

УДК 004.82

Система управления знаниями: задачи, категории продуктов, архитектура, обзор решений

Яроц Е.В., студентка,
Россия, 105005, г.Москва, МГТУ им.Н.Э.Баумана,
кафедра «Автоматизированные системы обработки информации и управления»

Научный руководитель: Гапанюк Ю.Е., к.т.н., доцент,
Россия, 105005, г.Москва, МГТУ им.Н.Э.Баумана
sfm2007@yandex.ru
vyaroz@bmstu.ru

В современном обществе знания являются главным источником устойчивых конкурентных преимуществ предприятий, а управление знаниями — основным механизмом их достижения [1]. Любая автоматизированная система затрагивает проблемы хранения информации, но только система управления знаниями ориентирована на это в явном виде, тем самым способствует сохранению такого ценного ресурса как знания. База знаний должна обеспечивать накопление и повторное использование знаний. База знаний является основным компонентом системы управления знаниями.

На рисунке 1 представлен весь спектр задач, который относится к понятию «управление знаниями».



Рис. 1. Перечень задач, относящийся к понятию «управление знаниями»

Можно выделить три категории программных продуктов, которым подвластно решение указанных выше задач, они представлены на рисунке 2.

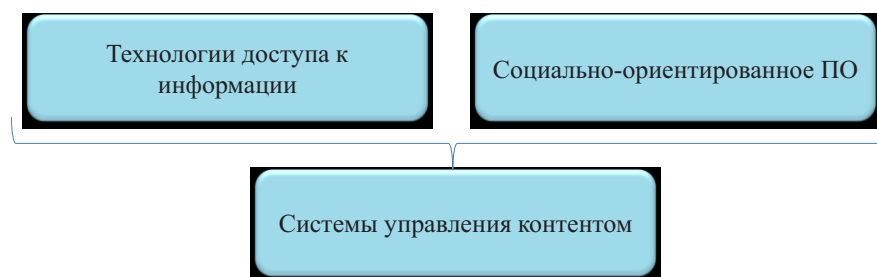


Рис. 2. Категории программных продуктов

Понятие «знание» может быть рассмотрено с двух точек зрения:

1. Знание как «запас». Согласно этому подходу знание является объектом, который способен существовать независимо от человека.
2. Знание как «поток». В этом подходе знание передается не как объект, а как продукт, в формировании которого участвует человек.

Подход «знание как запас» соответствует категории программных продуктов «Технологии доступа к информации», а подход «знание как поток» - «Социально-ориентированное ПО».

Категория «Технологии доступа к информации».

Из названия данной категории понятно, что она обеспечивает доступ к информации, а именно поиск по отдельным словам, запросы вида «Найти все документы проекта «Альта», которые составил Иванов?». Многие программные продукты рассматриваемой категории могут составлять классификации документов по тексту, а также аннотировать документ. Аннотирование, несомненно, очень полезная вещь, появляется возможность ознакомиться с кратким содержанием документа. Однако, ручное аннотирование эффективнее машинного, но в скором времени, благодаря развитию искусственного интеллекта, этот барьер может быть нарушен.

Категория «Социально-ориентированное ПО».

Программные продукты этой категории – это виртуальное единое пространство, включающее wiki, документооборот, средства для коммуникации пользователей. В таких программах каждый пользователь имеет свой профиль, в котором указаны его интересы, опыт работы, профессиональные навыки. Пользователи могут объединяться в группы для обсуждения отдельных тем или для совместной работы, например, для создания коллективного отчета.

Категория «Системы управления контентом».

Программы, которые относятся к данной категории, включают в себя различные инструменты для работы с документами, такие как сканирование, распознавание, управление версиями, создание архивов, наполнение, обновление, структурирование.

