

УДК 004.822

UDC 004.822

ОТКРЫТАЯ СЕТЬ ДОКУМЕНТИРОВАННЫХ ЗНАНИЙ ОРГАНИЗАЦИИ: КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ¹**OPEN NETWORK OF FORMAL CORPORATIVE KNOWLEDGE: A CONCEPTUAL MODEL**

Савченко Андрей Павлович

к. физ.-мат. н., доцент

Кубанский государственный университет, Краснодар, Россия

Savchenko Andrey Pavlovich

Cand.Phys.-Math.Sci., associate professor

Kuban State University, Krasnodar, Russia

В условиях формирования экономики знаний актуален вопрос эффективного использования и воспроизводства знаний в социально-экономических системах. Один из перспективных подходов к созданию действующей системы управления знаниями основан на организации открытой онтологической сети знаний. Сеть знаний есть модель процесса обмена знаниями, реализованная в виртуальном пространстве, наследующая достоинства этого процесса и устраняющая ряд его недостатков. Онтология – базовый компонент инфологической модели сети знаний, она описывает систему знаний и определяет структуру хранения информации в сети. Иерархия онтологий предполагает наличие трех уровней: метаонтология, базовая и предметная онтология. Использование онтологии решает две задачи: дает инструмент семантической разметки информационных ресурсов и раскрывает структуру информационного хранилища для пользователя. Вторая задача особенно важна для обеспечения эффективного поиска в больших информационных массивах. Технология создания открытой сети включает разработку ряда принципов и инструментов для наполнения и модификации предметной онтологии, быстрой семантической разметки ресурсов, поиска и выдачи ресурсов с учетом их семантических связей

Within emerging knowledge economy, the current issue is efficient use and reproduction of knowledge in socio-economic systems. One of the promising approaches to the creation of working knowledge management system is based on the open ontological knowledge network. The knowledge network is the model of the process of knowledge sharing, implemented in the virtual space, which inherits the advantages of this process and removes some of its shortcomings. Ontology is a basic component of the conceptual model of the knowledge network; it describes a system of knowledge and defines the structure of data storage in the network. The hierarchy of ontologies involves three levels: metaontology, basic and subject-specific ontology. The ontology use solves two tasks: providing a tool for semantic markup of information resources and disclosing the structure of the information storage for the user. The second task is especially important to ensure an efficient search in large data arrays. The technology of construction of open network includes a set of principles and tools for filling and updating of the domain ontology, fast semantic markup and retrieval of information resources taking into account their semantic links

Ключевые слова: ИНФОЛОГИЧЕСКАЯ МОДЕЛЬ, СЕТЬ ЗНАНИЙ, ОНТОЛОГИЯ ПРЕДМЕТНОЙ ОБЛАСТИ, СЕМАНТИЧЕСКАЯ РАЗМЕТКА, ИНФОРМАЦИОННЫЕ РЕСУРСЫ

Keywords: INFOLOGICAL MODEL, KNOWLEDGE NETWORK, DOMAIN ONTOLOGY, SEMANTIC MARKUP, INFORMATION RESOURCES

Введение

В условиях перехода к экономике, основанной на знаниях, особенно остро встает вопрос эффективного использования и воспроизводства корпоративных знаний в социально-экономических системах разных уровней. Корпоративными знаниями принято называть совокупность внешних и внутренних, явных и неявных знаний, циркулирующих в организации. К

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 14-07-31145 мол_a).

внутренним корпоративным знаниями традиционно относят знания сотрудников, ноу-хау, опыт прошлых проектов («корпоративная память»), знания о построении отношений с контрагентами и т.п.

Можно выделить две группы проблем, снижающих эффективное использование корпоративных знаний на практике:

- 1) разнородность и слабая структурированность информационных ресурсов организации;
- 2) организационные и личностные сложности в процессе неформального обмена знаниями между сотрудниками.

Одним из методов решения указанных проблем является использование информационной системы управления знаниями на базе открытой сети знаний. В данной статье обсуждается концептуальная модель открытой сети знаний, описываются технологии и принципы ее построения.

