объективно верных, проверенных практикой знаний, воспроизводящих факты, события и их предполагаемые причины в определенной логической связи; это система суждений и умозаключений, объясняющих определенный класс явлений и осуществляющих научное предвидение.

В области управления знаниями в определениях устойчиво присутствует связь знания с действием. Вот одно из распространенных определений: **Знание – результаты обобщения информации, которые позволяют ставить и решать задачи в этой области.**

Рассел Аккоф, один из классиков исследования операций, предложил следующую популярную иерархию:

данные – информация – знания – понимание – мудрость.

Данные по Р. Аккофу – это некоторые неупорядоченные символы, рассматриваемые безотносительно к какому-либо контексту.

Информация – это выделенная и упорядоченная часть базы данных, обработанная для использования, то есть отвечающая на вопрос: «Кто?, Что?, Где?, Когда?»

Знание – это выявленные тенденции или существенные связи между фактами и явлениями, представленные в информации.

Понимание – это осознание закономерностей, содержащихся в разрозненных знаниях, позволяющее ответить на вопрос: «Почему?».

Мудрость – взвешенное, оцененное понимание закономерностей с точки зрения прошлого и будущего.

Продвижение по уровням иерархии понятий от «данных» к «мудрости» не есть механическое суммирование данных, информации, знаний.

Каждое из приведенных понятий является основой для последующего, материалом для получения элементов нового более высокого качества знаний. При этом, как считает Р. Аккоф, первые четыре понятия имеют дело с прошлым или с тем, что уже известно, а «мудрость» касается будущего.

Встречается такое полезное определение знаний – *«информация в контексте»*. То есть именно контекст делает информацию знанием, если

она (информация) согласуется и полезна в ситуационном контексте.

Определение понятия «знание», приведенное в Схеме по управлению знаниями (Knowledge Management Framework) в рамках Европейского методического руководства (European Guide to good Practice in Knowledge Management): *«Знание – это комбинация данных и информации, к которым добавлено мнение, мастерство и опыт эксперта, что в результате дает ценный актив, который может быть использован для оказания помощи в принятии решений».*

Будем трактовать информацию, как наиболее общее понятие, а данные и знания – как разные степени организации этой информации.

«Данные» – это полученные эмпирическим путем и зафиксированные факты, характеризующие отдельные свойства объектов, процессов или явлений.

Знания связаны с данными, основываются на них, но представляют результат мыслительной деятельности человека.

«Знания» – это результаты обобщения фактов и установления определенных закономерностей в какой либо предметной области, которые позволяют ставить и решать задачи в этой области.

Если рассмотреть переход от «данных» к «знаниям» с точки зрения уровня организации информации, то этот переход характеризуется все большим усложнением информационных структур и появлением многочисленных связей между ними. Часто используются и такие определения знаний: *знания* – это хорошо структурированные данные, или данные о данных, или метаданные.

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

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

Например, интенционал понятия «мебель»: «предметы, предназначенные для обеспечения комфортного проживания человека и загромождающие дом».

Экстенционал понятия – это определение понятия через перечисление его конкретных примеров, т. е. понятий более низкого

уровня абстракции.

Экстенционал понятия «мебель»: «шкаф, диван, стол, стул, etc.»

Интенционалы формируют знания об объектах, в то время как экстенционал объединяет данные. Вместе они формируют элементы поля знаний конкретной предметной области.

1.2. Источники и виды знаний

Источники знаний, конечно, зависят от отраслей индустрии, но, как правило, к ним относятся корпоративные стандарты, методики, бизнес-правила, технологии, процедуры, информационные массивы, руководства, письма, новости, сведения о заказчиках и конкурентах, схемы, чертежи и другие документальные материалы, а также знания, навыки и умения персонала (рис. 1.1).



Рис. 1.1. «Цветок» информационных составляющих знаний в организации

Как видно из этого рисунка – источники существенно отличаются степенью формализации и доступности. Поэтому сразу имеет смысл выделить одну из основных классификаций знаний:

Явные (формализованные) знания – знания, которые можно найти в документах организации в форме сообщений, писем, статей, справочников, патентов, чертежей, видео- и аудиозаписей, программного обеспечения и т. д. Формализованные знания формируют структурный капитал, включающий в себя также «компьютерные программы, патенты, торговые марки и все остальные организационные аспекты, обеспечивающие производительный труд работников».

Скрытые знания – это персональные знания, неразрывно связанные с индивидуальным опытом. Его можно передать путем прямого контакта – «с глаза на глаз» или при помощи специальных процедур извлечения знаний. Считается, что именно скрытое практическое знание является ключевым для принятия решений и управления.

Скрытые знания формируют человеческий капитал, который воплощен в работниках компании или их коллективах как совокупность знаний, квалификации, новаторства каждого из сотрудников, как система ценностей, культура и философия компании.

Между явными и неявными знаниями происходит непрерывный обмен и трансформация, схема которых была предложена японскими специалистами Икуджиро Нонака и Хиротака Такеучи [Нонака И., Такеучи Х., 2003]. Этот циклический процесс показан на рис. 1.2.

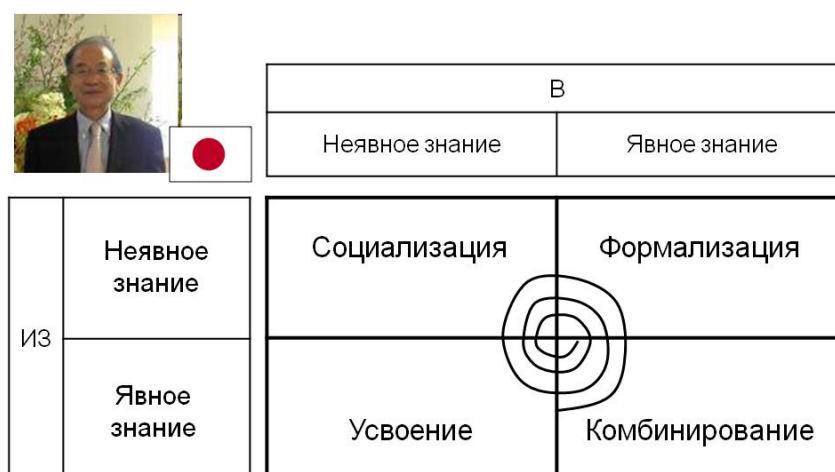


Рис. 1.2. Четыре модели трансформации знаний

Преобразование знаний в пределах одной формы или при переходах между формами происходит в результате следующих процессов:

- *Социализация* (неявное в неявное) включает формирование и передачу неявных знаний в коллективных формах (дискуссии, семинары, команды и т. п.). При этом чаще всего это происходит без создания явных знаний.

- *Формализация* (иногда говорят «экстернализация») (неявное в явное) – происходит путем концептуализации неявных знаний, извлечения и выявления их и, в конечном счете, их формулировании и фиксации в той или иной форме как итог дискуссий, семинаров, мозгового штурма и т. п.

- *Комбинирование* (явное в явное) – осуществляется в ходе интеграции различных источников информации, в процессе составления обзоров и сводных отчетов. Увеличение явных знаний происходит здесь за счет пополнения баз и хранилищ коллективного пользования, классификации и систематизации файлов и документов и т. п.

- *Усвоение* (интернализация) (явное в неявное) – осуществляется в процессе чтения и изучения документов из баз данных, журналов и книг. Это приводит к усвоению знаний, которые были созданы другими (возникновению новых неявных знаний у познающего субъекта), а также, возможно, к созданию новых знаний в результате мыслительной деятельности в процессе познания.

Перечисленные процессы не происходят изолированно, а выполняются вместе в различных сочетаниях в обычных производственных ситуациях. Например, знания создаются в результате взаимодействия сотрудников и явных знаний. Посредством разнообразных взаимодействий неявные знания материализуются и распределяются. Хотя индивидуумы (например, служащие) повседневно участвуют во всех этих процессах, наибольшая ценность возникает от их согласованного сочетания применительно к ключевым знаниям.

Рассмотрим некоторые другие полезные классификации знаний, используемые в рамках учебного пособия.

По уровню формализованности знания делятся на:

А. Знания в памяти человека как результат анализа опыта и мышления (неформализованные знания);

В. Знания, представленные в явном виде, с помощью произвольной знаковой системы (например, специальная литература, учебники или методические пособия на естественном языке);

С. Поле знаний — структурированное описание основных понятий и объектов предметной области, их атрибутов и закономерностей, их связывающих;

Д. Знания, записанные на языках представления знаний (продукционные языки, семантические сети, фреймы, дескрипционная логика);

Е. Базы знаний на машинных носителях информации.

В области разработки интеллектуальных систем распространено деление знаний на декларативные и процедурные:

- декларативные (знания предметной области) (domain knowledge) – знания о свойствах предметной области, закономерностях, характерных для нее, гипотезах о возможных связях между явлениями, процессами и фактами в ней.

- процедурные знания (procedural knowledge) – знания, воплощенные в компьютерных программах для решения тех или иных задач.

Классификация знаний по рассматриваемому аспекту предметной области, предложенная в работах Джона Захмана применительно к архитектуре предприятия и Татьяны Гавриловой применительно к анализу произвольной предметной области (табл. 1.1):

Таблица 1.1

Классификация знаний по рассматриваемому аспекту предметной области

| Аспект предметной области | Примеры знаний | Вид деятельности |
|----------------------------------|---|-------------------------------------|
| ЗАЧЕМ-знания | Знания о целях, назначении, функции рассматриваемой системы | <i>стратегический анализ/синтез</i> |
| ЧТО-знания | Основные концепты, понятийная структура | <i>концептуальный анализ</i> |
| КАК-знания | Алгоритмы, процедуры, процессы, модели принятия решения | <i>функциональный анализ/синтез</i> |

| Аспект предметной области | Примеры знаний | Вид деятельности |
|---------------------------|---|--|
| КТО-знания | Организационные роли или должности или конкретные люди являющиеся исполнителями функций, носителями знаний, ответственными за что-либо. | <i>организационный анализ/синтез</i> |
| ГДЕ-знания | Местоположение, окружение. | <i>пространственный анализ/синтез</i> |
| КОГДА-знания | временные параметры и ограничения, планы-графики | <i>временной анализ/планирование</i> |
| ПОЧЕМУ-знания | Обоснования, аргументы, причинно-следственные связи. | <i>каузальный или причинно-следственный анализ</i> |

Данная классификация позволяет правильно подбирать способы работы с каждым видом знаний, например, выбирать методы структурирования или модели представления знаний.

1.3. Знания и интеллектуальный капитал

Понятие интеллектуального капитала позволяет перейти к экономической точке зрения на знания, которыми обладает компания.

Появление термина «интеллектуальный капитал» относится к первой половине 90-х годов. В его популяризации заметную роль сыграла книга Тома Стюарта «Интеллектуальный капитал – главное богатство вашей компании». Следует отметить, что к тому времени и, тем более, в последующие годы интеллектуальный капитал ведущих национальных и транснациональных компаний стал составлять заметную и постоянно растущую долю в их рыночной стоимости.

Под рыночной стоимостью или стоимостью акционерного капитала понимается балансовая стоимость материальных активов компании плюс стоимость ее интеллектуального капитала. Последний состоит из

человеческого, структурного и капитала отношений. Часть интеллектуального капитала может быть юридически зафиксированна в виде интеллектуальной собственности (патенты, лицензии, торговые марки и т. п.). На рис. 1.3 в качестве иллюстрации приведены данные о структуре рыночной стоимости трех известных в мире компаний по данным за 1999 г.¹.

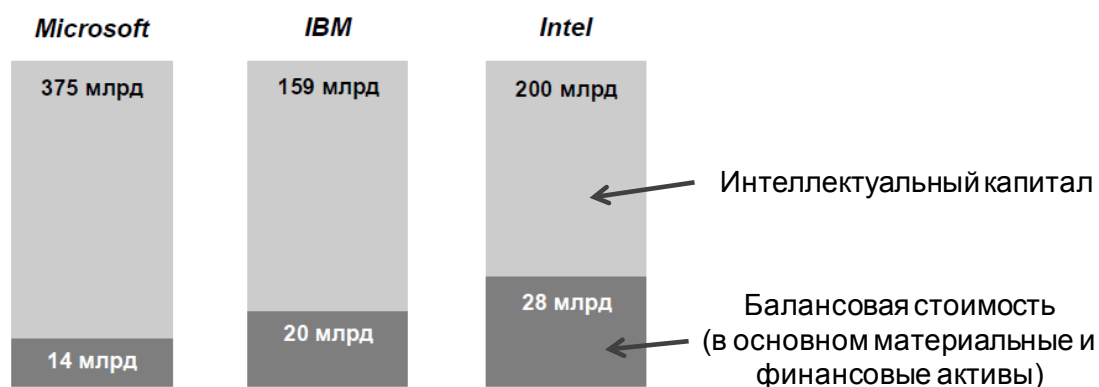


Рис. 1.3. Структура рыночной стоимости компании

Известно, что Microsoft, IBM и Intel – это компании, работающие в сфере высоких компьютерных, информационных и телекоммуникационных технологий, и высокий удельный вес интеллектуальных активов естествен для них.

Однако, как показывают экономические исследования, существенное превосходство рыночной стоимости компании над балансовой – это не исключение, а скорее правило, характерное для многих высокотехнологичных компаний. Не балансовая стоимость, а уровень текущих и будущих доходов определяет цену, которую готовы заплатить за нее покупатели пакетов акций либо бизнеса в целом.

Основные модели интеллектуального капитала представляют его в виде трех составляющих:

- Капитал отношений – совокупность наработанных связей с окружением, прежде всего с клиентами.
- Организационный (структурный) капитал – способности

¹ Tiwana A. The knowledge management toolkit: practical techniques for building a knowledge management system. Prentice Hall PTR. – 2000.

организации, проистекающие из формализованных знаний, деловых процессов, организационной культуры, ценностей и норм;

- Человеческий капитал – способность индивидуумов и команд организации удовлетворять потребности заинтересованных сторон благодаря своим компетенциям, лояльности и мотивации (включает неявные знания сотрудников).

С учетом такой структуры интеллектуального капитала, можно сказать, что управление знаниями направлено, в первую очередь, на повышение организационного и человеческого капиталов.

1.4. Определения и назначение управления знаниями

Управление знаниями (УЗ) – это междисциплинарная область науки и практики, имеющая корни в информационных технологиях, библиотечном деле, экономике, теории организации, педагогике, психологии, социологии. УЗ:

- это регулярная деятельность организации,
- с помощью которой организации удается извлечь прибыль из объема знаний или интеллектуального капитала, находящегося в ее распоряжении.

- направленная на предоставление нужных знаний, в нужной форме, в нужном количестве, в нужном месте, в нужное время и по приемлемой цене.

- рассматривающая явные и неявные знания в качестве стратегического ресурса организации,

- направленная на улучшение работы со знаниями на индивидуальном, групповом, организационном и межорганизационном уровнях,

- представляющая собой взаимосвязанную совокупность инструментов, техник и процессов для выявления, накопления, распределения, обмена, использования и оценки знаний на предприятии,

- обеспечивающая необходимый уровень технологических, организационных, культурных и правовых условий для эффективного использования и развития интеллектуального потенциала предприятия.

Примеры проблем, на решение которых направлено управление знаниями:

- «50 % работников уйдет на пенсию в течение пяти лет – нужно сохранить их знания»,
- «3 производственные линии делают похожую продукцию, часть работ схожая. Но они не используют наработки друг друга»,
- «В наше объединение входят три поликлиники. Между ними нет обмена знаниями и опытом. Доктора не хотят передавать лучшие практики»,
- «Высокая текучка персонала в службе поддержки оператора сотовой связи. Постоянно нужно обучать персонал. Даже ответы на типовые вопросы находятся только в головах сотрудников».

1.5. История возникновения и подходы к управлению знаниями

Важно понимать, что управление знаниями возникло не из набора формальных моделей и методов. Эта область имеет, как минимум, четыре основных источника возникновения.

Как это часто случается, примерно в одно и то же время в разных частях земли стали зарождаться близкие идеи. В середине 80-х годов XX века исследователи и практики стали интересоваться ролью знаний в бизнесе.

В США термин «управление знаниями» начал использоваться в контексте искусственного интеллекта в 1986 году. Проблема в области искусственного интеллекта заключалась в том, что большинство создаваемых систем искусственного интеллекта (как правило, экспертных систем) оказывалось ненужным через несколько месяцев после внедрения.

Группа консультантов компании Arthur D. Little начала изучать какую роль играют знания в различных областях промышленности и типах задач – в проектировании, в прогнозировании погоды, управлении технологическим процессом и т. д. Карл Вииг (http://www.krii.com/who_we_are.htm), который был руководителем той группы по искусственному интеллекту, сказал (электронное письмо, ноябрь 2001):

«Мы начали мыслить о создании, обучении, передаче, использовании и усилении знаний, как о социальных и динамических процессах, которые требуют управления (конечно, информационные

технологии иногда участвуют в этом, но они не являются центральным элементом, как мы выяснили позже). И мы не смогли найти лучшего термина, чем «управление знаниями». Сейчас я сожалею, что мы не смогли найти более подходящего словосочетания».

Карл Вииг впервые использовал термин УЗ в презентации 1986 года и опубликовал несколько книг, обобщающих опыт его группы. В 1990 году он написал, возможно, первую в мире статью с УЗ в заголовке:

«Knowledge Management: An Introduction» in Proceedings of IAKE Second Annual International Conference. Washington DC: International Association of Knowledge Engineers, 1990, pp. 13-41.

В Японии с начала 80-х годов XX века группа специалистов во главе с Икуджиро Нонака (http://www.ics.hit-u.ac.jp/faculty/profiles/ikujiro_nonaka.html) изучала процесс инноваций и возможности его ускорения в японских корпорациях. Нонака обратил внимание на неявное (неформализованное) знание сотрудников, а также на взаимные трансформации неявного и явного знаний. В 1995 Нонака совместно с Такеучи публикуют свою книгу «Компания – создатель знаний» («The Knowledge Creating Company»), которая переопределила область УЗ. Нонака противопоставляет свою концепцию создания знаний исходной концепции УЗ, которая, на его взгляд, была слишком связана с информационными технологиями.

В Швеции Карл-Эрик Свайби (<http://www.sveiby.com>) руководил своей собственной издательской компанией и пытался разработать стратегию для нее. Вопрос заключался в том, какие ресурсы имеются в его распоряжении. У него не было существенных материальных и финансовых ресурсов, а единственным фактором производства являлись знания и компетенции сотрудников. В результате, он разработал «стратегический подход, основанный на компетенциях», в 1986 г. опубликовал книгу «Kunskapsföretaget» («The Know-How Company», по-английски), а потом развил его в книге «Kunskapsledning» («Knowledge Management», «Управление знаниями», по-английски) в 1990. В названии книги Свайби ссылается на стратегический подход к управлению ресурсами знаний и работниками знаний. В это время он разработал теорию для оценки нематериальных активов и, в частности, капитала знаний/интеллектуального капитала (см. книгу «The Invisible Balance

Sheet», 1989). К 1991 году около 40 шведских компаний опубликовали свою отчетность в соответствии с предложенной Свайби методикой, а Skandia Insurance стала наиболее известным их представителем (см. «Навигатор компании Скандия», «Scandia Navigator»).

Достаточно близкие взгляды развивал и американский специалист Дэвид Тиис (<http://businessinnovation.berkeley.edu/Teece/>), занимавшийся вопросами разработки стратегии и исследованием конкурентных преимуществ в бизнесе. Он рассматривал знания, как стратегический ресурс и трудновоспроизводимый источник конкурентного преимущества. По его мнению, в основе конкурентных преимуществ компании лежат динамические способности – потенциал распознавания новых возможностей и реконфигурации знаний как активов, компетенций и комплиментарных активов и технологий для достижения устойчивого конкурентного преимущества [Тееце, 2003].

Каждый из указанных источников УЗ расставлял собственные акценты. Практика же требует комплексного применения указанных выше подходов. Поэтому сейчас доминируют синтетические подходы, в которых каждому аспекту УЗ уделяется необходимое внимание. Данные подходы представлены в работах Тома Дейвенпорта (Davenport Tom), Ларри Прусака (Prusak Larry), Букович У., Уильямс Р., см. рис. 1.4.



Рис. 1.4. Основоположники и практики УЗ

2-м различным взглядам на знание:

1. Знание как *«продукт»* (или как *«запас»*). Согласно этому подходу, знание – это актив, продукт, который может существовать независимо от человека и управление которым подобно управлению любыми другими ресурсами. Данному взгляду на знание соответствует стратегия (или подход) *«кодификации»* (Человек – Контент – Человек), которая предполагает обязательную формализацию тех знаний и информации, которые особенно важны для компании. В рамках данной стратегии сотрудники формализуют часть своих знаний (писать отчеты, например) и помещают их в базы и хранилища. В будущем они сами и их коллеги смогут воспользоваться ими.

2. Знание как *«процесс»* (или как *«поток»*). В таком подходе знание не может передаваться как некий объект, поскольку в формирование знания вносит существенный вклад человек-участник процесса передачи. Данный взгляд объясняется тем, что знание – результат интерпретации информации в определенном контексте. Знание не передается, а воссоздается в мозгу человека при получении информации в соответствии с исходными знаниями, опытом, ценностями и ожиданиями. Данному взгляду на знание соответствует стратегия (или подход) *«персонализации»* (Человек – Человек), которая подразумевает непосредственное общение сотрудников (например, на совещаниях по обсуждению концепции нового web-сайта компании) и совместную работу. В организациях выявляются сотрудники-лидеры в какой-то области знаний и создаются условия для того, чтобы они могли делиться опытом с остальными (например, наставничество), иногда формируются сообщества для обмена знаниями в определенной области.

Каждая компания может определить, какой из двух подходов к управлению знаниями для нее наиболее эффективен, для этого нужно учитывать множество факторов, например, сфера деятельности компании, ее стратегия, корпоративная культура и т. п. Различия взглядов на знание проявляется в различиях и степени использования информационных технологий. На практике оба взгляда часто сочетаются.

Например, в одной из крупнейших консалтинговых компаний для управления взаимоотношениями с клиентами использовался продукт группы CRM (Customer Relationship Management), в который все

сотрудники должны класть отчеты о встречах с клиентами, но при этом еженедельно проходили собрания сообщества этих же сотрудников, на которых обсуждались все детали этих контактов.

1.6. Обзор задач управления знаниями

При всей очевидности целей управления знаниями спектр задач, которые предлагают решать различные авторы публикаций и консультанты по управлению знаниями, очень многообразен и неоднозначен. Примеры различных подходов к формулировке задач управления знаниями, приведены в табл. 1.2:

Таблица 1.2

Подходы к определению задач УЗ

| Подход Карла Виига с коллегами [Wiig et al, 1997] | Методология Excalibur Technologies | Консалтинговая компания PricewaterhouseCoopers [Букович, Уильямс, 2002] |
|--|--|--|
| 1. Диагностика и оценка деятельности 2. Диагностика и анализ знаний 3. Выбор и планирование действий 4. Действия а. Развитие знаний б. Распространение знаний с. Комбинирование знаний d. Интеграция знаний | 1. Доступ/поиск 2. Совместная работа 3. Нововведение 4. Создание 5. Внедрение 6. Оценка 7. Усовершенствование 8. Хранение | Тактические процессы 1. Получение знаний 2. Использование знаний 3. Обучение 4. Распространение знаний Стратегические процессы 1. Оценка интеллектуального капитала (ИК) 2. Создание и поддержание ИК 3. Ликвидация активов знаний |

Задачи УЗ, выделяемые всеми авторами, образуют некие циклы УЗ.

Циклы, представленные в табл. 1.2, кажутся вполне логичными, если рассматривать их отдельно, но почему тогда возникают такие различия, как в формулировках задач, так и в содержании работ? Какой из предлагаемых циклов наиболее полно и точно отражает задачи, стоящие в предметной области? Для ответа на эти вопросы имеет смысл обратить внимание на объект управления, на который направлены задачи, решаемые в рамках представленных циклов.

Есть задачи, которые непосредственно направлены на работу с информацией – поиск, получение, распространение *знаний*. Однако другие задачи направлены на объекты, отличающиеся от знаний и информации: обучение и совместная работа *персонала*, оценка интеллектуального капитала / деятельности *компании*. Но можно ли утверждать, что такие задачи не относятся к области управления знаниями? Наверное, нет. Такой ответ, как и само многообразие задач, объясняется междисциплинарностью УЗ.

1.7. Междисциплинарность управления знаниями

УЗ – это область, находящаяся на пересечении трех дисциплин, см. рис. 1.5:

1. Экономика и менеджмент, область, с точки зрения управления знаниями отвечающая на вопросы «Как создавать добавленную стоимость и обеспечивать конкурентное преимущество на основе знаний?». Главным объектом управления в данной области является **предприятие** как открытая социально-экономическая система.

2. Науки об информации, которые с точки зрения управления знаниями отвечают на вопрос «Как эффективно работать с информацией?». Соответственно названию наук, главным объектом изучения и управления в данной области является **информация**.

3. Гуманитарные науки, которые с точки зрения управления знаниями отвечают на вопрос «Как обеспечить необходимое поведение людей?». Главным объектом изучения в данной области является **человек**.



Рис. 1.5. Междисциплинарность управления знаниями

Таким образом, эти три грани УЗ заставляют по-разному рассматривать и формулировать задачи в области управления знаниями с учетом рассматриваемого объекта.

2. ОБЗОР МЕТОДОЛОГИЙ УПРАВЛЕНИЯ ЗНАНИЯМИ

2.1. Методология Know-Net

Know-net – комплексная методология управления активами знаний (knowledge asset management) или, другими словами, создания и развития СУЗ [Mentzas et al, 2002]. Данная методология была разработана в 2000 г. в рамках одноименного международного проекта. В консорциум разработчиков методологии входили консультационная фирма Knowledge Associate, консультационная фирма «Planet» (сейчас «Planet Ernst & Young»), бизнес-школа INSEAD, исследовательское подразделение национального технического университета г. Афины, немецкий центр искусственного интеллекта (DFKI), швейцарский академический институт Fachhochschule Basel.

Методология Know-Net включает 3 компонента:

1. Концептуальная схема (Know-net Framework) – система понятий, позволяющая менеджерам интегрировать усилия по УЗ.

2. Методика (Know-Net Method) – последовательность действий и рекомендации по их выполнению, которые помогают организациям определить их стратегию УЗ, усовершенствовать процессы, оргструктуру, неформальные сети коммуникаций и технологии для реализации выбранной стратегии УЗ, а также измерить и оценить интеллектуальный капитал организации.

3. Инструмент (Know-Net Tool) – программное средство поддержки УЗ на 2-х уровнях: стратегическом (анализ имеющихся активов и объектов знаний в организации) и операционном (сбор и категоризация внешней и внутренней информации, навигация по имеющимся знаниям и развитый поисковый механизм, средства совместной работы).

Три компонента методологии взаимосвязаны (рис. 2.1).

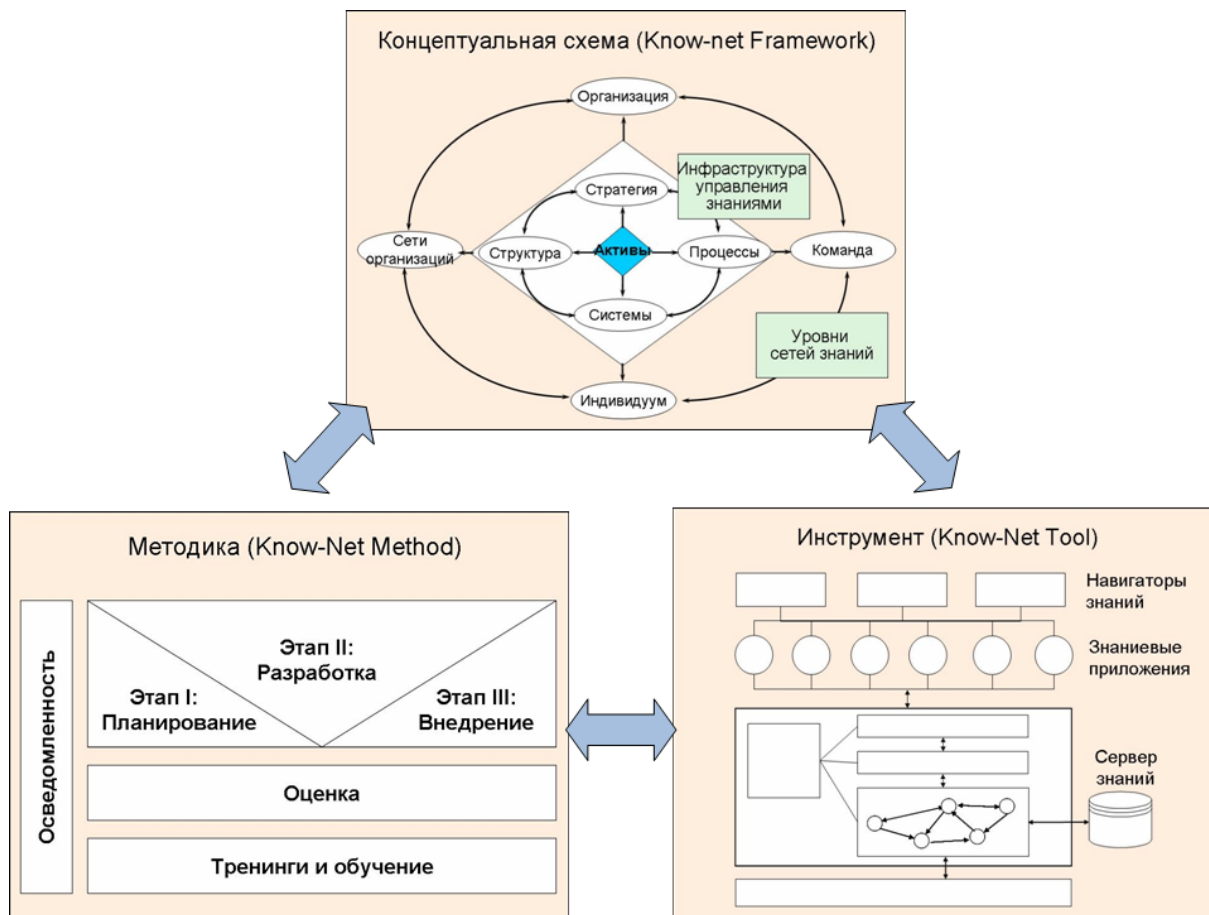


Рис. 2.1. Компоненты методологии Know-Net

Концептуальная схема позволяет ознакомиться с подходом, а также дает общий язык для коммуникации сотрудников предприятия и для работы в программном средстве УЗ. Схема позволяет обсудить, какие области знаний являются критичными, какие недоразвитыми и требуют улучшения, какие уже развиты, могут использоваться, представляют ценность и требуют охраны. Элементы схемы (стратегия, структуры, процессы, системы) используются для оценки уровня зрелости УЗ в организации и для разработки стратегии работы с активами знаний. Выбор уровня УЗ (индивид, группа, организация, межорганизационный) фокусирует анализ УЗ.

Реализация методики тесно связана с использованием, настройкой и развертыванием программного средства УЗ (Know-Net Tool). На этапе стратегического планирования программное средство позволяет описывать и структурировать основные элементы концептуальной схемы (активы и объекты знаний, стратегия, процессы, структуры, системы), на этапе

проектирования СУЗ начинается настройка приложений.

Концептуальная схема (Know-net Framework)

Активы знаний – ядро концептуальной схемы.

Другими элементами схемы являются:

Инфраструктура управления знаниями:

Стратегия управления знаниями включает в себя связанные со знаниями стратегические ценности; ссылки на знания в миссии, видении; связанные со знаниями качественные и/или количественные цели; связи стратегии управления знаниями с целями и задачами бизнес-стратегии. Например, для географического расширения (бизнес-стратегия) нужно уметь создавать новые продукты и осваивать новые компетенции (стратегии УЗ), для операционного совершенства (бизнес-стратегия) нужно развивать технологические инновации и обмен знаниями (стратегия УЗ), для близости к покупателю (бизнес-стратегия) нужно развивать знания о покупателе и бренде (стратегия УЗ).

Структура управления знаниями включает в себя следующие организационные роли: главный по знаниям (Chief Knowledge Officer), главный по обучению (Chief Learning Officer); менеджеры среднего звена: менеджеры по знаниям, координаторы знаний, интеграторы знаний; технические роли: инженеры по знаниям, контент-менеджеры, аналитики.

Процессы преобразования знаний включают в себя создание, получение, распределение, обмен и использование знаний.

Системы включают в себя сервисы распределения и публикации; сервисы совместной работы; сервисы индексирования и классификации; сервисы для хранения и обработки метаданных; сервисы интеграции; сервисы поиска и навигации; сервисы для передачи знаний.

Уровни управления знаниями:

Индивид, Команда, Организация, Сеть организаций.

Методика

Основной принцип УЗ в том, что это не только технологический вопрос.

Этап I. «Стратегическое планирование УЗ»: организация определяет (а) готовность и видение инициативы по УЗ, (б) содержание и реализуемость (scope and feasibility) проекта. Основные шаги этапа:

- Достижение поддержки целесообразности УЗ руководством,
- Установление связи УЗ со стратегией компании,
- Проведение анализа знаний,
- Оценка рисков и готовности к изменениям,
- Разработка бизнес-кейса (упрощенный бизнес-план) для УЗ,
- Принятие решения о внедрении.

Этап II. «Разработка архитектуры СУЗ»: архитектура комплексной СУЗ (включающей процессы, людей и технологии) итеративно разрабатывается, тестируется и оценивается (рис. 2.2): Анализ бизнес-процессов, улучшение знаний в бизнес-процессах, анализ сетей знаний (knowledge networks), улучшение сетей знаний, анализ технологий, улучшение технологий, разработка схемы активов знаний (knowledge asset schema), интеграция УЗ архитектуры.

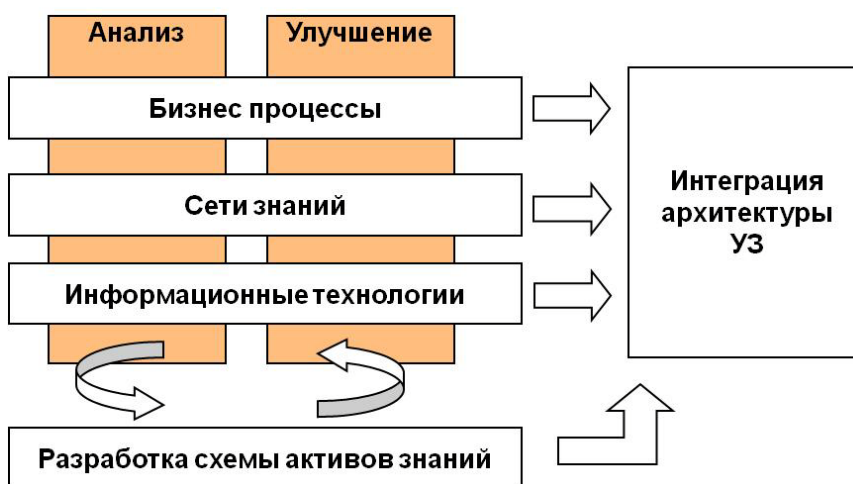


Рис. 2.2. Блоки работ по разработке архитектуры СУЗ

Этап III. «Внедрение»: полномасштабное внедрение разработанной архитектуры УЗ. Параллельно с указанными выше этапами проводится оценка активов знаний организации, а также тренинги и обучение. Ключевым объектом для оценки являются активы знаний, связанные с критическими факторами успеха.

Описание инструмента Know-Net может быть найдено в [Mentzas et al., 2002].

2.2. Создание знаний в подходе Нонака и Такеучи

В основе подхода Нонака и Такеучи стоит процесс создания знаний. На рис. 2.3 представлена пятифазная модель процесса создания знания организацией, предложенная в [Нонака И., Такеучи Х., 2003]. Данный процесс основан на предложенной авторами модели трансформации знаний представленной на рис. 1.2 и рассмотренной в разделе 1.2.

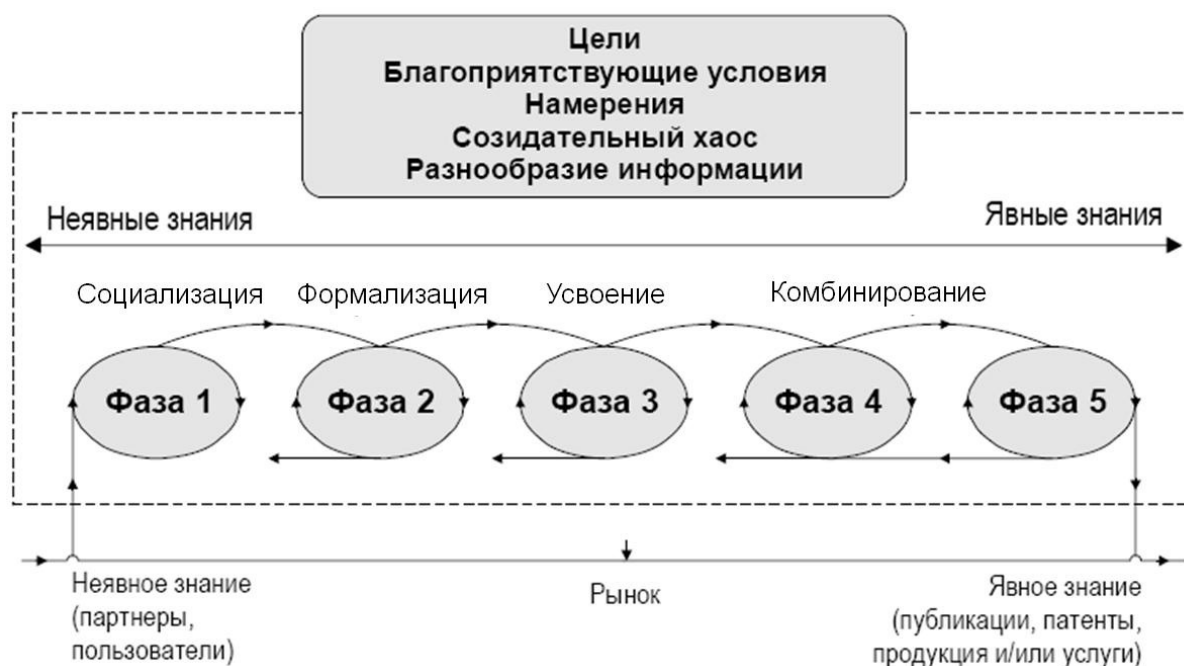


Рис. 2.3. Пятифазная модель процесса создания знания организацией:
фаза 1 – распространение неявного знания; фаза 2 – создание концепции;
фаза 3 – проверка концепции; фаза 4 – построение архетипа;
фаза 5 – переход знания на новый уровень

Процесс создания организационного (корпоративного) знания начинается с распространения неявных знаний (фаза 1), что в общих чертах напоминает социализацию, поскольку концентрированное и пока неиспользуемое индивидуальное знание должно, прежде всего, распространяться в пределах организации. Во второй фазе распространенное неявное знание преобразуется некоторой командой (стихийно или осознанно организованной) в виде новой концепции в явное знание. Процесс этот соответствует формализации знаний. В третьей фазе созданная концепция подвергается проверке, в ходе которой организация

определяет, действительно ли предложенная концепция имеет право на существование. Получившая «путевку в жизнь» концепция в четвертой фазе преобразуется в архетип, который может принять форму прототипа при разработке некоторого материального продукта или организационного элемента (структуры, процесса) в случае, если инновация носила нематериальный характер. Последняя фаза создания знания предусматривает распространение знания внутри подразделения, по разным подразделениям или среди дочерних компаний, потребителей, дистрибьюторов и даже в университетах. Тем самым осуществляется переход знания на новый уровень, и в том нет ничего удивительного, ибо создающая знания компания – это открытая система, постоянно обменивающаяся знаниями с актуальной средой.

Поскольку процесс создания нового знания сложен и, отчасти, таинствен, имеет смысл подробнее остановиться на описании каждой из упомянутых фаз.

Распространение неявного опыта (фаза 1) начинается со сбора неявных знаний, являющихся в большинстве организаций слабо задействованным, хотя и богатым источником нового знания. Неявные знания весьма сложно распространить или передать, поскольку они, в основном, получены из опыта и с трудом поддаются даже вербализации. Дело также не в том, что получателями неявных знаний являются множество индивидуумов с различным образованием и предшествующим опытом, отличающихся точками зрения и мотивацией. Для достижения успеха в обобществлении знаний эмоции и интеллектуальные модели индивидуумов следует сделать общими, то есть создать некоторое общее «поле», на котором индивидуумы смогут взаимодействовать в благоприятной среде, дискутировать, обмениваться опытом, синхронизировать свою физическую и интеллектуальную активность. Типичное общее поле взаимодействия – команда, члены которой собраны, а лучше сами пришли из различных отделов, но склонны работать сообща и стремятся к единой, увлекающей их цели.

Самоорганизующаяся команда создает новое организационное знание благодаря разнообразию своего состава, вовлекая избыточную информацию и разнообразие видения цели и средств ее достижения. Руководство организации преднамеренно создает созидательный хаос

постановкой труднодостижимых целей, представляя каждому члену команды значительную самостоятельность. Команда сама определяет целевые горизонты и границы задач и, как особый отдел с широкими полномочиями, начинает взаимодействовать с актуальной средой, накапливая неформализованное и формализованное знание.

Создание концепции (фаза 2) осуществляется в режиме наиболее интенсивного взаимодействия неявных и явных знаний. Как только в поле взаимодействия выкристаллизовывается некая общая интеллектуальная модель, команда приступает к ее формализации в процессе постоянного диалога, индивидуальных и коллективных размышлений. Эта фаза превращения неформализованной интеллектуальной модели в формализованную более всего соотносится с этапом формализации знаний. В этом процессе используются разнообразные методы мышления, такие, как дедукция, индукция и абдукция. Особая роль в данной фазе принадлежит абдукции, использующей язык метафоры и аналогий. Коэффициент полезного действия от взаимодействия членов команды повышается при использовании диалектики, ибо создание знаний – это идущий по спирали процесс синтеза, вытекающий из дискуссий, противоречий и парадоксов.

Автономия, а она является весьма ценным активом на этом этапе, позволяет членам команды свободно обмениваться мыслями с целью выявления направления интеллектуальной деятельности. Для создания концепции разработчикам порой приходится в корне пересматривать свои первоначальные представления. В этом им помогает разнообразие информации и множественность точек зрения членов команды. Встряски и хаос внешнего и внутреннего происхождения также активно воздействуют на изменение образа мыслей, что также способствует творчеству.

Проверка концепции (фаза 3) является важной фазой в создании знания, ибо новое организационное знание должно быть обоснованным истинным убеждением, имеющим определенную ценность, как для организации, так и для общества. Специалисты проверяют информацию, концепции или знания постоянно и даже бессознательно по мере их получения или в процессе их создания. Однако организация должна осуществлять проверку более строго и всесторонне для установления соответствия концепции своим намерениям и целям.

Для бизнес-организации общепринятыми критериями достоверности концепции являются затраты на ее реализацию, прибыль и вклад, вносимый результатами в развитие фирмы. Критерии оценки в зависимости от типов концепции могут быть количественными, качественными либо смешанными.

Определение системы критериев и желательных диапазонов их значения является прерогативой руководства компании. При этом оно должно учитывать не только корпоративные интересы и ценности, но и систему ценностей профессионального и регионального сообщества, общества в целом, его запросы и т. п.

Построение архетипа (фаза 4), которым могут быть прототип новой продукции или технология предоставления новой услуги, осуществляется путем сочетания только что созданного нового формализованного знания с уже существующим формализованным знанием, то есть этой фазе можно поставить в соответствие процесс комбинирования знаний.

Для построения, например, прототипа вместе собираются специалисты с различными знаниями и опытом (из отделов НИОКР, маркетинга, производства, технического контроля и др.) и совместно добиваются конкурентных характеристик нового продукта. Аналогично происходит и с моделью новой оргструктуры или технологией оказания услуг.

При извлечении знаний в процессе межличностного сотрудничества полезно использовать коммуникативные методы, а именно:

- пассивные (наблюдения, протоколы «мыслей вслух», лекции);
- активные (игры, диалоги, круглые столы);
- групповые («мозговой штурм», ролевые игры);
- индивидуальные (анкетирование, интервью, экспертиза).

Поскольку настоящая фаза носит комплексный характер, в ней необходимо обеспечить активное сотрудничество различных отделов и служб организации. Должна быть также обеспечена избыточность информации, разнообразие опыта, технологических инноваций, стимулирование межличностного и межструктурного сотрудничества.

Переход знания на новый уровень (фаза 5) – следствие того очевидного факта, что создание нового знания – процесс бесконечный и самовоспроизводящийся. Новая концепция создается, выверяется,

моделируется и затем восходит к новому циклу создания знаний, который находится в иной онтологической плоскости. Этот интерактивный, развивающийся по спирали процесс происходит как в пределах одной организации, так и на межорганизационном уровне. На данном этапе имеет место процесс усвоения знаний.

Реализованное или принявшее форму архетипа внутриорганизационное знание способно инициировать новый цикл создания знания, распространяясь в организации как по горизонтали, так и по вертикали. На межорганизационном уровне знание, созданное организацией, способно посредством активного взаимодействия мобилизовать знания в других дочерних компаниях, в клиентской среде, в среде поставщиков и конкурентов и в других компонентах среды. Например, введение нового способа бюджетирования в одной компании неизменно приводит к изменениям в бюджетировании дочерних компаний, а это, в свою очередь, может вызвать новый цикл инноваций. Положительный опыт рано или поздно станет достоянием других фирм, в том числе конкурентов.

Для эффективности процесса в этой фазе особенно важна автономность структурных единиц и отделов, позволяющая им поглощать и применять знания внешнего происхождения. Избыточность информации, ротация персонала здесь также весьма желательны. Все это может способствовать перекрестному обогащению знаниями в компании.

2.3. Методология CommonKADS

Методология CommonKADs [Schreiber et al, 2000] предоставляет структурированный подход к разработке систем на основе знаний, путем разработки набора моделей. Рис. 2.4 представляет состав CommonKADS моделей, которые подразделяются на три группы, так как есть три типа вопросов, на которые необходим ответ:

1. **Почему?** Почему система на основе знаний – потенциальная помощь или необходимое решение? Для каких проблем? Какие преимущества, затраты и влияния на организацию она имеет? Понимание организационного контекста и среды здесь – самый важный вопрос.

2. **Что?** Какова природа и структура вовлечённых знаний? Какова

природа и структура соответствующих связей? Концептуальное описание знаний, применяемых в задаче, - главная проблема этого вопроса.

3. **Как?** Как знания должны внедряться в компьютерную систему? Как выглядит программная архитектура и вычислительные механизмы? Технические аспекты компьютерной реализации – главный фокус вопроса.

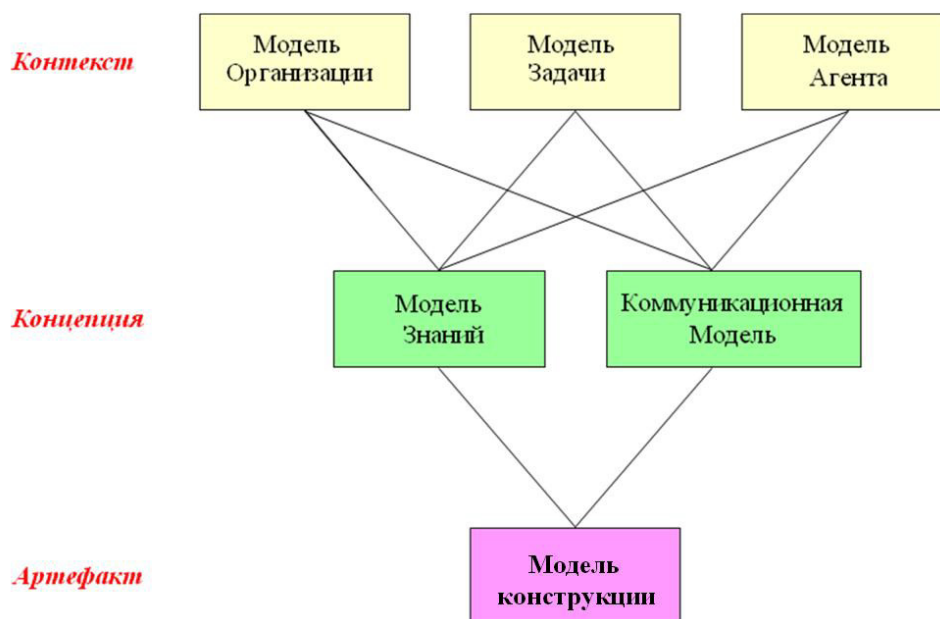


Рис. 2.4. Состав CommonKADs моделей

На все эти вопросы можно ответить путем разработки аспектных моделей. CommonKADS имеет определенный набор моделей, каждая из которых фокусируется на частном аспекте, но все вместе они дают полное представление:

- **Организационная модель.** Организационная модель поддерживает анализ основных характеристик организации для того, чтобы выявить проблемы и возможности для внедрения систем на основе знаний, определить осуществимость, и оценить воздействия на организацию потенциальных действий, воздействующих на знания.
- **Модель задачи.** Задачи – это релевантные составляющие бизнес-процесса. Модель задач анализирует место задачи в общей структуре деятельности, ее входы и выходы, предусловия и критерий эффективности, а также необходимые ресурсы и компетенции.

- **Модель агента.** Агенты – это исполнители задачи. Агентом может быть человек, информационная система или какая-либо другая сущность, способная выполнить задачу. Модель агента описывает характеристики агентов, в частности, их компетенции, полномочия на выполнение действий и ограничения в этом отношении. Кроме того, она включает коммуникативные связи между агентами при выполнении задачи.

- **Модель знаний.** Назначение модели знаний – описать в деталях типы и структуры знаний, используемых при выполнении задачи. Она предоставляет понятным человеку образом независимое от реализации описание роли, которую разные компоненты знаний играют в решении проблемы. Это делает модель знаний важным средством коммуникации с экспертами и пользователями по поводу аспектов решения проблемы системой на основе знаний, как во время разработки, так и во время работы системы.

- **Коммуникационная модель.** Поскольку несколько агентов может быть вовлечено в задачу, важно смоделировать коммуникативные транзакции между ними. Это делается посредством коммуникационной модели, концептуальным и независимым от реализации образом, так же как и модель знаний.

- **Модель конструкции (Design model).** Вышеупомянутые CommonKADS-модели в комплексе могут быть рассмотрены как составление спецификаций требований для системы на основе знаний, разбитых на различные аспекты. Основываясь на этих требованиях, модель конструкции предоставляет спецификацию технической системы в части архитектуры, внедряемой платформы, модулей программного обеспечения, репрезентативных конструкций и вычислительных механизмов, необходимых для внедрения функций, заложенных в основу модели знаний и коммуникационной модели.

Организационная модель, модель задачи и модель агента вместе анализируют организационную среду и соответствующие критические факторы успеха для системы на основе знаний. Модель знаний и коммуникационная модель разрабатывают концептуальное описание функций и данных решения проблемы, которые должны быть обработаны и предоставлены системой знаний. Модель конструкции преобразует всё

это в техническую спецификацию, являющуюся основой для внедрения программной системы. Однако стоит заметить, что не всегда необходимо разрабатывать все модели. Это зависит от целей проекта, а также от опыта, полученного во время выполнения проекта.

Для разработки СУЗ основной интерес представляет анализ контекста организации и задачи. Рассмотрим данный вопрос более подробно.

Модель контекста для определения требований к СУЗ

Методология CommonKADS предоставляет набор средств для достижения следующих важных целей:

Определения проблем и возможностей: *Поиск перспективных направлений, в которых система на основе знаний или иное решение в области управления знаниями может создать добавленную ценность в организации.*

Выбора решений и оценке их осуществимости: *Понимание, имеет ли проект смысл с точки зрения затрат и выгод, технологических возможностей, и необходимых ресурсов и обязательств в рамках организации.*

Улучшение решаемых задач и связанных с ними знаний: *Понимание природы задач, участвующих в выбранных бизнес-процессах с точки зрения знаний, используемых агентами при их выполнении, и того, какие улучшения могут быть достигнуты в этом отношении.*

Планирование необходимых организационных изменений: *Определение того, какие последствия внедрения предлагаемой системы на основе знаний проявятся в различных аспектах деятельности организации, а также подготовка плана действий, для принятия необходимых организационных мер.*

Анализ организации и задачи непосредственно связан с вышеуказанными целями.

Основные этапы анализа организации и задачи

Основными этапами анализа организации и задачи являются следующие:

Этап 1: Определение области применения и технико-экономическое обоснование, состоящего из двух частей:

- Выявление проблем / возможностей и потенциальных решений, и

рассмотрение их в широкой перспективе организации;

- Определение экономических, технических и проектных возможностей с целью выбрать наиболее перспективное направление и решение.

Этап 2: Определение влияния и улучшений для выбранного целевого решения, также состоящего из двух частей:

- Определение взаимосвязей между задачей, участвующими агентами, и используемыми знаниями для успешной работы, а также выявление улучшений, которые могут быть достигнуты здесь,

- Принятие решений об организационных мерах и изменениях с целью принятия и интеграции программного решения на основе знаний в деятельность организации.

Общий процесс анализа контекста организации и задачи

Для внедрения в организации некоторого решения по управлению знаниями нужно провести ее анализ, построить модель. Естественно, невозможно описать организацию полностью, к тому же в этом нет необходимости. Задача заключается в том, чтобы описать организацию с точки зрения знаний (Knowledge orientation).

Таким образом, идея организации модели CommonKADs заключается в том, чтобы собрать соответствующие элементы и накопленный опыт из различных источников – в том числе теории организации, анализа бизнес-процессов, управления информацией, – и интегрировать их в единый и всеобъемлющий пакет, ориентированный на использование знаний в деятельности организации.

Формы (Worksheets) и последовательность их создания показана на рис. 2.5. На этапе определения сферы применения и технико-экономического обоснования строится модель организации (МО), описывается и анализируется среда (контекст, миссия, цели, стратегия, окружение и др.) организации. На этапе определения влияния и улучшений строится модель задачи (МЗ) и модель агента (МА). На этом этапе организация описывается более детально, внимание фокусируется на наиболее интересных, с точки зрения выбранной проблемы, частях организации. Модель задачи (МЗ) описывает задачи, а также знания, непосредственно связанные с ней. Модель агента (МА) описывает агентов, участвующих в модели задач.

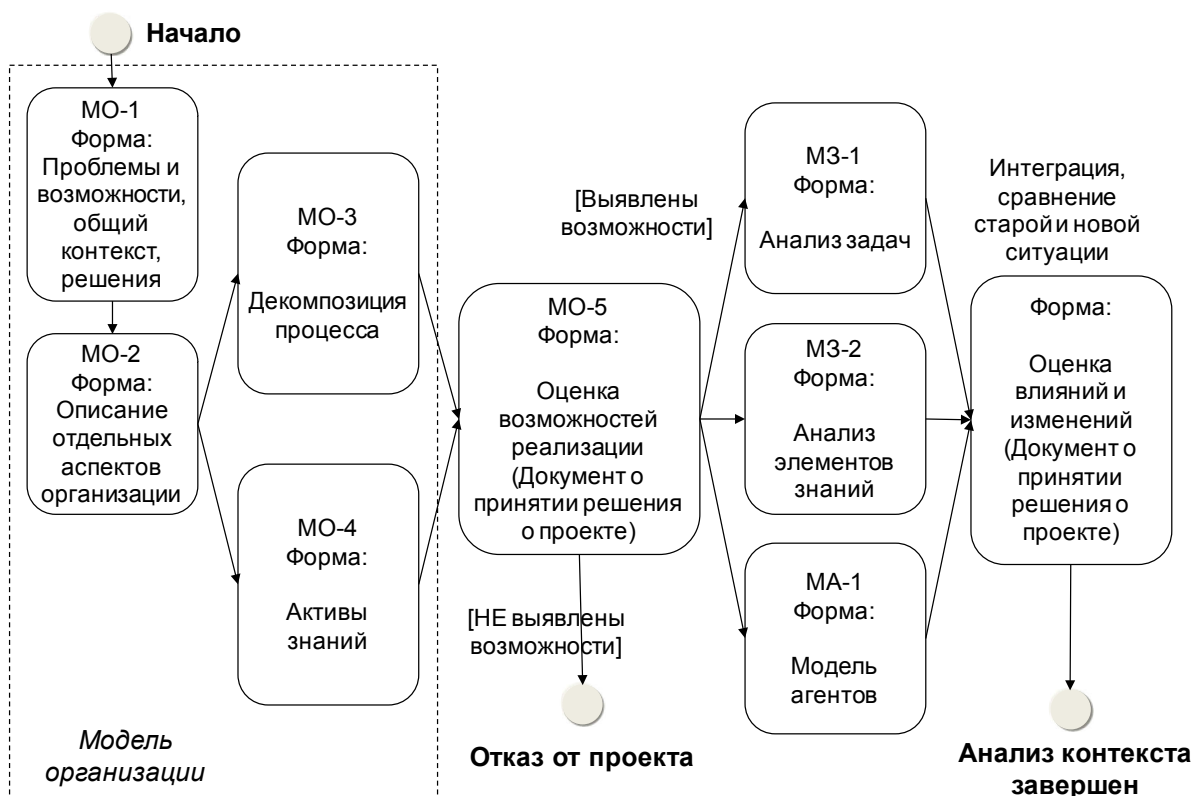


Рис. 2.5. Процесс анализа контекста организации и задачи

Моделирование процесса предполагает создание рабочих форм: МО: 5 рабочих форм; МЗ: 2 рабочие формы; МА: 1 рабочая форма; 1 итоговая модель: 1 форма – документ о принятии решения о проекте.

2.4. Методология DECOR

Путем слияния некоторых элементов из методов CommonKADS и IDEF5 методология DECOR² предоставляет методологическое руководство для ведения проекта по УЗ, ориентированному на бизнес-процесс, и включает следующие шаги:

Шаг 1: Идентификация бизнес-процесса. Эта операция включает в себя идентификацию самого важного процесса/ов, который/ые должен будет поддерживаться проектом по УЗ, ориентированному на бизнес-

² Abecker A., «Business process oriented Knowledge Management – Concepts, Methods and Tools», Phd thesis, 2003.

процесс. Бизнес-процессы с высокой сложностью и сильной интенсивностью знаний – те процессы, которые имеют высокий потенциал поддержки УЗ. Если они также являются центральными для создания целей организации, они становятся кандидатами для проекта по УЗ, ориентированному на бизнес-процесс.

Шаг 2: Анализ бизнес-процесса. Это действие включает в себя общее описание выбранного бизнес-процесса/ов с позиции (а) задач, составляющих бизнес-процесс; (б) вовлеченных ролей; и (в) ключевых людей и исходного материала.

Шаг 3: Анализ задачи. Это действие включает в себя более детализированное описание индивидуальных задач, например, их входные и выходные объекты. Контроль связей между задачами, и т. д. Более того, каждая задача в процессе оценивается через свой вклад в центральные действия по УЗ, т. е. генерирование, сохранение, распространение и применение знаний. Это может привести к отнесению некоторых задач к классу связанных со знаниями и/или наукоемких.

Шаг 4: Проектирование бизнес-процесса. Это действие включает в себя моделирование выбранного бизнес-процесса при помощи графического инструмента. Итогом этого шага должна стать модель бизнес-процесса, расширенная задачами УЗ для потока знаний в бизнес-процессе.

Шаг 5: Создание онтологии. В DECOR–методологии онтологии используются в качестве комплексной модели для структурирования индексов, используемых для описания содержания документа. Создание онтологии DECOR-методологии основывается на трех центральных концепциях: виды (объектная категория сущностей, разделяющих ряд свойств), характеристики (свойства, принадлежащие виду) и отношения (общие особенности, которые виды проявляют совместно, а не индивидуально).

Шаг 6: Улучшение/Совершенствование онтологии. Это действие включает в себя усовершенствование и утверждение разработанной онтологии.

Применение: ИКА (Греческий институт соцобеспечения)

Методология DECOR успешно применялась для бизнес-процесса предоставления полной пенсии по старости ИКА (Греческий институт

социального обеспечения). Зона обслуживания пенсий – сфера высокой степени значимости для ИКА, при этом также характеризующаяся как весьма наукоемкая. Пенсионный процесс имеет высокую значимость из-за большого числа бенефициаров, которое на момент внедрения методологии DECOR (2002) составляло 1 000 000 человек и возрастает с годовым уровнем в 10 %. К тому же пенсионные процессы высокоинтенсивны, поскольку требуют глубоких знаний соответствующего законодательства – во-первых, для составления решения, имеет ли застрахованное лицо право на получение пенсии, и, во-вторых, для подсчета размера пенсии. Довольно часто для одного конкретного случая могут подходить несколько правовых регулирований, и выявить все эти правила, а затем и выбрать среди них подходящее, будет вопросом знаний.

Система, установленная в ИКА и реализовавшая результаты применения DECOR-методологии для вышеупомянутых бизнес-процессов, привела к следующим результатам на количественном уровне: увеличению числа решений, принимаемых за день; увеличению процентного отношения количества решений, принимаемых в неделю, к количеству поданных в неделю заявок; предполагаемому сокращению процента обращений к решениям, принятым ИКА;

С другой стороны, на качественном уровне интервью с персоналом ИКА дали следующие ожидаемые преимущества: возможное использование инструмента в целях обучения новых работников; более согласованное рассмотрение случаев, так как все работники могут обращаться к одним и тем же правилам; более быстрое распространение новых правил, лучших упражнений или другой новой информации; основные преимущества от автоматизации процесса вместо ручной обработки документов.

2.5. Методология Карла Виига

Карл Вииг с коллегами [Wiig et al., 1997] различает уровень управления и уровень объектов знания, подобно управлению в любых других областях.

Уровень объектов знаний представлен на рис. 2.6. Его составляют три основных компонента: агенты, бизнес-процессы, активы знаний.

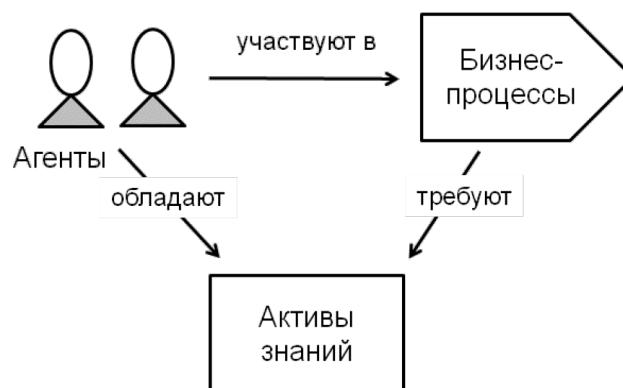


Рис. 2.6. Объекты управления знаниями

Действия по УЗ воздействуют на характеристики и взаимосвязи указанных компонент (например, в результате действий какие-либо активы знаний переводятся в другую форму, а какими-либо активами знаний начинают обладать новые агенты).

Уровень управления представляет собой цикл управления знаниями (рис. 2.7):

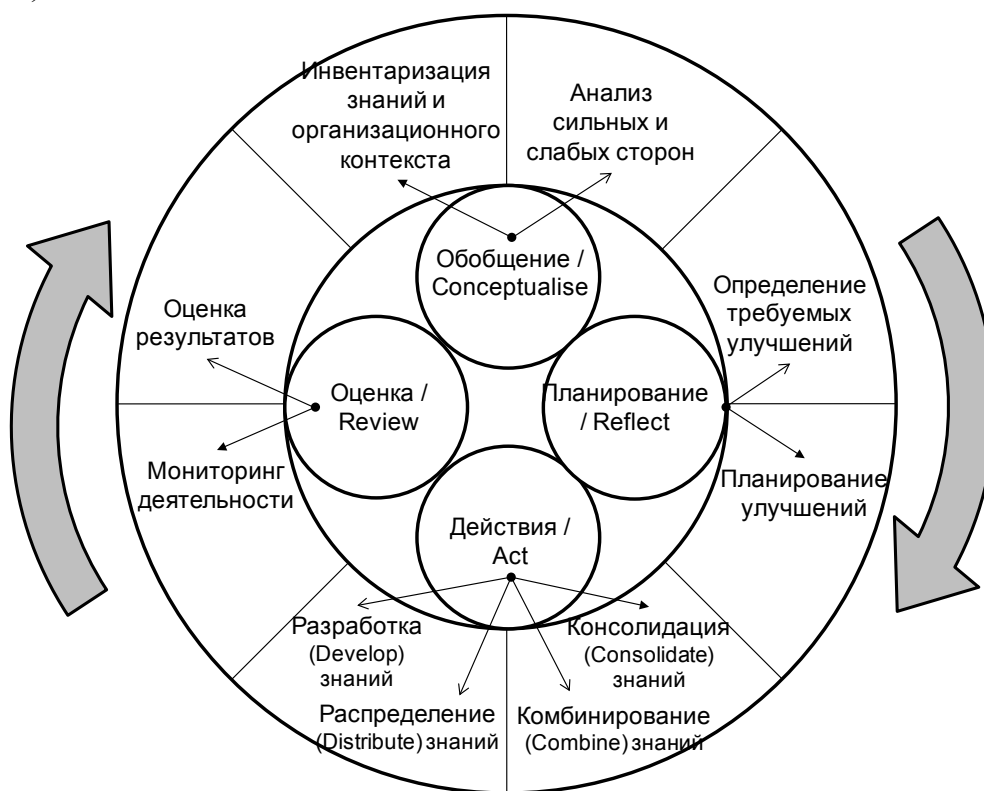


Рис. 2.7. Цикл управления знаниями

В нем Вииг выделяет четыре фазы: оценка (Review), обобщение (Conceptualise), планирование (Reflect), действия (Act). Каждая фаза детализируется более конкретными действиями – см. внешний круг. Рассмотрим фазы и составляющие их действия более подробно.

Оценка / Review

Хотя фаза оценки принята за начальную точку, это не значит, что на практике каждый цикл УЗ будет начинаться с нее. Выбор ее в качестве начальной точки удобен лишь для описания методологии. Фаза оценки состоит из 2-х составляющих: мониторинг деятельности и оценка результатов.

Мониторинг деятельности

Мониторинг эффективности деятельности организации с точки зрения УЗ требует наличия соответствующих процедур. Эти процедуры зависят от типа выбранных ранее метрик. В большинстве практических случаев метрики будут связаны с проектами улучшений, запущенными на фазе планирования. Однако не только проекты улучшений являются объектом мониторинга. Знания и связанные с ними объекты также должны быть объектами мониторинга. Особенно важно наблюдать за внешней средой и новыми событиями, которые могут повлиять на работу со знаниями.

Оценка результатов

Результаты мониторинга следует оценивать с точки зрения поставленных целей: привели ли реализованные планы улучшений к ожидаемым результатам? Для этого необходимо учитывать стратегию организации, а также стратегию управления знаниями.

Обобщение (Conceptualise)

Инвентаризация знаний и организационного контекста

Одним из основных элементов УЗ в организации является получение представления об имеющихся в организации знаниях. Это предполагает ответы на вопросы: в каких бизнес-процессах используются знания, какие знания используются, где и когда они используются, какие организационные роли предоставляют знания. Данные вопросы соответствуют рассмотренным выше объектам УЗ (рис. 2.6).

Анализ сильных и слабых сторон

Анализ сильных и слабых сторон организации с точки зрения УЗ может выполняться разными методами, например, с помощью SWOT-анализа или анализа узких мест. Название SWOT-анализ, объединяет первые буквы четырех английских слов – Strength – сила, Weakness – слабость, Opportunities – возможности, Threats – угрозы. Данный экспертный метод предполагает выявление сил и слабостей компании с точки зрения выбранной цели, а также сопоставление их с существующими возможностями и угрозами во внешней среде. Анализ узких мест состоит в анализе организации с точки зрения наличия или отсутствия в ней типовых узких мест, относящихся к знаниям. Примеры типовых узких мест – недостаток знаний в местах их приложения, ограниченное повторное использование готовых знаний и т. п.

Планирование (Reflect)

Основная цель фазы планирования – разработка планов улучшений, которые с высокой вероятностью приведут к успеху. Вииг разделяет улучшения и планы (внедрения) улучшений. Для внедрения улучшений необходимо выполнять комплекс действий, которые учитывают организационные и прочие препятствия. На практике, обычно, требуется несколько итераций для согласования состава улучшений и плана их внедрения. Улучшения направлены на повышение ценности активов знаний для организации. Планы же направлены на минимизацию рисков.

Определение требуемых улучшений

В результате фазы обобщения идентифицированы узкие места, проблемы, возможности, слабые стороны и т. п. Улучшения направлены на устранение проблем и узких мест, с одной стороны, и на реализацию возможностей, с другой. Кроме того, не все предлагаемые улучшения могут быть выполнены в один момент и существуют финансовые ограничения. Поэтому для предлагаемых улучшений следует определять приоритет. После этого происходит отбор наиболее приоритетных улучшений. Для этого могут использоваться методы принятия решений, например, многокритериальная оценка альтернатив и т. п.

Планирование улучшений

После того, как состав улучшений определен, их необходимо преобразовать в операционные планы. Обычно для реализации улучшений

необходимо запустить несколько проектов, поэтому операционные планы будут соответствовать планам этих проектов. Внимание же следует уделить стандартным объектам управления проектами, таким как время, бюджет, результаты проектов, человеческие и другие ресурсы, качество, команда проекта и распределение ответственности. Особое внимание следует уделять рискам проектов. Те планы улучшений, которые имеют высокую вероятность наступления рисков, а также высокую опасность их последствий должны быть пересмотрены, даже в случае высокой ожидаемой выгоды от улучшения.

Действия (Act)

Фаза действий предполагает выполнение планов улучшений. В соответствии с предлагаемой концепцией деятельность данной фазы не относится к УЗ. Она относится к самостоятельным областям, в которых есть свои методы, техники и инструменты:

Управление персоналом. Многие улучшения непосредственно связаны с сотрудниками организации. Вот некоторые из них: схемы оплаты труда, обучение, повышения, подбор персонала, мобильность персонала.

Информационные технологии. Они предоставляют целый ряд техник и инструментов для работы со знаниями. Неполный список: системы, основанные на знаниях, базы данных, машинное обучение, системы управления потоками работ, системы поддержки принятия решений и т. п. Для разработки таких приложений существуют специальные методы.

Организационное развитие. Часто необходимо пересмотреть способ организации работы. Реинжиниринг бизнес-процессов – типичный подход в данной области.

2.6. Методология On-To-Knowledge

Эта ориентированная на процесс методология [Staab et al., 2001] была разработана в Университете Карлсруэ в рамках одноименного проекта (<http://www.ontoknowledge.org>), использовавшего онтологию для улучшения качества управления знаниями в больших и распределенных организациях. У истоков ОТК лежит CommonKADS [Schreiber et al., 2000], известная методология в области разработки систем, основанных на

знаниях.

Методология ОТК предлагает разделять два процесса: первый – внедрение УЗ-решения в деятельность организации и его сопровождение (называемый авторами мета-процессом, «Knowledge Meta Process»), второй – это использование внедренного УЗ-решения (так называемый, цикл (преобразования) знаний, «Knowledge Process») (рис. 2.8).

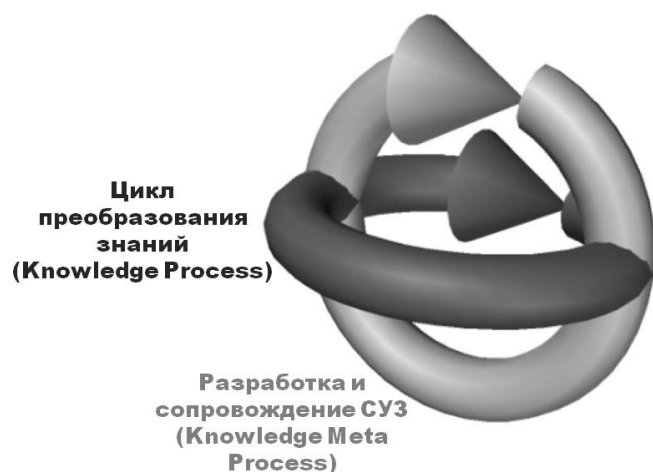


Рис. 2.8. Два ортогональных циклических процесса в области УЗ

Внедрение УЗ в организации (движение по мета-процессу, серое кольцо на рис. 2.8), обычно, идет по простому прагматичному пути работы с документами. Данный подход, как правило, позволяет собирать лишь «низковисящие плоды». Его недостатком является то, что отдельные этапы процесса преобразования знаний (рис. 2.9) – создание, импорт, получение, поиск/доступ и использование – очень слабо связаны, если, вообще, связаны. Причина этого в том, что для каждого этапа процесса преобразования знаний важны разные документы. Такая ситуация делает повторное использование знаний крайне затруднительным.



Рис. 2.9. Цикл преобразования знаний

Получение знаний, самый нагруженный по составу работ этап в цикле преобразования знаний. Он включает в себя аннотирование / формирование метаданных, извлечение неявных знаний сотрудников, получение знаний путем анализа документов (Text Mining) и баз данных (Data Mining), структурирование / классификацию. Сюда же попадает формирование организационной памяти, интеграция и хранение знаний. Использование знаний предполагает, что доступное знание используется сотрудниками для выполнения своей работы более результативно и эффективно, вновь созданное знание влияет на конкретные действия.

В основе процесса разработки и сопровождения СУЗ лежит разработка онтологии (будет подробно рассмотрена в следующем разделе). Данный процесс включает в себя пять стадий, с четко определенными входами и выходами на каждой из них (рис. 2.10): изучение осуществимости проекта (feasibility study), начало (kickoff), уточнение (refinement), оценка (evaluation), поддержка и эволюция.

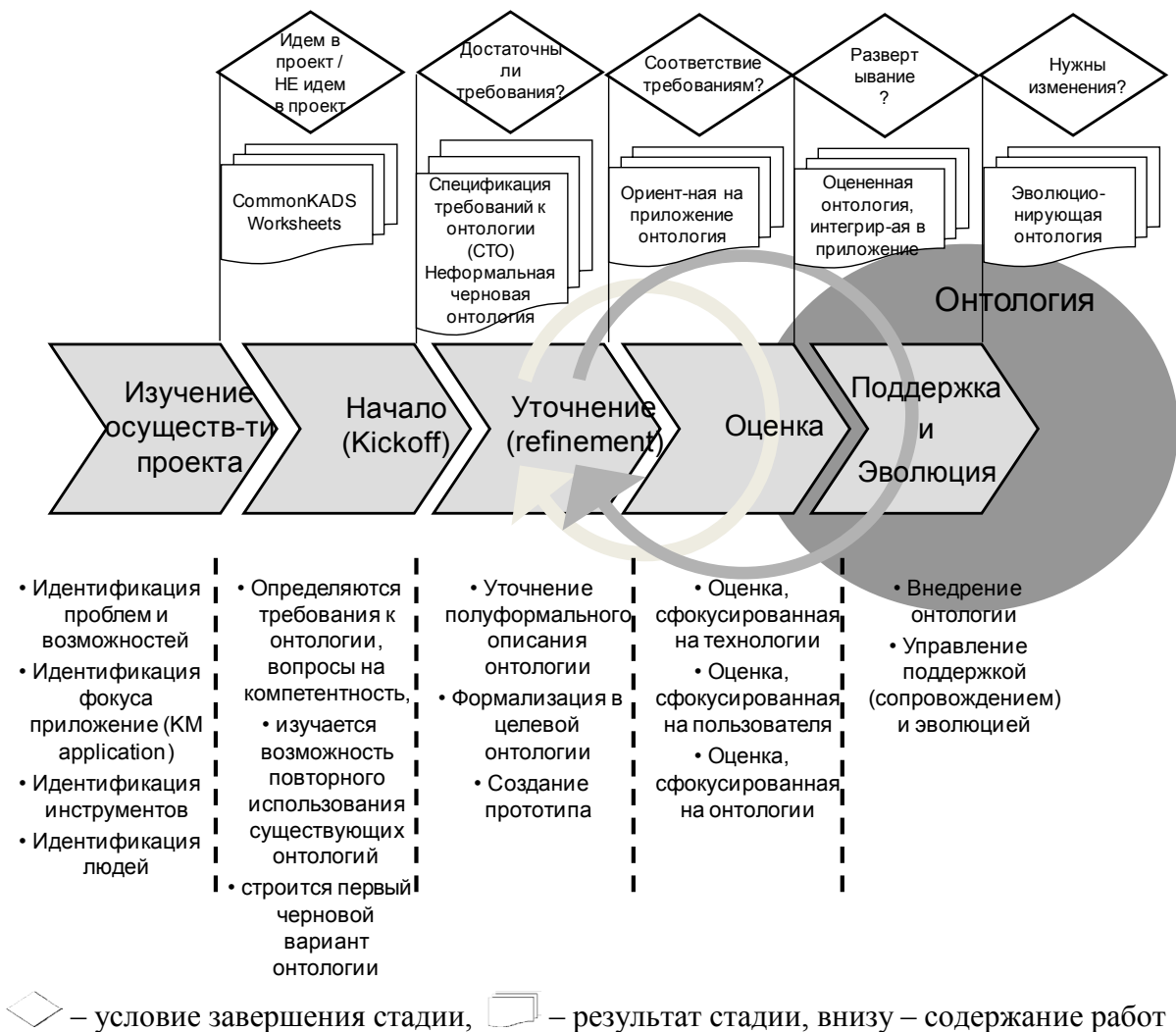


Рис. 2.10. Процесс разработки и сопровождения СУЗ

3. СЕМАНТИЧЕСКИЙ ПОДХОД К УПРАВЛЕНИЮ ЗНАНИЯМИ И ОНТОЛОГИИ

3.1. Потребность в общем языке

«Люди не могут обмениваться знаниями, если у них нет общего языка»
Том Дейвенпорт

Для одного и того же объекта реального мира различные агенты могут иметь различные абстрактные модели (рис. 3.1), что препятствует обмену знаниями между ними.

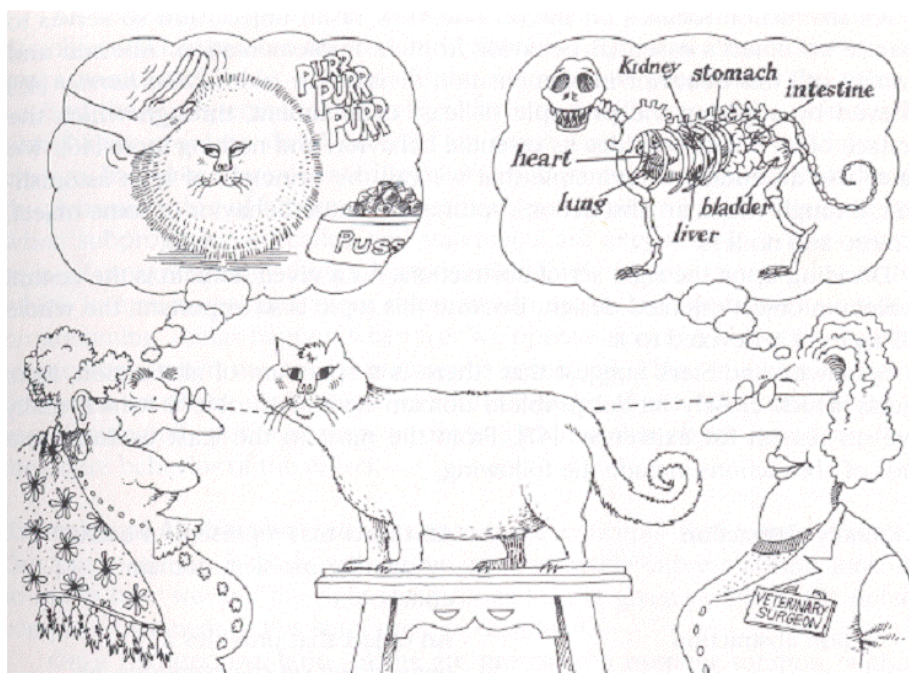


Рис. 3.1. Различие в абстрактных моделях разных субъектов при рассмотрении одного и того же объекта

Для понимания роли «общего языка» в УЗ рассмотрим роль знаковой системы в коммуникации.

Обычно коммуникацией называют: а) передачу информации, т. е. сообщение; б) взаимодействие, т. е. общение. Она может осуществляться в

процессе любой деятельности с помощью различных знаковых способов, например, речи, жестов, азбуки Морзе и др. Изучением таких знаковых систем занимается семиотика.

Семиотика, или семиология (от греч. Semeiotikon, от semeion – знак, признак) – наука, исследующая свойства знаковых систем и соответственно знаков (естественных и искусственных). С одной стороны, семиотика близка к кибернетике, исследующей процессы связи и управления в живом организме, природе и обществе, а с другой – к лингвистике, изучающей человеческий язык.

Знаковая система – это совокупность знаков, устроенная определенным образом. Знаковая система есть материальный посредник, который служит для обмена информацией между двумя другими материальными системами (например, человек – знаковая система – человек, человек – знаковая система – компьютер).

Три основных аспекта изучения знака и знаковой системы, выделяемые семиотикой:

1) синтактика, изучающая отношения между знаками, т. е. **внутренние свойства систем знаков** (иначе, правила построения знаков в рамках знаковой системы);

2) семантика, изучающая отношения между знаками и обозначаемым предметом, т. е. **содержанием знаков**;

3) прагматика, изучающая отношения между знаком и человеком-пользователем знака: говорящим, слушающим, пишущим, читающим, т. е. **назначение знаков**.

Примером знаков могут служить широко используемые в операционных системах ярлыки-иконки, например, ярлык с изображением принтера. Его изображение на экране является именем знака. Содержание знака может быть раскрыто с помощью перехода от имени к таблице, содержащей все нужные сведения о принтерах, а включение процедуры печати при щелчке по ярлыку принтера определяет назначение этого знака.

Синтактика изучает объективные законы устройства знаковых систем. Ее задачей является описание запаса правильно построенных текстов (составных знаков) для различных классов знаковых систем. В случае естественного языка формальное описание синтактики фрагментарно (например, в русском языке множество исключений из

правил) вследствие избыточности и динамичности естественного языка. Для искусственных языков синтактика, как правило, описывает все допустимые тексты. Для этого в теории формальных языков используется грамматика – любой конечный механизм задания языка.

Семантика изучает законы смысла. Центральным понятием семантики является понятие значения. Обычно для разъяснения структуры языкового значения используется «семантический треугольник» или «треугольник Фреге».

Строение знака – треугольник Фреге (рис. 3.2):

I. Вещь, предмет, явление действительности и т. д. Иное название – денотат,

II. Знак: в лингвистике, например, фонетическое слово или написанное слово; в математике – математический символ; иное название, принятое особенно в философии и математической логике, – имя.

III. Понятие о предмете, вещи. Иные названия: в лингвистике – сигнификат, десигнат, в математике – смысл имени, или концепт денотата.

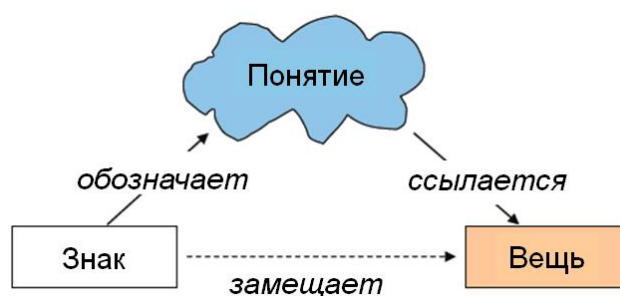


Рис. 3.2. «Семантический треугольник» или «треугольник Фреге»

Например, слово «стул» обозначает класс реально существующих предметов. Все существующие виды стульев отображаются в нашем сознании в виде понятия о стуле. Значение слова стул содержит минимальный набор характерных признаков (признаков понятия), позволяющих соотнести слово и обозначаемый объект.

Возможные расширения «треугольника Фреге»:

- Иногда вершиной треугольника обозначают не саму вещь (денотат), а ее восприятие или представление о ней, словом ее отражение в сознании человека, которое уже, в свою очередь, связано с

денотатом³;

- Иногда в схему вводится образ в качестве вершины, для разделения понятия о предмете на две части: логическую («левополушарную») и образную («правополушарную»)⁴.

В «треугольнике Фреге» часто возникает неоднозначность соответствия знак – понятие – денотат, выражающаяся в омонимии, полисемии и синонимии.

Омонимия (рис. 3.3) состоит в том, что существует несколько совершенно одинаковых по форме знаков, каждый из которых имеет свое значение, причем эти значения между собой абсолютно не связаны, так же как и соответствующие денотаты.

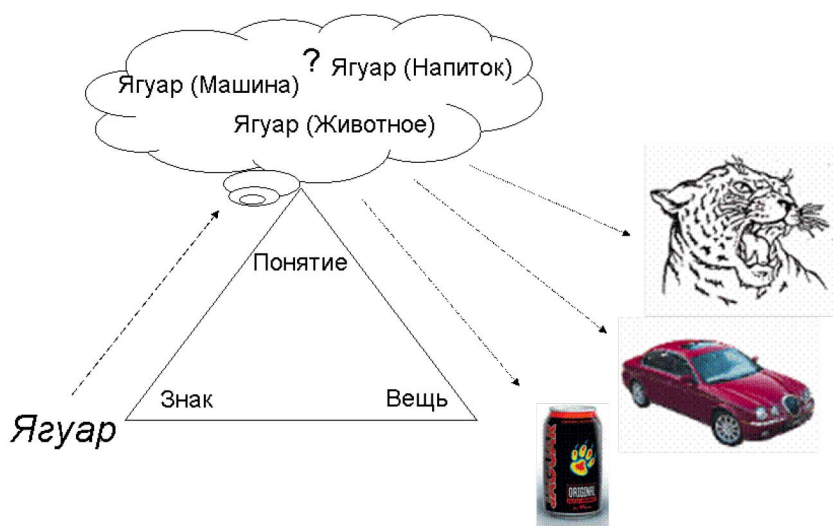


Рис. 3.3. Пример омонимии

Полисемия представляет собой наличие различных значений у одного и того же знака. Обычно данный термин применяется в ситуациях, когда эти различные значения в какой-либо мере связаны между собой (в отличие от омонимии).

³ Пospelов Д.А., Осипов Г.С. Прикладная семиотика // «Новости искусственного интеллекта» № 1. Москва 1999. С. 9-35.

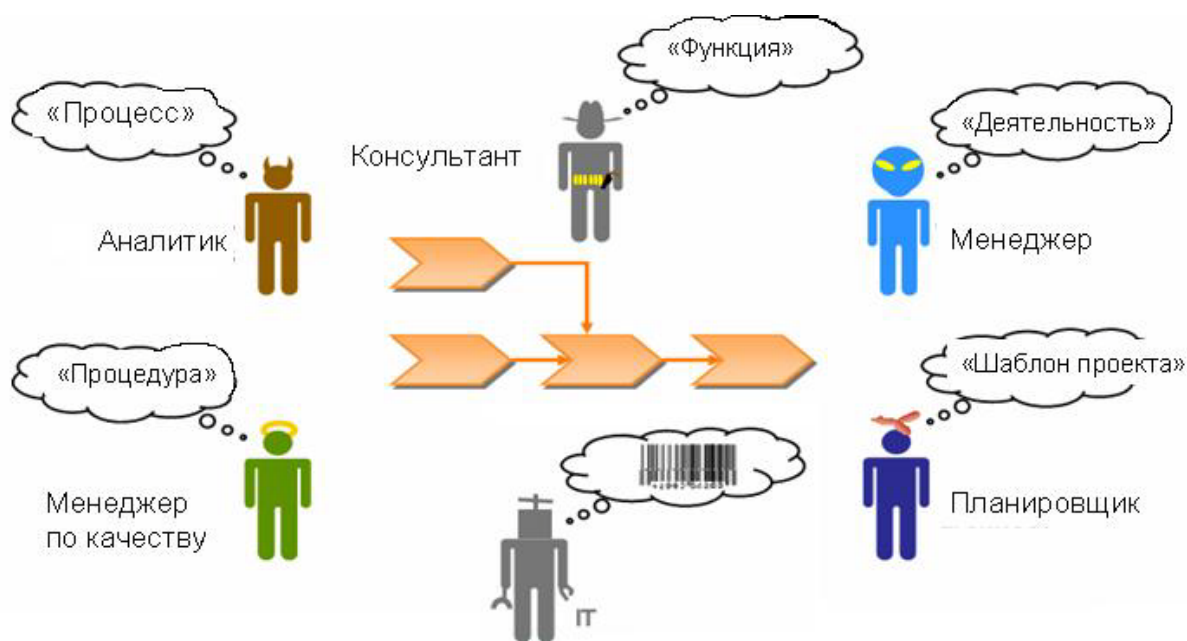
⁴ Валькман Ю.Р., Исмагилова Л.Р. О языке образного мышления // Труды Международного семинара Диалог'2004 «Компьютерная лингвистика и интеллектуальные технологии». Москва – 2004.

