

УДК 519.8

М.В. Новожилова, А.В. Ушеров-Маршак, Е.В. Латорец, И.А. Михеев

Харьковский государственный технический университет строительства и архитектуры

РАЗРАБОТКА ИНФОРМАЦИОННОЙ ТЕРМИНОЛОГИЧЕСКОЙ СИСТЕМЫ В ОБЛАСТИ БЕТОНОВЕДЕНИЯ

Построена информационная система «Лексикон Бетонovedение» для обеспечения поддержки принятия решений широким кругом специалистов строительных специальностей.

Ключевые слова: информационная система, терминологический словарь.

Введение

Одним из способов качественного повышения уровня подготовки высококвалифицированных специалистов в системе образования и науки является создание, внедрение и использование современных информационных технологий, интеллектуальных систем обработки данных и знаний в различных предметных областях.

В последние годы в области бетоноведения произошли значительные изменения. Появление новых составов, определяющих новые свойства бетона, технологий производства и сфер применения бетона как основного строительного материала обусловили необходимость проведения логического анализа, структурирования, создания терминологической системы в области бетоноведения.

Проведение такой работы усложняется слабой формализованностью предметной области, зависимостью от оценок экспертов, наличием множества различных трактовок, разрозненностью и зачастую противоречивостью сведений в имеющихся на сегодня немногочисленных публикациях в данной области.

Анализ имеющихся публикаций. Проведенный анализ перечня изданной за последние 10 лет в Украине специальной научной литературы показал, что на сегодня полноценные терминологические словари в области бетоноведения отсутствуют. В имеющихся словарях по общестроительной тематике данный раздел представлен слабо и не отражает всего многообразия современного состояния вопроса. На основании мониторинга информационных порталов в сети «Интернет» можно сделать вывод, что имеющаяся информация (например, портал «Бетон.ги») – носит скорее коммерческий, а не научно-образовательный характер и при этом сильно «зашумлена» рекламой.

Постановка задачи. Целью настоящего исследования является создание онтологии бетоноведения в виде, доступном для изучения и использования специалистами в данной предметной области, как в практической работе, так и в процессе обучения.

Пространственная распределенность научных и образовательных центров по территории стран СНГ, динамика развития бетоноведения, соответствие непрерывно повышающимся требованиям практических приложений и уровню международных стандартов – эти и другие факторы определили выбор инструментальных средств проектирования создания и сопровождения информационных технологий (ИТ) в качестве базовых.

Результатом функционирования ИТ непосредственно на уровне восприятия человека – пользователя выступает квант повышения уровня знаний или снижения неопределенности (энтропии) знаний пользователя в области бетоноведения.

Основной материал исследования

На теоретико-множественном уровне терминологический фундамент бетоноведения может быть представлен как целенаправленная система S вида:

$$S = \langle M, R, P \rangle, \quad (1)$$

где M – множество элементов системы: терминов и их определений, представляющих понятия (знания) предметной области; R – множество связей между элементами системы – терминами бетоноведения; P – множество свойств системы, позволяющее достичь заданной цели.

Термин, как элемент терминологической системы должен быть:

1) свободен от субъективности жизненного опыта. Недопустимо, чтобы разные исследователи при произнесении одного термина представляли себе разные вещи;

2) однозначен. Недопустимо, чтобы один и тот же термин одной и той же науки описывал в разных случаях разные объекты;

3) должен иметь точно определенную область значений, то есть должно существовать строго определенное множество объектов, описываемых этим термином.

Важным свойством элемента терминологической системы является то, что каждый термин имеет строгое определение, и для понимания термина необходимо знать как его собственное определение,

так и определения всех терминов, использованных в его определении, вплоть до базовых, неопределяемых, понятий.

Связи между элементами системы отражают иерархию понятий в их соотношении друг с другом. Связи между терминами, как и сами термины, определяются отраслевыми специалистами (экспертами) совместно с логиками и лингвистами. Характер установленных связей может быть различным.