Открытая сеть знаний: обсуждение концепции

На наш взгляд, наиболее реалистичный подход к созданию действующей системы управления знаниями заключается в организации открытой сети документированных знаний. Под *открытой сетью знаний* мы понимаем распределенное хранилище документированных знаний (данных, метаданных, регламентов, бизнес-правил, спецификаций бизнес-процессов и других элементов), снабженное инструментами поиска и формализации знаний и, самое главное, обеспечивающее не только использование, но и пополнение знаний пользователями сети.

Идея открытой информационной сети не нова. Наиболее успешным примером ее реализации может служить международный интернет-проект «Википедия», существующий уже 13 лет. На сегодняшний день это крупнейший энциклопедический ресурс в мире. С момента своего появления этот проект подвергается серьезной критике сторонниками «классических» энциклопедий, однако практика показывает, что начиная с 2009 г. Вики-

педия является самым посещаемым научно-информационным ресурсом и шестым по посещаемости среди всех ресурсов в Интернете [1]. Если еще пять-шесть лет назад на Википедию ссылались в основном в студенческих работах, то сегодня на Википедию как источник статистических и справочных данных ссылаются СМИ, авторы высоко цитируемых научных публикаций [2-4] и даже сайты правительственных учреждений. Отметим, что теоретические исследования, проведенные в 2008 г. показали, что рост Википедии является устойчивым [5], что также говорит в пользу перспективности подобного подхода к документированию знаний.

Среди главных объектов критики открытого подхода – потенциальная неточность и непроверяемость сведений, добавляемых пользователями, ввиду неподтвержденной квалификации авторов, бесконтрольного процесса наполнения и отсутствия традиционного научного рецензирования статей.

Однако, на наш взгляд, данные в открытой сети нельзя назвать непроверяемыми. Любая публикуемая информация просматривается другими пользователями, и в первую очередь теми, чьи интересы и компетенции лежат в данной области знаний. При обнаружении некорректной информации пользователь может удалить или исправить ошибку с минимальными усилиями. Вообще, развитие подобной системы будет устойчивым в том случае, если количество «активных вандалов» составляет менее половины аудитории проекта [5]. На наш взгляд, в случае корпоративной или вузовской сети логично предположить, что большая часть пользователей не будет умышленно искажать данные, тем более, что все действия в сети могут протоколироваться и выявить вредителя довольно просто.

Характерно, что открытая природа сети знаний в большей степени соответствует структуре социально-экономических и гуманитарных знаний, чем знаний естественнонаучных. Дело в том, что знание в общественных науках носит эмпирический, ситуативный и зачастую субъективный харак-

тер. Опыт становления естественных наук показал, что для формирования системы объективных (в некотором приближении) знаний требуются столетия научных исследований. Социальные системы имеют более сложную природу, чем системы физические, поэтому большинство гуманитарных исследований относятся к слабоформализованным областям. Например, единой теории управления организацией до сих пор не создано, хотя существует множество примеров успешного управления. Принципы управления социальными системами носят рекомендательный характер и справедливы только в определенных уникальных условиях, т.е. они изначально необъективны.

Использование открытой сеть знаний основано на доверии к знаниям пользователей, участвующих в наполнении сети, так же как рабочие процессы общения, обучения и неформального обмена опытом в коллективе основаны на доверии участников коммуникации к мнению друг друга. Пользователи открытой сети могут играть две роли:

- читатель – участник, который осуществляет поиск и изучение информационных ресурсов;

- автор – участник, который является создателем или редактором информационных ресурсов сети. Структура информационных потоков в сети знаний представлена на рис. 1.