Пример: определения понятия для словосочетания «Капитальный ремонт» встречающиеся на одном из отечественных предприятий:

– работы по техническому обслуживанию и ремонту оборудования, выполняемые в капитальный останов, т. е. останов производственной линии длительностью свыше 24-х часов.

– ремонт, выполняемый для восстановления исправности и полного или близкого к полному восстановлению ресурса изделия с заменой или восстановлением любых его частей, включая базовые.

Синонимия (рис. 3.4) указывает на равнозначность, но не тождественность знаков. Под равнозначностью понимается: 1. либо соотнесенность с одним и тем же денотатом (предметом); 2. либо соотнесенность с одним и тем же понятием, точнее с той его частью, которая содержит характеризующую информацию.



По материалам компании FutureModels

Рис. 3.4. Пример распространенной синонимии в организациях

Для снижения представленной выше неоднозначности соответствия знак – понятие – денотат необходим «общий язык», включающий в себя:

- Строго определенный словарь лексических единиц (знаков),

- Непротиворечивое понимание того, какие понятия обозначаются заданными лексическими единицами (знаками).

3.2. Потребность в интеграции информации

Кроме потребности в общем языке организации испытывают потребность в интеграции информации. Необходимые знания могут находиться в различных местах:

- В документах (например, отчет по маркетинговому исследованию),
- В головах сотрудников (например, неформализованные знания специалиста отдела продаж X о клиенте Y),
- В базах данных (например, формализованные знания о клиенте Y в базе данных системы управления отношениями с клиентами, CRM),
- В базах знаний (например, в аналитическом модуле системы управления отношениями с клиентами, CRM).

При появлении потребности в знаниях необходимо проводить раздельный поиск в различных источниках (рис. 3.5):

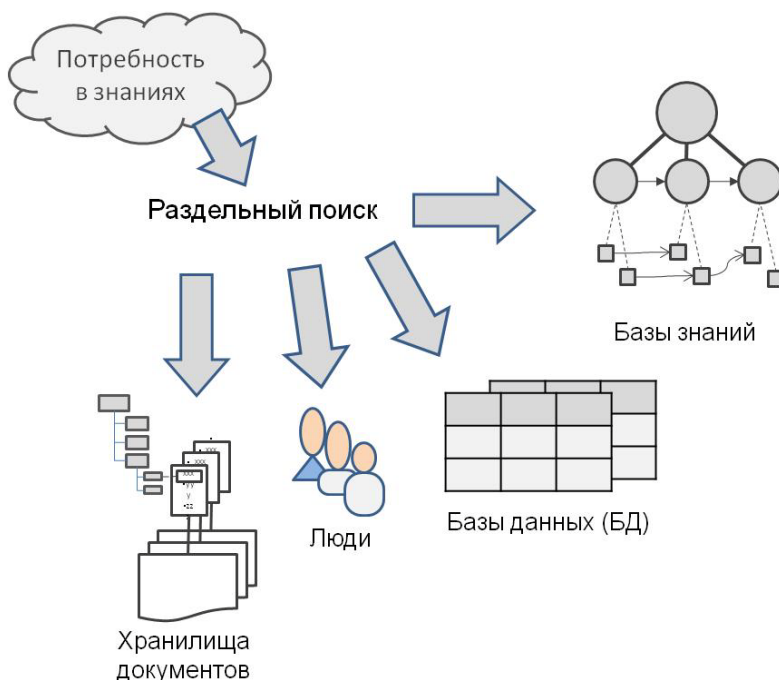


Рис. 3.5. В традиционных информационных системах потребность в знаниях нужно удовлетворять путем раздельного поиска по различным источникам

3.3. Онтология как общий язык в управлении знаниями

Онтология (от древнегреч. онтос – сущее, логос – учение, понятие) – термин, определяющий учение о сущем, бытии, в отличие от гносеологии – учение о познании. В философском смысле, а этот термин заимствован из философии, онтология есть определенная система категорий, являющихся следствием определенных взглядов на мир.

В настоящее время термин «онтология» переместился в область информационных технологий, где был использован рядом исследовательских сообществ по искусственному интеллекту вначале в области инженерии знаний, в обработке естественных языков, а затем в представлении знаний⁵. В конце 1990-х – начале 2000-х годов понятие онтологии также стало широко использоваться в таких областях, как интеллектуальная интеграция информации, поиск информации в Интернете и управление знаниями⁶. Позже онтологии стали рассматриваться в качестве ключевого элемента в проекте семантического веба – нового этапа развития сети WWW (Word Wide Web). Если существующая Web-сеть – это огромное множество документов, которые связаны перекрестными ссылками, то создаваемый семантический веб должен добавить к существующей сети множество онтологий и метаописаний знаний, содержащихся в документах Web-сети (включая стандарты и программные инструменты)⁷.

Понимание термина «онтология» зависит от контекста и целей его использования. В целом онтология, или концептуальная модель предметной области, состоит из иерархии понятий предметной области, связей между ними и законов, которые действуют в рамках этой модели. Онтология строится как сеть, состоящая из концептов и связей между ними. Связи могут быть различного типа (например, «является», «состоит

⁵ Uschold M, Jasper R. Framework for Understanding and Classifying Ontology Applications // Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5) Stockholm, Sweden, August 2, 1999.

⁶ Fensel D. Ontologies: a silver bullet for Knowledge Management and Electronic Commerce. – Berlin: Springer Verlag, 2001.

⁷ Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American, May 17, 2001. (перевод Евгения Золина http://ezolin.pisem.net/logic/semantic_web_rus.html)

из», «является исполнителем» и т. п.). Для выполнения роли общего языка будет включать в себя словарь, описывающий предметную область и набор эксплицитных (явных) интенциональных определений, задающих смысл элементов словаря на языке логической теории. Связи между словарем (знаками) и семантикой ограничивают множество возможных интерпретаций знаков. Например, после установки связи между словом «Ягуар» и интенциональным определением «хищное животное семейства кошек» мы исключаем из множества возможных интерпретаций этого слова машины и напитки. В результате онтологии позволяют правильно соотносить знаки, используемые людьми и компьютерами, с понятиями (семантическими моделями) и обозначаемыми объектами реального мира (рис.3.6).

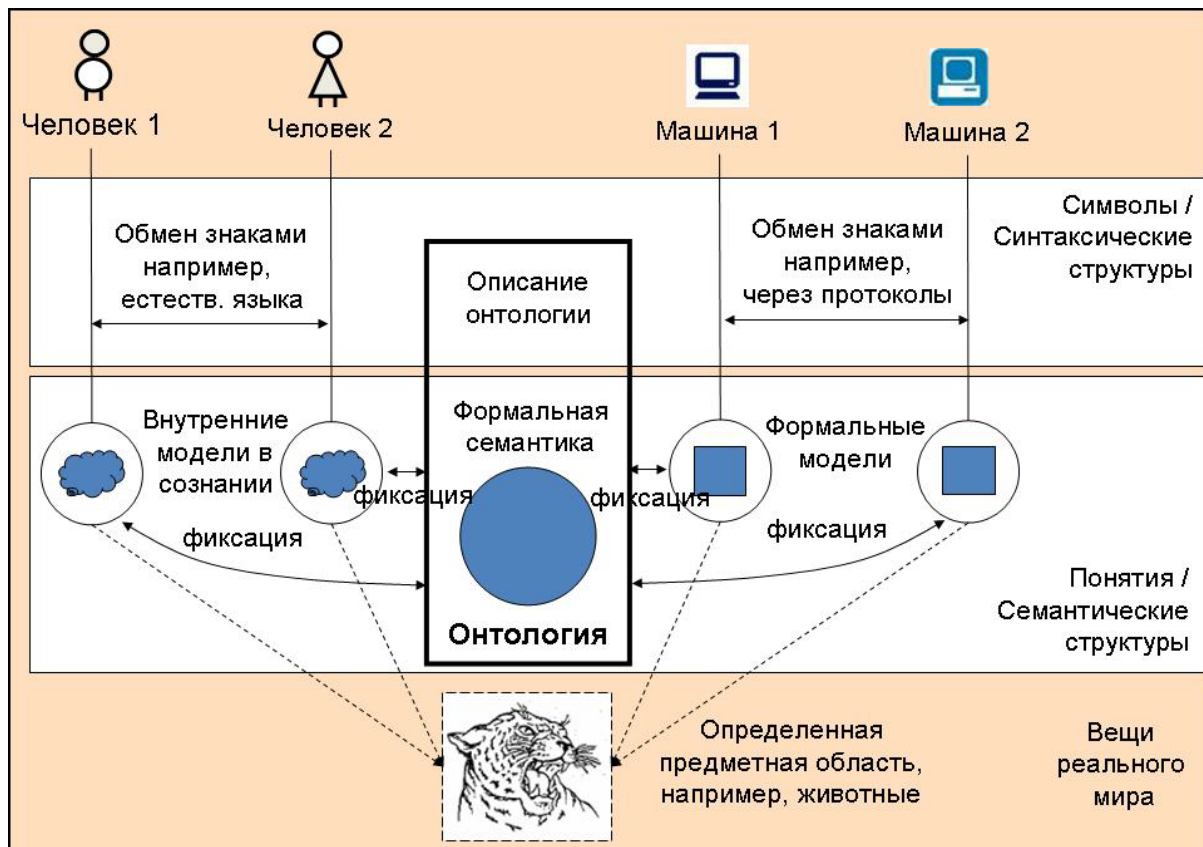


Рис. 3.6. Использование онтологий для коммуникаций между людьми и/или машинами

Штудер и его коллеги дают более развернутое определение онтологии в информатике (computer science)⁸ – это *формальная спецификация разделяемой концептуальной модели*, где

- под «концептуальной» моделью подразумевается абстрактная модель предметной области, описывающая систему понятий предметной области,
- под «разделяемой» подразумевается согласованное понимание концептуальной модели определенным сообществом (группой людей),
- «спецификация» подразумевает описание системы понятий в явном виде,
- «формальная» подразумевает, что концептуальная модель задана на формализованном языке.

Под формальной моделью онтологии O будем понимать:

$$O = \langle C, P, R, A \rangle,$$

где: C – конечное множество понятий (классов сущностей) предметной области; P – конечное множество свойств этих понятий (классов); R – конечное множество связей между понятиями (классами); A – множество аксиом, утверждений, построенных из этих понятий, их свойств и связей между ними.

Особо важным характеристическим свойством онтологии является наличие представление родовидовых отношений между классами объектов – таксономии классов. Например, в онтологии организации предложенной компанией «Бизнес Инжиниринг Групп» (<http://bigc.ru/>) таксономия понятия «деятельность» выглядит следующим образом (рис. 3.7):

⁸ Studer R., Benjamins R., Fensel D. Knowledge Engineering: Principles and Methods // Data and Knowledge Engineering, 25(1-2), 1998. P. 161-197.



Рис. 3.7. Пример таксономии классов для понятия «деятельность»

Именно, благодаря таким отношениям удается описать содержание понятия, дать его интенциональное определение. Некоторые примеры таких определений для понятий из таксономии деятельности:

Мероприятие – уникальная деятельность (родовое понятие), выполняемая для решения какой-либо задачи (характеристические признаки).

Операция – повторяющееся (характеристический признак) элементарное действие (родовое понятие), способность реализации которого поддерживается в организации;

Проект – уникальная деятельность (родовое понятие), направленная на получение определенного результата и/или достижения цели, состоящая из совокупности скоординированных и управляемых мероприятий с начальной и конечной датами, соответствующая конкретным требованиям, включающим ограничения по срокам, стоимости и ресурсам (характеристические признаки);

Процесс (деловой процесс) – регулярная деятельность (родовое понятие), представляющая собой повторяющуюся последовательность

взаимосвязанных операций, направленная на получение определенного результата (характеристические признаки);

Для пояснения роли онтологии в коммуникации можно расширить формальное определение онтологии, введя в него словарь [Maedche A., Zacharias V., 2002]:

$$O = \{C, R, A, L, F, G, H\},$$

где

- $L = L_C \cup L_R$ – словарь онтологии, содержащий набор лексических единиц (знаков) для понятий L_C и набор знаков для отношений L_R ;
- C – набор понятий онтологии;
- R – отношения между понятиями онтологии;
- A – набор аксиом онтологии;
- F и G – функции ссылок такие, которые связывают наборы лексических единиц $\{L_j\} \subset L$ с наборами понятий и отношений, на которые они соответственно ссылаются в данной онтологии. При этом одна лексическая единица может ссылаться на несколько понятий или отношений, и одно понятие или отношение может ссылаться на несколько лексических единиц;
- H – фиксирует таксономический характер отношений (связей), при котором понятия онтологии связаны нереклексивными, ациклическими, транзитивными отношениями $H \subset C \times C$. Выражение $H(C_1, C_2)$ означает, что понятие C_1 является подпонятием C_2 .

Кроме повышения эффективности коммуникаций, одним из ключевых преимуществ онтологии в СУЗ является возможность интеграции разнородной информации. Онтология задает структуру базы знаний организации, обеспечивая доступ к ее содержимому, предоставляет словарь для описания документов (контента) и профилей компетентности сотрудников, выступает общей схемой для интеграции различных баз данных (рис. 3.8). В результате, при возникновении потребности в знаниях сотруднику не требуется осуществлять поиск по разным источникам. Онтология позволит ему через один интерфейс получить доступ ко всем необходимым знаниям по требуемой теме, вне зависимости от формы представления знаний.

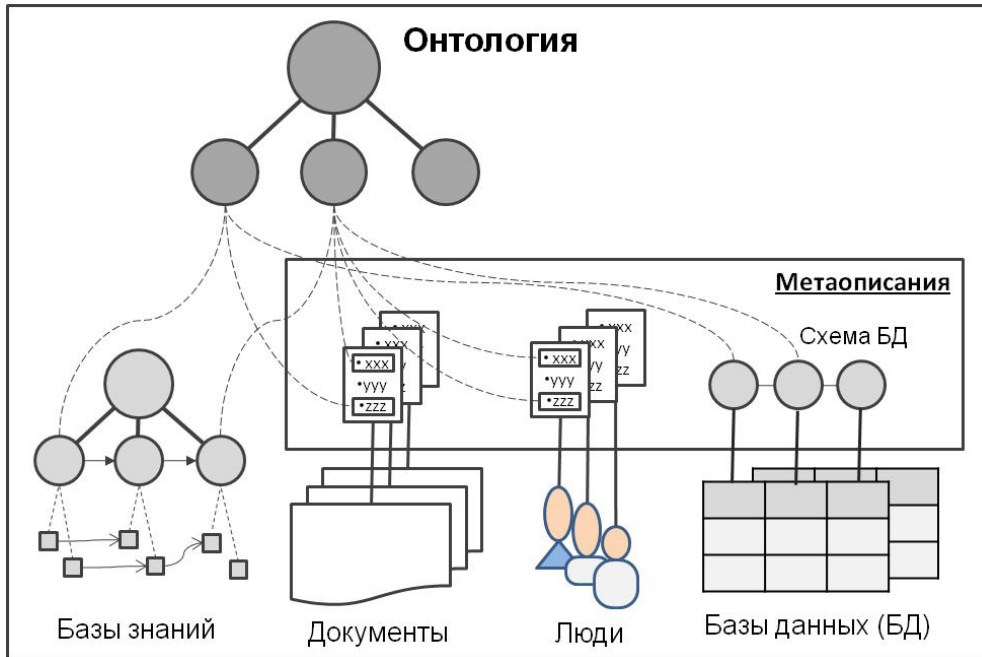


Рис. 3.8. Онтология позволяет интегрировать разнородные источники информации

3.4. Примеры онтологий

В рамках классификации по цели создания выделяют четыре уровня (см. рис. 3.9): онтология представления, онтология верхнего уровня, онтология предметной области и прикладная онтология.

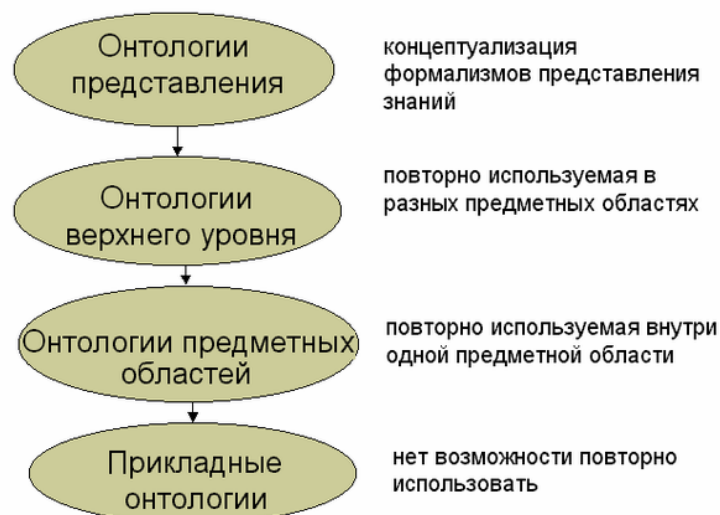


Рис. 3.9. Классификация онтологий по цели создания

Онтология представления

Цель ее создания – описать область представления знаний, создать язык для спецификации других онтологий более низких уровней. Пример: описание понятий языка OWL средствами RDF/RDFS (рис. 3.10). В данном описании определяются такие понятия, как «класс», «отношение», «ограничение на значение свойства», «домен», «диапазон» и т. п.

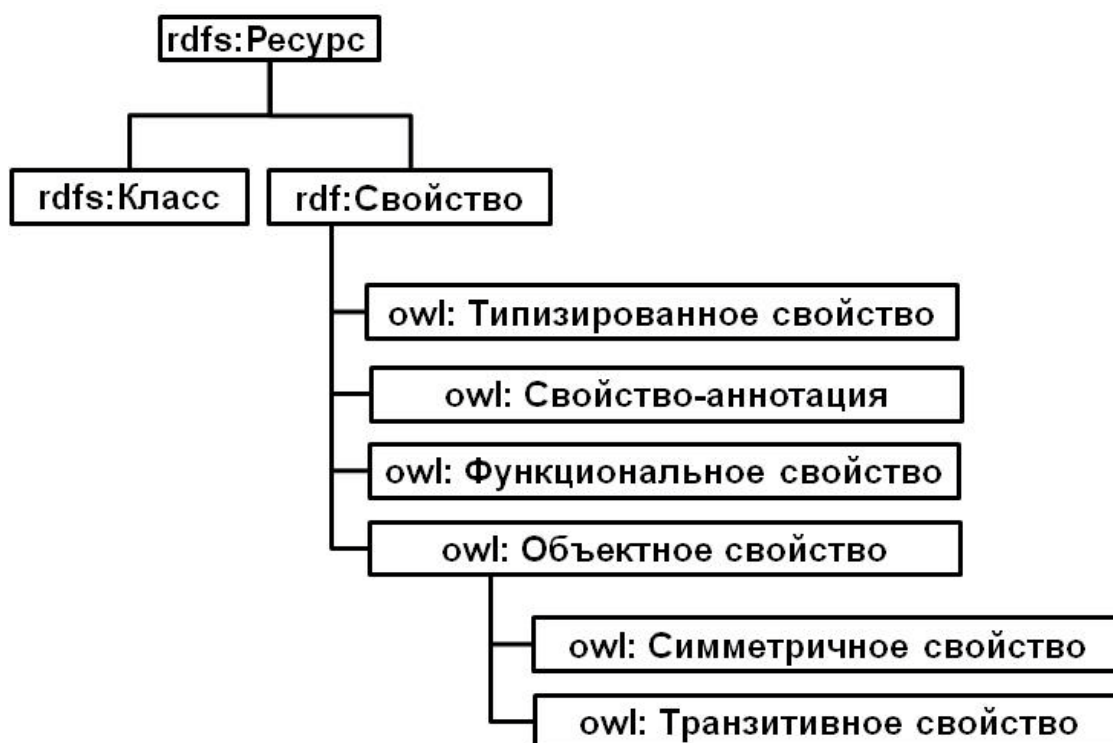


Рис. 3.10. Онтология представления для языка OWL

Онтология верхнего уровня

Ее назначение – в создании единой «правильной» онтологии, фиксирующей знания, общие для нескольких предметных областей, и в многократном использовании данной онтологии. Существует несколько крупных онтологий верхнего уровня: Сус, DOLCE, SUMO, онтология Джона Совы (J. Sowa) и другие. Но в целом попытки создать онтологию верхнего уровня на все случаи жизни пока не привели к ожидаемым результатам. Многие онтологии верхнего уровня похожи друг на друга. Они содержат одни и те же концепты: сущность, явление, процесс, объект,

роль, пространство, время, материя, событие, действие и т. п. Рассмотрим в качестве примера DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) – первую из онтологий в библиотеке базовых онтологий проекта WonderWeb (<http://www.loa-cnr.it/DOLCE.html>).

Онтологию DOLCE предполагается применять в Semantic Web для согласования между интеллектуальными агентами, использующими разную терминологию. При этом онтология не претендует на звание универсальной, стандартной или общей. Основная цель разработчиков – создать модель, помогающую при сравнении и объяснении связей с другими онтологиями библиотеки онтологий WonderWeb, а также для выявления скрытых допущений, лежащих в основе существующих онтологий и лингвистических ресурсов, таких как WordNet. DOLCE имеет когнитивный уклон, поскольку фиксирует онтологические категории естественного языка и знания «здравого смысла».

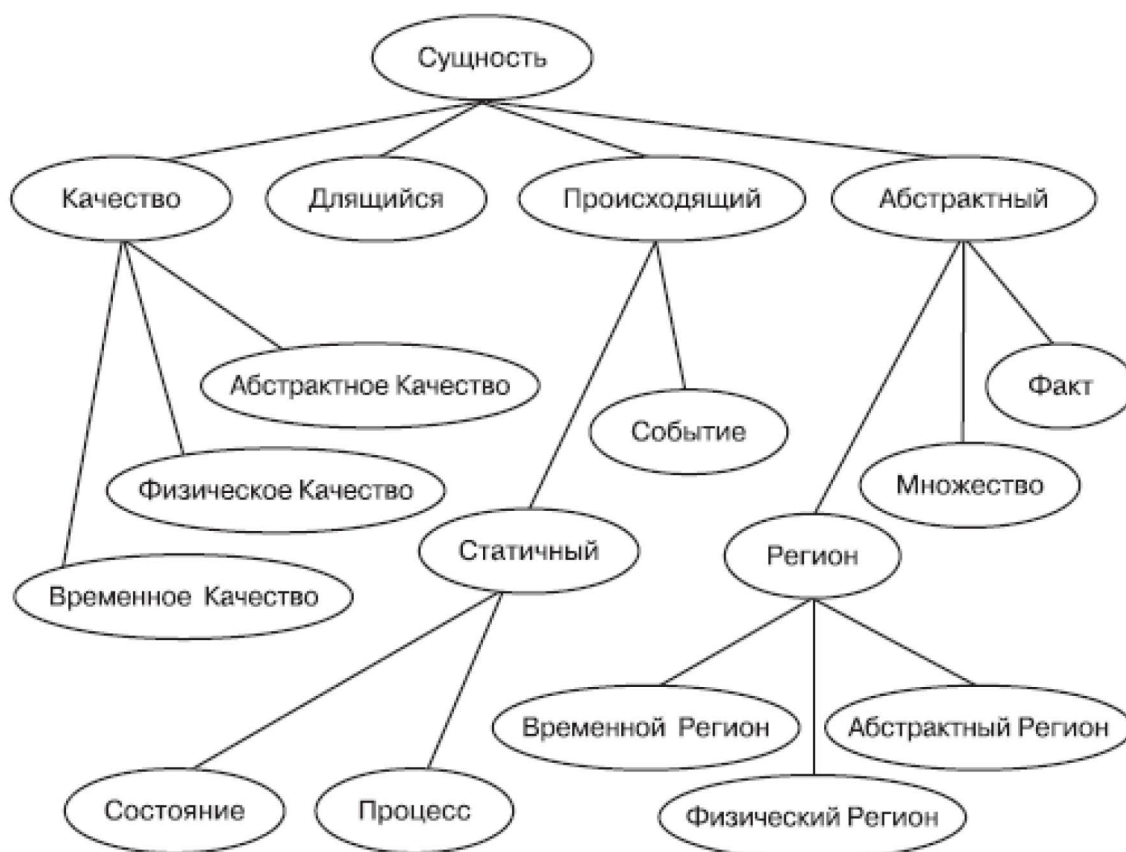


Рис. 3.11. Верхние уровни иерархии DOLCE

В основу процесса проектирования легло фундаментальное философское разделение всех сущностей на универсалии (сущности, потенциально или реально имеющие экземпляры) и индивиды (или частности), которые не имеют и не могут иметь экземпляров. DOLCE – онтология индивидов, в том смысле, что область описания ограничена только ими. В качестве примера универсалии можно привести понятие «собака» (оно имеет множество экземпляров, конкретных примеров в окружающем мире). В отличие от этого понятия, понятие «время» скорее рассматривается как индивид (едва ли кому-то понадобится трактовать «время» как множество различных сущностей, конечно, если речь не идет о параллельных мирах).

Еще одна черта DOLCE (также заимствованная разработчиками из философии) – явное разделение на «постоянные» и «происходящие» сущности. Различие между ними состоит в том, что «постоянные» сущности имеются в наличии целиком и неизменно в некотором фиксированном промежутке времени (например, стол или дом в течение периода своего существования). «Происходящие» разворачиваются во времени и в каждый момент в некотором временном интервале они могут быть различными, по-разному себя проявлять, иметь разный состав (например: ураган или период раннего Ренессанса), однако при этом их идентичность сохраняется.

Такое разделение на «объект» и «процесс» весьма условно, и здесь прослеживается когнитивный уклон DOLCE. Оно привело к тому, что в онтологии определены два типа отношения ЧАСТЬ-ЦЕЛОЕ. Первое никак не зависит от времени, второе имеет временной индекс, определяющий, в каких временных рамках отношение действует. Подобное «раздвоение» наблюдается и для отношения «КАЧЕСТВО – ОБЛАДАТЕЛЬ КАЧЕСТВА». Другие базовые отношения онтологии: УЧАСТНИК-ПРОЦЕСС, КОМПОНЕНТ-ЦЕЛОЕ (компонент входит в состав целого) и отношение зависимости имеют временной индекс. Для сравнения, в онтологии OpenCus нет явного деления на «Постоянные» и «Происходящие». Поэтому среди множества отношений в разделе «Части объектов» нет отношения, учитывающего временной аспект: возможное непостоянство данного отношения.

Онтология предметной области

Другое название – онтология домена. Назначение схоже с назначением онтологии верхнего уровня, но область интереса ограничена одной предметной областью (т. н. доменом), например, авиация, медицина, культура, дистанционное обучение, организация. Онтология предметной области обобщает понятия, использующиеся в некоторых задачах домена, абстрагируясь от самих задач (так, онтология автомобилей независима от любых особенностей конкретных марок машин). Во многих дисциплинах сейчас разрабатываются стандартные онтологии, которые могут использоваться экспертами по предметным областям для совместного использования и аннотирования информации в своей области.

Например в области медицины созданы большие стандартные, структурированные словари, такие как SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms – систематизированная номенклатура медицины – клиническая терминология) и UMLS (Unified Medical Language System – семантическая сеть Системы унифицированного медицинского языка). Также появляются обширные общецелевые онтологии. Так, программа ООН по развитию (United Nations Development Program) и компания Dun&Bradstreet объединили усилия для разработки онтологии UNSPSC, которая предоставляет терминологию товаров и услуг (www.unspsc.org).

Активно создаются онтологии организации (enterprise ontologies), которые включают в себя понятия характерные для любой организации, такие как, цель, процесс, проект, должностное лицо, подразделение, полномочия. Одной из самых первых и известных онтологий предприятия является разработка Учхолда, Кинга и других специалистов

Все термины их онтологии попадают в пять верхне-уровневых разделов, отражающих разные аспекты предприятия: метаонтология и время (Meta Ontology and Time), деятельность, план, способность и ресурс (Activity, Plan, Capability, Resource), организация (Organisation), стратегия (Strategy) и маркетинг (Marketing) (табл. 3.1).

Связи и отношения между приведенными и другими понятиями также содержатся в описании данной онтологии. В качестве иллюстрации рассмотрим связи терминов и понятий в секции «деятельность и процессы».

Таблица 3.1

**Некоторые понятия и отношения онтологии
предприятия Учходда**

| Деятельность и др. | Организация | Стратегия | Маркетинг | Время |
|---|-----------------------------|---------------------------------|-----------------------------|-----------------------|
| Деятельность | Лицо | Цель- назначение | Продажа | Ось времени |
| Спецификация деятельности | Машина | Удерживаемая цель-назначение | Потенциальная продажа | Временной интервал |
| Выполнять | Корпорация | Намечаемая цель-назначение | Для продажи | Момент времени |
| Спецификация выполняемой деятельности | Партнерство | Держатель цели- назначения | Коммерческое предложение | |
| Время начала | Партнер | Стратегическая цель | Вендор | |
| Время окончания | Юридическое лицо | Задача | Фактический покупатель | |
| Предварительное условие | Подразделение | Виденье | Потенциальный покупатель | |
| Эффект | Управлять | Миссия | Покупатель | |
| Исполнитель | Делегировать | Цель-результат | Дистрибьютор | |
| Под-активность | Управленческая связь | Обеспечивать достижение | Продукт | |
| Полномочия | Право собственности | Стратегия | Запрашиваемая цена | |
| Владелец активности | Незаконная собственность | Стратегическое планирование | Цена продажи | |
| Событие | Собственность | Стратегическое действие | Рынок | |
| План | Владелец | Решение | Переменная сегментации | |

Центральным термином здесь является ДЕЯТЕЛЬНОСТЬ (ACTIVITY). Он предназначен для фиксации обозначения чего-либо, что включает действительное выполнение, в частности действие. ДЕЯТЕЛЬНОСТЬ может происходить в прошлом, а может происходить и в настоящем. Термин может также использоваться для ссылки на гипотетическую

будущую ДЕЯТЕЛЬНОСТЬ. Однако существует потребность ссылаться явно на спецификацию (детальное описание) или план ДЕЯТЕЛЬНОСТИ. Это называется СПЕЦИФИКАЦИЕЙ ДЕЯТЕЛЬНОСТИ (ACTIVITY SPECIFICATION). Подобно рецепту, спецификация определяет на некотором уровне детальности одну или более возможных ДЕЯТЕЛЬНОСТЕЙ. Понятие ДЕЯТЕЛЬНОСТЬ близко связано с понятием ИСПОЛНИТЕЛЬ, который ВЫПОЛНЯЕТ СПЕЦИФИКАЦИЮ ДЕЯТЕЛЬНОСТИ путем выполнения детально определенных (специфицированных) ДЕЯТЕЛЬНОСТЕЙ. ИСПОЛНИТЕЛЬ может быть ЧЕЛОВЕКОМ, ОРГАНИЗАЦИОННОЙ ЕДИНИЦЕЙ или МАШИНОЙ.

Прикладная онтология

Назначение такой онтологии в том, чтобы описать концептуальную модель конкретной задачи или приложения. Прикладные онтологии описывают концепты, которые зависят как от онтологии задач, так и от онтологии предметной области.

Например, онтология организации для задач организационного проектирования может иметь свои особенности. Продолжим рассматривать онтологию предложенную компанией «Бизнес Инжиниринг Групп». Для коммерческих организаций состав задач оргпроектирования, для решения которых используется онтология, следующий:

1. Стратегическая фокусировка бизнес-системы: идентификация стратегий, «Контроль качества» стратегий, идентификация системы стратегических показателей;

2. Формирование оптимальной организационно-функциональной структуры: идентификация существующей организационно-функциональной структуры, анализ существующей организационно-функциональной структуры, разработка функционала и системы процессов бизнес-системы, разработка / оптимизация оргструктуры и распределения ответственности, выпуск и актуализация организационной документации по ОФМ;

3. Разработка, реинжиниринг и непрерывное совершенствование процессов: идентификация и описание процессов, анализ процессов, оптимизация процессов, выпуск и поддержание в актуализированном состоянии регламентов исполнения процессов.

Задачи, методы и, особенно, правила принятия решений требует

более детальных классификаций понятий (рис. 3.12), а также приводят к возникновению в онтологии классов-ролей, принадлежность к которым зависит от рассматриваемой задачи/метода (табл. 3.2):

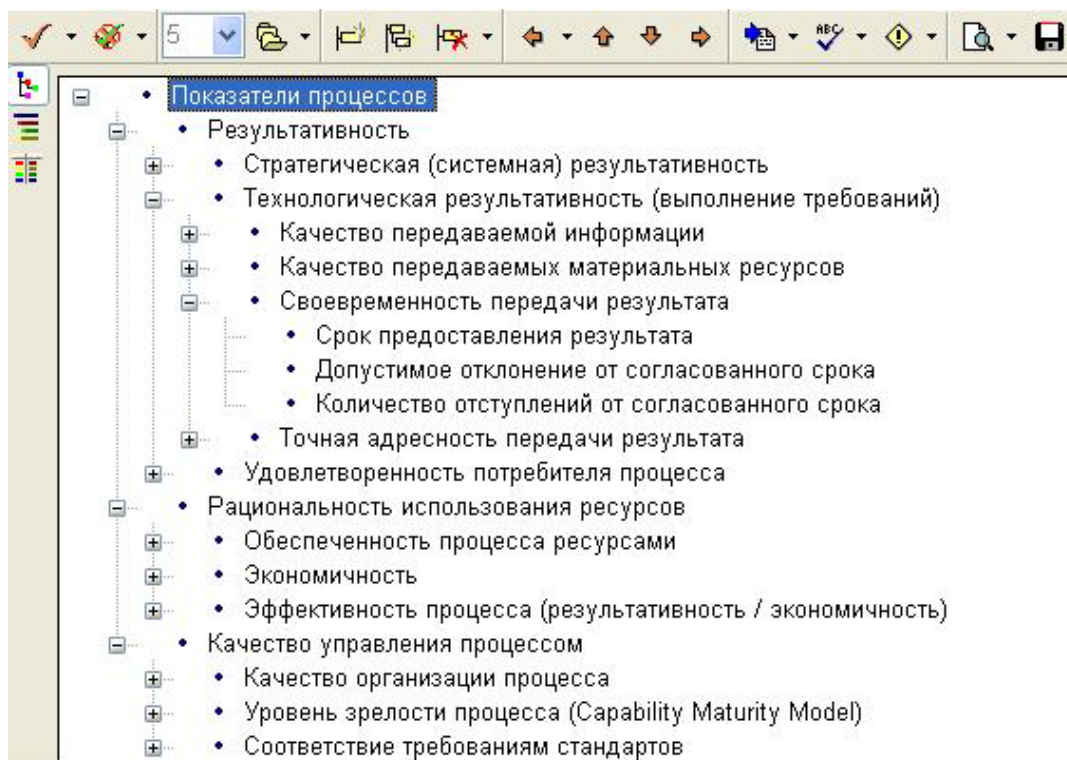


Рис. 3.12. Детализация таксономии «Показателей процессов» в прикладной онтологии организации

Таблица 3.2

Пример классов-ролей в прикладной онтологии

| Контекст задачи и метода | Контекстно-зависимые классы-роли | Классы-типы – исполнители ролей |
|---|----------------------------------|---|
| Прикладной контекст: 1. Анализ деятельности 1.1. Стратегический SWOT-анализ | Силы Слабости | Ресурсы организации (финансовые, материальные, компетенции) Позиция на рынке Ценностное предложение |
| Позиционный контекст: Вид деятельности | Возможности Угрозы | Текущее состояние и тенденции внешней среды (политика, экономика, технологии, общество) |

Данная таблица иллюстрирует, что в контексте стратегического

SWOT-анализа⁹ какие-либо финансовые ресурсы организации могут исполнять роль «Силы» организации, а, например, позиция на рынке – роль «Слабости».

3.5. Применение онтологий

В статье¹⁰ было выделено девять наиболее популярных типов примеров использования онтологий и проведен анализ распространенности того или иного типа примера.

(1) Общий словарь (Common Vocabulary)

Онтология предоставляет общий словарь для описания информационных ресурсов (рис. 3.13), а также общения сотрудников. Общий словарь снижает неоднозначность соответствия между знаком, понятием и денотатом. Данный пример использования был рассмотрен выше (см. раздел 3.1), и он является основой для обоснования целесообразности использования онтологий в УЗ.

(2) Поиск (Search)

При использовании онтологий поисковая машина будет выдавать только такие ресурсы, где упоминается в точности искомое понятие, а не произвольные страницы, в тексте которых встретилось данное многозначное ключевое слово. Сегодняшние поисковые системы зачастую выдают бесчисленное множество совершенно не относящихся к запросу «хитов», обрекая пользователя на длительный ручной отбор материала. Например, если вы ввели для поиска слово «ягуар», то компьютеру совершенно непонятно, имеете ли вы в виду автомобиль, животное или напиток. Вся проблема в том, что для компьютера слово «ягуар» не имеет четкого смысла, или другими словами, семантического содержания. Онтология позволяет указать контекст для поиска. Указывая в поисковой строке слово «ягуар» необходимо дополнительно указать, идет ли речь об автомобиле, животном или напитке. На техническом языке это может

⁹ Один из методов стратегического анализа деятельности организации. Подробнее рассмотрен в разделе «Процесс управления знаниями».

¹⁰ Kozaki K., Hayashi Y., Sasajima M., Tarumi S., Mizoguchi R. Understanding Semantic Web Applications. Proc. of the 3rd Asian Semantic Web Conference (ASWC 2008), February 2-5, Bangkok, Thailand, 2008. P. 524-539.

называться сужением области поиска. Онтологии дают возможность производить запросы на основе понятий, а не на основе совпадения строк. Например, если пользователь задаст вопрос: «*Какие транспортные средства производятся в России?*», то он получит из базы ответ, в который попадут автомобили (= подкласс транспортных средств), производимые во Всеволожске (= город, который находится в России).

Система на основе онтологии может предлагать пользователю возможные интерпретации его поискового запроса на основе анализа взаимосвязей понятий, а также распределять результаты поиска по различным интерпретациям запроса пользователя. Например, для запроса «организационное проектирование» система может уточнить интересуется ли пользователем вопросом проектирования организаций (область менеджмента) или проектированием любых объектов, но в рамках организации.

(3) Указатель (Systematized Index)

В данном применении онтологии используются в качестве структурированного указателя на ресурсы (рис. 3.13). Онтология предоставляет возможности для навигации, по аналогии с папками на домашнем компьютере или каталогом ресурсов. Часто пользователь не может точно сформулировать запрос, например, из-за новизны предметной области для него или нечеткости информационной потребности. В таком случае использование указателя (каталога) может быть более эффективной стратегией, нежели формирование запросов. Семантические описания также могут направлять пользователя в процессе поиска информации, предлагать детализацию, более общий взгляд, смежные области. В случае малого количества результатов поиска, иерархия понятий из онтологии может быть использована для расширения запроса добавлением более общих и более частных понятий, например, добавлением понятий «транспортное средство» и «внедорожник» к понятию «автомобиль».

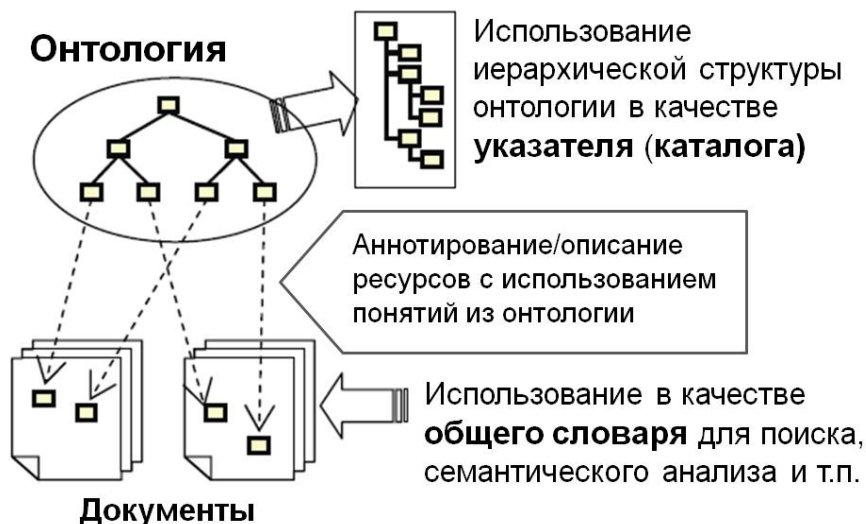


Рис. 3.13. Использование онтологии в качестве общего словаря и указателя

Конкретными примерами приложений данной группы являются семантическая система навигации Magpie¹¹ и система Ontogator¹², в которой указатель основан на представлениях (views) онтологии и интегрирован с контекстными структурами.

(4) Схема данных (Data Schema)

Онтология может использоваться для описания и спецификации схемы базы данных (рис. 3.14).

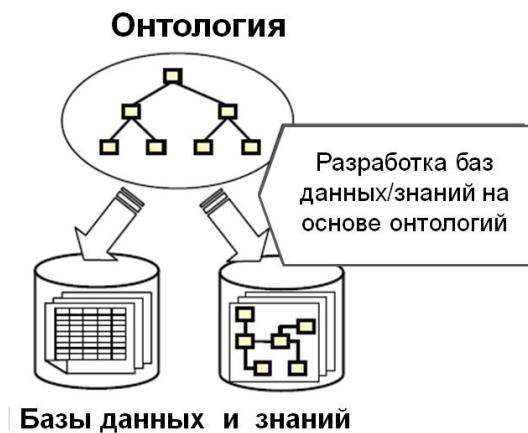


Рис. 3.14. Использование для описания схемы базы данных и структуры модели знаний

¹¹ Martin Dzbor, John Domingue and Enrico Motta: Magpie – Towards a Semantic Web Browser, In Proc. of ISWC2003. P. 690-705.

¹² Eetu Makela, Eero Hyvonen and Samppa Saarela: Ontogator — A Semantic View-Based Search Engine Service for Web Applications, In Proc. of ISWC2006. P. 847-860.

Конкретными примерами в данной группе является схема метаданных единого научного информационного пространства Российской академии наук¹³, а также масштабные базы данных по исследованиям¹⁴.

(5) Посредник для интеграции знаний (Media for Knowledge Sharing)

Онтологии позволяют интегрировать (обмениваться) знания, находящиеся в различных источниках, – документы, базы данных и знаний, сотрудники. Потребителями интегрированной информации могут быть как люди, так и программы (или программные агенты). Возможны два варианта интеграции (рис. 3.15): (а) потребители работают с одной общей онтологией, с помощью которой описываются все информационные ресурсы, (б) потребители имеют свои собственные частные онтологии, между элементами которых установлены соответствия. В целом, доступ к информации может быть организован за счет ее совместного использования потребителями или посредством обмена информацией между ними.

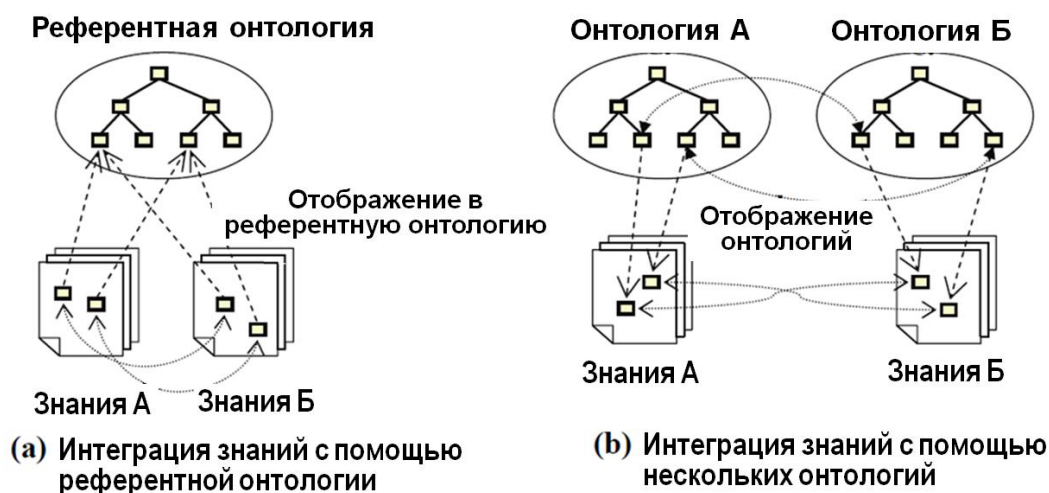


Рис. 3.15. Использование в роли посредника при интеграции знаний

Так как сейчас все большее внимание уделяется распределенным информационным системам моделирования организаций, которые обмениваются знаниями друг с другом, совместно используют знания,

¹³ Бездушный А.А., Бездушный А.Н., Нестеренко А.К., Серебряков В.А., Сысоев Т.М. Предложения по наборам метаданных для научных информационных ресурсов ЕНИИ РАН. Электронные библиотеки. 2004. Т. 7. № 5. С. 1-23.

¹⁴ Fox P. et al.: Semantically-Enabled Large-Scale Science Data Repositories, In Proc. of ISWC2006. P. 792-805.

реализуют возможности повторного использования знаний, то возникает вопрос о представлении знаний в формальной форме. Несмотря на то, что такие системы имеют стилистические и организационные различия в представлении знаний, необходимо, чтобы они сохраняли декларативное содержание этих знаний и, таким образом, могли формулировать запросы к другим информационным системам, обмениваться с ними данными и знаниями. В данном случае возникает проблема необходимости организации взаимодействия с учетом нескольких созданных онтологий. В результате, онтология становится средством предоставления возможности взаимодействия между компьютерными системами независимо от индивидуальных технологий систем, архитектуры информации и областей применения.

В качестве реального примера использования онтологий для интеграции знаний можно рассмотреть интеграцию знаний о бизнес-процессах, представленных в разных нотациях. В рамках проекта SUPER (<http://www.ip-super.org/>) были разработаны онтологии, соответствующие стандартным нотациям (EPC, BPMN, BPEL), обобщенная онтология моделирования бизнес-процессов, верхнеуровневая онтология процессов, а также схемы соответствия между ними (рис. 3.16).

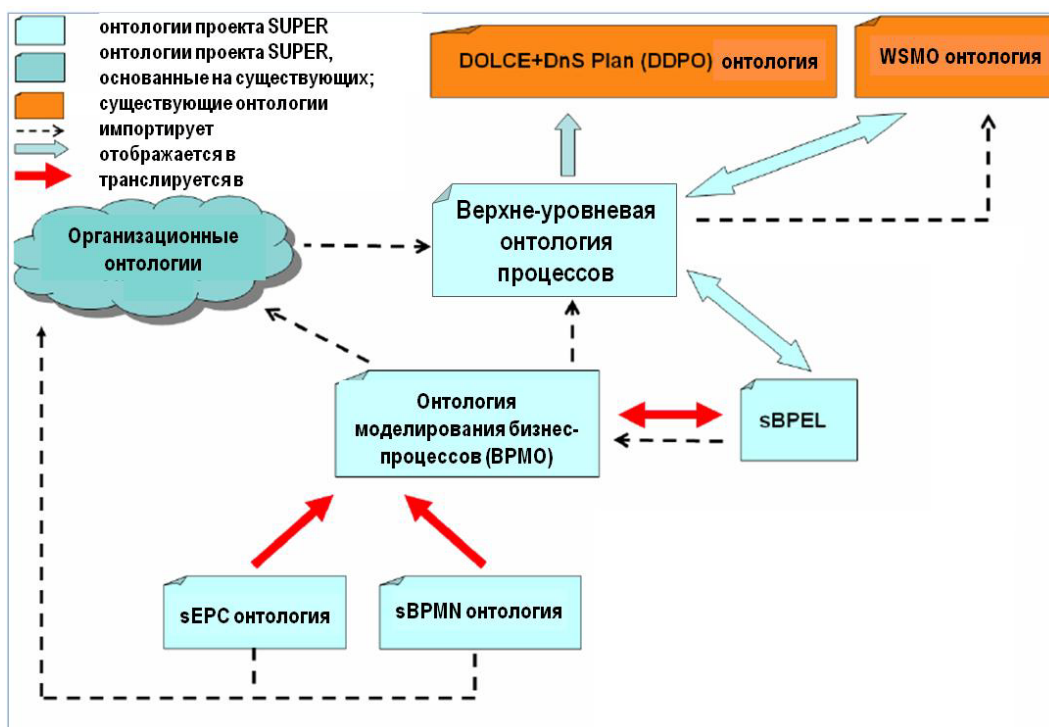


Рис. 3.16. Интеграция онтологий моделирования бизнес-процессов

Данные онтологии и схемы, заложенные в программную среду, позволяют преобразовывать модели процессов из одной нотации в другую, а также осуществлять интегрированный поиск по всем моделям.

(6) Семантический анализ (Semantic Analysis)

К знаниям, представленным с помощью онтологий, могут быть применены автоматизированный логический вывод и семантическая обработка, которые позволяют проводить семантический анализ и создавать новые знания. Одним из распространенных примеров вывода является автоматическая классификация объектов с помощью инструментов логического вывода (inference engine). Например, зная, что ингредиентами какого-либо блюда являются только овощи и крупа, система может автоматически отнести данное блюдо к классу вегетарианских.

Интеграция онтологий с правилами позволяет проводить автоматическую классификацию, проверку соблюдения политик и может использоваться для поддержки принятия решений. Такое применение онтологий детально рассматривается в проекте OntoRule (<http://ontorule-project.eu/>).

Некоторые приложения визуализируют информацию с помощью онтологий для поддержки работ по анализу¹⁵.

Использование информации программными агентами при решении рутинных умственных задач, которое было предложено в рамках концепции семантического веба, также относится к семантическому анализу на основе онтологий. Семантический веб (Semantic Web) привносит смысловую структуру в содержание веб-страниц. Это не какая-то отдельная сеть, а расширение уже существующей, такое, что в ней информация снабжена точно определенным смыслом, позволяющим человеку и машине успешно взаимодействовать.

Сценарий работы таких программных средств представлен в основополагающей статье идеологов семантического веба – Тима Бернерса-Ли и Джеймса Хендлера, опубликованной в 2001 году в журнале Scientific American:

«Пит поднял трубку, звонила его сестра Люси из кабинета врача:

¹⁵ Harith Alani et al.: Monitoring Research Collaborations Using Semantic Web Technologies, In Proc. of ESWC2005. P. 664-678.

«Моей маме нужно попасть на прием к врачу, а потом ей требуется пройти несколько сеансов физиотерапии. Примерно два раза в неделю. Я сейчас поручу своему агенту записать нас на прием к врачу». Пит сразу же согласился подвезти ее маму на своей машине. В кабинете врача Люси дала указания своему электронному агенту семантического веба через свой портативный веб-браузер. Агент сразу же получил информацию о *назначенном маме лечении* от агента ее врача, просмотрел несколько перечней *медицинских учреждений*, оказывающих подобные услуги, проверил, какие из них *подходят* [in-plan] по ее страховому полису и находятся в *окрестности 20 миль* от ее дома и при этом имеют *рейтинг* – «отличный» либо «очень хороший» согласно достоверным рейтинговым источникам. Далее агент попытался сопоставить имеющиеся *часы приема* врачей (предоставленные агентами соответствующих медучреждений на их веб-сайтах) с расписаниями Пита и Люси. (Слова, выделенные курсивом, указывают на термины, семантика или смысл которых была предоставлена агентам посредством семантического веба.) Уже через несколько минут агент представил им план. Питу он не понравился – университетская клиника расположена на другой стороне города, и ему пришлось бы возвращаться обратно как раз в час пик. Он попросил своего агента повторить поиск с более строгими предпочтениями относительно *места* и *времени*. Тем временем агент Люси, *полностью доверяя* агенту Пита в рамках данной задачи, автоматически помогал, предоставляя тому права доступа и ссылки на уже добытую им информацию. Практически мгновенно был предложен новый план: найденная клиника была уже гораздо ближе, и часы приема более ранние».

(7) Извлечение информации (Information Extraction)

Данное применение онтологии позволяет выявлять экземпляры класса и отношения между ними в тексте на естественном языке, а также записывать выявленные факты в семантическое хранилище (базу знаний) (рис. 3.17):

На рис. 3.17 указано, что в результате семантической разметки фразы «XYZ планирует инвестировать 150 млн долларов в строительство завода в Исландии» некоторые ее элементы позволяют наполнить семантическое хранилище. В свою очередь, информация в нем удобна для дальнейшей обработки и анализа (см. применения выше).



Рис. 3.17. Семантическая разметка текста на естественном языке

(8) Спецификация структуры моделей знаний (Usage as a Rule Set for Knowledge Models)

Онтология используется для спецификации структуры моделей знаний, то есть баз знаний. На рис. 3.18 приведена онтология, которая задает структуру для онтологической модели (модели знаний), представленной на рис. 3.19.

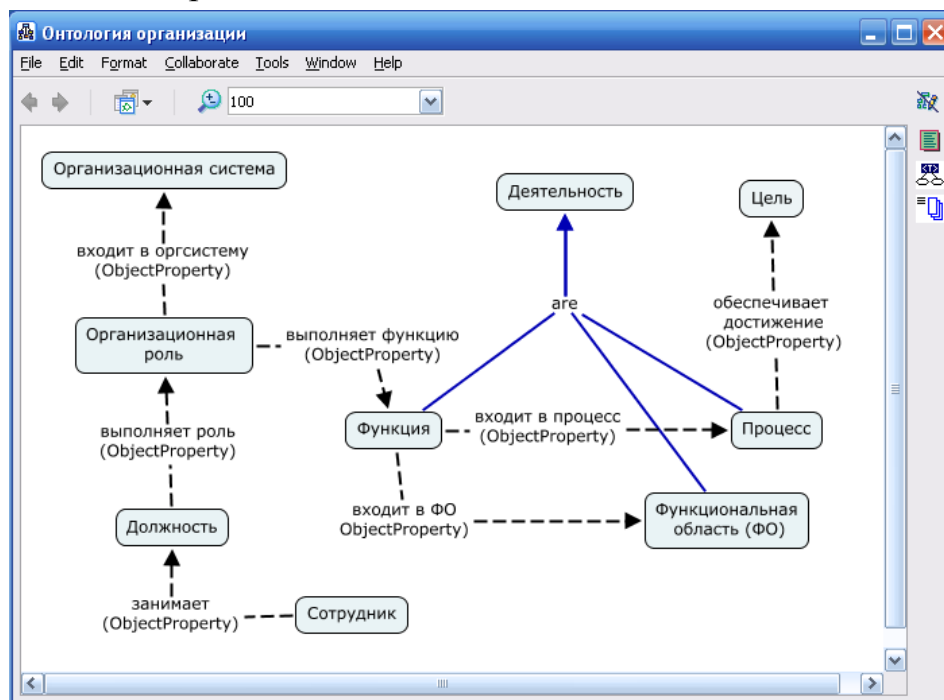


Рис. 3.18. Онтология, задающая структуру базы знаний

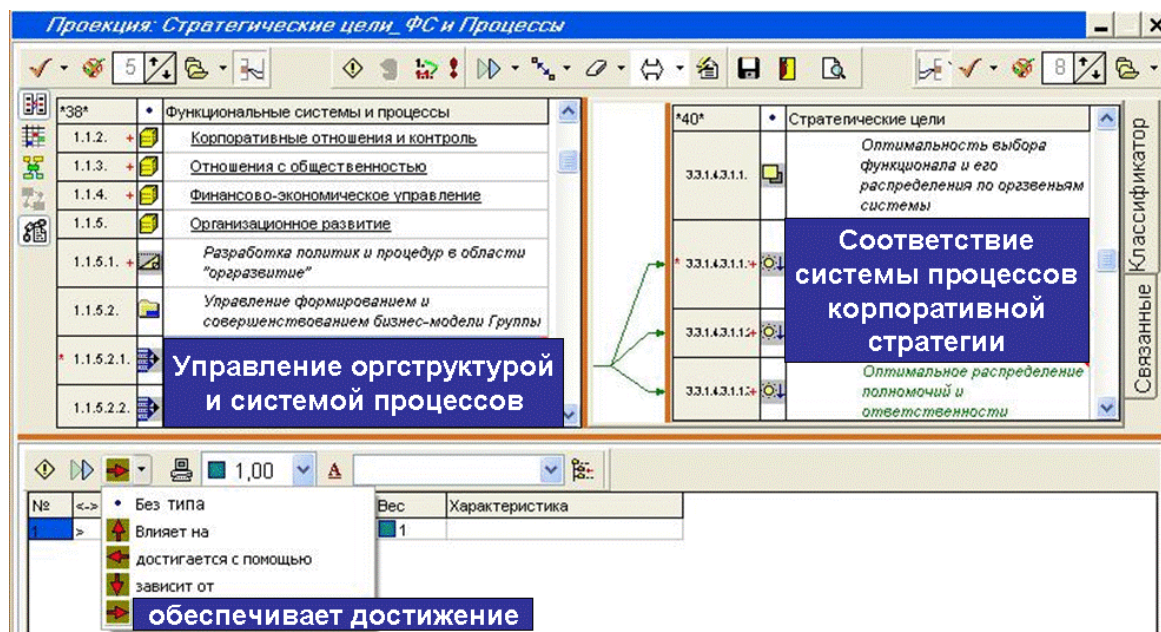


Рис. 3.19. База знаний на основе онтологий

(9) Систематизация знаний (Systematizing Knowledge)

Структурировать знания об окружающем пространстве – для отдела персонала это может быть каталог компетенций сотрудников, для службы техобслуживания и ремонта – справочник оборудования и инструментов, для маркетологов – сегменты рынка, для бизнес-аналитика – справочник бизнес-процессов. Ориентиром в этой области применения онтологий на предприятии является комплексное структурированное описание архитектуры предприятия (enterprise architecture).

На практике, приложение часто сочетают указанные выше примеры использования. Например, выполнения простых автоматических рассуждений и интеграция информации могут повышать качество результатов поиска информации. Онтологии дают возможность получения не заданных явно знаний из информационных хранилищ путем логического вывода – поиск «скрытой информации». Например, пользователь системы может задать вопрос: *Какие поставки продукции находятся сейчас в состоянии риска?* В ответ на такой вопрос система в одной онтологии тарифов определит, что с учетом текущих условий (например, географических или погодных) существуют риски связанные с перевозкой овощей и фруктов. А в другой базе или онтологии деклараций по перевозке груза определит, что в декларации № А345 указаны арбузы,

которые являются подклассом «Овощей и фруктов» (см. рис. 3.20). В результате, система сможет дать конкретный ответ на поставленный вопрос: *Поставка COSCO #A345*

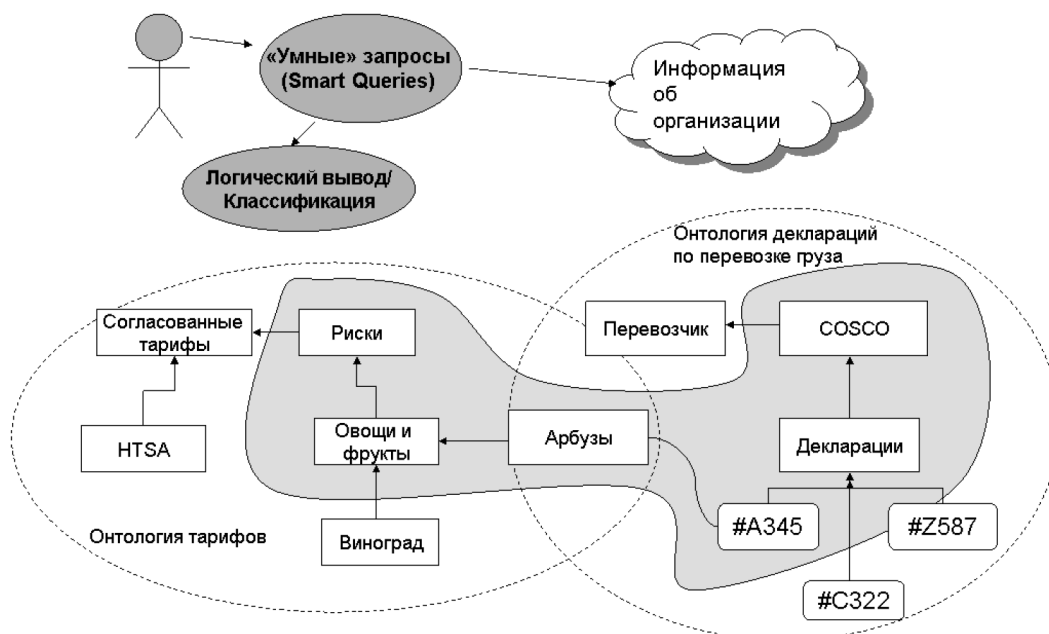


Рис. 3.20. Интеллектуальный поиск на основе логического вывода и интеграции информации

Исследования японских специалистов выявили распределение типов использования онтологий в семантических приложениях по частоте использования см. рис 4.18:



Рис. 3.20. Распределение типов использования онтологий в семантических приложениях

Конкретные примеры семантических приложений и ссылки на описывающие их статьи, структурированные по указанным выше видам применения онтологий, можно посмотреть здесь: <http://www.hozo.jp/OntoApps/>.

3.6. Согласование академической и бизнес- терминологий

Для интеграции бизнес и академического сообществ полезно установить соответствие между их терминологиями в области работы с онтологиями [Mika, Akkermans, 2004] – см. табл. 3.3.

Таблица 3.3

Схема соответствия бизнес- и академической терминологии

| Бизнес-терминология | Академический язык |
|---|--|
| Таксономия, модель предметной области, рубрикатор, сеть знаний, понятия и отношения, система категорий | Онтологии, терминология (T-box), понятия и их свойства |
| Метаданные, контекст, знания | База знаний, экземпляры (Instance data), утверждения (A-Box), сущности и факты, семантическая разметка, метаданные |
| Анализ документов, получение метаданных, категоризация, превращение неструктурированных данных в (ценные, применимые) знания | Аннотирование, извлечение информации (information extraction), классификация |
| Расширение метаданных (metadata enhancement), анализ/выявление (discovery) скрытых знаний/фактов, связывание релевантной информации | Рассуждения (reasoning) |

4. СТРУКТУРА И РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ

В данном разделе предлагается авторский подход, определяющий основные составляющие элементы системы управления знаниями и их взаимосвязи. Детальное описание этих элементов позволяет дать строгое определение самому понятию «система управления знаниями». Выявленные в данном разделе компоненты СУЗ последовательно рассматриваются в последующих главах. Предложенный подход, во многом, опирается на рассмотренные ранее методологии управления знаниями Know-Net [Mentzas et al, 2002], Нонака и Такеучи [Нонака, Такеучи, 2003], CommonKADs [Schreiber et al, 2000], DÉCOR [Abecker, 2003], On-To-Knowledge [Staab et al, 2001].

4.1. Два уровня деятельности: работа со знаниями и управление

Управление – воздействие на управляемую систему, нацеленное на обеспечение требуемого ее поведения.

Как правило, управление предполагает выделение управляемой системы из среды, ее анализ, выработку управленческих воздействий и оценку их последствий. Выделим два уровня деятельности: уровень управления и уровень знаний (см. методологию Карла Виига в разделе «Обзор методологий управления знаниями»), см. рис. 4.1. Данное разделение характерно для любой области деятельности – производство продукции относится к основной деятельности, а планирование и технологическая подготовка производства к управлению, закупки продукции и управление (планирование, учет, контроль) закупками и т. д.

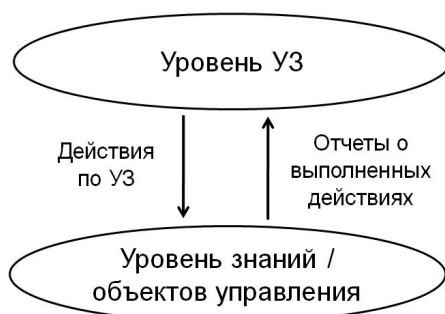


Рис. 4.1. Два уровня в СУЗ

Верхняя часть, уровень управления знаниями, охватывает задачи управления. Когда знания рассматриваются как ресурс, уровень должен управлять этим ресурсом так же, как и любым другим ресурсом. По сути, это означает, что ресурс должен быть доступным в надлежащее время, в надлежащем месте, в надлежащей форме, с необходимым качеством и с как можно меньшими затратами. То есть управление знаниями, как и другие задачи управления, может быть представлено как метадеятельность, которая влияет на объектный уровень. Действия по УЗ (рис. 4.1) воздействуют на характеристики и взаимосвязи элементов уровня знаний. По результатам выполнения этих действий формируются соответствующие отчеты, передаваемые обратно на уровень управления.

4.2. Уровень знаний и объекты управления

Центральным элементом данного уровня являются сами знания. Описание данного понятия и основные классификации были рассмотрены в разделе «Знание и его виды». Для описания знаний предлагается использовать два понятия интеллектуальные ресурсы (ИР) и объекты знаний, подобно методологии Know-Net [Mentzas et al, 2002]¹⁶. Для решения задач стратегического характера (бизнес-стратегия, стратегия УЗ и т. д.) удобнее понятие «интеллектуальных ресурсов», как экономической категории, для детального же анализа проектирования деятельности организации подходит «объект знаний». Для интеграции «стратегического» и «операционного» уровней УЗ производится отображение элементов ИР в объекты знаний.

Понятие «интеллектуальные ресурсы» имеет корни в ресурсном подходе к стратегической теории фирмы. Суть ресурсного подхода заключается в следующем. Различия в экономических результатах фирм, действующих на одном товарном рынке, связаны с различиями в составе ресурсов, которыми располагает фирма, и которые используются ею в

¹⁶ Хотя в [Mentzas et al., 2002] использовано «активы знаний», было принято решение не делать дословный перевод, а взять близкое по смыслу понятие ИР. Это объясняется тем, что в отечественной практике под активом знаний, как правило, понимаются объективированные знания, находящиеся в собственности фирмы, и не включают знания и компетенции персонала.

своей деятельности. Ресурсы понимаются при этом в самом широком смысле – как все то, что обеспечивает функционирование фирмы. Наличие у фирмы неких особых ресурсов, обозначаемых аббревиатурой VRIN (Valuable/Ценные, Rare/Редкие, Inimitable/Невоспроизводимые, Non-substitutable/Незаменяемые), обеспечивает ей конкурентное преимущество, т. е. получение сверхнормальной прибыли, или ренты. В рамках данного подхода знания рассматриваются как один из основных ресурсов фирмы, которые, как раз, обладают VRIN-свойствами.

На уровне предприятия интеллектуальные ресурсы [Климов, 2002] – комплексная категория, объединяющая интеллектуальный капитал людей и различные формы нематериальных активов, фиксирующие знания и профессиональные умения (объективированные знания). Интеллектуальные ресурсы бывают трех видов: *человеческие*, такие как люди, команды, сообщества, *структурные*, такие как бизнес-процессы, технологии и *отношенческие*, такие как знания о клиентах и клиентов о компании (рис. 4.2).



Рис. 4.2. Классификация и примеры интеллектуальных ресурсов

Использование понятия «объекты знаний» позволяет перейти от обобщенной работы на уровне стратегии к реализации задуманных идей, а также интегрировать подходы «знание как запас» и «знание как поток» (см. раздел [«ЗНАНИЕ И УПРАВЛЕНИЕ ЗНАНИЯМИ»](#)). В подходе «знание как запас» акцент делается на предоставлении доступа и упорядочении активов знаний, а подход «знание как поток» устанавливает прямые связи между активами знаний организации. Оба подхода используют некоторую форму представления знаний для их передачи либо между человеком и системой, либо между людьми.

Определим «объекты знаний» как средство выражения знаний, тогда следующее утверждение устанавливает связь между ИР и объектами знаний: *ИР создают, хранят и/или распространяют объекты знаний.*

Например:

Человек – это ИР, который может создавать новые идеи, техническую документацию, коммерческие предложения (объекты знаний),

Сообщество практики – это ИР, который может создавать новые идеи, обобщения опыта, лучшие практики (объекты знаний),

Процесс – это ИР, который может создавать, хранить или распространять лучшие практики, корпоративные стандарты, отчеты по НИОКР (объекты знаний).

Подход «знание как поток» рассматривает объект знаний как информацию, необходимую человеку для формирования знания в его голове. Знания, в данном случае, формируются благодаря ряду операций, которые осуществляются либо в сознании одного человека, либо в процессе взаимодействия людей. Такие типовые операции формирования знаний включают:

- Сравнение: как информация о текущей ситуации соотносится с другими известными ситуациями?
- Последствия: какое значение имеет информация для выполнения действия или принятия решения?
- Связывание: как этот элемент связан с остальными знаниями?
- Обсуждение: что другие думают об этой информации?

Объекты знаний активизируют данные операции, предоставляя людям необходимую информацию.

В подходе «знание как продукт» объект знаний используется для поиска, упорядочения и распространения контента.

В результате, объект знаний выступает связующим звеном как между двумя взглядами на знание (рис. 4.3):

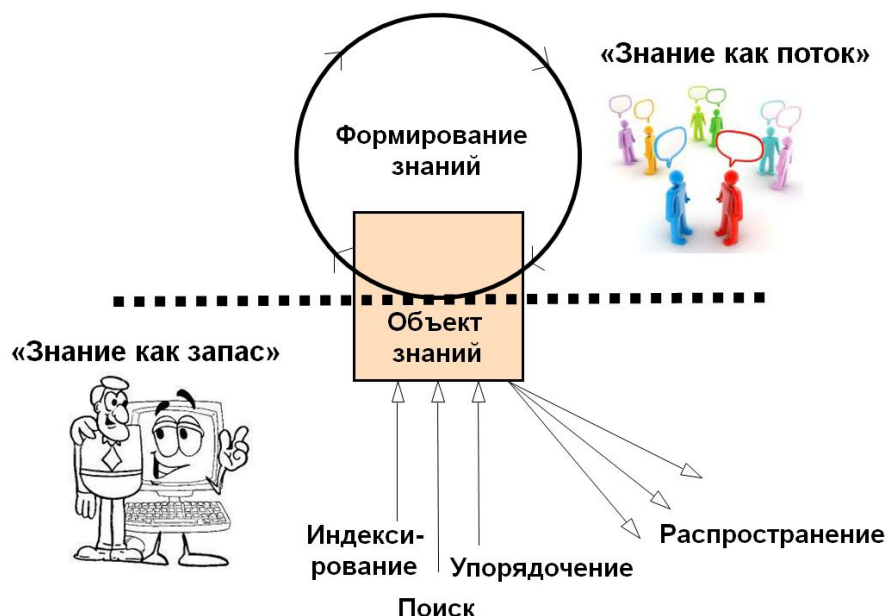


Рис. 4.3. Объект знаний – связующее звено между двумя подходами «знание как поток» и «знание как запас»

Кроме того, понятие объектов знаний отражает семантический подход к УЗ и способствует использованию онтологий, в которых как раз и фиксируются объекты знаний (см. раздел «Семантический подход к УЗ и онтологии»).

Множество явных (формализованных) знаний организации (включая знания о знаниях, внешних и внутренних) представляет собой память организации / корпоративную память (organizational memory / corporate memory).

Конечной целью управленческих воздействий являются характеристики знаний, которые необходимо изменить, поскольку они не соответствуют желаемому состоянию (рис. 4.4):

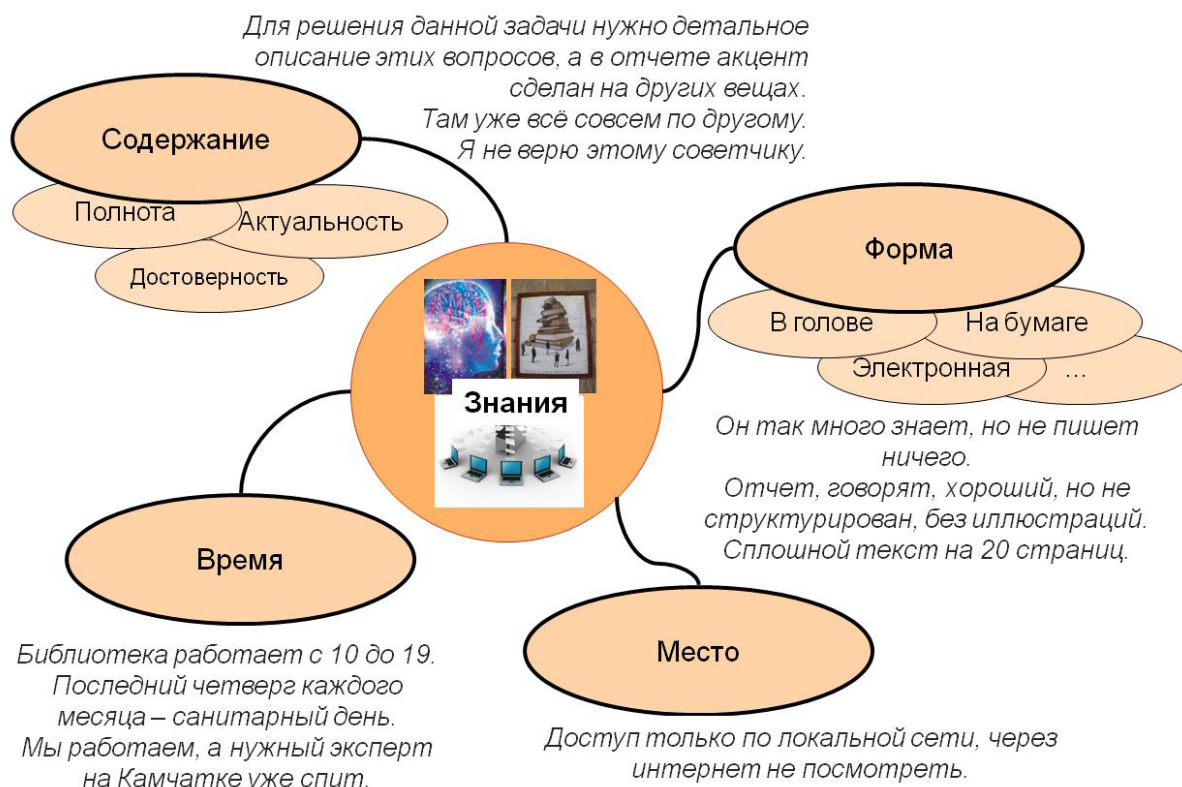


Рис. 4.4. Характеристики знаний и примеры проблем

Определим другие объекты УЗ, которые непосредственно связаны со знаниями и обеспечивают изменение его характеристик. Как уже отмечалось ранее, различная природа явных и неявных знаний определяет разные подходы к работе с ними. Подход «знания как продукт» ориентирован на работу с явными знаниями, подход «знание как процесс» – с неявными. Для подхода «знания как продукт» основным объектом управления будут процессы преобразования знаний (создание, накопление, распределение, обмен неявными знаниями, использование). Однако для подхода «знание как процесс» анализ процессов преобразования знаний не дает достаточной информации для принятия управленческих решений, поскольку изменение характеристик знаний происходит в процессе неформальных коммуникаций и совместной работы, поэтому основным объектом управления для данного подхода являются сети знаний, которые на индивидуальном уровне представляют собой сети сотрудников, на групповом и организационном – сообщества практиков, а на межорганизационном – альянсы.

4.3. Процессы преобразования знаний

Ниже дано описание процессов преобразования знаний. Рис. 4.5 иллюстрирует преобразование знаний с точки зрения состава задач и их взаимосвязей, а рис. 4.6 – с точки зрения получаемых результатов.

В описании используются ссылки на классификацию знаний по степени их формализованности (см. раздел «Знание и его виды»).

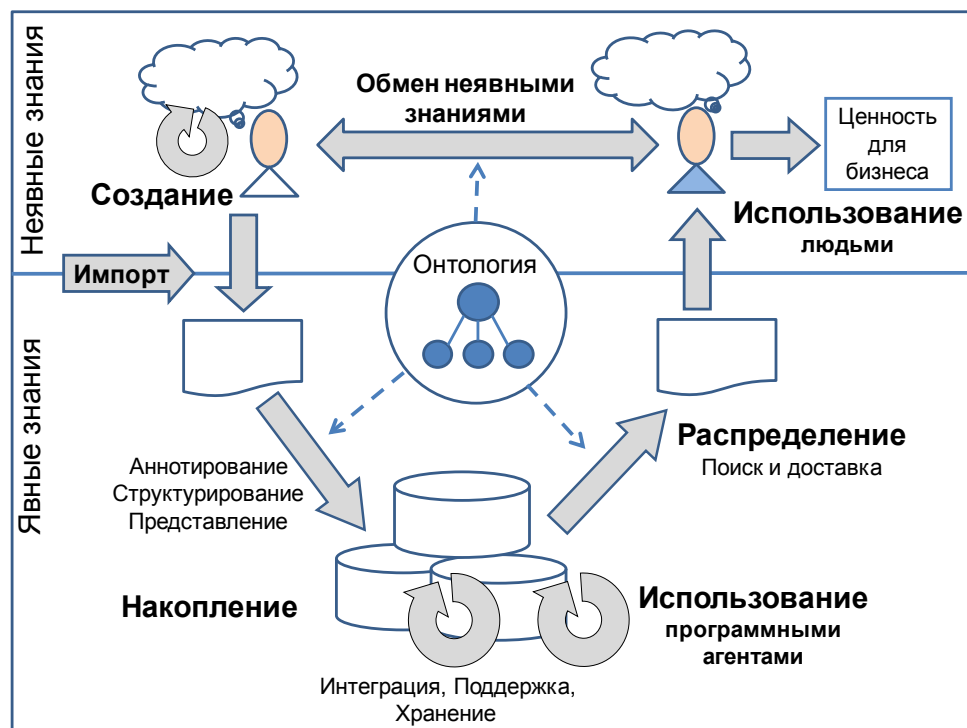


Рис. 4.5. Процессы преобразования знаний и их взаимосвязи

В процессе создания возникают явные и неявные знания, то есть знания в головах сотрудников, а также знания, выраженные и зафиксированные вне человеческого разума с помощью какой-либо знаковой системы, как правило, в форме бумажных или электронных документов или записей. В первом случае возникновение знаний происходит в результате обучения, рассуждений или инсайта. Во втором случае они возникают в результате самостоятельной формализации неявных знаний, процедур извлечения знаний или применения машинных методов (автоматизированный вывод, разведка знаний/knowledge discovery). В последнем варианте компьютерная программа находит новые связи или закономерности предметной области, например, с помощью логических или статистических методов. Кроме создания знаний внутри

организации, существует также импорт «внешних» интеллектуальных ресурсов, например, через статьи из интернета или через услуги внешних консультантов.

В основе процесса накопления знаний лежит описание и структурирование имеющейся в организации информации, внешних ресурсов, а также неявных знаний сотрудников. Именно описание и структурирование позволяет в дальнейшем найти и доставить нужные знания потребителям. Для документов и баз данных указывается их содержание, автор, место создания, область применения и прочие атрибуты. Для баз данных возникает централизованное описание их схем. У сотрудников появляются профили, описывающие их знания, умения, опыт работы. Происходит дополнительное структурирование/разметка внутри документов. Часть знаний формализуется с помощью с помощью специализированных языков представления в базах знаний. Накопление знаний предполагает разработку концептуальных структур, которые используются для описания ресурсов знаний. В традиционном УЗ для этого создаются классификационные схемы (или рубрикаторы, или таксономии), в семантическом – онтологии. Также к накоплению можно отнести стандартные задачи обработки информации – интеграция, эволюция, хранение.

В процессе распределения нужные знания в нужном виде в нужное время и место доставляются нужным потребителям знаний. Это осуществляется путем предоставления пользователям возможностей поиска (например, с помощью запросов, навигации по каталогу или онтологии) или с помощью проактивной доставки знаний (например, почтовые рассылки, подписки на обновления и т. п.) или благодаря работе программных агентов или путем визуализации концептуальных структур. Целью распределения знаний является удовлетворение информационной потребности, а результатами могут быть ответы на вопросы пользователя, рекомендации, один или множество релевантных документов, указание на человека-носителя знаний, комплексные пакеты знаний, а также целые концептуальные структуры, формирующие ментальные модели потребителей знаний.

Обмен неявными знаниями предполагает коммуникации и совместную работу. Участники обмена передают друг другу сообщения (информацию),

наблюдают за поведением более опытных участников, в результате чего в их сознании возникают элементы знаний. Участниками обмена неявными знаниями на индивидуальном и групповом уровне могут быть отдельные сотрудники, на уровне организации – отдельные группы, а также представители организаций на межорганизационном уровне. Онтология в данном случае может задавать общий язык для коммуникаций.

Использование знаний предполагает решение практических задач с помощью полученных знаний, повышение эффективности деятельности сотрудников и организации в целом, инновации. В условиях развития семантических технологий можно учитывать также использование накопленных знаний программными агентами. Использование замыкает цикл преобразования знаний, поскольку решение задач, намеренно или ненамеренно, будет приводить к созданию новых знаний.

Стоит отметить, что представленные выше формулировки процессов преобразования знаний соответствуют точке зрения субъекта управления (знаниями), а не участников данных процессов. Например, пользователь системы, который получает нужные знания в результате их распределения, будет называть данный процесс получением знаний.



Рис. 4.6. Результаты процессов преобразования знаний

4.4. Неформальные сети знаний

Понятие «неформальная сеть знаний», вначале, рассмотрим на трех примерах, взятых из книги [Лессер, Прусак, 2006], написанную специалистами по УЗ компании IBM. Первый пример иллюстрирует использование неформальной сети знаний на индивидуальном уровне, второй рассматривает неформальные группы для обмена знаниями (сообщества практиков), третий – рассматривает обмен знаниями на межорганизационном уровне.

Пример 1. Сеть знаний на индивидуальном уровне.

«В 2002 г. IBM приобрела PriceWaterhouseCoopers Consulting (PWCC), тем самым удвоив свое консалтинговое отделение, что привело к созданию одной из крупнейших в мире фирм по управленческому консалтингу. На первых этапах интеграции планировалось провести ряд официальных мероприятий, в результате которых объединенная компания вобрала бы все то лучшее, что имелось в каждой фирме. Предусматривалось создать команды из представителей обеих фирм и решить множество ключевых проблем новой организации, в том числе касающихся развития бизнеса, укомплектованности, стимулирования и методологий.

Вскоре после заключения соглашения с нами связалась старший консультант практики управления знаниями PWCC: она хотела побольше узнать о том, как IBM работает в сфере управления знаниями и чем IBM Institute for Business Value мог бы помочь ей. Мы хорошо ее знали, так как раньше вместе работали в другой фирме. Хотя мы несколько лет не общались, мы быстро нашли общий язык и поделились с ней своим мнением о перспективах слияния. Прошло немного времени, и мы с вновь обретенной коллегой представили друг друга членам наших консалтинговых групп во время телеконференции, таким образом, расширив число контактов между нашими организациями. В течение недели эти связи принесли несколько коммерческих предложений, над которыми стали работать сотрудники обеих фирм. Так один-единственный разговор принес множество новых контактов, раскрывших способности и умения членов коллектива, и заложил основы будущего расширения

продаж.

Эта история иллюстрирует один важный момент: несмотря на всю значимость официальных мер по интеграции двух фирм, невозможно переоценить значение неформальных социальных сетей как средства распространения знаний. Они быстро и эффективно действуют при поиске людей с нужным опытом, обмене идеями и мнениями, доступе к рабочей информации и «расшифровке» культурных норм и ценностей».

Пример 2. Неформальные группы для обмена знаниями (сообщества практиков).

«Во многих организациях ведущую роль в области управления знаниями играют сообщества практиков. Вслед за Всемирным банком и такими компаниями, как Ford, British Petroleum и Shell, многие государственные и частные компании стали осознавать важность неформальных или полуструктурированных групп сотрудников, которые регулярно обмениваются знаниями и идеями. Том Бойл из компании British Telecom, один из первых участников нашего проекта, как-то сравнил успешное сообщество практиков с ирландским пабом в небольшом городке. Как и паб, сообщество объединяет людей со схожим опытом (однако с потенциально разными убеждениями), которые собираются, чтобы обмениваться информацией и завязать отношения. Именно здесь можно услышать: «Заходите почаще, нам вас не хватает».

Подобно хорошему питейному заведению, сообществам практиков необходимо наличие как минимум трех ресурсов: времени и места для общения, а также энтузиазма в обсуждении предмета дискуссии. Если нет времени на общение, сообщество не сможет строить и развивать социальные связи, а также формировать коллективную память, которой могли бы пользоваться его члены. При отсутствии места для общения людям потребуется больше сил на поиск других членов сообщества и нужных им документов, т. е. члены сообщества будут зря тратить время, которое было бы можно использовать с толком. Наконец, при отсутствии энтузиазма пропадает интерес к проблеме, и обсуждение теряет всякий смысл».

Пример 3. Сеть знаний на межорганизационном уровне

«Деятельность организаций в области управления знаниями главным образом сосредоточена на обмене и применении знаний внутри фирмы.

Однако все большее признание получает тот факт, что значительный объем знаний, имеющих определенную ценность для фирм и правительственных агентств, находится за их пределами. Это подтверждает и ряд тенденций, наблюдаемых в настоящее время. Число совместных предприятий и альянсов в современной корпоративной среде за последние годы выросло в несколько раз: фирмы признают ценность партнерских отношений в таких сферах, как исследовательская деятельность, разработка стандартов и снабжение комплектующими. Последние события в мире требуют от правительственных организаций, чьи интересы сосредоточены в области разведки и обороны, признания важности сотрудничества и обмена знаниями (в том числе и с иностранными партнерами), в решении проблем, связанных с угрозой национальной безопасности. Одним словом, теперь важно не только то, какими знаниями обладает ваша организация, но и то, каким образом она может максимально эффективно использовать знания своих основных партнеров и клиентов.

Хотя значимость обмена знаниями за пределами организации продолжает расти, существует ряд проблем, связанных с организацией этого процесса. Поиск нужных знаний может оказаться непростой задачей, если ими владеют внештатные сотрудники. Нередко сложно найти человека, владеющего информацией по конкретной теме в крупной фирме или правительственном агентстве, а особые трудности может вызвать поиск знаний за пределами организации. Кроме того, порой трудно определить, какими знаниями организации могут и должны обмениваться. Проблемы, связанные с правами интеллектуальной собственности, возможностью тайного сговора, а также потенциальной утечкой знаний, могут снизить скорость и эффективность обмена знаниями между организациями».

Концептуальная карта на рис. 4.7 иллюстрирует состав и взаимосвязи понятий, характеризующих понятие «сеть знаний».

Как видно из рис. 4.7, сеть знаний состоит из людей, которые передают друг другу информации, общаются, знают или не знают о знаниях других, являются доступными или недоступными для других. Люди, входящие в сети знаний обладают объектами знаний (если анализ ведется на оперативном уровне) или интеллектуальными ресурсами (если

на стратегическом). Знания отдельных людей переходят в знания сети.

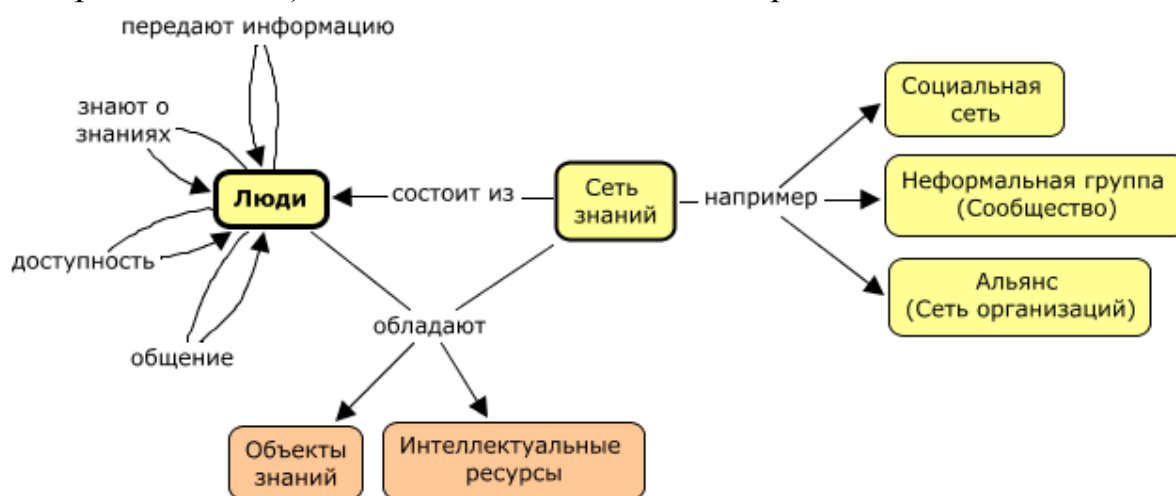


Рис. 4.7. Структура и примеры сети знаний

Примерами сети знаний являются социальная сеть (индивидуальный уровень), сообщество (уровень группы или организации) или сеть организаций (например, альянс).