Самое важное в проблеме организации управления знаниями, разработать единые правила представления информации, которые помогут её дальнейшему структурированию. Сразу дадим определение онтологии [2]. Онтология – это структура, задающая отношения между понятиями. Таким образом, онтология определяет классы и отношения между ними. Введем понятие семантики. Семантика изучает законы смысла, её центральным понятием является значение. Благодаря развитию направления «управление знаниями» удастся структурировать не только информация внутри одной компании, но и информацию в глобальной сети Интернет. Самым эффективным инструментом для реализации последнего, является Семантический Веб. Семантический Веб – это расширение существующей глобальной сети, в которой информация снабжена определенным смыслом, что позволяет пользователю и компьютеру эффективно взаимодействовать. В основе Семантического Веба лежит использование онтологий. Задание классов, подклассов, а также отношений между объектами является мощным инструментом в Вебе. В простейшем случае, онтологии используют для увеличения точности поиска в Вебе — поисковая машина будет выдавать только такие сайты, где упоминается в точности искомое понятие, а не произвольные страницы, в тексте которых встретилось данное многозначное ключевое слово. В полную силу Семантический Веб будет реализован тогда, когда люди создадут множество программ - агентов, которые, знакомясь с содержимым Веба из различных источников, обрабатывают полученную информацию и обмениваются результатами с другими программами. Семантический Веб стимулирует подобного рода синергию: даже те агенты, которые не были созданы специально для работы сообща, могут передавать информацию друг другу, коль скоро эта информация будет снабжена семантикой. Использование онтологий позволяет осуществлять запросы на основе понятий, а не по совпадающим словам. Например, если пользователь напишет запрос, «Какие виды фруктов растут в России?», то он получит ответ, в котором будут сливы, яблоки (=подкласс фруктов), которые растут в Сочи (=город России). Также решается проблема использования в запросе слов, одинаковых по правописанию, но имеющих разные смысловые значения (например, слово «ключ»), благодаря присвоению словам семантического содержания. С помощью онтологий возможен поиск скрытой информации. Например, пользователь задает вопрос «Какие поставки находятся в зоне риска?». В ответ на запрос в онтологии тарифов система определит, что с учетом погодных условий (сейчас лето, жарко) существуют риски с перевозкой кондитерских изделий. В другой онтологии, где хранится информация по перевозкам, определит, что в декларации № A123 написаны конфеты, которые являются

подклассом класса «Кондитерские изделия». В результате система даст ответ: поставка № A123.

Для большего понимания, что такое система управления, приведем ее архитектуру. Она представлена на рисунке 3.

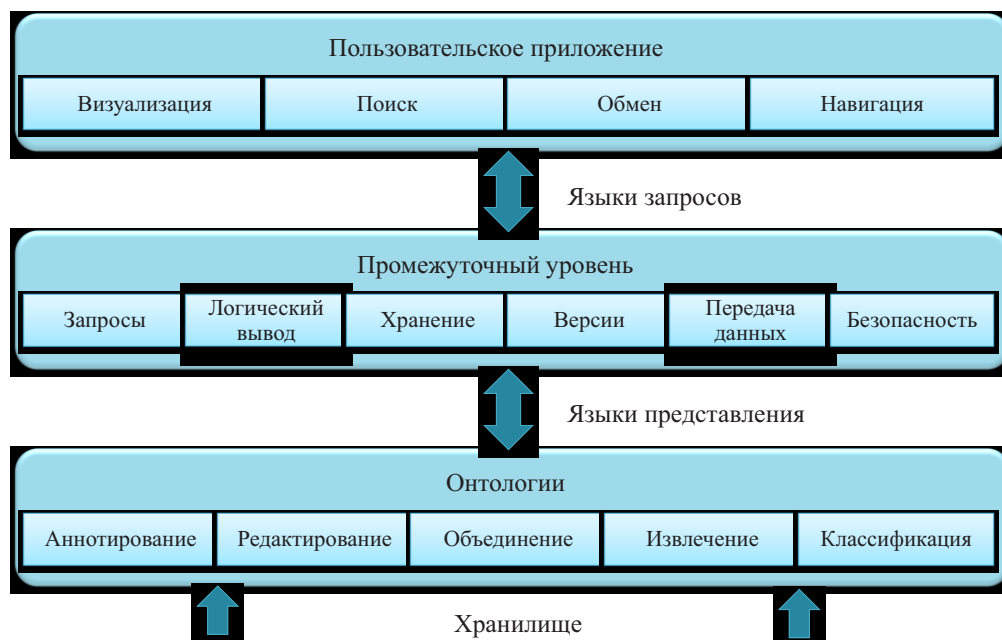


Рис. 3. Общая архитектура системы управления знаниями

Уровень «Онтологии» обеспечивает получение знаний из Хранилища. Под знаниями здесь подразумеваются аннотации из разных источников информации, например документов [3]. Аннотации создаются автоматически, документ не хранит в себе краткое описание. Источники информации могут быть структурированные (таблицы, диаграммы) или неструктурированные (документы). «Промежуточный уровень» обеспечивает хранение онтологий, управление версиями, передачу данных, взаимодействие с Хранилищем, логический вывод и обработку запросов. Уровень «Пользовательское приложение» обеспечивает доступ пользователя к знаниям. Доступ может осуществляться путем обмена информацией между пользователями, с помощью запроса к системе. В реальности, архитектура системы управления знаниями (рис. 3) может быть воплощена частично, например, акцент будет сделан на разработку уровня «Пользовательское приложение». В этом случае система будет являться «порталом знаний». Если сделать упор на разработку «Промежуточного уровня», то такие системы будут носить название «базы знаний» или «корпоративной памяти». Далее введем понятие «корпоративная память». Корпоративная память – это хранилище, в котором находится база данных, содержащая все версии базы данных. За счет такого хранения изучение и анализ связей между объектами происходят наиболее эффективно, так как все хранятся в