Терминология является открытой системой; происходит постоянное наполнение ее новыми элементами при угасании или полном исчезновении других элементов.

Каждое из частных свойств $\{p_1, p_2, \dots, p_n\}$ характеризует локальное функциональное качество (например, p_1 – полнота, p_2 – открытость, p_3 – точность, p_4 – избыточность словаря), а вместе они достаточно полно характеризуют систему как целое.

В данной работе в качестве основного инструментального средства ИТ выступает информационная система (ИС). ИС рассматривается как системная совокупность программных и технических средств, а также реализованного банка данных (банка знаний), позволяющих с помощью специально разработанных в рамках системы методов, методик и нормативных ограничений (стандартов) эффективно в интересах и по запросам пользователя автоматически и однозначно поддерживать сбор, поиск, распознавание, получение, хранение, защиту, обработку и передачу информации.

Определена трехуровневая архитектура системы (рис. 1), отражающая и реализующая модель, структуру, выполняемые функции и взаимосвязь всех компонентов ИС, а также самой ИС с окружающей ее информационной средой.



Рис. 1. Архитектура ИС «Лексикон «Бетонovedение»

Первичной структурной единицей в системе является целостный информационный объект, а именно статья словаря (комплексная семантическая единица). При этом часть объекта, например, заголовок статьи (элементарная семантическая единица – ЭСЕ – неделимая идентифицируемая уникальным именем информационная семантическая единица), может физически храниться в одной базе данных (файле данных) нижнего уровня, а текст статьи или разрозненные фрагменты этого текста могут быть

размещены и быть найдены в другой базе (файле) данных, но на запрос внешнего интерфейса вся статья вместе с заголовком выдается целиком и одновременно в ее исходном виде.

Рассматриваемую систему можно классифицировать как вертикальный портал – портал узкой тематической направленности, предъявляющий различные сервисы для определенной целевой аудитории.

Реализация Уровня 3 архитектуры системы. Концептуальная модель базы знаний ИС построена на основании представления (1). Для обеспечения единого пространства знаний созданы онтологии в виде сопочиненности концептов (множества M терминов, понятий) предметной области бетоноведения и некоторого количества сопровождающих знаний (common knowledge), представленных системой аксиом-правил, задающих взаимосвязь терминов друг с другом.

Анализ различных моделей представления данных (семантические сети, сети фреймов, морфологическая энтропийная модель и др.) позволил обосновать выбор неоднородных **семантических сетей** как инструментального средства описания онтологии (представления знаний) на логическом уровне проектирования информационной системы.

Семантическую сеть рассматривают как маркированный ориентированный граф с помеченными узлами и дугами. Узлам соответствуют некоторые объекты, а дугам – семантические отношения между этими объектами. Метки, приписываемые узлам, выделяют множество рассматриваемых объектов и выступают в качестве их имен, в роли которых могут быть слова естественного языка. Метки, приписываемые дугам, соответствуют элементам множества отношений, заданных на графе.

Реализация семантической сети осуществлена на основе построения **тезауруса**, в котором отражена иерархия понятий. Применительно к тезаурусу рассмотрены 5 типов отношений: более широкий термин – выше; более узкий термин – ниже; связанный термин – ассоциация; целое для термина – часть; часть для термина – целое.

Отметим, что в общем случае онтологические соглашения по отраслям знаний и отраслям профессиональной деятельности имеют тенденцию развития до уровня отраслевых стандартов и руководящих технических материалов.

Построение Уровня 2 архитектуры системы основано на модели поиска текстовой информации, включающей:

- представление документов и запросов;
- критерий смыслового соответствия;
- методы ранжирования результатов запросов;
- механизм обратной связи для оценки релевантности документов.

Реализация данной модели связана с полнотекстовым поиском по всей коллекции документов. В качестве критериев эффективности поиска определены релевантность по отношению к запросу пользователя, актуальность поисковых индексов, надежность, быстродействие обслуживания пользователя.