4.5. Инструменты управления знаниями

Для поддержки процессов преобразования знаний и неформальных сетей используются инструменты УЗ:

- информационные технологии,
- организационные и административные механизмы,
- корпоративная культура,
- техническая инфраструктура,
- правовые аспекты.

Для получения предварительного представления об *информационных технологиях УЗ* можно обратиться к рис. 4.8. Существующие классы программных продуктов структурированы на нем по типу подхода к УЗ, на который они, в первую очередь, ориентированы. Как видно из рисунка, не все классы программных продуктов могут быть однозначно отнесены к тому или другому подходу. Скорее они предоставляют функциональные возможности и характеристики, которые ближе тому или иному взгляду на знания.



Источник: Ovum, 1998

Рис. 4.8. Систематизация информационных технологий УЗ по доминирующему подходу

Раздел «Программная реализация СУЗ» детально рассматривает и систематизирует существующие программные продукты для УЗ, также в тексте разделов 6-9 основной акцент сделан на возможностях информационных технологий.

Организационные и административные механизмы позволяет воздействовать на преобразование знаний и сети знаний путем: реорганизации бизнес-процессов, создания дополнительных должностей для работы со знаниями (менеджер знаний, координатор знаний и т. п.), перераспределение ответственности, создания системы метрик, связанных со знаниями, изменения системы мотивации сотрудников.

Корпоративная культура позволяет воздействовать на преобразование знаний и сети знаний путем: создания атмосферы доверия – ключевого фактора успеха для УЗ, культивации ценностей обучающейся организации – системное обучение, совершенствование личности, групповое обучение, формирования нематериальной мотивации – поощрение обмена знаниями, развития лидерства и вовлеченности, развития персональной культуры работы с информацией и знаниями

(информационная грамотность).

Техническая инфраструктура позволяет воздействовать на преобразование знаний и сети знаний путем: создания и оборудования специальных помещений, предназначенных для обмена знаниями (комната для совещаний, круглый стол, доска, маркеры), учета особенностей умственной работы при проектировании / приобретении / аренде помещений и мебели.

Правовые аспекты позволяют воздействовать на преобразование знаний и сети знаний путем защиты и гарантии прав интеллектуальной собственности.

Выбор и определение требований к инструментам УЗ происходит в рамках процесса УЗ.

4.6. Процесс управления знаниями

В процессе управления знаниями можно выделить следующие задачи: диагностика и анализ знаний, проектирование СУЗ, планирование внедрения СУЗ, контроль и оценка.

На этапе диагностики и анализа знаний (иногда используется термин Аудит знаний) выявляются существующие ключевые знания организации, путем установление связи знаний с целями организации, с ее стратегией, деятельностью, а также с существующими проблемами и возможностями. Производится описание (картирование) этих знаний, позволяющее определить направления для улучшений. В заключение данного этапа может быть разработана политика работы со знаниями, задающая основные направления для разработки мер по улучшения на следующем этапе.

На этапе проектирования СУЗ разрабатываются, оцениваются и отбираются меры по улучшению существующего состояния. Планируемые результаты мер по улучшению интегрируются в перспективную архитектуру СУЗ. Данная архитектура описывает перспективные характеристики знаний и перспективное состояние процессов преобразования знаний, структур, неформальных сетей знаний, а также инструментов УЗ.

На этапе планирования выбранные на предыдущем этапе меры по

улучшению получают статус проектов развития, разрабатываются планы для этих проектов, а также общая программа мероприятий по созданию (если «с нуля») или развитию СУЗ.

В рамках контроля и оценки производится мониторинг исполнения мероприятий по созданию или развитию СУЗ (сроки, бюджет, качество...), оценивается с помощью специальных качественных и количественных показателей состояние знаний, характеристик СУЗ, ее влияние на основную деятельность предприятие и достижение целей.

4.7. Определение понятия «система управления знаниями»

Рассмотренные выше объекты, инструменты и процесс УЗ сведены вместе на рис. 4.9.



Рис. 4.9. Общая структура СУЗ

В результате, можно разделить СУЗ в широком смысле (или комплексную СУЗ) и программную реализацию СУЗ, относящуюся к ИТ-инструментам УЗ:

Комплексная СУЗ – это совокупность процессов и инструментов УЗ, обеспечивающих целенаправленное развитие и эффективное функционирование процессов преобразования и сетей знаний.

Программная реализация СУЗ – это комплекс информационных технологий, обеспечивающий целенаправленное развитие и эффективное функционирование процессов преобразования и сетей знаний.

Стоит отметить, что разработка СУЗ может осуществляться на разных уровнях – межорганизационном, предприятия в целом, группы сотрудников или индивидуальном (разный масштаб).

4.8. Структура проекта разработки и внедрения системы управления знаниями

На рис. 4.10 представлена возможная структура проекта внедрения корпоративной памяти (КП) и СУЗ на предприятии. Состав и логика работ данного проекта соответствуют процессу УЗ.

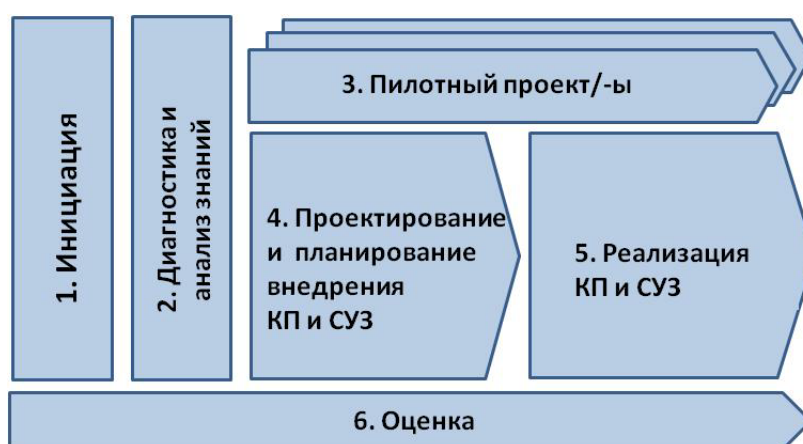


Рис. 4.10. Основные этапы проекта разработки и внедрения СУЗ во времени

Этап 1. Инициация проекта

Задачи:

- 1.1. Проведение ознакомительного семинара,
- 1.2. Формирование команды, устава и плана проекта,
- 1.3. Предварительная оценка уровня зрелости УЗ в компании,
- 1.4. Определение масштаба проекта (организация в целом, бизнес-направление, функциональная область/отдел, отдельные сотрудники).

Результат:

- Знакомство с тематикой и вовлеченность в проект руководства и ключевых специалистов заказчика,
- Команда, устав и план проекта,
- Оценка уровня зрелости управления знаниями в компании,
- Выбранный масштаб проекта.

Этап 2. Диагностика и анализ знаний

Задачи:

2.1. Идентификация ключевых знаний (документированных и в головах персонала) путем анализа целей в области знаний, ключевой деятельности, а также существующих проблем и возможностей.

2.2. Определение носителей и источников ключевых знаний,

2.3. Анализ основных характеристик ключевых знаний (содержание, форма, размещение и доступность) и выявление проблемных областей.

Результат:

Стратегическая карта знаний (классификатор ключевых знаний с их описанием, которое включает в себя стратегические требования, области использования, носителей/источники, характеристики знаний, а также проблемные области – направления для улучшений).

Этап 3. Разработка и реализация пилотного/-ых проекта/-ов

Задачи:

3.1. Разработка перечня потенциальных пилотных проектов по формированию и/или улучшению корпоративной памяти и СУЗ,

3.2. Оценка потенциальных проектов и выбор пилотного проекта/-ов для реализации,

3.3. Разработка плана/-ов реализации для пилотного проекта/-ов,

3.4. Реализация пилотного проекта/-ов.

Результат:

- Разработанные и используемые объекты УЗ (знания, процессы преобразования знаний, сети знаний),
- Разработанные и внедренные инструменты УЗ.

Этап 4. Проектирование и планирование

Задачи:

4.1. Определение и отбор инициатив по работе со знаниями и построению СУЗ,

4.2. Разработка стратегии УЗ, которая определяет основные цели в данной области и принципы их достижения – *опционально, предполагает высокий уровень зрелости управления,*

4.3. Архитектурное проектирование корпоративной памяти,

4.4. Архитектурное проектирование СУЗ (как комплексной, так и ее программной реализации) – *опционально, предполагает высокий уровень зрелости управления и/или обилие частных решений, требующих интеграции,*

4.5. Разработка целевой программы по построению СУЗ и развитию корпоративной памяти.

Результаты:

- Перечень инициатив по построению СУЗ,
- Стратегия УЗ – *опционально, предполагает высокий уровень зрелости управления,*
- Стратегическая карта знаний «как надо»,
- Архитектура СУЗ – *опционально, предполагает высокий уровень зрелости управления,*
- Целевая программа по построению СУЗ и развитию корпоративной памяти.

Этап 5. Реализация мероприятий по разработке и внедрению СУЗ

Задачи:

5.1. Реализация мероприятий по развитию корпоративной памяти,

5.2. Реализация мероприятий по разработке и внедрению СУЗ.

Результат:

- Корпоративная память и СУЗ, удовлетворяющие требованиям.

Этап 6. Оценка деятельности в области знаний

Задачи:

6.1. Разработка критериев, методов и процедур оценки деятельности в области знаний,

6.2. Оценка деятельности в области знаний в соответствии с критериями, методами и процедурами.

Результат:

- Критерии, методы и процедуры оценки деятельности в области знаний, а также результаты оценки.

5. ПРОЦЕСС УПРАВЛЕНИЯ ЗНАНИЯМИ

В разделе 4 было дано укрупненное описание процесса УЗ. Данный раздел детализирует описание каждого этапа.

5.1. Диагностика и анализ

На данном этапе необходимо идентифицировать и проанализировать ключевые знания организации. Для выявления ключевых знаний следует рассмотреть стратегические требования к интеллектуальным ресурсам¹⁷, вытекающим из целей предприятия, требования к знаниям со стороны ключевой деятельности предприятия, а также проблемы и возможности организации, связанные со знаниями.

Анализ и диагностика начинается с определения стратегических требований к СУЗ, источниками которых являются цели. В зависимости от уровня диагностики и анализа знаний в качестве источника стратегических требований могут выступать корпоративная стратегия (уровень корпорации), бизнес-стратегия (уровень бизнес-направления), функциональная стратегия (уровень функциональной системы/области).

Для примера можно рассмотреть упрощенную карту бизнес-стратегии¹⁸ вымышленного предприятия, занимающегося производством упаковки (например, коробка для телевизора, коробка пива и т. п.) (рис. 5.1):

¹⁷ В контексте стратегического управления знания являются одним из ресурсов организации – интеллектуальным ресурсом.

¹⁸ Карта стратегий – это инструмент стратегического менеджмента, используемый для визуализации стратегии организации, на которой указываются цели, связанные причинно-следственными связями.



Рис. 5.1. Карта стратегий – источник целей в области знаний

Данная карта говорит, что:

для достижения удовлетворенности заказчика необходимо оптимизировать упаковочные решения, а для достижения этого персонал компании должен обладать знаниями в области функций упаковки (примеры функций упаковки – защита продукции, продвижение продукции),

для быстрого освоения производственных технологий необходимы компетентность персонала в области технологий производства, а также навыки обучения и самообучения,

и т. д.

Для анализа требований к знаниям со стороны деятельности (функциональных областей, процессов, проектов, операций) необходимо предварительно ее описать/структурировать. Для этого можно обратиться к справочникам типовой деятельности, например, Американский центр по продуктивности и качеству (American Productivity & Quality Center, APQC) выпускает и регулярно обновляет свою схему классификации процессов, основанную на изучении лучшей практики работы компаний США.

«Среднестатистический» справочник деятельности американских компаний иногда сложно применить для анализа деятельности российской компании, кроме того, диагностика и анализ знаний может осуществляться на разных уровнях (корпорация, бизнес-направление, функциональная область/подразделение, процесс/группа), поэтому для систематизации деятельности может быть полезен шаблон (образец, pattern) «системы деятельности»¹⁹ (рис. 5.2):



Рис. 5.2. Шаблон «системы деятельности»

Данный шаблон указывает, какая деятельность должна осуществляться в любой области и на любом уровне обобщения для того, чтобы эффективно создавать ценность для потребителя результатов (продукции или услуг) данной деятельности. С помощью данного шаблона можно рассматривать деятельность корпорации в целом (создание

¹⁹ Система деятельности – совокупность взаимосвязанных элементов деятельности, объединенных для реализации определенного предназначения или/и получения результата/продукта на выходе;

ценности для акционеров), бизнес-направления (создание ценности для клиентов), функциональной области/системы (создание ценности для других функциональных областей/систем).

В основе шаблона системы деятельности лежит *цепочка создания ценности (основная деятельность)*. Именно для обеспечения этой деятельности и управления ею возникают все остальные элементы шаблона. Цепочка создания ценности меняет состав и свойство входящих ресурсов²⁰, преобразуя их в объекты нового качества (новых свойств). Если рассмотреть последовательность действий по получению конечного результата на выходе, то в цепочке создания ценности можно выделить элементы, соответствующие этапам или стадиям жизненных циклов продукции и услуг, поставляемых организацией.

В различных общих и отраслевых стандартах существуют многочисленные образцы «жизненных циклов продукции» (либо общих, либо отраслевых). Их надо «увидеть» на своем предприятии и ориентироваться на них при формулировании названий таких ФС см. табл. 5.1, рис. 5.3:

Таблица 5.1

Стандарты жизненных циклов продукции и услуг

| Общие (межотраслевые) жизненные циклы | |
|--|--|
| «Петля качества» | ГОСТ ИСО 9001:2004 |
| 1. Маркетинг | 1. Процессы, связанные с потребителями |
| 2. Проектирование и разработка | 1.1. Определение требований, относящихся к продукции |
| 3. Подготовка и разработка технологии | 1.2. Анализ требований, относящихся к продукции |
| 4. Материально-техническое снабжение | 1.3. Связь с потребителями |
| 5. Производство | 2. Проектирование и разработка |
| или | 2.1. Стадии проектирования и разработки (разрабатываются организацией) |
| предоставление услуг | 2.2. Анализ проекта и разработки |
| 6. Контроль и испытания | 2.3. Верификация проекта и разработки |
| 7. Распределение и реализация | 2.4. Валидация проекта и разработки |
| 8. Монтаж | 3. Закупки |
| 9. Эксплуатация | 4. Производство и обслуживание |
| 10. Техпомощь и обслуживание | 5. Управление устройствами для мониторинга и измерений |
| 11. Утилизация | |
| использования | |
| после | |

²⁰ Сырье, материалы, комплектующие изделия.

| Жизненные циклы для программного обеспечения (ПО) | |
|---|--|
| Методология Rational Unified Process (RUP): 1. Анализ требований к ПО 2. Проектирование ПО 3. Реализация (программирование) ПО 4. Тестирование ПО 5. Внедрение ПО 6. Сопровождение ПО | V-модель: 1. Разработка концепции 2. Анализ требований и разработка архитектуры 3. Детальное проектирование 4. Реализация 5. Интеграция, тестирование и верификация компонент 6. Верификация и валидация системы 7. Сопровождение |



Рис. 5.3. Примеры отраслевых жизненных циклов продукции

Обычно названия элементов цепочки создания ценности формулируются в виде следующего выражения: («Наименование этапа ЖЦ» + «Наименование продукта/услуги»). Например, («Проектирование» + «Программного обеспечения»), («Внедрение» + «Программного обеспечения») или («Строительство» + «Зданий»).

Кроме элементов жизненных циклов, цепочка создания ценности включает в себя снабжение преобразуемыми ресурсами, а также сбыт продукции/услуг клиенту.

Следующим компонентом шаблона системы деятельности является обеспечение основной деятельности ресурсами:

- Обеспечение персоналом,
- Обеспечение финансовыми ресурсами,
- ИТ-обеспечение,

- Обеспечение оборудованием,
- Обеспечение энергоресурсами,
- Обеспечение инфраструктурой (земля, здания, сооружения).

Цепочка создания ценности (основная деятельность) и обеспечение ресурсами требуют управления. Управление предполагает реализацию традиционного управленческого цикла (сбор информации, планирование, организация, учет, анализ, регулирование) для различных объектов управления (продукция, вид ресурса) на различных горизонтах (день, месяц, год...).

Изменяющиеся внешние условия, а также внутренние изменения системы требуют развития и совершенствования, что и отражено в шаблоне системы деятельности.

Структуру всей системы деятельности задает концепция, на разработку и актуализацию которой направлена соответствующая деятельность.

Таким образом, справочники и шаблон позволяют идентифицировать и систематизировать деятельность организации. С точки зрения УЗ, это необходимо для выявления ключевых элементов деятельности (функциональных областей/систем, процессов, проектов, операций). Для каждого ключевого элемента деятельности определяются наиболее значимые используемые знания (табл. 5.2):

Таблица 5.2

Примеры знаний, требуемых для выполнения деятельности

| Деятельность | Требуемые знания |
|-----------------------------------|--|
| Производство стульев | Знание технологии производства стульев, Знание требований к стульям (нормативы качества), Знание требований к сырью – пиломатериалам, клею, шурупам, Знание требований к работникам, Владение инструментами – пилой, рубанком, молотком, рулеткой. |
| Планирование производства стульев | См. выше, + знание методов управления, + знание программных средств для управления. |
| Снабжение пиломатериалами | Знание требований к пиломатериалам, Знание поставщиков. |

Еще одним способом выявления ключевых знаний является анализ проблем и возможностей организации, связанных со знаниями.

Проблемы и возможности могут быть получены либо просто путем рассмотрения типовых проблем и возможностей в области знаний, либо с помощью проведения стратегического SWOT-анализа²¹.

Примерами типовых проблем в области знаний являются [Wiig et al., 1997]:

- Недостаток знаний в местах их приложения,
- Не используется обратная связь (например, между участниками цепочки добавленной ценности),
- Ограниченный обмен и распределение знаний.

Результатом SWOT-анализа для вымышленной компании по производству упаковки будут:

Силы

1. Накопленный опыт работы с предприятиями разных отраслей (*область знаний*)
2. Наличие системы подготовки кадров
3. Отлаженные производственные технологии (*область знаний*)

Слабости

1. Низкая осведомленность о закупочном комитете заказчика (*область знаний*),
2. Плохая координация действий разных отделов,
3. Система мотивации не соответствует целям компании,
4. Отсутствие регулярной и плановой НИОКР (*область знаний*),
5. Низкий уровень качества продукции.

Возможности

1. Плохая осведомленность заказчиков о возможностях упаковки (*область знаний*),
2. Опережающий рост спроса на высокотехнологичную продукцию (*область знаний*).

²¹ SWOT-анализ – базовая техника систематизации и качественного анализа информации для выработки стратегии, путем сопоставления сильных и слабых сторон организации с возможностями и угрозами, существующими во внешней среде. Примечание: SWOT это аббревиатура из английских слов: Strengths – силы, Weaknesses – слабости, Opportunities – возможности, Threats – угрозы.

Угрозы

1. Быстрое изменение потребностей заказчиков,
2. Долгосрочные договора заказчиков с конкурентами.

Выявленные факторы позволяют определить потенциальные проблемы и возможности в области знаний (потенциальные, поскольку они могут быть уточнены далее после проведения дополнительного анализа активов знаний и выполнения следующего этапа SWOT-анализа).

Полученные цели, проблемы и возможности в области знаний позволяют определить перечень ключевых знаний:

- Знания и компетенции в области функций упаковки (например, защита продукции, продвижение продукции и т. п.),
- Компетентность персонала в области технологий производства,
- Знания о (основных) бизнес-процессах,
- Знания заказчиков о возможностях упаковки,
- и т. п.

Для выявленных ключевых знаний определяются их носители. В зависимости от масштаба анализа, это могут быть подразделения, сотрудники, информационные системы, базы данных/знаний, архивы, группы документов. Результаты идентификации ключевых знаний сводятся в карту знаний.

Карта знаний – модель, как правило, графическая, отражающая распределение объектов знаний между различными элементами организации, такими как организационная единица, функция, процесс, местонахождение и т. п. Дополнительно на карте устанавливается степень покрытия (coverage) элементом знаний соответствующей потребности.

Руди Раглс предлагает два направления использования карт знаний, которые определяют их структуру:

«...они (карты знаний) разрабатываются, чтобы помочь людям понять, куда идти для получения необходимых им знаний, вне зависимости от того, что является конечным пунктом назначения человек, документ или место. Другим способом применения карт знаний является представление потоков знаний внутри бизнес-процессов, начиная с создания / получения, продолжая совершенствованием и хранением, и заканчивая распределением и использованием. Такие карты не должны

охватывать все знания в организации, а должны концентрироваться на ключевых вопросах, обеспечивающих достижение определенных результатов. Например, Hoffmann-LaRoche создал карту знаний вокруг процесса одобрения лекарств, которая позволила увеличить время заполнения необходимых документов с 18 месяцев до 90 дней, а окончательное одобрение лекарств комиссией с ожидаемых 3 лет до 9 месяцев».

Первый вариант использования ориентирован в первую очередь на конечных пользователей знаний, которые используют карту знаний для поиска необходимых им знаний. Основной целью карты знаний в данном случае является формализация всех запасов знаний в организации и мест их нахождения, а также создание удобной навигации по этим запасам. Пользователь в результате получит быстрый доступ к эксперту или носителю явного знания.

Второе направление применения карты знаний ориентировано на управление организацией. Данное применение карты знаний интересно в большей степени для менеджеров и аналитиков, помогая им оценивать знания во взаимосвязи с их применением и важностью для достижения целей бизнеса. Такое применение карт знаний является синергетической комбинацией управления знаниями и улучшения/реинжиниринга бизнес-процессов, которая основана на том факте, что именно бизнес-процессы создают платформу для знаний организации. То есть знания создаются и используются в бизнес-процессах, а приобретение нового и поддержание существующего знания актуально/целесообразно, если оно необходимо для оптимального выполнения бизнес-процессов.

Для описания знаний в рамках диагностики и анализа знаний используется второй вариант карты знаний. В случае создания двух типов карт знаний в одной организации, карты второго типа можно называть стратегическими или управленческими, рис. 5.4.

Ключевые знания или интеллектуальные ресурсы могут структурироваться на карте:

- По видам интеллектуальных ресурсов (структурные, человеческие, отношенческие),
- По видам продукции или услуг,
- По структуре деятельности.

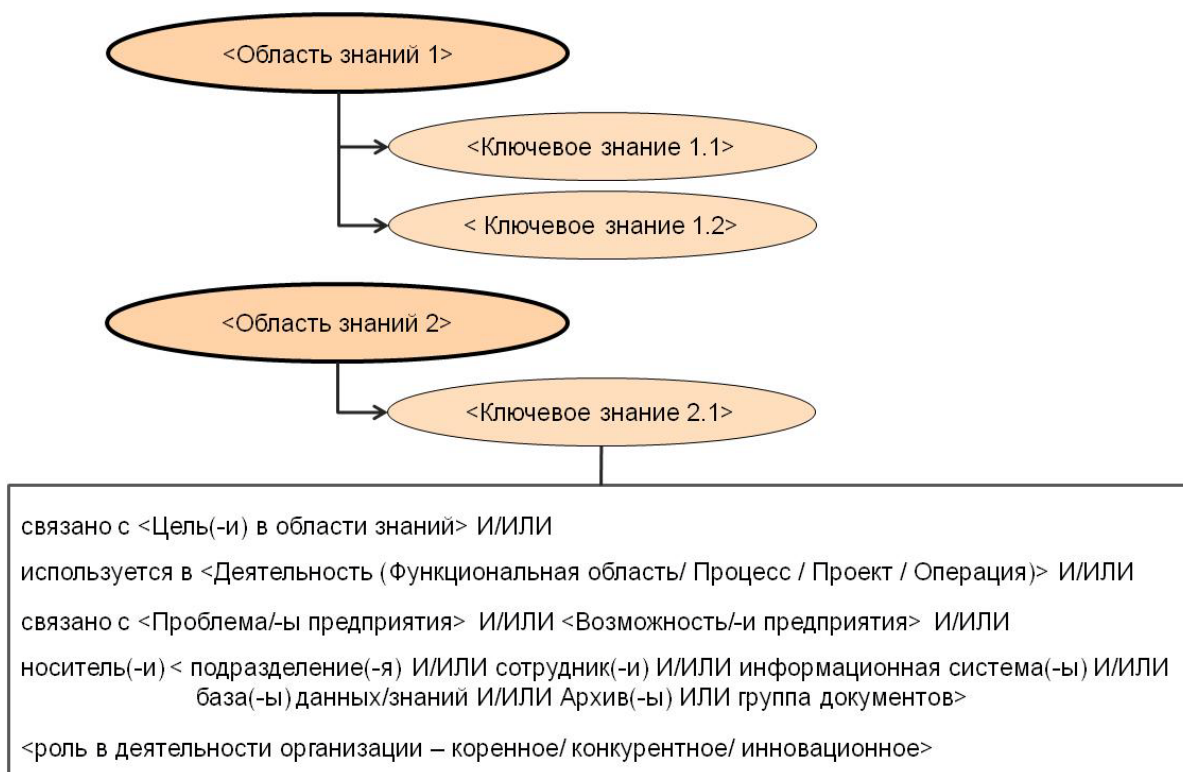


Рис. 5.4. Стратегическая карта знаний (интеллектуальных ресурсов)

Также классификатор знаний может иметь смешанную структуру, например, рис. 5.5:

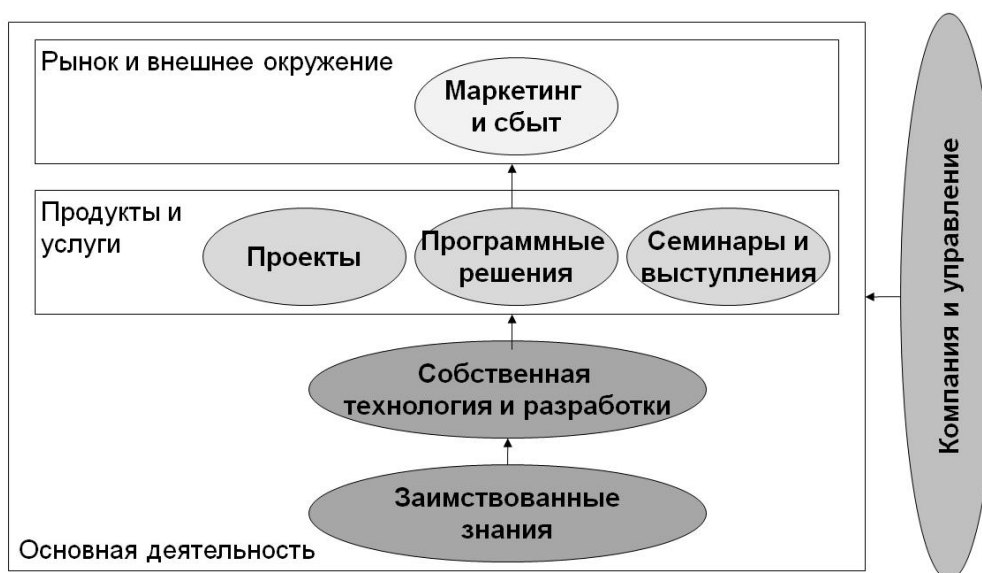


Рис. 5.5. Классификатор знаний для консультационно-внедренческой фирмы

Характеристика «Роль в деятельности организации» на рис. 5.4 основана на классификации знаний с точки зрения их роли в стратегии организации. Знания в данном случае делятся на коренные, обеспечивающие успех и инновационные²².

Коренные знания представляют минимум, и их уровень обеспечивает «участие в игре». Обладание ими не может обеспечить долговременную конкурентную значимость фирмы, но создает определенный барьер входу в отрасль. Коренные знания обычно имеют все участники отрасли и, следовательно, они обеспечивают определенное преимущество перед фирмами, желающими войти в отрасль.

Знания, обеспечивающие успех, снабжают фирму конкурентным потенциалом. Фирма может иметь в общем тот же самый уровень, кругозор или качество знаний как и ее конкуренты, хотя специфические знания могут помочь ей использовать стратегию дифференциации. Эти фирмы могут выбрать конкуренцию по знаниям в сходной конкурентной позиции, надеясь, что они знают больше конкурентов.

Инновационные знания дают фирме возможность лидировать в отрасли. Они часто предоставляют фирме возможность изменить «правила игры».

Знания не статичны и то, что сегодня является инновационным знанием, завтра неотвратно станет коренным. Таким образом, защита и улучшение конкурентной позиции требуют постоянного обучения и восприятия знаний. Способность фирмы обучаться, аккумулировать знания из опыта может обеспечить ей стратегическое преимущество.

Хотя знания динамичны, карта позиционирования в области ключевого (стратегического) знания (рис. 5.6) дает возможность сделать моментальный снимок того, где фирма находится сегодня, ее желательный стратегический профиль и профили конкурентов. Дополнительно она может использоваться для воссоздания исторической и прогнозной траекторий фирменных знаний. Схема может использоваться применительно к области конкуренции, к бизнес-направлению, отделению, продуктовой линии, функции или рыночной позиции.

²² Гольдштейн Г.Я. Стратегические аспекты управления НИОКР. – Таганрог: Изд-во ТРТУ, 2000.

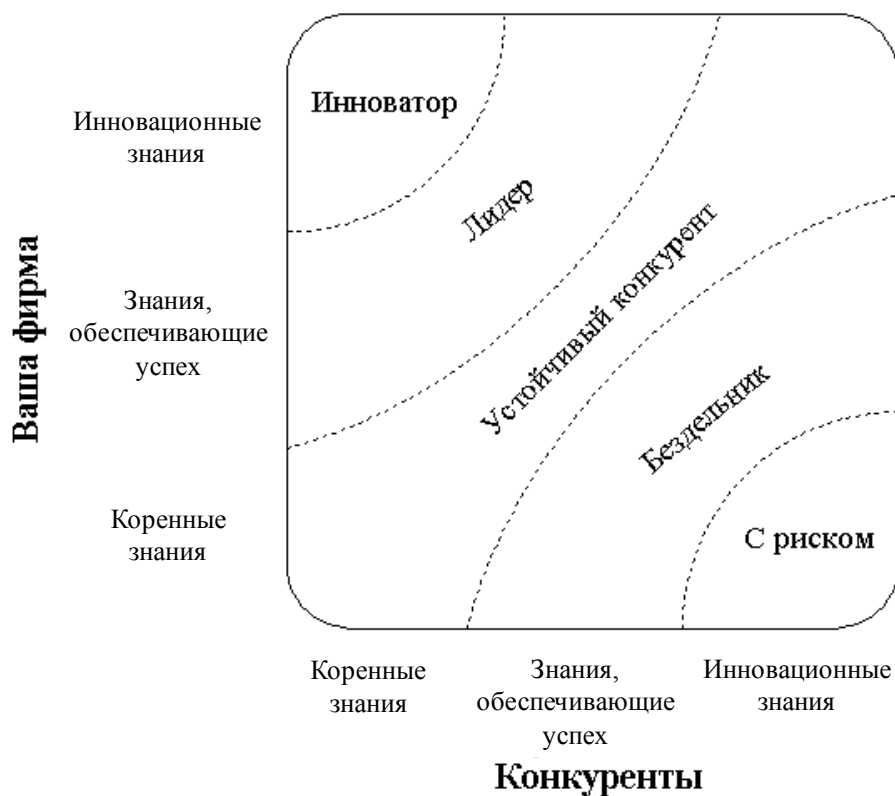


Рис. 5.6. Карта стратегического позиционирования с точки зрения ключевых знаний

Имея отраженную на карте конкурентную позицию фирмы по знаниям, она может приступить к анализу разрыва (бреши) в знаниях. Разница в том, что фирма должна делать в конкуренции, и в том, что она действительно делает, представляет собой стратегическую брешь. В то же время выявление стратегических брешей определяет и потенциальную брешь знаний. Основываясь на карте стратегического позиционирования с точки зрения знаний, организация может идентифицировать, в каких категориях существующее знание организации находится в соответствии со стратегическими требованиями. Результатом такого анализа является выявление ряда потенциальных брешей знания. В некоторых случаях организация может знать даже больше, чем требуется для поддержки конкурентной позиции. Тем не менее, стратегия знаний должна рассматривать любые возможные рассогласования. Чем больше число, изменения, размеры текущих и будущих брешей знаний и больше непостоянство базы знаний из-за динамики и неопределенности конкурентного окружения, тем более агрессивная стратегия знаний требуется. Организация, которая не может реализовать требуемую

стратегию, должна или подогнать свою стратегию под свои возможности, или овладеть способностями выполнять нужную стратегию (рис. 5.7).

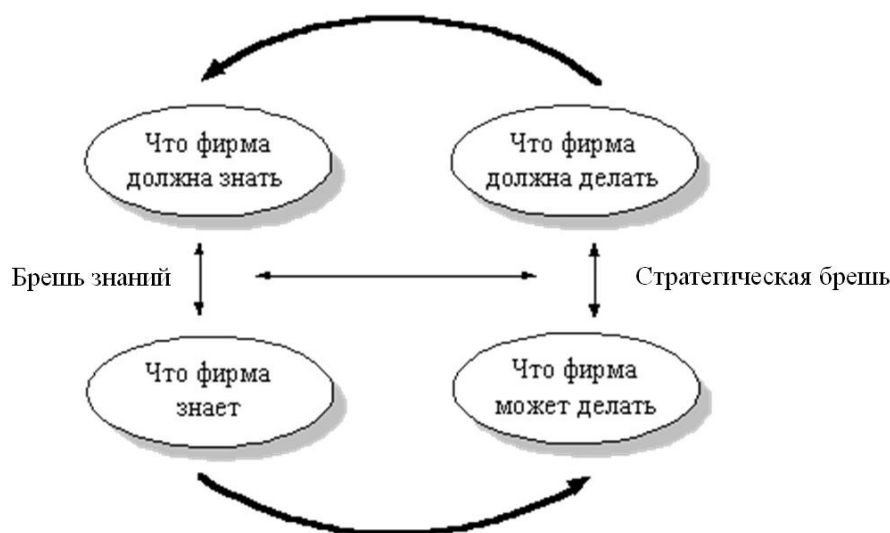


Рис. 5.7. Взаимосвязь бреши в знаниях и стратегической бреши

Для выявленных ключевых знаний, особенно, для тех, по которым есть стратегический разрыв (брешь), производится более детальный анализ, направленный на определение направлений улучшений. В процессе данного анализа проверяется, соответствуют ли стратегическим требованиям основные характеристики знаний: содержание, форма, место, время. Рис. 5.8 поясняет смысл указанных характеристик и направляет процесс анализа. В результате нужно определить есть ли проблемы/возможности с той или иной характеристикой знаний для каждого ключевого знания, а также прокомментировать проблему/возможность в случае наличия.

Характеристики *содержания* знаний могут анализироваться на предмет наличия и критичности НЕ-факторов²³. Термин «НЕ-факторы» был предложен более двадцати лет назад для обозначения комплекса свойств, характерных для человеческой системы знаний о реальном мире, но плохо представленных в формальных системах (неполнота, неточность, недоопределенность, некорректность и т. п.).

²³ Нариньяни А.С. НЕ-факторы и инженерия знаний: от наивной формализации к естественной прагматике // Труды 4-й Нац. конф. «Искусственный интеллект – 94», Рыбинск, 1994, т. 1, с. 9-18.

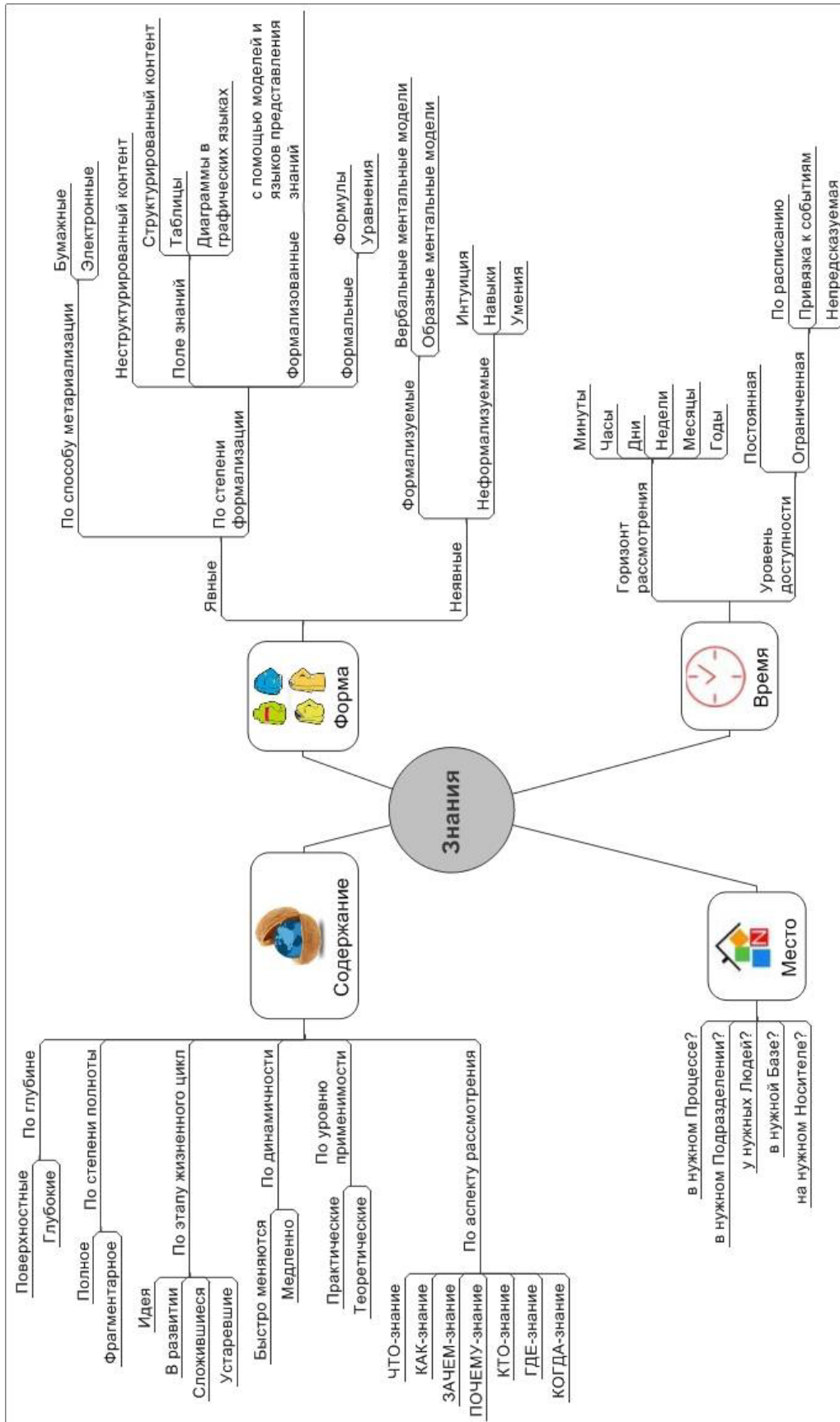


Рис. 5.8. Характеристики знаний

Для выявления проблем или возможностей, связанных с местом размещения знаний, применяется **анализ процессов преобразования знаний**. В соответствии со схемой, предложенной Стромайером, цикл преобразования знаний об организации деятельности может быть представлен в следующем виде (рис. 5.9). Данная схема позволяет идентифицировать недостающие этапы (процессы) преобразования знаний. Например, знания могут создаваться и накапливаться, но не использоваться, либо создаются и используются, но не накапливаются. Также данная схема позволяет связать процессы преобразования знаний с бизнес-процессами и носителями знаний. На рис. 5.9 показано, что создание знаний о клиентах происходит в процессе продаж продукции отделом продаж, в процессе технической поддержки службой поддержки, в процессе маркетинговых исследований отделом маркетинга. Накопление осуществляется только в процессе маркетинговых исследований (например, путем сохранения результатов маркетингового исследования в базе). Распределение в процессе разработки новой продукции, например, путем поиска в базе исследований и извлечения из нее требуемых знаний. Обмен неявными знаниями происходит в рамках неформальных встреч между отделами продаж и маркетинга. Использование – при разработке требований к новой продукции силами отделов маркетинга и продаж, а также в процессе продаж продукции отделом продаж.

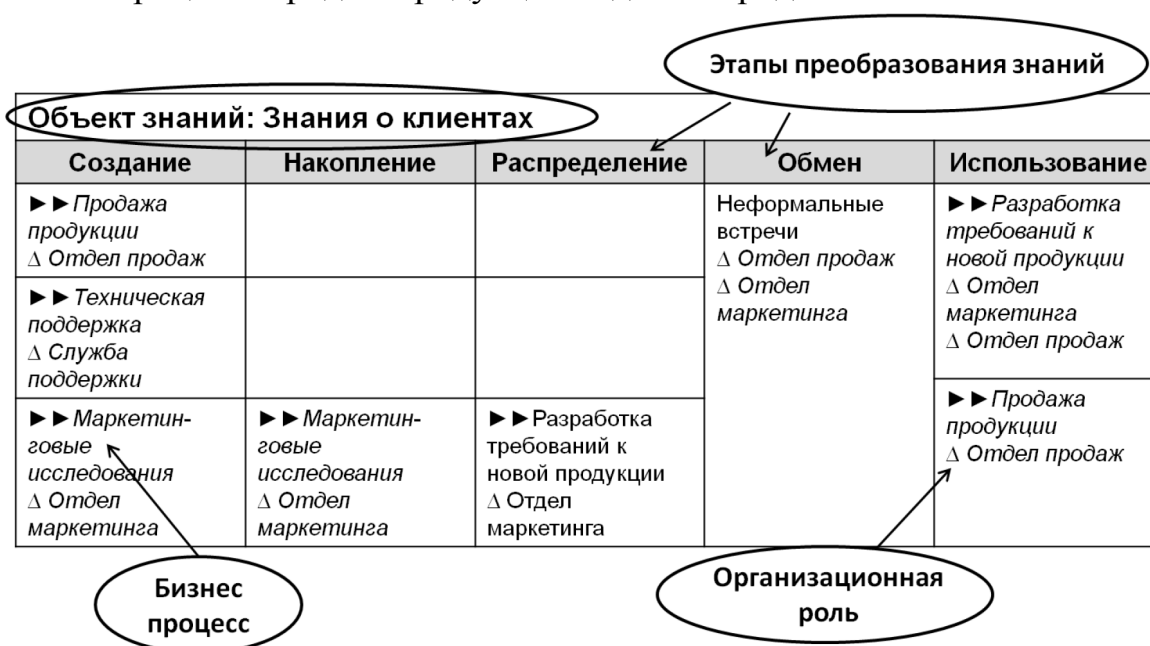
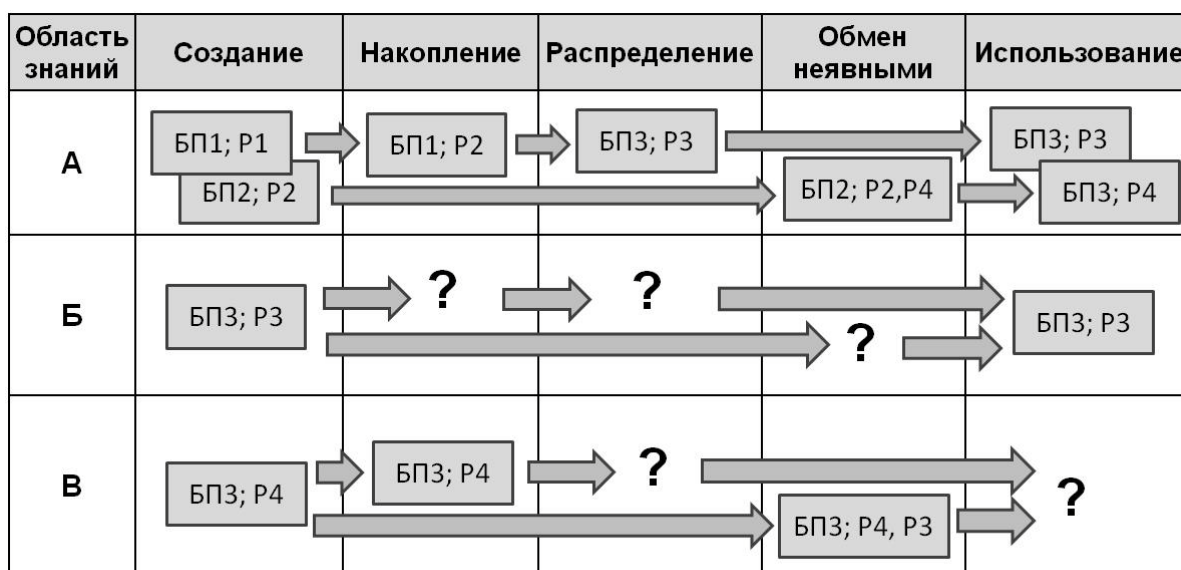


Рис. 5.9. Таблица анализа цикла преобразования знаний

Данный рисунок позволяет увидеть, что знания, создаваемые в процессе технической поддержки, не накапливаются ни в какой базе или хранилище, не передаются в рамках неформальных встреч и обсуждений и, в конечном итоге, не используются отделами маркетинга и продаж. Данный факт определяет одно из возможных направлений улучшения – передачу знаний службы поддержки о клиентах к месту их использования в бизнес-процессах разработки требований к новой продукции и продажи продукции. Другой типовой проблемой, которую легко выявить с помощью представленного анализа, является создание и использование знаний без их накопления, распределения и обмена, что увеличивает риски компании и зависимость от отдельных незаменимых исполнителей. Анализ процессов преобразования знаний производится по каждому ключевому знанию (рис. 5.10):



БП – бизнес-процесс; Р – организационная роль

Рис. 5.10. Анализ процессов преобразования знаний по нескольким ключевым областям знаний

Данный вид анализа может быть расширен путем включения в таблицу дополнительной информации, например, относительно вида знаний (явные/неявные). В случае доминирования неявных знаний в процессах преобразования знаний, можно дополнительно указывать конкретных сотрудников в таблице анализа цикла преобразования знаний,

а также провести анализ социальной сети знаний.

Для выявленных проблем с ключевыми знаниями производится анализ причин этих проблем. Для ключевых знаний, по которым существуют проблемы или возможности улучшений, определяется текущее состояние связанных с ними объектов и инструментов УЗ: носителей знаний, информационных технологий, оргструктуры и процессов, корпоративной культуры, технической инфраструктуры, правовых аспектов.

5.2. Проектирование системы управления знаниями

На основе понимания проблем и возможностей, связанных с ключевыми знаниями, а также текущего состояния связанных с ними объектов и инструментов УЗ предлагаются инициативы (меры) по улучшению текущей ситуации – кандидаты в проекты улучшений.

При высоком уровне зрелости УЗ в организации могут быть разработаны стратегия УЗ, определяющая основные цели в данной области и принципы их достижения, а также политики в области УЗ, относящиеся к той или иной области знаний.

Предложенные инициативы (меры) необходимо оценить, определить целесообразность их реализации и выбрать необходимое их количество. Для этого для каждой инициативы (меры) следует заполнить карточку (табл. 5.3), позволяющую ответить на следующие вопросы:

- Какая область для внедрения/улучшения СУЗ наиболее перспективна, и какое решение лучше принять?
- В чем суть предлагаемой инициативы (содержание)?
- Какие преимущества будут получены, и какие для этого потребуются расходы (бизнес – возможности реализации)?
- Являются ли доступными технологии для реализации решения (технические возможности реализации)?
- Какие дальнейшие действия проекта могут успешно осуществляться (проектные возможности реализации)?

Для инициатив, требующих масштабных инвестиций может проводиться дополнительный более детальный анализ.

Таблица 5.3

Карточка инициативы для СУЗ

| <Наименование инициативы> | |
|--|---|
| <i>Область внедрения</i> | <p>Описываются:</p> <ul style="list-style-type: none"> • области знаний для улучшений И/ИЛИ • инструмент УЗ (ИТ, организация, культура) для улучшения, • текущие проблемы в данной области |
| <i>Содержание</i> | <p>Описываются:</p> <ul style="list-style-type: none"> • ожидаемые результаты реализации инициативы, • основная идея/концепция. |
| <i>Бизнес-возможности реализации</i> | <p>Описываются:</p> <ul style="list-style-type: none"> • ожидаемые преимущества – «материальные» (экономические) и нематериальные; • ожидаемая добавленная ценность; • ожидаемые затраты; • сравнение с альтернативными решениями; • требуемые организационные изменения; • риски и неопределенности. |
| <i>Технические возможности реализации</i> | <p>Описываются:</p> <ul style="list-style-type: none"> • сложность процесса обработки знаний; • доступность современных методов; • присутствие критических аспектов; • время, качество, необходимые ресурсы; • измеримые факторы успеха. <p>Важно понять, делалось ли подобное раньше?</p> |
| <i>Проектные возможности реализации</i> | <p>Анализируются и оцениваются:</p> <ul style="list-style-type: none"> • вовлеченность агентов и заинтересованных сторон; • наличие ресурсов (время, бюджет, оборудование, персонал); • наличие необходимых знаний и компетенций; • реалистичность ожиданий; • адекватная организация проекта и коммуникации. |
| <i>Предлагаемые действия по реализации решения</i> | <p>Описываются:</p> <ul style="list-style-type: none"> • требуемые действия по проекту, • сопутствующие меры: какие меры, кроме основных, нужно еще принять (тренинги и др.)? |

На основе полученных карточек производится отбор инициатив.

Отобранные инициативы необходимо согласовать между собой и интегрировать в существующую СУЗ. Инструментом такого согласования и интеграции является архитектура предприятия (для комплексной СУЗ) и ИТ-архитектура (для программной реализации СУЗ). То есть результаты отдельных инициатив интегрируются и в результате этого получается перспективная архитектура предприятия и/или ИТ-архитектура (рис. 5.11).

ИТ-архитектура определяет совокупность методологических, технологических и технических решений для обеспечения информационной поддержки деятельности предприятия, определяемой его бизнес-архитектурой, и включает в себя архитектуру приложений, архитектуру данных и техническую архитектуру²⁴.

Архитектура приложений, в свою очередь, включает в себя:

- собственно прикладные системы, поддерживающие исполнение бизнес-процессов;
- интерфейсы взаимодействия прикладных систем между собой и с внешними системами и источниками или потребителями данных;
- средства и методы разработки и сопровождения приложений.

Архитектура данных включает в себя:

- базы данных и хранилища данных;
- системы управления базами данных или хранилищами данных;
- правила и средства санкционирования доступа к данным.

Техническая архитектура состоит из сетевой архитектуры и архитектуры платформ. Сетевая архитектура включает в себя локальные и территориальные вычислительные сети; используемые в сетях коммуникационные протоколы, сервисы и системы адресации. Архитектура платформ включает в себя аппаратные средства вычислительной техники – серверы, рабочие станции, накопители и другое компьютерное оборудование; операционные и управляющие системы, утилиты и офисные программные системы; аварийные планы по обеспечению бесперебойной работы аппаратуры (главным образом, серверов) и баз данных в условиях чрезвычайных обстоятельств.

²⁴ Калянов Г.Н. Архитектура предприятия и инструменты ее моделирования // Автоматизация в промышленности. № 7, 2004.

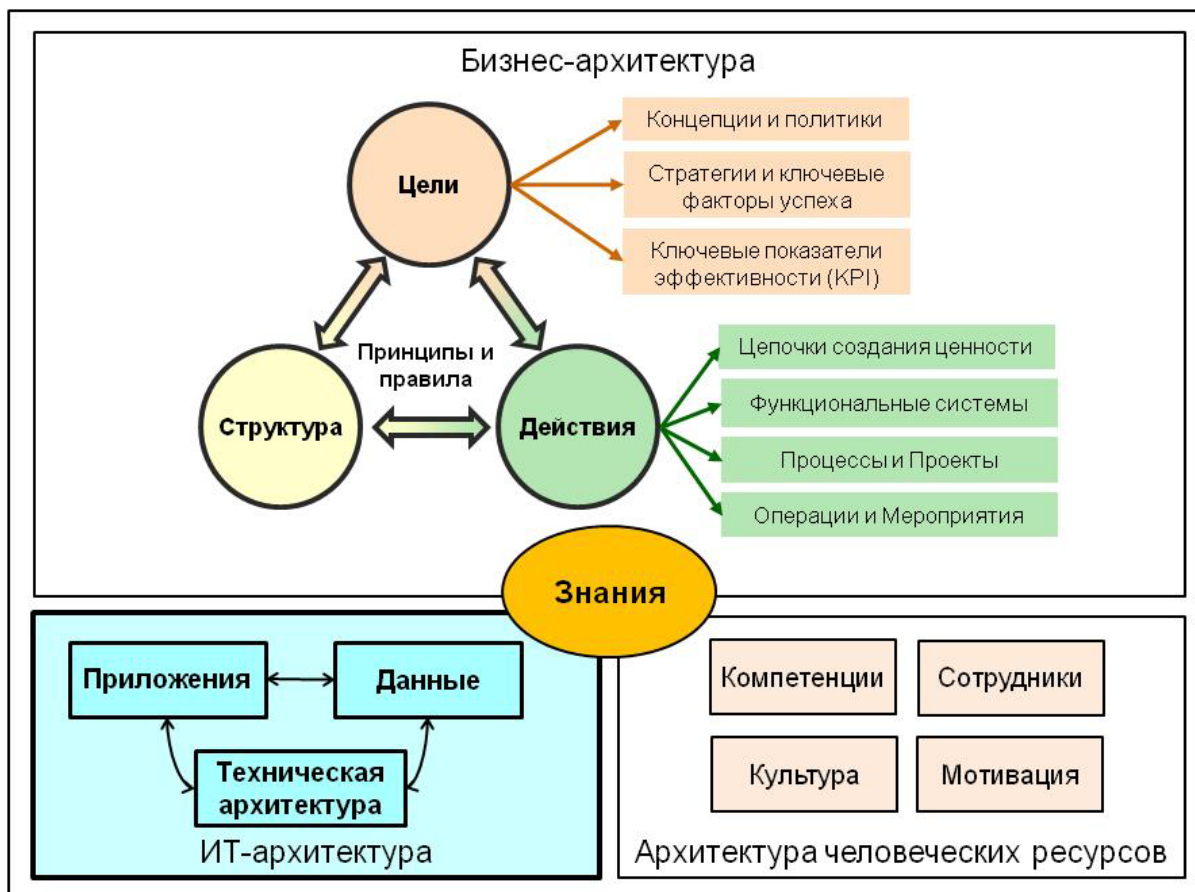


Рис. 5.11. Компоненты архитектуры предприятия²⁵

5.3. Планирование внедрения системы управления знаниями

Важная задача этапа планирования внедрения СУЗ – отбор и распределение ресурсов между проектами. Формирование программы внедрения СУЗ происходит из наиболее приоритетных проектов в области инструментов УЗ, с учетом допустимой хронологии выполнения работ и ресурсных ограничений. Проектами могут быть «разработка и внедрение дополнительной функциональности X информационной системы», «Разработка онтологии для области Y», «Реорганизация процессов управления проектными работами».

Приоритеты проектов были определены на предыдущем этапе процесса УЗ, исходя из роли результатов проекта в комплексной СУЗ.

Учет допустимой хронологии работ предполагает, что

²⁵ Указаны приоритетные для СУЗ компоненты, некоторые другие компоненты не отражены.

анализируются технологические зависимости проектов и решается, как выстроить последовательность работ по проекту, чтобы полученные результаты могли использоваться в других проектах. Например, обучение сотрудников навыкам системного мышления должно предварять работы по разработке корпоративных онтологий, а не наоборот. В результате анализа технологических» зависимостей получаем предварительную программу внедрения СУЗ, которую можно представить в виде «дорожной карты», на которой каждая полоса объединяет определенную группу проектов, а по горизонтальной оси отложено время (рис. 5.12). На дорожной карте может также указываться взаимозависимость между проектами.

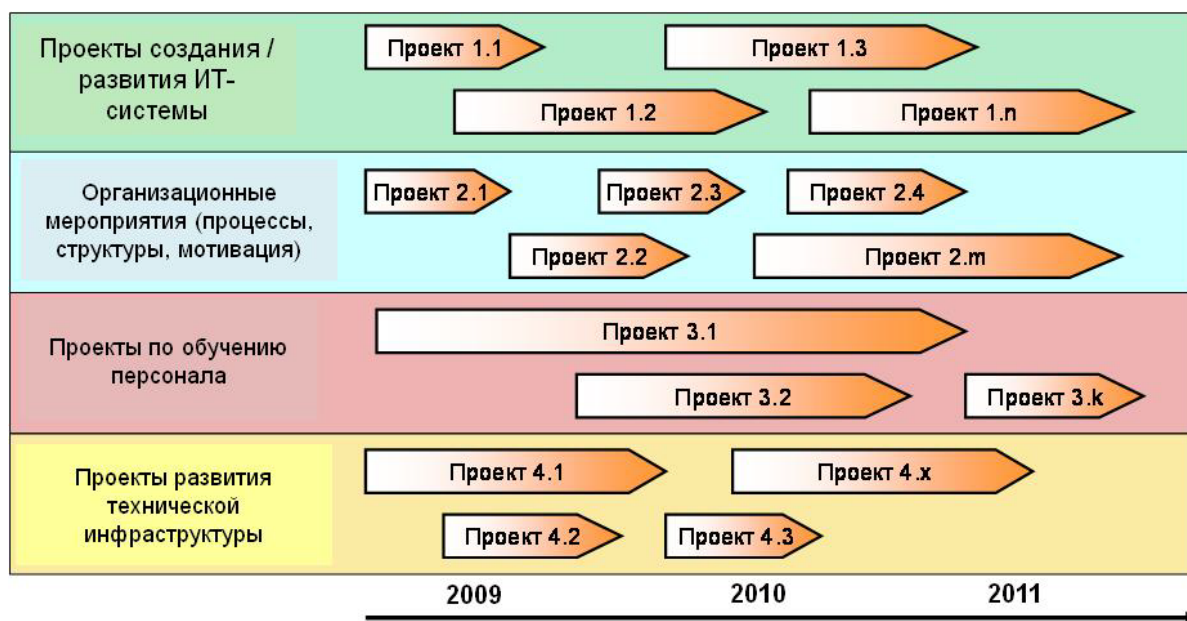


Рис. 5.12. Схема программы внедрения СУЗ в виде «дорожной карты»

Теперь полученную программу внедрения СУЗ необходимо согласовать с ресурсами, которыми располагает компания. Иными словами, перейти к этапу «учета ресурсных ограничений»

В последнее время наиболее распространенным ограничением на реализацию проектов становятся люди. В компании может не хватать необходимых компетенций для реализации проектов (рис. 5.13).

Профиль компетенций для команды внедрения СУЗ

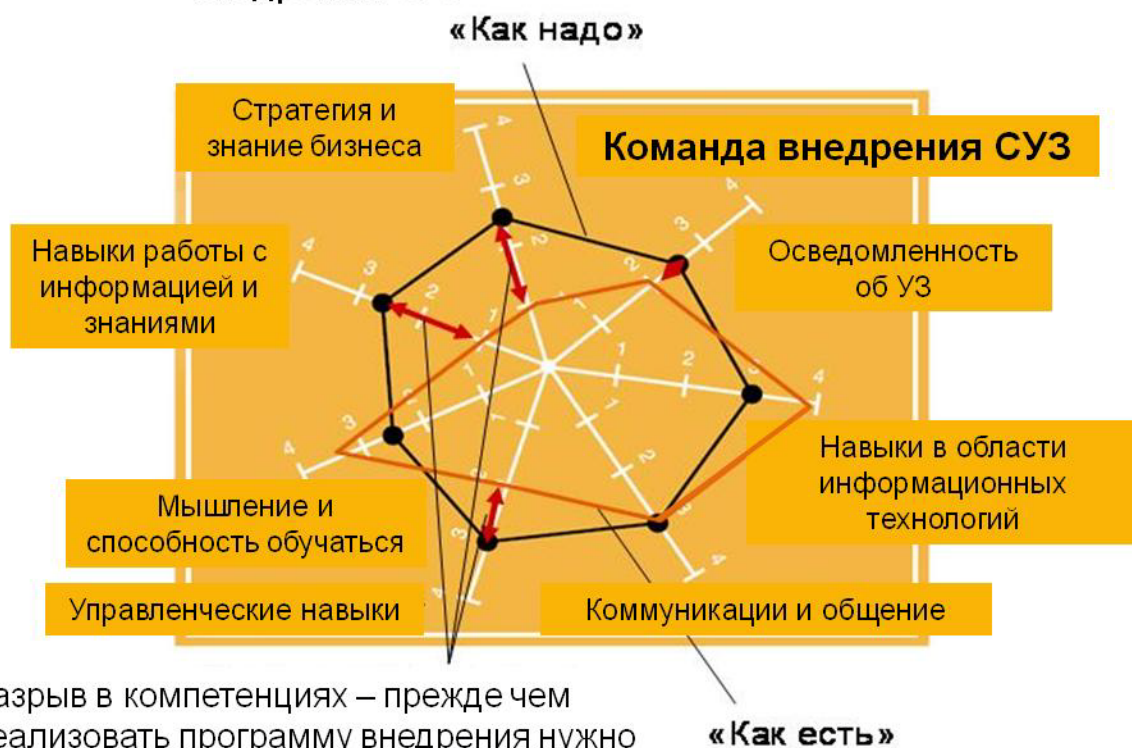


Рис. 5.13. Анализ требований к компетенциям команда внедрения СУЗ

Сотрудникам компании может не хватать времени на реализацию проектов, особенно с учетом наличия операционной деятельности (участие в процессах компании, а не только в проектах развития). Материальные и информационные ресурсы также могут выступать в качестве ограничений на выполнение мероприятий по реализации стратегии. И конечно, компании необходимо согласовать состав проектов по УЗ с имеющимися финансовыми ресурсами. Для этого следует составить бизнес-планы всех планируемых проектов, получить инвестиционный бюджет и согласовать его с операционным бюджетом, иначе может оказаться, что премии сотрудникам платить будет нечем. Это в свою очередь ставит под вопрос дальнейшие проектные работы и внедрение СУЗ.

В итоге разрешения всех указанных противоречий компания должна получить программу внедрения СУЗ, в котором будут:

- указаны стратегические цели, показатели их достижения и

плановые значения показателей;

- определены программы и проекты по достижению этих целей;
- распределены ресурсы между отобранными программами и проектами.

5.4. Контроль и оценка

В рамках *контроля и оценки* производится мониторинг исполнения мероприятий по созданию или развитию СУЗ (сроки, бюджет, качество...), оценивается с помощью специальных качественных и количественных показателей состояние знаний, характеристик СУЗ, ее влияние на основную деятельность предприятия и достижение целей.

При построении системы контроля и оценки знаний очень важно дать ответ на следующие вопросы: *зачем* необходимо оценивать знания? *Как* их оценивать? *Что* конкретно оценивать (какие компоненты, аспекты и процессы знаний)? *Кто* будет оценивать? *Когда* оценивать?

Зачем оценивать знания?

Знания оцениваются, чтобы ими управлять. Нельзя управлять тем, что нельзя измерить. Здесь главная заинтересованная сторона – это менеджеры и сотрудники компании. Им оценка знаний необходима для точной идентификации объектов улучшения (оптимизации) в организации. Кроме того, оценка знаний компании поможет в оценке же работы сотрудников и их объективной мотивации. Она будет также способствовать сбалансированному распределению и планированию ресурсов (финансы/инвестиции, люди, время). Но главный результат заключается в том, что оценка знаний поможет в разработке и уточнении стратегии компании.

Кроме того, мы оцениваем знания компании для ее подготовки к выходу на фондовый рынок и с последующим обеспечением покупки и продажи ее акций. В этом случае главными потребителями результатов оценки знаний становятся инвесторы, банки и т. п.

Как оценивать знания?

Знания оцениваются с помощью комбинации качественных и количественных показателей. Количественные показатели имеют численное выражение, а качественные показатели передают информацию в

текстовой или описательной форме, которая может включать в себя как утверждения о фактах, так и выражать определенные мнения. Как правило, качественные измерители преобладают в системах оценки знаний на ранних этапах их построения. По мере роста опыта использования этих систем в них начинают преобладать количественные показатели.

Стоит отметить, что существуют различные формы оценки знаний. Это мониторинг эффективности или результативности деятельности, оценка в форме аттестации, контроль качества и внутренний контроль.

Что оценивать?

После того, как мы определились, зачем нам оценивать знание и как его оценивать, нам стоит понять, что собственно оценивать. Во-первых, следует оценивать ***осведомленность*** сотрудников организации об управлении знаниями, о наличии и размещении знаний, о существующих средствах поддержки работы со знаниями. Во-вторых, оценивается ***деятельность в области УЗ и ресурсы, затрачиваемые на данную деятельность***. Здесь производится мониторинг мероприятий (проектов) по УЗ, оценивается наличие необходимых процессов, политик, должностей, сообществ, программных средств, а также размер бюджета, количество вовлеченных сотрудников и т. п. Следующий, третий, объект оценки – ***непосредственные результаты УЗ***: количество баз, отчетов, обобщающих опыт, обсуждений в форумах, методических материалов, использование средств УЗ. Последним объектом оценки в области УЗ являются ***конечные результаты или стратегические последствия*** данной деятельности: повышение продаж новой продукции, увеличение качества, снижение затрат, рост прибыли и т. п.

Каждый из данных объектов измерения включает в себе группу показателей. Все группы показателей взаимосвязаны (рис. 5.14). Так показатели осведомленности являются драйверами для показателей деятельности и ресурсов, которые в свою очередь являются показателями результата для показателей осведомленности и драйверами для показателей непосредственных результатов и т. д.

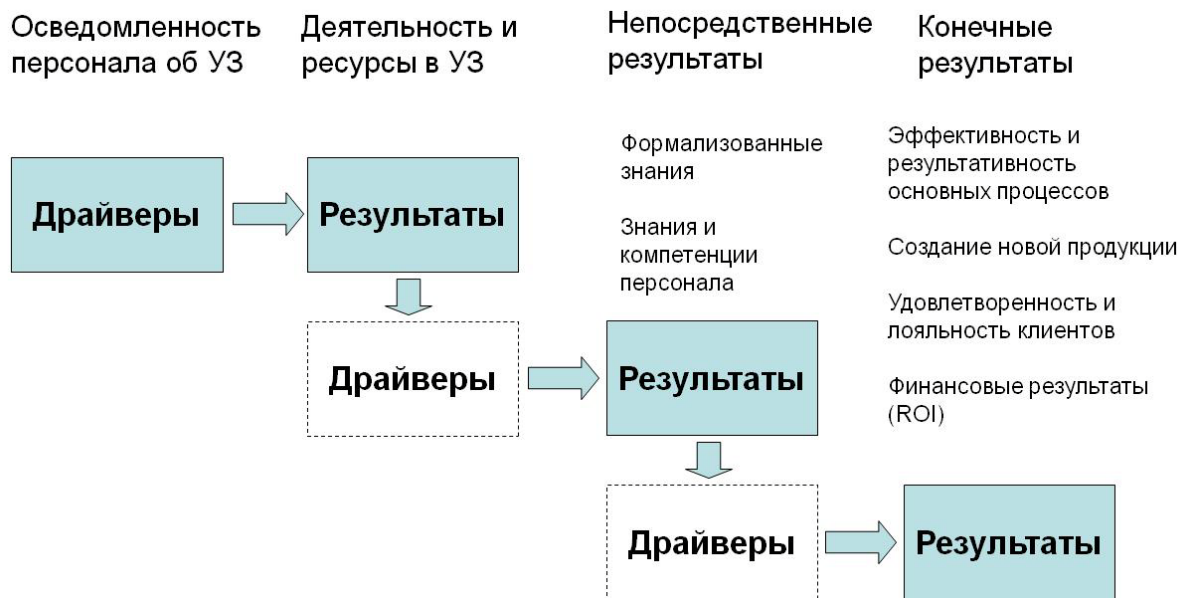


Рис. 5.14. Взаимосвязь показателей оценки знаний и СУЗ

Для поиска стандартных показателей и критериев оценки для каждой группы полезно соотнести их со стандартными подходами к оценке.

Группы 1 и 2, осведомленность в области УЗ, а также деятельности в данной области, характеризуются уровнем зрелости организации в УЗ.

Группы 3 и 4, непосредственные и конечные результаты УЗ, традиционно относятся к развитой сфере оценки интеллектуального капитала.

Оценка зрелости организации в УЗ включает формирование определенного набора стандартов, а также выбор или построение метода сравнения со стандартами.

На основе набора стандартов выделяются *пять степеней зрелости*:

1. Первоначальный уровень. Организация, находящаяся на первом уровне зрелости, имеет слабое представление о работе и управлении знаниями. Положение вещей в такой компании может иметь следующие характеристики:

- Индивидуальные, нескоординированные попытки использования методологий, инструментов и техники УЗ,
- Ограниченный масштаб инициатив, связанных с УЗ,
- Минимальное вовлечение работников,
- Невысокое использование внешних источников знаний и опыта, связанных с УЗ (например, учебники, статьи, консультанты).

2. Уровень осознания. Компания на данном уровне уже осознала необходимость работы и выработки определенных подходов к УЗ и, как правило, имеет какой-то минимальный опыт в этой области. Скорее всего, в ближайшем будущем такая компания будет наращивать компетенции в этой области и увеличивать количество людей, которые смотрят на свою компанию как на совокупность интеллектуальных ресурсов. Обычно такая компания обладает определенной комбинацией следующих характеристик:

- Осознание важности описания и управления знаниями,
- Ключевые интеллектуальные ресурсы компании идентифицированы,
- Вовлечение топ-менеджмента в УЗ,
- Активное использование простых моделей и техник инженерии и управления знаниями,
- Первые попытки использования структурированных методологий и общих стандартов,
- Увеличивающееся использование внешних источников знаний и опыта, связанных с УЗ.

3. Уровень измеряемости. Компания на третьем уровне зрелости постоянно наращивает свои компетенции в области УЗ и увеличивает количество людей, которые знакомы с концепцией УЗ и рассматривают свою компанию уже с новой перспективы.

- Использование разработанного инструментария УЗ,
- Использование комбинаций методов и инструментов УЗ,
- Проведение обучения персонала в области УЗ,
- Снижение использования внешних источников информации и опыта по УЗ.

4. Уровень управляемости. Как правило, компания на четвертом уровне зрелости УЗ обладает хорошо развитой системой УЗ и может быть охарактеризована следующими признаками:

- Наличие методов оценки знаний и СУЗ,
- Построение ИТ инфраструктуры и структуры организации в соответствии со стратегией на основе знаний,
- Широкое применение методов и инструментов УЗ,
- Ориентация на использование интеллектуальных ресурсов компании для достижения стратегических целей компании,

- Минимальная зависимость от внешних источников знаний и опыта по УЗ.

5. *Уровень оптимизации (совершенствования)*. На данном уровне компания имеет полноценную СУЗ компании, которая разумно встроена в другие элементы деятельности организации. Обычно такая компания обладает определенной комбинацией следующих характеристик:

- УЗ является частью управленческой деятельности, отчетности и измерения производительности,
- Широкое принятие и использование стандартных методов и технологий УЗ.

Другими словами, пятый уровень зрелости компании – это тот желаемый уровень развития, который необходим для того, чтобы компания смогла достичь своих стратегических целей.

Для *оценки интеллектуального капитала* за последние 10-15 лет было предложено множество систем, при этом особое внимание в них уделялось нефинансовым показателям. Все эти методологии можно разделить на четыре группы.

1. *Методы прямой оценки интеллектуального капитала* (Direct Intellectual Capital Methods, DICM) оценивают денежную ценность нематериальных активов при помощи выявления различных ее компонентов. После определения этих компонентов они могут быть оценены непосредственно по отдельности или при помощи агрегирующего коэффициента.

2. *Методы рыночной капитализации* (Market Capitalization Methods, MCM) рассчитывают ценность интеллектуального капитала как разницу между рыночной капитализацией компании и ценностью ее акционерного капитала.

3. *Методы определения доходности активов* (Return on Assets Methods, ROA) рассчитывают среднюю прибыль компании до уплаты налогов и делят ее на усредненную ценность материальных активов компании. В результате получают коэффициент ROA, который потом сравнивается со средними показателями по отрасли. Разница умножается на усредненную ценность материальных активов для расчета средней прибыли от нематериальных активов. Разделив полученную вышесреднюю прибыль на средневзвешенную ценность капитала компании или ставку

процента, можно получить приблизительную ценность ее интеллектуального капитала.

4. Методы балльных оценок (Scorecard Methods, SC) выявляют различные составляющие интеллектуального капитала, а затем определяются индикаторы и индексы и представляются в виде оценочных карт или графиков. Методы SC схожи с методами DIC за исключением того, что не производится оценка денежной ценности интеллектуального капитала. При этом может также рассчитываться сводный индекс.

Для примера можно рассмотреть один из методов оценки интеллектуального капитала – *система сбалансированных показателей* (Balanced Scorecard, BSC), которая входит в группу методов балльных оценок (SC). Система была разработана двумя учеными Нортон и Капланом в 1992 году и до сих пор активно цитируется и используется на практике. В данной модели эффективность работы компании оценивается при помощи индикаторов, отражающих четыре основных фокусных направления: 1) финансовые результаты; 2) рынки и внешнее окружение (клиенты); 3) внутренние процессы; 4) обучение (см. рис. 5.15).

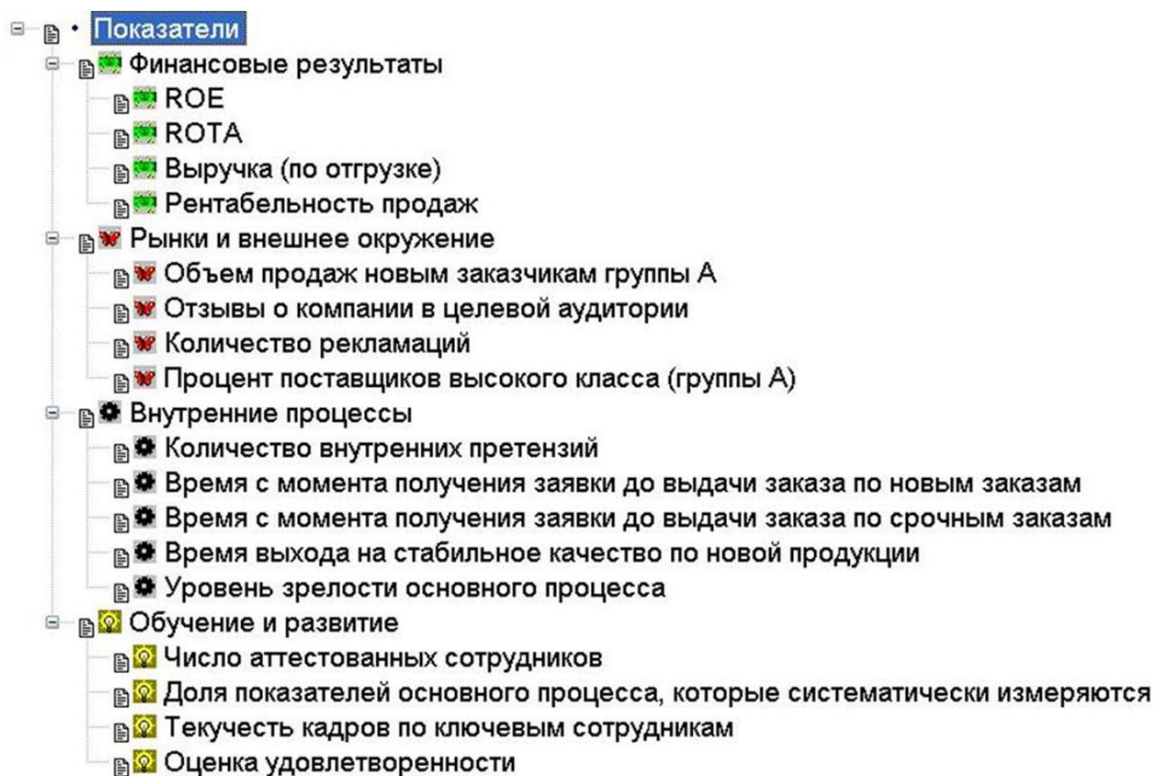


Рис. 5.15. Пример системы сбалансированных показателей

Под сбалансированностью понимается, что учитываются не только финансовые результаты, но и нефинансовые показатели. Индикаторы определяются на основе стратегических целей фирмы.

Кто будет оценивать знания?

Как правило, оценкой знаний в компании занимаются сами сотрудники. В этом случае мы получаем оценку внутренней деятельности компании. Менеджеры оценивают эффективность, результативность и инновационность компании, а также созданные активы знаний. Клиенты и поставщики осуществляют внешнюю оценку.

Когда осуществлять оценку знаний?

В идеальном варианте оценка должна проходить периодически для всей организации. Плюс в конце и в начале проекта по управлению знаниями.

Пример Hewlett-Packard

Объекты измерений:

- Профили выполненных проектов – позволяют использовать результаты выполненных проектов в получении и реализации новых проектов (продажи);
- Сообщества практиков – позволяют сконцентрировать детальные материалы в «горизонтальных» и «вертикальных» областях;
- Форумы – обеспечивают персональное взаимодействие участников сообществ;
- Отчеты – короткие технические статьи для передачи знаний;
- Совместная работа групп – обмен информацией и знаниями с использованием 2-х инструментов, предложенных отделом УЗ.

Уровни измерений: страна, консалтинговая практика (направление услуг), область знаний.

Перечень показателей:

7 индексов для ежемесячного анализа

3 индекса для оценки накопления знаний:

- $PPR\text{-index} = \text{РПП-индекс (Репозиторий профилей проектов)} = (\# \text{ завершенных профилей проектов с начала года} / \# \text{ новых проектов с начала года}) * 100$

- KB-index = Индекс создания отчетов = (# созданных отчетов / # консультантов в регионе)*12 (для годового анализа)
 - Forum contribution index = Индекс участия в форумах = (# постов в форуме / # консультантов в регионе)*12 (для годового анализа)
- 4 индекса для оценки использования инструментов УЗ:
 Инструмент (WSS²⁶/Groove²⁷/Сообщество практиков – CoP/Форум)

индекс

$$= (\# \text{ активных пользователей} / \text{общее } \# \text{ консультантов}) * 100\%$$

где «активный пользователь» инструмента – человек, который хоть раз просмотрел/скачал/загрузил что-либо в инструменте УЗ в последнем месяце.

²⁶ Windows SharePoint Services – часть SharePoint, предоставляющая функциональность для совместной работы и управления документами с помощью веб-портала.

²⁷ Microsoft Office Groove – инструмент для совместной работы

6. СОЗДАНИЕ ФОРМАЛИЗОВАННЫХ ЗНАНИЙ

В процессе создания возникают явные и неявные знания, то есть знания в головах сотрудников, а также формализованные²⁸ знания, выраженные и зафиксированные вне человеческого разума с помощью какой-либо знаковой системы, как правило, в форме бумажных или электронных документов или записей. Неявные знания сотрудников создаются в рамках процессов обучения, практической деятельности, мозговых штурмов, а также инсайтов и озарений. Данная тема требует отдельного детального рассмотрения и не рассматривается в рамках данного пособия. В данном разделе рассматривается создание формализованных знаний, которое может осуществляться как людьми, так и с помощью компьютерных технологий (рис. 6.1).



Рис. 6.1. Способы создания формализованных (явных) знаний

²⁸ В данном случае «формализованные» трактуется более широко, по сравнению с трактовкой, используемой в рамках искусственного интеллекта и представления знаний.

Создание formalизованных знаний людьми может осуществляться непосредственно специалистами-носителями неформализованных знаний с помощью структурированного ввода информации. При наличии нескольких специалистов формализующих свои знания важной становится технология совместного формирования контента. Кроме самостоятельного создания formalизованных знаний, может применяться их опосредованное создание с привлечением аналитиков или инженеров по знаниям, которые извлекают и формализуют знания экспертов. Также создание formalизованных знаний может осуществляться с помощью записи и распознавания человеческой речи.

В случае создания знаний с помощью компьютера, в основном, путем выявления знаний в больших массивах данных с использованием методов интеллектуального анализа данных. В последнем варианте компьютерная программа находит новые связи или закономерности предметной области, например, с помощью статистических или логических методов.

6.1. Структурированный ввод информации

Создаваемые formalизованные (явные) знания, пригодные для дальнейшей компьютерной обработки, могут, как относиться к двум крайностям – сильно-структурированные и неструктурированные, так и занимать промежуточное положение (что чаще). Создание сильно-структурированных знаний происходит при наполнении базы знаний (онтологии) экспертной системы ее экземплярами или при вводе записей в базу данных. Наполнение баз данных может осуществляться как при непосредственной работе с ними, так и при работе с приложениями, например, с системой автоматизированного проектирования. К неструктурированным знаниям относятся документы, в которых отсутствует какое-либо явное структурирование (например, как в художественном романе). Создание знаний в структурированном виде позволяет упростить и снизить трудоемкость работы по структурированию в дальнейшем. В свою очередь, чем лучше структурировано знание, тем лучше оно подходит для поиска, адресной доставки, усвоения. Для повышения структурированности создаваемых знаний используются

формы и шаблоны документов, которые могут заполняться специалистами в рамках выполнения своих текущих обязанностей внутри бизнес-процессов. Формируемое таким образом знание имеет среднюю степень структурированности, занимая промежуточное положение между базами данных/знаний и сплошным текстом (см. табл. 6.1).

Таблица 6.1

Степени структурированности информации

| Степень | Примеры | Технологии |
|-----------------------------------|-------------------------------------|---|
| Сильно-структурированная | База данных, база знаний | Реляционные базы данных, хранилища RDF-троек, формы для ввода |
| Частично-структурированная | Форма документа Шаблон документа | XML-документы |
| Неструктурированная | Сплошной текст | Текстовый файл |

Форма и шаблон документа – очень близкие понятия. Однако форма относится к более структурированным областям, и ее поля обычно связаны с моделью предметной области. Шаблоны документов оставляют больше свободы для авторов, элементы структуры в шаблоне будут более обобщенными. Примерами форм являются форма заказа продукции, заявка на работы, паспорт процесса/проекта. Примерами шаблонов могут являться типовая структура бизнес-плана, отчета по научно-исследовательской работе и т. п. Например, в соответствии с «ГОСТ 7.32-2001 Отчет о научно-исследовательской работе (НИР). Структура и правила оформления» структурными элементами отчета являются: титульный лист; список исполнителей; реферат; содержание; определения; обозначения и сокращения; введение; основная часть; заключение; список использованных источников; приложения. Обязательные структурные элементы выделены полужирным шрифтом. Остальные структурные элементы включают в отчет по усмотрению исполнителя НИР с учетом требований разделов 5 и 6. Указанные структурные элементы могут иметь свою внутреннюю структуру, так текст реферата должен отражать: объект исследования или разработки; цель работы; метод или методологию проведения работы; результаты работы и их новизну; основные

конструктивные, технологические и технико-эксплуатационные характеристики; степень внедрения; рекомендации по внедрению или итоги внедрения результатов НИР; область применения; экономическую эффективность или значимость работы; прогнозные предположения о развитии объекта исследования.

В СУЗ часто сочетаются различные степени структурированности знаний, например, к неструктурированному документу может быть присоединена форма для ввода стандартных метаданных из Дублинского ядра (см. подраздел «Метаданные» в разделе «Накопление знаний»). Применение форм и шаблонов для создания знаний может существенно повысить уровень структурированности создаваемых (вводимых) знаний может быть получен без дополнительных усилий. Документы предприятий, как правило, не являются текстами в свободной форме, чаще они имеют предустановленную структуру, которая определяется корпоративными стандартами, требованиями системы менеджмента качества или требованиями к инженерной документации.

Работе со структурированными документами посвящены работы Боба Глушко²⁹.

Интеграция форм для ввода информации и семантических технологий выполнена в рамках исследовательского проекта X-Media (<http://www.x-media-project.org/>) в виде так называемых K-форм (K от англ. Knowledge – знание)³⁰. Данные формы являются частью инструментария работы со знаниями (<http://www.k-now.co.uk/>). Данные формы легко настраиваются и позволяют авторам вводить в них свои знания в веб-интерфейсе. Получаемые знания сразу же представляются с помощью RDF/OWL-моделей.

6.2. Совместное создание контента и использование вики

Наиболее распространенными в настоящее время способами совместного создания контента являются совместная работа с

²⁹ Document Engineering: Analyzing and Designing Documents for Business Informatics and Web Services. The MIT Press. 2005.

³⁰ Chapman S., Lanfranchi V., Bhagdev R. *K-Tools: Towards Semantic Knowledge Management*. In Proceedings of ESWC'2009. P. 806-810.

документами и использование вики.

Совместная работа над документами означает, что над отдельным документом или набором документов работают несколько авторов. Авторы могут редактировать документ одновременно или рецензировать спецификацию в рамках структурированного делового процесса. Совместное редактирование документа, являющееся частью совместной работы над документами, означает одновременную работу над документом вместе с еще одним или несколькими пользователями. Существует ряд методов совместной работы над документами и их совместного редактирования, которые отличаются друг от друга степенью структурированности и контроля. Чтобы выбрать оптимальный метод совместной работы над документами и оптимальный программный продукт, необходимо сопоставить свои требования с возможными вариантами.

Спектр задач совместной работы над документами:

1. Полуформальное совместное редактирование: несколько авторов одновременно вносят изменения в произвольные места документа. Примеры для программы OneNote: повторяющиеся протоколы, сеансы мозгового штурма и справочные материалы; примеры для Excel: совместно разрабатываемые финансовые модели, бюджеты и списки учета активов.

2. Формальное совместное редактирование: несколько авторов одновременно редактируют документ в рамках управляемого процесса, сохраняя материалы, когда они готовы к выпуску. Примеры для Word: бизнес-планы, информационные бюллетени и исковые заявления; примеры для PowerPoint: маркетинговые презентации и доклады для конференций.

3. Добавление примечаний и рецензирование: первоначальный автор обращается к другим пользователям с просьбой внести в документ изменения и примечания (это можно организовать в виде обсуждений), направляя документ в рабочий процесс. При этом за окончательную публикацию документа также отвечает первоначальный автор. Примерами таких документов являются справка в Интернете, технические документы и спецификации.

4. Наборы документов: в рамках рабочего процесса нескольким авторам назначаются отдельные документы, которые затем публикуются в едином наборе. Примеры: литература по новому продукту и брошюра по

продажам.

Вики (англ. wiki) – веб-сайт, структуру и содержимое которого пользователи могут сообща изменять с помощью инструментов, предоставляемых самим сайтом. Известнейший вики-сайт – Википедия. Впервые термин «вики» для описания веб-сайта был использован в 1995 году Уордом Каннингемом, разработчиком первой вики-системы WikiWikiWeb, который заимствовал слово гавайского языка, означающее «быстрый».

Уорд Каннингем и его соавтор Бо Леуф в их книге «The Wiki Way: Quick Collaboration on the Web» описали сущность концепции вики следующим образом:

- Вики предлагает всем пользователям редактировать любую страницу или создавать новые страницы на вики-сайте, используя обычный веб-браузер без каких-либо его расширений.
- Вики поддерживает связи между разными страницами за счет почти интуитивно понятного создания ссылок на другие страницы и отображения того, существуют данные страницы или нет.
- Вики не является тщательно изготовленным сайтом для случайных посетителей. Напротив, вики стремится привлечь посетителей к непрерывному процессу создания и сотрудничества, который постоянно меняет вид сайта.