одном месте. Существуют корпоративные порталы знаний, которые обеспечивают доступ к корпоративной информации. Часто на порталах применяются карты знаний. Карта знаний служит навигационным помощником для поиска. Карта знаний описывает источники, потоки и границы знаний организации. Карты знаний могут указывать как на людей, так и на документы и базы данных. Очевидные преимущества карт знаний заключаются в том, что достигается минимизация объема информации, большая ясность, наглядность и доступность. В качестве базы знаний можно привести известный пример – это wiki. Благодаря вики-движкам, расширениям, например, Semantic MediaWiki [4], это уже не просто база знаний, а скорее система управления знаниями. Перечислим функции расширения Semantic MediaWiki:

1. Оно позволяет поддерживать согласованность данных на вики.
2. С его помощью можно обращаться с вики как с базой данных.
3. С его помощью можно генерировать красивые графики, диаграммы и графы на основе вики-данных.
4. Оно позволяет пользователям возможность вводить данные с помощью форм, а не вики-разметки, понижая тем самым порог вхождения для пользователей.

Semantic MediaWiki позволяет использовать wiki как составную часть системы управления знаниями, с его помощью можно создавать системы, которые сочетают простоту wiki и возможности управления знаниями. В основу Semantic MediaWiki положена концепция семантических свойств (для создания данных) и семантических запросов (для использования данных). Пользователи аннотируют статьи категориями (создают тэги) и свойствами для того, чтобы информация становилась доступной для запросов. В качестве примера рассмотрим wiki, посвященную путешествиям, в которой будут описаны города, достопримечательности и кафе. Wiki содержит множество статей на разные темы.

Страницы wiki, соединенные гиперссылками, представим в виде графа (рис. 4).

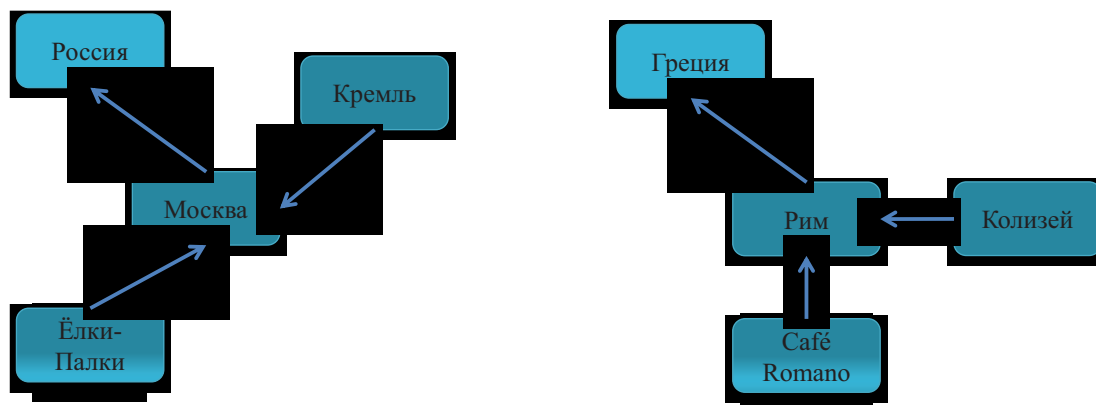


Рис. 4. Граф, отражающий страницы wiki, соединенные гиперссылками

С помощью категорий MediaWiki, мы можем объединить статьи из примера в несколько множеств. Категории могут вкладываться друг в друга, образуя иерархии. Для того чтобы отнести страницу к определенной категории добавим следующий код в текст статьи:

```
[[Категория:Название_категории]]
```

Например, выделим следующие категории: Город, Страна, Населенная область, Кафе, Достопримечательность, показанные на рисунке 5.