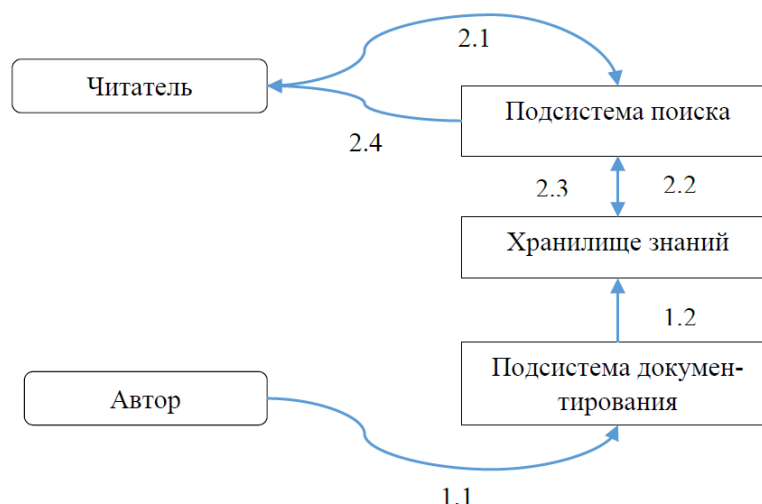


Рисунок 1 – Структура информационных потоков в сети знаний
 (1.1 – знания эксперта; 1.2 – документированные знания (информационные ресурсы); 2.1 – поисковый запрос; 2.2 – обработанный поисковый запрос; 2.3, 2.4 – релевантные информационные ресурсы)

Таким образом, открытая сеть знаний по сути представляет собой информационную модель процесса обмена знаниями в коллективе, наследуя большинство его достоинств. При этом предлагаемый подход имеет ряд преимуществ перед традиционными механизмами коммуникации в коллективе.

1. Всеобщий охват. Наличие посредника в виде сети знаний между источником и реципиентом знания предоставляет всем пользователям равные шансы получить интересующие их сведения, при этом существенно снижается влияние субъективных психологических и личностных качеств участников, таких как усталость, личная неприязнь, боязнь обращения к авторитетному эксперту и др.

2. Постоянная доступность. Документированные знания доступны в любое время и практически из любой точки. Использование сетью знаний избавляет пользователя от необходимости договариваться о встрече с экспертом, отрывать его от основной работы и т.д.

3. Стандартизация языка информационного обмена. Зачастую процесс обмена информацией затруднен из-за того, что глоссарий пользователя не совпадает с глоссарием эксперта, т.е. пользователь не может правильно сформулировать вопрос к эксперту, а эксперт затрудняется в объяснении сложной информации пользователю. Соблюдение синтаксических и стилевых стандартов при документировании знаний и их поиске позволит решить эту проблему.

Таким образом, реализация в организации открытой сети знаний позволит усовершенствовать коммуникационные процессы в коллективе, обеспечит условия для создания единой информационной среды, способствующей творческому мышлению и развитию человеческого и интеллектуального капитала организации. Рассмотрим подробнее технологию организации открытой сети знаний.

Инфологическая модель открытой сети знаний

Выбор концептуальной модели информационной сети и средств ее реализации определяется требованиями, предъявляемыми к этой сети. Учет недостатков традиционных коммуникаций позволил сформулировать следующие требования к сетевой модели коммуникации:

- унифицированное представления разнородных знаний и данных;
- учет связанности информационных единиц;
- обеспечение эффективного поиска данных и знаний.

Для выполнения поставленных условий предлагается концептуальная модель, базирующаяся на многослойной онтологии предметной области.

Под онтологией в данном случае понимается структурная спецификация предметной области, т.е. формализованное представление основных понятий и связей между ними [6]. Базовым компонентом информационной модели сети знаний служит многослойная онтология, она описывает систему знаний и определяет структуру хранения информации в сети знаний.

Авторская концепция многослойной онтологии подробно описана в работе [7].

Структура онтологии предполагает наличие трех уровней. Верхний уровень – метаонтология Дж. Сова, моделирующая основополагающие понятия модели знаний, такие как «объект», «система», «процесс» и т.п.

Главным ее отличием от других онтологий верхнего уровня выступает отсутствие жесткой структуры корневых понятий и связей между ними. Вместо этого онтология предлагает совокупность базовых свойств (характеристик), которые можно использовать для описания этих корневых сущностей. Выбирая тот или иной набор характеристик, создатель онтологии может формировать собственную структуру сущностей, а классифицирующие связи между ними определяются структурой связей между базовыми характеристиками. Благодаря такому подходу реализуется возможность совмещения на верхнем уровне онтологий, созданных разными исследователями.