Казалось бы, подобная «вседозволенность» должна приводить к полному уничтожению материалов и превращению любой вики в свалку спамерских ссылок. Тем не менее, этого не происходит, причем далеко не только из-за совестливости посетителей. В вики сохраняются все изменения, произошедших со всеми статьями с момента их создания. В любой момент можно сравнить две версии одной и той же статьи, увидеть, какие изменения были внесены, откорректировать страницу снова или просто вернуться к старой версии. Исправить нанесенный ущерб проще, чем его нанести. Ну и, конечно же, у администраторов всегда есть возможность заблокировать изменение той или иной страницы или разрешить редактирование только определенным пользователям.

Вики характеризуется следующими признаками:

- Возможность многократно править текст посредством самой вики-среды (сайта), без применения особых приспособлений на стороне

редактора. Для этого используется особый язык разметки – так называемая вики-разметка, которая позволяет легко и быстро размечать в тексте структурные элементы и гиперссылки; форматировать и оформлять отдельные элементы.

- Учет изменений (версий) страниц: возможность сравнения редакций и восстановления ранних.
- Проявление изменений сразу после их внесения.
- Разделение содержимого на именованные страницы.
- Гипертекстовость: связь страниц и подразделов сайта через контекстные гиперссылки.
- Множество авторов. Некоторые вики могут править все посетители сайта.

Для создания вики-среды необходимо особое ПО – движок вики. Это частный вид систем управления сайтом, довольно простой в своем устройстве и функциональности, ибо почти все действия по структурированию и обработке содержимого делаются пользователями вручную. Появились компании, занимающиеся разработками вики-приложений для корпоративной среды – такие, как Confluence, Jot или Near-Time. Wiki занимают и в Microsoft, например, SharePoint Portal Server 2007. На основе JotSpot построен крупнейший коммерческий вики-сайт, принадлежащий eBay. Сегодня Wiki используется в таких компаниях, как Lockheed Martin, Telenor, Citigroup, Sony или Toshiba.

Что такое вики и зачем она может использоваться, мы, надеюсь, разобрались. Давайте теперь поговорим о выборе. На самом деле, выбор – дело очень и очень сложное. На сегодня существует множество решений, приспособленных для выполнения определенных задач. Как и при выборе любого ПО, требуется, прежде всего, определить, что же требуется компании, какой набор функций и возможностей, на какой платформе всё это будет работать и т. д. В настоящее время вариантов множество, и выбор может оказаться непростым. В выборе конкретного решения вам может очень помочь сайт WikiMatrix.org, в базе которого собраны данные более чем по шести десяткам вики-движков.

Для выбора вики-движка, прежде всего, необходимо определить требуется ли размещение вики на собственном сервере организации

или же подойдет вики-сервис, предлагаемый той или иной компанией. На Западе подобных вики-сервисов достаточно много. Многие компании, занимающиеся корпоративными wiki-решениями, предлагают бесплатно завести вики у них на сайте. Для корпоративного пользователя это будет не более чем тест-драйв, а для пользователя рядового, частного, может оказаться вполне достаточным. Другой вопрос, что на русскую локализацию можно и не рассчитывать – у нас пока просто нет рынка корпоративных вики. У нас, если и делают корпоративную вики, в большинстве случаев используют открытую и бесплатную WackoWiki, правда сейчас этот проект находится в подвешенном состоянии.

В любом случае, вики-хостинг поможет, как минимум, понять, подходит ли wiki для решения ваших задач, и какие функции вам действительно необходимы.

Ну а дальше вам предстоит выбор платформы и основных функций. Решите, нужен ли вам контроль версий, ограничения доступа, RSS-ленты с информацией об изменении страниц, подписка на изменения конкретных страниц и т. д. Количество параметров для сравнения внушительно. Впрочем, если интересно, можно посмотреть на матрицу на http://en.wikipedia.org/wiki/Comparison_of_wiki_software.

Семантическая вики – вики, модель знаний которой описана на ее страницах.

Обычные вики заполняются структурированным текстом и нетипизированными гиперссылками (такими, как в этой статье). Семантические вики позволяют указывать тип ссылок между статьями, тип данных внутри статей, а также информацию о страницах (метаданные).

Цель семантической вики: обеспечить машинную обработку вики с минимальными усилиями со стороны пользователей. Это может быть реализовано в виде расширения MediaWiki, например Semantic MediaWiki. В этом расширении для запросов используется язык SPARQL. Система Semantic MediaWiki написана с помощью механизма расширений MediaWiki. Это упрощает интеграцию в существующие приложения MediaWiki.

Семантическая вики предоставит следующие элементы для разметки статей:

- типизированные ссылки,

- категории,
- атрибуты – свойства содержимого статей.

Типизированные ссылки (триплеты) описывают свойства, связываемых RDF ресурсов (статей). Вот пример типизированной ссылки в статье London: `[[capital of::England]]`. Это утверждение означает, что Лондон является столицей Англии:

субъект – имя страницы (London);

отношение – столица (capital of);

объект – страна (England).

Число типов ссылок не ограничено, пользователь может добавлять новые (принцип аналогичен категориям Википедии), решение о добавлении нового типа принимает сообщество. Поиск будет возможен, только если разные типы ссылок используются повторно (как и категории – то есть, не уникальны). Так же можно указать несколько типов отношений для одной ссылки, например `[[тип_1::тип_2:: ... ::тип_т::целевая статья]]`. Семантическая разметка помещается внутри текста. Сохраняется возможность указывать в ссылке видимый текст, отличный от имени статьи. Например, такая ссылка `[[is capital of::United Kingdom|UK]]` будет выглядеть как UK.

Типизированные ссылки позволяют выполнять: прямой запрос (столица Англии?), логический вывод (Лондон – столица Англии) => (Лондон находится в Англии), агрегирование поисковых критериев в запросе (Лондон – город, Англия – в Европе) => (Лондон – европейский город).

Работа поисковой базы данных триплетов (triplestore) включает три этапа: извлечение типизированных ссылок из текста статьи, преобразование в RDF триплеты; обновление базы данных триплетов.

Атрибуты описывают отношения между статьей и значениями переменных. Например, в той же статье London атрибут area позволяет указать площадь Лондона в квадратных милях: `[[area:=609 square miles]]`. В отличие от типизированных ссылок, значения атрибутов должны дополнительно распознаваться (на этапе парсинга). Еще одна задача заключается в приведении к общим единицам измерения. При этом желательно небольшое число разных единиц измерения. В расширение Semantic Wikipedia встроен преобразователь для известных единиц

измерения.

Движки семантической вики – Semantic MediaWiki, POWL, WikiOnt и ряд других.

6.3. Извлечение знаний

Если структурированный ввод предполагал непосредственное создание формализованных знаний самим автором, то извлечение знаний позволяет формализовать экспертные знания опосредованно с помощью инженера по знаниям (аналитика). Инженер по знаниям путем взаимодействия с экспертом(-ами) и анализа текстовых источников информации создает первичное формализованное знание, которое в дальнейшем «обретает форму» с помощью структурирования.

В работе Гавриловой и Червинской³¹ была предложена классификация методов извлечения знаний (рис. 6.2), позволяющая инженерам по знаниям в зависимости от конкретной задачи и ситуации выбирать наиболее подходящий метод.

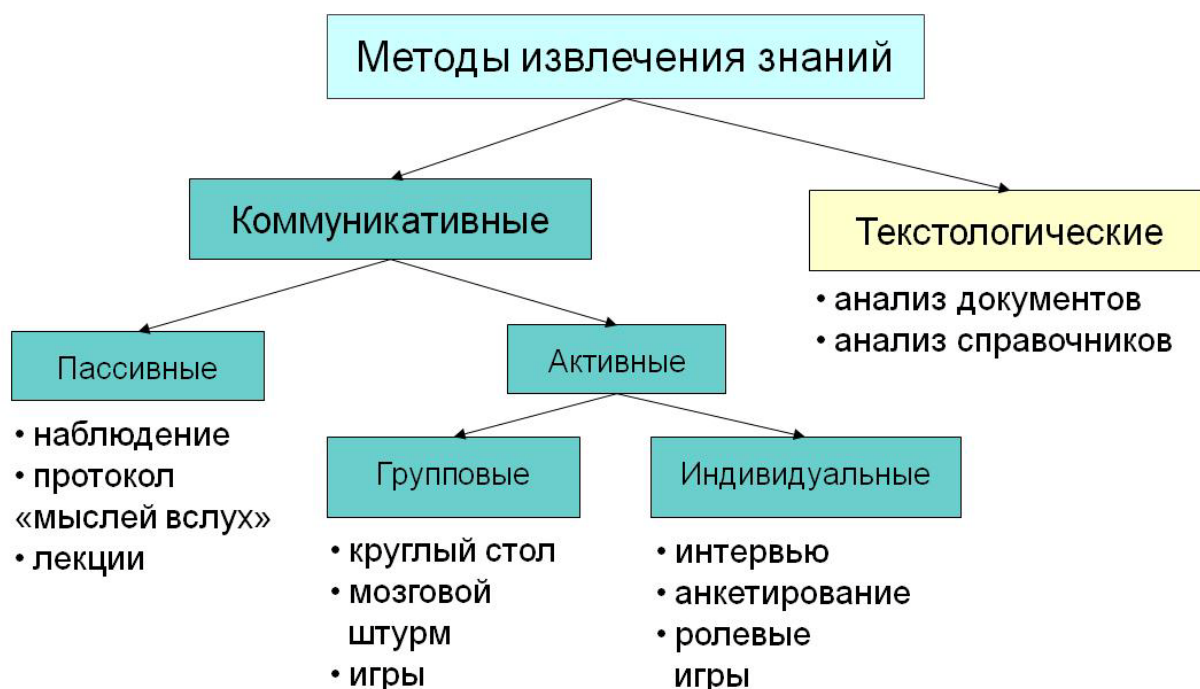


Рис. 6.2. Методы извлечения знаний

³¹ Гаврилова Т.А., Червинская К.Р. Извлечение и структурирование знаний для экспертных систем. М.: Радио и связь. 1992.

Рассмотрим подробнее соответствующие группы методов и рекомендации по их использованию. Из предложенной схемы классификации видно, что основной принцип деления связан с источником знаний. Коммуникативные методы извлечения знаний охватывают методы и процедуры контактов инженера по знаниям с непосредственным источником знаний – экспертом, а текстологические включают методы извлечения знаний из документов (методик, пособий, руководств) и специальной литературы (статей, монографий, учебников).

Разделение этих групп методов на верхнем уровне классификации не означает их антагонистичности, обычно инженер по знаниям комбинирует различные методы, например, сначала изучает литературу, затем беседует с экспертами, или наоборот.

В свою очередь, коммуникативные методы можно также разделить на две группы: активные и пассивные. Пассивные методы подразумевают, что ведущая роль в процедуре извлечения как бы передается эксперту, а инженер по знаниям только протоколирует рассуждения эксперта во время его реальной работы по принятию решений или записывает то, что эксперт считает нужным самостоятельно рассказать в форме лекции. В активных методах, напротив, инициатива полностью в руках инженера по знаниям, который активно контактирует с экспертом различными способами – в играх, диалогах, беседах за «круглым столом» и т. д.

Следует еще раз подчеркнуть, что и активные, и пассивные методы могут чередоваться даже в рамках одного сеанса извлечения знаний. Например, если инженер по знаниям застенчив и не имеет большого опыта, то вначале он может использовать пассивные методы, но постепенно, ближе знакомясь с экспертом, захватывать инициативу и переходить «в наступление».

Пассивные методы на первый взгляд достаточно просты, но на самом деле требуют от инженера по знаниям умения четко анализировать поток сознания эксперта и выявлять в нем значимые фрагменты знаний. Отсутствие обратной связи (пассивность инженера по знаниям) значительно ослабляет эффективность этих методов, чем и объясняется их обычно вспомогательная роль при активных методах.

Таблица 6.2

Сравнение пассивных методов извлечения знаний

| Метод | Достоинства | Недостатки |
|--------------------|---|--|
| Наблюдения | Нет влияния аналитика Погружение аналитика | Отсутствие обратной связи |
| Мысли вслух | Свобода самовыражения Обнаженность рассуждений Нет влияния аналитика | Фрагментарность Отсутствие обратной связи «Зашумленность» деталями |
| Лекции | Свобода самовыражения Нет влияния аналитика Структурированное изложение Концентрация | Слабая обратная связь Нет хороших лекторов среди экспертов-практиков |

Активные методы можно разделить на две группы в зависимости от числа экспертов, отдающих свои знания: индивидуальные и групповые. В табл. 6.3 и 6.4 анализируются достоинства и недостатки отдельных активных методов извлечения знаний.

Таблица 6.3

Сравнение активных индивидуальных методов извлечения знаний

| Метод | Достоинства | Недостатки |
|-------------------------|--|--|
| Анкетирование | Стандартизация опроса экспертов Легко для аналитика | Требует умения составления анкет Нет обратной связи |
| Интервью | Обратная связь, возможность уточнения деталей | Длительное время на подготовку |
| Свободный диалог | Гибкость Сильная обратная связь Возможность изменения сценария | Трудно для аналитика Отсутствие методик Трудно протоколировать |

Таблица 6.4

Сравнение групповых методов извлечения знаний

| Метод | Достоинства | Недостатки |
|------------------|--|--|
| «Круглый стол» | Объективные фрагменты знаний Оживляет процедуру Обмен знаний между участниками | Большие организационные затраты Сложность проведения |
| «Мозговой штурм» | Глубинные пласты знаний Активизация экспертов Получение нового знания (гипотезы) | Только для новых исследовательских проблем Низкая эффективность |

6.4. Распознавание речи

Человеческая речь является источником знания. Средства распознавания речи предназначены главным образом для преобразования неявных знаний сотрудников в явные. Так, рабочая дискуссия, обсуждение, в процессе которых с большой вероятностью открываются неявные знания сотрудников, могут быть зафиксированы (например, в виде аудио- или видеозаписи). Затем применяется средство автоматического распознавания речи для преобразования обсуждения в текстовый аналог.

Приложения, распознающие речь, могут принадлежать одной из трех категорий:

- системам диктовки или создания документов;
- системам навигации или транзакций (системы с речевым ответом);
- мультимедийным индексным системам.

В системах диктовки все слова, произносимые пользователем, переводятся в текст. В системах навигации слова, произносимые пользователем, могут использоваться для перехода по ссылкам в Интернете или для навигации в приложении. В транзакционных системах слова, произносимые пользователем, используются для осуществления транзакций, таких, как покупка акций, банковские транзакции и т. д. В мультимедийных индексных системах речь используется для

преобразования аудиозаписи (например, извлеченной из видеозаписи) в текст.

Хотя наиболее удачные приложения в этой области связаны с относительно ограниченными областями и словарями и требуют высококачественных аудиозаписей, в последнее время точность систем автоматического распознавания речи увеличилась.

6.5. Специализированные службы по созданию знаний

Распространенной формой создания знаний является организация специализированных служб, таких, например, как отделы исследований и разработок. Их главная задача – создавать новые знания и новые способы выполнения работ.

В литературе приводятся примеры таких служб, созданных в некоторых консалтинговых фирмах. Компания *Ernst & Young* создала для этих целей *Центр бизнес-инноваций*; *Andersen Consulting* организовала *технологические исследовательские центры* в Силиконовой долине и во Франции; *IBM Consulting* сформировала *группы разработок по ключевым направлениям консультирования*. Крупные корпорации (*Motorola, Merck* и др.) также создают службы, которые в большей или меньшей степени связаны с созданием знаний. Некоторые *корпоративные библиотеки* также представляют собой варианты научно-исследовательских отделов, которые предназначены для снабжения организации новым знанием.

Серьезной проблемой специализированных служб является сложность передачи полученных знаний в те подразделения корпорации, где они могут быть применены. Часто это связано с тем, что создатели и пользователи знаний говорят как бы на разных языках, а истинную ценность новых знаний трудно измерить. Решение этой проблемы зависит от менеджеров, которые должны своевременно принимать меры к тому, чтобы знания, созданные специализированными службами, стали доступными организации. Для этого используются и формальные средства передачи знаний определенному кругу подразделений, и регулярно проводимые на высоком уровне собрания, конференции, семинары, целью которых являются оценка и интеграция новых знаний в знания организации.

6.6. Выявление новых знаний в базах данных

Потребность в аналитической обработке сверхбольших объемов информации, накапливаемой в современных хранилищах данных, обусловила появление и активное развитие технологий интеллектуального анализа данных. Этот термин является приближенным переводом с английского языка терминов Knowledge Discovery in Databases и Data Mining.

Интеллектуальный анализ данных (ИАД) обычно определяют как технологию извлечения новых знаний из хранилищ данных путем выявления закономерностей в накопленной информации. Эта технология не заменяет аналитиков и менеджеров, а дает им мощный и современный инструмент для поддержания принятия решений. Обзор технологий ИАД дан на основе материала А. В. Блинниковой.

Ключевой возможностью применения технологий ИАД стало снижение за последние несколько лет цены на устройства хранения информации – с десятков долларов за хранение мегабайта информации до десятков центов. Это существенно удешевило и увеличило возможности сбора и хранения больших объемов информации.

Падение цен на процессоры с одновременным увеличением их быстродействия также способствует развитию технологий, связанных с обработкой огромных массивов информации. В результате было преодолено множество барьеров, стоявших на пути нахождения нового знания в хранилищах информации.

Клиент-серверная архитектура также является необходимым атрибутом технологий ИАД. Она предоставляет возможность выполнять наиболее трудоемкие процедуры обработки данных на высокопроизводительном сервере.

Процессы ИАД можно разделить на три большие группы: поиск зависимостей, прогнозирование, анализ аномалий.

Поиск зависимостей состоит в просмотре базы данных с целью автоматического выявления зависимостей. Проблема заключается в отборе действительно важных зависимостей из огромного числа существующих в БД.

Прогнозирование предполагает, что пользователь может предъявить системе записи с незаполненными полями и запросить недостающие значения. Система сама анализирует содержимое базы и делает правдоподобное предсказание относительно этих значений.

Анализ аномалий – это процесс поиска нетипичных данных, сильно отклоняющихся от устойчивых зависимостей.

Системы ИАД применяются в научных исследованиях и образовании, в работе правоохранительных органов, производстве, здравоохранении и многих других областях. Особенно широко технология ИАД используется в деловых приложениях.

В основу технологии ИАД положен не один подход, а несколько принципиально различных подходов, причем использование некоторых из них невозможно без специальной подготовки. Выбор подхода нередко требует привлечения специалиста по ИАД.

Получение данных включает итеративно создаваемые модели на основе подготовленного множества данных, а затем применение одной или нескольких моделей. Поскольку создание моделей на больших множествах данных может оказаться весьма дорогостоящим, аналитики часто сначала работают с несколькими выборками множества данных.

Интеллектуальный анализ данных состоит из следующих этапов (рис. 6.3):

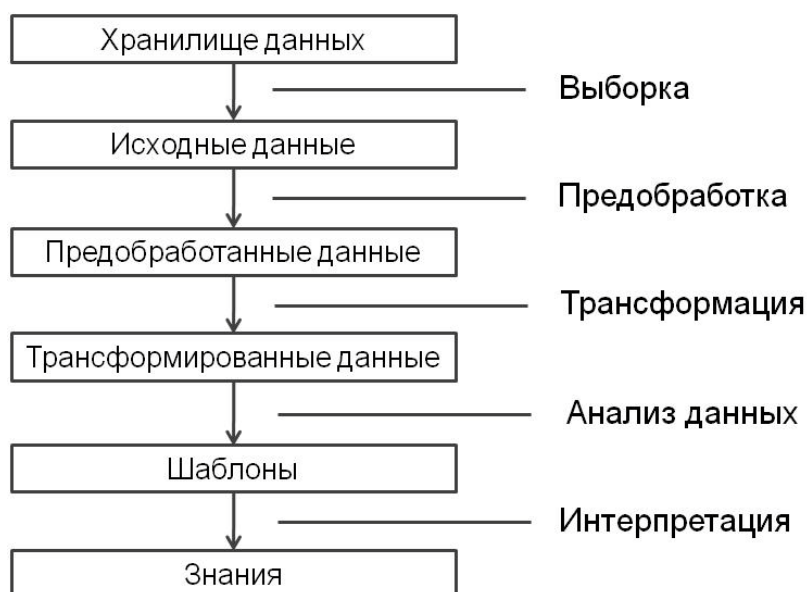


Рис. 6.3. Этапы интеллектуального анализа данных

1. Подготовка исходного набора данных. Этот этап включает создание набора данных, в том числе из различных источников, выбор обучающей выборки и т. д. Для этого должны существовать развитые инструменты доступа к различным источникам данных.

2. Предобработка данных. Для того чтобы эффективно применять методы ИАД, следует обратить серьезное внимание на вопросы предобработки данных. Данные могут содержать пропуски, шумы, аномальные значения и т. д. Кроме того, данные могут быть избыточны, недостаточны и т. д. В некоторых задачах требуется дополнить данные некоторой априорной информацией. Данные должны быть качественны и корректны с точки зрения используемого метода. Более того, иногда размерность исходного пространства может быть очень большой, и тогда желательно применение специальных алгоритмов понижения размерности.

3. Трансформация, нормализация данных. Этот этап необходим для тех методов, которые требуют, чтобы исходные данные были в каком-то определенном виде. Нейронные сети, скажем, работают только с числовыми данными, причем они должны быть нормализованы.

4. Анализ данных. На этом этапе применяются различные алгоритмы для нахождения знаний: нейронные сети, деревья решений, алгоритмы кластеризации и установления ассоциаций и т. д.

5. Постобработка данных. Это интерпретация результатов и применение полученных знаний в бизнес-приложениях.

Целью технологии ИАД является производство нового знания, которое пользователь может в дальнейшем применить для улучшения результатов своей деятельности.

ИАД – это процесс обнаружения в исходных данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности. ИАД использует сложный статистический анализ и моделирование для нахождения моделей и отношений, скрытых в базе данных, – таких отношений, которые не могут быть найдены обычными методами.

У профессионалов обработки данных часто возникает вопрос о разнице между средствами ИАД и средствами оперативной аналитической

обработки данных (OLAP). OLAP – это часть технологий, направленных на поддержку принятия решения. Технология OLAP используется для объяснения предполагаемых взаимосвязей. При этом пользователь сам формирует гипотезу о данных или отношениях между ними и после этого использует серию запросов к базе данных для подтверждения или отклонения этой гипотезы.

Средства ИАД отличаются от средств OLAP тем, что они позволяют выдвигать новые гипотезы о характере неизвестных, но реально существующих отношений между данными. Кроме того, средства ИАД вместо проверки предполагаемых взаимосвязей на основе имеющихся данных могут формировать модели, позволяющие количественно оценить степень влияния исследуемых факторов. В целом инструментарий оперативной добычи и аналитической обработки данных позволяет проводить развернутый анализ информации, что, в свою очередь, позволяет получить более эффективный результат процесса обнаружения новых знаний в хранилищах данных.

OLAP-технологии обычно применяются на ранних стадиях процесса ИАД. Это позволяет лучше понять полученные данные, помогает аналитикам выявить актуальные порции данных, а технологии ИАД обогащают эту функциональность.

Можно выделить шесть стандартных типов отношений данных, которые позволяют выявлять методы ИАД: классификацию, регрессию, прогнозирование временных последовательностей (рядов), кластеризацию, ассоциацию, последовательность. Первые три используются главным образом для предсказания, в то время как последние удобны для описания существующих закономерностей в данных.

1. С помощью *классификации* выявляются признаки, характеризующие группу, к которой принадлежит тот или иной объект. Это делается посредством анализа уже классифицированных объектов и формулирования некоторого набора правил. Предполагается, что классы заранее известны.

2. *Регрессионный анализ* используется в том случае, если отношения между переменными могут быть выражены количественно в виде некоторой комбинации этих переменных. Полученная комбинация далее используется для предсказания значения, которое может принимать

выходная (зависимая) переменная, вычисляемая на заданном наборе значений входных (независимых) переменных. В простейшем случае для этого используются стандартные статистические методы, такие, как линейная регрессия. К сожалению, большинство реальных моделей не укладываются в рамки линейной регрессии.

3. *Прогнозирование временных последовательностей* позволяет на основе анализа поведения временных рядов оценить будущие значения прогнозируемых переменных. Основой для всевозможных систем прогнозирования служит историческая информация, хранящаяся в БД в виде временных рядов. Если удастся построить или найти шаблоны, адекватно отражающие динамику поведения целевых показателей, есть вероятность, что с их помощью можно предсказать и поведение системы в будущем.

4. *Кластеризация* отличается от классификации тем, что сами группы заранее не заданы. С помощью кластеризации средства ДМ самостоятельно выделяют различные однородные группы данных. Объекты внутри кластера должны быть похожими друг на друга и отличаться от объектов, вошедших в другие кластеры. Чем больше похожи объекты внутри кластера и чем больше отличий между кластерами, тем точнее кластеризация.

5. *Ассоциация* имеет место в том случае, если несколько событий связаны друг с другом, что позволяет выявить закономерности между ними. Примером такой закономерности служит правило, указывающее, что из события X следует событие Y. Такие правила называются ассоциативными. Впервые эта задача была предложена для нахождения типичных шаблонов покупок, совершаемых в супермаркетах, поэтому иногда ее еще называют анализом рыночной корзины.

6. Если существует цепочка связанных во времени событий, то говорят о *последовательности*.

Интеллектуальные средства анализа возникли и развиваются на стыке таких дисциплин, как статистика, распознавание образов, искусственный интеллект, теория информации, теория баз данных и др. Отсюда обилие методов и алгоритмов, реализованных в различных действующих системах ИАД. Многие из таких систем интегрируют сразу несколько подходов.

Обратимся к основным методам, которые используются для проведения интеллектуального анализа данных. К ним относятся: нейронные сети, деревья решений, индукция правил.

Кроме этих методов существует еще несколько дополнительных: системы рассуждения на основе аналогичных случаев, нечеткая логика, генетические алгоритмы, алгоритмы определения ассоциаций и последовательностей, анализ с избирательным действием, логическая регрессия, эволюционное программирование, визуализация данных.

Нейронные сети. Это большой класс систем, архитектура которых имеет аналогию с построением нервной ткани из нейронов. Нейросетевые технологии предоставляют сегодня широкие возможности для решения задач прогнозирования, обработки сигналов и распознавания образов. По сравнению с традиционными методами математической статистики, классификации и аппроксимации эти технологии обеспечивают достаточно высокое качество решений при меньших затратах. Они позволяют выявлять нелинейные закономерности в сильно зашумленных неоднородных данных, дают хорошие результаты при большом числе входных параметров и обеспечивают адекватные решения при относительно небольших объемах данных.

Деревья решений являются одним из наиболее популярных подходов к решению задач ИАД. Они создают иерархическую структуру классифицирующих правил типа «если... то...» (if – then), имеющую вид дерева. Для принятия решения, к какому классу отнести объект или ситуацию, требуется ответить на вопросы, стоящие в узлах этого дерева, начиная с его корня. Вопросы имеют вид: «Значение параметра L больше x ?» Если ответ положительный, осуществляется переход к правому узлу следующего уровня, если отрицательный – к левому узлу; затем снова следует вопрос, связанный с соответствующим узлом. Популярность подхода связана как бы с наглядностью и понятностью. Но деревья решений принципиально не способны находить «лучшие» (наиболее полные и точные) правила в данных. Они реализуют примитивный принцип последовательного просмотра признаков и «цепляют» фактически осколки настоящих закономерностей, создавая лишь иллюзию логического вывода.

Индукция правил создает неиерархическое множество условий,

которые могут перекрываться. Она осуществляется путем генерации неполных деревьев решений, а для того чтобы выбрать, какое из них будет применено к входным данным, используются статистические методы.

Идея *систем рассуждения на основе аналогичных случаев* (прецедентов) крайне проста. Для того чтобы сделать прогноз на будущее или выбрать правильное решение, эти системы находят в прошлом близкие аналоги наличной ситуации и выбирают тот же ответ, который был для них правильным. Поэтому данный метод называют еще методом «ближайшего соседа» (nearest neighbour). Системы рассуждения на основе аналогичных случаев показывают очень хорошие результаты при решении самых разнообразных задач. Главный их минус заключается в том, что они вообще не создают каких-либо моделей или правил, обобщающих предыдущий опыт, – в выборе решения они основываются на всем массиве доступных исторических данных, поэтому невозможно сказать, на основе каких конкретно факторов эти системы строят свои ответы.

Системы визуализации многомерных данных. В той или иной мере средства для графического отображения данных поддерживаются всеми системами ДМ. В подобных системах основное внимание сконцентрировано на дружелюбности пользовательского интерфейса, позволяющего ассоциировать с анализируемыми показателями различные параметры диаграммы рассеивания объектов (записей) базы данных. К таким параметрам относятся цвет, форма, ориентация относительно собственной оси, размеры и другие свойства графических элементов изображения. Кроме того, системы визуализации данных снабжены средствами, удобными для масштабирования и вращения изображений. Стоимость систем визуализации может достигать нескольких сотен долларов.

Ключом успеха в применении методов ИАД являются качество данных, мощность используемого программного обеспечения и мастерство аналитика, который участвует в процессе построения модели. На эффективность обработки данных положительно влияют такие параметры программного обеспечения, как развитые средства формирования запросов и визуализации результатов, наличие графического инструментария и т. д.

Сами методы ИАД являются лишь основой для реализации типовых видов анализа, используемых аналитиками и лицами, принимающими

решения в бизнесе. К этим видам анализа относятся:

1. Ранжирование – упорядочение факторов по степени их влияния на целевой показатель. С каждым фактором связывается весовой коэффициент, дающий количественную оценку степени влияния.

2. Сегментация – разделение области значений фактора на сегменты для проведения дальнейшего нисходящего анализа. Например, успешные показатели по продажам в целом могут скрыть от исследователя неудачи в некоторых регионах. Проведение сегментации по географическим районам позволит выявить, что успешные продажи имели место только в двух регионах – А и В. Последующий нисходящий анализ может показать, что объем продаж в регионе А зависит от затрат на рекламу, а в регионе В на него влияют главным образом сезонные колебания.

3. Профилирование наилучших достижений – выявление основных характеристик наиболее успешных регионов, филиалов, клиентов и т. д. для планирования дальнейшей деятельности компании. Данный вид анализа используется также в банках и страховых компаниях для оценки риска при выдаче кредитов.

4. Выявление ассоциаций – поиск ассоциированных групп значений факторов, т. е. значений, почти всегда появляющихся вместе. Этот вид анализа необходим в первую очередь при планировании продаж, усовершенствовании работы складов, проведении рекламных кампаний.

5. Выявление исключений – поиск элементов, выпадающих из общей картины. Появление подобных элементов может быть вызвано как ошибками в данных, которые следует исправить, так и необычными ситуациями в работе компании, требующими немедленного вмешательства со стороны руководства.

Назовем типовые задачи интеллектуального анализа данных: прогнозирование, маркетинговый анализ, анализ работы персонала, анализ эффективности продаж, профилирование клиентов, оценка потенциальных клиентов, анализ результатов маркетинговых исследований, анализ работы региональных отделений компании, сравнительный анализ конкурирующих фирм.

Очевидно, что перечисленные виды задач актуальны практически для всех отраслей бизнеса: банковского дела и страхования (выявление злоупотреблений с кредитными карточками, оценка кредитных рисков,

оценка закладных, выявление профилей пользователей, оценка эффективности региональных отделений, вероятность подачи заявки на выплату страховки и др.), финансовых рынков (прогнозирование, анализ портфелей, моделирование индексов), производства (прогнозирование спроса, контроль качества, оценка дизайна продукции), торговли и т. д.

7. НАКОПЛЕНИЕ ЗНАНИЙ

В основе процесса накопления знаний лежит описание и структурирование имеющейся в организации информации, внешних ресурсов, а также неявных знаний сотрудников. Именно описание и структурирование позволяет в дальнейшем найти и доставить нужные знания потребителям.

7.1. Структурирование знаний

Стадия концептуального анализа или структурирования знаний традиционно является (наряду со стадией извлечения) «узким местом» в жизненном цикле разработки интеллектуальных систем. Методологии структурирования в основном опираются на современную теорию больших систем или сложных систем, где акцент ставится на стадии проектирования. Достаточно условно методы структурирования знаний можно разделить на визуальные и символьные. Символьные методы опираются на математическое, либо текстовое, либо табличное описание знаний. Математическое описание знаний, базирующееся на дескрипционной логике, будет рассмотрено далее в разделе, посвященном представлению знаний с помощью языков семантического веба. К текстовым методам структурирования можно отнести описания информационно-поисковых языков, которые также будут рассмотрены далее. В данном разделе остановимся на табличных и визуальных методах структурирования знаний.

Таблица решений (таблица принятия решений) – способ компактного представления модели со сложной логикой. Аналогично условным операторам в языках программирования, она устанавливает связь между условиями и действиями. Но, в отличие от традиционных языков программирования, таблицы решений в простой форме могут представлять связь между множеством независимых условий и действий. Таблицы принятия решений, как правило, разделяются на четыре квадранта: *Условия* – список возможных условий, *Варианты выполнения условий* – комбинация из выполнения и/или невыполнения условий из этого списка. *Действия* – список возможных действий, *Необходимость действий* –

указание надо или не надо выполнять соответствующее действие для каждой из комбинаций условий (рис. 7.1).

| Условия | Варианты выполнения условий | | |
|-------------------------------|-----------------------------|-----|-----|
| | Да | Нет | Нет |
| Свет в соседней комнате горит | Да | Нет | Нет |
| Свет у соседей горит | - | Да | Нет |
| Поменять лампочку | X | | |
| Проверить пробки | | X | |
| Позвонить электрику | | X | X |
| Позвонить диспетчеру | | | X |

Действия
Необходимость действий

Рис. 7.1. Пример таблицы решений

Например, для ситуации «свет погас», но «свет в соседней комнате горит» необходимо «поменять лампочку». Вариантов выполнения условия может быть не два: да или нет, а несколько, например цвет может быть красным, оранжевым, синим. В более сложных таблицах может применяться нечеткая логика. Действия могут быть элементарными или ссылаться на другие *таблицы принятия решений*. Необходимость выполнения действий может быть неупорядоченной, как в данном примере, или упорядоченной. В последнем случае если при определенной комбинации выполнения условий возможно выполнение нескольких действий, то в таблице решений указывается их приоритет.

Другим примером табличного метода структурирования знаний является QFD (акроним от англ. Quality Function Deployment), или структурирование (развертывание) функции качества. Данный метод помогает передавать знания и принимать решения при разработке продуктов или услуг. По мнению создателей, QFD может помочь организации сосредоточить внимание на важнейших характеристиках новых или существующих продуктов или услуг с точки зрения отдельного клиента, сегмента рынка, компании, или технологии развития. Результатами применения методики являются понятные схемы и матрицы, которые могут быть повторно использованы для будущих товаров либо услуг. QFD трансформирует потребности клиентов (голос клиента) в

инженерные характеристики продукции, расставляет приоритеты для каждого продукта / услуги и одновременно определяет задачи в области развития продукции или услуги.

Основа QFD – построение фигурной матрицы, названной в соответствии со своей формой «Дом качества» (см. рис. 7.2), в рамках которой фиксируется информация о качестве продукта и принимаемых решениях.

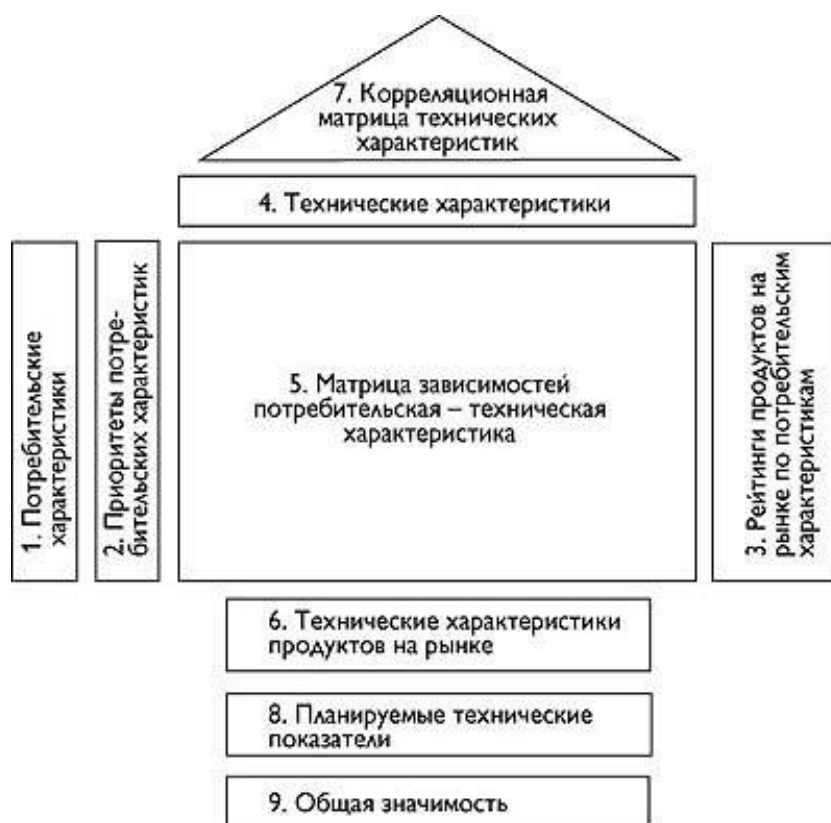


Рис. 7.2. Схема дома качества

Центральная часть дома – это таблица, столбцы которой соответствуют техническим характеристикам, а строки потребительским. В клетках отмечается уровень зависимости, если она есть. Крышу дома представляют сведения о корреляции между техническими характеристиками. Левое крыло – столбец приоритетов пользовательских характеристик. Правое крыло – таблица рейтингов потребительских характеристик (с точки зрения пользовательского восприятия) для существующих на рынке подобных продуктов. Подвал дома содержит результаты анализа технических характеристик конкурирующих продуктов, результаты выработки стратегии изменения технических

характеристик своего продукта (планируемые показатели для первоначальной разработки), оценки абсолютной и относительной важности.

Предпосылками QFD являются маркетинговые исследования, определяющие, что хочет пользователь, насколько важны те или иные качества (левое крыло, шаги 1 и 2), а так же, как решают подобные проблемы другие поставщики (правое крыло, шаг 3). Каждому продукту, включая свой текущий, конкурентов, свой перспективный по каждому требованию присваивается рейтинг. После определения набора технических характеристик (шаг 4), заполняется центральная часть дома – определяются зависимости между потребительскими и техническими характеристиками (шаг 5). На шестом шаге анализируется уровень реализации в конкурирующих продуктах. После анализа взаимной корреляции технических характеристик (шаг 6), исходя из полученных сведений, формируются целевые показатели для разрабатываемого продукта (шаг 7). Наконец для сопоставления значимости технических характеристик (шаг 8), формируется общая значимость, как сумма, в которой каждой непустой клетке в столбце матрицы зависимостей сопоставляется произведение приоритета соответствующего пользовательского показателя на уровень зависимости (рис. 7.3).

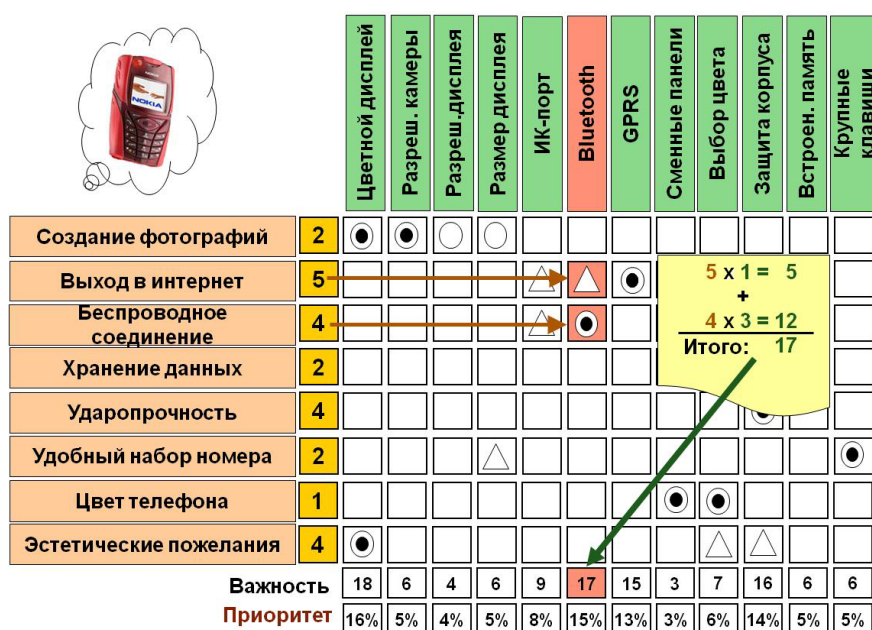


Рис. 7.3. Определение важности технических характеристик мобильного телефона с помощью QFD

QFD позволяет в компактной форме представить данные о разнообразных характеристиках продукта, а также отследить их влияние на принимаемые технические решения. В развернутом виде QFD включает четыре фазы, и на каждой из них строится свой дом качества. После преобразования потребительских характеристик в технические, последние преобразуются в характеристики компонентов, и далее: в операции процессов, а затем в параметры технологических процессов производства (рис. 7.4).

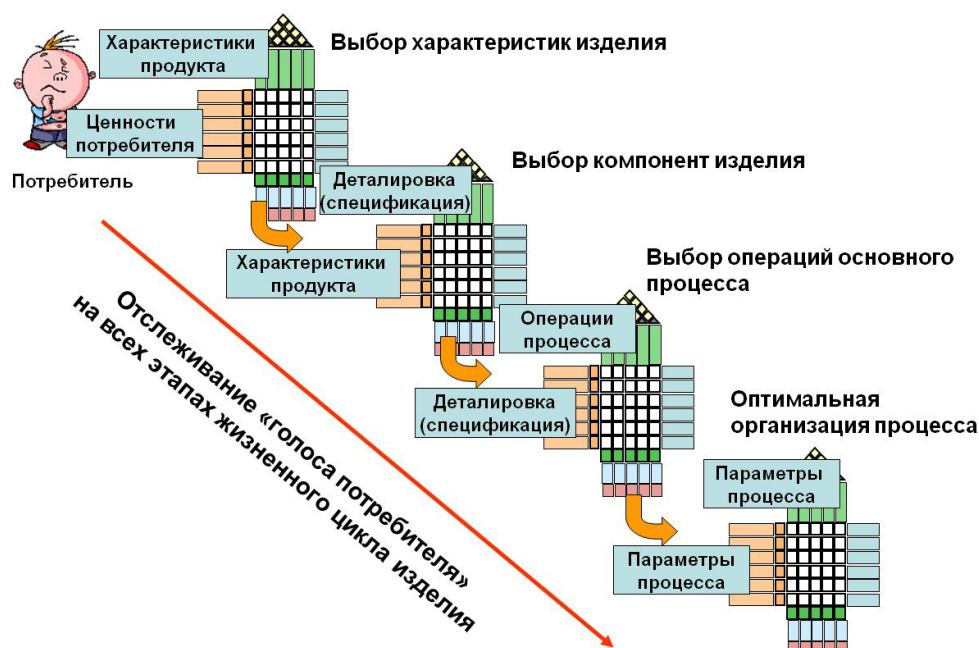


Рис. 7.4. Последовательное развертывание функции качества

QFD один из инструментов Бережливого производства и Системы менеджмента качества.

Визуальные методы структурирования знаний

Визуализация всегда считалась мощным инструментами познания, т. е. средством, предназначенным для организации и облегчения процесса познания. На практике, используются более сотни методов визуального структурирования – от интеллект-карт (mind maps), диаграмм бизнес-процессов до стратегических и «дорожных» карт (roadmaps). Такое многообразие обусловлено существенными различиями в природе, особенностях и свойствах знаний различных предметных областей.

Существующие методы можно структурировать по аспектам рассмотрения предметной области – см. рис. 7.5.

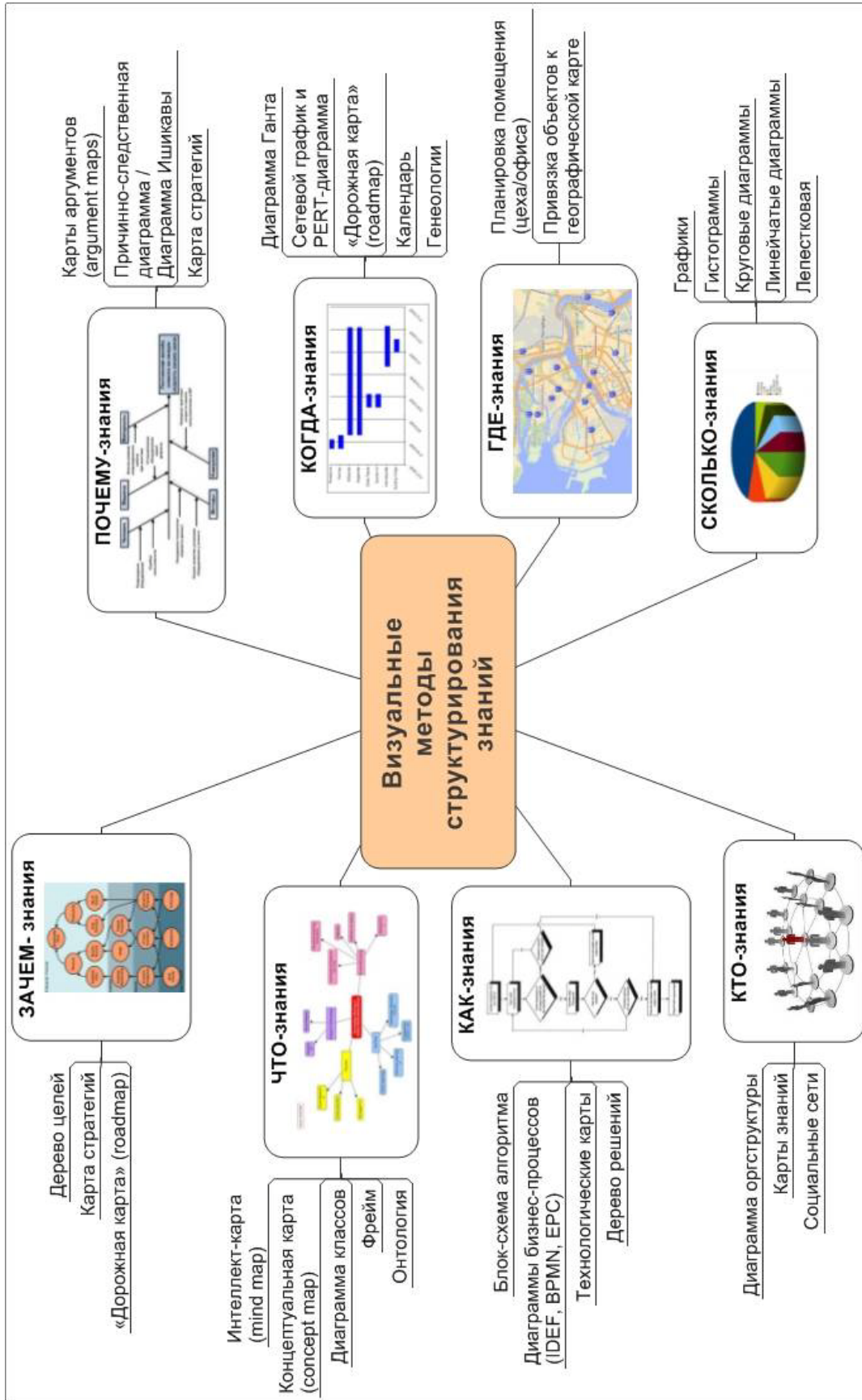


Рис. 7.5. Методы визуального структурирования знаний

Более подробно визуальные методы структурирования знаний рассмотрены в работах Гавриловой Т.А.

Интеграция табличных и визуальных методов структурирования знаний

Другим примером форм для наполнения онтологий являются, так называемые, классификаторы и проекции³², которые предоставляют формат для визуальной работы с экземплярами онтологии.

Визуальный язык классификаторов и проекций содержит следующие элементы:

$L ::= \langle \text{TYPES}, \text{CL}, \text{PR}, \text{TASKS}, \text{SPEC} \rangle$, где:

TYPES – описание типов, позволяющее представить таксономию верхнеуровневых классов, типы связей между классами, значения свойств классов. Каждому типу ставится в соответствие пиктограмма (графический знак).

CL – классификатор, формат ввода нижеуровневых классов и экземпляров онтологии, их свойств и иерархических связей между ними. Основные типы связей: «класс-подкласс», «часть-целое», «подчиняется» и т.п.;

$\text{CL} = \{ \text{CL}_i^{\text{name}}, \text{Onto-CL}_i, \text{Base-CL}_i \}$ – классификатор, где $\text{CL}_i^{\text{name}}$ – имя классификатора; Onto-CL_i – свойства классификатора – часть классификатора, задающая состав типов элементов, перечисляемых в классификаторе, перечень их свойств с областями допустимых значений и типы иерархических связей между элементами; Base-CL_i – содержание классификатора – часть классификатора, в которой перечисляются и типизируются моделируемые нижеуровневые классы, значения их свойств и задается древовидная система связей между классами.

PR – проекция, формат ввода связей между экземплярами онтологии, перечисленными в классификаторах. Примеры типов связей: «выполняет», «обеспечивает достижение», «отвечает за»;

$\text{PR}_j = \{ \text{PR}_j^{\text{name}}, \text{Onto-PR}_j, \text{Base-PR}_j \}$ – проекция, где $\text{PR}_j^{\text{name}}$ – наименование проекции; Onto-PR_j – свойства проекции – часть проекции, определяющая состав связей между типами, перечисляемых в

³² Кудрявцев Д.В., Григорьев Л.Ю. Визуальный язык классификаторов и проекций для разработки баз знаний // Труды 12-й Национальной конференции по искусственному интеллекту (КИИ-2010), 20-24 сентября, Тверь, Россия, 2010. – Т. 2 – С. 222-230.

классификаторе элементов; Base-PR_j – содержание проекции – часть проекции, содержащая множество связей между элементами, связываемых классификаторов.

TASKS – формат спецификации задач по разработке и использованию онтологии (не рассматривается в рамках данной статьи);

SPEC – формат спецификации запросов к онтологической модели (не рассматривается в рамках данной статьи).

Данный язык используется для структурирования и представления знаний, а также для формирования запросов к онтологической модели.

Язык визуального проектирования и наполнения баз знаний реализован в системе ОРГ-Мастер.

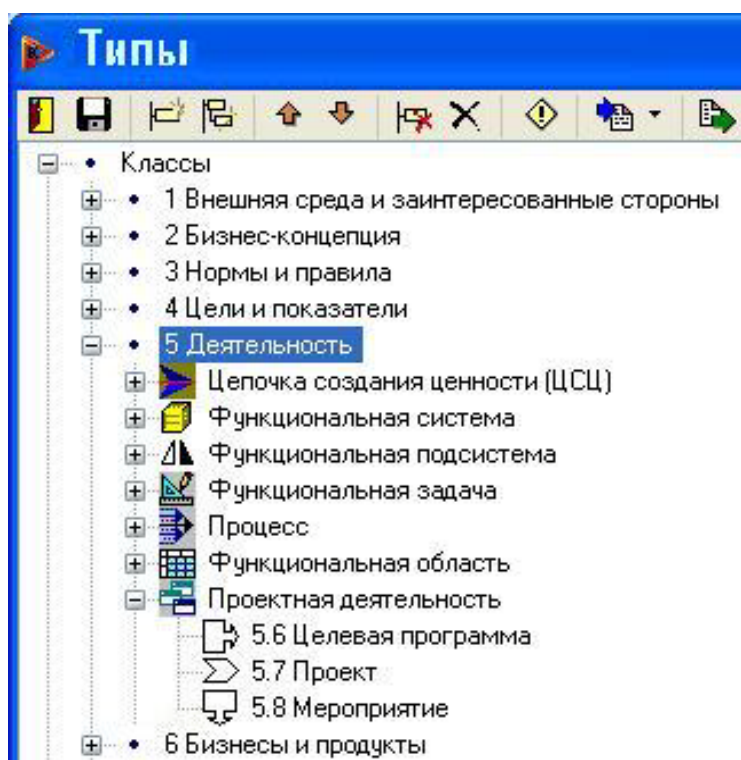


Рис. 7.6. Типы (TYPES) для разработки и визуализации верхнеуровневых классов

При вводе экземпляров онтологии (левое окно на рис. 7.7) происходит типизация элементов с помощью верхнеуровневой онтологии (правое окно на рис. 7.7). С точки зрения механизма визуализации, устанавливается связь между элементами 2-х иерархических списков.

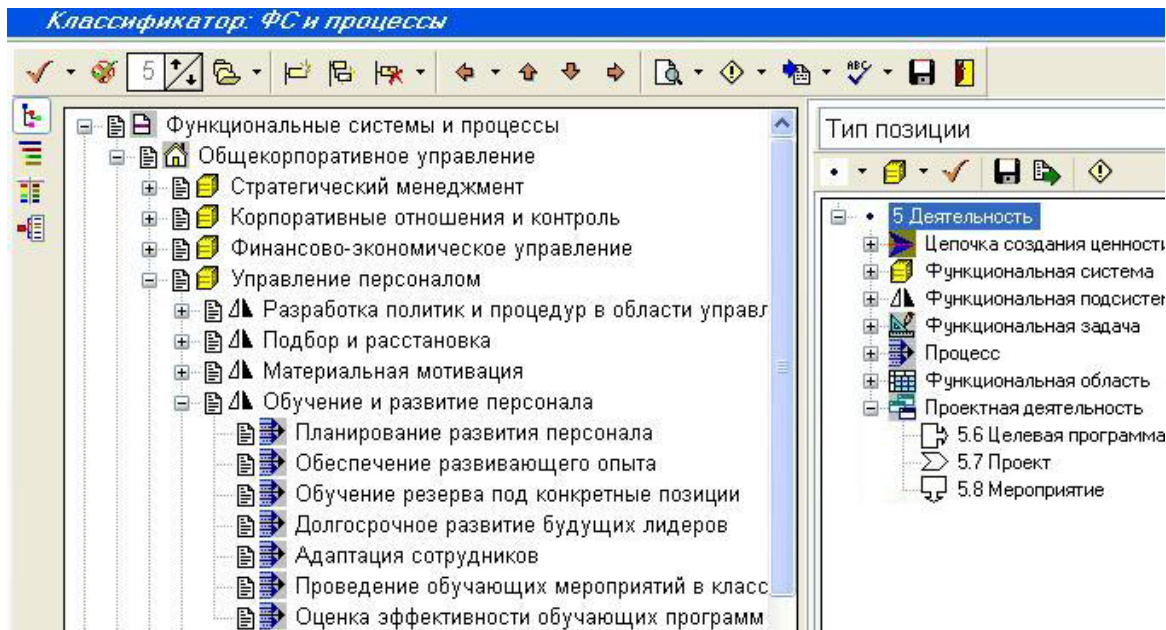


Рис. 7.7. Пример классификатора (CL) функциональных систем и бизнес-процессов предприятия

Для визуализации перекрестных связей между классами используются проекции (рис. 7.8). Связи в проекции могут быть направленными и ненаправленными, типизированными и нетипизированными.

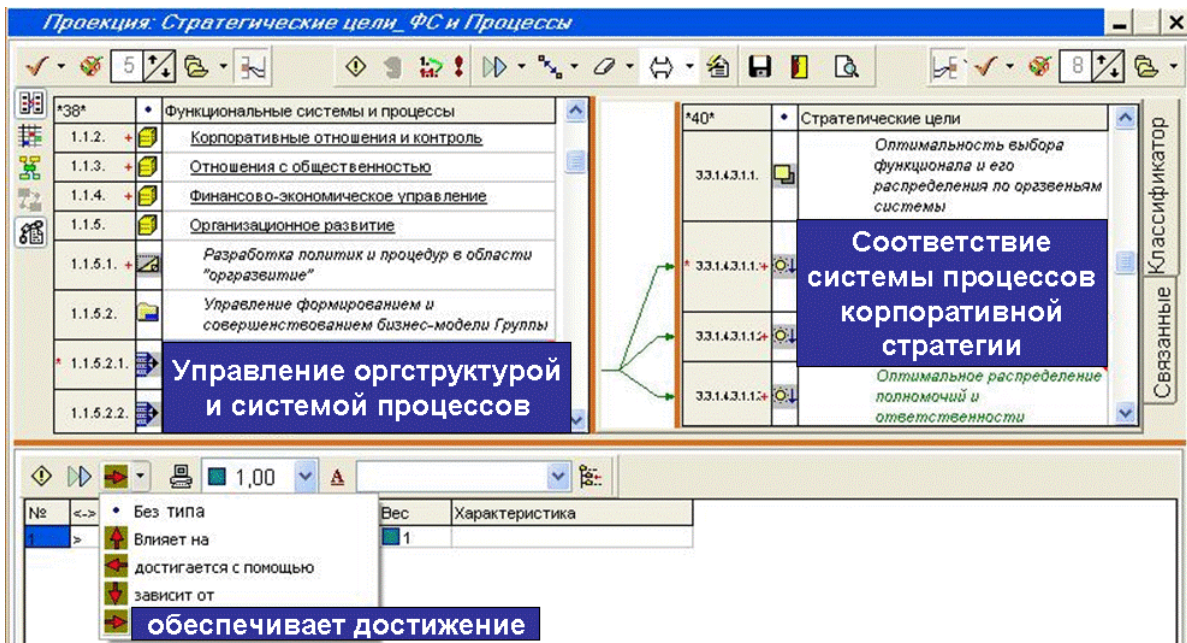


Рис. 7.8. Пример проекции (PR), связывающей бизнес-процессы предприятия с целями

Предложенный в статье визуальный язык представления онтологий основан на использовании иерархических списков и связей между списками. Иерархические списки используются как для представления классов, так и для представления экземпляров. Основные преимущества визуализации иерархическим списком – простота реализации и привычность для пользователя, поскольку такой же подход используется в многочисленных файловых браузерах (проводник Microsoft Windows, Total Commander...). Данный метод дает ясное представление имен классов и их иерархии. В отличие от других методов визуализации, в иерархическом списке имена классов не перекрывают друг друга и чтобы их увидеть, не нужно наводить курсор на соответствующий объект. Возможность свернуть/развернуть узел – полезное свойство для рассмотрения определенной части иерархии, особенно для масштабных онтологий. Простота таких интерфейсов делает их удобными для быстрого обзора. Основная проблема визуализации иерархическим списком – представление только древовидных структур, но не сетевых. В существующих редакторах онтологий, использующих визуализацию иерархическим списком, нет визуализации перекрестных связей. Они видны только как слоты классов. Однако данная проблема устранена в предложенном языке и технологии визуализации с помощью специального конструкта – проекции.

Методы, модели и инструменты структурирования знаний более детально рассматриваются в курсе «Инженерия знаний» или «Проектирование интеллектуальных систем» [Гаврилова, Муромцев, 2008].

7.2. Метаданные

Для представления сведений о ресурсах, в частности, о документах, базах данных, экспертах, используется понятие метаданных, описывающих содержимое ресурса в виде набора именованных значений, в том числе связей с другими ресурсами. Метаданные используются для автоматизированного анализа содержимого ресурса, построения поисковых индексов и позволяют обеспечить достаточно высокую точность и эффективность поиска разнородной информации.

Данный термин принято толковать, как «данные о данных», однако его значение распространяется помимо описания состава данных, их структуры (формата) представления, места хранения и других признаков

описания также на поддерживающие их информационные системы, технологии, пользователей, методы доступа и т. д. Особенно широко этот термин стал использоваться в последние годы в связи с развитием электронных библиотек, поскольку метаданные стали важнейшим средством обеспечения навигации и поиска информации. Однако до настоящего времени значение этого термина до конца четко не определено.

Существуют различные категории метаданных, например, описательные метаданные (descriptive metadata) (в том числе библиографические данные); метаданные о структурах и форматах (structural metadata); административные метаданные (administrative metadata), содержащие данные для управления доступом; идентификационные метаданные (identifier metadata), которые однозначно идентифицируют объекты внешнего мира и т. п. Существующие схемы организации метаданных могут рассматриваться в трех взаимосвязанных аспектах: семантическом, синтаксическом и структурном. Различие между данными и метаданными часто условно и зависит от контекста. Например, реферат с точки зрения терминологии электронных библиотек относится к метаданным, хотя в электронном каталоге или реферативной базе данных его содержание рассматривается как данные. Существуют различные концепции, стандарты и системы представления метаданных. Ниже приводятся только некоторые из них.

1. Dublin Core (DC) – «Дублинское ядро»: разработка ведется с 1995 г. рабочей группой с одноименным названием. Ею предложена простая структура описания документов, которая, по мнению разработчиков, должна заменить сложные системы существующей каталогизации документов. Текущая версия (DC 1.1) спецификации «Дублинского ядра» была принята в июле 1999 г.

В соответствии с нею структура описания данных включает 15 элементов:

1. Заголовок (Title) – название, присвоенное ресурсу его создателем;
2. Создатель (Creator) – лицо, организация или служба, ответственная за подготовку ресурса (автор, исполнитель, фотограф);
3. Предмет (Subject) – тема ресурса, выраженная ключевыми словами, элементами тематического справочника или короткой фразой;

4. Описание (Description) – текстовое описание ресурса;
5. Издатель (Publisher) – лицо, организация, или служба, обеспечивающая доступ к ресурсу;
6. Участник создания материала (Contributor) – человек или организация, которые не являются авторами, однако внесли в создание ресурса значительный интеллектуальный вклад помимо указанного в разделе «Создатель»;
7. Дата (Date) – дата, указывающая на создание или появление ресурса в доступном для использования виде;
8. Тип (Type) – жанр, категория ресурса, например домашняя страница, роман, статья и т. п.;
9. Формат (Format) – способ представления ресурса, например тип программного обеспечения и персонального компьютера, необходимых для отображения ресурса;
10. Идентификатор (Identifier) – например URL;
11. Источник (Source) – сведения о первичном источнике, из которого был продуцирован данный ресурс;
12. Язык (Language) – язык представления ресурса;
13. Связь (Relation) – ссылка на ресурс, связанный с данным; идентификатор вторичного ресурса и его связь с настоящим ресурсом, например издание книги и глава книги;
14. Охват (Coverage) – области времени, пространства и т. п., к которым относится содержание ресурса;
15. Авторские права (Rights) – права интеллектуальной собственности на ресурс.

Как указывается в RFC2413 (см. [RFC2413, Dublin Core Metadata for Resource Discovery](#)), эти элементы можно условно разбить на три группы (табл. 7.1).

1. Содержание (Content) – элементы, в основном относящиеся к содержанию ресурса;
2. Интеллектуальная собственность (Intellectual Property) – элементы, в основном рассматриваемые с позиции интеллектуальной собственности;
3. Экземпляр (Instantiation) – элементы, в основном относящиеся к данному экземпляру ресурса.

Таблица 7.1

Три группы элементов метаданных Дублинского ядра

| Содержание | Интеллектуальная собственность | Экземпляр |
|------------|--------------------------------|---------------|
| Название | Создатель | Дата |
| Тема | Издатель | Формат |
| Описание | Участник создания материала | Идентификатор |
| Тип | Авторские права | Язык |
| Источник | | |
| Связь | | |
| Охват | | |

В настоящее время ведется разработка версии DC 2.0. Рабочая группа Dublin Core работает в контакте с разработчиками RDF. Рассмотрим пример описания формате XML с использованием схемы метаданных Dublin Core статьи «Информационные технологии управления знаниями»:

```
<metadata>
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title> Информационные технологии управления знаниями </dc:title>
  <dc:creator> Гаврилова Татьяна Альбертовна </dc:creator>
  <dc:creator> Кудрявцев Дмитрий Вячеславович </dc:creator>
  <dc:subject xsi:type=dcterms:UDC> 004.82 </dc:subject>
  <dc:date> 2009 </dc:date>
  <dc:type> Article </dc:type>
  <dc:identifier>http://bigc.ru/publications/bigspb/km/itkm/ </dc:identifier>
</metadata>
```

Пояснение

Атрибут xmlns является ссылкой на пространство имен:

<http://www.w3.org/2001/XMLSchema-instance> – ссылка на описание языка XML,

<http://purl.org/dc/elements/1.1/> – ссылка на описание элементов Dublin Core,

<http://purl.org/dc/terms/> – дополнительный словарь Dublin Core.

Заметим, что у тэга dc:subject есть атрибут «UDC» из дополнительного словаря. Данный атрибут указывает на использование универсальной десятичной классификации для описания темы (предмета) ресурса.

Другие распространенные форматы метаданных:

LOM – модель учебного объекта – система метаданных для образовательных ресурсов, разработанная Международной ассоциацией IEEE, принятая в качестве стандарта в России для Единой информационно-образовательной среды.

GEM – версия «Дублинского ядра» метаданных, доработанная для образовательных информационных ресурсов, принятая в качестве стандарта для образовательных ресурсов США.

VCard – система метаданных для представления сведений о физических лицах и организациях, разработанная Международной ассоциацией IEEE.

MARC – машиночитаемый каталог – семейство форматов, широко распространенных для представления библиографической информации, главным образом в мировом библиотечном сообществе. В полном варианте включает до 800 элементов данных. Является стандартом де-факто для российских библиотечных сетей.

ONIX – формат описания информационных товаров в системах электронной торговли (книги, периодика, компактные диски, аудио- и видеокассеты и др.). Основан на системе стандартов ЭДИФАКТ. GILS – Глобальная (правительственная) служба поиска информации. Система метаданных

GILS представляет собой расширение форматов MARC для описания информационных ресурсов широкого профиля.

EDIFACT – одна из старейших международных систем метаданных, содержащих правила и структуру описания торгово-транспортных и других коммерческих документов. С 1998 г. поддерживается языком XML. Реализация применительно к торговле книгами и другими информационными товарами – система ONIX .

Формат Государственного регистра баз и банков данных. Российский стандарт, действующий с конца 1980-х гг. и содержащий систему метаданных для баз данных и других электронных наборов данных. Утвержден Правительством РФ, поддерживается НТЦ «Информрегистр», применяется в ряде отраслей и регионов, а также в странах СНГ для ведения баз метаданных.

Имеется еще несколько десятков менее известных систем

метаданных, в той или иной степени претендующих на широкое применение. Среди большого количества разнообразных систем метаданных, используемых в информационных ресурсах в сфере науки, культуры и образования и рассмотренных выше, наиболее распространенными и признанными являются системы метаданных семейства MARC, DC, а также быстро распространяющаяся в информационно-образовательной среде система метаданных LOM. Поэтому при создании сводных каталогов электронных ресурсов разработчикам чаще всего приходится иметь дело с интеграцией описаний ресурсов, представленных в этих системах метаданных.

7.3. Виды информационно-поисковых языков

Для улучшения поиска информационных ресурсов метаданные целесообразно выразить с помощью специализированного информационно-поискового языка (ИПЯ). Тема ИПЯ имеет многовековую историю в рамках библиотковедения и научно-технической информации. Основные положения в данной области зафиксированы в виде системы стандартов по информации, библиотечному и издательскому делу (СИБИД), которая объединяет в себе общетехнические и организационно-методические стандарты, регламентирующие правила представления данных, описания документов, функционирования библиотечных фондов, оформления печатных и электронных изданий и многое другое (<http://www.gsnti-norms.ru/norms/norms/stands2.htm>).

Большинство представленных в данном разделе понятий и определений основано на ГОСТ 7.74–96 «СИБИД. Информационно-поисковые языки. Термины и определения» (http://www.gsnti-norms.ru/norms/common/doc.asp?0&/norms/stands/7_74.htm).

Информационно-поисковый язык, ИПЯ (information retrieval language, indexing language) – искусственный язык, предназначенный для формализованного описания смыслового содержания документов, данных, отдельных понятий или терминов и обеспечения последующего их поиска в информационно-поисковых массивах. Формализация лексики и создание различных ИПЯ вызвано необходимостью устранения «избыточности» и «недостаточности» естественного языка для целей информационного

поиска, а также ликвидации присутствующих в нем синонимии и омонимии для реализации однозначности информационного поиска.

Поисковый образ документа (document description, search document image) – описание содержания документа на ИПЯ, отражающее важные признаки его содержания и вида для реализации поиска данного документа.

Индексирование: выражение содержания документа и/или смысла информационного запроса на ИПЯ. В исследованиях по семантическим технологиям данному термину соответствует *аннотирование (annotating)*.

ИПЯ делятся на *классификационные, предметизационные и дескрипторные*³³.

Классификационный ИПЯ (Classificational information retrieval language) – ИПЯ, предназначенный для индексирования документов (частей документов) и информационных запросов посредством понятий и кодов какой-либо классификационной системы.

Предметизационный ИПЯ (язык предметных рубрик, Subject headings language) – ИПЯ, предназначенный для индексирования документов (частей документов) и информационных запросов посредством предметных рубрик, представляющих собой краткую формулировку темы на естественном языке.

Дескрипторный ИПЯ (Descriptor language) – ИПЯ, предназначенный для координатного индексирования документов и информационных запросов посредством дескрипторов и/или ключевых слов.

В соответствие с видами ИПЯ различаются процессы индексирования:

Классификационное индексирование, библиографическая систематизация, классификационная систематизация или просто систематизация: Присвоение данным или документам классификационных индексов в соответствии с правилами какого-либо классификационного ИПЯ.

Классификационный индекс: Поисковый образ, построенный

³³ Сукиасян Э. Логика развития информационно-поисковых языков // Научные и технические библиотеки – 2004. – № 4.

Сукиасян Э. Каталогизация и классификация. Электронные каталоги и автоматизированные библиотечные системы: избр. ст. СПб.: Профессия, 2009. – 536 с.

средствами классификационного ИПЯ.

Предметное индексирование, предметизация: индексирование предметного содержания документов

Координатное индексирование: индексирование, предусматривающее многоаспектное выражение основного смыслового содержания документа или смыслового содержания информационного запроса множеством ключевых слов или дескрипторов.

Классификационные ИПЯ

Описание и применение классификационных ИПЯ строится на основе следующих понятий:

Классификационная система (КС) – средство формализованного представления содержания документов, данных и информационных запросов посредством кодов или описаний классов логически упорядоченного множества понятий, где: *код класса* – это обозначение класса средствами нотации (индексации) классификационной системы; *описание класса, наименование классификационного деления* – это обозначение класса на естественном языке. *Классификационная запись, классификационное деление:* совокупность элементов, обозначающая в классификационной таблице класс классификационной системы и состоящая из кода класса, описания класса и методических указаний.

Классификационный индекс – это поисковый образ, построенный средствами классификационного ИПЯ. Классификационные индексы бывают простыми, комбинированными и сложными. *Простой индекс* – классификационный индекс, содержащий один код класса без сочетания с кодами других классов. *Комбинированный индекс* – классификационный индекс, образованный из двух или более кодов класса основной таблицы классификации. *Сложный индекс* – классификационный индекс, образованный из сочетания кода класса основной таблицы классификации с кодами классов вспомогательных таблиц классификации.

По содержанию различаются КС универсальные, охватывающие все области знаний, и отраслевые для отдельных предметных областей.

По структуре КС можно разделить на перечислительные, комбинированные и фасетные. До конца XIX – начала XX вв. создавались КС перечислительного типа, деления которых располагались в простой последовательности единственного иерархического ряда. Наиболее

известны среди них «Десятичная классификация» М. Дьюи (1876), «Классификация Библиотеки конгресса США» (1904). С появлением метода типизации, в соответствии с которым вне основной классификационной таблицы создавались стандартизированные типовые деления, отражающие форму издания, географические, хронологические и др. признаки, на смену перечислительным системам пришли системы комбинационного типа, позволяющие включать в индекс несколько классификационных характеристик. Наиболее последовательной КС этого типа является Универсальная десятичная классификация (УДК). Дальнейшее развитие теории и практики КС привело к возникновению аналитико-синтетических, или фасетных КС. Теория их построения разработана индийским ученым и библиотековедом Ш.Р. Ранганатаном («Классификация двоеточием», 1933). В аналитико-синтетических КС отсутствуют готовые индексы, классификационная схема предусматривает лишь основные деления и дополняется огромным количеством отдельных таблиц, разработанных для каждой отрасли знания в соответствии с характерными для нее категориями (объектами, процессами и т. п.). Присоединяя к основным делениям индексы таблиц категорий, можно получить обозначения для самых сложных понятий. Состав категорий и их последовательность в каждой отрасли знаний регламентируются фасетной формулой.

Фасетная классификация вместо единого ряда делений в каждом основном классе имеет несколько «фасетов», соответствующих аспектам классифицируемого понятия или предмета. Все существенные термины данного класса распределяются по фасетам и образуют их «фокусы». При индексировании документов их содержание выражается цепочкой фокусов, последовательность которых определяется специальной «фасетной формулой». Примером построения индексов по системе Ш. Ранганатана могут служить фасеты и фокусы из области медицины, представленные в табл. 7.2. По этому фрагменту фасетной классификации индекс документа по диагностике инфекционных заболеваний кишечника – 23:41:3, по лечению туберкулеза легких – 41:411:4.

Преимущества этого вида классификаций в том, что они облегчают многоаспектное индексирование документов, позволяя собирать в одном месте все аспекты рассмотрения какого-либо предмета или темы, они легче

поддаются изменениям при введении новых понятий, допускают большую глубину индексирования при более коротких индексах.

Таблица 7.2

Фрагмент фасетной классификации по медицине

| ФАСЕТЫ | | |
|---|--|--|
| Органы тела | Проблемы медицины | Уход и лечение |
| 1. Органы в целом 2. Органы пищеварения 21. Пищевод 22. Желудок 23. Кишечник 3. Кровеносная система 4. Органы дыхания 41. Легкие | 1. Общие проблемы 2. Морфология 3. Физиология 4. Болезни 41. Инфекционные 411. Туберкулез | 1. Питание 2. Этиология 3. Диагностика 4. Лечение |

Их применение особенно эффективно при поиске в небольших по объему узкоспециализированных собраниях документов. Видный английский информатик Д. Фоскетт так обосновал достоинства фасетных классификаций: «От схемы не требуется более, чтобы она указывала *место* для каждого документа, включая любой термин или набор терминов в явном виде в классификационные таблицы по каждой предметной области. Эти схемы могут задать набор правил или рабочих процедур, с помощью которых такие контексты можно, по мере надобности, формулировать на основе тех же самых схем».

Предметизационные ИПЯ

В библиотеках распространены *предметные* каталоги, сущность которых заключается в том, что содержание документа кратко формулируется при помощи одного или нескольких типовых ключевых слов, получивших название *предметных рубрик*, а затем предметные рубрики располагаются по алфавиту, и под каждой из них как под заголовком³⁴ собираются библиографические описания документов.

³⁴ Предметные рубрики также называются предметными заголовками (*subject headings*), и этот вариант термина лучше отражает природу термина, выбранного для обозначения содержания ряда документов в каталоге. Собственно предметными рубриками следует называть совокупности документов, объединенные одним предметным заголовком.

Составление предметных рубрик и распределение по ним документов называется *предметизацией*. Задача предметизации обычно состоит в том, чтобы указать главный объект рассмотрения в документе и, может быть, основные его аспекты и основные отношения к другим предметам.

В отличие от КС, предметизация распределяет документы по предметам или понятиям, не соотнося их с какими-либо областями знания. Это различие делает классификационный и предметизационный принципы организации документов независимыми, дополняющими друг друга, предназначенными для поиска документов по разным типам запросов. Предметизация дает возможность собирать в одном месте документы по таким комплексам, как конкретный материал, свойство, изделие, явление природы или общества, род деятельности, географическое понятие и т. п., собирая под каждым предметным заголовком весь массив знаний, безотносительно к тому, какой области науки эти знания принадлежат.

Другим отличием предметизации от классификации является то, что заранее составленный список предметных заголовков не ограничивает подробности анализа содержания документа. Если документ посвящен вопросу, не отраженному в списке предметных заголовков, всегда имеется возможность сформулировать новый заголовок самостоятельно. Обычно же при выборе предметной рубрики для документа руководствуются заранее составленным списком предметных заголовков. Но не представляет труда внести в него вновь образованную предметную рубрику, которая займет свое надлежащее место, определяемое алфавитным порядком. Этого обычно не удается делать в языках классификационного типа, где введение новых классов зачастую влечет преобразование большей части классификационных связей.

У каждого предметного заголовка могут в принципе быть подзаголовки, делящие документы в рубрике на подрубрики. В некоторых случаях к предметному заголовку могут быть даны ссылки на другие рубрики, где могут находиться документы по сходному предмету. Таким образом, в списке предметных заголовков одна запись может иметь довольно сложный характер. Вот пример фрагмента словаря предметных рубрик:

...

самовар

самолеты

– *военные*

– *гражданские*

– – *грузовые*

– – *пассажирские*

см. также авиалайнеры

саморезы

см. винты-саморезы

самоходные баржи

см. теплоходы грузовые речные

Традиционная каталожная техника не позволяет раскрывать содержание документа предметными рубриками с достаточной полнотой. Предметные заголовки отражают только основной предмет документа, и даже документы многопланового, обзорного характера могут быть отражены в каталоге лишь в ограниченном числе предметных рубрик. Недостаток места, трудоемкость составления картотечных каталогов заставляли библиографов разрабатывать довольно сложные правила оптимального выбора используемых предметных заголовков. Эти правила вместе со словарями рекомендуемых предметных заголовков составляют ИПЯ предметных рубрик, предметные ИПЯ.

Языки предметных рубрик используются не только как основа для предметного каталога документов, но также и как вспомогательное средство для пользователя систематическим каталогом, основанным на тематической классификации. В тех случаях, когда пользователь знает предмет своего интереса, но не знает, к какой отрасли знания он относится, он может обратиться к алфавитно-предметному указателю, в котором для каждой предметной рубрики указывают подходящий раздел (разделы) тематической классификации, в котором собраны знания по данному предмету. В этом случае мы имеем дело с языком предметных рубрик в функции поиска не документов, а тематических разделов систематического каталога, или, что то же самое, для поиска классов тематической классификации.

ИПЯ координатного индексирования

Современная компьютерная техника снимает ограничения по объему каталогов и снижает трудоемкость их составления. Поэтому получила

распространение идея приписывать документам все термины, существенные для выражения содержания документа, и в электронном каталоге иметь инверсный файл записи адресов документов, использовавших каждый такой термин, получивший наименование «ключевое слово».

Под ключевыми словами в данном случае понимаются наиболее существенные для выражения содержания документа однозначные слова и словосочетания, обладающие назывной (номинативной) функцией. Поиск документа при этом должен происходить, как правило, не по одному ключевому слову (не по одной предметной рубрике, как в случае языка предметных рубрик), а по формулировке поисковой потребности, содержащей ряд ключевых слов, полно описывающих тему поиска. В процессе поиска по записям инверсного файла, соответствующим ключевым словам запроса, должны производиться логические операции над множествами адресов документов в соответствии с указаниями поискового предписания.

Ключевые слова образуют новый способ описания и поиска документов – язык ключевых слов (ИПЯ ключевых слов). Если в поисковой системе ключевые слова всей предметной области поиска сведены в словарь, где они связаны смысловыми отношениями так, что выражают структуру данной области знания, онтологию рассматриваемых явлений, то такой словарь называют информационно-поисковым тезаурусом (см., например, общественно-политический тезаурус русского языка используемый в УИС Россия <http://uisrussia.msu.ru/docs/ips/n/techno/index.htm>), а термины, включенные в него, называют *дескрипторами*.

Дескрипторные информационно-поисковые системы открыли принципиально новую возможность поиска документов и содержащейся в них информации по любому сочетанию заранее не предвиденных признаков. В настоящее время производятся попытки сочетания уже имеющихся поисковых средств со вновь создаваемыми. Отсюда вытекает и желание найти общие черты в этих разных системах и лежащих в их основе ИПЯ: языке предметных рубрик и дескрипторном языке, что обычно сочетается с поисками путей их совместимости.

Всегда можно найти такую удаленную позицию, такое основание

деления, при которых эти языки попадут в один общий класс. В ряду искусственных языков они принадлежат к классу информационных, в ряду информационных – к подклассу информационно-поисковых. В них используются в качестве индексов слова естественного языка. При построении этих языков применяются внешне схожие приемы: перечень предметных заголовков и словарная часть тезауруса упорядочиваются в алфавите слов. Тем не менее, учитывая эти общие и сходные черты, нельзя забывать и о принципиальных различиях данных языков. Основной словарный состав языка предметных рубрик – это имена сложных классов, построенных до индексирования документов, т. е. этот язык принадлежит к типу *предкоординируемых*. Deskрипторный же язык является *посткоординируемым*, т. е. строится из имен простых классов, которые образуют необходимые понятия при их пересечении (логическом умножении) в момент индексирования и/или поиска документов, т. е. после создания координатной сетки описания документов.

Для того, например, чтобы индексировать статью о производстве и экспорте вычислительных и пишущих машин в США, Японии и Великобритании достаточно дескрипторов *производство, экспорт, компьютер, пишущая машина* и названий трех этих стран. Тогда при поиске мы сможем сразу (за один шаг) найти этот документ при любом порядке перечисления дескрипторов (а таких комбинаций может быть $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 = 5040$, т. е. число перестановок из семи признаков). Если же пользоваться языком предметных заголовков, то в его словаре должны быть предусмотрены все осмысленные комбинации этих терминов, чтобы был возможен тот же результат. Число осмысленных комбинаций, конечно, меньше указанного выше, но оно все равно велико, и в реальные словари предметных рубрик включается только малая часть их. Это, в свою очередь, приводит к необходимости дополнительного поиска именно той комбинации, которая включена в словарь, а чаще влечет за собой значительные поисковые потери. Так в реальный словарь в соответствии сданной тематикой реально могут быть включены предметные заголовки типа:

компьютеры – производство – Великобритания

компьютеры – производство – США

компьютеры – производство – Япония

пишущие машины – экспорт – Великобритания

пишущие машины – экспорт – США

пишущие машины – экспорт – Япония.

Однако трудно предположить, что в нем будет представлен хотя бы один предметный заголовок со всеми шестью терминами. Предметные заголовки, наподобие желаемого нами «компьютер – пишущая машина – производство – экспорт – Великобритания – США – Япония», не только практически не реальны, но и теоретически недопустимы в языках предметных рубрик.

Координатное индексирование в том и состоит, что для характеристики содержания документа или запроса перечисляются такие ключевые слова или дескрипторы, пересечение (логическое умножение) которых выражает основное смысловое содержание (главную тему, предмет) этого документа или запроса, тогда как в предметизации для данной цели используются заранее сформулированные по определенным правилам заголовки и подзаголовки.

При индексировании, т. е. выражении основного смыслового содержания документа в терминах ИПЯ, процессы информационного анализа и синтеза совершаются в два этапа. Первый этап является общим для всех языков. Содержание документа анализируется как с позиций того, какие идеи и факты заложены в него автором, так и с позиций научных и практических интересов большинства его потенциальных читателей. (Если не иметь в виду узкоспециальные интересы, то обе точки зрения чаще всего совпадают.) Результаты этого анализа синтезируются в виде субъективного представления индексатора об основном содержании документа. Второй этап зависит от языка индексирования. Если это *предкоординированный* алфавитно-предметный язык, то свое представление о содержании документа индексатор сверяет с потенциальными запросами читателей, отраженными в перечне предметных заголовков. Для *посткоординируемого* дескрипторного языка аналогичному анализу подвергается тезаурус (не связанный непосредственно с потенциальными запросами потребителей) и сам текст индекслируемого документа. Синтез в данном случае выражается в выборе соответствующих предметных заголовков или дескрипторов (ключевых слов).

Другими словами, при всей внешней схожести процедур индексирования посредством этих разных типов ИПЯ, характер их использования различен. В одном случае мы пользуемся готовыми продуктами в виде заголовков и подзаголовков, обозначающих класс документов определенного содержания. В другом случае – это лишь исходный материал, дескрипторы и ключевые слова, при перемножении которых образуется класс, соответствующий данному содержанию. Вот почему перечень предметных заголовков и словарная часть тезауруса, при всей их внешней схожести, при том, что определенная часть слов в них может совпадать, на самом деле являются совершенно отличными друг от друга списками, слова для которых отбираются на основе разных критериев и играют различную роль.

Информационно-поисковые тезаурусы

Теория и практика информационного поиска свидетельствуют, что наиболее перспективными и универсальными являются ИПЯ, основанные в той или иной мере на тезаурусе понятий, который отражает структуру онтологии предметной области поиска.

Информационно-поисковый тезаурус (ИПТ) – это контролируемый словарь терминов на естественном языке, явно указывающий отношения между терминами и предназначенный для информационного поиска. ИПТ относится к ИПЯ координатного типа.

Основными целями разработки традиционных ИПТ являются следующие: обеспечение перевода естественного языка документов и пользователей на контролируемый словарь, применяемый для индексирования и поиска; обеспечение последовательного использования единиц индексирования; описание отношений между терминами; использование как поискового средства при поиске документов.

Основной единицей ИПТ являются термины, которые разделяются на дескрипторы (авторизованные термины) и недескрипторы (аскрипторы). Большинство версий стандартов по тезаурусам указывают на связь терминов с понятиями предметной области. По американскому стандарту термин – это слово либо словосочетание, обозначающее понятие. Стандарт ISO подчеркивает, что индексирующий термин – это представление понятия предпочтительно в форме существительного или именной группы. При этом понятие рассматривается как единица мысли, которая

формируется мысленно для отражения всех или некоторых свойств конкретного или абстрактного, реально существующего или мысленного объекта. Понятия существуют как абстрактные сущности, независимо от терминов, которые их выражают. Российский ГОСТ рассматривает понятие как форму мышления, отражающую существенные свойства, связи и отношения предметов и явлений, а термином в определении ГОСТа является слово или словосочетание, являющееся точным обозначением определенного понятия какой-либо области знания.

ГОСТ 7.74-96 определяет единицы тезауруса как лексические единицы информационно-поискового языка – то есть обозначения отдельного понятия, принятые в информационно-поисковом языке и неделимые в этой функции. Таким образом, разработчики тезаурусов предполагают, что понятие предметной области обычно имеет несколько возможных вариантов лексического представления в тексте, которые рассматриваются как синонимы. Среди таких синонимов выбирается дескриптор – термин, который рассматривается как основной способ ссылки на понятие в рамках тезауруса. Другие термины из синонимического ряда, включенные в тезаурус, называются аскрипторами или недескрипторами. Они используются как вспомогательные элементы, текстовые входы, помогающие найти подходящие дескрипторы.

Отношениями в информационно-поисковом тезаурусе в соответствии с ГОСТ 7.25 являются следующие: род – вид; часть – целое; причина – следствие; сырье – продукт; административная иерархия; процесс – объект; функциональное сходство; процесс – субъект; свойство – носитель свойства; антонимия. Такие содержательные типы связей между дескрипторами чаще всего не отражаются в подробном перечне отношений тезауруса, а записываются с помощью небольшого набора отношений, которые обычно разделяются на два типа: иерархические и ассоциативные.

Иерархические отношения по ГОСТу 7.25-2001 обладают свойствами транзитивности и антисимметричности, которые могут быть использованы при избыточном индексировании в интересах повышения эффективности информационного поиска. Предпочтительно указывать связи между дескрипторами как отношения иерархического вида, если они обладают этими свойствами. Применяемые в ИПТ иерархические

отношения могут дифференцироваться на отдельные виды. Основным иерархическим отношением, используемым в ИПТ, является родо-видовое отношение (оно же – отношение НИЖЕ-ВЫШЕ). Родо-видовая связь устанавливается между двумя дескрипторами, если объем понятия нижестоящего дескриптора входит в объем понятия вышестоящего дескриптора. Также в качестве иерархического отношения в ИПТ может устанавливаться отношение ЧАСТЬ-ЦЕЛОЕ. Многие руководства и стандарты подчеркивают, что иерархические отношения в ИПТ должны устанавливаться в тех случаях, когда отношения истинны независимо от контекста – только в таких случаях дескрипторы ИПТ могут быть организованы в иерархии. Эта рекомендация связана с тем, что обычно в информационном поиске очень трудно четко определить контекст употребления термина и понять, применимо ли в данном контексте то или иное отношение. Так, для мышей можно указать, что они грызуны, поскольку это внутренняя характеристика мышей. В то же время неправильно указывать, что мыши – вредители, поскольку имеются лабораторные мыши и домашние мыши, которые вредителями не являются. Рекомендуется использовать тест «все-некоторые». Например, «все мыши являются грызунами, но некоторые мыши являются вредителями».