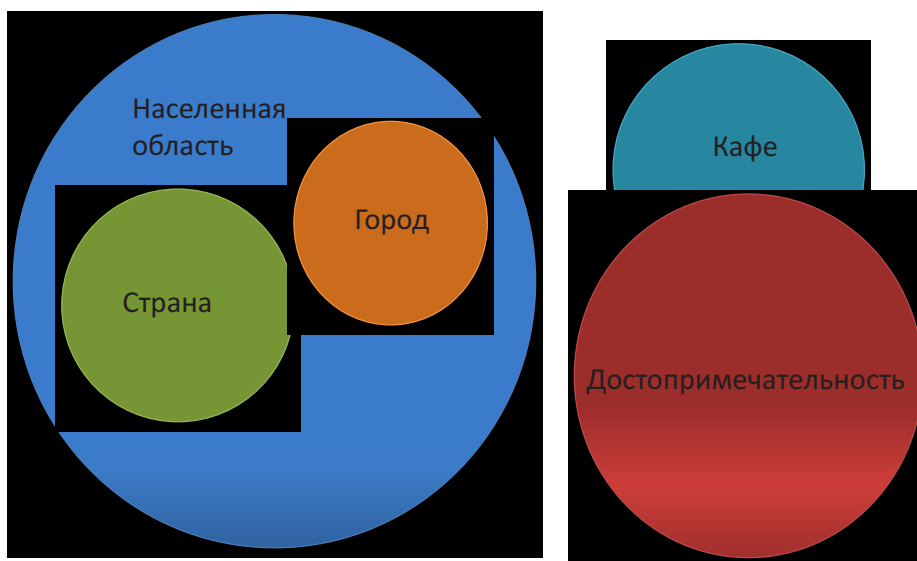


Рис. 5. Категории

Каждой категории соответствует свой цвет. В итоге наш граф преобразуется в следующий, представленный на рисунке 6.

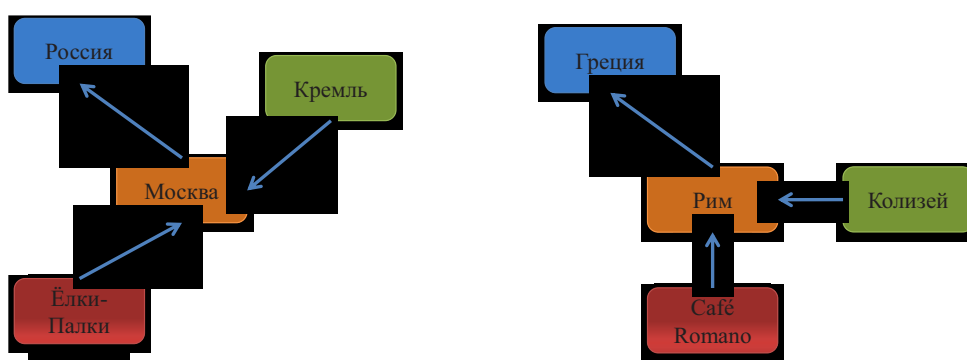


Рис. 6. Преобразованный граф

Категории являются базовым механизмом структурирования контента в MediaWiki. Можно делать динамические выборки по категориям, например, вывести все кафе, которые также являются достопримечательностями [5].

Информационные технологии не решают всех проблем управления знаниями, напротив, они могут породить и дополнительные сложности – при разработке и внедрении. Однако, при умелом использовании, программные средства управления знаниями могут помочь перевести инфраструктуру знаний компании на новый материально-технологический уровень, и превратить ее в мощный инновационный ресурс.

Список литературы

1. Кудрявцев Д.В. Системы управления знаниями и применение онтологий: учеб. пособие / Д.В. Кудрявцев. СПб.: Изд-во Политехн. ун-та, 2010. 344 с.
2. Джанетто К., Уилер Э. Управление знаниями. Руководство по разработке и внедрению корпоративной стратегии управления знаниями / Пер. с англ. Е.М. Пестеровой М.: Добрая книга, 2005. 195 с., илл.
3. Гаврилова Т.А. Модели структурирования и инструменты управления знаниями // Системы управления знаниями»: сб. докладов XI-й научно-практической конференции «Реинжиниринг бизнес-процессов на основе современных информационных технологий. М., 2008. С.47-50.
4. Интерфейс прикладного программирования. Категории MediaWiki. Режим доступа: <http://www.mediawiki.org/wiki/Help:Categories/ru.html> (дата обращения 03.03.2014).
5. Бизнес в паутине социальных сетей // Инновации в технологиях и бизнесе, №5, 2007. Режим доступа: <http://www-306.ibm.com/software/ru/mag/5/article.html> (дата обращения 03.03.2014).