Онтология первого уровня представляет собой модель методологических знаний, описывает структуру и содержание процесса научного исследования. В качестве методологической основы для создания этой онтологии используется теоретико-познавательный подход В.Г. Щедровицкого [8].

Онтология второго уровня – предметная онтология, описывающая знания выбранной предметной области (ПрО).

Использование разделенной онтологии упрощает настройку сети на выбранную предметную область, позволяет разрабатывать универсальные онтологии высшего и первого уровня и впоследствии применять их для решения различных задач.

Онтологии всех уровней имеют объектно-ориентированную структуру и соответствуют единой концептуальной модели (рис. 2).



Рисунок 2 – Инфологическая модель онтологической сети знаний

Основными сущностями онтологии служат классы и экземпляры классов – объекты. Между сущностями существует два типа связей:

- отношения «класс-класс» отражают связи между классами, этот тип отношений характерен, прежде всего, для онтологии первого уровня;
- отношения «объект-объект» выражают свойства конкретного объекта и могут быть двух видов: обязательные отношения, унаследованные от класса-предка (классовые признаки) и персональные связи, характерные только для данного экземпляра класса. Подробно методика построения онтологии описана в работе [7]

Онтология сети используется для обеспечения связности информационного контента сети. Кроме того, онтология имеет собственную ценность, поскольку содержит тезаурус предметной области и систему связей между концептами ПрО, т.е. онтология, наряду с контентом сети, несет в себе сведения о предметной области.

Известно, что правильно сформулированный вопрос содержит в себе половину ответа. Однако во многих случаях пользователи информационных систем испытывают затруднения как раз с формулировкой правильного запроса. Под правильным в данном случае понимается запрос, позволяющий получить результат, наиболее удовлетворяющий потребностям пользователя. Так как релевантность результатов во многом зависит от технологии организации данных в базе и методов их поиска, то правильным запросом считается *запрос, наиболее соответствующий структуре хранилища знаний информационной системы*. Трудности неопытных пользователей во многом объясняются отсутствием знаний об организации системы знаний ПрО в информационной системе (в том числе: информационно-поискового языка, терминов предметной области и связей между ними и т.д.) Использование онтологий позволяет открыть пользователю принципы организации информационных ресурсов сети. Онтология доступна для просмотра пользователям, благодаря чему упрощается навигация в сети, становится возможным поиск информации даже при отсутствии четкого запроса. Пользователь начинает поиск с рассмотрения общих понятий, концептов и в результате перехода по связям онтологии (т.е. постепенного уточнения запроса) добирается до нужной информации. Такой подход сопоставим с традиционным системным подходом к анализу проблемы, и потому представляется нам наиболее соответствующим принципам научного познания и обучения.

Технологии построения открытой сети знаний

В структуре сети знаний можно выделить четыре уровня (слоя) (рис. 3). Каждый уровень соответствует определенному набору технологий, обеспечивающих реализацию сети знаний.



Рис. 3. Иерархия технологий построения открытой сети знаний

Технологический слой представляет собой стек сетевых технологий, образующий инфраструктуру сети знаний. Состав этого слоя определяется современными стандартами построения вычислительных сетей, имеет высокую степень однородности на уровне протоколов. В зависимости от масштабов и характера сети знаний в качестве технологической базы могут служить как корпоративные, так и глобальные сети. На наш взгляд, обозначенным принципам открытости и стандартизации при создании сети знаний вполне отвечает глобальная сеть Интернет на базе стека протоколов TCP/IP, именно ее рекомендуется использовать в большинстве случаев.

Информационный слой не так унифицирован, как технологический, здесь существует множество стандартов представления данных, хотя наиболее распространенным многие годы остается стандарт передачи данных HTTP и прикладные сервисы на его основе.