Отношение ассоциации предназначено для указания на дополнительные дескрипторы, полезные при индексировании или поиске. Отношение ассоциации является неиерархическим и ассоциативным. Ассоциативное отношение наиболее трудно определить. Российский стандарт на создание ИПТ указывает, что «ассоциативное отношение является объединением отношений, не входящих в иерархические отношения или в отношения синонимии. Допускается включать в ассоциативное отношение все виды отношений, кроме синонимии и отношения РОД-ВИД». Другие источники стараются изложить более подробные принципы установления ассоциативных отношений.

Разные типы ИПЯ имеют свои достоинства и ограничения, которые делают их особо пригодными для решения разных поисковых задач. Возможности дескрипторного языка эффективно реализуются при узко тематическом поиске по произвольной комбинации признаков. Широкий тематический поиск по традиционным отраслям знаний и поиск по

конкретным предметам, дисциплинам и их разделам в фондах документов за многие годы и в условиях одновременного обращения к ним большого числа читателей по-прежнему хорошо обеспечиваются библиотечными каталогами, основанными на иерархических и алфавитно-предметных классификациях. Выпуск информационных изданий требует разработки специальных рубрикаторов с небольшим числом уровней иерархии и подвижной, быстро меняющейся рубрикацией.

Рассмотренные выше ИПЯ и связанные с ними принципы индексирования систематизированы в табл. 7.3. Классификационная система и тезаурус могут рассматриваться как разновидности онтологии разного уровня формальности³⁵. Однако, на взгляд автора, при наличии и достаточности альтернативных терминов не следует вводить понятие «онтология», а минимальным требованием для его использования считать наличие родовидовых отношений.

Таблица 7.3

Обзор принципов индексирования

| Принцип индексирования | Классификационный | Предметизационный | Координатного индексирования |
|-------------------------------|---|--|---|
| ИПЯ | Классификационный | Предметизационный | Дескрипторный |
| Представление ИПЯ | Классификационная система | Словарь (список) предметных рубрик | Тезаурус |
| Термины индексирования | Классификационные индексы | Предметные рубрики | Дескрипторы |
| Характер поиска | Систематический (по иерархии понятий классификационной системы) | Алфавитный (по алфавиту предметных рубрик) | Многоаспектный (по основной, алфавитной, части тезауруса и по указателям) |
| Название процесса | Систематизация | Предметизация | Координатное индексирование |
| Название специалиста | Систематизатор | Предметизатор | Индексатор |

³⁵ См. классификацию онтологий <http://www.intuit.ru/department/expert/ontoth/2/>

7.4. Разработка информационно-поискового языка на предприятии

Преобразование информационного хранилища в корпоративную память происходит за счет внедрения информационно-поискового механизма, основанного на ИПЯ и описании с помощью него ресурсов организации, а также полезных внешних ресурсов. В бизнес-языке, особенно в англоязычных текстах, ИПЯ называется таксономией. В отличие от исходного значения таксономии, как родовой иерархии классов, таксономия на языке поставщиков информационных систем (IBM, Convera, Microsoft) и аналитиков (Gartner, Forrester, Delphi Group) – это структура, отражающая способ классификации документов по группам для облегчения поиска, обучения и размещения.

Популярность использования данного слова в таком значении подтверждается множеством книг с соответствующими названиями:

- *Stewart D. Building Enterprise Taxonomies. Mokita Press, 2008.*
- *Lambe P. Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness (Chandos Knowledge Management). Neal-Schuman Publishers. 2007.*
- *Hedden H. The Accidental Taxonomist. Information Today, Inc. 2010.*

В традиционных информационно-поисковых системах предприятий в качестве исходного ИПЯ/таксономии используется классификационная система (КС), традиция использования которой является, скорее всего, следствием распространенной иерархической файловой системы.

Далее предлагается методика создания КС для предприятия, основанная на фасетной классификации (см. раздел «Виды информационно-поисковых языков»). Создание КС является одним из этапов создания корпоративной памяти:

- Определение области и масштаба (scope) корпоративной памяти,
- Создание КС,
- Наполнение корпоративной памяти и индексирование контента,
- Тестирование корпоративной памяти.

В настоящее время существует два подхода к созданию КС: ручной и автоматический. Необходимо отметить, что, несмотря на указываемые поставщиками ПО возможности автоматического создания КС, создание

эффективной КС требует интеграции автоматических методов с ручными, что подтверждается исследованиями компаний Delphi Group и Gartner. Далее предлагается процесс / алгоритм ручной разработки КС, см. рис. 7.8.

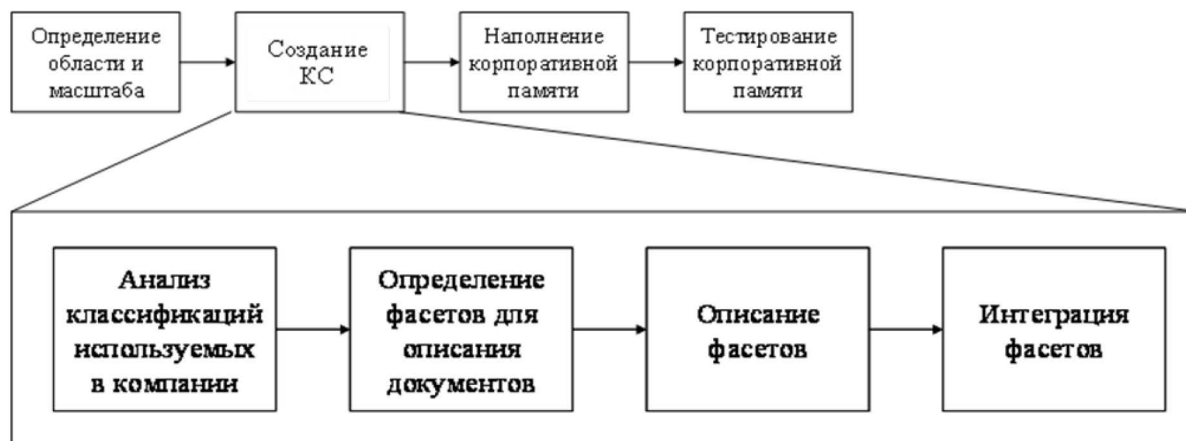


Рис. 7.8. Процесс создания классификационной системы и его контекст

Шаг 1: Анализ классификаций, используемых в компании

Начать создание таксономии полезно с анализа классификаций, которые можно найти в организации. Во-первых, можно обратиться к ведущим специалистам и попросить их показать способ *структурирования информации на своих персональных компьютерах*. Во-вторых, можно обратиться к имеющимся в наличии общекорпоративным классификациям – структура сайта, структура интранет портала (который, как ни странно, может создаваться раньше базы знаний), дерево папок на корпоративном сервере, КС в отделе научно-технической информации (НТИ) компании.

Необходимость этого шага объясняется, с одной стороны, экономией усилий (использование уже созданных наработок), а с другой стороны, необходимостью «не выплеснуть ребенка», олицетворяющего накопленные за долгие годы жизни корпоративную культуру и способ мышления сотрудников компании, которые могут являться одним из конкурентных преимуществ компании. (Хотя, иногда корпоративная культура и способ мышления могут, напротив, препятствовать развитию компании, и в таком случае создание новой таксономии позволяет проводить изменения неконкурентоспособной культуры и способа мышления, определяя новый взгляд на старые вещи.)

Шаг 2: Определение фасетов для описания документов

Фасет – классификационный признак, используемый при создании КС. На данном шаге необходимо определить признаки, по которым будут классифицироваться материалы в базе знаний. Возможные признаки классификации:

- По продуктам (услугам) компании,
- По этапам жизненного цикла продукции,
- По функционалу компании (областям деятельности, бизнес-процессам),
- По этапам проектов (для проектных компаний),
- По оргструктуре (область создания и использования материалов),
- По рыночным сегментам,
- По типу документа,

а также другие элементы архитектуры предприятия и их признаки классификации³⁶.

Кроме «нисходящего» подхода, формирование фасетов может происходить и «снизу-вверх» путем анализа документов и анализа потребностей будущих пользователей.

Необходимо учитывать, что не нужно пытаться сразу выявить все фасеты – достаточно вначале определиться с 2-3 (а иногда хотя бы с 1). Выбранные фасеты позволят перейти к следующему шагу создания КС, в процессе выполнения которого могут непроизвольно «всплывать» новые фасеты (например, при возникновении таких описаний документов, которые важны, но не соответствуют ни одному из выбранных фасетов).

Шаг 3: Описание фасетов

После того как определены «стартовые» фасеты нужно переходить к их структуризации – созданию таксономии внутри каждого фасета. Однако прежде чем заниматься самостоятельным созданием таксономии внутри каждого фасета целесообразно воспользоваться имеющимся в мире опытом – посмотреть на уже созданные внешние (по отношению к компании) таксономии (см. табл. 7.4, 7.5).

³⁶ Григорьев Л. (редактор), Горелик С., Кудрявцев Д., Корышев И., Кислова В., Макаревич М., Каменская М., Якубовская Т., Заблоцкий А., Панфилов М. Менеджмент по нотам. Технология построения эффективных компаний. – М: Альпина Паблшерз, 2010.

Таблица 7.4

Использование внешних классификаций

| Фасет | Источник таксономии | Описание и фрагмент таксономии | | | |
|---|--|---|--------|----------|----------|
| Процессы компании | The American Productivity & Quality Center (APQC) http://www.apqc.org/ | Process Classification Framework – Универсальная таксономия /классификация бизнес-процессов, На верхнем уровне процессы объединены в следующие группы: 1. Изучение рынков и потребителей; 2. Разработка видения и стратегии; 3. Разработка продуктов и услуг; 4. Маркетинг и продажи; 5. Производство и поставка продуктов и услуг; 6-13. ... | | | |
| Жизненный цикл продукции | Стандарты качества ISO | Plan | Do | Check | Act |
| | | Планируй | Пробуй | Проверяй | Действуй |
| Типовые этапы проектов (для проектных компаний) | Международный стандарт управления проектами Project Management Body of Knowledge (PMBoK) | Примеры декомпозиции работ по проекту (Work Breakdown Structure): Типовые фазы проекта:  <pre> graph TD A[Разработка программы] --> B[Анализ потребностей] A --> C[Дизайн] A --> D[Программирование] A --> E[Тестирование] </pre> | | | |

Таблица 7.5

Использование существующих внутренних классификаций

| Фасет | Источник информации |
|---------------------------|--|
| Продукты и услуги | Прайс-лист |
| Организационная структура | Организационно-распорядительная документация |
| Функционал компании | |
| Сегменты рынка | Рыночная стратегия |

При использовании представленных в табл. 7.5 внутренних классификаций требуется тесная интеграция процессов создания таксономии с организационным управлением в части состава работ, методик и конечных результатов. Если, например, в качестве фасета выбрана организационная структура компании, то процесс разработки таксономии будет согласован с процессом разработки организационной.

Для выполнения таких работ могут использоваться соответствующие методики из области организационного управления (например, технология процессного описания компании), а также применяться специальные программные продукты, поддерживающие организационное управление, – системы бизнес-моделирования. Компоненты моделей, созданных в системах бизнес-моделирования, могут стать исходным вариантом КС.

Особенности использования готовых классификаций:

- Каждая компания должна адаптировать типовые таксономии под свои потребности, корректируя, расширяя, детализируя.
- Дается возможность «быстрого старта».
- Потенциальная экономия ресурсов.

При создании иерархической КС можно использовать несколько подходов:

- Процесс **нисходящей** разработки («сверху-вниз»),
- Процесс **восходящей** разработки («снизу-вверх»).

Процесс комбинированной разработки – это сочетание нисходящего и восходящего подходов: Сначала определяются более заметные понятия, а затем они соответствующим образом обобщаются и детализируются.

В рамках данной статьи не будем подробно рассматривать методику создания и требования к классификации внутри фасета.

Шаг 4: Интеграция фасетов

В результате выполнения предшествующих шагов по созданию таксономии мы получили ряд фасетов, которые имеют свою внутреннюю классификацию. Если информационная система предприятия поддерживает многоаспектную (фасетную) классификацию и поиск, то создание КС может завершиться уже на этом этапе. Ниже приведен пример использования фасетной классификации в консалтинговой компании. Для получения какой-либо информации пользователь будет в разных фасетах указывать различные элементы, например, для получения *Информации X*

необходимо в *Фасете 1* выбрать «Проекты», а в *Фасете 2* выбрать тип документа «Планы работ» (рис. 7.9).

| | | Фасет 1: Продукты и услуги | | |
|--------------------------|---------------------------|----------------------------|----------|----------------------|
| | | Проекты | Семинары | Программные продукты |
| Фасет 2: Типы документов | Предложения | | | |
| | Договорные документы | | | |
| | Планы работ | X | | |
| | Отчеты | | | |
| | Переписка | | | |
| | Отзывы | | | |
| | Семинары | | | |
| | Раздаточные | | | |
| | Версии ПП | | | |
| | Статьи и описания | | | |
| | Презентации и демо-ролики | | | |
| | Документация на ПП | | | |

Рис. 7.9. Поиск информации в фасетной классификации

Такое описание размещения информации в фасетной классификации: *{Проект X: Планы работ}* называется фасетной формулой (по синтаксису фасетной формулы видно, почему используется термин «Классификация с двоеточием»). Подход к поиску информации в данном случае соответствует использованию фильтров в Excel. Использование созданной фасетной классификации позволяет точно и быстро находить любую необходимую информацию, однако реализация такой таксономии имеет свои недостатки:

- Необходима высокая осведомленность пользователей о предметной области: для того, чтобы сформировать фасетную формулу необходимо заранее знать наполненные («непустые») ячейки, которые на рис. 7.9 выделены серым. Для этого нужно заранее четко

представлять результат поиска, чего часто нет при поиске по таксономии.

- Для реализации фасетной классификации требуется специальное программное обеспечение, которое обеспечит как (1) механизм поиска, так и (2) механизм настройки прав доступа

Возможным способом устранения таких недостатков может быть создание иерархической таксономии путем интеграции разработанных фасетов.

В представленной на рис. 7.10 таксономии явно видны фасеты: 1. *Продукты* и 2. *Тип контента*, которые интегрированы в рамках одной таксономии.

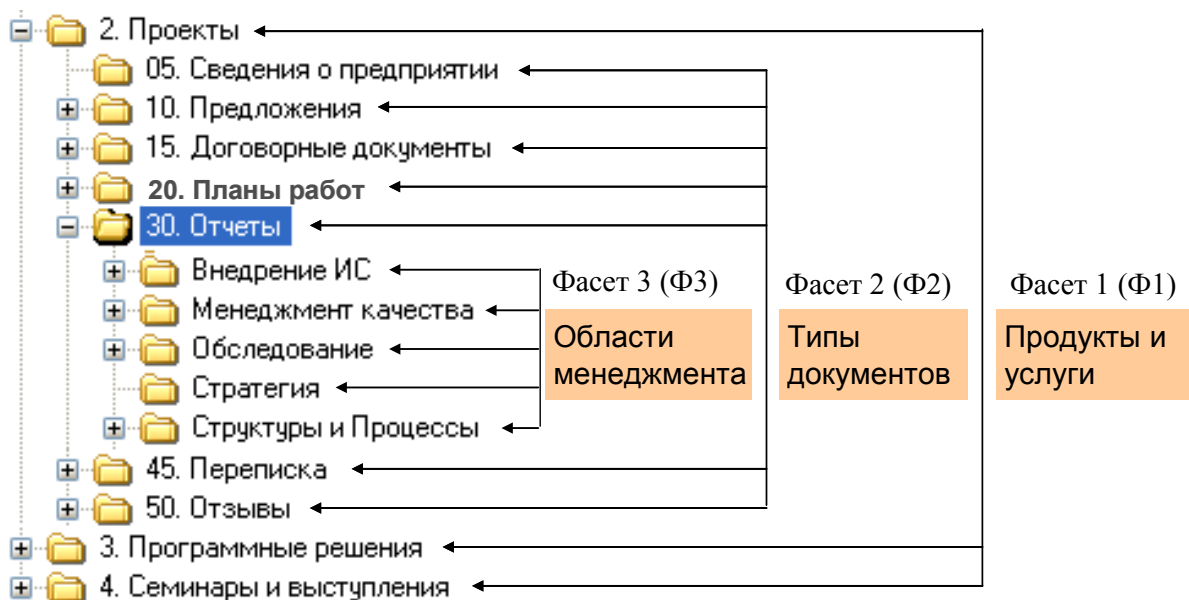


Рис. 7.10. КС корпоративной памяти консалтинговой компании

В представленной на рис. 7.10 КС интегрированы следующие фасеты: 1. *Продукты и услуги (консалтинговой компании)* 2. *Типы документов* 3. *Области менеджмента*.

7.5. Использование онтологий и языков семантического веба

Использование онтологий в процессе накопления знаний может выполнять роль координатного ИПЯ (см. информационно-поисковый

тезаурус), задавать структуру базы знаний и выступать общей схемой для интеграции информации. Для представления онтологий и индексирования информационных ресурсов в СУЗ используются языки семантического веба. В 1997 году консорциум W3C определил спецификацию RDF (Resource Description Framework). RDF предоставляет простой, но мощный язык описания ресурсов, основанный на триплетях (triple-based) «Субъект-Предикат-Объект» и спецификации URI. В 1999 году RDF получает статус рекомендации. Этот шаг в направлении улучшения функциональности и обеспечения интероперабельности (т. е. возможности обмениваться данными, несмотря на их разнородность) в Сети считается одним из важнейших. Концептуально RDF дает минимальный уровень для представления знаний в Сети. Спецификация RDF опирается на ранние стандарты, лежащие в основе Web:

- Unicode служит для представления символов алфавитов различных языков,
- URI используется для определения уникальных идентификаторов ресурсов,
- XML и XML Schema – для структурирования и обмена информацией и для хранения RDF (XML синтаксис RDF).

Кроме RDF был разработан язык описания структурированных словарей для RDF – RDF Schema (RDFS). Он предоставляет минимальный набор средств для спецификации онтологий. RDFS получил статус рекомендации W3C в 2004 году. Однако препятствием для Semantic Web стало то, что документов, написанных на языке RDF/RDFS, было относительно мало. В период с 2001 по 2004 годы шла интенсивная работа по созданию программных средств для обработки и автоматической генерации RDF-документов.

Результатом в 2004 году стал язык GRDDL (Gleaning Resource Descriptions form Dialects of Languages). Его назначение состоит в предоставлении средств для извлечения RDF-триплетов из XML и XHTML данных (в особенности это относится к документам, автоматически генерируемым из закрытых баз данных). Развивалось и программное обеспечение для Semantic Web. В области создания библиотек классов и построения логических выводов над RDF-графами была создана библиотека Jena Framework, в области создания модулей расширения для

браузеров – Simile для Firefox. В области создания визуальных сред редактирования большое число редакторов онтологий стали поддерживать RDF.

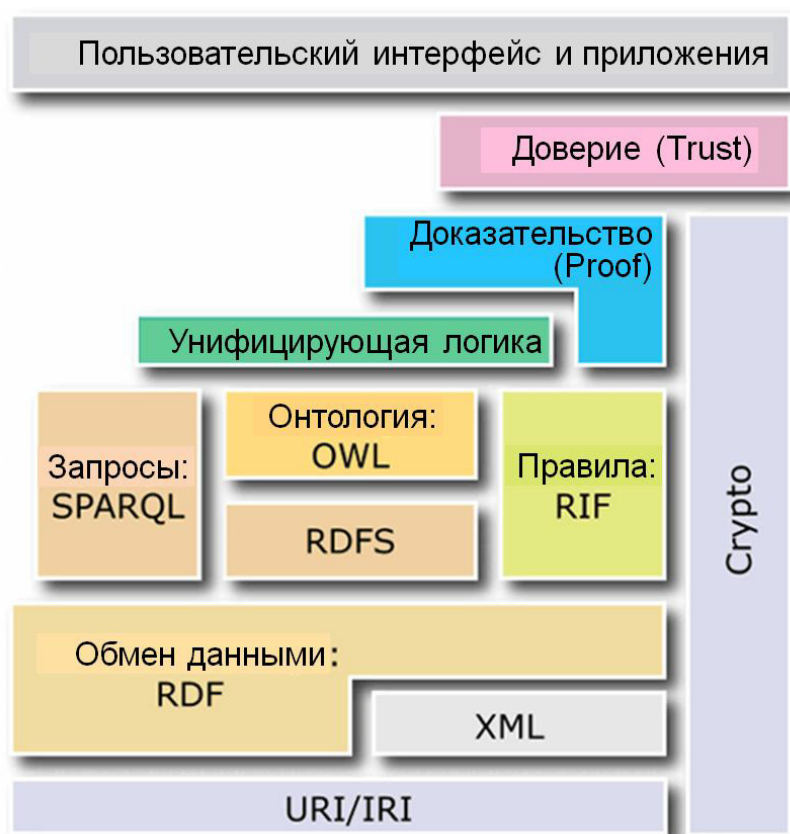
В 2004 году статус рекомендации получил язык OWL (Web Ontology Language). Он имеет три диалекта (три множества структурных единиц), используемых в зависимости от требуемой выразительной мощности. OWL фактически является надстройкой над RDF/RDFS и поддерживает эффективное представление онтологий в терминах классов и свойств, обеспечение простых логических проверок целостности онтологии и связывание онтологий друг с другом (импорт внешних определений). Многие формализмы описания знаний могут быть отображены на формализм OWL (два из его диалектов – OWL Lite и OWL DL – соответствуют двум дескриптивным логикам, имеющим разную выразительную силу). Большое число создаваемых в настоящее время онтологий кодируются на OWL; уже существующие онтологии транслируются в него. На этом работа по обеспечению Semantic Web необходимыми стандартами не остановилась. В 2005 году началась работа над форматом обмена правилами – RIF (Rule Interchange Format). Его назначение – соединить в одном стандарте несколько формализмов для описания правил, по которым может осуществляться нетривиальный логический вывод: логику клауз Хорна, логики высших порядков, продукционные модели и т.п.

Язык SPARQL – язык запросов к RDF-хранилищам – в январе 2008 года приобрел статус официальной рекомендации Консорциума W3C. Синтаксически он очень похож на SQL. Он уже широко используется разработчиками информационных систем.

На рис. 7.11 представлена диаграмма, называемая иногда стеком (или даже «слоеным пирогом») Семантического веба (Semantic Web), которая интегрирует отдельные языки.

Все основные уровни диаграммы были описаны выше. Уровням «Ontology vocabulary» и «Logic» соответствуют OWL и RIF. Уровни Доказательств (Proof) и Доверия (Trust) на данный момент еще не затронуты никакими стандартами. Здесь и возникает одно из существенных препятствий к реализации всей идеи: поддержка автоматической проверки корректности и правдивости информации. В

самом деле, у многих поставщиков семантических описаний может возникнуть соблазн «обмануть» программу-агента, предоставив информацию, не соответствующую действительности, либо навязчивую рекламу, как это в настоящее время прделывается с поисковыми машинами, спам-фильтрами и т. п.



Источник: <http://www.semanticfocus.com/>

Рис. 7.11. «Слоеный пирог» семантического веба

Рассмотрим более подробно отдельные стандарты.

RDF (Resource Description Framework)

RDF – язык представления информации о ресурсах WWW. В частности, RDF служит для представления метаданных, связанных с ресурсами Сети, таких как «заголовок», «автор», «дата последнего изменения страницы». Но RDF может использоваться и для представления информации о ресурсах «второго типа», на которые можно только ссылаться (или идентифицировать в Сети при помощи URI), но

невозможно непосредственно получить к ним доступ через Сеть.

Может оказаться, что в некоторых случаях для управления метаданными достаточно использовать XML и XML Schema (либо вообще ограничиться подэлементом HEAD элемента HTML). Но этот подход слабо масштабируется: при увеличении объема метаданных, усложнении их структуры управление метаданными, построенными на основе XML Schema, становится трудоемкой задачей, для решения которой и предназначен RDF.

Модель данных RDF. RDF-граф

Базовой структурной единицей RDF является коллекция троек (или триплетов), каждая из которых состоит из субъекта, предиката и объекта (S,P,O). Набор триплетов называется RDF-графом. В качестве вершин графа выступают субъекты и объекты, в качестве дуг – предикаты (или свойства). Направление дуги, соответствующей предикату в данной тройке (S,P,O), всегда выбирается так, чтобы дуга вела от субъекта к объекту (рис. 7.12).



Рис. 7.12. RDF-тройка

Каждая тройка представляет некоторое высказывание, увязывающее S, P и O.

Первые два элемента RDF-тройки (субъект и предикат) идентифицируются при помощи URI. Объектом же может быть как ресурс, идентифицируемый при помощи URI, так и RDF-литерал (значение).

Пример представления информации с помощью RDF:

- 1) Тройки {субъект, предикат, объект}
{KR, имя, «Онтологии и представление знаний»}
{KR, читается, БК}
{KR, homepage, <http://logic.pdmi.ras.ru/csclub/courses/ontology>}
- 2) RDF-граф (рис. 7.13)

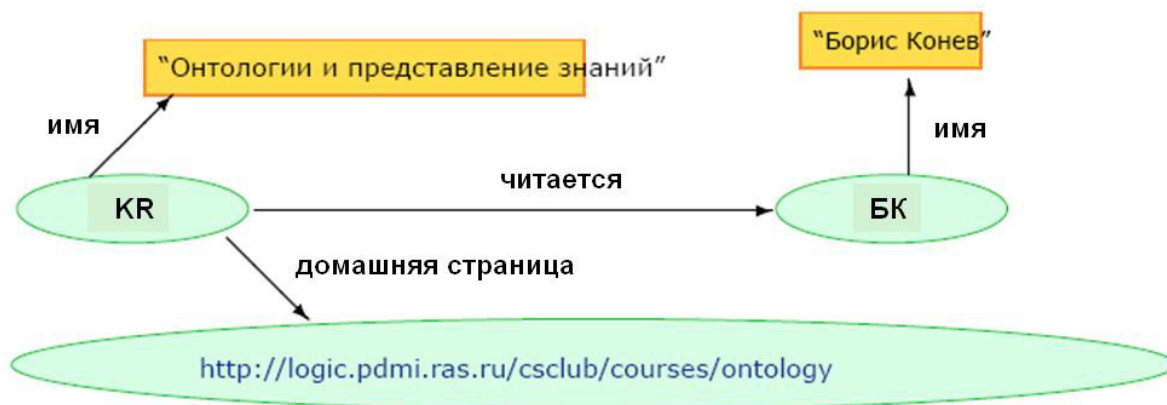


Рис. 7.13. Пример RDF-графа

3) В XML-синтаксисе:

Пояснение: Каждый тег description описывает ресурс и каждый вложенный элемент или атрибут является свойством ресурса

```

<rdf:Description rdf:about="#KR">
  <csclub: читается rdf:resource="#БК"/>
  <csclub:имя>Онтологии и представление знаний</csclub:name>
  <csclub: домашняя страница
rdf:resource="http://logic.pdmi.ras.ru/csclub/courses/ontology"/>
</rdf:Description>
<rdf:Description rdf:about="#БК">
  <csclub:имя>Борис Конев</csclub:имя>
</rdf:Description>

```

4) В N3-синтаксисе (удобный для чтения человеком и расширяющий исходную модель данных RDF)

```

<ex:index.html> < csclub:имя > «Онтологии и представление знаний».
<ex:index.html> < csclub:читается> «БК»".
<ex:index.html> < csclub:домашняя страница >
«http://logic.pdmi.ras.ru/csclub/courses/ontology».

```

Словари метадаанных и онтологии

RDFS

RDFS является семантическим расширением RDF. Он предоставляет механизмы для описания групп связанных ресурсов и отношений между этими ресурсами. Все определения RDFS выражены на RDF (поэтому RDF называется «самоописывающимся» языком). Новые термины, вводимые

RDFS, такие как «домен», «диапазон» свойства, являются ресурсами RDF.

Система классов и свойств языка описания RDF-словарей похожа на систему типов объектно-ориентированных языков программирования, например, Java. Но RDF отличается от большинства таких систем тем, что здесь центральным аспектом является определение свойства, а не класса. Свойства в RDF определяются как пары (домен, диапазон). При этом домен представляет некоторое множество классов RDF, к которым данное свойство применимо, диапазон определяет допустимое множество ресурсов – значений свойства. Для сравнения: в Java определение класса имеет законченную форму (свойства класса выражаются в полях и методах класса). В RDF, напротив, описание класса всегда остается открытым (набор свойств класса определяется вне самого класса).

Пример. Определим свойство «автор» с доменом «Документ» и диапазоном «Человек». В случае появления дополнительной информации о свойствах «Документа» нет необходимости изменять описание класса «Документ». Достаточно добавить новое свойство с соответствующим доменом.

Пример a-la RDF:

Класс ("Документ");

Класс ("Человек");

Свойство ("Автор", "Документ", "Человек").

Пример "a-la Java":

Класс "Документ"

{"Человек" "Автор"}

Можно заметить, что при изменении смысла свойств изменять придется именно их. При этом все классы, зависящие от изменяемых свойств, косвенно изменяют свою семантику.

Основное преимущество такого подхода – в легкой расширяемости: добавление/удаление свойств интуитивно проще, чем управление множеством классов, обладающих каждый своим индивидуальным набором свойств (как в ООП). Фактически, любой может расширять описание существующих ресурсов (лозунг Web: «Кто угодно может сказать что угодно о чем угодно!»).

Классы

Ресурсы могут объединяться в группы, называемые классами. Члены

класса (здесь наиболее близкий термин – «экземпляры» или «объекты» ООП) называются экземплярами класса. Сами классы также являются ресурсами и идентифицируются ссылками RDF-URI. Чтобы указать, что ресурс является экземпляром класса, используется свойство `rdf:type` ("`rdf`" здесь применен как префикс пространства имен).

В RDF определение класса или свойства (т.н. интенционал) отделено от множества экземпляров класса и значений свойства (т.н. экстенционала). Так, два класса с одинаковыми экстенционалами считаются различными, если они имеют разные наборы свойств (интенционалы). Рассмотрим множества

$$A = \{0, 2, 4, 6, 8\},$$

$$B = \{x \mid x = 2k, k = 0..4, k - \text{целое}\},$$

C – множество неотрицательных четных чисел, меньших 10.

В этом примере множество A полностью описывается своим экстенционалом, множества B и C описываются интенционалами, т.е. с использованием характеристических свойств данного множества. Множества, имеющие бесконечное число элементов, могут быть описаны только своим интенционалом. Однако при использовании интенционала могут возникнуть парадоксы³⁷. Чтобы избежать их, в теории множеств вводятся дополнительные аксиомы. Примечательно, что RDF нарушает эти аксиомы. Классу RDF не запрещено быть экземпляром самого себя.

Группа ресурсов, являющихся классами, в RDFS описывается термином `rdfs:Class`.

На множестве классов определено отношение ПОДКЛАСС-НАДКЛАСС, описываемое RDFS-свойством `rdfs:subClassOf`. Семантика данного отношения состоит в том, что экстенционал любого подкласса данного класса C целиком содержится (как множество) в экстенционале самого класса C. Другими словами, если ресурс *i* является экземпляром класса C*, а класс C* является подклассом класса C, то *i* является экземпляром класса C.

Любой класс RDFS по определению является подклассом самого

³⁷ Например, парадокс Рассела: пусть множество M – множество всех множеств, не содержащих себя. Содержит ли M само себя? Если содержит, то оно не удовлетворяет своему определению – интенционалу; если M не содержит себя, то оно удовлетворяет определению и, следовательно, должно себя содержать.

себя.

В спецификации по RDFS определены также списки, коллекции и контейнеры ресурсов, текстовые пометки и комментарии для создания удобных для чтения примечаний к ресурсам.

Проблемы с RDFS:

1. Не достаточно выразителен:

- Нет локализованных ограничений на домен и множество значений (множество значений свойства `hasChild` является людьми, когда применяется к человеку и слонами когда применяется к слонам)
- Нет ограничений на существование (*у всех людей есть мать, которая тоже человек*) и мощностность (*у матери-героини больше двух детей*),
- Не определяются транзитивные, обратные или симметричные свойства (свойство `isPartOf` транзитивное, свойство `hasPart` обратно `isPartOf`, свойство *прикасаться* симметрично)
- ...

2. Трудности с логическим анализом:

- Нет средств, поддерживающих нестандартную семантику
- Структурные алгоритмы

OWL

OWL (Web Ontology Language, в аббревиатуре буквы намеренно переставлены местами, чтобы получилось английское слово «сова») – язык представления онтологий в Web. Фактически это словарь, расширяющий набор терминов, определенных RDFS. OWL-онтологии могут содержать описания классов, свойств и их экземпляров. Создание OWL – это ответ на необходимость представления знаний в Сети в едином формате. Исторически предшественником OWL был язык DAML+OIL, объединивший 2 инициативы: проект DAML (DARPA Agent Markup Language) и проект OIL (Ontology Inference Layer). Наиболее ранним проектом представления онтологий в Web был SHOE (Simple HTML Ontology Extensions). OWL с 2004 года является рекомендацией W3C и объединяет лучшие черты своих предшественников.

Язык OWL имеет 3 диалекта (подмножества терминов).

- OWL Lite – имеет наименьшую выразительную мощностность из всех, но для решения простых задач его может быть достаточно. Данный

диалект языка OWL эквивалентен некоторой дескриптивной логике (разрешимой части логики предикатов первого порядка). OWL Lite обладает важнейшим свойством – разрешимостью (т. е. задача вывода следствий из утверждений, сформулированных в этом языке, является вычислимой). Именно разрешимость (и относительно невысокая вычислительная сложность) является главной причиной использования OWL Lite для создания многочисленных практических онтологий (в медицине, биоинформатике и т. п.).

- OWL DL – обладает большей выразительной мощностью, чем OWL Lite, но тоже эквивалентен некоторой (более выразительной) дескриптивной логике. Для большинства задач, встречающихся при проектировании онтологий, выразительности этого диалекта достаточно. OWL DL тоже обладает свойством разрешимости, однако вычислительная сложность у него выше, чем OWL Lite. Разрешимость достигается, в частности, наложением ограничений на синтаксис языка; так, в OWL DL классу запрещено быть экземпляром.
- OWL Full – наиболее выразительный диалект. Эквивалентен RDF. При использовании OWL Full нет никаких гарантий по вычислимости заключений.

Каждый из этих диалектов (кроме OWL Lite) является расширением предыдущего. Как следствие, любая OWL Lite онтология является OWL DL онтологией, а любая OWL DL онтология является OWL Full онтологией.

Базовые элементы OWL

Любая онтология имеет заголовок и тело. В заголовке содержится информация о самой онтологии (версия, примечания), об импортируемых онтологиях. За заголовком следует тело онтологии, содержащее описание классов, свойств и экземпляров.

Классы

В OWL введен новый термин – класс (`owl:Class`). Необходимость этого объясняется тем, что не все классы диалектов OWL DL и OWL Lite являются RDFS-классами (в этом случае `owl:Class` является подклассом `rdfs:Class`). В диалекте OWL Full подобных ограничений нет, и `owl:Class` фактически является синонимом `rdfs:Class`.

Для организации классов в иерархию используется свойство `rdfs:subClassOf`.

Особое место занимают два взаимодополняющих класса – `owl:Thing` и `owl:Nothing`. Первый из них является надклассом любого класса OWL, второй – подклассом любого класса OWL. Экземпляр любого класса OWL входит в экстенционал класса `owl:Thing`. Экстенционал класса `owl:Nothing` является пустым множеством.

OWL-класс может быть описан шестью способами:

1. идентификатором класса (URI);
2. перечислением всех экземпляров класса;
3. ограничением на значение свойства;
4. пересечением 2-х и более определений классов;
5. объединением 2-х и более определений классов;
6. дополнением (логическим отрицанием) определения класса.

Только первый способ определяет именованный класс OWL. Все оставшиеся определяют анонимный класс через ограничение его экстенционала. Способ 2 явно перечисляет экземпляры класса, способ 3 ограничивает экстенционал только теми экземплярами, которые удовлетворяют данному свойству. Способы 4-6 используют теоретико-множественные операции (объединение, пересечение и дополнение) над экстенционалами соответствующих классов, чтобы определить экстенционал нового класса.

Описания класса являются строительными блоками для определения классов посредством аксиом.

Простейшая аксиома, определяющая именованный класс:

```
<owl:Class rdf:ID="Human"/>
```

Всё, что постулирует эта аксиома, – существование класса с именем `Human`.

В OWL определены еще 3 конструкции, комбинируя которые, можно определять более сложные аксиомы классов:

- `rdfs:subClassOf` говорит о том, что экстенционал одного класса (подкласса) полностью входит в экстенционал другого (надкласса);
- `owl:equivalentClass` говорит о том, что экстенционалы двух классов совпадают;
- `owl:disjointWith` говорит о том, что экстенционалы двух классов

не пересекаются. Иногда говорят, что таким образом определяются дизъюнктивные классы.

Свойства

В OWL выделяют две категории свойств: свойства-объекты (или объектные свойства) и свойства-значения. Первые связывают между собой индивиды (экземпляры классов). Вторые связывают индивиды со значениями данных. Оба класса свойств являются подклассами класса `rdf:Property`.

Для определения новых свойств как экземпляров `owl:ObjectProperty` или `owl:DatatypeProperty` используются аксиомы свойств.

Пример аксиомы:

```
<owl:ObjectProperty rdf:ID="hasParent"/>
```

Все, что постулирует данная аксиома, – существование некоторого свойства `hasParent`, связывающего экземпляры класса `owl:Thing` друг с другом.

Кроме того, OWL поддерживает следующие конструкции для построения аксиом свойств:

- Конструкции RDFS: `rdfs:subPropertyOf` (определяет подсвойство данного свойства), `rdfs:domain` (определяет домен свойства) и `rdfs:range` (определяет диапазон свойства).
- Отношения между свойствами: `owl:equivalentProperty` (определяет эквивалентное свойство) и `owl:inverseOf` (определяет обратное свойство).
- Ограничения глобальной кардинальности: `owl:FunctionalProperty` (определяет однозначное свойство – однозначное отображение домена свойства на диапазон) и `owl:InverseFunctionalProperty` (обратно функциональное свойство, т. е. определяет, что свойство, обратное данному свойству, является однозначным).
- Логические характеристики свойства: `owl:SymmetricProperty` (определяет свойство как симметричное) и `owl:TransitiveProperty` (определяет транзитивное свойство).

Индивиды (экземпляры классов)

Индивиды определяются при помощи аксиом индивидов (т. н. фактов). Рассмотрим два вида фактов:

1. факты членства индивидов в классах и факты о значениях

свойств индивидов;

2. факты идентичности/различности индивидов.

Пример аксиом индивидов первого вида:

```
<Балет rdf:ID="ЛебединоеОзеро">
```

```
<имеетКомпозитора rdf:resource="#Чайковский"/>
```

```
</Балет>
```

Данная аксиома постулирует сразу 2 факта: (1) существует некоторый индивид класса Балет, имеющий имя ЛебединоеОзеро; (2) этот индивид связан свойством имеетКомпозитора с индивидом Чайковский (который определен где-то в другом месте). Первый факт говорит о членстве в классе, второй – о значении свойства индивида.

Аксиомы второго вида необходимы для суждения об идентичности индивидов. Дело в том, что в OWL не делается никаких предположений ни о различии, ни о совпадении двух индивидов, имеющих различные идентификаторы URI. Подобные утверждения выражаются аксиомами идентичности с помощью следующих конструкций:

- owl:sameAs постулирует, что две ссылки URI ссылаются на один и тот же индивид;
- owl:differentFrom постулирует, что две ссылки URI ссылаются на разные индивиды;
- owl:AllDifferent предоставляет средство для определения списка попарно различных индивидов.

7.6. Индексирование/Аннотирование ресурсов (создание метаданных)

Использование метаданных, в особенности контентных (семантических), позволяет эффективно решать такие задачи работы со знаниями как поиск, категоризация и рекомендация знаний. Индексирование или аннотирование – это создания метаданных, поэтому 3 данных термина можно считать синонимами. Термин «индексирование» обычно используется в области информационного поиска и information science, «аннотирование» и «создание метаданных» в области семантического веба. Индексирование/аннотирование может выполняться как с участием человека (там, где это необходимо), так и без его участия – автоматически. Однако в связи с тем, что задача понимания текстов на естественном языке до сих пор в полной мере не решена, не

представляется возможным составление качественных контентных метаданных без участия человека. В лучшем случае, этот процесс является полуавтоматическим, когда программы предлагают варианты утверждений для контентных метаданных, а человек анализирует их и либо принимает, либо редактирует или отвергает.

Анализ состава информации современной организации показывает, что основная ее часть содержится в виде текстов на естественном языке – более 80 %, в бумажной и электронной формах. В связи с этим, одной из наиболее сложных задач в построении СУЗ является разработка методов составления достаточно точных контентных метаданных для текстовых документов.

Онтологический (или семантический) подход к решению задачи аннотирования документов предполагает в качестве содержимого метаданных использование элементов онтологии. Задача аннотирования в таком случае заключается в создании семантических метаданных, т. е. в формировании множества утверждений (триплетов), на основе некоторой онтологии и соответствующей ей базе знаний. Возможны ручной и полуавтоматический варианты решения данной задачи.

Ручной вариант реализации заключается в создании редактора метаданных, который позволяет пользователю с помощью специального интерфейса выбрать элементы утверждений, используя онтологию некоторой предметной области и свои знания об аннотируемом объекте (документе, специалисте и т. п.). Основной задачей интерфейса является предоставление возможности конструирования метаданных с одновременной навигацией по онтологии, в том числе и с интерактивной визуализацией отдельных ее частей.

Полуавтоматический вариант реализации предполагает создание подсистемы, которая анализирует объект знаний, имеющий текстовое содержание, а после этого предоставляет пользователю «начальный вариант» семантического метаописания, который пользователь может отредактировать. При этом экономится время специалиста на ознакомление с содержанием объекта.

Автоматизированное формирование метаданных основано на 2-х группах методов:

1. Методы машинного обучения:

опора на статистические (вероятностные) методы, необходим размеченный корпус для «обучения» системы.

2. Методы, построенные на правилах:

опора на языки описания правил-шаблонов (и действий), правила пишутся экспертами; процесс написания правил может занимать много времени.

Автоматизированная поддержка индексации/аннотирования информационных ресурсов осуществляется инструментами, поддерживающими обработку естественного языка, обзор которых выполнен в работе В. Хорошевского³⁸. В настоящее время основное внимание исследователей и разработчиков сосредоточено на системах для извлечения информации из текстов (Information Extraction) – системах, названных по аналогии с разработкой месторождений, системами «разработки» текстов (Text Mining), а также на системах семантической классификации и кластеризации (Semantic Classification/Clustering). «Номенклатура» функционалов, продуктов и сервисов, разрабатываемых коллективами и организациями, представленными в настоящем обзоре, концентрируется, в основном, вокруг выделения из текстов достаточно ограниченного набора именованных сущностей. И только некоторые из них, во-первых, имеют дело с серьезным спектром объектов (Inxight – более 35 типов сущностей, SRA International – 7 основных и более 70 подтипов важных сущностей), а во-вторых, идут на выделение из текстов семантических отношений между сущностями и/или событиями и других артефактов, связанных с объектами (SRA International, Teragram). Из прикладных функционалов достаточно «частотными» являются реферирование (CognIT, Delphes, Megaputer Intelligence, SRA International, Teragram), а также кластеризация и классификация (CognIT, Convera, Megaputer Intelligence, Ontotext, TEMIS, SRA International). Несколько коллективов и организаций, представленных в настоящем обзоре, имеют серьезные разработки в области семантической навигации и визуализации (ClearForest, Compris Intelligence, Inxight, Megaputer Intelligence, Ontotext, SRA International). Таким образом, для сравнения инструментов автоматизированной обработки естественного языка можно выделить

³⁸ Хорошевский В. Пространства знаний в сети Интернет и Semantic Web (Часть 1) // Искусственный интеллект и принятие решений. 2008. № 1.

следующие задачи: выделение объектов и отношений (извлечение сущностей / information extraction), реферирование, классификация, кластеризация (табл. 7.6).

Таблица 7.6

Сравнительный анализ инструментов обработки естественного языка

| Коллективы, организации | Спектр обр. языков | Средства предв. обработки | Инстр. | Спектр выд. объектов (отн.) | Рефер. | Кластер., классиф. | Сем. навиг. и визуал. |
|---------------------------|--------------------|---------------------------|--------|-----------------------------|--------|--------------------|-----------------------|
| BASIS Tech | 16 | + | + | +(-) | | | |
| ClearForest | 1 | + | | +(+) | | | + |
| CognIT | 1+? | + | | +(-) | + | + | |
| Compris Intelligence | 2+? | + | | +(-) | | | + |
| Convera | много | + | | +(-) | | + | |
| Delphes | 4 | + | | +(-) | + | | |
| Megaputer Intelligence | 1 | + | | +(-) | + | + | + |
| Insightful Corp. & InFact | 1+? | + | | +(-) | | | |
| Inxight Software | 15 | + | + | 35(+) | | | + |
| MITRE | 1+? | + | + | +(-) | | | |
| Ontotext | | + | + | +(+) | + | + | + |
| SRA International | много | + | | 7 осн. и 70 подтипов (+) | + | + | + |
| TEMIS | 12 | + | + | +(-) | | + | |
| Teragram | Все евр. | + | + | +(-) | + | | |

Для примера, рассмотрим более подробнее продукты и решения компании CognIT, которые концентрируются вокруг следующих основных областей применения: инновационные машины поиска, Content Management Support (CMS) и Knowledge Management and Best Practice (KM). Если говорить более конкретно, продукты компании, разработанные на основе технологий, созданных исследовательскими лабораториями Норвегии, зарегистрированы под маркой CORPORUM® и могут использоваться для интеллектуального поиска и индексирования, структуризации контента на порталах, аннотирования документов по контенту, реферирования и сжатия информации, а также для извлечения имен и отношений из текстов.

Линейка продуктов CORPORUM® Intranet Search & Navigation включает CORPORUM® SLATEWEB и CORPORUM® engine. Инсталляция CORPORUM® SLATEWEB поступает к пользователю с несколькими готовыми для использования приложениями, такими как web crawler, file indexer и др. В нее входят также анализаторы для нескольких форматов документов (в частности, PDF и Microsoft Word). И, наконец, SLATE WEB включает последнюю версию CORPORUM® engine,

поддерживающего английский, норвежский, немецкий и шведский языки. Основные функционалы SLATEWEB связаны с поиском (точные образцы и множественные термы, полная поддержка булевских выражений, неограниченная вложенность скобок), сортировкой (по релевантности и дате), интервальным поиском (фильтрация результатов по дате, релевантности и предметной области), лингвистическим анализом (автоматическое определение языка документа, NLP, реферирование и извлечение концептов, сущностей, связанных концептов), таксономией и классификацией (категоризация на основе правил), автоматической генерацией таксономии, а также поддержкой краулинга по многим источникам (Internet, Intranet, IMAP, Exchange, Oracle, MSQl) и с некоторыми другими функционалами.

В основе CORPORUM® engine лежит технология CognIT Mimir, с помощью которой тексты интерпретируются на основе онтологий, отражающих модель интересов пользователя, а сама модель интересов рассматривается как база знаний для определения контекстного и тематического соответствия обрабатываемых документов этой модели. При этом модель интересов пользователя управляет процессом извлечения информации, а результаты сохраняются в базе данных для дальнейшего использования. Для представления и обработки знаний используется специальная среда The CORPORUM® portfolio, включающая, по последним данным, CORPORUM® KnowledgeFactory, CORPORUM® Knowledge Server, CORPORUM® Summarizer и CORPORUM® Best Practice. Более развернутая информация о детальных характеристиках отдельных составляющих CORPORUM® portfolio практически недоступна.

Другой масштабной разработкой в области семантического анализа текстов является продукт компании IBM – UIMA (Unstructured Information Management Architecture)³⁹. Данная платформа предназначена для использования ее в качестве основы для построения сложных систем семантического анализа неструктурированной информации.

Архитектура UIMA унифицирует процесс обработки не только текстов на естественных языках, но и позволяет анализировать

³⁹ UIMA Tutorial and Developers' Guides. <http://www.research.ibm.com/UIMA>

мультимедиа файлы.

UIMA позволяет создавать сложные приложения, где каждый компонент выполняет определенные функции: идентификация языка; синтаксический анализ или непосредственное аннотирование текста. Для каждого элемента системы четко определены входные параметры и формат результирующих данных на выходе, что позволяет компонентам с легкостью использовать для своего алгоритма данные результаты других блоков. Четкая спецификация взаимодействия составляющих системы предоставляет разработчику возможность повторного использования компонентов других авторов.

Синтаксис системы типов позволяет описывать иерархии одиночного наследования. Каждый пользовательский тип расширяется набором свойств (features). Таким образом, система типов может интерпретироваться как таксономия понятий предметной области. Система типов отображается утилитой автоматической генерации кода в набор Java_классов. Каждый аннотатор внедряет специальную логику идентификации набора символов в тексте как экземпляр того или иного класса. Результаты аннотирования индексируются и могут быть доступны третьим компонентам, отвечающим непосредственно за предоставление определенной функциональности пользователю. Функциональная специализация отдельных блоков обуславливает их относительную самодостаточность и возможность повторного использования, а система, построенная из отдельных компонентов, является масштабируемой.

Типичная схема построения приложений на базе UIMA представлена на рис. 7.14.

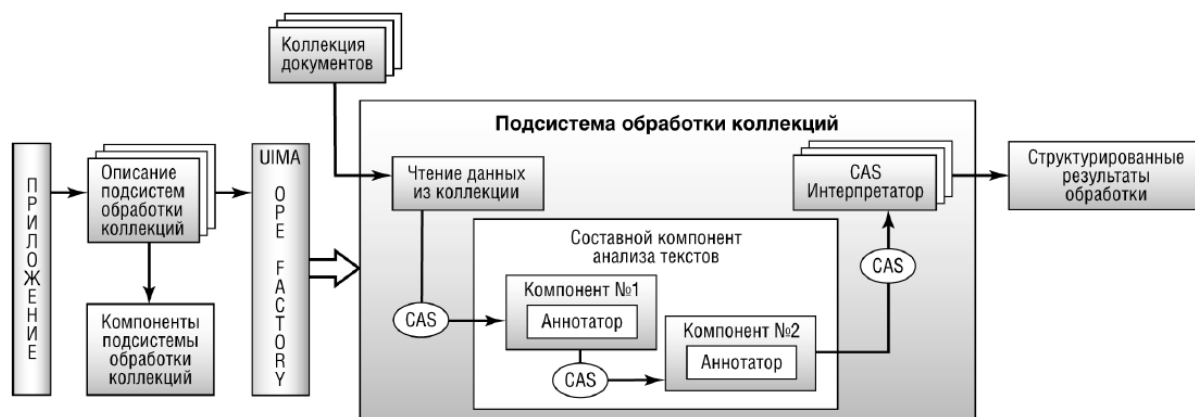


Рис. 7.14. Архитектура системы семантической обработки текстов

Представленная на рисунке подсистема обработки коллекций (CPE – Collection Processing Engine) состоит из набора компонент-аннотаторов. Обмен данными между различными компонентами системы осуществляется посредством общей структуры для анализа (CAS, Common Analysis Structure), представляющей результаты аннотирования текста. Применение технологии IBM для построения приложений семантического анализа данных не ограничивается научно-исследовательскими проектами. Данная концепция нашла свое применение в ряде коммерческих продуктов, перечень которых варьируется от интеллектуального поиска (OmniFind от IBM) до систем автоматизированной обработки заявок в службе поддержки пользователей.

7.7. Референтные модели, справочники, типовые элементы

Распространенными инструментами для накопления знаний в любой области являются справочники, референтные модели, типовые элементы (типовые проекты, типовые архитектуры или конструкции, типовые компоненты, типовые ошибки) и их описания, шаблоны (patterns), библиотеки.

Справочники – инструменты передачи знаний, которые, как правило, описывают один объект (например, справочник оборудования, справочник процессов, справочник показателей и т. д.). Косвенно справочники могут нести информацию и о других объектах, например, когда показатели (основной объект) структурированы по процессам. На рис. 7.15 приведен фрагмент списка справочников (библиотечных классификаторов), используемых в системе автоматизированной поддержки организационного проектирования.

Преимущества использования референтных моделей и справочников:

- Проверенные на практике способы организации деятельности;
- Возможность сравнения с другими компаниями (бенчмаркинг);
- Легкость и скорость объединения компаний при слияниях и поглощениях, благодаря повышению вероятности сходства систем процессов объединяемых компаний;
- и прочие преимущества стандартизации.

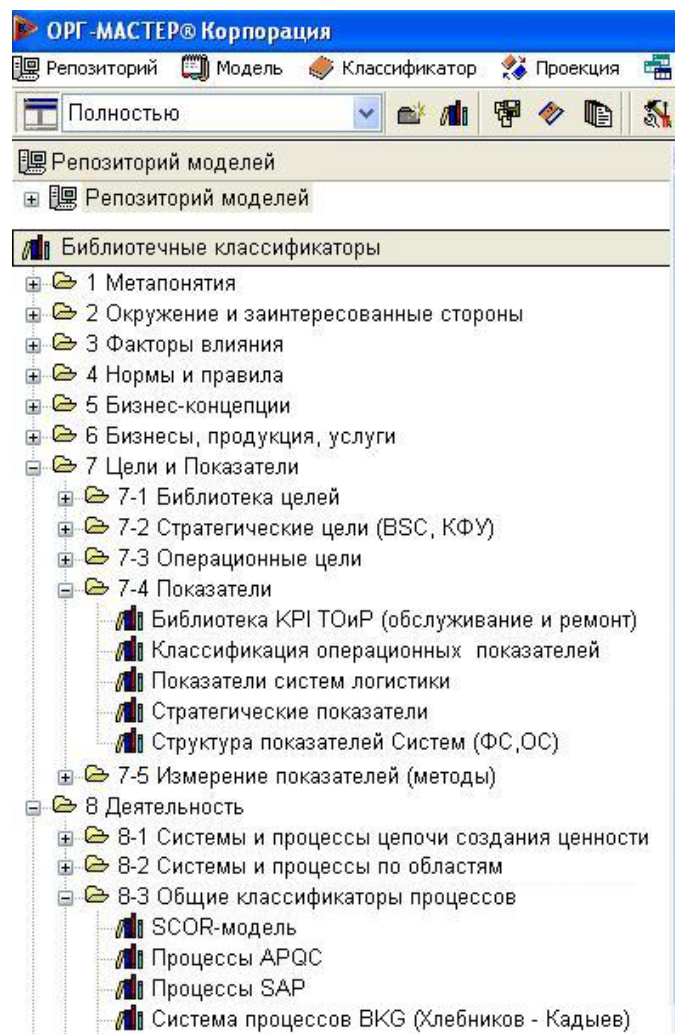


Рис. 7.15. Библиотечные классификаторы (справочники) для поддержки организационного проектирования

Под референтной моделью будем понимать модель деятельности абстрактного предприятия, которая направлена на фиксацию лучших практик организации деятельности. Традиционно референтные модели включают в себя систему процессов, их цели, результаты, показатели, состав работ по процессам, краткое описание методик выполнения процессов и отдельных операций. Референтная модель может описывать как предприятие в целом (например, референтная модель деятельности компании сферы телекоммуникаций e-TOM), так и его часть (например, ITIL – библиотека инфраструктуры информационных технологий, описывающая лучшие из применяемых на практике способов организации работы подразделений или компаний, занимающихся предоставлением услуг в области информационных технологий).

8. РАСПРЕДЕЛЕНИЕ ЗНАНИЙ

8.1. Тянущий (поиск) и толкающий (доставка) подходы в распределении знаний

В процессе распределения⁴⁰ нужные знания в нужном виде в нужное время и место доставляются нужным потребителям знаний. В распределении знаний можно выделить два основных подхода: «Вытягивание» («pull») и «Выталкивание» («push»), см. рис. 8.1. Согласно основам теории коммуникации термины pull (англ. тянуть) и push (англ. толкать) используются для обозначения инициатора коммуникации. Если коммуникация инициирована получателем, то это – pull-коммуникация. Если коммуникация инициирована источником, то это – push-коммуникация. Далее для первого подхода будет использоваться понятие «Поиск» (иногда его называют «Активным поиском»), а для второго «Доставка» (иногда его называют «Пассивным поиском»).



Рис. 8.1. Подходы к распределению знаний

⁴⁰ Формулировка «Распределение знаний» соответствует точке зрения субъекта управления (знаниями), а не участников данных процессов. Пользователь системы, который получает нужные знания в результате их распределения, может называть данный процесс получением знаний.

Поиск (или поиск информации, или информационный поиск) – процесс нахождения, отбора и выдачи определенной заранее заданными признаками информации (в том числе документов, их частей и/или данных) из массивов и записей любого вида и на любых носителях. Побудительной причиной осуществления поиска является информационная потребность,

Поиск осуществляется либо с помощью запросов, либо с помощью навигации, либо путем их комбинации, например, используя навигацию по тезаурусу или онтологии для отбора понятий и их связей, которые будут включены в поисковый запрос.

Запрос (*query, request, interrogation*) – входное сообщение в автоматизированную систему, содержащее требование на выдачу информации или (в более общем случае) на выделение ресурсов. Для формирования запросов может использоваться информационно-поисковый язык, рассмотренный ранее в разделе «Накопление знаний». Поиск может осуществляться по ключевым словам, по атрибутам или, другими словами, по элементам метаданных (например, по теме и году создания материала; по автору и названию журнала и т. п.),

Для формирования запросов, как правило, используется специальный язык запросов. В табл. 8.1 рассмотрен фрагмент памятки по использованию языка запросов системы «Яндекс» (<http://help.yandex.ru/search/?id=1111313>).

Таблица 8.1

Пример языка запросов

| Пример | Значение |
|--|-------------------------------------|
| «К нам на утренний рассол» | Слова идут подряд в точной форме |
| «Прибыл * посол» | Пропущено слово в цитате |
| полгорбушки & мосол | Слова в пределах одного предложения |
| снаряжайся && добудь | Слова в пределах одного документа |
| глухаря куропатку кого-нибудь | Поиск любого из слов |
| title:(в стране) | Поиск по заголовкам документов |
| url:ptici.narod.ru/ptici/kuropatka.htm | Поиск по URL |
| mime:pdf | Поиск по одному типу файлов |
| ... | ... |

Навигация – метод поиска информации, при котором пользователь для получения информации перемещается по элементам информационно-поискового языка, либо переходит по гипертекстовым ссылкам от одной страницы к другой. Для навигации в традиционных СУЗ используются классификационные системы, предметные рубрики, информационно-поисковые тезаурусы, в семантических онтологиях. Описание разновидностей ИПЯ и методика разработки классификационной системы были приведены выше в разделе «Накопление знаний».

Запросный поиск обычно сочетают с навигацией, например, ограничивая область поиска по запросу.

Выбор метода поиска зависит от стратегии поиска (рис. 8.2), среди которых выделяют:

1. *Поиск известного объекта*, когда пользователь знает чего хочет, где это находится и как сформулировать запрос;

2. *«Исследовательский» поиск*, когда пользователь имеет расплывчатое представление о желаемых результатах поиска, не может подобрать слова для запроса, не знает где искать, но, возможно, узнает то, что ищет, если увидит;

3. *Повторный поиск*, когда пользователь ищет то, что уже когда-то видел;

4. *«Незнание незнания»*, когда пользователь ищет одно, а нужно ему другое (или «еще и» другое).

В случае *«выталкивания» (push) или доставки* пользователь получает информацию без своего активного участия

Технологии доставки можно разделить на три основные группы:

- Автоматизированные рассылки,
- RSS,
- Технологии «Смышленное вытягивание» (Smart pull) и «Реальной доставки» (Real push).

В общем случае можно выделить следующие особенности технологий доставки:

1. Независимость от пользователя. В первом приближении можно сказать, что технологии доставки работают независимо от пользователя. Т. е. информация доставляется пользователю без его непосредственного (активного) участия.

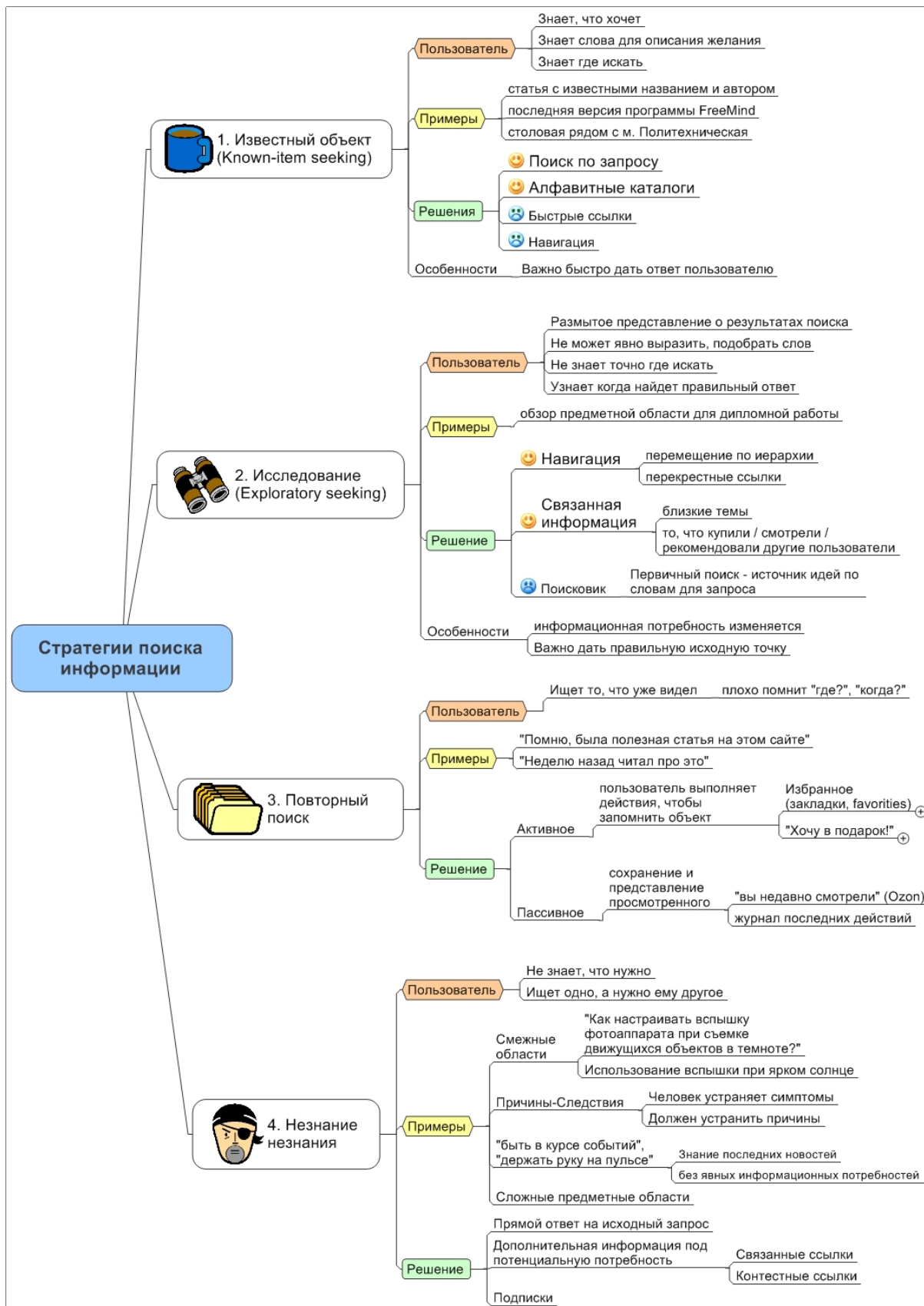


Рис. 8.2. Стратегия поиска информации определяет требования к системе

2. Адресность. Технологии доставки позволяют предоставлять информацию определенной группе пользователей, той, для которой эта информация предназначена и которые заинтересованы в ее получении.

3. Периодичность. Технологии доставки предназначены для многократного предоставления актуальной информации пользователям.

Автоматизированные рассылки (электронная почта)

Автоматизированная рассылка для пользователя это не что иное, как обычная электронная почта. Тем не менее, на сервере предприятия должно быть программное обеспечение, обеспечивающее рассылку электронных писем на различные адреса электронной почты. При этом используется база данных, где содержатся параметры сотрудников (в частности его e-mail) и информация о принадлежности сотрудников к некоторым группам пользователей.

Возможны такие реализации рассылок, когда письма генерируются в зависимости от предпочтений, указанных пользователем. Т. е., например, сотрудник консалтинговой фирмы может указать, что его интересуют только отчеты по выполненным проектам в области разработки бизнес-стратегии в Азии.

Достоинства: отсутствие необходимости обучения сотрудников новым технологиям; простота технической реализации и использования.

Недостатки: автоматизированные рассылки совершенно не подходят для областей, где необходима гарантированная доставка информации в реальном времени. Дело в том, что невозможно предсказать, когда информация дойдет по электронной почте до конкретного пользователя (и дойдет ли вообще).

Рекомендуемой областью применения является распространение информации (статей, документации, новостей и т. п.), не требующей сиюминутного внимания и для которой требование гарантированной доставки не является определяющим фактором. В автоматизированные рассылки можно включать: информацию о состоянии выполнения проектов; информацию о новых изделиях и новых услугах; документацию, рекомендации по использованию продукции; новости отрасли, информацию об изменении стратегии фирмы и т. п.; формы для опросов; статьи, которые могут заинтересовать сотрудников. Почти всегда имеет смысл создать несколько групп рассылки.

RSS

RSS – семейство XML-форматов, предназначенных для описания лент новостей, анонсов статей, изменений в блогах и т. п. Информация из различных источников, представленная в формате RSS, может быть собрана, обработана и представлена пользователю в удобном для него виде специальными программами-агрегаторами.

В разных версиях аббревиатура RSS имела разные расшифровки:

Rich Site Summary (RSS 0.9x) – обогащенная сводка сайта;

RDF Site Summary (RSS 0.9 и 1.0) – сводка сайта с применением инфраструктуры описания ресурсов;

Really Simple Syndication (RSS 2.x) – очень простой сбор сводной информации.

Обычно с помощью RSS 2.0 дается краткое описание новой информации, появившейся на сайте, и ссылка на ее полную версию. Интернет-ресурс в формате RSS называется RSS-каналом, RSS-лентой или RSS-фидом.

Многие современные браузеры, почтовые клиенты и интернет-пейджеры умеют работать с RSS-лентами, среди них Safari, Maxthon, Miranda, Mozilla Firefox, Mozilla Thunderbird, Opera, Opera mini, Windows Internet Explorer (начиная с 7-й версии), Google Chrome. Кроме того, существуют специализированные приложения (RSS-агрегаторы), собирающие и обрабатывающие информацию RSS-каналов. Также очень популярны веб-агрегаторы, представляющие собой сайты по сбору и отображению RSS-каналов, такие как Яндекс. Лента, Google Reader, Новотека и Bloglines.

Технологии «Смышленное вытягивание» (Smart pull) и «Реальной доставки» (Real push)

В технологии Smart-pull все действия по отслеживанию изменений в документах на сервере, их скачиванию в фоновом режиме и оповещению пользователя берет на себя его собственное программное обеспечение (off-line браузер). В некоторых случаях она представляет собой загрузку новой информации с сайтов и последующий ее просмотр. В другом случае эта технология предоставляет сервисы аналогичные технологии Real-push. Отличие состоит в том, что программное обеспечение, управляющее процессом передачи информации, расположено на компьютере

пользователя. В технологии Real-push сервер формирует в сети поток информации, к которому по мере необходимости подключаются пользователи. Доставка информации реализуется как фоновая, низкоприоритетная задача предоставления данных. Существует большое количество реализаций этой технологии, как с аппаратной, так и с программной точки зрения. Поэтому невозможно обобщить все ее положительные и отрицательные стороны.

В настоящее время компании (BackWeb, Magimba и др.) предлагают комплексные решения с использованием push-технологий для различных областей коммерческой деятельности (B2C, B2B, обслуживание вызовов по телефону, брокерское обслуживание, связь с реселлерами, информационное обеспечение консалтинговых центров, распространение обновлений программного обеспечения и т. д.).

Таблица 8.2

Сравнительная характеристика методов «вытягивания» (pull, поиска) и «выталкивания» (push, доставки) информации

| | Метод вытягивания («пулл») | Метод «выталкивания» («пуш») |
|-----------------------|--|--|
| Цель | Поиск информации (навигация) | Поглощение (захват) информации |
| Пользователь | Активен: инициативно ищет информацию | Пассивен: некий источник посылает информацию |
| Пригоден для | Одноразовой потребности в знаниях Проведения исследования Получения детальной информации | Долговременного спроса на знания получения срочной или быстроустаревающей информации |
| Основное преимущество | Позволяет получить доступ к информации в нужное время | Привлекает внимание к важным материалам |
| Основной недостаток | Пользователь тратит много времени (на овладение технологией поиска) | Перегрузки пользователя, так как он не контролирует процесс поступления информации |
| Примеры | Поисковик, каталоги, указатели, браузер (просмотровая система) | E-mail |

Эффективность распределения знаний зависит от многих факторов.

1. *Необходимо определить потребность в информации.* При этом следует обеспечить понимание работниками, как будет использоваться информация, и правильное направление информационных запросов.

2. *Персонал должен быть осведомлен о наличии информации,* т. е. должны быть указатели, каталоги, телефонные справочники, навигаторы и т. д., а также возможность обращения к профессионалам для получения сведений о накопленных в организации знаниях. Содействие в поиске информации обеспечивается введением в организации новых должностей специально для помощи тем, кто занят поиском информации, а также привлечением экспертов для фильтрации информации.

3. *Обеспечение доступа к информации.* Для получения информации методами «pull» и «push» необходимо обеспечить пользователю наиболее удобные средства и поддерживать баланс между этими методами (табл. 8.2).

8.2. Оценка качества поиска и доставки знаний

Пользователь для удовлетворения определенной информационной потребности формализует ее средствами, которые ему дает тот или иной информационно-поисковая система. Если пользователь имеет дело с информационно-поисковой системой Гугл или Яндекс, информационная потребность будет формализована в виде последовательности слов или в виде расширенного запроса. Если поиск происходит по каталогу, формализацией информационной потребности будет выбранная рубрика каталога.

После того, как информационная потребность пользователя формализована, происходит поиск.

Пользователь получает список интернет-ресурсов, соответствующих его запросу. Информационно-поисковые системы используют для этого свой индекс – набор проиндексированных ранее интернет-ресурсов. Интернет-каталог выдает все хранящиеся в нем ссылки, относящиеся к выбранной рубрике. Выдаваемый список интернет-ресурсов упорядочен по степени соответствия поисковому запросу. Таким образом, в нашей модели информационного поиска имеем: пользователя, желающего удовлетворить

определенную информационную потребность, формализованный запрос к поисковой системе и выдачу поисковой системы – упорядоченный набор интернет-ресурсов. Исходя из этой модели, сформулируем параметры качества информационного поиска. Наиболее распространенные параметры качества: релевантность, пертинентность, точность, полнота.

Пертинентность – мера качества поиска; насколько хорошо результат поиска удовлетворяет информационную потребность пользователя, т. е. это соответствие информации, полученной в результате поиска, потребности пользователя. Пертинентность определяется субъективным восприятием пользователя: в какой степени документ удовлетворяет его информационную потребность. Информационная потребность пользователя может быть выражена в формализованном запросе с той или иной степенью полноты и точности.

Релевантность – мера того, как хорошо список результатов отвечает на запрос; определяет порядок, в котором результаты поиска представлены пользователю. Когда есть большое количество найденных ресурсов, поисковая машина сортирует список результатов так, чтобы более релевантные страницы оказались в списке раньше, чем менее релевантные.

Релевантность – более узкое понятие, чем пертинентность. Документ может быть релевантен формализованному запросу, но при этом может не удовлетворять информационную потребность пользователя. Различают содержательную и формальную релевантность – по методу ее определения. Формальная релевантность – соответствие, определяемое алгоритмически сравнением поискового предписания и поискового образа документа, на основании применяемого в информационно-поисковой системе критерия выдачи. Содержательная релевантность – соответствие документа информационному запросу, определяемое неформальным путем.

Пертинентность – степень соответствия между ожиданиями пользователя и результатами поиска, отношение объема полезной для пользователя информации к общему объему полученной информации, найденной поисковой системой. Достижение высокой пертинентности – основная задача современных поисковых систем.

Существуют еще две характеристики информационного поиска, более глубоко раскрывающие релевантность.

Точность – мера эффективности поиска, выраженная в виде отношения числа найденных релевантных ресурсов к общему количеству ресурсов, содержащихся в выдаче поисковой системы в ответ на формализованный запрос.

Полнота – мера эффективности поиска, выраженная в виде отношения числа релевантных ресурсов, извлеченных поисковой системой из Интернета в ответ на формализованный запрос, к общему количеству релевантных ресурсов, содержащихся в Интернете.

Полнота показывает, насколько хорошо поисковая система находит то, что нужно пользователю; точность показывает, насколько хорошо поисковая система отфильтровывает то, что пользователю не нужно.

В приведенных выше определениях точности и полноты поиска мы можем заменить «релевантный» на «пертинентный». В таком случае получим, что точность – это доля пертинентных ресурсов среди всех ресурсов, присутствующих в выдаче, а полнота – это доля пертинентных ресурсов, присутствующих в выдаче, среди всех пертинентных ресурсов, имеющих в сети Интернет. Очевидно, что при таком изменении определений полноты и точности изменятся и их числовые значения для конкретных результатов обработки конкретных запросов. Есть некоторый произвол в определении мер эффективности информационного поиска.

Чтобы сделать наше исследование точнее, потребуется ввести более формальное определение параметров эффективности поиска (рис. 8.3).

| | Оценка РЕЛЕВАНТНЫ | Оценка НЕ РЕЛЕВАНТНЫ |
|-------------------------------|---|----------------------------------|
| Объекты НАЙДЕНЫ | а. число объектов релевантных и найденных | б. Не релевантных и найденных |
| Объекты НЕ НАЙДЕНЫ | в. Релевантных и не найденных | г. не релевантных и не найденных |

$$\text{Точность} = \frac{а}{а + б}$$

$$\text{Полнота} = \frac{а}{а + в}$$

Высокая точность = максимизация а, минимизация б

Высокая полнота = максимизация а, минимизация в

см. Nenad Stojanovic «Information Retrieval»

Рис. 8.3. Тематическая релевантность

8.3. Применение онтологий для поиска информации

А. Расширенный с помощью онтологии поиск по ключевым словам / Augmenting Keyword Search with Ontology

Многие инструменты поиска по ключевым словам используют тезаурус (как разновидность онтологии) для расширения запросов. Особенно распространено и известно применение тезауруса WordNet (<http://wordnet.princeton.edu>). Все инструменты расширения поиска данного класса работают примерно по одной схеме: сначала заданные в запросе ключевые слова находятся в тезаурусе/онтологии, далее путем обхода графа выявляются смежные понятия, в заключении термины, связанные с выявленными понятиями, используются для расширения или сужения запроса.

Для поиска дополнительной релевантной для запроса информации на основе полученных стартовых результатов полнотекстового поиска предложен специализированный алгоритм. Вначале к коллекции документов применяется традиционный поиск по тексту. Далее начинается процесс обхода RDF графа, который начинается с RDF-аннотаций, найденных на предыдущем шаге документов. Целью является поиск связанных понятий, таких как автор документа, проект, на который ссылается документ и т. д.

Система CIRI предоставляет внешний интерфейс на основе онтологии для последующего полнотекстового поиска. Поиск производится с помощью браузера онтологий, который визуализирует родивидовые иерархии онтологий, из которых могут быть выбраны понятия для ограничения поиска. Сам поиск производится традиционным поисковиком по ключевым словам, которые связаны с выбранным понятием онтологии и его разновидностями.

В. Многоаспектный (фасетный) поиск на основе онтологии

Мощным инструментом поиска является многоаспектный (фасетный) поиск, а также его сочетание с онтологиями. Данный метод поиска реализован в порталах, использующих Ontogator⁴¹, и OntoViews⁴². В

⁴¹ Hyvönen E., Saarela S., Viljanen K. Ontogator: Combining view- and ontology-based search with semantic browsing // In Proc. XMLFinland '03, Oct. 2003.

многоаспектном (фасетном) поиске используются представления (views) для доступа к данным, см. рис. 8.4, 8.5.

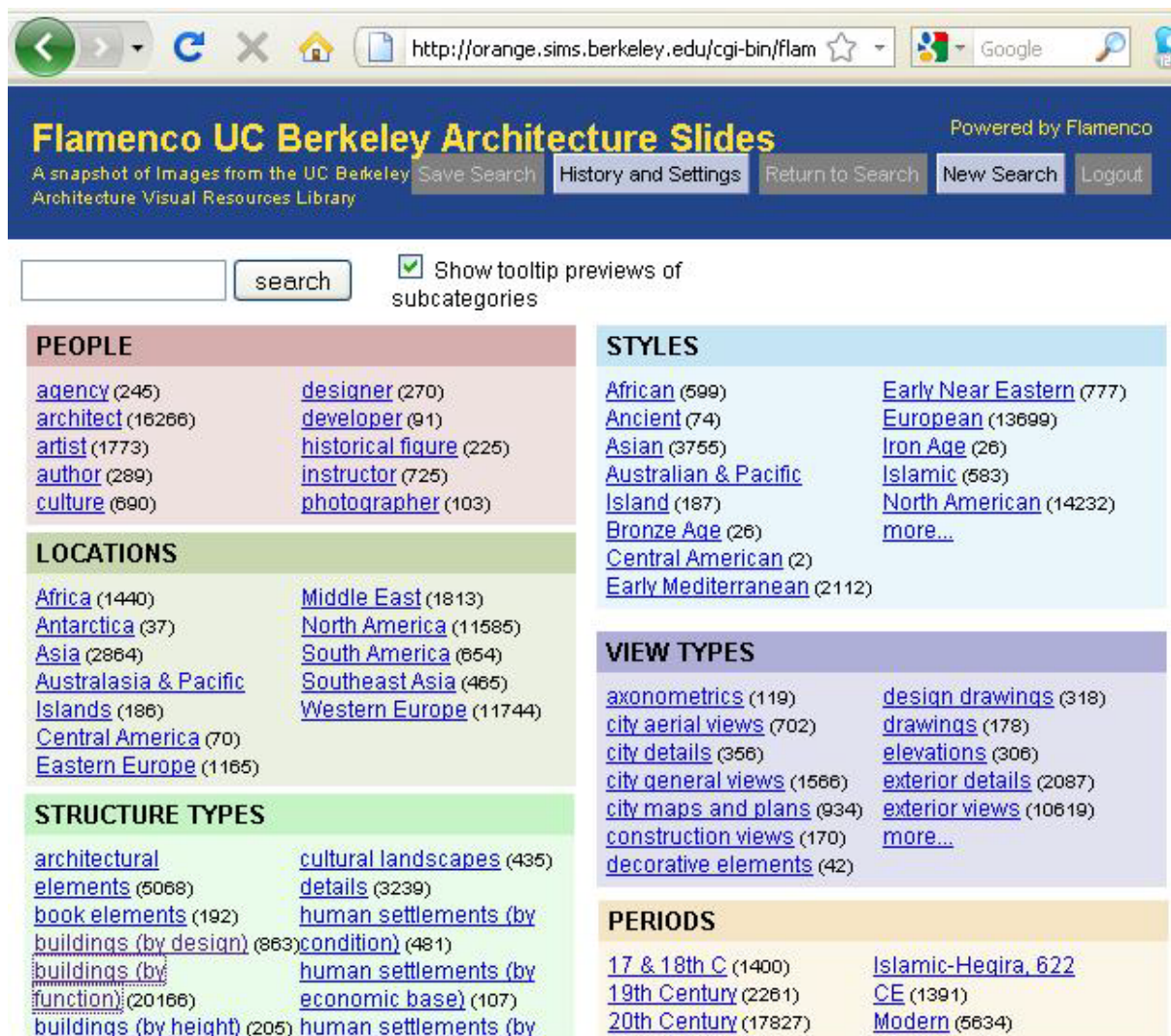


Рис. 8.4. Пользовательский интерфейс для многоаспектного поиска в библиотеке изображений по архитектуре

⁴² Mäkelä E., Hyvönen E., Saarela S., Viljanen K. OntoViews — a tool for creating semantic web portals // In Proc. 3rd Int. Conf. Semantic Web, Springer Verlag, 2004.

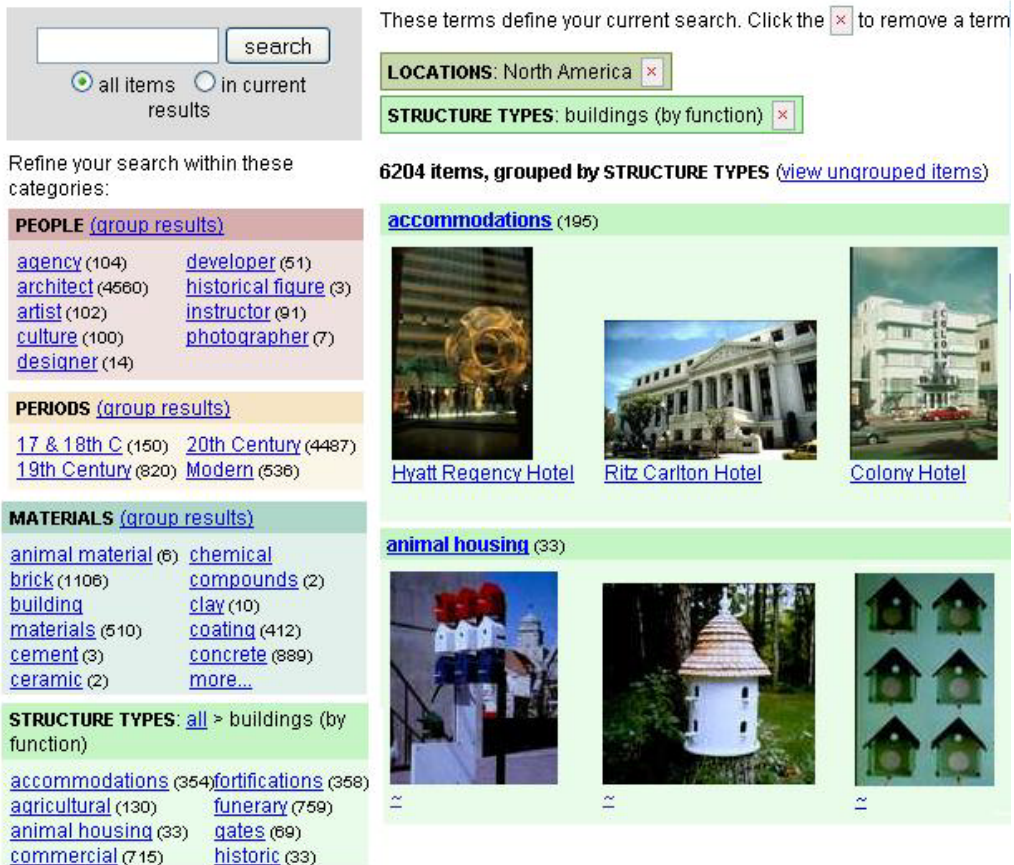


Рис. 8.5. Сужение области поиска в процессе многоаспектного поиска

Представления создаются путем проецирования онтологии (если она используется) на аспекты (точки зрения, фасеты), см. рис. 8.6.

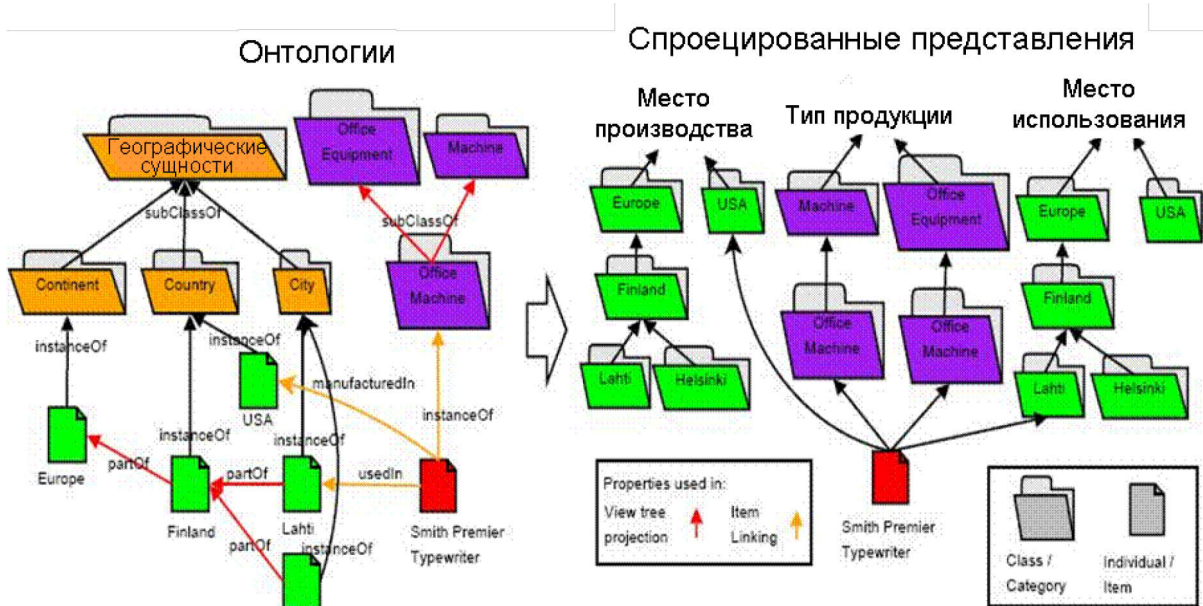


Рис. 8.6. Настройка многоаспектных представлений с помощью онтологий

В Ontogator-е онтология предметной области отображается на аспекты (фасеты, точки зрения) и поддерживает многоаспектный (фасетный) поиск. Когда пользователь находит интересующую информацию с помощью многоаспектного поиска, онтогатор рекомендует ему связанные ресурсы благодаря использованию онтологии и аннотаций (метаданных) ресурсов.

Другим примером интеграции многоаспектного поиска с онтологиями является Longwell браузер, разработанный в рамках Simile проекта⁴³.

Преимуществом многоаспектного (фасетного) поиска является то, что пользователь может накладывать такие ограничения на информационно-поисковый массив и в такой последовательности, которые ему более удобны.

С. Поиск по онтологии (Native Ontology-based Search)

Информация в семантической базе знаний состоит из двух частей: онтологическая часть и экземпляры (соответствующие конкретным объектам предметной области). Как правило объектом поиска являются сущности, принадлежащие классу, однако знания предметной области обычно состоят в описании взаимосвязей между классами онтологии. В системе SHOE⁴⁴ пользователю предлагается визуализация родовидового дерева онтологии, в котором он может выбрать класс интересующих его объектов. Далее ему предлагаются связи и свойства, ассоциированные с выбранным классом. С помощью данных связей и свойств пользователь может ограничить множество экземпляров путем наложения фильтра на их значения. Пример запроса, который может быть получен таким образом: «найти все публикации автора с конкретным именем, относящиеся к конкретному проекту». Подобная функциональность предоставляется в SEAL портале⁴⁵.

⁴³ <http://simile.mit.edu/wiki/Longwell>

⁴⁴ Heflin J., Hendler J. Searching the web with SHOE, In Papers from the AAAI Workshop, pp. 35-40, 2000.

⁴⁵ Maedche A., Staab S., Studer R., Sure Y., Volz R. SEAL – Tying Up Information Integration and Web Site Management by Ontologies // IEEE Data Engineering Bulletin 25 (1), 2002. P. 10-17.

Поисковый механизм системы OntoLoger⁴⁶ записывает поведение пользователя в онтологию и воспроизводит его. Эта система основана на анализе использования информации, размещенной на портале, который создан на основе онтологии. Структура онтологии отражает потребности пользователей. Используя ее, OntoLoger помогает пользователю точнее сформулировать свой запрос.

Система Knowledge Sifter использует онтологии совместно с агентами. В данной системе запрос пользователя уточняется путем «консультирования» с «онтологическим агентом» (ontology agent), который предоставляет концептуальную модель предметной области.