Построение Уровня 1 архитектуры системы включает реализацию интерактивных средств доступа к данным, основанных на представлении информации в виде гипертекста и гипермедиа с использованием языка разметки HTML, а также более сложные процедуральные структуры, взаимодействующие с пользователем на уровне локальных приложений с достаточно развитой функциональностью.

В процессе выполнения заданий данного исследования было разработано техническое задание

на программную реализацию ИС и определены основные требования: удобство использования, эргономичный дизайн, масштабируемость, модульность, открытость, наличие поисковой системы, наличие справочной системы, наличие средств защиты информации, возможность подключения и просмотра: нормативной документации, библиотек материалов (учебники, монографии, диссертации и др.), интерфейсов других образовательных ресурсов различных уровней.

Данная работа проведена в соответствии с международным стандартом "ISO 12207:1995 – Процессы жизненного цикла программных средств" и другими регламентирующими документами.

На рис. 2 представлена концептуальная модель ИС «Лексикон «Бетонведение».

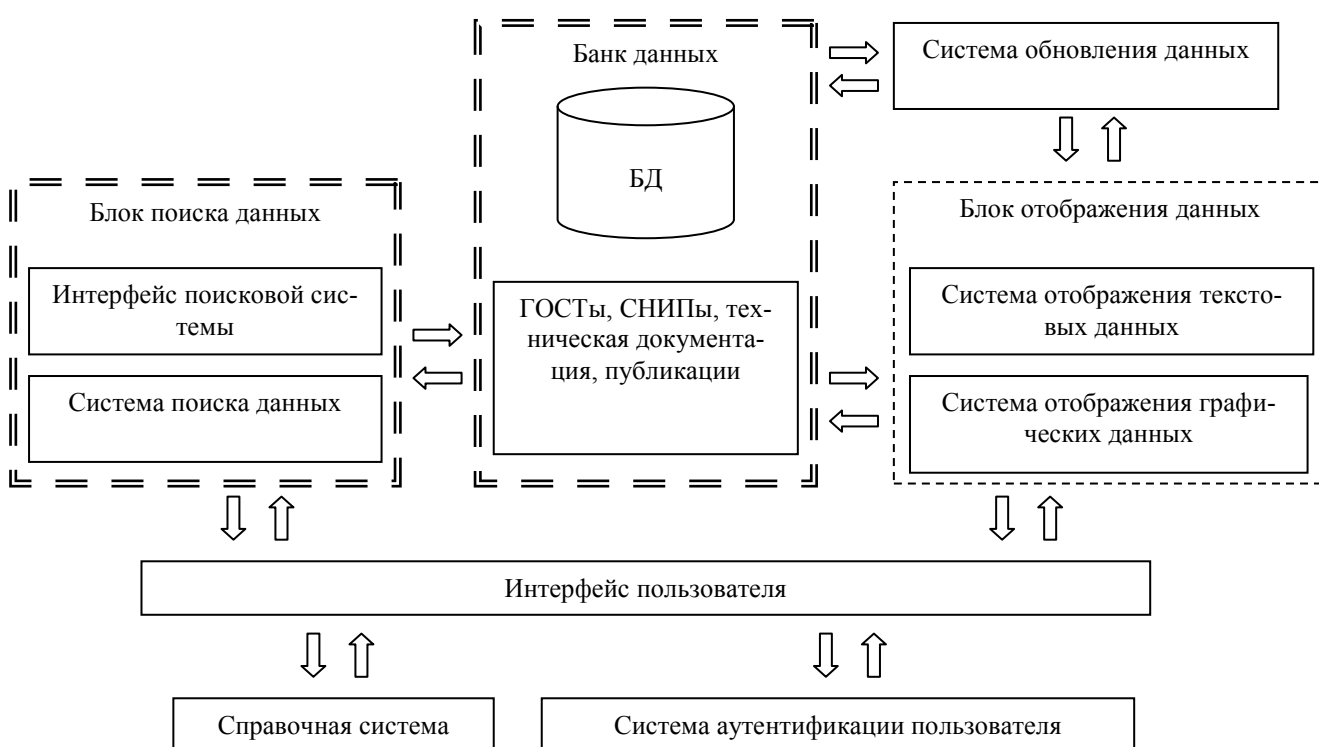


Рис. 2. Концептуальная модель информационной системы