В области стандартов семантического и онтологического моделирования можно выделить как минимум две укрупненных группы технологий разработки онтологической, основанных на языках описания онтологий OWL и Schema.org. Оба подхода подробно регламентированы, стандартизированы и имеют многочисленных последователей в среде разработчиков информационных ресурсов. Однако они имеют ряд общих недостатков, основным среди которых, на наш взгляд, является сложность процесса создания и наполнения онтологий. Для устранения этой проблемы автором предложена упрощенная методика построения онтологии на базе языка SXML [7], которая позволяет снизить сложность процесса описания предметной области и распределить его во времени и в пространстве, сделать процесс наполнения онтологии поэтапным и удобным для пользователя.

Заключение

Итак, создание корпоративной открытой сети знаний позволит усовершенствовать процессы обмена знаниями в коллективе за счет создания единой информационной среды организации.

Важным элементом сети знаний и информационной системы управления знаниями на ее служит онтология предметной области. Использование онтологии решает две задачи: дает инструмент семантической разметки информационных ресурсов и раскрывает структуру информационного хранилища для пользователя. Вторая задача особенно важна для обеспечения эффективного поиска в больших информационных массивах.

Технологии построения сети знаний разделены на 4 уровня: технологический, информационный, семантический и прикладной. Реализация двух первых уровней обеспечивается использованием стандартных сетевых протоколов, для реализации третьего уровня предлагается авторская методика создания онтологии с помощью упрощенного языка онтологического моделирования SXML. Разработка прикладных инструментов рабо-

ты с документированными знаниями является перспективной задачей, решение которой является следующим этапом авторской научно-исследовательской работы.

Библиографический список

1. Alexa: статистика посещения сайта Wikipedia.org [Электронный ресурс]. – Режим доступа: <http://www.alexa.com/siteinfo/wikipedia.org>.
2. Назарчук А.В. Сетевое общество и его философское осмысление // Вопросы философии. 2008. № 7. С. 61-75.
3. Подвальный С.Л. Многоальтернативные системы: обзор и классификация // Системы управления и информационные технологии. 2012. Т. 48. № 2. С. 4-13.
4. Nakamura K. et al. Review of particle physics // Journal of Physics G: Nuclear and Particle Physics. 2010. Vol. 37. № 7A. P. 075021.
5. Spinellis D., Louridas P. The collaborative organization of knowledge // Communications of the ACM. 2008. Vol. 51. № 8, P. 68–73.
6. Gruber T. Ontology // Encyclopedia of Database Systems / Ed. by Ling Liu, Tamer Özsu M. Springer-Verlag, 2009.
7. Савченко А.П. Упрощенная методика построения многослойной онтологической модели предметной области // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №100. – Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/21.pdf>.
8. Щедровицкий Г.П. Проблемы методологии системного исследования. М., 1964.

References

1. Alexa: statistika poseshcheniya sayta Wikipedia.org [Elektronnyy resurs]. – Rezhim dostupa: <http://www.alexa.com/siteinfo/wikipedia.org>.
2. Nazarchuk A.V. Setevoe obshchestvo i ego filosofskoe osmyslenie // Voprosy filosofii. 2008. № 7. S. 61-75.
3. Podval'nyy S.L. Mnogoal'ternativnye sistemy: obzor i klassifikatsiya // Sistemy upravleniya i informatsionnye tekhnologii. 2012. T. 48. № 2. S. 4-13.
4. Nakamura K. et al. Review of particle physics // Journal of Physics G: Nuclear and Particle Physics. 2010. Vol. 37. № 7A. P. 075021.
5. Spinellis D., Louridas P. The collaborative organization of knowledge // Communications of the ACM. 2008. Vol. 51. № 8, P. 68–73.
6. Gruber T. Ontology // Encyclopedia of Database Systems / Ed. by Ling Liu, Tamer Özsu M. Springer-Verlag, 2009.
7. Savchenko A.P. Uproshchennaya metodika postroeniya mnogoslnoynoy ontologicheskoy modeli predmetnoy oblasti // Nauchnyy zhurnal KubGAU [Elektronnyy resurs]. – Krasnodar: KubGAU, 2014. – №100. – Rezhim dostupa: <http://ej.kubagro.ru/2014/06/pdf/21.pdf>.
8. Shchedrovitskiy G.P. Problemy metodologii sistemnogo issledovaniya. M., 1964.