Система Haystack развивает подходы рассматриваемого класса, обеспечивая навигацию от ресурса к ресурсу. В основе данной системы лежит идея о том, что пользователи часто не знают или не помнят точных характеристик искомого объекта, но примерно представляют, с какими другими объектами он должен быть связан. Процесс поиска состоит в навигации по ресурсам, которые как-то связаны с искомым и знакомы пользователю до тех пор, пока не будет собрана информация, достаточная для нахождения искомого ресурса.

Формулирование запросов

Многие комплексные запросы могут быть преобразованы в задачу нахождения группы объектов определенного типа, которые связаны определенным типом отношений. Для решения такой задачи с помощью онтологии формируется шаблон графа с ограничениями на типы узлов и дуг. Например,

– пусть есть запрос «Найти все статьи, опубликованные в тезисах конференций IEEE с 2003 по 2005 на тему семантического веба, на которые ссылаются статьи 2009 года», тогда:

– «статьи», «тезисы конференций IEEE», «семантический веб», годы – классы и экземпляры онтологии, ограничивающие выбор узлов графа;

– «опубликована в», «ссылаться на», «дата публикации» – требуемые дуги в шаблоне.

Хотя такие шаблоны легко превратить в запрос на стандартном

⁴⁶ Stojanovic N., Gonzalez J., Stojanovic L. Ontologer – A system for usage-driven management of ontology-based information portals. In Proc. L-CAP '03 Conf., 2003.

языке (например, SPARQL), использование их затруднено, поскольку пользователю сложно их сформулировать. Поэтому возник ряд подходов к разработке интерфейса для максимально интуитивного формирования таких шаблонов/запросов.

Например, в системе Knowledge Sifter⁴⁷ существует графический пользовательский интерфейс для построения запросных шаблонов графа, основанный на навигации по онтологии (GRQL). Вначале пользователь выбирает класс, который станет начальной точкой. Все имеющиеся свойства выбранного класса предлагаются пользователю для расширения шаблона/запроса. Выбирая какое-либо свойство, пользователь расширяет шаблон графа и переводит фокус внимания (выделение) на класс, задающий область допустимых значений для выбранного свойства. Например, если пользователь, стоя на классе «Художник», выбрал свойство «нарисовал», то он получает шаблон «Художник → нарисовал → Картина», а фокус внимания переносится на класс «Картина», и система предоставляет свойства этого класса для дальнейшего расширения шаблона/запроса. Кроме удлинения навигационного пути, другие операции могут быть применены к шаблону/запросу. Шаблон можно сузить, чтобы он выбирал только определенные подклассы заданного ранее класса, например, для предыдущего примера «Художник → нарисовал → Пейзаж». Подобным образом, ограничения на свойства могут быть сужены до их разновидностей (подсвойств). Более сложные запросы могут создаваться путем возврата к уже выбранному узлу и создания дополнительной ветви, например, с «Натюрмортом» (рис. 8.7).

| Навигация | Шаблон | RQL-запрос |
|-----------|---|---|
| | <p>Художник → нарисовал → Пейзаж, Натюрморт</p> | <p>Select x2 from {x1} нарисовал {x2; Пейзаж}, Натюрморт {x2}</p> |

Рис. 8.7. Использование графического интерфейса и навигации по онтологии для построения запросов

⁴⁷ Kerschberg L., Chowdhury M., Damiano A., et al., Knowledge Sifter: Ontology-Driven Search over Heterogeneous Databases. In Proc. 16th Int. Conf. Scientific and Statistical DB Management, 2004.

Также примером интерфейса для формирования запросов может являться многоаспектный (фасетный) поиск, при котором на каждом шаге в шаблон запроса вводятся дополнительные ограничения.

Уточнение запроса

Данная задача предполагает выявление неопределенности в запросе (Query ambiguity discovery) и непосредственно само уточнение запроса (Query refinement).

Таблица 8.3

Таблица сравнения систем поиска на основе онтологий

| Системы | Метод поиска | | Уточнение запроса | Формулирование запроса с помощью онтологии | Использование аннотированных ресурсов |
|------------------|-------------------------------------|--------------------|-------------------|--|---------------------------------------|
| | Улучшение поиска по ключевым словам | Поиск по онтологии | | | |
| Ontogator | + | | + | | + |
| OntoLoger | | + | + | + | |
| Knowledge Sifter | | + | + | | |
| OntoViews | + | + | + | + | + |
| OntoDoc | + | + | | + | |
| Ontobroker | | + | | + | + |
| SHOE | | + | | + | |
| SEAL | | + | | + | |
| Haystack | | + | + | + | |
| SemanticSearch | + | | | | + |

8.4. Формирование пакетов знаний

Представление части базы знаний соответствует распространенному в англоязычной литературе понятию «view» (например, ontology view, document-oriented views of knowledge bases). Синонимичными терминами для представления также могут выступать англоязычные segments, modules, extracts, partitions, islands, packages онтологии/знаний/базы знаний.

Идея извлечения фрагмента из онтологии или базы знаний не нова и представлена под разными названиями (представление/views, сегменты/segments, модули/modules, выдержка/extracts, раздел или сектор/partitions, «острова»/islands, пакеты или блоки/packages) в работах многих специалистов (Acker, 1994; Volz, 2003; Samson, 2005; Seidenberg, 2006; Dellschaft, 2006), см. рис. 8.8.

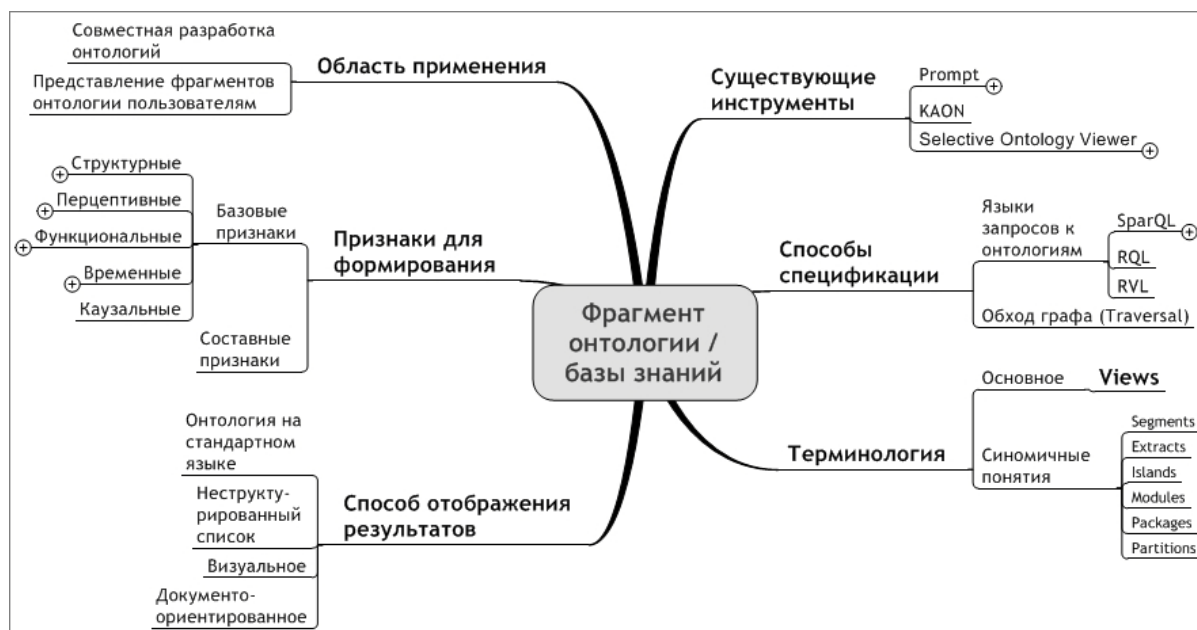


Рис. 8.8. Технология распределения знаний

Применение фрагментов онтологий / баз знаний можно обобщить в задачу кастомизации информации, адаптации содержания к потребностям и задачам пользователей.

Распределение знаний в данном случае соответствует задаче представления фрагментов онтологии или базы знаний. Существующие методы настройки (спецификации) содержания фрагментов онтологии / базы знаний можно разделить на 3 категории⁴⁸:

1. Методы, основанные на языках запросов,
2. Разделение сети на фрагменты,
3. Извлечение путем обхода графа (traversal).

⁴⁸ Seidenberg J. and Rector A. Web Ontology Segmentation: Analysis, Classification and Use // WWW 2006, May 23-26, Edinburgh, Scotland, 2006.

8.5. Использование контекста для повышения качества поиска

С самого своего возникновения передовые исследования в области управления знаниями были направлены на разработку методологий и систем, которые бы обеспечивали предоставление необходимой информации в нужном месте в нужное время нужному человеку. Однако в последние 10 лет, в связи с революционными изменениями в бизнес- и информационных системах, содержание понятий «нужный» и «необходимый» существенно преобразилось из четко-определенного описания в крайне контекстно-зависимое понятие. На самом деле, в связи с:

- Увеличением объемов информации, «нужность информации» зависит от предпочтений и потребностей потребителя знаний,
- Ростом мобильности рабочей среды (пространства), «нужное место» превращается в «бегущую мишень»,
- Увеличением скорости изменений в бизнесе и постоянной потребностью в инновациях, «нужное время» превратилось в постоянное проактивное предоставление информации,
- Огромными изменениями возможностей коммуникационных технологий, стимулирующих командную работу и групповое принятие решений, «нужный человек» может быть вытеснен (нужным/-ой) сообществом /группой.

Таким образом, потребность в предоставлении знаний в любое время, в любое место для любого сотрудника преобразует однобокий взгляд, оценивающий знания с точки зрения содержания, в многомерный, который включает в себя все факторы, характеризующие связь между элементом знаний и потребителем знаний в контексте окружающей действительности. Действительно релевантность знаний определяется не только их достоверностью, динамичностью и точностью, но и связью с другими элементами знаний, рабочей ситуацией, для которой они предназначаются, и когнитивным контекстом конкретного человека – потребителя знаний.

По сути, сейчас можно говорить о контекстно-ориентированном управлении знаниями, в котором разные контексты выступают в роли посредников между элементом знаний и сотрудником.

Можно сказать, что контекст, определяющий информационные потребности человека, характеризуется тремя основными факторами:

1. Решаемая задача, требующая знаний;
2. Позиция человека в организации;
3. Индивидуальные предпочтения.

Для описания ресурсов знаний и их поиска в корпоративной памяти целесообразно использовать три вида онтологий⁴⁹:

1. Онтологию информации для описания метасвойств и доступа к информации;
2. Онтологию предприятия для описания контекста создания и применения информации;
3. Предметную онтологию для описания контента (содержания) информации.

Общая схема реализации корпоративной памяти с применением данных онтологий представлена на рис. 8.9:

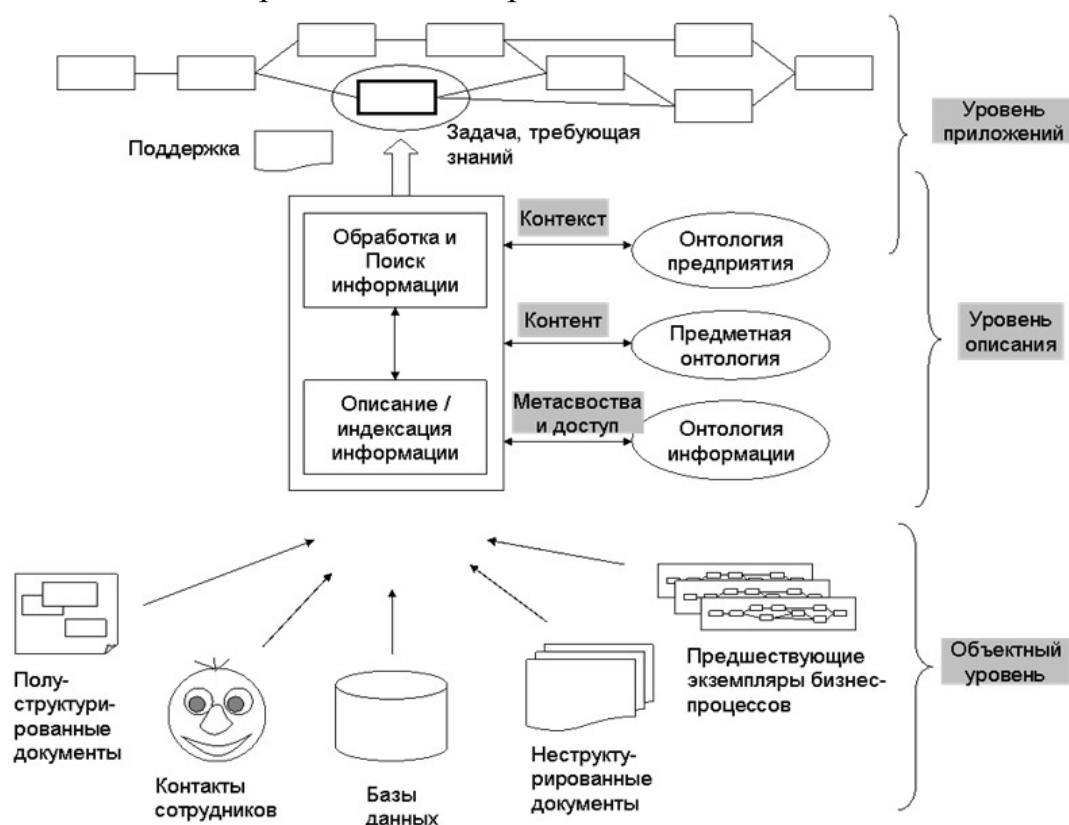


Рис. 8.9. Реализация корпоративной памяти на основе онтологий

⁴⁹ Abecker A., Bernardi A., Hinkelmann K., Kuhn O., Sintek M. Toward a Technology for Organizational Memories IEEE Intelligent Systems. – 1998. – № 3, 40-48.

В случае представленного выше подхода к обработке ресурсов знаний отдельный элемент знаний может быть описан с применением онтологий следующим образом:

Метасвойства и доступ

Название: «Как получить выгодные условия оплаты?», Автор: Фома Берлиоз, Жанр: Совет, Тип: Эвристика, на основе опыта, Форма: Текст на русск. языке, MS Word, Источник файла: E:\home\experiences\ds9-12-99rp.doc, Доступность: Всегда, Затраты: Нет.

Контент

Продукт: 20 SUN Ultra, Поставщик: Borg Inc., Персона для контакта: Д-р Лютер Кинг

Контекст

Процесс-создатель ИЛИ Проект-создатель: Проект Оптик для «Пятерочки», декабрь '99, Функция (activity)-создатель: Согласование цены с поставщиком комплектующих, Подразделение-создатель: Отдел закупок

Особое внимание при разработке поисковых механизмов в корпоративной памяти предлагается уделять именно контексту. Основная задача корпоративной памяти состоит в том, чтобы обнаруживать информационную потребность в течение выполнения производственного процесса и определять уместное знание в специфическом контексте задачи, поэтому авторы предлагают использовать информацию контекста задачи для информационной фильтрации⁵⁰ (рис. 8.10).

Моделирование и использование контекста для поиска информации в корпоративной памяти создает потенциал для применения результатов бизнес-моделирования.

При моделировании контекста рассматриваются две стороны:

1. Контекст потенциального применения знаний,
2. Контекст, в котором знание было создано.

⁵⁰ Oleg Rostanin; Harald Holz; T. Suzuki; K. Maeda; Andreas Dengel; K. Kanasaki Task-Based Process Know-How Reuse and Proactive Information Delivery in TaskNavigator. In: Proceedings CIKM. P. 521-533, 11/2006.

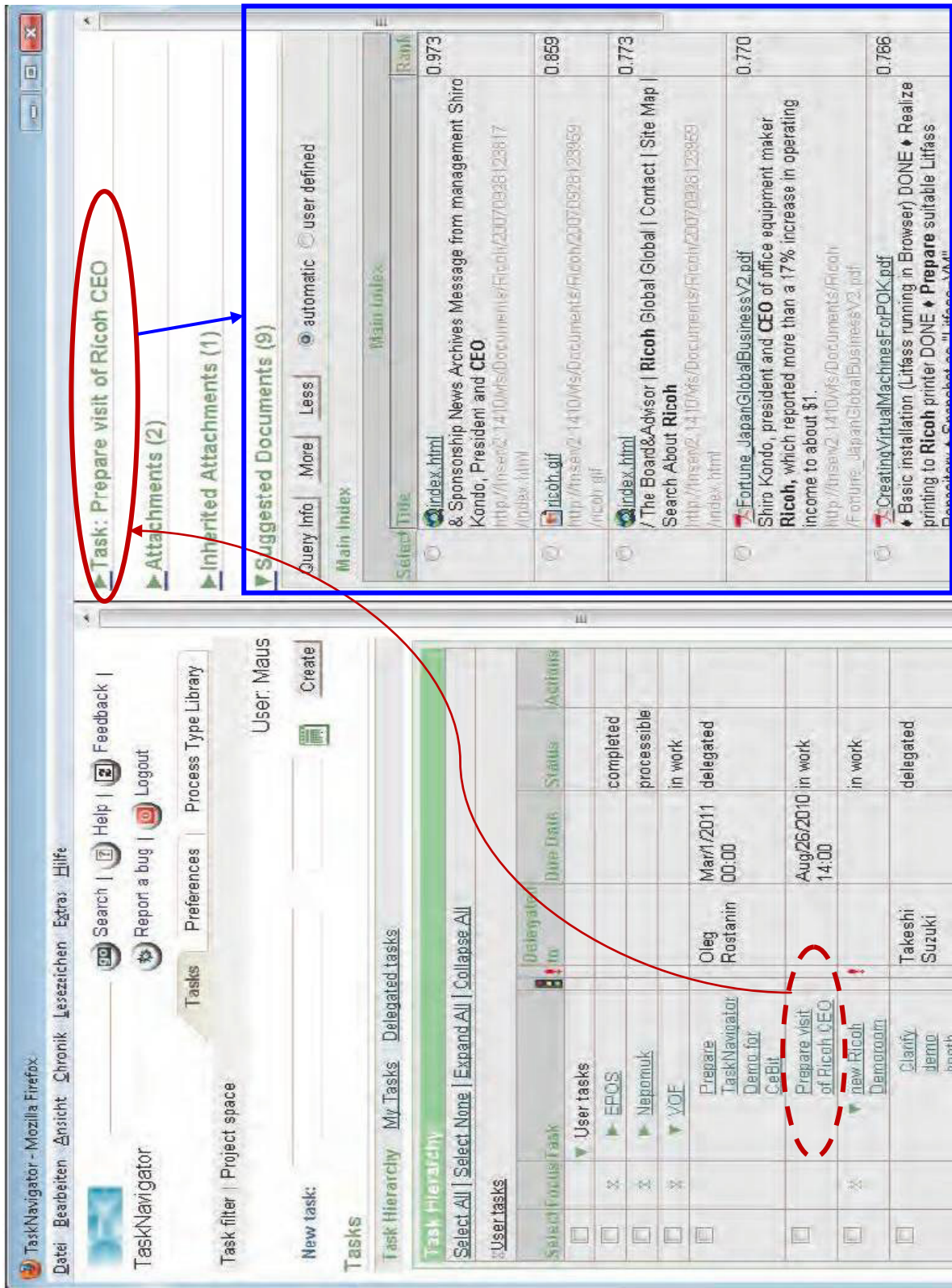


Рис. 8.10. Пример рекомендуемых документов (справа) для задачи «Prepare visit of Ricoh CEO» в Task Navigator

Если рекомендация (замечание) о каком-либо покупателе или поставщике была создана в определенном бизнес-контексте, например, согласование цены, то такая информация может быть очень важна при определении релевантности рекомендации (замечания) в новом контексте применения. Для некоторых видов информации, таких как отчеты о лучшей практике, полученные уроки или регламентирующая документация, контекст применения могут быть определены заранее. Контекст применения рекомендуется представлять в терминах организационной структуры, процессов, функций, которые, как раз, составляют основу онтологии организации и онтологической модели, которые создаются при моделировании.

8.6. Визуализация знаний на основе онтологий

Визуализация знаний в процессе их распределения может использоваться для повышения эффективности процессов поиска и доставки информации, а также для представления их результатов.

С учетом акцента на разработке семантических СУЗ, рассмотрим методы визуализации онтологий, детальный обзор которых был сделан в 2007 году в работе Катифори и коллег⁵¹, в рамках которой выделяется шесть категорий методов визуализации онтологий и рассматриваются их достоинства и недостатки:

Иерархический список (Indented List): многие инструменты визуализации онтологий, в том числе Protege, OntoEdit, Каон и некоторые другие, предлагают в качестве основного способа визуализации представления в стиле «Проводника» в Windows.

Узлы-связи и деревья (Node-link and Tree): методы этой категории отображают онтологию как набор соединяющихся узлов, представляющих иерархию классов в виде вертикально- или горизонтально-размещенного дерева. Пользователю обычно предоставляется возможность разворачивать и сворачивать узлы и поддеревья для управления детализацией информации. Двухмерная реализация данного метода осуществляется в OntoViz, IzaVis (рис. 8.11), SpaceTree, трехмерная в OntoSphere (рис. 8.12).

⁵¹ Katifori A., Halatsis C., Lepouras G., Vassilakis C., Giannopoulou E. Ontology visualization methods – a survey. ACM Comput. Surv. 39(4), 2007.

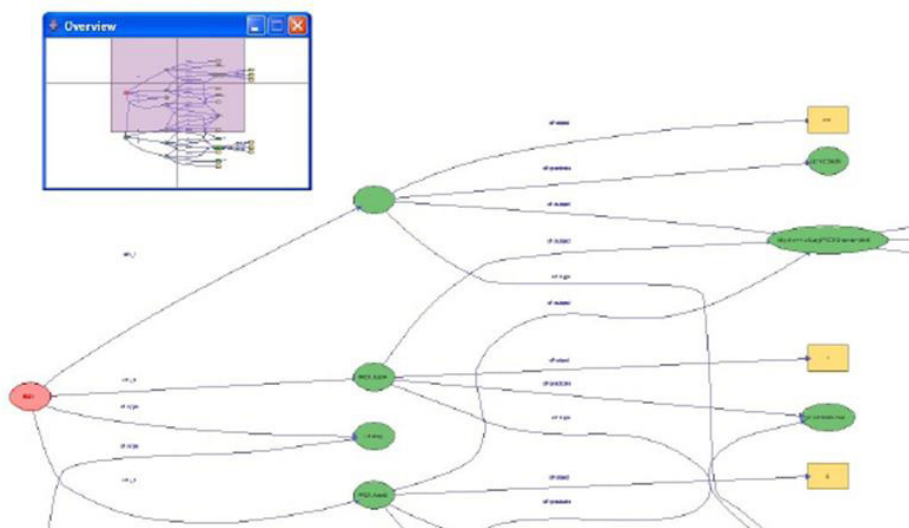


Рис. 8.11. Двухмерная визуализация с помощью инструмента IzaVis

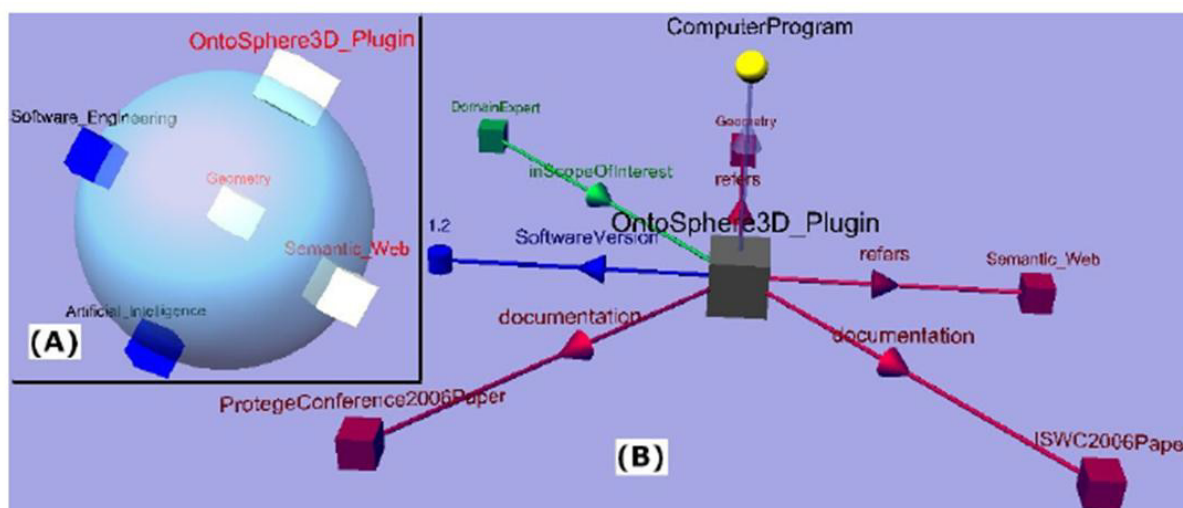


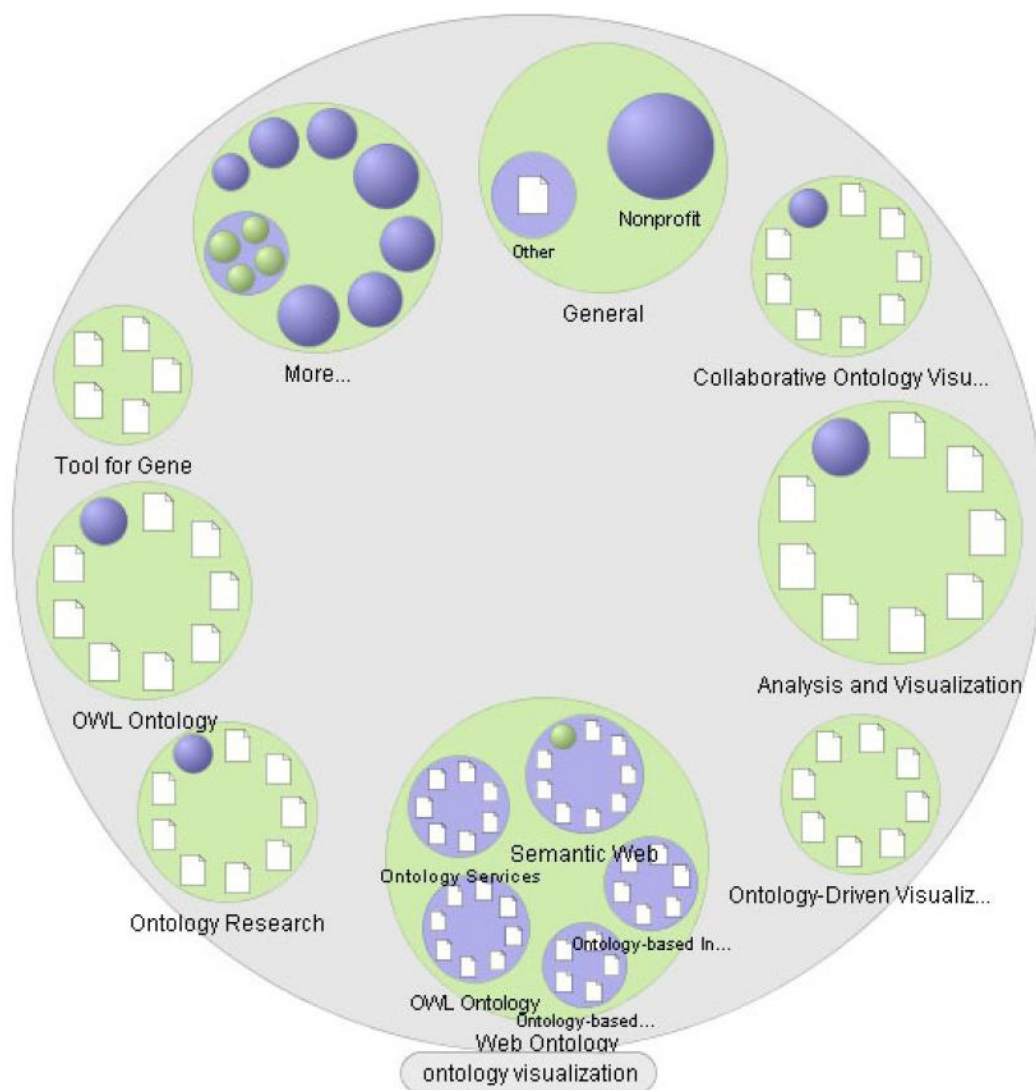
Рис. 8.12. Визуализация с помощью OntoSphere (2 возможных вида А и В)

Масштабируемые (Zoomable): методы данной категории отражают иерархию классов и экземпляры через визуальную вложенность объектов (диаграммы Вена, «меньшие коробочки размещаются в больших»). При выборе объекта он увеличивается в размере (*zoom-in*), позволяя более детально рассмотреть свои компоненты (подклассы и экземпляры). Примеры систем и соответствующих методов: Grokker (рис. 8.13), Jambalaya, CropCircles.

Заполнение пространства (Space-filling): идея данной группы методов заключается в использовании всего пространства экрана путем деления пространства каждого узла между его потомками в

зависимости от свойств узла, например, количества потомков и т. п. Примеры: TreeMaps, Information Slices.

Фокусирование+контекст (Focus + Context): данная группа техник основана на понятии искажения (distorting) графа с целью фокусировки на определенном узле и представления его контекста (окружении). Фокусный узел отображается в центре, а связанные с ним элементы располагаются вокруг него и их размер уменьшается по мере удаления от фокуса. Для увеличения какого-либо узла пользователь должен сделать его фокусным.



Примечание: кружки – классы, «белые документы» – найденные ресурсы.

Рис. 8.13. Визуализация результатов поиска по запросу «ontology visualization» (визуализация онтологий)

Примеры: двухмерные Touchgraph и TGVizTab (рис. 8.14), 3D Hyperbolyc Tree (рис. 8.15).

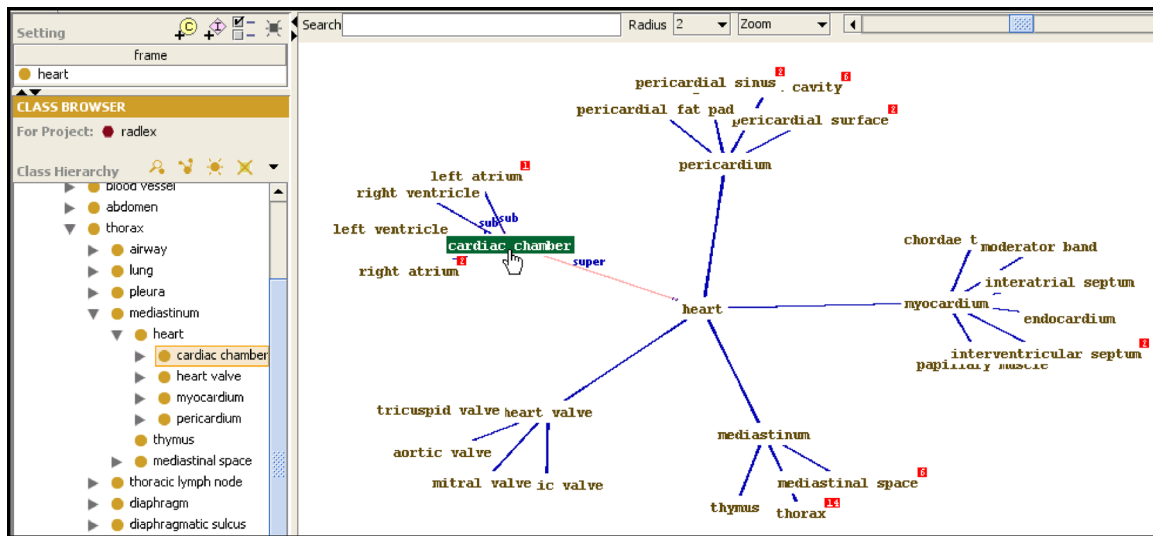


Рис. 8.14. Пример визуализации взаимосвязей понятий в TGVizTab

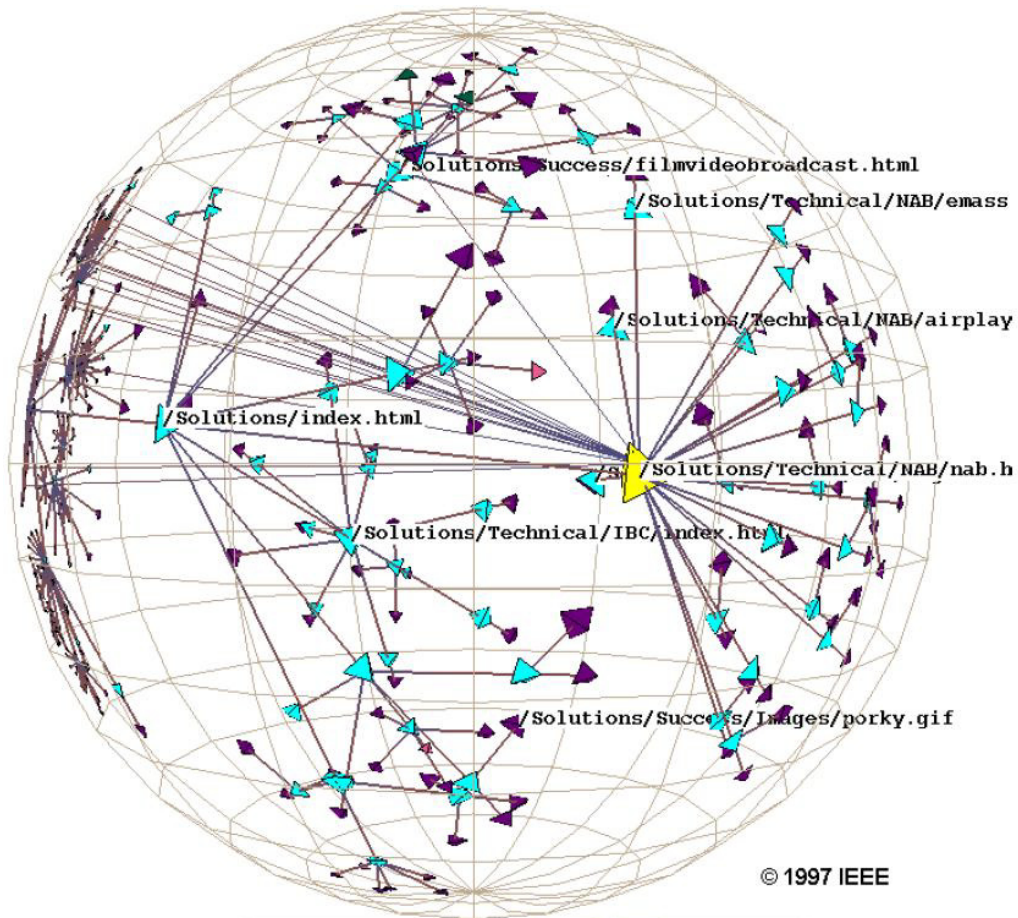


Рис. 8.15. Пример 3d-визуализации взаимосвязей между ресурсами с помощью гиперграфа

Трёхмерные информационные ландшафты (3D Information Landscapes).

По результатам обзора можно сделать заключение, что в настоящий момент отсутствует визуальный язык или метод представления онтологии, который может быть использован как на этапе разработки онтологии, так и на этапе ее наполнения, который обеспечит также машинную обработку онтологий и позволит работать с масштабными онтологиями (более 1000 элементов).

8.7. Настройка прав доступа

В контексте управления знаниями на предприятиях распределение знаний предполагает настройку прав доступа сотрудников к информации (табл. 8.4).

Таблица 8.4

Пример настройки прав доступа

| <u>Рубрики</u> | Генеральный директор | Ведущий консультант | Консультант | Консультант-стажер | Служба маркетинга и сбыта |
|-----------------------|----------------------|---------------------|-------------|--------------------|---------------------------|
| Маркетинг и сбыт | + | + | - | - | + |
| Проекты | + | +/- | +/- | - | +/- |
| Программные решения | + | + | +/- | - | + |
| Семинары | + | + | +/- | - | + |
| Методология | + | + | +/- | +/- | +/- |
| Библиотека | + | + | + | + | + |
| Компания и управление | + | +/- | +/- | +/- | +/- |
| Целевые пакеты знаний | + | +/- | +/- | +/- | +/- |

+ есть доступ – нет доступа +/- доступ уточняется при развертывании рубрик

9. ОБМЕН НЕЯВНЫМИ ЗНАНИЯМИ И ПОДДЕРЖКА СЕТЕЙ ЗНАНИЙ

9.1. Средства совместной работы

Наиболее типичный способ, с помощью которого формируется и распространяется неявное знание, – личные встречи и обмен опытом, часто неформальные. При этом информационные технологии играют минимальную роль. Тем не менее, увеличивается доля встреч и других межличностных взаимодействий, для проведения которых используют on-line-средства, известные как средства совместной работы или программное обеспечение коллективного пользования (англ. collaborative software, groupware, workgroup support systems, group support systems). Эти программные продукты поддерживают коммуникации (мгновенные сообщения, почту, IP-телефонию, видео-конференции) и предоставляют единую виртуальную рабочую среду, совместную работу с документами, включающую wiki, рабочие доски и пр. (табл. 9.1). Они используются для того, чтобы либо дополнить традиционные встречи, либо в некоторых случаях заменить их. Они могут упростить формулирование и преобразование неявных знаний.

Таблица 9.1

**Классификация средств совместной работы
по характеру взаимодействия**

| | Одно время (синхронное) | Разное время (асинхронное) |
|---------------------|---|---|
| Одно место | Комнаты для совещаний, для занятий... | Доски объявлений, электронные «классные доски»... |
| Разные места | Чаты, аудио и видео конференции, on-line игры | E-mail, форумы, распределенные документы, wiki |

Богатые возможности обмена опытом предоставляют приложения, которые поддерживают **on-line-встречи в реальном времени**. Это

наиболее последние разработки в группе средств Groupware. On-line-встречи могут включать конференции, основанные на видео и на тексте, а также синхронные коммуникации и чат. Однако не все системы организации on-line-встреч дают возможности тесного общения. В настоящее время вместо замещения непосредственных встреч многие системы проведения on-line-встреч дополняют существующие системы совместной работы и телефонные конференции.

Чат создает возможности для проведения диалога, при котором может формироваться неявное знание и преобразовываться в явное. Основанный на тексте чат считается способным поддерживать обмен знаниями между людьми в разговорной форме. Мгновенный обмен сообщениями обладает свойствами, характерными для личных встреч и телефонного разговора.

Системы проведения on-line-встреч расширяют телефонные конференции, предоставляя участникам возможность наблюдать и использовать один и тот же экран открытого приложения, а также возможность проведения встреч в виртуальной реальности (например, Second Life). Одно из направлений исследований состоит в интеграции on-line-встреч с классическими приложениями совместной работы, которые поддерживают совместное использование документов и асинхронные конференции: результаты on-line-встреч могут фиксироваться в виде документов и представляться на непосредственной встрече.

Сетевые конференции и форумы открыты для всех и имеют некоторые характеристики типичных групповых обсуждений: так же можно поставить вопрос и получить на него ответ. Хотя участники могут быть незнакомы, тем не менее, как показывают наблюдения, многие люди, участвующие в сетевых конференциях, склонны оказывать поддержку и давать советы. Архив конференции и форума становится репозиторием полезных знаний, и люди, озадаченные схожими проблемами, могут найти решение, используя при этом архив. Архивы конференций и форумов можно рассматривать также и как одно из средств создания формализованных знаний.

Средства для совместной работы с документами или вики – важная часть средств совместной работы. Обычного Word-а или Excel-я недостаточно, когда речь идет о коллективной работе. Система

рецензирования, хоть и удобна и позволяет вносить правки разными людьми, но не дает работать с документом одновременно нескольким пользователям. При отсутствии локальной сети и централизованного сервера (например, в распределенных командах, объединенных в рамках общего проекта), на котором будет лежать редактируемый файл, необходимо дополнительно пользоваться почтой для пересылки документа. Это приводит к путанице с версиями документов. Для решения данной проблемы используются средства он-лайн работы с документами (например, Google Docs, Zoho Writer) или вики (рис. 9.1). Средства совместной работы с документами и вики позволяют создавать контент, поэтому также должны рассматриваться в контексте создания формализованных знаний (см. раздел 6).

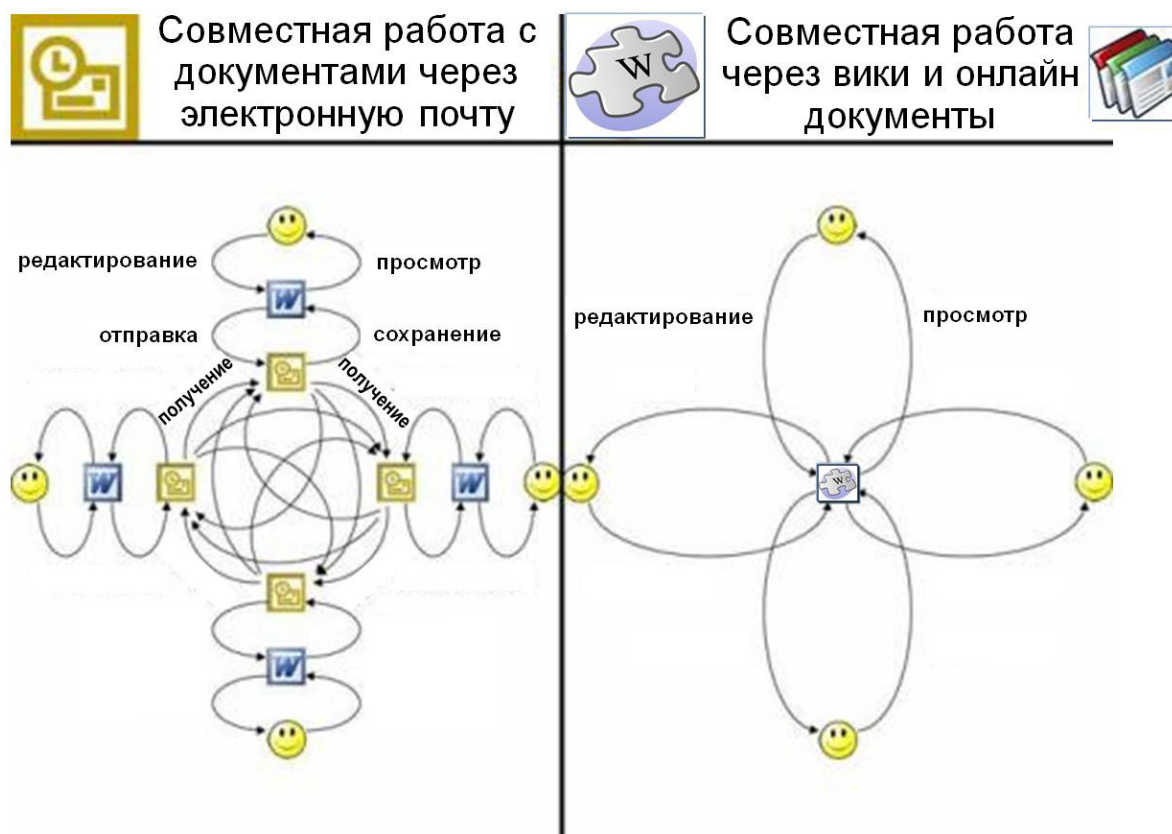


Рис. 9.1. Сравнение способов совместной работы с информацией

Другой подход к обмену неявным знанием связан с системами, которые находят людей с общими интересами. Для этого может использоваться анализ сходства используемых ими документов.

Обнаружение людей с общими интересами – это функция, которую можно добавить к системам персонализации, цель которых – направить входящую информацию заинтересованным в ней индивидам.

Системы обнаружения экспертов предназначены для поиска специалиста по определенному вопросу. Они предлагают список имен людей, которые имеют знания в определенной области. В самой простой форме такие системы представляют собой поисковые машины, но их качество зависит от характера фактов, исходя из которых они делают заключение об области знаний эксперта. Некоторые возможные источники таких фактов следующие:

- профиль или форма, заполняемые пользователем;
- существующая в компании база данных, например в отделе по человеческим ресурсам;
- ассоциации, связанные с названиями документов;
- ответы на вопросы, которые уже давал специалист.

Проблема, связанная с использованием явного профиля, состоит в том, что у пользователей может отсутствовать стимул поддерживать его в актуальном состоянии, поскольку это требует периодического заполнения новой формы. Таким образом, предпочтительнее собирать информацию автоматически, если возможно, из существующих источников. Например, в базе данных компании может храниться резюме или список проектов, в которых участвовал специалист. Другой автоматический подход состоит в том, чтобы сделать заключение об области специализации, исходя из содержания документов, с которыми ассоциируется имя пользователя.

Например, авторство документа предположительно отражает знакомство с содержащимися в нем вопросами, в то время как чтение его пользователем отражает интерес к данной проблеме. Таким образом выделяются два подхода к использованию документальных фактов с целью обнаружения эксперта: либо документы классифицируются согласно некоторой схеме, таким образом классифицируя их авторов; либо, когда пользователь формирует запрос к системе обнаружения экспертов, поиск осуществляется по документам, при этом запрос преобразуется в список авторов (с соответствующим «весом»), который предоставляется пользователю.

Существующее состояние данной области предполагает

использование трех источников фактов: явные профили (рис. 9.2); факты, содержащиеся в существующих базах данных; факты, выведенные из ассоциации между людьми и документами.

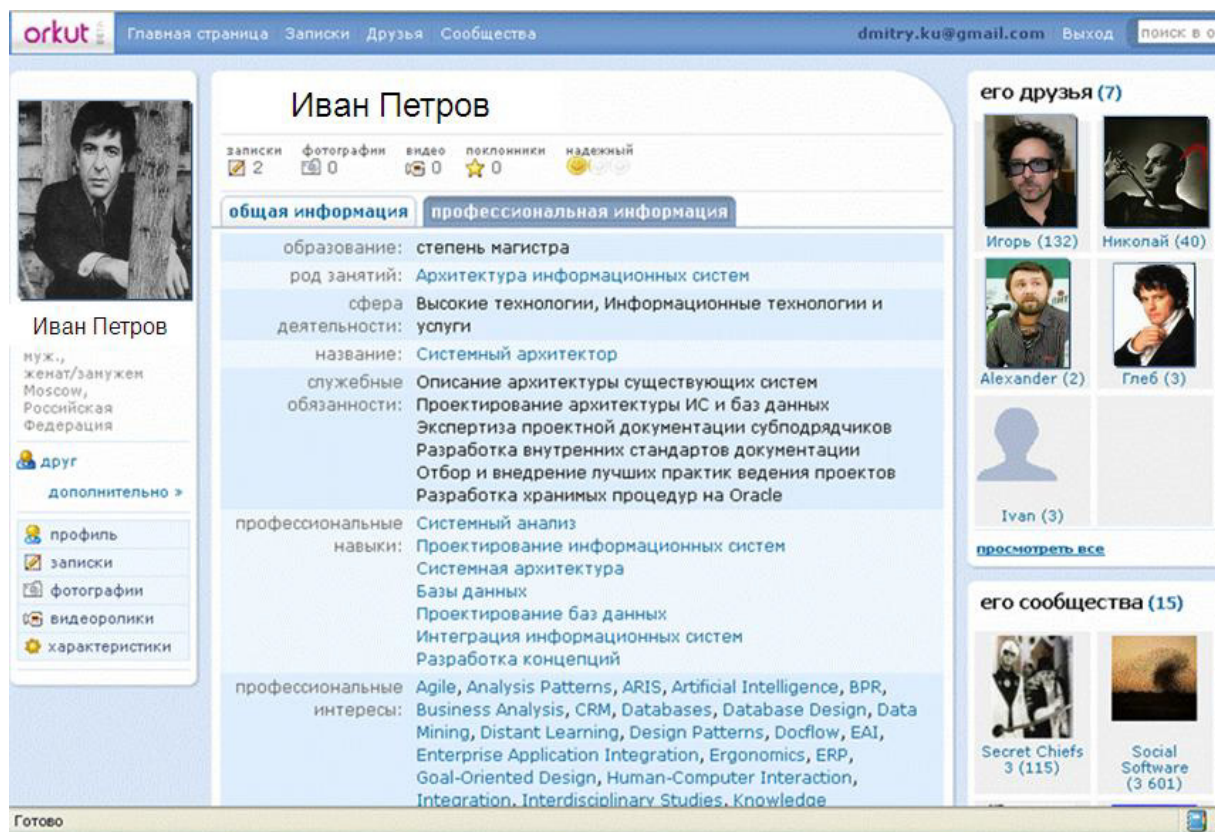


Рис. 9.2. Пример профиля эксперта в «профессиональной» социальной сети

Использование онтологий для управления компетенциями

В системах управления компетенциями распространено использование онтологий, которые формируют дерево компетенций организации и позволяют связывать элементы данного дерева с сотрудниками. Остальные функциональные возможности варьируются от приложения к приложению. Базовые возможности применения онтологий в данной области можно рассмотреть на *прототипе системы Управления компетенциями в компании Swiss Life (УК-SL), который был разработан в 2002 году.*

Swiss Life Group – одна из крупнейших страховых компаний Швейцарии. Компетенции сотрудников – это наиболее ценный ресурс

компании для решения интеллектуальных задач, включающие в себя неформализованные знания, умения, навыки. Компании Swiss Life был необходим репозиторий компетенций и электронный доступ к нему.

Задачами данной системы управления компетенциями были: поиск людей с необходимыми знаниями, анализ разрывов в компетенциях, развитие и обучение персонала и, в конечном итоге, рост интеллектуального капитала компании.

Преимущества публикации знаний и умений: легче находить, что человек знает и умеет; проще находить экспертов, людей, способных помочь; улучшение продуктивности организации; упрощение назначения сотрудников на проект; возможность планирования профессионального развития.

Разработанная система сочетала онтологический поиск с классическим текстовым поиском.

При создании персональной страницы каждого сотрудника компетенции выбираются из онтологии. При описании компетенций сотрудника указывается уровень владения: базовые знания, опыт практического применения, глубокие знания, ведущий специалист. Дополнительно на странице сотрудника указывается: образование, участие в текущих проектах, участие в предыдущих проектах, интересы.

Основные элементы архитектуры УК-SL (Рис. 9.3):

- *Редактор онтологий*, включающий в себя *OntoEdit* – основной инструмент для создания и изменения онтологии, а также *Mind Manager*, который удобен при мозговом штурме;
- *Sesame* – хранилище RDF/RDFS, Онтология компетенций в RDFS, Связи компетенций и сотрудников – RDF facts, RDF Annotations
- *QuizRDF*: Онтологический и текстовый поиск.

Онтология в УК-SL включала компетенции (700 концептов) в области ИТ, страхования физлиц, управления персоналом, рабочие обязанности (180 единиц), образования (130 единиц). Личный опыт сотрудников сложно структурировать – использовались текстовые поля. Онтология навыков не может быть полной, поэтому пользователи предлагают дополнения, а администратор их утверждает.

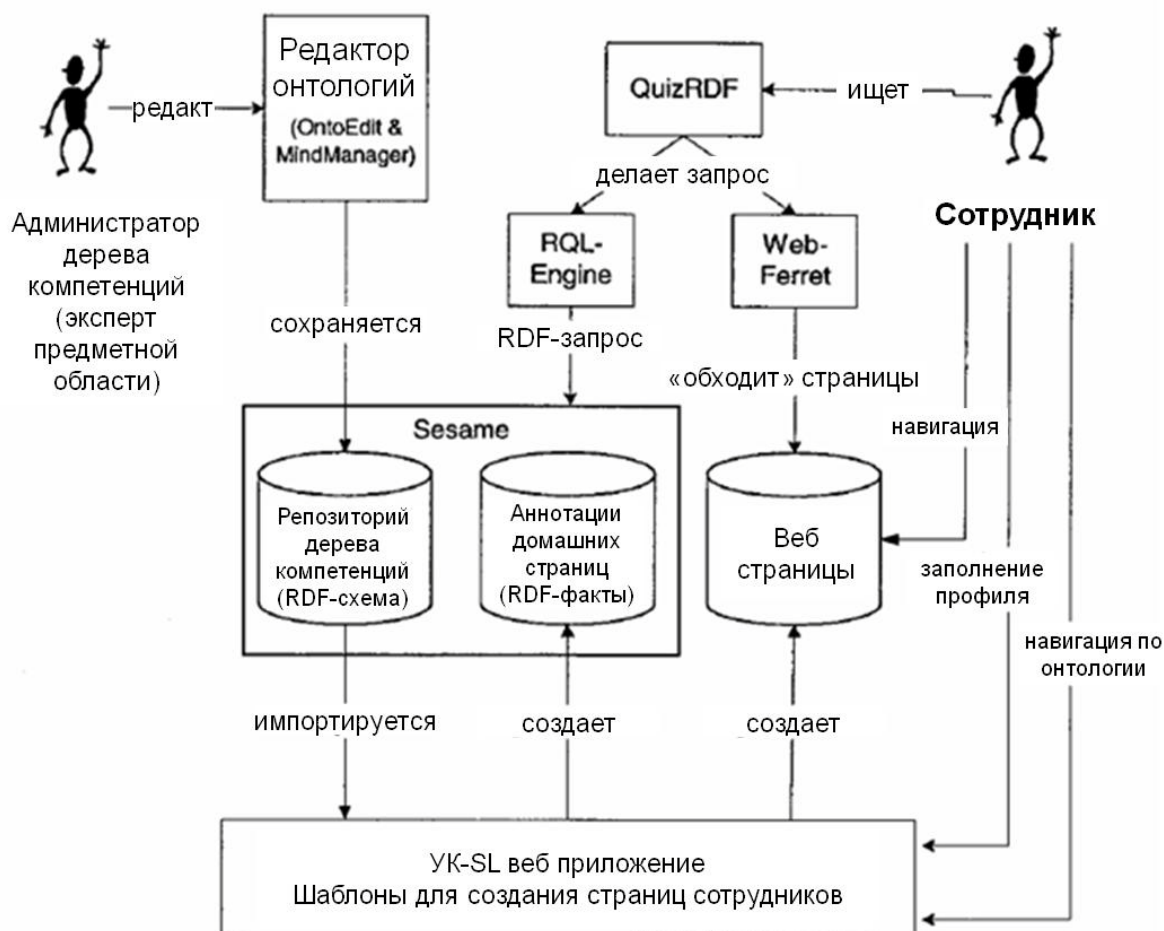


Рис. 9.3. Архитектура УК-SL

Более подробно применение онтологий для управления компетенциями рассмотрено в следующих статьях, доступных в Интернете:

Моделирование способностей и управление знаниями (Capability Modelling and Knowledge Management, Stader J., Macintosh A., Artificial Intelligence Applications Institute, Division of Informatics, The University of Edinburgh, Edinburgh, Scotland).

Аннотация: Организации осознают, что очень важно «знать кто что знает» и максимально использовать такие знания. Управление знаниями решает подобные задачи. Институт применения искусственного интеллекта (ИПИИ) вовлечен в деятельность по управлению знаниями. ИПИИ также работает с онтологиями и, в особенности, с онтологиями

навыков и компетенций в контексте workflow систем, и сейчас применяет техники управления знаниями для более эффективного использования онтологии навыков и компетенций в управлении организацией. Программный продукт для поддержки управления компетенциями и основанный на онтологии навыков и компетенций может помочь организациям согласовать навыки и компетенции существующих и будущих сотрудников с целями бизнеса.

Интегрированные процессы и инструменты для развития персонала (Integrated Processes and Tools for Personnel Development /Ernst Biesalski, Andreas Abecker).

Аннотация: Одним из основных компонентов такого управления является создание Каталога компетенций – карты, устанавливающей связь между сотрудниками организации и их компетенциями, которая позволяет пользователю находить людей, имеющих требуемые знания, навыки и умения. Кроме создания каталога, компетенции необходимо интегрировать в деятельность организации и связать с процессами, проектами, функциями, подразделениями. Такая интеграция позволит определять необходимые организации компетенции, а также идентифицировать / находить компетенции на основе контекста их применения (решаемые задачи, процессы, проекты...). Преимущества управления компетенциями на основе онтологий: Легкость определения требуемых компетенций благодаря анализу уже имеющегося списка; возможность сопоставления компетенций сотрудников из разных отделов / структурных единиц; возможности дополнительного анализа в разных разрезах (карта компетенций/знаний отделов, наглядный и удобный для сопоставления анализ развития персонала); высокая точность подбора сотрудников под необходимые задачи (проекты); ускорение и повышение эффективности обмена опытом.

Для эффективного обмена неформализованными знаниями недостаточно только предоставить всем сотрудникам доступ к средствам совместной работы и указать на носителей знаний. Сотрудники еще должны знать друг друга и доверять друг другу. Для обеспечения данных условий необходимо анализировать и координировать сети знаний (см. раздел 4.4).

9.2. Сети знаний индивидуального уровня

В условиях глобальной автоматизации интеллектуальный багаж и креативность участников бизнес-процессов становятся залогом конкурентоспособности компаний. Однако в формализованном представлении знаний не учитываются столь важные компоненты интеллектуального капитала организаций, как неявные знания сотрудников и интеллектуальные связи, сформировавшиеся в процессе их общения. Взаимосвязи в обществе анализируются в дисциплине социальных сетей, возникшей намного раньше технологий компьютерных сетей, но приобретающей в век компьютеризации особое значение. В основе данного раздела лежит статья Дубовой Н. «Социальная сеть знаний», опубликованная в журнале «Открытые системы» в 2005 году.

В энциклопедии Wikipedia социальная сеть определяется как социальная структура, объединяющая отдельных людей или даже целые организации. Социальная сеть показывает, каким образом ее участники связаны друг с другом теми или иными отношениями – от случайных знакомств до тесных семейных связей. Сам термин введен в 1954 году английским антропологом Джоном Барнсом. Он развил изобретенный еще в 30-е годы в Америке подход к исследованию взаимосвязей между людьми с помощью социограмм, то есть визуальных диаграмм, в которых отдельные лица представлены в виде точек, а связи между ними – в виде линий. К 70-м годам окончательно сформировался комплекс социологических и математических методов исследований, которые составляют научный фундамент современного анализа социальных сетей (social network analysis, SNA).

Предметом анализа социальных сетей, в отличие от большинства традиционных социологических исследований, являются не атрибуты отдельных личностей (умный или глупый, дружелюбный или агрессивный и т. д.), а структура их взаимосвязей в рамках того или иного сообщества или рабочей группы. В социальной сети, где отношения между людьми представлены с помощью узлов и линий с определенным направлением (рис. 9.4), могут исследоваться отношения родства, социальные роли, привычные совместные действия, проявления определенных чувств, материальные отношения, сходство в поведении или вкусе и т. д.

Объектами исследования обычно становятся разномасштабные группы людей (от 25 до 150 человек), которые могут образовывать отдельные группы (например, распределенную команду разработки программного обеспечения, группу специалистов по продажам, сообщество с едиными интересами, бизнес-подразделение компании или целую организацию).

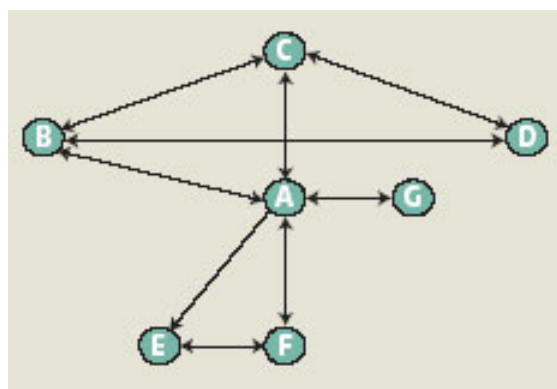


Рис. 9.4. Диаграмма социальной сети

Оценить значимость происходящего в социальной сети для достижения определенных целей, например для повышения эффективности работы сотрудников, помогает категоризация связей между узлами сети. Такие связи могут быть прямыми или косвенными, сильными или слабыми, односторонними или двусторонними. Прямая связь в социальной сети – это непосредственные отношения между двумя людьми.

В сети могут фигурировать персоны с большим количеством прямых связей (например, А на рис. 9.4). Очевидно, что они играют важную роль в структуре взаимоотношений данной группы, часто демонстрируют большую продуктивность работы и большую удовлетворенность ею, чем участники сети с меньшим числом прямых связей. Эти персоны выполняют функции основных посредников между людьми, соединенными косвенными, непрямыми связями. Такое посредничество может оказывать как положительное, так и отрицательное влияние на взаимоотношения в социальной сети, помогая или, наоборот, препятствуя распространению информации либо налаживанию отношений между различными подгруппами сети.

Сильные связи характеризуются высокой частотой и разнообразием взаимодействий между людьми, формируют ощущение близости между

ними. Поэтому они могут сопровождаться эмоциональной поддержкой и постоянным обменом информацией, имеющей отношение к общей работе или общим интересам. Однако источником новых идей оказываются не сильные, а слабые связи, обычно объединяющие людей из разных подгрупп одной сети или разных социальных сетей.

Близкие друзья или исследователи, работающие в тесном взаимодействии над одной проблемой, постоянно обмениваются информацией. Но источником озарения и открытий для них, скорее всего, будут данные, поступающие извне (скажем, полученные при случайном обсуждении изучаемой проблемы с исследователями, работающими в той же области в другой организации). Для бизнеса важно соблюдать баланс между слабыми и сильными связями в социальных сетях – первые дают дополнительные источники инноваций, а вторые помогают создавать и поддерживать необходимый рабочий микроклимат.

На рис. 9.4 стрелками показаны односторонние и двусторонние связи между участниками социальной сети. Новичок в рабочей группе будет постоянно обращаться за советом и помощью к опытным сотрудникам, не ожидая аналогичных обращений к нему. Это пример односторонней связи. Наличие двусторонних взаимодействий часто помогает повысить эффективность работы группы. Так, эксперт, который отвечает за принятие решений, не только является источником информации для остальных членов группы, но и сам черпает нужные ему сведения, обращаясь к коллегам.

Исследование социальных сетей в бизнесе (объединяющих рабочую группу предприятия, организацию в целом, несколько компаний из одной отрасли, заказчиков и поставщиков и т. д.) открывает резервы повышения эффективности совместной работы, позволяет лучше использовать таланты и знания сотрудников, определять связанные с персоналом узкие места при реализации тех или иных решений, внедрении новаторских подходов и технологий.

Социальные сети могут показывать так же, как компании взаимодействуют между собой, обеспечивать выявление неформальных связей между руководящими кадрами и рядовыми сотрудниками. Такие данные могут оказаться полезными, например, при решении проблем конкуренции или налаживании более тесных связей между поставщиками

и клиентами.

Структуру организаций принято представлять с помощью иерархических схем, отображающих отношения «начальник–подчиненный» в разных подразделениях. Однако использование диаграмм социальных сетей может оказаться более ценным источником информации о реальном положении дел в компании. Например, исследователи из IBM провели анализ коллектива крупной нефтяной компании, причем в одном из ее подразделений анализ социальной сети сотрудников был соотнесен с имевшимся в ней иерархическим представлением о структуре организации (рис. 9.5). Выяснилось, что один из менеджеров среднего звена (Коул), занимающий в иерархии этого департамента далеко не главную позицию, фактически играет центральную роль во взаимосвязях между членами разных рабочих групп. В то же время формальный лидер подразделения (Джонс) оказался на периферии этих отношений. Такой анализ позволяет принять соответствующие управленческие решения, переместив на формальные руководящие позиции фактических лидеров рабочего процесса.



Рис. 9.5. Организационная структура и социальная сеть компании

Для идентификации социальных сетей с целью последующего анализа используются, как правило, интервью, опросы, специальные методы наблюдений. С развитием Web интерес к анализу социальных сетей возрос – появились новые типы сообществ, объединяющих с помощью Сети людей в самых разных уголках земного шара. За последние годы возникло немало сайтов, которые поддерживают виртуальные сообщества (Facebook, ВКонтакте, МойКруг, LinkedIn и т. п.).

Развиваются средства для поддержки социальных сетей в организации. Компания IBM, предлагает продукт IBM Lotus Connections.

IBM Lotus Connections – это социальное программное обеспечение, позволяющее использовать коллективные знания организации, партнеров и заказчиков, динамически создавая новые связи между пользователями, информацией, которой они обладают, и выполняемыми задачами. Это часть общей концепции IBM по использованию принципов Web 2.0 для реализации социально-ориентированных технологий.

Компоненты Lotus Connections: Профайлы (Profiles), Блоги (Blogs), Сообщества (Communities), Обмен закладками (Dogear), Операции (Activities), Домашняя страница (The homepage).

Доступ к компонентам осуществляется с помощью Web-браузера или других приложений, таких как IBM Lotus Notes, IBM Lotus Sametime, IBM WebSphere Portal и Microsoft Office.

Профайлы (Profiles) – предоставляет адресный справочник для нахождения экспертов по различным вопросам, тем самым, обеспечивая пользователям возможность быстро находить людей, которые могут помочь в решении возникших проблем или в реализации проектов.

- Поиск пользователей в масштабах организации по имени, номеру телефона, местоположению или должности, а также просмотр данных визитной карточки,
- Поиск пользователей по сферам деятельности с помощью ключевых слов в профилях, отражающих их знания и обязанности,
- Отправка мгновенных сообщений одним щелчком по контактной информации пользователя,
- Отображение в профиле пользователя общих закладок, сообществ, в которые входит пользователь, сообщений в блогах и общих задач.

- Поля профиля можно заполнить данными из внутренних систем с помощью IBM Tivoli Directory Integrator.

Развиваются и специальные программные средства визуализации и анализа социальных сетей. Среди них можно упомянуть инструментарий Social Network Fragments, разработанный MIT Media Lab, систему InFlow, которую продвигает компания Orgnet.com, ArnetMiner (рис. 9.6) и Ontos (рис. 9.7).

Среди наиболее перспективных направлений анализа социальных сетей для УЗ является анализ информационных потоков организации, определение способов обмена знаниями между совместно работающими людьми как части процессов управления знаниями. Кто именно является в компании источником знаний по определенным проблемам, с кем и по каким причинам сотрудники постоянно делятся информацией, к кому чаще всего обращаются за советом?

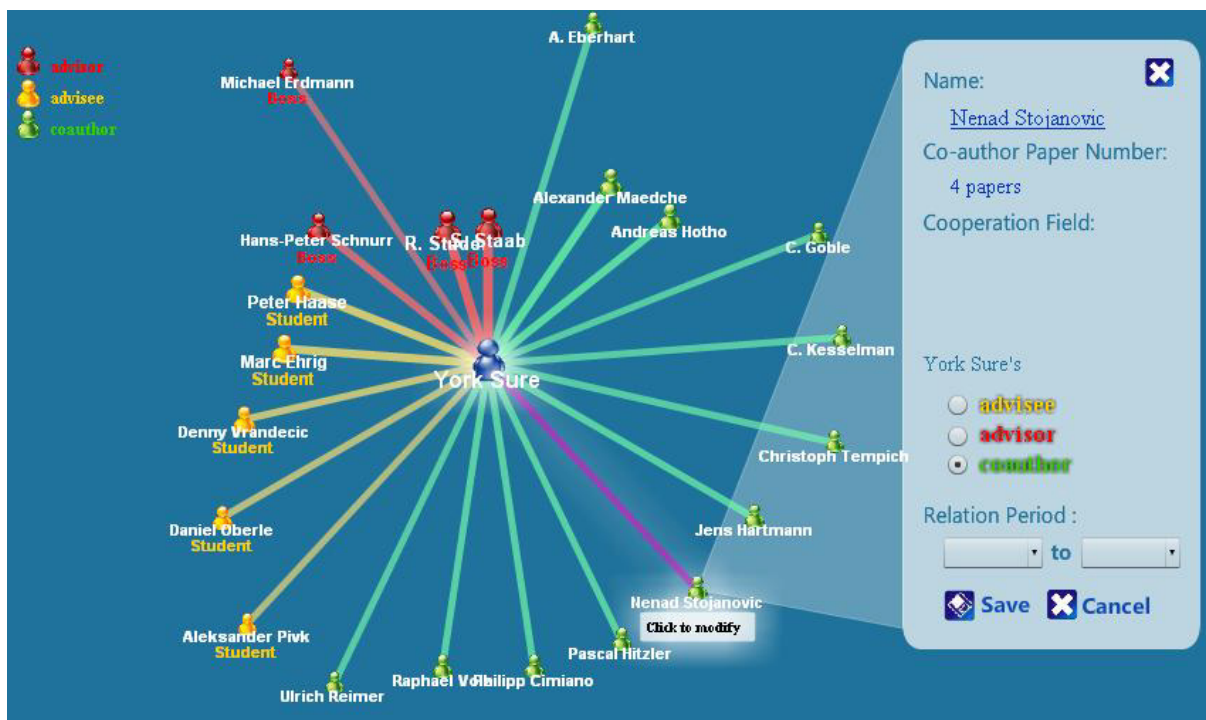


Рис. 9.6. Система Arnetminer для анализа академического сообщества

Решения Ontos для Semantic Web.

Обработка текстов под управлением онтологий

Предметная онтология «Скрытые коллективы»

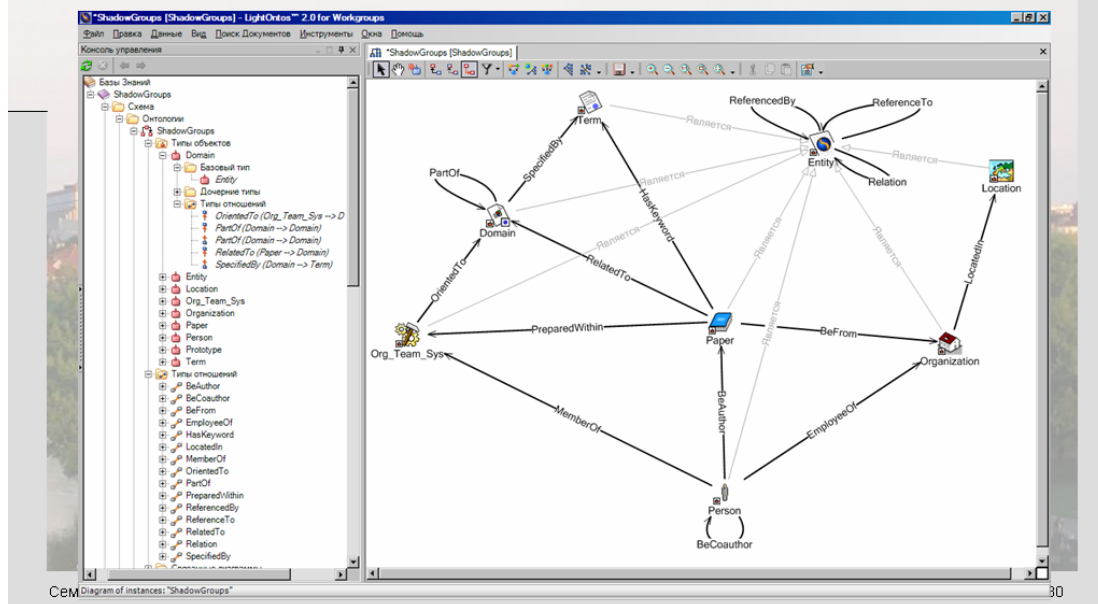


Рис. 9.7. Применение системы Ontos для выявления и анализа научных связей

9.3. Сообщества практиков как пример сети знаний

Сообщество практиков – группа людей, находящихся в неформальных связях, которые разделяют желание или стремление к чему-то, что они делают; по мере регулярного взаимодействия они учатся делать это лучше". Венгер⁵², который предложил данное понятие, описывает три важных компонента для создания сообществ практиков.

Предметная область (домен). Сообщество практиков – это не клуб друзей или социальная сеть. Оно четко определяется общей сферой интересов. Участие в сообществе подразумевает приверженность определенному домену и, следовательно, совместную компетенцию, которая отличает членов сообщества от других людей. Это может быть какое-то техническое знание, например, ремонт оборудования, некая профессиональная дисциплина, например, инженерные специальности, или же какая-то тема, например, креативность. Что бы ни объединяло

⁵² Wenger E. Communities of practice learning as a social system // The Systems Thinker, Vol. 9, No. 5. 1998.

сообщество, оно должно обладать фокусной точкой и должно быть достаточно хорошо определено, чтобы люди могли идентифицировать себя с ним.

Сообщество. Любое сообщество – это сложнейшая сеть личных взаимоотношений. Люди в сообществе знают друг друга и доверяют друг другу. Вместе они делают разные вещи. Некоторые из них по своей природе направлены на социальные отношения, другие представляют собой трудовую деятельность, Преследуя свой интерес в определенной области, члены сообщества участвуют в совместной деятельности и дискуссиях, помогают друг другу и обмениваются информацией. Они налаживают связи, которые позволяют им учиться друг у друга.

Практика. Участники сообщества практиков являются практическими специалистами. Когда сообщество действует в своей области, его члены создают инструменты, документы, процессы, общий словарь и делятся друг с другом методами выполнения работ. Многие сообщества решают проблемы, возникающие в повседневной работе. Многие разрабатывают и документируют лучшие практики, лучшие методы выполнения той или иной задачи.

Сообщество практиков одна из наиболее эффективных форм обмена знаниями в СУЗ как внутри отдельной организации, так и в сети Интернет (рис. 9.8).

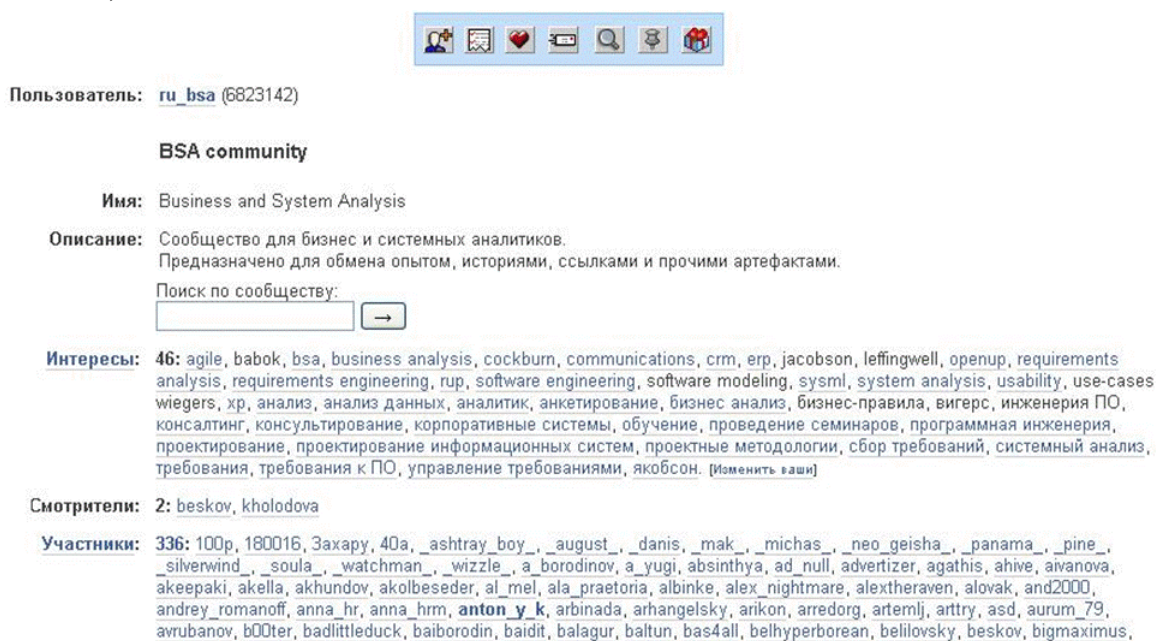


Рис. 9.8. Пример сообщества бизнес-аналитиков

Сообщества практиков могут находиться на разных этапах жизненного цикла: 1. Зарождение, 2. Создание, 3. Самостоятельное функционирование, 4. Активность, 5. Адаптивность. Каждый этап требует соответствующих инструментов поддержки⁵³, см. табл. 9.1, 9.2, 9.3, 9.4 и 9.5.

Таблица 9.1

Инструменты поддержки для этапа «Зарождение» сообщества

| Функция | Установление контакта |
|-------------------------------|--|
| Что делают участники | Находят друг друга и знакомятся Организация может быть не заинтересована в потенциальном сообществе или не знать о нем ничего ИЛИ Организация может оказать некоторую поддержку в поиске и людей и их знакомстве |
| Как поддержать процесс | Определить потенциальных членов сообщества Найти потенциальных членов сообщества Помочь людям собраться в одно целое |
| Технологии поддержки | Электронные системы связи: электронная почта, чаты, рассылка Телефон, телеконференции Онлайн форумы Справочная информация |

Таблица 9.2

Инструменты поддержки для этапа «Создание» сообщества

| Функция | Накопление информации и контекста |
|-----------------------------|---|
| Что делают участники | Основные члены Узнают друг о друге больше Делятся знаниями и опытом Вырабатывают общий язык Выбирают для себя роли и нормы общения Формируют совместную историю сообщества и записывают ее Начинают подбирать примеры из своей практики Сообщество получает признание. |

⁵³ Gongla P., Rizzuto C. R. Evolving communities of practice: IBM Global Services experience // IBM Systems Journal, vol. 40, NO. 4, 2001.

| Функция | Накопление информации и контекста |
|-------------------------------|--|
| Как поддержать процесс | Классифицировать информацию и сохранить ее Разработать способы поддержания всех этапов накопления и использования знаний Планировать деятельность сообщества Начало работы сообщества |
| Технологии поддержки | Общее хранилище информации Инструменты для первичной классификации и сортировки информации Системы управления документами и библиотекой Рабочая среда, способствующая сотрудничеству |

Таблица 9.3

**Инструменты поддержки для этапа
«Самостоятельное функционирование» сообщества**

| Функция | Обеспечение доступа и обучение |
|-------------------------------|--|
| Что делают участники | Члены сообщества <ul style="list-style-type: none"> • Строят доверительные отношения между собой • Осознают свою причастность к сообществу • Способствуют привлечению новых членов • Строят модель обмена знаниями и информацией • Делятся своим опытом • Активно ищут и собирают материалы для копилки знаний сообщества • Развивают обмен знаниями Организация взаимодействует с сообществом и лучше узнает его возможности |
| Как поддержать процесс | Привлекать новых членов Управлять объемами работы Осуществлять процесс накопления информации и знаний и работы с ними Поддерживать обмен неявными знаниями Развивать обмен информацией Собирать и обрабатывать отзывы Решать проблемы и исправлять ошибки Пересмотреть определение сообщества и масштаб его деятельности Обеспечить самоуправление сообществом |

| Функция | Обеспечение доступа и обучение |
|-----------------------------|---|
| Технологии поддержки | Порталы «Желтые страницы» Перевод на другие языки Обследования, анкетирование в электронном виде, /другие инструменты получения и обработки обратной связи |

Таблица 9.4

Инструменты поддержки для этапа «Активность» сообщества

| Функция | Сотрудничество |
|-------------------------------|---|
| Что делают участники | Члены сообщества работают <i>друг с другом</i> : решают проблемы и выполняют реальную работу Сообщество создает рабочие группы, Сообщество связывается с другими сообществами и взаимодействует с ними, Организация активно поддерживает сообщество и нормирует его работу, Организация уже может положиться на знания, которыми оперирует сообщество, поскольку они являются ценными для бизнеса |
| Как поддержать процесс | Решать проблемы, принимать решения Оценивать рабочую среду сообщества Развивать обучение и обратную связь Осуществлять интеграцию с организационными процессами Взаимодействовать с другими сообществами |
| Технологии поддержки | Встречи в электронном формате Обсуждения по темам Работа в группах Инструменты для анализа и принятия решений Интеграция технологий сообщества и технологии организации |

Таблица 9.5

Инструменты поддержки для этапа «Адаптивность» сообщества

| Функция | Инновации и рост |
|-------------------------------|--|
| Что делают участники | Сообщество изменяет среду, в которой существует, путем создания новой проекции, рынков, предприятий. Члены сообщества расширяют масштабы знаний и раздвигают горизонты деятельности. Сообщество спонсирует создание новых сообществ. Организация использует сообщество для новых возможностей и новых условий рынка. |
| Как поддержать процесс | Адаптироваться к окружающей среде, проявлять динамическую стабильность Развивать процессы в смежных областях Контролировать образование новых сообществ Уделять ключевое внимание инновациям. |

Пример технической поддержки сообщества практиков IBM Rational

Как и в большинстве сообществ практиков, участники сообществ IBM Rational находятся в разных географических местоположениях. У них нет возможности общаться лицом к лицу. Поэтому необходима технология, помогающая общаться на расстояниях глобального масштаба. На рис. 9.9. показаны средства ИТ-поддержки сообществ IBM Rational. Основная цель полученной инфраструктуры – соединить участников сообщества. Примеры:

- Рабочие места малых групп, где используются вики-страницы, сайты Lotus® Quickr™ и продукт IBM Rational Team Concert (на базе платформы IBM Rational Jazz).
- Дискуссионные онлайн-форумы с использованием IBM Forums.
- Синхронные соединения типа IBM Lotus Sametime® и IBM Lotus Sametime Unyte.™
- Новости, календари, информационные каналы с

использованием RSS-каналов и подкастов.

- Онлайн-пространства обучения, например IBM Learner Portal.
- Практическая библиотека примеров, ресурсов, сред, руководств и «дорожных карт» (в настоящее время мигрирует на IBM Rational Asset Manager).
- Доступ к экспертам по различным темам.



Рис. 9.9. Инфраструктура сообщества практиков IBM Rational предоставляет целый ряд различных способов связи между участниками

Для поддержки сообществ создаются специальные порталы, а для унификации структуры данных порталов разработаны онтологии сообществ (рис. 9.10).

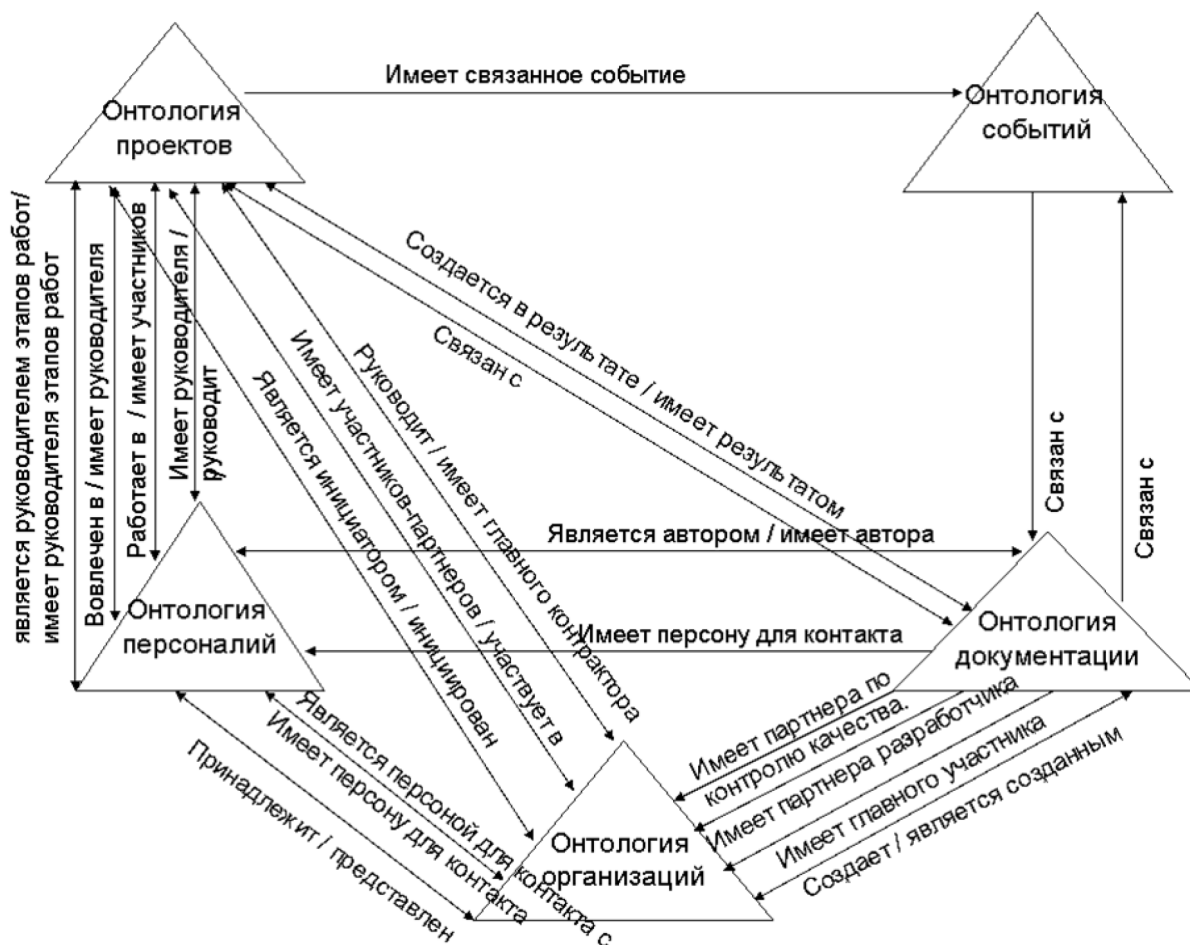


Рис. 9.10. Онтология портала проекта KnowledgeWeb (сообщества исследователей в области семантического веба)

Пример семантического портала, основанного на данной онтологии дан на рис. 9.11. На данном рисунке представлена страница с информацией об одном из участников проекта, а в виде гиперссылок указаны все связанные с ним проекты, отчеты по проектам, организации.

Asunciyn GymeZ-PiRez



Full Name: Asunciyn GymeZ-PiRez
e-mail: Asunciyn.GymeZ-PiRez
Homepage: <http://www.oeq-upm.net>
Date of Birth: 03/09/1967
Role: Scientific Vicedirector
Country: Spain
City: Boadilla del Monte, Madrid
Zip code: 28660
Street Address: Campus de MONTegancedo sn
Telephone: 34.91.3367439
Fax: 34.91.352.4819
Gender: Female

Принадлежит: [Universidad Politicnica de Madrid](#)
Является персоной для контакта с: [Universidad Politicnica de Madrid](#)

Является руководителем этапов работ в:

- [WP1.2: Evaluation for technology selection](#)
- [WP1.6: Semantic Portal Structure](#)

Работает в:

- [WP1.3: Technology Recommendations](#)
- [WP1.6: Semantic Portal Structure](#)
- [WP2.1: Scalability](#)
- [WP2.2: Heterogeneity](#)
- [WP2.6: Weaving the European Semantic Web Research](#)
- [WP3.1: EASE](#)
- [WP3.2: Educational Contents and Event Provision](#)
- [WP4.1: Operational Management](#)
- [WP4.2: Financial Management](#)
- [WP4.3: Technical Management](#)
- [WP4.4: Legal and Knowledge Management](#)
- [WP4.5: Society and Gender Issues](#)
- [WP4.6: Self-Assessment, Risk Analysis and Market Watch](#)

Вовлечена в:

- [T3.2.3 Educational Events: Summer schools](#)
- [T3.2.4 Towards the shared masters program: Joint curriculum](#)
- [T4.1.1 Communication with the EC and co-ordination of reporting](#)
- [T4.2.1 Accounting](#)
- [T4.3.1 Activity report](#)
- [T4.3.2 Joint Programme of Activity revision and expansion](#)
- [T4.5.1 Overseeing the science and society issues related to Knowledge Web](#)
- [T4.6.1 Self-assessment, risk analysis and market watch](#)

Является автором:

- [Benchmarking ontology tools. A case study for the WebODE platform](#)
- [Crete - WP1.6 - Semantic Portal](#)
- [D1.3.2: Identification of standards on metadata for ontologies](#)
- [D1.6.1: Portal requirements analysis and system design](#)
- [D1.6.2: Portal ontology](#)
- [D3.1.3: First version of the EASE association statutes](#)
- [D4.5.1: Report on Gender Action Plan](#)
- [Ontological Engineering](#)
- [Why Evaluate Ontology Technologies? Because it Works!](#)

Рис. 9.11. Пример страницы семантического портала

Российским аналогом представленной онтологии сообществ (хотя и более общим) является схема метаданных ЕНИП⁵⁴ (рис. 9.12). Единое научное информационное пространство РАН (ЕНИП РАН) – это инициатива, ставящая своей задачей интеграцию научных данных различных учреждений РАН и построение единой распределенной среды с целью обеспечения активных научных коммуникаций и эффективного использования научной информации, более эффективных средств поиска информации, сотрудничества и совместной работы.

⁵⁴ *Бездушный А.А., Бездушный А.Н., Серебряков В.А.* Схемы метаданных ЕНИП: практика применения OWL в ЕНИП // Информационное обеспечение науки: новые технологии (ред. Н. Е. Калёнов) – М., 2005. С. 155–182.

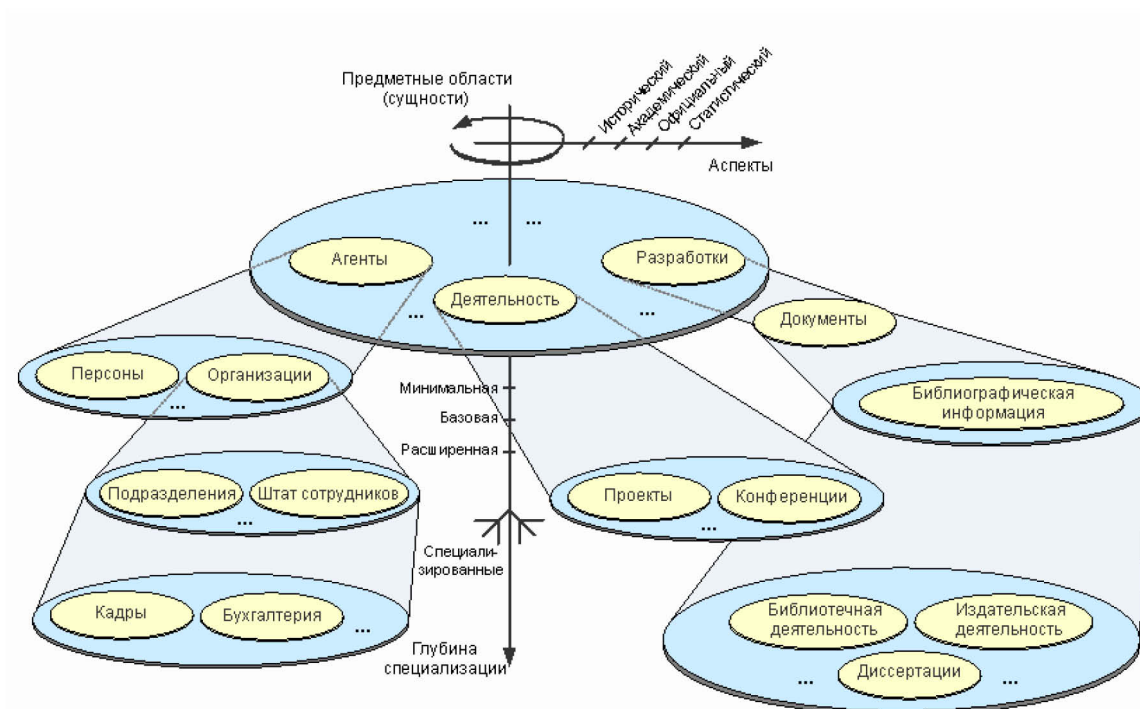


Рис. 9.12. Способ структурной организации элементов метаданных ЕНИП

10. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ

10.1. Традиционные программные средства для управления знаниями

Программные средства для УЗ можно разделить средства общего назначения и специализированные (профессиональные). Под средствами общего назначения будем понимать те, которые поддерживают работу с информацией в любой предметной области. Под специализированными те, которые поддерживают работу с информацией в определенной области, например, работа с клиентами (CRM), проектирование (CAD), управление (ERP, BPM...). Рассмотрим вначале средства общего назначения.

В обзоре будем опираться на отчеты известной аналитической компании Gartner Group (<http://gartner.com/>). Можно выделить три категории программных средств, выделяемых компанией Gartner, которые относятся к УЗ:

А. Технологии доступа к информации (Information Access Technology).

В. Совместная работа и социально-ориентированное ПО (Collaboration and Social Software).

С. Системы управления контентом (Enterprise Content Management), которую также можно рассматривать как инфраструктуру для представленных выше групп

Две первые категории соответствуют двум различным взглядам на знание: знание как «*запас*», знание как «*поток*».

Лидеры программных средств для управления знаниями

Рассмотрим некоторых типичных представителей и функциональные возможности программных средств, относящихся к сформированным категориям.

Категория А. «Технологии доступа к информации»

Что включает понятие доступа? Конечно, в первую очередь – поиск по ключевым словам и сложные запросы (например: «Найти все отчеты отдела маркетинга, которые писал Иванов?»). Также важен поиск с

использованием классификации (или таксономии) документов по группам.

Некоторые программы могут сами аннотировать любой входной документ и автоматически строить классификации по текстам документов.

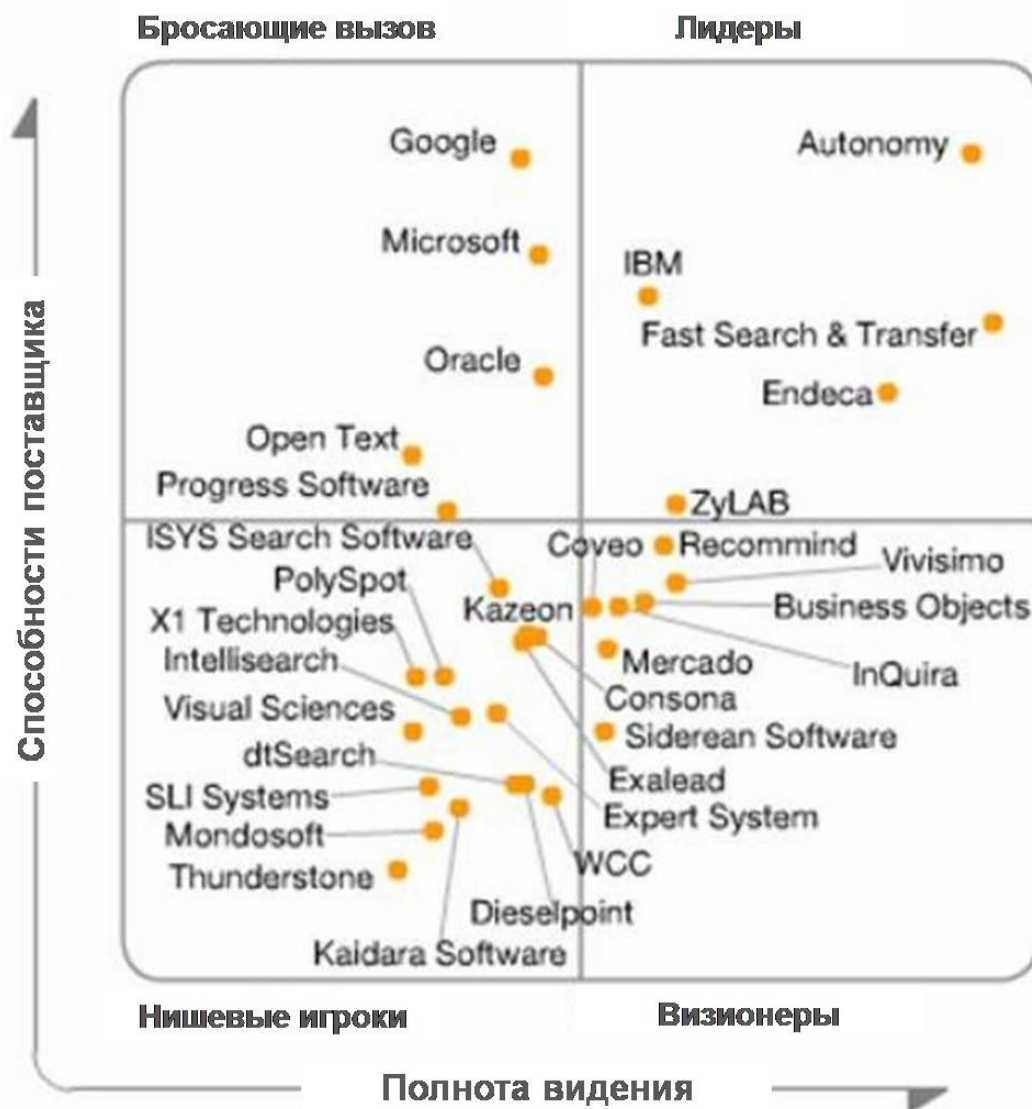
Типовой перечень функциональных возможностей продуктов данной категории приведен в обзоре Gertner и представлен на рис. 10.1.



Рис. 10.1. Функциональные возможности для категории «Технологии доступа к информации»

Наглядный обзор поставщиков программных средств дают диаграммы Magic Quadrant компании Gartner. Позиционирование компаний на диаграмме производится по двум координатам (рис. 10.2):

1. «Способности поставщика», оценивающие поставщиков по их способностям и достигнутым результатам в воплощении своих концепций;
2. «Полнота виденья», оценивающая поставщиков по их способности убедительно формулировать заключения о движении рынка, инновациях, потребностях клиентов, а также по их пониманию того, как использовать рыночные силы для создания преимуществ.



Источник: Gartner Group, 2007

Рис. 10.2. Обзор рынка программных средств в категории «Технологии доступа к информации»

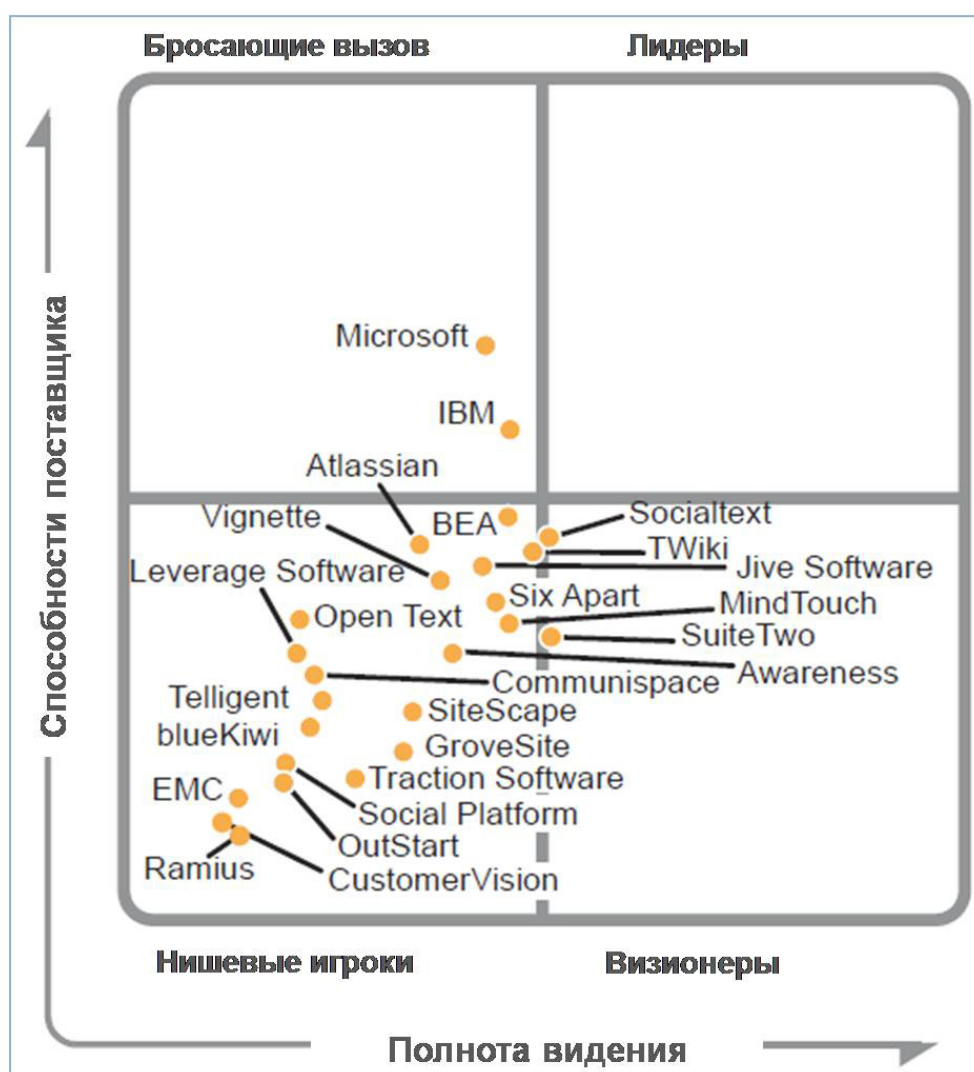
Как видно из рисунка, лидерами в данной категории являются продукты компаний Autonomy (IDOL™ Server), Fast Search&Transfer (FAST ESP, Convera RetrievalWare), Endeca, IBM.

Категория В. «Совместная работа групп и социально-ориентированное ПО»

Эти программные продукты поддерживают коммуникации (мгновенные сообщения, почту, IP-телефонию, видео-конференции) и предоставляют единую виртуальную рабочую среду, включающую wiki,

совместную работу с документами, рабочие доски и пр. Так можно организовать совместную разработку отчета в едином пространстве, выделив сообщество участников, или организовать коллективное обсуждение корпоративного праздника. Программы этой категории создают персональные профили сотрудников с описанием интересов, компетенции, опыта работ и т. п., а также каталоги экспертов.

Лидерами в данной группе являются продукты компаний BEA (BEA Aqualogic), SpikeSource (SuiteTwo), Socialtext (Socialtext), IBM и Microsoft, см. рис. 10.3:



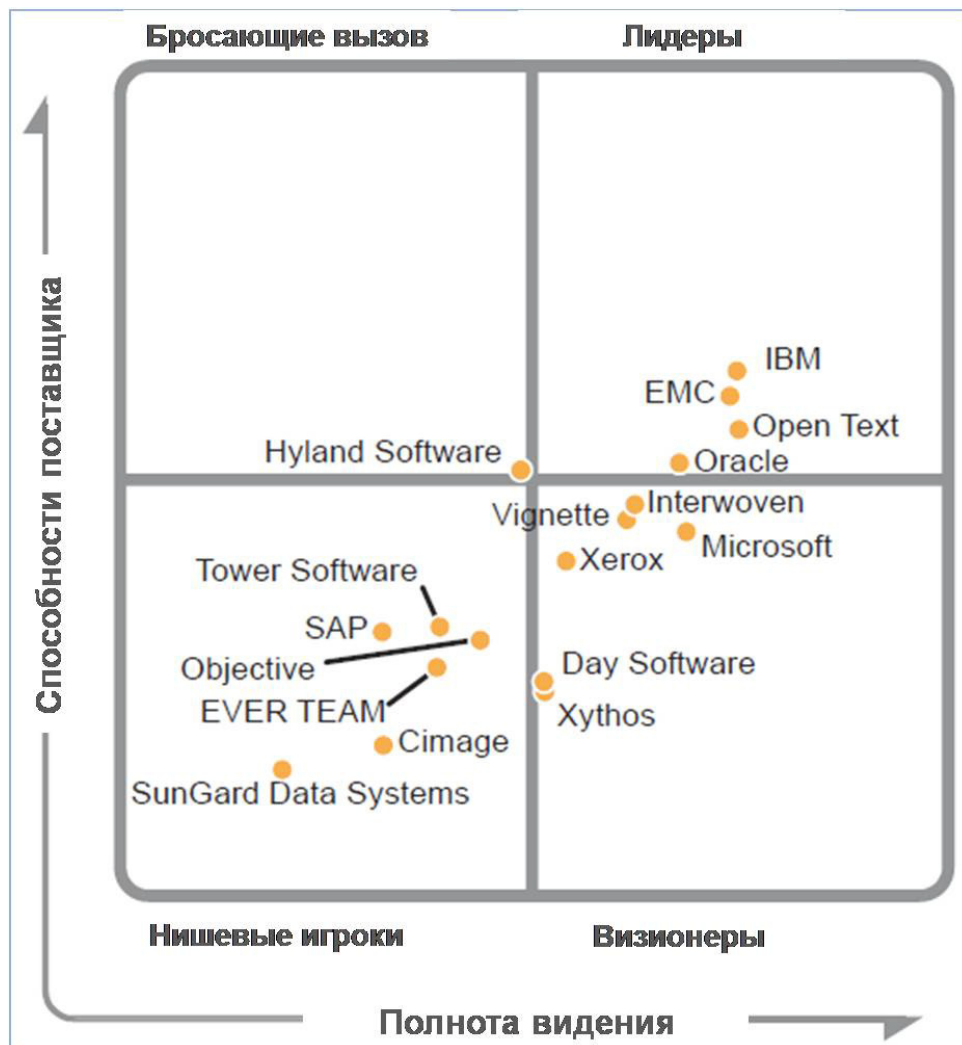
Источник: Gartner Group, 2007

Рис. 10.3. Обзор рынка программных средств в категории «Совместная работа групп и социально-ориентированное ПО»

Категория С. «Системы управления контентом (Enterprise Content Management)»

Эта группа программ включает различные инструменты для гибкой и надежной работы с документами, например, регистрацию, управление версиями, безопасность, сканирование и распознавание. Также сюда относят создание архивов, наполнение, обновление и структурирование содержимого сайтов, организация взаимодействия с посетителями сайтов.

Лидерами в данной категории являются продукты компаний IBM (Content Manager, Filenet P8), EMC (EMC Documentum), Opentext (Livelink ECM 10), см. рис. 10.4:



Источник: Gartner Group, 2007

Рис. 10.4. Обзор рынка программных средств в категории «Системы управления контентом»

Основными поставщиками интегрированных решений для разработчиков программных средств управления знаниями являются компании IBM, EMC, Opentext и Microsoft. Эти компании стараются покрыть функциональными возможностями своих программных продуктов весь комплекс задач управления знаниями. Компания IBM занимает лидирующие позиции во многих классах программных продуктов, поэтому далее будет детально рассмотрена ее продуктовая линейка. Программы других поставщиков имеют сходный функционал, отличаясь лишь технологиями реализации, а также качеством и эффективностью работы.

Продукты категории А – «Технологии доступа к информации»

IBM OmniFind

OmniFind ориентирован на поиск информации для работников организации и клиентов. Поиск может одновременно проводиться по всем ресурсам – в Интернете, по внешним корпоративным сайтам компании, в базах данных, и т. д.

OmniFind предоставляет полнотекстовый поиск по ключевым словам и фразам для начинающих пользователей, а также расширенные виды поиска для продвинутых пользователей. Система позволяет также задавать правила для автоматической классификации документов.

IBM Classification Module

Данный модуль автоматически классифицирует неструктурированный контент по заданным категориям, а также позволяет создавать и настраивать таксономии. Эта программа выступает дополнением к IBM OmniFind, которое позволяет повысить эффективность и точность поиска с помощью навигации (browsing) и наложение ограничений на результаты поиска (например, поиск только в определенной категории).

WebSphere Portal

WebSphere Portal – это программная среда, реализующая единую точку индивидуального взаимодействия людей и программ. С точки зрения управления знаниями особенно важно в портале сочетание персонализации контента (взгляд на знание как на «запас») с возможностями совместной работы и обмена неформализованными знаниями (взгляд на знание как на «поток»).

Продукты категории В – средства совместной работа

IBM Lotus Notes and Domino

IBM Lotus Notes – это главная часть программного обеспечения IBM Lotus Domino, который обладает функциями планирования (календарь) и работы с электронной почтой. Продукт является платформой для совместной работы и повышения гибкости предприятия в целом. Близок известному продукту от Microsoft – почтовой системе Outlook.

IBM Lotus Sametime

IBM Lotus Sametime – это программа для общения и совместной работы с возможностью проведения Web-конференций (телеконференций через Интернет). Lotus Sametime позволяет передавать голосовые и видео-сообщения, определять местонахождение участников, и поддерживает связь с мобильными устройствами (например, мобильными телефонами).

IBM Lotus Quickr

IBM Lotus Quickr – это программный продукт для обеспечения совместной работы нескольких пользователей, который помогает организовать совместную работу сотрудников в интерактивном режиме как внутри организации, так и за ее пределами.

В интерактивные рабочие области пользователи могут включать календарь, дискуссионные форумы, блоги, wiki и другие средства совместной работы для управления проектами и создания общего контента.

IBM Lotus Connections

IBM Lotus Connections представляет собой решение для социализации бизнес-процессов, расширяющее инновационные возможности и ускоряющее работу с помощью объединения сотрудников, партнеров и заказчиков в динамические сети.

«Если Lotus Quickr – это управление контентом в контексте проекта, то Lotus Connections предоставляет средства для объединения людей, которые должны сотрудничать в работе над проектом. Система содержит пять основных компонентов: Profiles, Blogs, Communities, Activities и средство создания закладок Bookmarks (Dogear). Используя их, я и мои коллеги в моей организации и, возможно, за ее пределами, знают: что я из себя представляю – Profiles; над чем я сейчас работаю – Activities; чем я интересуюсь – Bookmarks (Dogear); что хочу обсудить – Blogs; с кем нахожусь в контакте – Communities», – говорит Дэвид Фаррел, вице-

президент IBM по программному обеспечению Lotus в регионе «Европа, Ближний Восток и Африка».

Продукты категории С – системы управления контентом

Далее представлены не все компоненты системы управления контентом, а только те, которые используются для реализации системы управления знаниями.

IBM Content Manager

Это программный продукт для хранения, контроля, распределения и интеграции информации. Интегрирует данные различных форматов (документ, изображение, Web, медиа и другие) и сопрягается с различными профессиональными системами сканирования.

IBM WebSphere Information Integrator

WebSphere Information Integration обеспечивает доступ к данным, хранящимся в разнородных системах, интегрируя содержимое порталов, бизнес-приложений и других программ. Этот интегратор делает так, что множество разнородных хранилищ данных выглядят и функционируют как единый архив.

Некоторые отечественные программы

Среди отечественных продуктов, позиционирующихся как системы для поддержки УЗ, следует отметить несколько:

1) Система «Айкумена Аналитик» компании IQMen (http://www.iqmen.ru/our_proposal/) обеспечивает информационно-аналитическую поддержку принятия решений с помощью таких современных технологий как:

- Автоматический сбор информации из разнородных источников: электронные СМИ, аналитические агентства, специализированные базы данных.
- Автоматическое составление и рассылка индивидуальных тематических отчетов экспертам и руководству.
- Средства графического анализа событий по любой тематике, оперативное выявление в новостном потоке горячих тем.
- Мгновенная подготовка досье по любому вопросу.
- Выборка наиболее важной фактической информации – дат, финансовых данных, персон и организаций.
- Автоматическое выявление физических и юридических лиц,

имеющих отношение к любому вопросу и анализ взаимосвязей между ними.

2) Система «**Галактика ZOOM**» – инструмент для создания хранилища текстовой информации, который обладает возможностями для проведения поиска и аналитических исследований.

- Поиск информации в больших информационных массивах.
- Выявление значимых слов и словосочетаний документа, отражающих содержание текста.
- Сравнение документов – обнаружение сходства, различия и аномалий изучаемых объектов.
- Динамика во времени – отслеживание изменений во времени слов или словосочетаний характеризующих изучаемый объект.

3) Система **DocsVision** также предназначена для задач управления документами и процессами (<http://www.docsvision.com/>). Система DocsVision «Делопроизводство» для создает архивы документов, автоматизирует делопроизводство в организации.

4) Система **ЕВФРАТ-Документооборот** (<http://www.evfrat.ru>) – комплексное решение по организации электронного документооборота на предприятии. Внедрение системы обеспечивает:

- систематизацию учета и хранения документов;
- оперативный доступ к документам и отчетной информации;
- управление процессами движения и обработки документов;

5) **Yndex.Server** – программа для поиска информации в интернете или в локальной сети с учетом морфологии русского языка.

Веб 2.0 инструменты для работы со знаниями

В 2005 году Тим О’Рейли описал концепцию «эволюционировавшей всемирной паутины», которую он обозначил термином Web 2.0, подчеркнув тем самым ее значимость как следующего поколения интернет-систем. Под этим понятием он объединил уже сформировавшиеся к тому времени признаки веб-сайтов «новой волны», основным из которых, с пользовательской точки зрения, можно считать использование «коллективного разума» или, точнее, «коллективной деятельности». Пользователи Интернета сами становятся создателями информации – создаваемый в сети контент, его описание и структурирование, оценка и ранжирование становятся результатом

деятельности рядовых пользователей. Технологически вовлечение пользователей обеспечивается за счет интерактивности веб-страниц и простоты их интерфейсов – для размещения информации в Интернете пользователю не нужно никаких специальных знаний.

Вот неполный список инструментов Веб 2.0:

- Интернет-почта (mail.ru, yandex.ru, gmail.com),
- средства обмена сообщениями (ICQ, google-чат)
- блоги (например, Живой Журнал, корпоративные блоги Яндекса, Google-a) – он-лайн журналы и дневники, позволяющие обмениваться мнениями, размещать новости. Обновления блогов могут рассылаться с помощью RSS-потоков – см. далее,

- социальные сети (например, vkontakte.ru, moikrug), позволяющие узнать о знаниях, опыте, интересах других людей и представить свои,

- средства интеграции и чтения новостных RSS-потоков (Яндекс Лента, Google Reader). RSS (Really Simple Syndication) – средство подписки на обновления новостей сайтов, записей блогов, результатов поиска по определенному запросу и на любые другие обновляемые материалы – фотографии, видео, веб-закладки и т. д.

- Peer-to-peer сети (иногда P2P-сети) – средство эффективного обмена файлами (музыкой, видео, текстами) либо через Интернет, либо между ограниченным числом пользователей. Данное средство открывает доступ к содержимому персональных компьютеров пользователей и обеспечивает высокую скорость передачи информации, благодаря специальной технологии разделения файла на отдельные фрагменты.

- веб-закладки (del.icio.us, БобрДобр), позволяющие пользователям не только сохранять ссылки на интересные для них ресурсы, но и сопровождать их ключевыми словами, обмениваться этими ссылками, анализировать популярность ресурсов (количество ссылок) и тенденции, находить других пользователей со схожими интересами.

- средства для размещения фотографий (photosight.ru, flickr.com), презентаций (slideshare) и видео (youtube), с богатыми возможностями описания материалов (теги, категории), поиска, обсуждения.

- wiki-средства (pbwiki.com, wikispaces.com) – средства создания и функционирования веб-сайтов по типу Википедии (см. ru.wikipedia.org). Страницы вики-сайта совместно редактируются сотрудниками организации и выступают альтернативой письмам и вложениям электронной почты.

- программы для коллективного редактирования документов в онлайн (Google Docs),

- гибридные приложения (mashups), объединяющие несколько инструментов Web 2.0 между собой или с другими приложениями (например, интеграция географических карт от Google, Yahoo со средствами размещения фотографий – Panoramio, flickr позволяет пользователям привязывать фотографии к местности)

- средства планирования – календари (Google Календарь) и т. п.

К Web 2.0 можно также отнести «ПО как сервис» (Software as a Service, SaaS). Такое ПО работает через Интернет, а компания ежемесячно оплачивает услуги сервис-провайдера. Это дешево и удобно. Не нужно покупать дорогие лицензии, устанавливать и настраивать сложный софт.

Более полумиллиона западных компаний успешно используют онлайн-платформу AppExchange от Salesforce.com, что позволяет подписаться на любое из нескольких сотен бизнес-приложений, доступных для подписки (CRM-системы, бухгалтерские программы и др.). В данный момент SaaS значительно уступает по популярности традиционным моделям использования ПО, но в будущем ситуация должна измениться. Gartner прогнозирует, что к 2011 г. доля SaaS в общем объеме рынка программного обеспечения вырастет до 25 %, то есть в пять раз по сравнению с 2005 г.

Инструменты Web 2.0 подходят для поддержки совместной работы в системах управления знаниями: они просты, эффективны и уже знакомы некоторым пользователям. По оценкам компании McKinsey основной целью использования инструментов Web 2.0 в бизнесе является поддержка совместной работы в компании – 75 % опрошенных менеджеров (70 % – взаимодействие с покупателями, 51 % – взаимодействие с поставщиками и партнерами). При этом большая часть респондентов рассматривает

совместную работу как инструмент управления знаниями⁵⁵. Рисунки 10.5 и 10.6 отражают мнения пользователей, полученные в рамках других исследований, о пользе и препятствиях таких инструментов.



Источник: InformationWeek Research, 2006

Рис. 10.5. Польза инструментов Веб 2.0 для бизнеса



Источник: InformationWeek Research, 2006

Рис. 10.6. Препятствия для Веба 2.0 в бизнесе

⁵⁵ Bughin J., Manyika J. How businesses are using Web 2.0: A McKinsey Global Survey // McKinsey Quarterly, March, 2007

Вряд ли инструменты Web 2.0 вытеснят с рынка стандартные системы управления документами и софт для групповой работы, но они определенно могут его дополнить и стать неотъемлемой их частью. Собственно, такая интеграция происходит уже сейчас (см. выше описание IBM Lotus Quickr, Sametime и др.).

В России внедрение в корпоративные системы блогов, wiki и социальных сетей идет несколько медленнее, чем на Западе. Причины заключаются в худшем состоянии технической инфраструктуры. Широкополосный Интернет, который является главным условием для развития публичных сервисов Веб 2.0, распространен только в Москве. По этой причине блоги, подкастинг, wiki и прочие современные технологии еще не стали достаточно популярными среди пользователей. Корпоративная культура только начинает приспосабливаться к использованию социально-ориентированных сервисов в бизнесе.

Мы рассмотрели программные средства, которые могут использоваться для создания СУЗ в произвольной предметной области. Однако почти в любой специализированной области существуют свои программные средства автоматизации. Например, дизайнеры работают в продуктах для обработки изображений, проектировщики в системах автоматизированного проектирования, программисты в средах программирования, менеджеры и бизнес-аналитики в системах организационного моделирования и проектирования. Как правило, такие системы имеют свои встроенные компоненты для накопления и передачи знаний, например, шаблоны, типовые узлы, справочники и библиотеки в САПР; скрипты, плагины и паттерны в Photoshop и т. д. В результате программные продукты для СУЗ могут быть сгруппированы так, как показано на рис. 10.7.

Задача формирования СУЗ на программном уровне требует укращения всего «зоопарка» отдельных программных систем и сопряжения их в единую систему, ядром которой является некоторая база знаний.

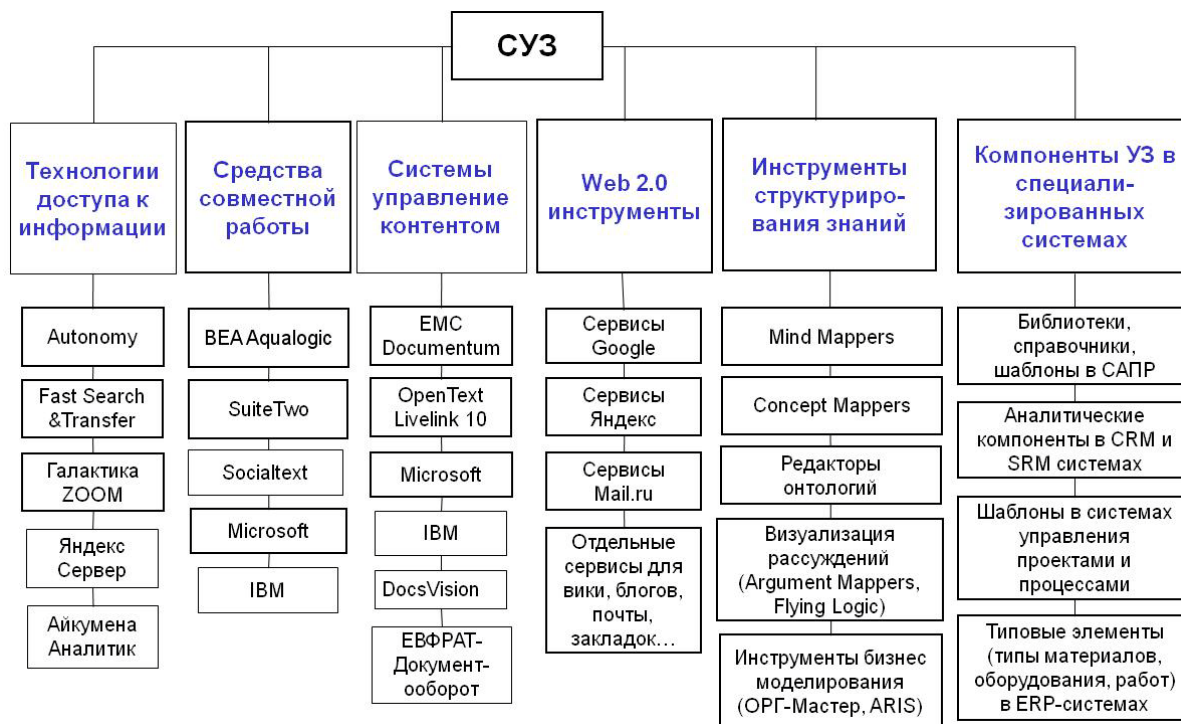


Рис. 10.7. Систематизация программных средств для СУЗ

Интеграция отдельных инструментов УЗ с целью создания ценности для пользователя и предприятия превращает их в интегрированную СУЗ предприятия. СУЗ может носить имя «портал знаний», «информационная система корпоративной памяти» и т. п. В случае портала знаний акцент делается на поддержке задач уровня пользовательских приложений – поиск, визуализацию знаний, в случае «корпоративной памяти» – на накоплении знаний.

А) Информационная система корпоративной памяти

Одним из первых инструментов УЗ стали *хранилища данных*, которые работают по принципу центрального склада. Как правило, хранилища содержат многолетние версии обычной базы данных (БД), физически размещаемые в той же самой базе. Когда все данные содержатся в едином хранилище, изучение и анализ связей между отдельными элементами могут быть наиболее эффективны. В дальнейшем идея хранилища была развита в понятие *корпоративной памяти (corporate memory)*⁵⁶, которая по аналогии с человеческой памятью позволяет

⁵⁶ Kühn O., Abecker A., Corporate Memories for Knowledge Management in Industrial Practice: Prospects and Challenges. 1998.

накапливать информацию из предыдущего опыта и якобы избежать повторения ошибок, что является чисто декларативным утверждением.

Корпоративная память хранит информацию из различных источников предприятия и делает эту информацию доступной специалистам для решения производственных задач.

Основные функции информационной системы корпоративной памяти:

- *Сбор и систематическая организация информации,*
- *Интеграция с существующими автоматизированными системами,*
- *Обеспечение нужной информации по запросу (пассивная форма) и при необходимости (активная форма).*

Б) Порталы знаний и карты знаний

Корпоративный портал – это система, которая объединяет все имеющиеся у организации информационные ресурсы (приложения, базы и хранилища данных, аналитические системы и пр.) и, используя web-интерфейс, предоставляет пользователям единый защищенный доступ к корпоративной и внешней информации⁵⁷ (рис. 10.8).



Рис. 10.8. Концептуальная схема корпоративного портала

Некоторые полезные свойства порталов, упоминаемые в различных источниках:

⁵⁷ Collins H. Enterprise Knowledge Portals. AMACOM, 2003.

- систематизация контента и предоставление удобных средств поиска и навигации,
- управление контентом для различных групп сотрудников (управление доступом).
- внутрикорпоративный обмен знаниями и совместная работа за счет наличия различных конференций, форумов и единого рабочего пространства,
- индивидуальный персонализированный интерфейс.

10.2. Архитектура системы управления знаниями на основе онтологии

Описание применения онтологий для создания и поддержки порталов знаний можно найти в работах Стааба, Медхе, Штудер (Staab, Maedhe, Studer). В данных работах указывается, что поддержание любого портала требует с одной стороны управления порталом, а с другой интеграции и индексации информации. Основная идея, предлагаемая авторами, заключается в создании и использовании единой онтологии, которая будет с одной стороны обеспечивать управление порталом, а с другой интеграцию и индексацию информации, см. рис. 10.9.

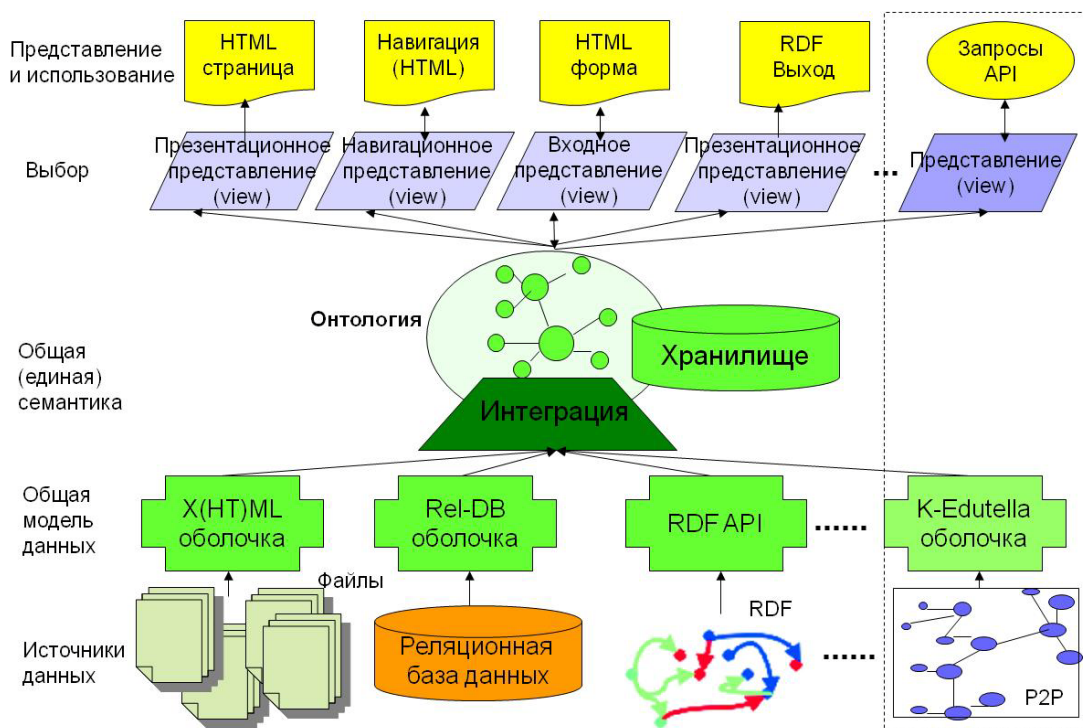


Рис. 10.9. Архитектура семантического портала на основе онтологии

Такой подход с одной стороны позволяет значительно снизить усилия на создание и поддержку портала, а с другой стороны повысить гибкость и адаптивность портала – его становится легко перестраивать и это доступно бизнес-пользователям, благодаря соответствующему интерфейсу. В предлагаемом авторами подходе для создания и перестройки онтологии, которая является системообразующим элементом портала, используется специальное программное обеспечение – редактор онтологий.

Для описания типичной современной архитектуры СУЗ на основе онтологий мы несколько адаптировали рисунок из работы Аккерманса и Мики, см. рис. 10.10.



Рис. 10.10. Пример архитектуры СУЗ

Нижний уровень архитектуры (рис. 10.10) обеспечивает получение знаний из разнородных источников структурированной (базы данных и знаний, таблицы, формы) и неструктурированной (документы)

информации. Получение знаний подразумевает аннотирование разнородных источников информации с помощью онтологии, которая позволяет отразить содержание информации. Для получения онтологии используются программные средства автоматического формирования онтологии, поддержки ручной разработки и редактирования онтологии, средства слияния, объединения и оценки качества онтологий. Для аннотирования используются средства автоматической классификации документов по онтологии, извлечения информации из текста и баз данных. Полученные онтологии и метаданные представляются с помощью специальных языков – OWL, RDF.

Промежуточный уровень обеспечивает хранение онтологий и метаданных, управление версиями, доступом, передачу данных для взаимодействия с внешними системами и хранилищами, обработку запросов и логический вывод.

Пользовательские приложения предоставляют пользователям доступ к знаниям, находящимся в системе. Доступ может осуществляться с помощью поискового механизма, средств навигации и визуализации, а также путем непосредственного обмена знаниями между пользователями (например, путем обмена ссылками на ресурсы). Взаимодействие пользовательских приложений с промежуточным уровнем осуществляется с помощью языков запросов и манипулирования данными (например, SPARQL).

10.3. Программные средства управления знаниями на основе онтологий

В настоящее время для всех компонент архитектуры СУЗ на основе онтологий (раздел 10.2) существует множество различных программных реализаций, которые сведены в табл. 10.1.

Семантические модели, как правило, должны быть интегрированы с существующими средствами автоматизации предприятия:

- Система управления документами и контентом,
- Информационный портал,
- Интегрированные бизнес-коммуникации,
- Разведка бизнеса и анализ клиентов,
- Системы электронного обучения.

Таблица 10.1

**Примеры программных продуктов для реализации СУЗ
на основе онтологий**

| Компоненты семантической СУЗ | Конкретные продукты |
|-------------------------------------|---|
| Редакторы онтологий | Altova Semanticworks, NeOn Toolkit, DOODLE (Domain ontology rapid development environment), DOE, DOME, Fenfire, Graphl, GrOWL, IBM Integrated Ontology Development Toolkit, Infered, IsaViz, KAON OI Modeler, Linkfactory, Ontotrack, Powl, Protégé, Rhodonite, SemTalk, SWOOP, Topbraid composer, WebODE, DogmaModeler, ICOM |
| Репозиторий онтологий | KAON2, Jena, Sesame, Ontology Server, RDF Server, Knowledge zone, Onthology, OntoSelect, DAML Ontology Library, SchemaWeb, ONTOSEARCH2, Protégé Ontologies Library, OntStore, RDFPeer, RDF2GO |
| Ответы на запросы | AJAX Client for SPARQL, Bor, Corese, KAONP2P, KAONWeb, Oyster2 |
| Обработка семантических запросов | AeroText, Sesame |
| Редактор запросов | Ontogator, SemSearch, Knowledge Sifter, Haystack, OntoViews |
| Инструменты визуализации онтологий | Brownsauce, BrowseRDF, Drive RDF Browser, Disco, Horus, Longwell, OINK, RDF Gravity, Tabulator, Welkin, Jambalaya, Ontosphere 3D, OntoViz, OWLViz, TGVizTab |
| Автоматизированный вывод | HermiT, Pellet, FaCT++, Racer Pro, CEL, CB |
| Инструменты оценки онтологий | ARP: Another RDF Parser, CLEANONTO, ConsVISor, Eyeball, ODEVal, OWL API, Semantic Web RDF Library for C#/.NET |

| Компоненты семантической СУЗ | Конкретные продукты |
|---|--|
| Инструменты отображения онтологий | AMV, AUTOMS, CMS, CtxMatch, eTuner/iMap/Glue/LSD, Falcon-AO, NOM, QOM, APFEL, H-Match, LOM, MapOnto, MetaQuerier, MoA, OLA, S-Match, SAMBO, ToMAS/Clio, OntoBuilder, OntoMerge, Aligment API & Aligment server, COMA & COMA++, FOAM, PROMPT, Rondo, Chimaera, MAFRA, Mapping Discovery |
| Ручное аннотирование | OCAT, OntoMat-Annotizer, M-OntoMat-Annotizer, PhotoStuff (Mindswap), AKTive Media – Ontology based annotation system, Ontolog, Magpie |
| Автоматическое аннотирование | KIM, AKTiveAgent, GATE ML, • OntoOffice, OntoText |
| Автоматическое наполнение (Ontology populator) | CLIE, AKTive Futures, ALVIS |
| Поиск и ранжирование онтологий | Watson, Swoogle |
| Управление эволюцией онтологий | KAON, DOME, MarcOnt Portal, Linkfactory, Powl |
| Инструменты создания частных представлений (ontology views) | Longwell, TGVizTab, OntoViz, Jambalaya, OWLViz, /facet, mSpace, VIKI, CropCircles, CS Aktive Space, SpaceTree, TreeMap, Spotlight, IsaViz |
| Редакторы экземпляров (наполнение онтологий) | GATE Ontology Editor, OCAT |
| Поиск веб-сервисов | Hybrid OWL-S Web Service Matchmaker – OWLS MX, The TUB OWL-S Matcher (The OWLSM), WSMX Discovery Framework, OWL Semantic Search Services (owl-semsearch) |
| Выбор веб-сервисов | WSMX Selector and Ranking Prototype |
| Композиция веб-сервисов | Kweb Semantic Web Service Composition, Semantic Web service composition through Cusal Link Composition, Composer, Semantic web services browser and composer, Web service Composition, Service Composition Engine (Developed within ASG) |
| Хореография веб-сервисов | WSMX Choreography Engine, IRS-III |

10.4. Процесс разработки систем управления знаниями на основе онтологий

Разработка СУЗ на основе онтологий разделяется на две основные части – разработка онтологии (сети онтологий) и программная реализация СУЗ. Программная реализация СУЗ основана на компонентном подходе к разработке информационных систем. То есть вначале для конкретной программной реализации СУЗ нужно определить ее архитектуру и состав компонента (см. раздел 10.2), потом для каждого компонента выбрать программный продукт для реализации (см. раздел 10.3), а потом интегрировать их. Разработка же онтологий может быть основана на какой-либо из существующих методологий онтологического инжиниринга. Методология On-To-Knowledge была представлена в разделе 2.6. В данном разделе будут рассмотрены две наиболее современные методологии разработки онтологий: DILIGENT и NeOn.

Основную часть процесса разработки СУЗ на основе онтологий составляют работы по онтологическому инжинирингу, поэтому рассмотрим основные методологии в данной области.

Методологии инженерии онтологий

Методология DILIGENT – методология разработки и сопровождения (evolution) онтологии в условиях множественности и распределенности экспертов. Процесс разработки онтологии в соответствии с данной методологией включает в себя следующие шаги:

1. **Создание.** На данном этапе быстро создается первая версия онтологии, которая может быть передана для использования заинтересованным сторонам.
2. **Локальная адаптация.** Пользователи применяют исходную онтологию и локально адаптируют ее под свои потребности.
3. **Анализ.** Контрольная комиссия оценивает изменения, предлагаемые пользователями онтологии - анализирует локальные онтологии, запросы и предложения пользователей. В результате такого анализа комиссия пытается выявить сходства в онтологиях пользователей.
4. **Ревизия.** Комиссия принимает решения об изменениях в онтологии. Комиссия должна регулярно проводить ревизию разделяемой онтологии, чтобы локальные онтологии существенно от нее не

отличались.

5. **Локальное обновление.** Пользователи онтологии обновляют свои локальные онтологии с учетом измененной версии общей онтологии.

Методология NeOn (www.neon-project.org)

Особенности методологии NeOn:

- Повторное использование знаний: неонтологических ресурсов (гlossарии, тезаурусы, таксономии, классификационные схемы...), готовых онтологий, шаблонов (ontology design patterns),
- Контекстуализация онтологий (кастомизация, модуляризация, локализация),
- Поддержка ВСЕГО жизненного цикла работы с онтологиями,
- Динамика (эволюция онтологий),
- Совместная работа.

Методология NeOn основана на следующих элементах:

Сеть онтологий (Ontology network) – комплекс интегрированных онтологий, созданный для конкретного приложения.

Типовые задачи онтологического инжиниринга – в рамках методологии NeOn выделено 53 типовые задачи работы с онтологиями (концептуализация онтологии, отображение онтологии, эволюция онтологии...), для которых разработаны согласованные определения предлагаются специальные методы, техники и инструменты. (русский перевод см. http://ontology.ipi.ac.ru/index.php/Категория:Список_терминов),

Модель жизненного цикла сети онтологий (ontology network life cycle model) – обобщенная схема, выбранная организацией для систематизации и отбора типовых задач онтологического инжиниринга с целью получения собственного (для организации) жизненного цикла онтологической системы.

Возможные модели жизненного цикла онтологий:

- Каскадная (Waterfall),
- Пошаговая (Incremental),
- Итеративная (Iterative),
- Развития (эволюция) прототипа (Evolving Prototyping),
- Спиральная (Spiral).

Жизненный цикл сети онтологий (ontology network life cycle) –

последовательность задач онтологического инжиниринга для реализации конкретного проекта, выполняемая инженерами по знаниям и разработчиками программных продуктов. Данная последовательность получается путем отображения типовых задач онтологического инжиниринга на выбранную модель жизненного цикла онтологической системы.

В методологии онтологического инжиниринга NeOn предлагается создавать жизненный цикл онтологической системы под конкретную ситуацию.

Последовательность разработки жизненного цикла сети онтологий:

1. Определение требований к разработке сети онтологий.

2. Выбор модели жизненного цикла сети онтологий.

– Выявление характеристик и ограничений для процесса разработки сети онтологий;

– Выявление всех возможных моделей жизненного цикла сети онтологий (для выбора модели жизненного цикла сети онтологий предлагается дерево решений);

– Оценка различных моделей жизненного цикла сети онтологий на основе прошлого опыта и возможностей организации;

– Выбор модели жизненного цикла сети онтологий, которая наиболее соответствует характеристикам и ограничениям для процесса разработки онтологической системы.

3. Выбор требуемых задач онтологического инжиниринга из списка типовых задач. Часть задач является обязательной для любой ситуации (required), а часть задач выполняется при определенных условиях (If applicable). Для выбора задач используются диагностические вопросы.

4. Отображение выбранных задач онтологического инжиниринга на выбранную модель жизненного цикла сети онтологий.

5. Определение порядка выполнения задач – жизненный цикл сети онтологий под конкретную задачу конкретной организации.

Наиболее распространенные сценарии создания сети онтологий (ontology network), см. рис. 10.11:

Сценарий 1: Создание сети онтологий «с нуля» без использования существующих ресурсов

Сценарий 2: Создание сети онтологий путем повторного

использования «неонтологических» ресурсов (поле знаний, не доведенное до уровня онтологии или ее фрагментов);

Сценарий 3: Создание сети онтологий путем повторного использования существующих онтологий или модулей;

Сценарий 4: Создание сети онтологий путем повторного использования и реинжиниринга существующих онтологий или модулей;

Сценарий 5: Создание сети онтологий путем повторного использования и объединения существующих онтологий или модулей;

Сценарий 6: Создание сети онтологий путем повторного использования, объединения и реинжиниринга существующих онтологий или модулей;

Сценарий 7: Создание сети онтологий путем реструктуризации существующих онтологий или модулей;

Сценарий 8: Создание сети онтологий путем локализации существующих онтологий или модулей.

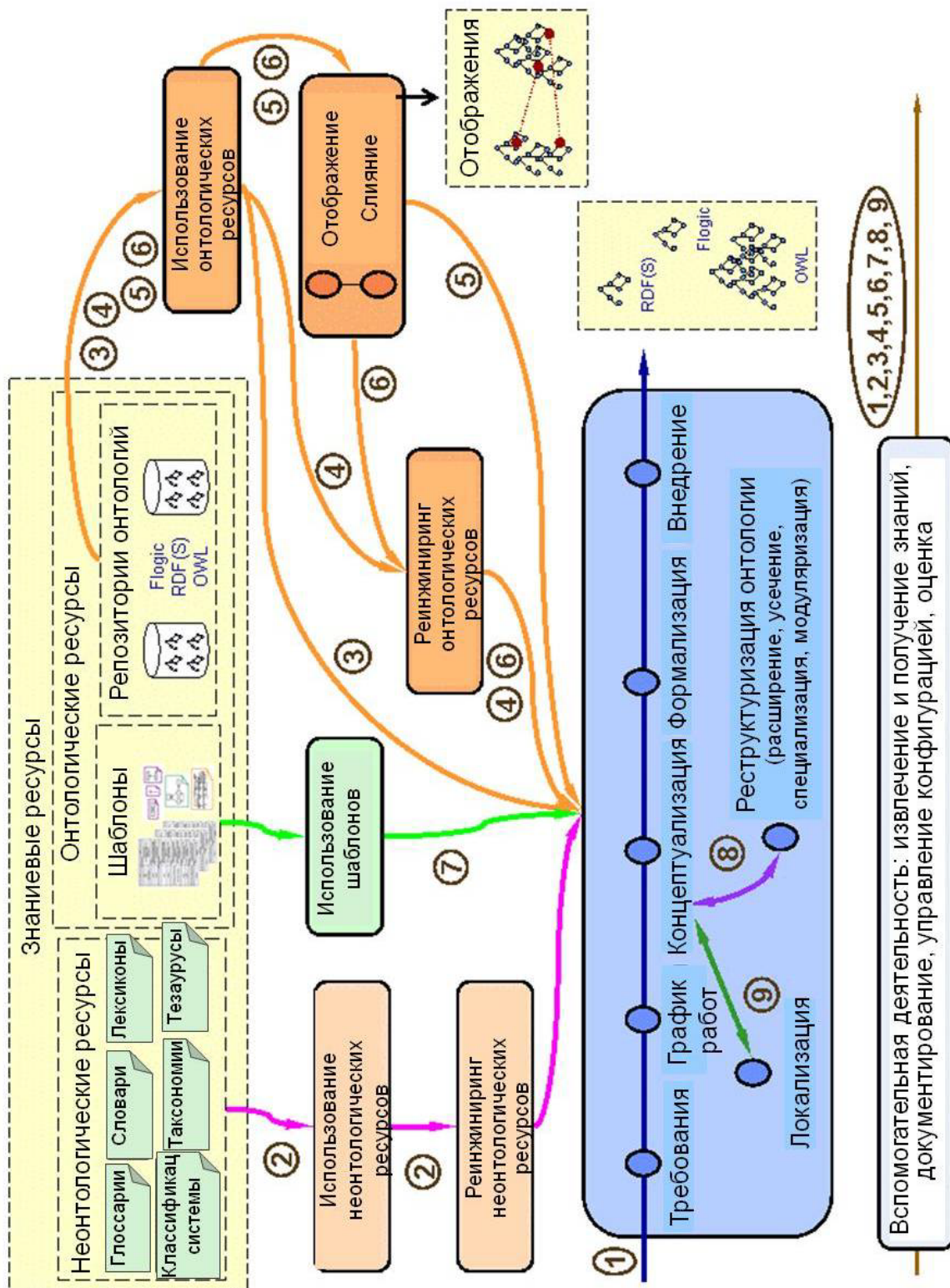


Рис. 10.11. Сценарии создания сети онтологий в методологии NeOn

11. ПРАКТИЧЕСКИЕ ПРИМЕРЫ СИСТЕМ УПРАВЛЕНИЯ ЗНАНИЯМИ

11.1. База знаний компании IBS

Компания IBS (<http://ibs.ru/>) основана в 1992 году. Основной сферой деятельности IBS является реализация проектов в сфере управленческих и информационных технологий для крупных российских и международных корпораций и государственных организаций.

Представленный в данном разделе материал основан на докладе Григория Ципермана на семинаре «Интеллектуальная корпорация» 21 сентября 2004 в г. Москва и обобщает его практический опыт работы в должности главного технолога компании IBS (приводится с минимальным редактированием исходного материала автора).

Цели

Цели создания базы знаний (БЗ) компании IBS

Основной целью создания БЗ Компании является сохранение инвестиций компании в интеллектуальный капитал путем:

- Создания инструмента отторжения опыта работы от персонала с целью обеспечения профессиональной устойчивости компании,
- Обеспечения постоянного роста качества работ, на основе профессионального роста сотрудников компании,
- Обеспечение снижения издержек на выполнение стандартных работ за счет повторного использования наработанных материалов.

Цели использования БЗ

Различные специалисты компании имеют разные цели использования БЗ:

- *Руководитель компании:* Оценка готовности, востребованности и ресурсного состояния продуктовой линейки,
- *Руководитель проекта:* Оптимальная организация проектных ресурсов,

- *Продавец*: Получение сведений о продуктах, проектах и экспертах компании,
- *Специалист*: Получение сведений о лучших решениях,
- *Стажер*: Освоение продуктов и технологий компании,

Кроме указанных выше целей БЗ стоит рассматривать как инструмент позиционирования сотрудников в компании, демонстрирующий их реальные профессиональные достижения, а также как механизм, реализующий каналы информирования профессионального сообщества компании о работах коллег.

В качестве *культурологических принципов* создания БЗ были выделены: использование апробированных решений (не изобретать велосипед); анализ и обобщение опыта; развитие собственной экспертизы; указание источника (отказ от плагиата); существование в интеллектуальной среде компании.

Основные информационные ресурсы БЗ:

- Эксперт – субъект (персона или организационная единица), обладающий знаниями, умениями и навыками в продвижении и реализации продуктов компании,
- Продукт – вещественный или нематериальный результат проектов и работ (предмет, научное открытие, идея),
- Проект – ограниченная во времени и по ресурсам (людским и финансовым) деятельность компании по решению задачи (задач) заказчика,
- Глоссарий – собрание употребляемых в деятельности компании терминов с их толкованием (определением).

Концептуальная модель БЗ представлена на рис. 11.1.

Для описания документов в базе используется профиль документа, в котором присутствует рецензия на размещенный документ, а также его оценка (рис. 11.2).

Оценки документа:

1. Не рекомендуется использовать документ
2. В документе нет интересного материала
3. В документе встречается интересный материал
4. Документ можно брать в качестве образца
5. Документ относится к лучшим практикам компании

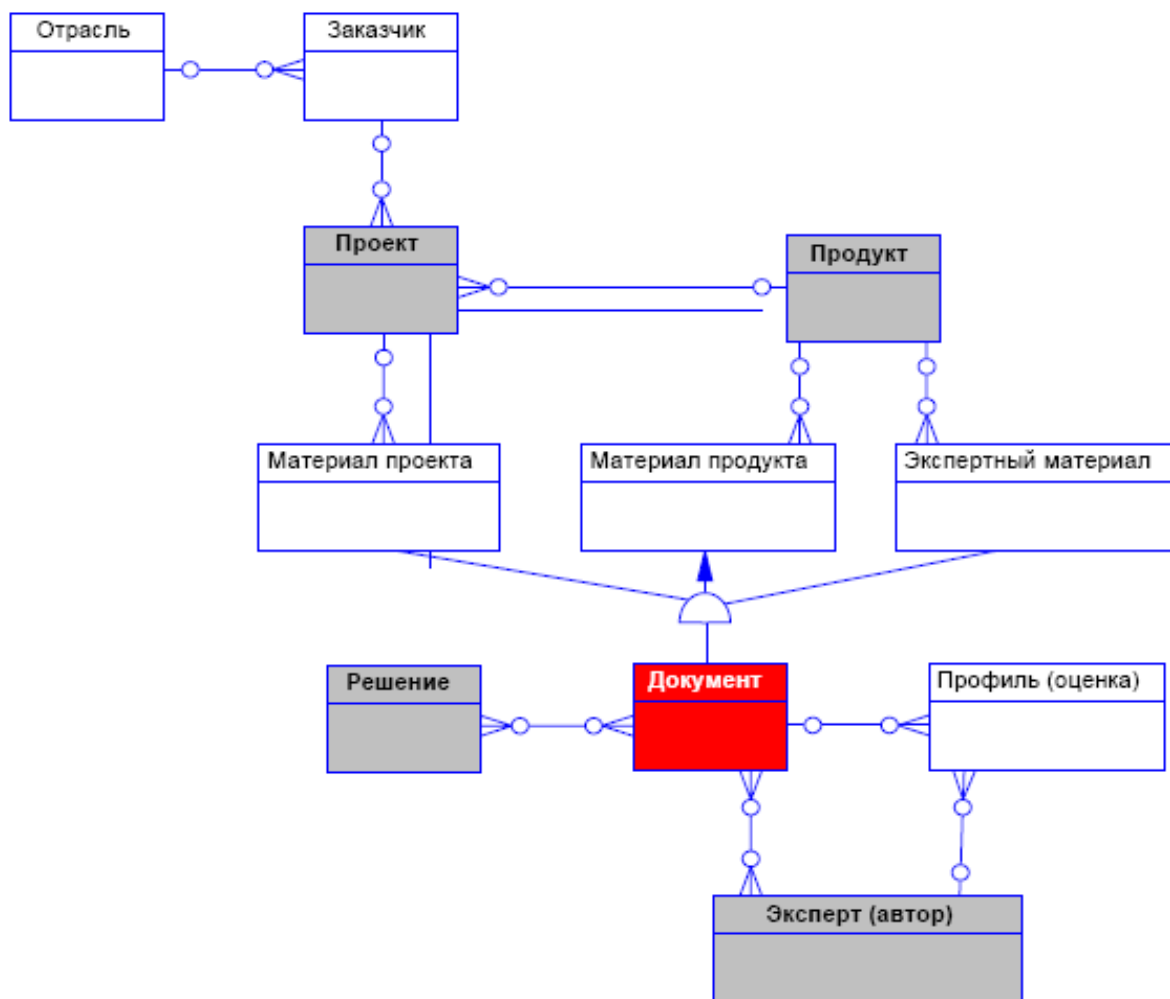


Рис. 11.1. Концептуальная модель БЗ

ОЦЕНКА ДОКУМЕНТА

Информация о документе

| | |
|-------------|--|
| Ваша оценка | 4 |
| Рецензия | Система регистрирует факт попытки просмотра документа, при этом, если просмотр осуществляется пользователем первый раз, система создает для пользователя задачу по оценке просмотренного |
| сохранить | |

Рис. 11.2. Профиль документа в БЗ

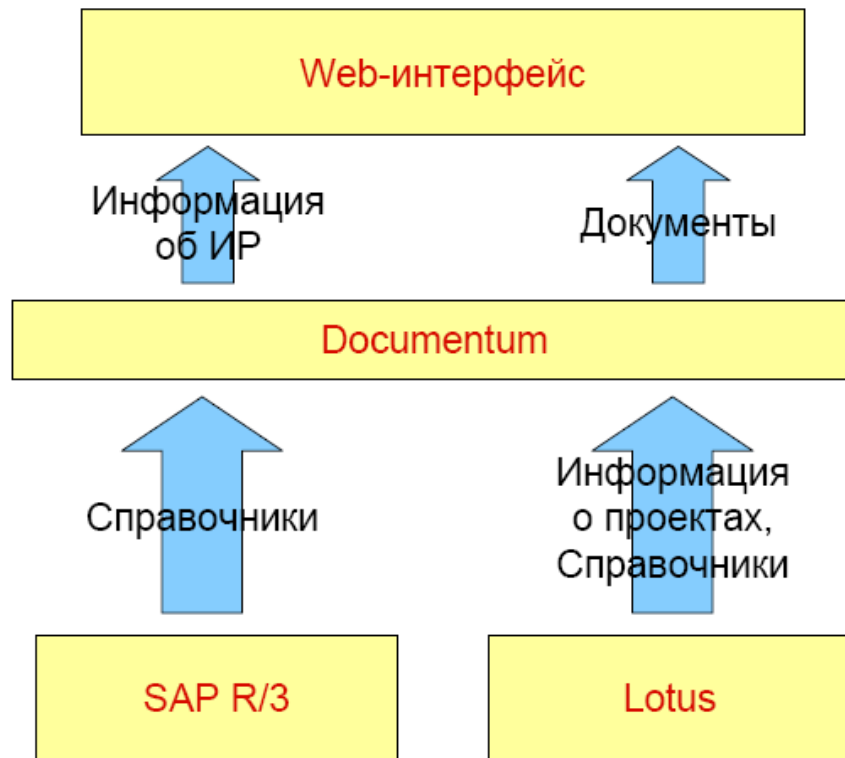


Рис. 11.3. Архитектура базы знаний

Размещение пакета документов в БЗ и атрибутирование (привязку метаданных) делает менеджер проекта, служба качества согласовывает материалы и метаданные, а куратор проекта утверждает.

Режимы работы с БЗ:

- Навигация – выборка объектов БЗ при помощи просмотра различных деревьев, представляющих содержимое БЗ (рис. 11.4),
- Поиск – выборка объектов БЗ (проект, продукт, документ) при помощи полнотекстового, атрибутивного (рис. 11.5) и/или атрибутивно-полнотекстового поиска,
- Оповещение – автоматическое информирование о появлении новых материалов,
- Статистика – представление статистических данных о содержимом БЗ.

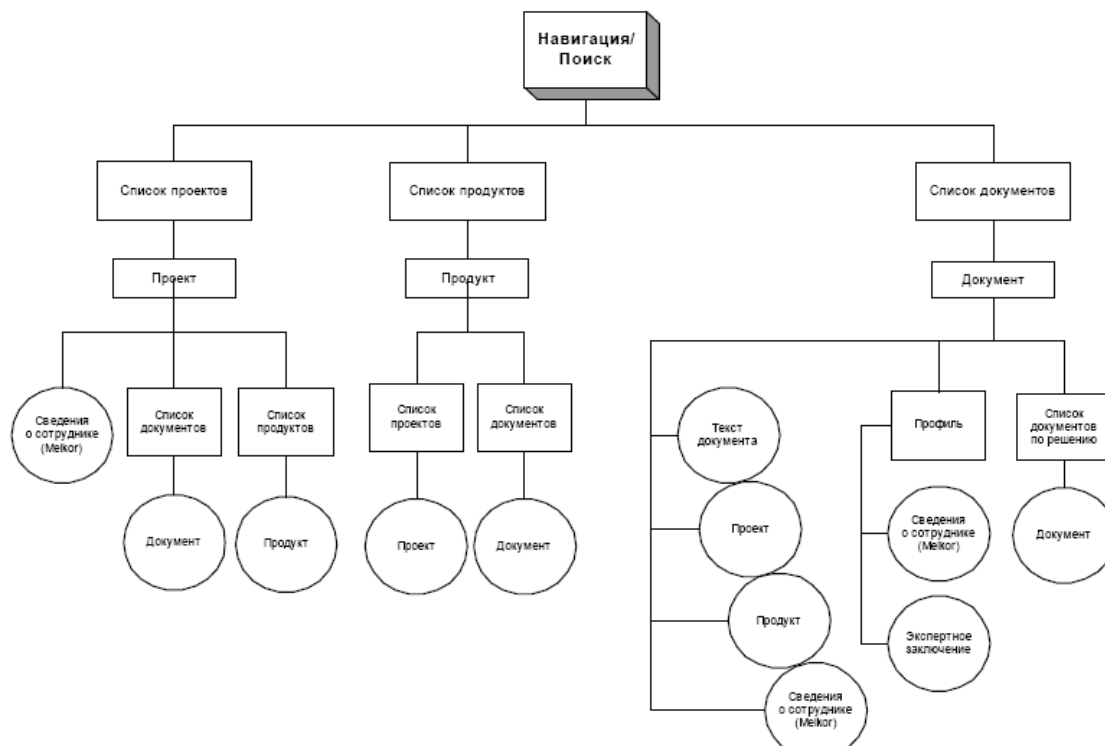


Рис. 11.4. Общая схема навигации по БЗ



Рис. 11.5 Атрибуты документа в БЗ, позволяющие проводить поиск

Для оценки использования и востребованности БЗ используются соответствующие показатели. Одним из наиболее значимых сводных показателей является *показатель активности работы с документом* =

$$APD = Ca \cdot A + Ce \cdot E + Cr \cdot R + Cs \cdot S, \text{ где:}$$

A (количество просмотров документа),

E (количество оценок документа),

R (количество рецензий документа),

S (количество ссылок на документ).

$C_a = 1, C_e = 10, C_r = 30, C_s = 100$ – веса вышеуказанных частных показателей;

Показатель активности работы с документом используется для морального стимулирования (авторский рейтинг востребованности материалов, рейтинг востребованности материалов подразделений),

материального стимулирования (конкурс на самый востребованный материал), для аттестации сотрудников, а также как критерий чистки БЗ.

Также для оценки использования БЗ используется профиль материала:

- Количество профилированных менеджером материалов – показатель знания менеджером предметной области деятельности компании,
- Инструмент менеджера для постановки заданий на разработку документов,
- Аттестационный показатель менеджера.

Безопасность БЗ достигается через анализ основных источников угроз (рис. 11.6, табл. 11.1) и соответствующее управление рисками.

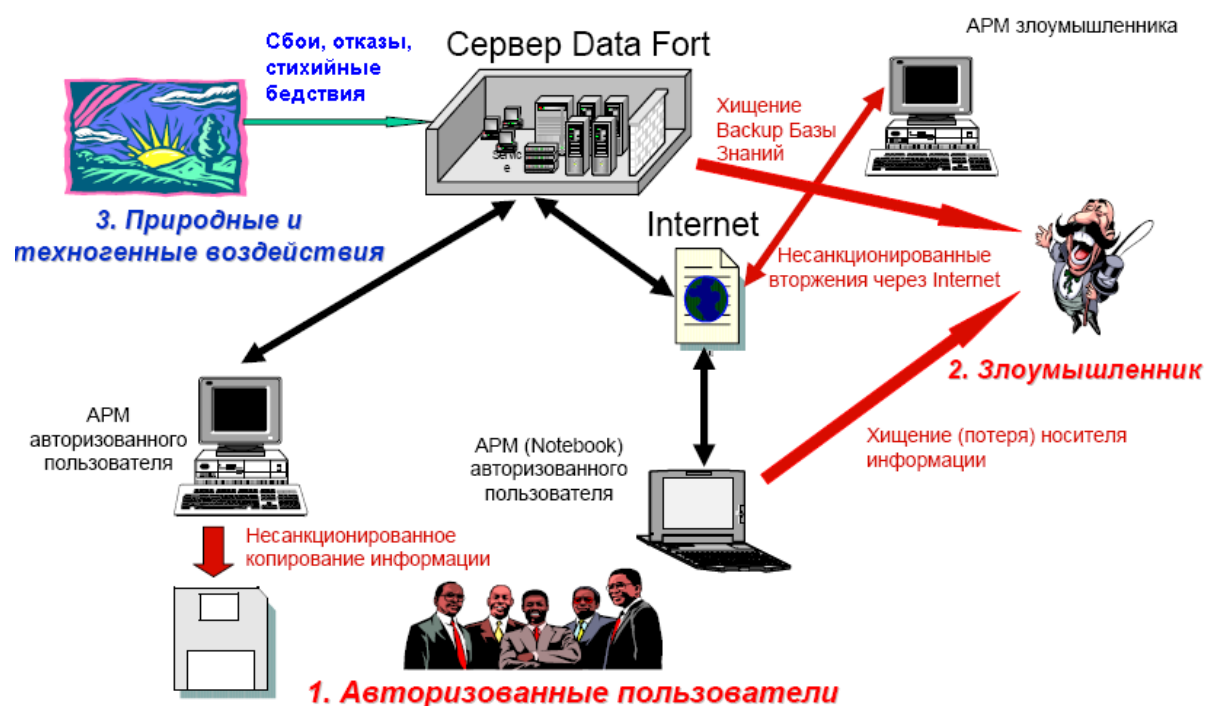


Рис. 11.6. Направления доступа к БЗ и основные источники угроз

Таблица 11.1

Анализ основных угроз и рисков

| Степень опасности | Источник угрозы | Характеристика угрозы | Последствия (риски) |
|--------------------|-------------------------------------|--|---|
| Высшая | Авторизованный пользователь | Несанкционированное копирование информации | Получение третьей стороной всей информации из Базы знаний |
| | Злоумышленник | Кража Backup Базы знаний | |
| Высокая | Авторизованный пользователь | 1. Утрата или передача аутентификационной информации для доступа к Базе знаний 2. Потеря Notebook'a | Получение третьей стороной значительной части информации из Базы знаний |
| | Злоумышленник | Кража Notebook'a | |
| Значительная | Природные и техногенные воздействия | Разрушение основной информации Базы знаний | Уничтожение Базы знаний (необходимость восстановления) |
| Существенная | Природные и техногенные воздействия | Разрушение копии информации Базы знаний | Необходимость восстановления копии Базы знаний |
| | Злоумышленник | Перехват информационного обмена при работе с удаленных терминалов | Получение третьей стороной некоторой части информации из Базы знаний |
| Менее существенная | Злоумышленник | Вирусные и хакерские атаки | Нарушение нормального режима работы Базы знаний |

Развитием и сопровождением БЗ занимаются главный технолог, координатор знаний, менеджер знаний и технолог.



Рис. 11.7 Организационная структура отдела, сопровождающего БЗ

Как видно из рис. 11.8 экономический эффект БЗ существенно возрастает с увеличением числа выполняемых проектов, которые создают новые и используют существующие знания.

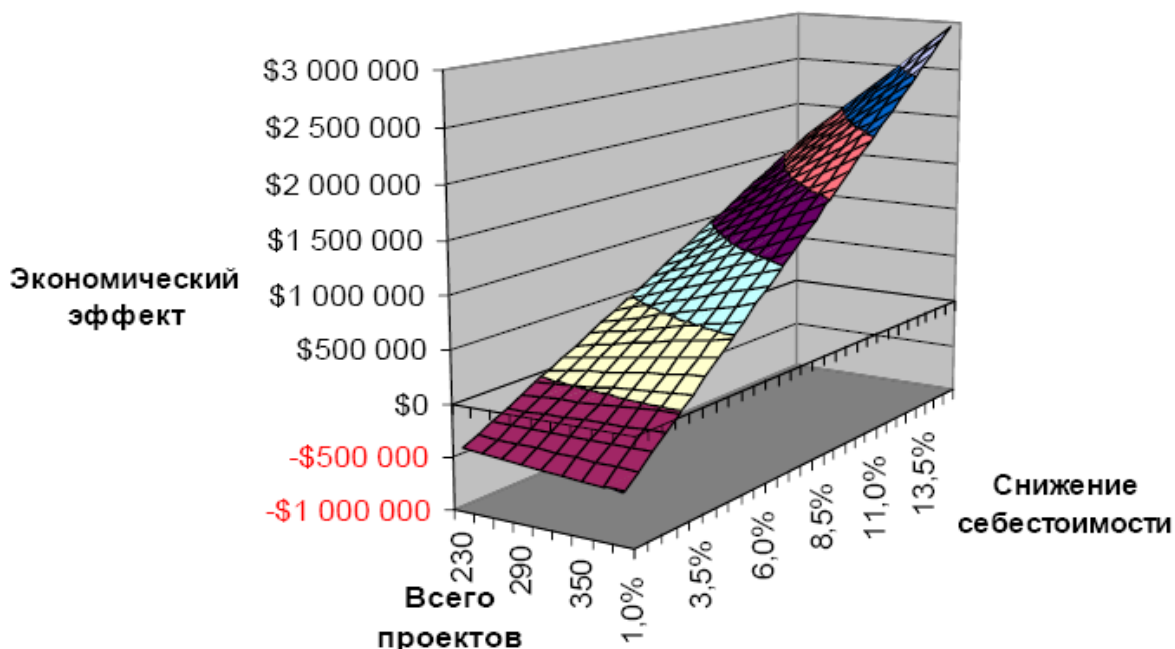


Рис. 11.8. Оценка экономического эффекта БЗ

11.2. Создание базы знаний проектов по управлению знаниями (на основе онтологий и рассуждений по прецедентам)

Обмен и повторное использование знаний и передового опыта может существенно повысить результативность, обеспечить экономию и снизить риски в процессах

- совершенствования деятельности (реинжиниринг бизнес-процессов, реструктуризация);
- разработки / конструирования новых решений (законотворческих, организационных, информационно-технических);
- выбора альтернатив;
- планирования.

Обмен и передача знаний в рамках одной организации обычно принимает форму создания баз знаний, обеспечивающих накопление и повторное использование знаний. Основная проблема, возникающая при попытках повторного использования знаний и передового опыта, заключается в сложности выбора успешных решений, опыт которых применим к конкретной решаемой задаче, поэтому обеспечение обмена и повторного использования знаний и передового опыта основано на *систематизации* успешных решений. Для систематизации успешных решений в практике бизнеса и государственных организаций традиционно используются справочные модели. Справочные модели позволяют осуществлять поиск и отбор успешных решений. Далее представлены существующие примеры из бизнеса и государственного управления.

Рассмотрим технологию разработки справочной модели для обмена опытом на примере задачи обмена знаниями и опытом в области внедрения систем управления знаниями⁵⁸.

Внедрение систем управления знаниями в организациях может наталкиваться на разнообразные барьеры – технические, организационные, культурные. Один из способов преодоления этих барьеров – обучение на основе опыта внедрения систем управления знаниями в других

⁵⁸ Hefke M., Jager K., Abecker A. Best Practice Cases for Knowledge Management and Their Portability to Other Organisations, 6th International Conference on Knowledge Management, 6-8 September, 2006, Graz, Austria.

организациях. Это может быть реализовано путем поиска и анализа успешных прецедентов (УП, best practice cases) внедрения систем управления знаниями и последующей адаптации таких УП к собственной организации.

Данный пример основан на распространенном подходе работе со знаниями – рассуждениям по прецедентам (case-based reasoning, CBR). CBR – это метод формирования умозаключений, основанный на поиске и анализе случаев формирования подобных умозаключений в прошлом. Разумеется, такие умозаключения не являются достоверными и требуют верификации. Проверка корректности умозаключения может являться частью CBR-процесса.

Практическая проблема такого подхода заключается в том, что описания существующих УП (обычно представленных в виде success stories в учебниках или на веб-сайтах) плохо структурированы, обычно несопоставимы и, соответственно, их применимость с конкретной организации очень сложно оценить. Для решения этой проблемы предлагается использовать справочную модель, которая позволит структурировать УП. Использование справочной модели позволит организациям разрабатывать собственные решения на основе анализа структурированного и сопоставимого опыта других организаций.

Технология разработки справочной модели состоит из следующих этапов:

1. Выявление значимых признаков для описания и выбора УП;
2. Оценка и анализ выявленных признаков УП;
3. Разработка справочной модели для описания и выбора УП.

В процессе *выявления значимых признаков* использовались методические материалы, учебники, часто-посещаемые Интернет-страницы тематических ресурсов, результаты опроса.

В результате был получен полный список признаков для описания УП.

На основе результатов опроса были проведены *оценка и анализ признаков*, в результате которых был получен отранжированный список признаков (табл. 11.2) и установлены некоторые взаимосвязи между ними:

Таблица 11.2

**Индикаторы для описания и выбора успешных решений
(накопленного опыта)**

| Признак / Критерий | Возможные значения | Важность |
|--|--|-----------------|
| Проблемы в области знаний | Напр., идентификация знаний, приобретение знаний, хранение знаний | высокая |
| Методы и инструменты | Например, интеллект-карты (mind maps), желтые страницы, SWOT-анализ, storytelling, рынок знаний, сценарии... | высокая |
| Цели в области знаний | Напр., улучшение обмена знаниями, Создание новой продукции... | высокая |
| Функциональные области и процессы | Финансы, управление персоналом, маркетинг | высокая |
| Количественные выгоды | Рост выручки, снижение затрат | высокая |
| Уровень зрелости | Зависит от выбранной модели | высокая |
| Длительность внедрения | Натуральное число | высокая |
| Уровень затрат на внедрение | Натуральное число | средняя |
| Размер компании | Натуральное число | средняя |
| Тип организационной структуры компании | Например, матричная | средняя |
| Уровень воздействия | Например, команда, отдел, предприятие | средняя |
| Использованное программное обеспечение | Например, Lotus-notes, технологии semantic-web | средняя |
| Количество вовлеченных специалистов по управлению знаниями | Натуральное число | низкая |
| Размер выручки компании | Натуральное число | низкая |
| Организационно-правовая форма компании | Например, ЗАО, АО, ООО | низкая |

На основе результатов оценки и анализа признаков была разработана справочная модель, основанная на онтологии, для описания и выбора успешных решений. Онтология поддерживает описание понятий, иерархии понятий и свойств, которые могут быть либо атрибутами (для более детального описания понятий), либо связями между понятиями.

Полученная справочная модель разделяет признаки, описывающие общий профиль компании (размер компании, отрасль, выручка, используемые технологии) и специфичные для предметной области – в данном случае для управления знаниями (проблемы управления знаниями, цели в области знаний, методы и инструменты управления знаниями). Все признаки представляются в онтологии либо в виде понятий, либо в виде атрибутов (рис. 11.9).

Кроме того, признаки структурированы с использованием иерархии понятий (рис. 11.9). Например:

- проблемы в области знаний делятся на организационные, технологические и связанные с культурой компании;
- цели в области знаний делятся на стратегические, тактические и операционные цели в области знаний;
- функциональные области и процессы на верхнем уровне разделяются на операционные и обеспечивающие и дальше классифицируются с помощью справочной классификации процессов APQC (The American Productivity & Quality Center, <http://www.apqc.org/>).

Установлены взаимосвязи между признаками, например, признак для описания и выбора успешных решений «проблемы в области знаний» связан с признаком «решения» типом связи «имела решение» (рис. 11.10).

На следующем шаге были определены измеримые и сопоставимые диапазоны допустимых значений для каждого признака (что позволило обеспечить автоматизированный поиск успешных решений).

После разработки справочной модели для обмена знаниями необходимо обеспечить регулярно-воспроизводимую работу с данной моделью и поддержание базы (репозитория) прецедентов, это требует внедрения в практику процесса работы со справочными моделями, представляющего собой замкнутый цикл, состоящий из 4-х этапов:

1. Поиск и выбор наиболее похожих прецедентов под возникшую

- проблему;
2. Повторное использование информации и знаний из найденного прецедента для решения возникшей проблемы;
 3. Оценка эффективности предложенного решения проблемы;
 4. Сохранение вновь возникшего успешного прецедента в базе (репозитории) для обмена знаниями и решения подобных проблем в будущем.

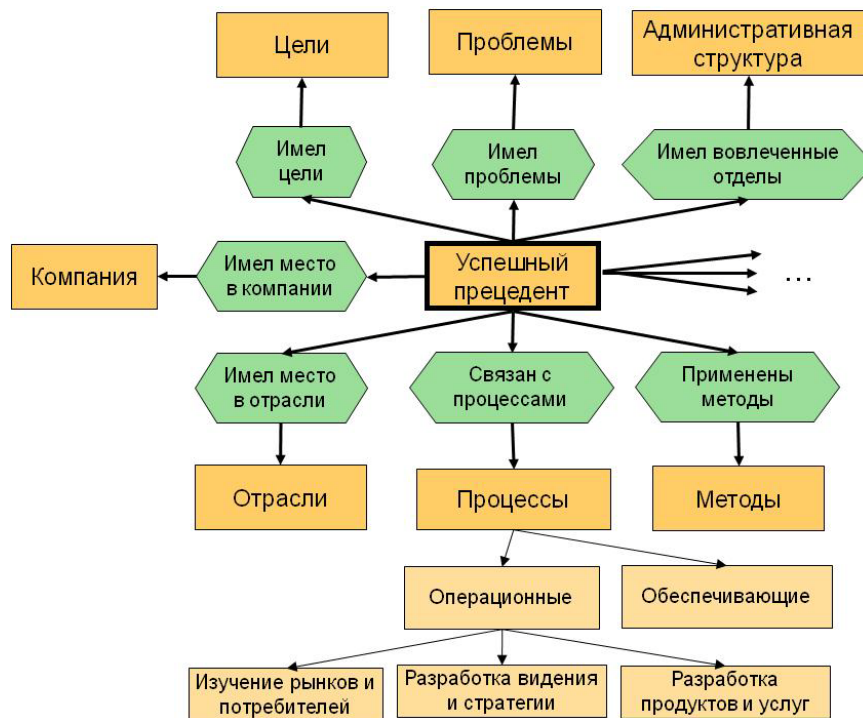


Рис. 11.9. Онтологическое представление признаков УП и их детализация

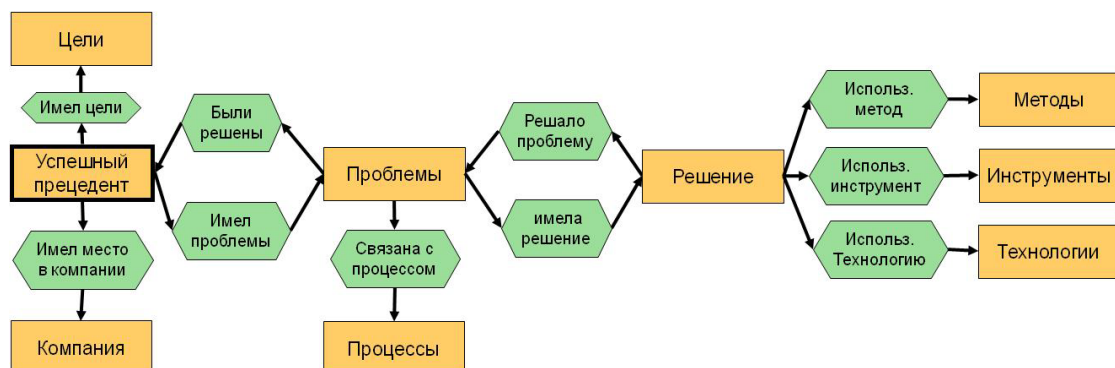


Рис. 11.10. Взаимосвязи между признаками УП

В качестве справочной модели для поиска и выбора успешных прецедентов выступает представленная выше онтологическая модель. При этом классы объектов модели будут выступать в роли признаков успешных прецедентов, экземпляры классов – в качестве значений признаков, отношения – будут описывать связи между признаками.

Описываемые с помощью справочной модели успешные прецеденты включают в себя с одной стороны успешное решение, а с другой стороны описание контекста его применения. В зависимости от решаемых задач состав успешного решения может варьироваться.

Для задачи внедрения управления знаниями, представленной в приведенном выше описании технологии разработки справочной модели, структура *успешного решения (отдельно или в совокупности)* приведена в табл. 11.3.

Таблица 11.3

Компоненты успешного решения

| Элемент решения | Возможная форма представления |
|---|---|
| 1. Вариант организационных изменений (процессы, распределение ответственности...) | модель организации (описание процессов и структур) |
| 2. Применяемая методика | текстовая документация |
| 3. Применяемый программный продукт и его технологическая поддержка | модель организации (элементы информационной архитектуры – системные компоненты, программные продукты,...) текстовая документация |
| 4. Применяемые способы изменения культуры компании | текстовая документация |

За основу возможной функциональной архитектуры системы поддержки работы со справочными моделями можно взять архитектуру, состоящую из следующих компонент:

1. База (репозиторий) прецедентов, систематизированная с использованием онтологии,

2. Модуль (web-based) для структурированного описания вновь возникающего прецедента, который не только позволяет описать новый прецедент для поиска подобных прецедентов, но и является некой базовой схемой для диагностики и аудита существующей ситуации в процессе

поиска решения,

3. Модуль сопоставления прецедентов, ищущий наиболее похожие прецеденты по описанию, сделанному на предыдущем этапе,

4. Модуль рекомендации, формирующий рекомендации по возможным решениям в существующем прецеденте, путем презентации наиболее похожих прецедентов, методов и решений,

5. Модуль обучения, который сохраняет повторно использованные и уточненные прецеденты в виде новых прецедентов в базу (репозиторий) прецедентов.

11.3. Технология обработки знаний в области инженерии организации

В настоящее время предприятия должны быть гибкими, быстро реагировать на изменения внешней среды, рынка и технологий, обеспечивать соответствие внутренним и внешним стандартам (ИСО, СММ, SOX...). Гибкость и адаптивность реализуются через непрерывные трансформации конструкции бизнеса, что подразумевает интегрированное изменение системы целей и показателей, политик, бизнес-процессов, организационной структуры, а также ИТ-решений. Такое интегрированное изменение требует согласованной работы сразу нескольких областей менеджмента: стратегического управления, управления эффективностью, управления бизнес-процессами, организационного проектирования, проектирования информационных систем и др. В ответ на вызовы времени для поддержки управления такими интегрированными изменениями совсем недавно на западе сформировалась и быстрыми темпами развивается отдельная дисциплина – инженерия организаций (enterprise engineering). Данная дисциплина основана на понимании организации – как целенаправленно спроектированной и реализованной на практике системы. Соответственно, ее перепроектирование может осуществляться систематическим и управляемым образом. В качестве основы для проектирования предлагается использовать инженерный подход, а точнее, системную инженерию, опирающуюся на создание точных моделей устройства организации. То есть целенаправленное изменение организации предполагает вначале изменение модели этой организации, а затем ее внедрение.

В основе инженерии организации лежит понятие архитектуры организации (enterprise architecture) – «описание компонент организации, связей их между собой и со средой». Поскольку понятие архитектуры организации изначально возникло в ИТ-сообществе (было предложено Джоном Захманом), то в ней стало принято разделять ИТ-архитектуру, включающую описание информационных систем, информации и технических средств, и бизнес-архитектуру (условно – всё, что не относится к области ИТ). Однако управление другими базовыми ресурсами организации, подобно информационным, целесообразно осуществлять с помощью архитектурного подхода, рис. 11.11.

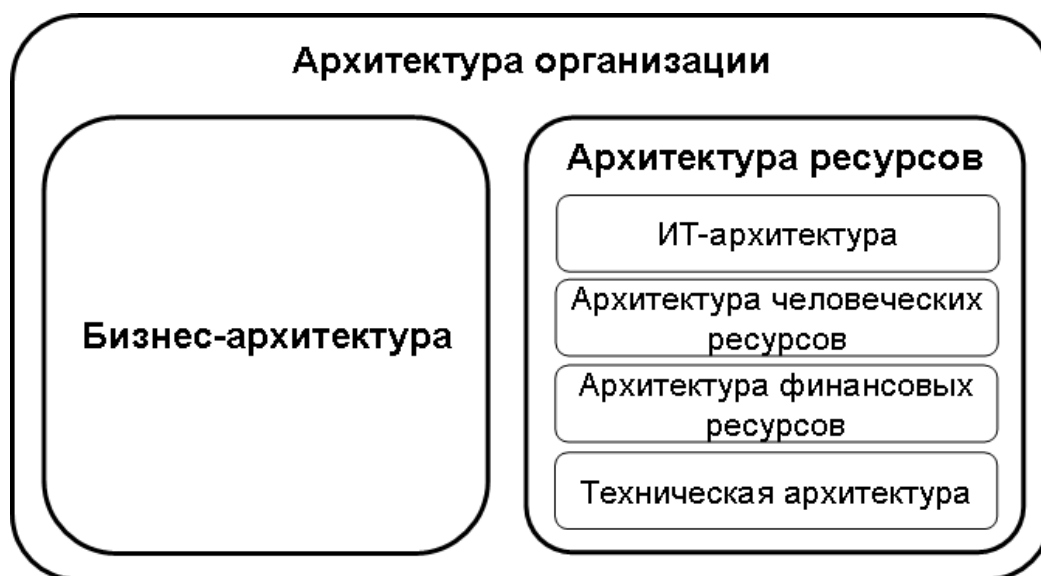


Рис. 11.11. Компоненты архитектуры организации

Для инженерии организации необходимо использовать программные средства, подобные CASE-средствам, которые используются при проектировании информационных систем. Такие программные средства позволяют создавать модели организации, которые применяются как для анализа, так и для синтеза архитектуры организации. Требования к данным средствам есть следствие процессов преобразования знаний в рамках работ по инженерии организаций.

В проекте по инженерии организации бизнес-аналитики (либо сотрудники компании, либо внешние консультанты) совместно с экспертами в предметных областях (руководителями отделов и

менеджерами) *создают* знание об организации деятельности, *накапливают* и *распределяют* его среди остальных сотрудников компании.

Применение знаний об организации деятельности происходит во всех бизнес-процессах компании – основных, управленческих и обеспечивающих. Организационные роли сотрудников, применяющих эти знания, варьируются от рабочего до директора. В соответствии со схемой предложенной Стромайером и рассмотренной в разделе 5, процесс преобразования знаний об архитектуре организации может быть представлен в следующем виде (рис. 11.12):

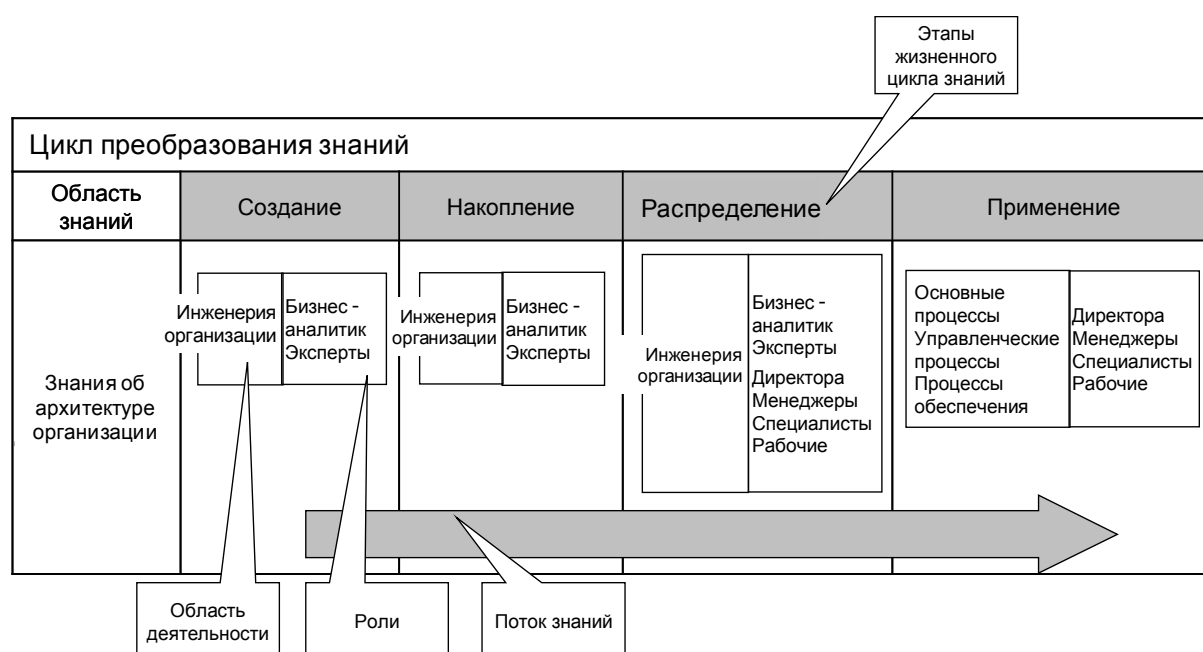


Рис. 11.12. Процессы преобразования знаний об архитектуре организации

Наиболее важной характеристикой этого жизненного цикла знаний является наличие разных организационных ролей, вовлеченных в него, в особенности на этапе передачи знаний. В процессе распределения знаний бизнес-аналитики передают знания сотрудникам из разных предметных областей, находящимся на разных уровнях иерархии, и главным фактором его успешного усвоения сотрудниками является способ представления знаний. Традиционно способ представления знаний потребителям определяется тем, в каком виде знания создаются и накапливаются.

Основными создателями (интеграторами) знаний об организации деятельности являются бизнес-аналитики. Обычно это люди, которые

обладают развитыми компетенциями в области организационного управления и системного анализа. Они обладают детальным пониманием бизнеса с разных точек зрения и работают с различными бизнес-объектами и их взаимосвязями (организационные единицы, функции, процессы, цели и т. д.). Наиболее удобным средством представления знаний для бизнес-аналитиков являются диаграммы в разных нотациях (IDEF0, eEPC и т. п.).

С другой стороны использовать полученные знания об организации деятельности должны все сотрудники компании и большая часть персонала обладает недостаточными компетенциями в области организационного управления. Поэтому для восприятия знаний в области организационного управления сотрудникам требуются привычные средства представления таких знаний – должностные инструкции, положения о подразделениях, регламенты процессов, выполненные либо в текстовом виде, либо в табличном. Кроме того, часто регламентирующие документы должны быть выполнены в соответствии с требованиями национальных, отраслевых или корпоративных стандартов. Именно эти регламентирующие документы смещают приоритеты от неформальных методов управления к регулярному менеджменту.

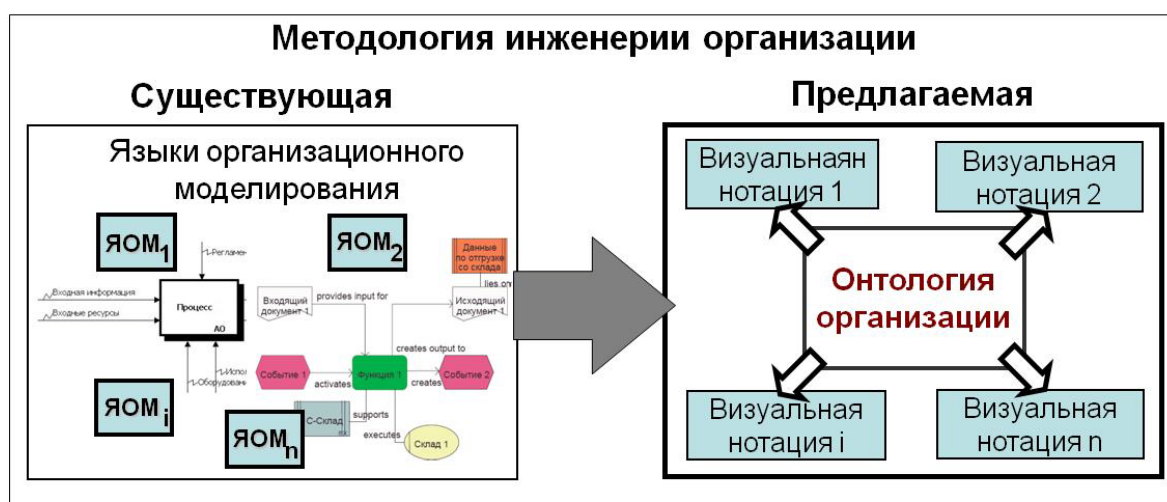
Таким образом, необходимость поддержки взаимодействия людей, говорящих на разных языках, предопределяет требования к методам и инструментам инженерии организации. В настоящее время основной проблемой формализации и представления знаний об организации является многообразие слабо согласованных между собой языков организационного моделирования, которые опираются на различные графические нотации, создаваемые для отражения многочисленных аспектов деятельности предприятия. Такое многообразие языков и гигантские объемы графических документов приводят к сложности управления масштабными моделями организации, особенно в условиях дефицита квалифицированных специалистов-аналитиков. Для решения данной языковой проблемы может использоваться онтология.

Если в традиционных системах модель организации – это диаграммы, созданные с использованием различных графических языков организационного моделирования, то в автоматизированной системе инженерии организации на основе онтологии модель – это наполненная экземплярами онтология, а заменой языкам организационного

моделирования выступают визуальные нотации, семантика которых задается онтологией организации (рис. 11.13).

Онтология организации выполняет следующую роль в системе:

- Задает структуру содержания модели организации;
- Используется для настройки структуры содержания и наполнения выходных документов и диаграмм;
- Представляет информацию на разных уровнях обобщения / детализации;
- Повышает эффективность коммуникации между сотрудниками.



ЯОМ – язык организационного моделирования

Рис. 11.13. Онтология – основа методологии инженерии организации

Процесс преобразования знаний в автоматизированной системе инженерии организации на основе онтологии представлен на рис. 11.14.

В данном процессе используются различные модели и методы, рассмотренные в пособии, например, на этапе создания знаний используется структурированный ввод информации (см. раздел 6.1), на этапе накопления – визуальный язык классификаторов и проекций для структурирования и внутреннего представления знаний (см. раздел 7.1), а на этапе распределения формируются пакеты знаний (см. раздел 8.4).

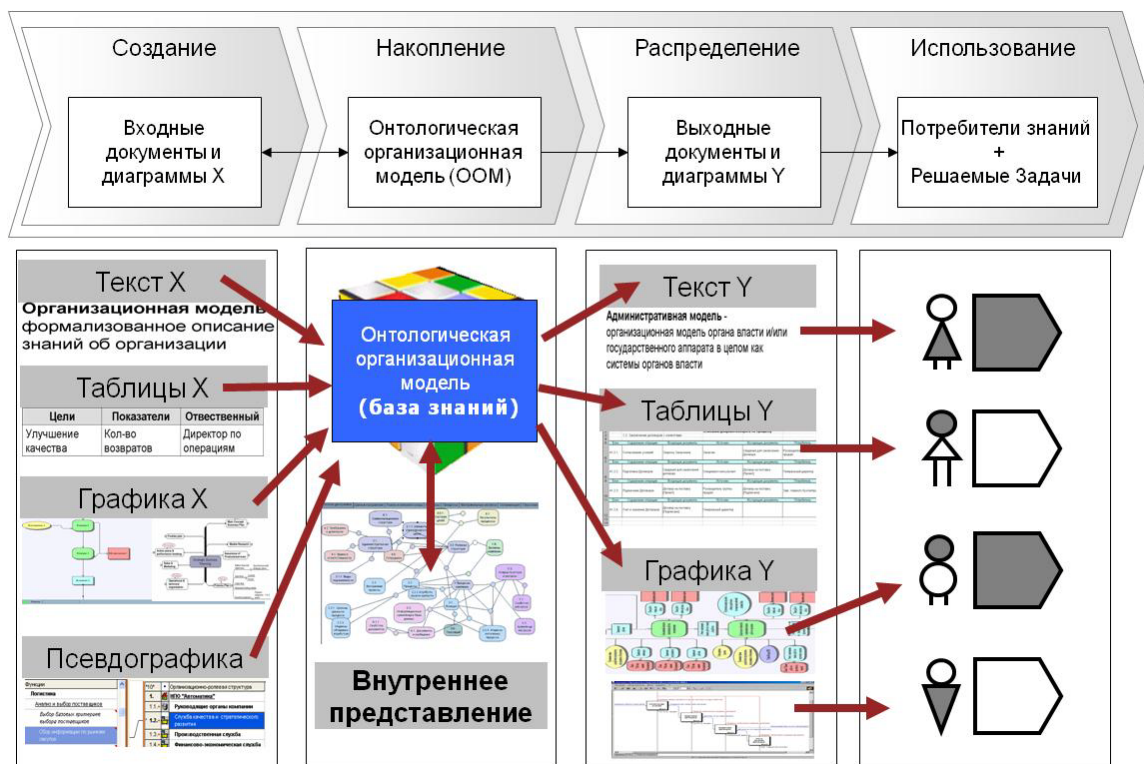


Рис. 11.14. Преобразование знаний в процессе инженерии организации

Программа ОРГ-Мастер – инструмент поддержки инженерии организации на основе онтологии

Методология инженерии организации на основе онтологии реализована в программном инструменте ОРГ-Мастер. Модели и методы инженерии организации и представления знаний совместно с реализующим их программным инструментом составляют автоматизированную систему поддержки инженерии организации.

Входом автоматизированной системы поддержки инженерии организации являются слабоструктурированные представления об организации, ее целях (разного уровня), показателях их достижения, функциях и ответственности за их исполнение, процессах и ответственности за их результаты (см. рис. 11.15). Изначально эти представления находятся в разрозненных документах и памяти сотрудников организации. Автоматизированная система переводит слабоструктурированные представления в формализованную модель организации «как есть». Далее она позволяет специалисту по инженерии организации проанализировать системные несоответствия (с учетом

нормативной модели) и создать модель перспективного состояния организации и необходимые промежуточные модели, позволяющие перейти в новое состояние. На основе модели система формирует организационные документы (положения о подразделениях, процедуры выполнения процессов, должностные инструкции...), а также аналитические отчеты для лиц, принимающих решения, ориентированные на проверку соблюдения различных нормативных правил построения оптимальной организации.



Рис. 11.15. Входы и выходы для автоматизированной системы поддержки инженерии организации

Аналитики рынка информационных технологий (Gartner, Forrester, ...) относят программные инструменты, подобные ОРГ-мастеру, к пересекающимся классам Enterprise Architecture (EA) tools (инструменты архитектуры организации) и Business Process Analysis (BPA) tools (инструменты анализа бизнес-процессов).

В общей информационной системе управления предприятием среда инженерии организации исполняет роль средства настройки:

- системы управления исполнением процессов (workflow) и проектов,
- системы управления ресурсами (ERP, MRP, CRM, SCM и т. п.),
- системы мониторинга и анализа показателей деятельности (PMS, BAM).

Применение методологии и инструмента ОРГ-Мастер

Представленная методология и инструмент инженерии организации имеет 10-тилетнюю историю применения в проектах по организационному развитию и совершенствованию процессов коммерческих организаций и органов государственной власти.

ОРГ-Мастер используется в коммерческих организациях России и стран СНГ, при этом размер клиентов-пользователей ОРГ-Мастера варьируется от небольших компаний до холдинговых структур с количеством работников до 10 000 человек.

Методология и инструмент инженерии организации ОРГ-Мастер были использованы при выполнении более 50 проектов по управленческому консультированию. Примеры организаций: Киришская ГРЭС, ИЛИМ, Северо-Западное пароходство, Группа предприятий ГОТЭК, Концерн «Силловые машины», ЗАО «Евросиб», ОАО «Иркутскэнерго», ОАО «Газаппарат», «АСТРА Холдинг», Торговый дом «Петровский», Холдинг ПЕКАР и др.

Например, в группе предприятий ГОТЭК были достигнуты следующие цели и задачи:

Цели:

Совершенствование и формализация управления деятельностью группы компаний

Задачи:

- Идентификация, описание и совершенствование бизнес-процессов;
- Организационное разделение процессов между управляющей компанией, центром предоставления услуг и предприятиями;
- Установление связи между стратегией и деятельностью компании;
- Внедрение системы управления проектами;
- Разработка организационной документации – процедур выполнения бизнес-процессов, положений о подразделениях, должностных

инструкций, документации для сертификации системы менеджмента качества.

В процессе инженерии организации в группе ГОТЭК в системе ОРГ-Мастер разрабатывались следующие документы (неполный перечень):

Нормативно-регламентирующие документы:

- «Стратегия функциональной системы»,
- «Паспорт бизнес-процесса»,
- «Концепция бизнес-процесса»,
- «Регламент бизнес-процесса»,
- «Рабочая инструкция сотрудника».

Аналитические документы:

- «Анализ состава стратегических показателей»,
- «Показатели ответственности и мотивационные показатели»,
- «Контрольный отчет по описанию процесса».

Представленная методология и инструментарий позволяют получить следующие системные эффекты:

- Усиление динамических способностей организации по воплощению стратегий,
- Снижение времени реакции организации на изменения внешней среды,
- Рост управляемости и контроля на всех уровнях компании,
- Накопление и способность передачи организационных знаний,
- Рост стоимости компании – за счет высокой оценки указанных выше факторов.

11.4. Стандартизация ремонтных операций

В данном разделе приводится статья из газеты «Котласский бумажник» № 44 от 27 ноября 2010 года. Статья посвящена проекту по управлению знаниями, который идет на Котласском целлюлозно-бумажном комбинате⁵⁹ при участии автора пособия.

⁵⁹ http://ru.wikipedia.org/wiki/Котласский_целлюлозно-бумажный_комбинат

«Можно ли знания сделать интеллектуальным капиталом?»

В рамках реформирования системы технического обслуживания и ремонта (ТО иР) стартовал проект «Управление знаниями». Его разработчики уверены: профессиональные знания – это интеллектуальный ресурс, при помощи которого предприятие тоже может получать прибыль. Главное – научиться его грамотно использовать.

Не первый год на основном производстве и в ремонтной службе говорят о том, что вместе с опытными рабочими, которые выходят на заслуженный отдых, утрачиваются ценные знания по управлению технологическими процессами или ремонту уникального оборудования. Чтобы сохранять и передавать «интеллектуальный багаж», на предприятии действует «Школа профессиональной преемственности», но и она не в силах сделать знания достоянием многих поколений рабочих. Отношения «наставник-ученик» носят индивидуальный, неформализованный характер и, как правило, не документируются.

Консалтинговая компания «Бизнес Инжиниринг Групп» (БИГ) предложила управлять ключевыми знаниями (пока в области ТОиР), чтобы предоставлять их «в требуемом виде и количестве в нужном месте в установленное время». По словам специалиста БИГ Дмитрия Кудрявцева, идея родилась в начале года, а летом определились, в каком формате ее воплощать.

– Совместно со специалистами филиала Группы «Илим», участниками проекта ТОиР было решено заняться стандартизацией операций, – комментирует Дмитрий. – Понятно, что на производстве таких операций тысячи. Для «обкатки» мы выбрали несколько часто повторяющихся операций, например замену подшипника, а также редкие и крайне важные для производства, к примеру, ремонт гидропривода DD-промывателя. На первом этапе проекта составлено 16 инструкций, в основном для работы слесарей по ремонту и электромонтеров.

Чем еще, кроме сохранения и передачи знаний опытных специалистов, может быть полезна стандартизация? Во-первых, она должна повысить скорость и качество выполнения операций. Во-вторых, снизить риск допустить ошибку. Это особенно важно для тех участков, где используется уникальное дорогостоящее оборудование.

В-третьих, хорошо составленная инструкция поможет не только в обучении новичков на производстве и учащихся ПУ, но и при повышении квалификации (разряда) специалистов. Дополнительно работы по стандартизации позволят выявить ключевых «носителей знаний» в области ТОиР – тех специалистов, чьи профессиональные знания и навыки крайне важны для предприятия.

Каким образом создаются инструкции? И насколько они отличаются от тех, что уже действуют в ремонтной службе? По словам разработчиков, основные компоненты новой инструкции – описание состава и последовательности действий, ключевые моменты по каждому действию, требования к инвентарю, ТМЦ, исполнителям, охране их труда, экологии. Примечательно, что новая инструкция должна быть проста и наглядна, как современная инструкция по эксплуатации бытовой техники. При описании действий используются иллюстрации, фото, схемы. А в примечаниях особыми значками выделены советы, как лучше выполнить то, или иное действие, в какой момент обратить внимание на безопасность, какие инструменты лучше использовать.

– Создан экспертный совет, в который вошли ведущие специалисты по ремонту, – поясняет Дмитрий Кудрявцев. – Они решают, какие операции наиболее важны, а значит, должны быть стандартизированы. Далее составители инструкций, среди них есть и старшие рабочие, по определенному шаблону делают описание операции. Затем к работе подключается специалист по стандартизации, который, встречаясь с экспертами, дорабатывает документ, готовит иллюстрации. К слову, некоторые фото удалось сделать только в капремонт, когда необходимое оборудование было остановлено и вскрыто. Готовые инструкции проходят проверку и согласование на экспертном совете.

Как пояснил специалист БИГ, «на бумаге» готовые инструкции доступны рабочим в мастерских. В электронном виде они будут размещены на корпоративном портале «Е-Мастер», где также хранятся описания процессов, обрабатываются заявки. Таким образом, вся исходная информация по ремонту сосредоточится в единой базе данных.

Проект «Управление знаниями» «обкатывается» в сервисах ПСБЦ и ПБ, КБП. Специалистам ясно, что база стандартизированных операций станет хорошим подспорьем в работе ремонтных служб. В настоящее

время начинается разработка следующей партии инструкций. По мнению участников проекта, на этом этапе важно, чтобы создаваемые материалы не лежали мертвым грузом, необходимо их активно изучать и использовать в работе.

11.5. Навигатор по методологии проектирования организаций – пример использования контекста для персонализации информации

В разделе 8.5. было указано на целесообразность использования контекстной информации для повышения качества поиска. Систематизация знаний предметной области в соответствии с контекстом их использования была выполнена автором пособия в рамках разработки *навигатора* по технологии проектирования организаций, разработанной в компании Бизнес Инжиниринг Групп (bigc.ru). Данная методология частично представлена в электронном виде зафиксирована во внутренней базе знаний упомянутой компании, а также в виде книги:

Григорьев Л. (редактор), Горелик С., Кудрявцев Д., Корышев И., Кислова В., Макаревич М., Каменская М., Якубовская Т, Заблоцкий А., Панфилов М. Менеджмент по нотам. Технология построения эффективных компаний. - М: Альпина Паблишерз, 2010.

В полном объеме *навигатор* представлен в самой книге, а также по электронному адресу <http://bigc.ru/publications/mn/navigator/>

Описание данного *навигатора* с точки зрения инженерии знаний было выполнено в работе: Kudryavtsev D. From context to knowledge: consecutive mapping ontologies and contexts // 6th International Conference on Knowledge Management, 6-8 September, 2006, Graz, Austria. – Graz: J.UCS, 2006. – P. 97-104.

URL: http://bigc.ru/publications/bigspb/km/context_knowledge.php

Далее в пособии приведены некоторые фрагменты *навигатора*.

Навигатор состоит из трех инструментов, которые дополняют друг друга:

Ролевой навигатор (КТО-навигатор) – позволяет выбрать материал для изучения в зависимости от Вашей роли в организации

Целевой навигатор (ЗАЧЕМ-навигатор) – позволяет выбрать материал для изучения под решение типовых задач (типовых элементов организационных проектов)

Смысловый навигатор (ЧТО-навигатор) включает:

- Расширенное оглавление – позволяет более ясно представить, о чем рассказывается в каждой главе,
- Поиск по ключевым словам – позволяет определить главу для чтения под конкретный информационный запрос,
- Глоссарий с указателем – уточняет поиск по ключевым словам, поясняя трактовку используемых в книге слов.

Смысловый навигатор здесь не рассматривается.

Ролевой навигатор (КТО-навигатор)

Данный навигатор позволяет выбрать материал для изучения в зависимости от Вашей роли в компании. Причем такой роли, которая, так или иначе, связана с организацией деятельности. Мы пытались подобрать материал под собирательные роли, который у нас возникли в результате работы с разными компаниями. Наименование роли, которое принято у Вас в компании может изначально не совпадать с нашим. Поэтому на Шаге 1 мы даем описание роли, которое позволит Вам узнать себя или те роли, которые Вас интересуют. Шаг 2 позволяет определить разделы для чтения в зависимости от выбранных ролей. Ваша деятельность может подпадать под несколько собирательных ролей, тогда Ваш материал – это сумма материалов для этих ролей.

Кому, как мы считаем, может быть адресована наша Книга:

Владельцам и акционерам компании, которые заинтересованы в устойчивом развитии и повышении стоимости компании, освоении новых подходов к выбору и реализации конкурентоспособных бизнес-идей.

Генеральному директору, несущему ответственность перед акционерами, как за результаты операционной деятельности, так и стратегические перспективы бизнеса, за успех сегодня и в будущем.

Директору по стратегическому развитию...

Директору (специалисту) по организационному развитию...

Директору (специалисту) по стратегическому контроллингу...

Директору (специалисту) по менеджменту качества...

Директору по персоналу...

Бизнес-аналитикам и специалистам по бизнес-моделированию...

Шаг 1: Узнайте себя ...

Собственник не только владеет бизнесом, но и... определяет направление развития бизнеса, участвует в разработке стратегии компании, ставит цели перед менеджерами, определяет права и полномочия менеджмента.

Генеральный директор управляет компанией и отвечает за рост стоимости компании, выбор и реализацию стратегии, выбор и внедрение методов управления, обеспечение прав акционеров и инвесторов

...

Шаг 2: Выберите разделы – см. табл. 11.4

Таблица 11.4

Выбор тем для изучения в зависимости от роли в организации

| Темы \ Роли | Собственник | Ген. директор | Директор по стратегическому развитию | Директор по организационному развитию | Специалист по организационному развитию | Функциональные руководители | Директор по персоналу | Директор по ИТ |
|--|-------------|---------------|--------------------------------------|---------------------------------------|---|-----------------------------|-----------------------|----------------|
| Часть I СОВРЕМЕННОЕ УПРАВЛЕНИЕ ОРГАНИЗАЦИЕЙ БИЗНЕСА | | | | | | | | |
| Глава 1 Бизнес-инжиниринг | ○ | | | ● | ● | | ● | ● |
| Глава 2 Корпоративная архитектура и ее модель | ● | ● | ○ | ● | ● | ○ | ○ | ● |
| Глава 3 Процессный подход к управлению | ○ | ● | | ● | ● | ○ | ○ | ● |
| Глава 4 Средства моделирования организационных систем | | | | ● | ● | | ● | ● |
| Глава 5 Система организационных управленческих регламентов | ○ | ● | | ● | ● | ● | ● | ● |
| Часть II ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ БИЗНЕС-ИНЖИНИРИНГА | | | | | | | | |
| ... | | | | | | | | |

Легенда: ● – важно, ○ – полезно

Целевой навигатор (ЗАЧЕМ-навигатор)

Данный навигатор позволяет выбрать материал для изучения под решение типовых задач (для выполнения типовых проектов).

Отметим, что мы сделали этот раздел по материалам обращений компаний в «Бизнес Инжиниринг Групп». Сохранены даже отдельные формулировки.

Шаг 1: Узнайте проблему / задачу...

Тема 1. Первичная регуляризация (Наведение порядка, Прояснение ситуации, Постановка регулярного менеджмента)

- «Надоело уже, у нас кавардак, мы героически с ним боремся». Если посмотреть на протоколы оперативок, как на вести с «передовой линии военных действий», то оказывается, что в компании поднимаются, но так и не решаются одни и те же вопросы. О стратегии в такой ситуации думать некогда, нужно думать о том, как конкретный заказ выполнить, потому что одно не поставили, другое забыли и т.д. Топ-менеджеры чувствуют что-то не то. У них ощущение, что они не справляются с потоком информации. Они хотят структурировать этот поток, задокументировать все необходимое, и таким образом навести порядок.
- Хочется регулярно выполнять то, что сделано один раз, а способом закрепления накопленного опыта являются организационные регламенты. При этом необходимо, чтобы бумаги отражали текущее положение вещей.
- ...

Тема 2. Систематизация работы с персоналом

- Перед менеджерами по персоналу во многих компаниях стоит задача работать с организационной структурой, с процессами, с должностными инструкциями. Они сталкиваются с рядом неудобств: если вводятся изменения в направлениях деятельности фирмы, набор функций работников меняется. Нужно быстро изменить должностные инструкции, параллельно нужно вручную вносить изменения в положение о подразделении и т.д. Неизбежны неувязки, ручная корректировка бумаг вызывает раздражение и хочется получить инструмент для облегчения работы. Да и вообще,

занимаясь этим регулярно они приобретают новую специализацию – «организационный менеджмент»

Тема 3. Внедрение системы менеджмента качества (СМК),

Тема 4. Подготовка к автоматизации,

Тема 5. Создание базы организационных знаний,

Тема 6. Постановка процессного управления,

Тема 7. Изменение структур под новую стратегию,

Тема 8. Уточнение и реализация стратегии компании,

Тема 9. Поиск новых идей организации бизнеса.

Шаг 2: Выберите разделы – см. табл. 11.5.

Таблица 11.5

Выбор тем для изучения в зависимости от проблемы/решаемой задачи

| Темы \ Проблемы/задачи | Первичная регуляризация | Систематизация работы с персоналом | Внедрение СМК | Подготовка к автоматизации | Создание базы организационных знаний | Постановка процессного управления | Изменение структуры под новую стратегию | Уточнение и реализация стратегии компании | Поиск новых идей организации бизнеса |
|--|-------------------------|------------------------------------|---------------|----------------------------|--------------------------------------|-----------------------------------|---|---|--------------------------------------|
| Часть I СОВРЕМЕННОЕ УПРАВЛЕНИЕ ОРГАНИЗАЦИЕЙ БИЗНЕСА | | | | | | | | | |
| Глава 1 Бизнес-инжиниринг | ○ | ○ | ○ | ● | ● | ● | ● | ● | |
| Глава 2 Корпоративная архитектура и ее модель | ○ | ○ | ○ | ● | ● | ● | ● | ● | ○ |
| Глава 3 Процессный подход к управлению | ○ | ○ | ● | ● | ● | ● | ● | ● | ○ |
| Глава 4 Средства моделирования организационных систем | ○ | ○ | ○ | ● | ● | ● | ● | ● | |
| Глава 5 Система организационных управленческих регламентов | ● | ● | ● | | ● | ● | ● | ● | |
| Часть II ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ БИЗНЕС-ИНЖИНИРИНГА | | | | | | | | | |
| ... | | | | | | | | | |

Легенда: ● – важно, ○ – полезно

12. НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ В ОБЛАСТИ ОНТОЛОГИЙ И УПРАВЛЕНИЯ ЗНАНИЯМИ

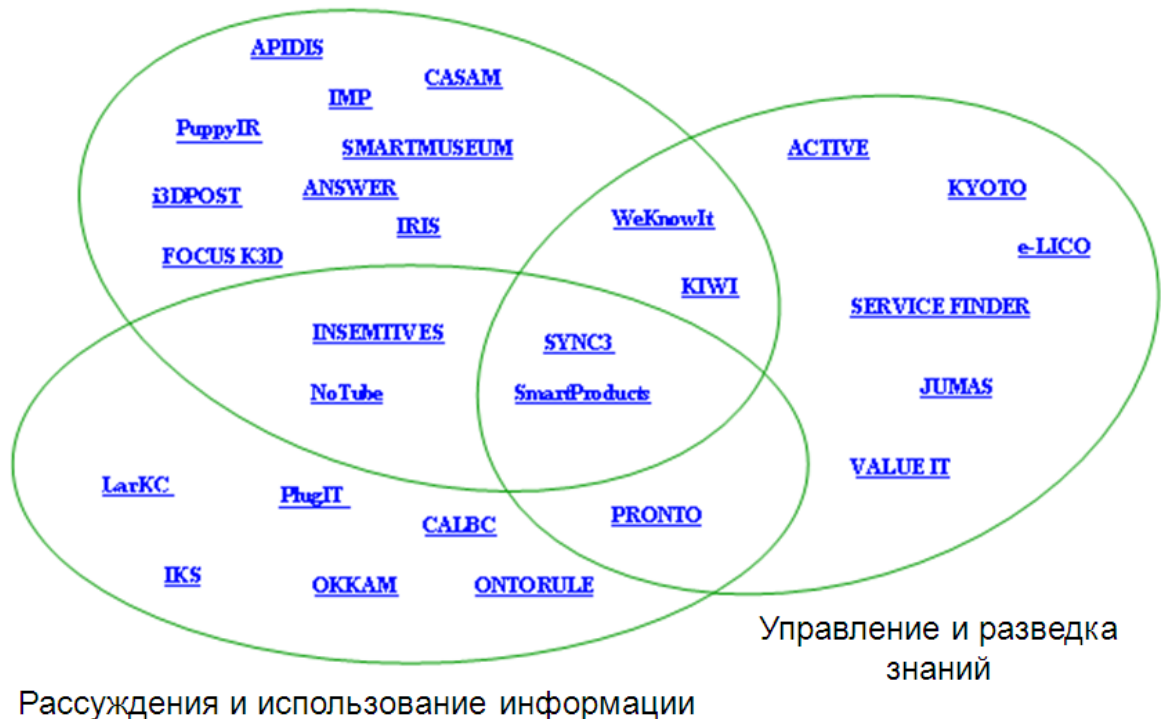
Исследования в области управления знаниями и онтологий занимают существенное место среди европейских научно-исследовательских проектов, относящихся к представлению знаний и применению онтологий в 2006 г. (рис. 12.1.):



Источник: <http://cordis.europa.eu/>, 2006

Рис. 12.1. Проекты в области управления знаниями и онтологий 6-й рамочной программы Евросоюза

Онлайн-контент, взаимодействие
и социальные медиа



Источник: <http://cordis.europa.eu/>, 2010

Рис. 12.2. Проекты в области управления знаниями и онтологий 7-й рамочной программы Евросоюза

Далее дано описание некоторых европейских проектов в области управления знаниями на основе онтологий. В описании каждого проекта дана ссылка на соответствующий сайт проекта, на котором можно получить детальную информацию о проекте и его результатах (результаты госпроектов полностью являются достоянием общественности и, как правило, размещаются в свободном доступе).

12.1. Проект ACTIVE

Сайт проекта: <http://active-project.eu/>

ACTIVE обращается к необходимости повышения производительности работника умственного труда с помощью трех исследовательских направлений:

1. Более легкий обмен информацией с помощью *синергии*

неформальных методов, таких как совместное тегирование ресурсов и вики, *и формального подхода*, основанного на онтологиях (последний используется для улучшения поиска информации);

2. Обмен и повторное использование *неформальных процессов работы со знаниями*, путем извлечения этих процессов из поведения пользователя;

3. *Понимание контекста пользователя*, таким образом, чтобы приспособить информацию, предоставляемую пользователю, к его текущей задаче.

Результаты ACTIVE применимы ко всем отраслям деятельности; апробацию они проходят в области консалтинга, телекоммуникаций и инженерии.

Краткая информация о деятельности

Проект ACTIVE стартовал в марте 2008 года, и начальные усилия были сосредоточены на достижении общего понимания технической базы проекта, на разработке архитектуры, и на понимание в деталях, как технология ACTIVE может принести пользу в трех кейсах (на трех предприятиях, выбранных в качестве полигона). Основные научные результаты проекта получены в следующих областях:

- развитие языка облегченных (lightweight) онтологий;
- понимание того, как осуществлять перевод из неформальных процессов работы со знаниями в восприятии пользователей в события низкого уровня, которые видны машине (получение письма в почтовом клиенте, открытие приложения X и т. п.);
- понимание того, как контекст может быть представлен для пользователя, и как он может направлять предоставление информации.

Архитектура ACTIVE была задумана как набор взаимодействующих программных сервисов, ACTIVE Knowledge Workspace (AKWS). Главной особенностью является то, что ACTIVE не заменяет обычные приложения пользователя, а поддерживает их; конечные пользователи могут извлечь выгоду из технологии ACTIVE без необходимости использования каких-либо новых программных средств. В рамках трех кейсов акцент был сделан на понимании поведения пользователей и разработку вариантов использования технологии ACTIVE, которая может предложить

поддержку для пользователей.

Онтологии и фолксонмии (Folksonomies) – слияние формального и неформального

Предприятия понимают важность обмена знаниями для работников умственного труда. Проблема систем, которые предоставляются сотрудникам, частов том, что они поглощают много времени и являются громоздкими для использования из-за формальной схемы. В последнее время работники умственного труда начали использовать неформальные методы, такие как теги и вики для совместного использования документов и мультимедийных объектов. Преимущество таких неформальных методов заключается в том, что они просты в использовании, недостатком является то, что результирующей схеме представления знаний не хватает выразительности формального подхода и, соответственно, не доступны логический вывод и семантический поиск. Работа в ACTIVE включает в себя разработку и использование облегченных онтологий, которые вычислительно разрешимы, то есть, могут быть обработаны за полиномиальное время, и автоматически формируются системой из фолксонмий. Работы с легкими онтологиями в настоящее время способствуют развитию языка OWL2. Неформальные теги перестают быть личными для отдельного пользователя, система ACTIVE позволяет более эффективно совместно их использовать.

Неформальные процессы работы со знаниями – обучение, повторное использование, обмен

Все работники умственного труда используют неформальные процессы работы со знаниями. В отличие от формальных бизнес-процессов, которые принадлежат организации, неформальные знания являются собственностью работников умственного труда, которые создают и используют их. Примерами являются: поиск информации о клиенте до визита к клиенту или письменное предложение по проекту. Недостатком является то, что эти знаниевые процессы редко распространяются и могут быть быстро забыты даже своими создателями. Результатом является то, что они постоянно создаются заново. Целью ACTIVE является использование методов искусственного интеллекта для изучения таких процессов. В рамках проекта ACTIVE внимание было сосредоточено на разработке модели для перевода из неформальных процессов работы со

знаниями в восприятии пользователей в события низкого уровня, которые видны машине.

Использование контекста для управления доставкой информации

С одной стороны, работникам умственного труда необходима информация в соответствии с их текущей задачей. С другой стороны, приход новой информации часто стимулирует работника умственного труда к переключению контекста, и постоянное переключение контекста является одним из основных факторов понижения производительности. С технологией ACTIVE, пользователь может установить свой собственный контекст, или же машина будет автоматически определять контекст пользователя. В любом случае, зная контекст пользователя, разработанная технология будет управлять и расставлять по приоритетам поток информации для пользователя. Благодаря использованию контекста, система поможет пользователю найти нужную информацию быстрее и гарантировать, что информация, представленная пользователю, будет соответствовать его текущим потребностям. Когда пользователь вынужден переключить контекст, технология ACTIVE позволит облегчить и ускорить это переключение, сохраняя текущий контекст для его восстановления в будущем.

Модуль «Менеджер контекста» отображает возможные контексты, при этом текущий контекст выделен. Чтобы изменить контекст пользователь может щелкнуть на нужном новом контексте. В MS Word после выбора нужного контекста в разделе меню «Последние документы» будут показываться не все документы, а только те, которые относятся к выбранному контексту. В приложении электронной почты пользователь увидит только сообщения, имеющие отношение к текущему контексту; альтернативой может быть ситуация, когда показываются все электронные письма вместе с теми, которые касаются текущего контекста, при этом последние будут показаны в верхней части списка.

Знаниевое рабочее место – делая все вместе

Рассматриваемый проект объединяет вместе целый ряд современных разработок. Все инновации проекта доступны пользователю в интегрированном решении – знаниевом рабочем месте (Knowledge Workspace) ACTIVE. Существующие приложения пользователя не будут

заменены. Наоборот предложенные в проекте функциональные возможности интегрируются с повседневными приложениями пользователя.

Конкретные примеры (кейсы)

ACTIVE имеет три кейса, которые используются для апробации разработанных в проекте методов и инструментов. Они находятся в трех различных секторах:

Телекоммуникации – повторное использование знаний и навыков поиска. Технология ACTIVE поможет техническим и торговым специалистам ВТ (BritishTelecom) обмениваться и повторно использовать их опыт, а также находить навыки, необходимые для быстрого реагирования на потребности клиентов. Основное внимание уделено процессу разработки сотрудниками ВТ предложений для клиентов.

Консультирование – снижение барьеров для обмена знаниями. Проблемы, решаемые в компании Accenture, очень похожи на те, что отмечены выше для компании ВТ. Также областью для апробации решения стала разработка предложений для клиентов.

Проектирование в микроэлектронике – технология ACTIVE будет вести (guide) инженеров Cadence в сложном процессе проектирования интегральных схем.

12.2. Проект NEPOMUK

Сайт проекта: <http://nepomuk.semanticdesktop.org/>

Проект NEPOMUK поддерживается фондом Европейского Союза IST, грант номер FP6-027705.

Введение

Большая доля повседневной деятельности тем или иным образом концентрируется вокруг обработки информации. Поиск информации, ее усвоение, запись новых идей и обмен результатами с другими людьми являются ключевыми формами деятельности, как на работе, так и во время досуга. Обилие компьютеров и Интернета в сегодняшнем мире приводит к новому количеству и качеству обмена информации и взаимодействий, которые воспринимаются пользователями и как шанс, и как угроза. Поддержка обработки личной и общей информации, таким образом, является крайне необходимой, но еще нерешенной проблемой.

В традиционных компьютерных архитектурах приложения представляют собой изолированные острова данных, каждое приложение имеет свои собственные данные, не зная о связанных и соответствующих данных в других приложениях. Индивидуальные производители приложений могут решить, чтобы их приложения взаимодействовали, так чтобы, например, почтовый клиент «знал» об адресной книге. Однако на сегодняшний день нет согласованного подхода, который бы обеспечивал взаимодействие и общесистемный обмен данными между приложениями. Аналогичным образом компьютеры различных пользователей также являются изолированными островами; нет унифицированной архитектуры для взаимодействия и обмена данными между компьютерами. Пользователи могут обмениваться данными, посылая электронные письма или загружая их на сервер, но до сих пор не существует способа связи между приложением, используемым одним человеком на своем компьютере с приложением, которое используется другим человеком на другом компьютере.

Проблемы компьютеров аналогичны проблемам Интернета, где мы также сталкиваемся с изолированными островами данных, при этом не существует общего способа интеграции и обмена данными между различными веб-приложениями (например, веб сервисами). Видение семантического веба предлагает решения для обеих задач. Язык RDF (см. раздел «Использование онтологий и языков семантического веба») представляет собой общий формат данных, который строит мосты между островами, а технология семантических веб-сервисов предлагает средство для интеграции приложений в Интернете.

Парадигма социального семантического персонального компьютера (ССПК, англ. – Social Semantic Desktop) применяет идеи парадигмы семантического веба для персонального компьютера. Формальные онтологии охватывают как общие концептуализации данных на компьютере, так и личные ментальные модели. RDF представляет собой общий формат представления данных. Веб-сервисы (приложения в Интернете) могут описать свои возможности и интерфейсы в стандартной форме и, таким образом, стать семантическими веб-сервисами. На персональном компьютере приложения (или, вернее: их интерфейсы) поэтому будут смоделированы аналогичным образом. Вместе эти

технологии создают средства для построения семантических «мостов» необходимых для обмена данными и интеграции приложений. ССПК превратит обычный компьютер в сетевую рабочую среду путем сглаживания границ между отдельными приложениями и физическим рабочим пространством различных пользователей.

Создав ССПК, проект Peromuk внес вклад в несколько аспектов эффективной личной обработки информации:

- Индивидуальным пользователям предложен системный подход к структуре информационных элементов на их личном компьютере. Используя стандартную технологию для описания и хранения структур и отношений, пользователи могут легко воспроизводить и выражать все, что играет важную роль в их персональной сфере;

- Стандартизированные интерфейсы позволяют интегрировать все виды имеющихся приложений в личную информационную сеть. Инвестиции в программы, сбор данных, а также с трудом выученные стили работы не будут потеряны, а наоборот расширены и связаны во всеобъемлющее информационное пространство;

- Становятся возможны все виды автоматической и полуавтоматической поддержки, основанной на технологиях семантического веба, например, сервисы по классификации текстовых документов и изображений, оценки релевантности документов и т. д.

- Обмен стандартными форматами данных между отдельными рабочими пространствами поддерживается не только на техническом уровне (например, через стандартные протоколы связи), но и на семантическом уровне (через обмен и отображений (alignment) онтологий и соответствующих элементов аннотированной информации). Интеграция с формальными онтологиями облегчает обмен и взаимопонимание между разными лицами;

- В конечном счете, проект делает вклад в решение проблемы инициирования семантического веба. Когда отдельный пользователь будет получать непосредственную выгоду от семантического аннотирования в личном рабочем пространстве, его мотивация будет достаточно высока, чтобы инвестировать свое время и усилия в необходимые работы по структурированию и формализации. Поскольку используемые стандарты

позволяют без усилий обмениваться результатами такой работы, количество семантически аннотированной информации, которая может быть доступна в сети стремительно вырастет, что, в свою очередь, сделает целесообразной разработку новых сервисов на основе семантического веба.

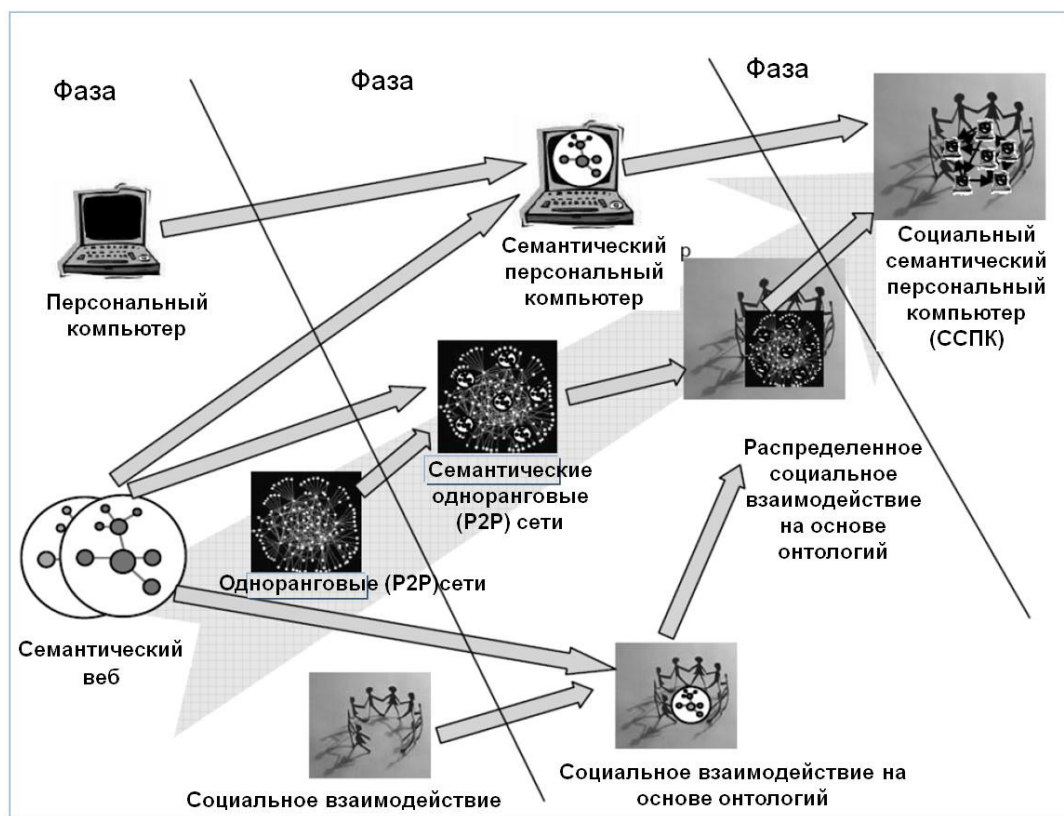


Рис. 11.1. Формирование концепции социального семантического персонального компьютера (ССПК)

В своей статье «ПК с мозгами» в журнале «Открытые системы» Леонид Черняк пишет «Совместные усилия европейских специалистов направлены на создание "семантического ПК" (Semantic Desktop), то есть устройства, которое будет хранить информацию индивида, тексты, фото и видео – все что угодно, все, что встретилось когда-либо в его жизни. Сохраненная информация интерпретируется как совокупность ресурсов в духе Semantic Web, каждый ее фрагмент снабжается своим идентификатором Uniform Resource Identifier, а доступ к нему обеспечивается в соответствии с Resource Description Framework. Кроме того, средствами онтологий пользователь может выражать собственные ментальные модели и представления, задавать семантические связи

между фрагментами информации. Можно сказать, что Semantic Desktop расширяет возможности человеческой памяти.

Семантический ПК создается как открытое программное обеспечение в рамках рассчитанного на три года проекта Nepomuk Project, работа над которым координируется Германским центром исследований в области искусственного интеллекта (German Research Center for Artificial Intelligence, DFKI). Исходно Nepomuk был рассчитан на свободную среду рабочего стола для Unix-подобных операционных систем KDE (K Desktop Environment). Но программное обеспечение Nepomuk может быть загружено и Windows, Macintosh и Linux, а с некоторыми ограничениями можно подключиться и через Web-страницы, используя Firefox.

Сама идея семантического ПК не нова. Ее, так или иначе, можно проследить со времен гипотетической машины метех Ванневару Буша до нашего времени, однако многочисленные попытки были переусложненными. На их фоне представляется, что авторы Nepomuk нашли наиболее практичное решение. Оно сводится к тому, что программное обеспечение автоматически добавляет дополнительную семантическую информацию к той, что уже хранится в компьютере, и создает условия к тому, чтобы пользователь мог вручную дополнять и расширять семантическую составляющую. Исходная семантическая информация формируется с использованием программ-«обходчиков» (crawler), обследующих доступные им файлы. Когда она собрана, Nepomuk увязывает разные источники между собой. В отличие от поисковых систем, где поиск сводится к совпадению по ключевым словам, здесь может быть организована контентная связь, скажем «имя», «дата» или что-то подобное.

Таким образом, функционирование «ПК с мозгами» сводится к следующим трем этапам. Сначала аннотируется и увязывается любая информация, хранящаяся на компьютере, независимо от форматов и приложений, где она применяется. Для этого используются структуры данных и методы, отработанные при создании Semantic Web. Затем артикулируются и визуализируются идеи пользователя, чтобы потом трансформировать их в семантическую информацию, на этом этапе используются технологии, аналогичные вики. На третьем этапе интегрируются результаты, полученные автоматически и вручную.

Практическую направленность проекта Nерomuk высоко оценили коллеги, в том числе Нова Спивак, основатель компании Radar Networks, который отметил: «В этом решении огромный потенциал. Nерomuk создан в расчете на пользователей и разработчиков». Есть мнение, что развитие семантических ПК сможет вывести из застойного состояния и Semfntic Web, работа с семантикой на собственном компьютере в последующем приучит людей к осмысленной работе в более широком контексте.

Функциональные возможности решения

ССПК является платформой, которая используется для разработки различных видов социальных и семантических приложений. Эти приложения имеют общие функциональные возможности, которые должны поддерживаться ССПК. Они разделены на пять групп, которые могут быть рассмотрены как различные аспекты ССПК. В табл. 12.1 приведены пять различных аспектов и отдельных функциональных возможностей внутри каждой группы.

Таблица 12.1

Функциональные возможности ССПК

| | |
|----------------|--|
| Компьютер | Аннотация, внешний доступ к персональному компьютеру в автономном режиме, совместное использование персонального компьютера (Desktop Sharing), управление ресурсами, интеграция приложений, управление уведомлениями |
| Поиск | Поиск, нахождение связанных элементов |
| Социальный | Социальное взаимодействие, совместное использование ресурсов, управление правами доступа, публикация/подписка, управление группой пользователя |
| Профилирование | Обучение, настраивание, доверие, ведение журнала |
| Анализ данных | Рассуждение, извлечение ключевых слов, сортировка и группировка |

Компьютер. На уровне персонального компьютера, семантическая функциональная возможность, общая для большинства приложений, заключается в возможности добавлять информацию о любых ресурсах. Функциональная возможность **Аннотация** предоставляет инструмент для сохранения и последующего поиска семантических отношений о чем-либо в компьютере. Когда пользователь аннотирует свои фотографии из

поездки, он делает это с помощью своего самого любимого приложения по работе с фотографиями (например, Picasa или Iphoto), аннотации затем сохраняются в ССПК. Разработчики называют эту функцию **Интеграцией приложений**; приложения взаимодействуют с ССПК с помощью различных служб. Когда пользователь получает уведомление о поездке в другой город, то это работает служба **Управление уведомлениями**. ССПК управляет различными видами механизмов, таких как электронная почта, RSS или текстовые сообщения. Когда пользователь создает новый концепт в персональной онтологии или даже новый файл на ССПК, приложение, которое он использует, взаимодействует с функциональной возможностью **Управлению ресурсами ССПК**, создавая необходимые семантические элементы (структуры) в соответствии с текущим контекстом и установками. Некоторая информация, необходимая пользователю при бронировании поездки, хранится в компьютере его подруги. Если она не подключена к сети, с помощью функции **Автономного доступа** экспортируется соответствующая информация на другой компьютер. Совместное использование ресурсов (**Desktop Sharing**) является возможностью для различных пользователей ССПК работать, используя одни и те же ресурсы. Два пользователя могут совместно писать отчет о совместной поездке: управление ресурсами осуществляется на компьютере одного, но второй может получить доступ и изменять его удаленно.

Поиск. Семантические сети, созданные на компьютере, раскрывают совершенно новый способ поиска на ССПК. **Поиск** использует семантические отношения, а также социальные отношения для получения соответствующих результатов. Как только объект будет найден, пользователь сможет также **найти связанные элементы**. Например, когда пользователь ищет билеты в другой город, он может также выполнить поиск связанных элементов и на основе опыта своих социальных контактов выяснить, что билеты другой компании дешевле.

Социальный. ССПК обеспечивает различные средства **социального взаимодействия**, например, вложение семантической информации в электронные письма или текстовые сообщения или предоставление возможности аннотировать ресурсы другого пользователя. Для некоторых функциональных возможностей на уровне компьютера (совместное использование ресурсов и автономный доступ) требуется, чтобы в ССПК

использовалась функциональная возможность **совместное использование ресурсов**. Для совместной работы над докладом о поездке, доступ к нему может быть предоставлен всем участникам рабочей группы, путем добавления его к общему информационному пространству. При совместном использовании ресурсов или информации в сети, **управление правами доступа** ССПК позволяет определить конкретные отношения между правами пользователей, групп и ресурсов. Система **управления группой пользователей** ССПК делает легким быстрое создание новых групп из списка пользователей. Эти группы могут затем быть использованы для изменения прав доступа или для совместного использования ресурсов в общем информационном пространстве. Например, некоторые из друзей пользователя, возможно, подписались на получение уведомлений о новых фотографиях, которые он аннотирует и делает доступными. Механизм ССПК **Публикация / Подписка** облегчает создание каналов соответствующей информации.

Профилирование. Если включена функциональная возможность ССПК **Ведение журнала**, то она регистрирует активности пользователя, которые помогут обнаружить текущий контекст пользователя. Профилирование ССПК может быть сделано автоматически с помощью **Обучения**: ССПК учится предсказывать поведение пользователя. Пользователь может также **Настраивать** интеллектуальное поведение ССПК: некоторые изученные контексты могут стать нерелевантными и, возможно, должны будут повторно адаптированы или удалены. Понятие **Доверия** на ССПК между людьми или источниками информации является также результатом **профилировании** компьютера. Пользователь А может определить, что он доверяет информации своего друга Пользователя Б, или ССПК Пользователя Б может определить, что Пользователь А является надежным источником информации относительно проекта X.

Анализ данных. Для поддержки обучения ССПК или запросов связанных элементов, ССПК предоставляет различные механизмы *анализа данных*, например, такие как **Рассуждение**. Например, когда пользователь аннотирует фотографию метками «река Вычегда» и «Россия», ССПК может сделать вывод, что Вычегда находится в России. Эта информация впоследствии может быть использована для поиска. **Сортировка и**

группировка применяются в приложении, которое выполняет поиск. ССПК извлекает элементы из многих источников и людей, а потом сортирует и группирует эти элементы, в соответствии с различными критериями, используя семантику, определенную на этих ресурсах. **Извлечение ключевых слов** из ресурсов, из текстовых, например, полезно для выполнения автоматического аннотирования или синтеза кратких описаний (summarizing).

12.3. Проект OrganiK

Подход OrganiK к управлению знаниями: на пути к социально-техническому соответствию

Предлагаемый подход заключается в том, чтобы дать возможность работникам умственного труда в малых интеллектуальноемких (knowledge-intensive) компаниях эффективно сотрудничать и использовать организационные знания благодаря использованию СУЗ. Предлагаемый в проекте подход основан на социально-техническом взгляде и рассматривает эффективность взаимодействия между людьми и технологиями как задачу, требующую максимума усилий. Основными компонентами предлагаемой СУЗ являются следующие:

- *Сконцентрированная на человеке* концептуализация УЗ сфокусирована на социальных процессах, ситуативных методах работы и организационных структурах (например отдельный сотрудник, группа, подразделение). Ситуативные процессы *инновационного менеджмента*, развитие *сообществ практиков* и *процедуры адаптации* проектов – основные компоненты этого социально-ориентированного процессного подхода.

- *Сконцентрированная на технологиях* концептуализация УЗ сфокусирована на интеграции социального программного обеспечения предприятий (wikis, блоги, инструменты совместных закладок и поисковые системы) с семантическими технологиями (аннотация на основе онтологий, семантический анализ текста, рассуждения на основе логики).

Предлагаемая архитектура

Архитектура, которая предлагается в проекте OrganiK, необходима для сокращения социотехнического разрыва между методами работы в

малых интеллектуальноемких фирмах и современными СУЗ. Она сочетает в себе ключевые элементы из областей «Предприятие 2.0» и Semantic Web технологий. В части области «Предприятие 2.0», системная архитектура использует структуру SLATES (аббревиатура от английских слов):

- *Поиск (Search)* предоставляет механизмы для обнаружения информации.
- *Ссылки (Links)* помогают работникам умственного труда в обнаружении необходимых знаний и обеспечивают возникающие структуры для онлайн-контента.
- *Авторская разработка (Authoring)* дает возможность работникам умственного труда поделиться своим мнением с широкой аудиторией.
- *Теги (Tags)* предоставляют альтернативный способ навигации, использующий неиерархические классификации содержимого интрасети.
- *Расширения (Extensions)* используют коллективный разум (collaborative intelligence) и рекомендуют работникам умственного труда контент, который соответствует их контексту.
- *Сигналы (Signals)* автоматически оповещают работников умственного труда о новом доступном и релевантном контенте.

Цель заключается в предоставлении работникам умственного труда совместной рабочей области, которая включает в себя набор интегрированных приложений Web 2.0 (вики, блоги, система закладок и механизм поиск/рекомендация), дополненных обработкой естественного языка и возможностями семантической интеграции информации, которые обеспечивают комбинированное использование фолксономий и ситуативных тегов с тезаурусами и общими онтологиями. Использование семантических технологий в предусмотренных решениях включает в себя следующие основные функции:

- *Семантическое представление знаний*: представление знаний в формальной, понятной для компьютеров форме.
- *Семантическое аннотирование ресурсов*: аннотирования артефактов знаний и других ресурсов с помощью ссылок на понятия, определенные в онтологической модели.

- *Семантический вывод*: осуществление автоматизированных, основанных на логике рассуждений для выявления нового, неявного знания, основанного на том, что уже утверждалось в явном виде.
- *Семантический поиск и разведка*: использование онтологических терминов для описания поискового запроса, а также применение рассуждений, основанных на логике, чтобы получить соответствующие результаты.

Каждая из вышеупомянутых социальных функций программного обеспечения соответствует одному или более компоненту структуры SLATES и, как показано в таблице 12.2, соответствует определенному компоненту в предлагаемой архитектуре СУЗ.

Таблица 12.2

Соответствие между компонентами в структуре SLATES и в предлагаемой архитектуре

| Структура SLATES | Предлагаемая архитектура |
|-------------------------|---|
| Поиск | Семантический поиск |
| Ссылки | Совместные закладки |
| Авторская разработка | Вики и блог пространства |
| Теги | Совместные закладки, вики и блог пространства |
| Расширения | Рекомендующие системы |
| Сигналы | Информационные каналы (RSS) |

12.4. Проект X-Media

Сайт проекта: <http://www.x-media-project.org/>

Проект X-Media обращается к проблеме управления знаниями в сложных распределенных средах. В нем изучаются, разрабатываются и реализуются методологии и инструменты крупного масштаба, которые в состоянии поддерживать обмен и повторное использование знаний, распределенных в различных медиаформатах (изображения, документы и данные) между хранилищами. Это 4-х летний и шестой по счету рамочный проект IST, который начался в марте 2006 года. 15 академических и промышленных партнеров сотрудничают в проекте. X-Media координируется Университетом Шеффилда.

Увеличение размера архивов организации имеет серьезные последствия для того, каким образом может быть реализовано УЗ. Сложность задачи складывается из нескольких составляющих – присутствия нескольких медиа-форматов в распределенных архивах, трудности сосредоточения и интеграции знаний из различных источников, а также динамичной и неопределенной природы знаний. X-Media предоставляет основу для нового поколения служб и инструментов: (1) он разрабатывает новое поколение односредовых (single medium) **технологий получения знаний**, более мощных, чем те, которые существуют в настоящее время; (2) он определяет и реализует **мульти- и кросс-медиа-технологии** для получения и повторного использования знаний во всех медиаформатах; (3) он определяет **новые способы обмена и повторного использования знаний** в крупных организациях на основе семантического веба.

Проект X-Media исследует, проектирует и разрабатывает:

- Новые парадигмы для поиска, обмена и повторного использования знаний, которые позволяют пользователям определять и параметризовать взгляды (views) на имеющиеся знания.
- Надежные и масштабируемые инструменты получения знаний и анализа данных, преодолевающие границы различных медиаформатов,
- Методы объединения (fusion) знаний для помощи работникам умственного труда в принятии решений, когда они сталкиваются со знаниями, возможно, противоречивыми, полученными из различных источников.
- Техники для представления и управления (I) неопределенностью, (II) доверием и происхождением (источников), а также (III) динамическими аспектами знаний.
- Методологию и техническую инфраструктуру способную справиться со сложностью мультимедийных хранилищ больших масштабов.
- Обобщенную и гибкую архитектуру, которая позволяет конечным пользователям легко ее настроить и интегрировать с практиками УЗ.

Данная работа обусловлена потребностями реального мира, взятыми

из конкретных случаев в области производства: Rolls-Royce (реактивные двигатели) и Fiat (легковые автомобили). В качестве «полигона» для разработки и внедрения предлагаемых технологий взяты мониторинг жизненного цикла продукта (Rolls-Royce) и анализ конкурентов (Fiat).

Способность справляться со знанием из различных мультимедийных объектов (тексты, изображения и данные) хорошо подходит среде, в которой: (1) датчики и камеры обеспечивают основные данные для интерпретации; (2) текстовые документы дополняют, описывают и помогают интерпретировать данные и изображения; (3) онтологии описывают области и приложения.

Направления деятельности

Техническая деятельность разделена на четыре области:

Получение знаний: автоматизированное получение и извлечение знаний из документов, изображений и исходных данных. Изучаются и разрабатываются методы и технологии для аннотирования мультимедийных ресурсов (например текст, изображения, данные), а также соотношение знания в различных мультимедийных ресурсах; аннотирование документов должно быть в значительной степени автоматизировано и поддержано технологиями получения знаний.

Обмен знаниями: обмен знаниями и повторное их использование: изучаются и разрабатываются технологии и методики для легкого и интеллектуального доступа и повторного использования формализованных и не формализованных знаний.

Инфраструктура: инфраструктура, в которой реализована среда для получения знаний, их интеграции и совместного использования. Так как проект X-Media представляет собой комплексный проект, ориентированный на разработку приложений, интеграция должна быть осуществлена не только на концептуальном уровне, но и на уровне программной реализации.

Применение: применение и тестирование тех технологий, которые разработаны в перечисленных выше областях, используются для определения наглядных примеров и прототипов приложений для промышленного тестирования. Проект нацелен на передачу результатов проекта в область бизнеса, по этой причине две основных области для анализа определяются двумя крупными промышленными потребителями

(Rolls-Royce и Fiat).

В X-Media были разработаны следующие технологии и инструменты в рамках указанных выше направлений деятельности:

Получение знаний

- обработка документов нескольких форматов;
- технологии для извлечения нескольких типов семантической информации из хранилищ документов;
- интеграция технологий получения знаний в общую архитектуру X-Media.

Обмен знаниями

- инфраструктура управления метазнаниями;
- прототип инструмента аннотации текста;
- системы, выполняющие объединение (fusion) знаний;
- системы, осуществляющие поиск и просмотр (browsing) знаний.

Инфраструктура

Окончательный вариант архитектуры X-Media был представлен, чтобы проверить на опыте и определить, какие концептуальные основы и архитектура программного обеспечения подходят для потребностей X-Media. Этот компонент был решающим для обеспечения возможностей интеграции программного обеспечения среди партнеров и стал ключом в успешной оценке технологий.

Применение

Для мониторинга жизненного цикла продукта на Rolls-Royce:

- Разрешение проблемных вопросов (issue resolution): что включает в себя поддержку выявления и разрешения возникающих проблемных вопросов в реактивных двигателях; цели – это приобретение и обмен знаниями, как для решения открытых вопросов, так и для использования их в будущем.

- Экспериментальная вибрация: управление знаниями о вибрации двигателя в целях обеспечения безопасности двигателей.

Для конкурентного анализа на Fiat:

- Комфорт водителя: анализ и оценка кривых шума, необходимые инженерам в создании лучших решений с точки зрения аэродинамического комфорта;

- Прогноз деятельности конкурента: направлен на поддержку публикаций календарей прогноза на новые автомобили конкурентов.

12.5. Проект Mymory

Контекстуализированное приобретение знаний в личной семантической вики.

Сайт проекта: <http://www.dfki.uni-kl.de/mymory/>

В отличие от традиционных систем, основанных на знаниях, где знание некоторых экспертов формализуется в целях поддержки работы других (менее подготовленных) людей, сегодняшний работник умственного труда – это зачастую эксперт, инженер по знаниям и пользователь системы в одном лице. Технологии для управления личными знаниями должны поддерживать этот тип работы. Рабочая среда Mymory предоставляет образцовую систему управления личными знаниями для работы, сконцентрированной вокруг документов. Система вики выступает в качестве инструментального средства для генерации и использования документов. Семантическое расширение этой вики позволяет обогащать вики-текст формализованными аннотациями (метаданными) на основе модели личной информации и динамически загружаемых онтологий. Аннотации создаются вручную, а также автоматически.

Возможны два типа автоматизированных аннотаций:

1. Контекстная аннотации, основанная на наблюдении за работой пользователя и фиксации потока контекстных признаков, описывает ситуацию работника умственного труда, используя словарь (лексику) его личной информационной модели;

2. Аннотации внимания, которые автоматически формируются на основе показаний устройства для слежения за глазами человека (ай-трекера, eye-tracker). Такие аннотации позволяют различать прочитанные, бегло просмотренные, непрочитанные части текста.

Аннотации могут увеличить понимание документов и управление ими (ограниченными страницами вики в проекте Mymory), особенно когда документы используются более чем одним человеком и в более чем в одном контексте. Далее аннотированные документы расширяют поисковые возможности в тех случаях, когда поиска на основе ключевых слов не

достаточно, чтобы найти документ, который просматривался уже давно. Кроме того, аннотации позволяют использовать дополнительные фильтры при поиске, которые могут повысить его точность. Это особенно характерно для контекстной аннотации и, соответственно, для контекстной фильтрации. В проекте с помощью сценариев продемонстрировано, как хранение пользовательского контекста обеспечивает контекстный поиск или фильтрацию.

Система Мугогу полностью реализована и включена в физическое демонстрационное рабочее место. Однако есть еще некоторые технические ограничения, которые должны быть решены в будущем: Аннотации могут быть созданы только в рамках вики. Хотя существует средство импорта HTML и другого веб-контента, желательно иметь возможность аннотировать внешние веб-сайты и другие типы документов, такие как PDF-файлы, без необходимости импортировать их. Поскольку это требует серьезных технических усилий, поднимающих много нетривиальных вопросов (что делать, если изменяется внешнее содержание аннотированного?), проект был сконцентрирован на аннотировании контента, размещенного только в самой системе. Одной из перспектив развития проекта является исследование повторного использования аннотаций между контекстами, например, между пользователями или для одного пользователя, но в разное время или для различных задач, должно быть выполнено.

В настоящее время система применяется в двух реальных ситуациях: во-первых, она заменяет стандартную вики, которая уже используется в течение нескольких лет для управления личными знаниями в исследованиях. Кроме того, контекстно-ориентированные представления документа применяются в сценарии сервисного центра, где ситуационные взгляды (view) на документы должны быть созданы для различных групп пользователей. В этом сценарии участниками проекта также накапливается опыт по приобретению формальных знаний о документе с помощью анализа ручных пометок в документах/ вики-страницах (выделение цветом, подчеркивание).

12.6. Проект NEON

NeOn (www.neon-project.org) объединяет 14 европейских партнеров в рамках 6-й программы. Старт в марте 2006.

Цель – развитие онтологий для применения в больших семантических приложениях в распределенных организациях, так называемые сетевые онтологии с поддержкой совместной работы. NeOn должен перевести «теоретическую возможность» в конкретное эффективное программно-методическое решение для разработки и сопровождения крупномасштабных гетерогенных семантических приложений. Проект NeOn основан на следующих предпосылках, определяющих его новизну: 1. вновь создаваемые онтологии обычно интегрированы в систему уже существующих онтологий; 2. онтологии и метаданные в приложениях постоянно изменяются и требуют регулярного обновления. Проект NeOn создает методы, техники и инструменты для поддержки эволюции, контекстной адаптации и совместной работы с онтологиями.

Результаты проекта NeOn делятся на 11 групп:

1. Динамика онтологий,
2. Совместная работа,
3. Контексто-ориентированность,
4. Человеко-машинное взаимодействие,
5. Методология NeOn,
6. Справочная архитектура и инструментарий NeOn,
7. Приложение 1: Система оценки истощения запасов на основе онтологии,
8. Приложение 2: Поддержка совместной работы в фармацевтической отрасли,
9. Исследование рынка и разработка стратегии применения результатов NeOn в бизнесе,
10. Распространение результатов проекта NeOn,
11. Управление самим проектом NeOn.

12.7. Проект KnowledgeWeb

Сайт проекта: www.knowledgeweb.semanticweb.org

Группы результатов:

1. Бизнес-ориентированные результаты:
 - 1.1. Анализ потребностей в промышленных приложениях,
 - 1.2. Оценка для выбора технологии,
 - 1.3. Технологические рекомендации,
 - 1.4. Продвижение онтологической технологии,
 - 1.5. Межсетевая кооперация,
 - 1.6. Структура семантического портала;
2. Исследовательские результаты:
 - 2.1. Масштабируемость,
 - 2.2. Гетерогенность,
 - 2.3. Динамичность,
 - 2.4. Семантические веб-сервисы,
 - 2.5. Расширения языка семантического веба,
 - 2.6. Создание сообщества (сети) исследователей в области семантического веба;
3. Образовательные результаты
 - 3.1. Европейская ассоциация по образованию в области семантического веба (EASE, the European Association for Semantic Web Education),
 - 3.2. Образовательный контент и организация мероприятий,
 - 3.3. Образовательная семантическая платформа;
4. Результаты по управлению самим KnowledgeWeb-проектом.

12.8. Проект SEKT (Semantically-enabled Knowledge Technologies)

Сайт проекта: <http://www.sekt-project.com/>

Начало: 01.01.2004, Длительность: 36 months

Цель: Разработка технологий управления знаниями следующего поколения Next Generation Knowledge Management (NGKM). Интеграция технологий онтологического инжиниринга и работы с метаданными, технологиями обработки естественного языка и машинного обучения для

автоматического и полуавтоматического формирования и использования онтологий и метаданных.

Группы результатов проекта:

1. Генерация онтологии,
 2. Генерация метаданных,
 3. Управление онтологией и метаданными,
 4. Отображение (mapping, alignment) и объединение (merging) онтологий,
 5. Доступ к знаниям,
 6. Интеграция,
 7. Методология DILIGENT,
 8. Пользовательские потребности, Бенчмаркинг, Usability,
- Выгоды для бизнеса,
9. Кейс: Улучшение индивидуальной продуктивности,
 10. Кейс: Система поддержки принятия решения для юристов,
 11. Кейс: Сокращение накладных расходов на создание и сопровождение знаний,
 12. Распространение и использование.

ЗАКЛЮЧЕНИЕ

Создание СУЗ или отдельных ее компонент позволяет улучшить ключевые показатели работы предприятия, такие как, прибыль, качество продукции, скорость вывода новых продуктов и выполнения заказов, затраты, благодаря более эффективной работе со знаниями и интеллектуальными ресурсами.

Вот некоторые эффекты от внедрения СУЗ или ее компонент:

- Если велик риск ухода ключевых специалистов, то СУЗ позволит сохранить знания в организации,
- Если сотрудники слишком много времени тратят на поиск нужной информации и не всегда находят то, что искали, то СУЗ позволит сотрудникам получать нужную информацию в нужном месте в нужном формате своевременно и с минимальными затратами усилий,
- Если сотрудники делают ошибки в работе или принимают неоптимальные решения вследствие недостатка информации и знаний, то СУЗ позволит уменьшить количество ошибок сотрудников и увеличить качество принимаемых решений,
- Если компания сталкивается с проблемами внутренних или внешних коммуникаций, то методы и инструменты СУЗ позволят улучшить коммуникации, снизить информационные потери и искажения,
- Если отсутствует передача знаний между отдельными группами специалистов, и они постоянно изобретают «колесо», то СУЗ позволит стимулировать обмен знаниями, лучшими практиками и наработками.

В учебном пособии предложено описание основных компонент СУЗ, которое позволило более четко определить данное понятие. Комплексная СУЗ – это совокупность процессов и инструментов УЗ, обеспечивающих целенаправленное развитие и эффективное функционирование процессов создания, накопления, распределения, обмена и использования знаний, а также сетей знаний. Процессы УЗ в предложенном подходе включают диагностику и анализ знаний, проектирование, планирование внедрения и оценку эффективности СУЗ. Инструменты УЗ – это информационные

технологии, административно-организационные механизмы, управление персоналом, техническое и правовое обеспечение деятельности. Программная реализация СУЗ – это комплекс информационных технологий, обеспечивающий целенаправленное развитие и эффективное функционирование процессов создания, накопления, распределения, обмена и использования знаний, а также сетей знаний.

Для разработки программной реализации СУЗ необходимо:

- Выявить ключевые знания, необходимые для существования предприятия и/или реализации его стратегии,
- Определить источники и потребителей ключевых знаний, а также проблемы и возможности, связанные с этими знаниями,
- Определить модели, методы и программные средства, которые позволят улучшить процессы создания, накопления, распределения, обмена и использования ключевых знаний,
- Продумать интеграцию отдельных моделей, методов и программных средств в единую систему – разработать архитектуру СУЗ,
- Обосновать и донести до основных заинтересованных сторон необходимость внедрения предлагаемых элементов СУЗ (например, с помощью описаний примеров использования, технико-экономических обоснований),
- Спланировать процесс разработки и внедрения СУЗ – состав и результаты работ, последовательность работ, исполнителей, сроки, бюджет (формально такой план может быть представлен в виде программы внедрения СУЗ и т. п.),
- Определить метрики результативности и эффективности для оценки деятельности и инструментов УЗ.

Для разработки комплексной СУЗ необходимо дополнительно определить модели, методы и инструменты в области организации деятельности, управления персоналом, хозяйственного и юридического обеспечения деятельности, которые позволят улучшить процессы создания, накопления, распределения, обмена и использования ключевых знаний.

Важно отметить, что для создания работающей и полезной СУЗ необходимо выполнить все указанные выше задачи разработки СУЗ,

интегрировать различные виды инструментов УЗ, выбрать правильную последовательность их внедрения, а также учесть уровень зрелости предприятия в области УЗ.

Применение онтологий в СУЗ позволит:

- Улучшить внутри- и межорганизационные коммуникации, благодаря четкому определению значений слов,
- Интегрировать информацию, распределенную по различным хранилищам документов, базам данных и знаний, а также в головах сотрудников,
- Обобщить и систематизировать имеющуюся информацию, выступив в качестве метамоделей,
- Использовать автоматизированный логический вывод для улучшения результатов поиска, получения новых знаний и анализ информации,
- Использовать более эффективные механизмы получения, визуализации и поиска знаний.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Букович У., Уильямс Р. Управление знаниями: руководство к действию. – М: Инфра-М, 2002. – 504 с.
2. Гаврилова Т. А., Муромцев Д. И. Интеллектуальные технологии в менеджменте: Учеб. пособие. – СПб.: «Высшая школа менеджмента», Издат. дом СПбГУ, 2008. – 488 с.
3. Гиляревский Р. С. Информационный менеджмент. Управление информацией, знаниями, технологией. Серия: Библиотека. Изд-во: Профессия, 2009 г. – 304 с.
4. Мильнер Б. З., Румянцева З. П., Смирнова В. Г., Блинникова А. В. Управление знаниями в корпорациях. Под ред. Мильнера Б.З., М: «ДЕЛО», 2006. – 304 с.
5. Мильнер Б. З. (редактор) «Инновационное развитие: экономика, интеллектуальное развитие, управление знаниями», М: Инфра-М, 2009. – 624 с.
6. Климов С. М. Интеллектуальные ресурсы общества СПб: ИВЭСЭП, Знание, 2002. – 199 с.
7. Коллисон К., Парселл Д. Учитесь летать: Практические уроки по управлению знаниями от лучших обучающихся организаций. Серия: Библиотека ИКСИ. Изд-во: Институт комплексных стратегических исследований, 2006 г. – 296 с.
8. Лессер Э., Прусак Л. (составители) Как превратить знания в стоимость: Решения от IBM Institute for Business Value. Пер. с англ. – М.: Альпина Бизнес Букс, 2006. – 248 с.
9. Нонака И., Takeuchi Х. Компания – создатель знания. Зарождение и развитие инноваций в японских фирмах. – М.: Олимп-Бизнес, 2003. – 384 с.
10. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с.
11. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии) – Томск: Изд-во НТЛ, 2005. – 260 с.
12. Управление знаниями: Хрестоматия / Пер. с англ. под ред. Андреевой Т. Е., Гутниковой Т. Ю. СПб.: Изд-во «Высшая школа менеджмента», 2009. – 514 с.
13. Davies J., Grobelnik M., Mladenic D. (Eds.) Semantic Knowledge Management. Springer. 2009. – 252 p.
14. Davenport T., Probst G. Knowledge Management Case Book: Siemens Best Practices. Wiley-VCH. 2002. – 336 p.
15. Mentzas, G., Apostolou D., Abecker A., Young R. Knowledge Asset Management: Beyond the Process-centred and Product-centred Approaches, Springer London, 2002. – 200 p.
16. Schreiber G., Akkermans H., Anjewierden A., R. de Hoog, Shadbolt N., W. van de Velde, Wielinga B. Knowledge Engineering and Management: The CommonKADS Methodology, The MIT Press, Cambridge, MA, 2000. – 455 p.
17. Staab S., Schunurr H-P., Studer R., Sure Y. Knowledge processes and ontologies // IEEE Intelligent Systems, Special Issue on Knowledge Management, 2001. 16(1), P. 26-34.
18. Wiig, K., de Hoog, R., van der Spek, R. Supporting knowledge management: a selection of methods and techniques // Expert Systems with Applications, Vol. 13 No.1, 1997. Pp. 15-27.
19. Mika, P., Akkermans H. Towards a New Synthesis of Ontology Technology and Knowledge Management // Knowledge Engineering Review, Cambridge University Press, Volume 19, № 4, 2004. P. 317-345.

ГЛОССАРИЙ

Данные – это полученные эмпирическим путем и зафиксированные факты, характеризующие отдельные свойства объектов, процессов или явлений.

Индексирование – процесс выбора и присвоения документам, их частям, данным и/или отдельным понятиям (терминам) индексов – лексических единиц ИПЯ (в том числе цифровых или символьных кодов, если они предусмотрены).

Синонимы: Аннотирование, Формирование метаданных.

Интеллектуальный капитал (ИК) – форма капитализации интеллектуального потенциала. ИК может быть разделен на *Человеческий капитал, Капитал отношений* и *Организационный (или структурный) капитал*.

Интеллектуальные ресурсы (ИР) – комплексная категория, объединяющая интеллектуальный капитал людей и различные формы нематериальных активов, фиксирующие знания и профессиональные умения (объективированные знания).

Интеллектуальный анализ данных (ИАД) – Data Mining, Knowledge Discovery in Databases – технология извлечения новых знаний из хранилищ данных путем выявления закономерностей в накопленной информации.

Информационная потребность [Information demand, Information need, Information requirement] – состояние отдельного лица, коллектива или системы, характеризующееся необходимостью получения информации для успешного достижения каких-либо целей или выполнения работы.

Информационно-поисковый язык, ИПЯ [information retrieval language, indexing language] – искусственный язык, предназначенный для формализованного описания смыслового содержания документов, данных, отдельных понятий или терминов и обеспечения последующего их поиска в информационно-поисковых массивах.

Формализация лексики и создание различных ИПЯ вызвано необходимостью устранения «избыточности» и «недостаточности» естественного языка для целей информационного поиска, а также

ликвидации присутствующих в нем синонимии и омонимии (см. далее) для реализации однозначности информационного поиска.

Онтология выполняет роль ИПЯ в системах управления знаниями.

Запрос (query, request, interrogation) – входное сообщение в автоматизированную систему, содержащее требование на выдачу информации или (в более общем случае) на выделение ресурсов.

Знания – это результаты обобщения фактов и установления определенных закономерностей в какой либо предметной области, которые позволяют ставить и решать задачи в этой области.

Капитал отношений – совокупность наработанных связей организации с окружением, прежде всего с клиентами. Является частью *Интеллектуального капитала*.

Каталог компетенций – карта, устанавливающая связь между сотрудниками организации и их компетенциями, которая позволяет пользователю находить людей, имеющих требуемые знания, навыки и умения.

Карта знаний – модель, как правило, графическая, отражающая распределение элементов знаний между различными объектами организации, такими как организационная единица, функция, процесс, местонахождение и т. п. Дополнительно на карте устанавливается степень покрытия (coverage) элементом знаний соответствующей потребности.

Компетенции – способности применять знания, умения и личностные качества для успешной деятельности в определенной области.

Контент – содержимое (или наполнение) информационного ресурса (например, корпоративного портала, хранилища, документа), включающее в себя текст, графику, мультимедиа.

Корпоративная память – множество формализованных и неформализованных знаний организации, включая знания о знаниях, включающее в себя аннотированные документы, аннотированные базы данных, базы знаний, компетенции персонала и их описание в виде профилей или каталогов. Часть корпоративной памяти с формализованными знаниями может называться *базой знаний (в широком смысле слова), репозиторием, хранилищем*.

Лучшая практика (best practice) – результат обобщения успешного практического опыта. Часто лучшая практика оформляется стандартным

образом и выполняет рекомендательную или нормативную функцию. В литературе по менеджменту английский термин часто переводится на русский как *передовой опыт*.

Метаданные – структурированные данные, представляющие собой характеристики описываемых сущностей для целей их идентификации, поиска, оценки, управления ими.

Мета-знания – знания о знании. Примерами мета-знаний являются *Диаграмма знаний, Карта знаний и Каталог компетенций*.

Навигация – метод поиска информации, при котором пользователь для получения информации перемещается по элементам информационно-поискового языка, либо переходит по гипертекстовым ссылкам от одной страницы к другой. Один из методов *поиска и распределения знаний*.

Объект знаний – средство выражения знаний.

Онтология – формальная спецификация разделяемой концептуальной модели, где:

под концептуальной моделью подразумевается абстрактная модель предметной области, описывающая систему понятий предметной области, под «разделяемой» подразумевается согласованное понимание концептуальной модели определенным сообществом (группой людей), «спецификация» подразумевает явное описание системы понятий, выполненное по определенным правилам, «формальная» подразумевает, что концептуальная модель понимается компьютером, то есть он может работать с ней по законам логики.

Организационная память – см. *Корпоративная память*.

Организационный капитал – способности организации, проистекающие из формализованных знаний, деловых процессов, организационной культуры, ценностей и норм. Синоним – *Структурный капитал*. Является частью *Интеллектуального капитала* организации.

Поиск (или поиск информации, или информационный поиск) – процесс нахождения, отбора и выдачи определенной заранее заданными признаками информации (в том числе документов, их частей и/или данных) из массивов и записей любого вида и на любых носителях. Предоставление возможностей поиска является одним из способов *распределения знаний*.

Поисковый образ документа [document description, search document

image] – описание содержания документа на *ИПЯ*, отражающее важные признаки его содержания и вида для реализации поиска данного документа.

Семантика [semantics] – часть определенного языка, касающаяся указания смысла и действия текста, составленного в соответствии с синтаксическими правилами этого языка.

Семантические технологии – информационные технологии, основанные на использовании семантической модели предметной области.

Синтаксис [syntax] – раздел грамматики, изучающий способы организации слов в словосочетаниях и предложениях, а также типы предложений, их значения и условия использования.

Структурный капитал – см. *Организационный капитал*.

Система управления знаниями, Комплексная – это совокупность процессов и инструментов УЗ, обеспечивающих целенаправленное развитие и эффективное функционирование процессов преобразования и сетей знаний.

Система управления знаниями, Программная реализация – это комплекс информационных технологий, обеспечивающий целенаправленное развитие и эффективное функционирование процессов преобразования и сетей знаний.

Управление знаниями – это система инструментов, техник и процессов для выявления, накопления, распределения, обмена, использования и оценки знаний на предприятии, направленная на предоставление нужных знаний, в нужной форме, в нужном количестве, в нужном месте, в нужное время и по приемлемой цене.

Фолксономия – (англ. folksonomy, от folk – народный + taxonomy таксономия, от гр. расположение по порядку + закон) – народная классификация, практика совместной категоризации информации (веб-страниц, документов, мультимедиа и т. п.) посредством произвольно выбираемых меток, называемых тегами.

Кудрявцев Дмитрий Вячеславович



Доцент кафедры «Компьютерные интеллектуальные технологии в проектировании» Санкт-Петербургского Государственного Политехнического Университета. Руководитель направления управления знаниями «Бизнес Инжиниринг Групп» (БИГ-СПб). Кандидат технических наук, тема диссертации: «Разработка моделей и методов обработки знаний в области организационного проектирования на основе онтологий». С отличием окончил Балтийский государственный технический университет ("Военмех") по специальности «Экономика и управление на предприятии машиностроения». MBAE (Master of

Business Administration and Engineering) – совместная программа Регионального Университета г. Будо (Норвегия) и БГТУ «Военмех». Прошел российско-американскую программу по бизнесу, организованную ESI (Educational Services International).

Имеет успешный опыт реализации проектов по стратегическому управлению, организационному развитию и управлению знаниями в компаниях и на промышленных предприятиях в Санкт-Петербурге и РФ, а также выполнения научно-исследовательских проектов федерального и регионального уровней.

Автор более 30 публикаций (журналы «Intelligent Enterprise», «Вопросы государственного и муниципального управления Public Administration», «Научно-технические ведомости СПбГПУ», «Information Theories & Applications» и др.; главы в 2-х монографиях; сайт bigc.ru). Ряд публикаций на английском языке. Выступал на научных и научно-практических конференциях в России, Украине, Болгарии, Италии, Австрии по управлению знаниями, организационному управлению, когнитивным наукам, системному анализу и искусственному интеллекту. Участвовал в четвертой международной летней школе по онтологическому инжинирингу и semantic web (SSSW'06). В 2004 году участвовал в создании интернет-портала по технологиям управления знаниями (<http://kmtec.ru/>). Преподает в Санкт-Петербургском государственном политехническом университете курсы «Системы управления знаниями» и «Проектирование интеллектуальных систем и инженерия знаний».

E-Mail: dmitry.ku@gmail.com