Программная реализация концептуальной модели ИС реализована в среде объектно-ориентированного программирования Borland Delphi 6.0, работающей под управлением операционной системы Windows, при этом использованы стандартный редактор текстовых документов в Windows Notepad.exe, Webbrowser, поддерживающий технологию i-фреймов: Microsoft Internet Explorer.

Взаимосвязь клиентского приложения и базы данных осуществлена с помощью стандартных компонентов среды объектно-ориентированного программирования Borland Delphi Webbrowser.

Интерфейс пользователя содержит средства обеспечения работы конечного пользователя как в локальном так и в удаленном режимах (рис. 3).

На физическом уровне программный продукт состоит сегодня из 26 каталогов, содержащих 681 отдельный файл.

Выводы и направления дальнейших исследований

Проект «Лексикон Бетонведение» является первым шагом в разработке программного обеспечения поддержки принятия решений для широкого круга специалистов строительных специальностей. Реализация задачи позволяет приблизить создание специализированной базы знаний с детерминированными и вероятностными моделями, методами исследований и способами формализации качественных знаний экспертов бетоноведения.

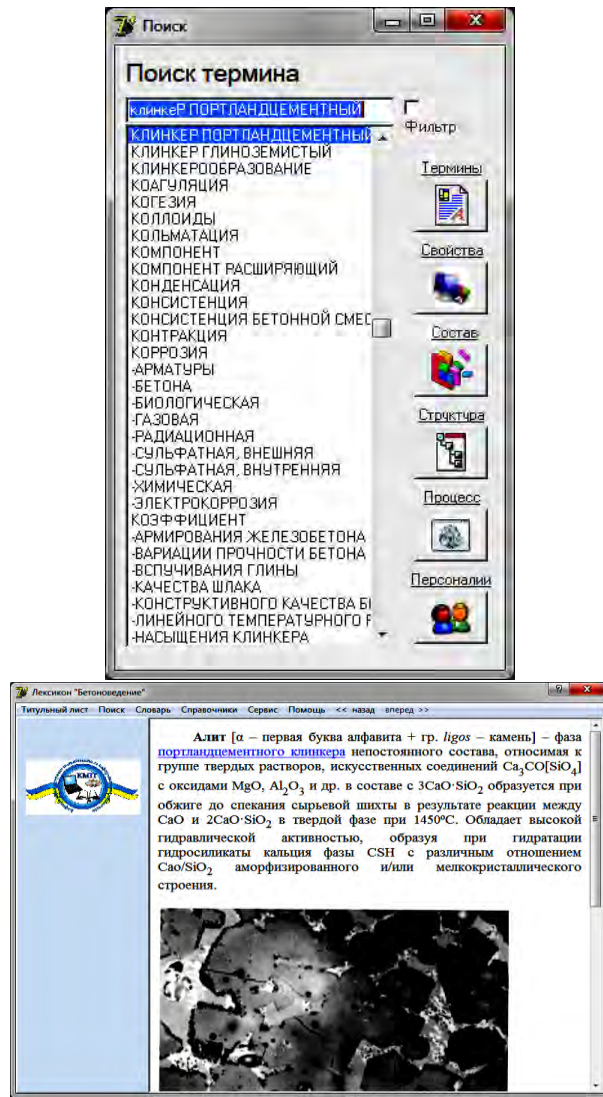


Рис. 3. Вид программы с результатом запроса пользователя

Построенная ИС может служить базой для применения технологий искусственного интеллекта в компьютерных сетях по следующим двум взаимосвязанным аспектам: обмен знаниями в данной предметной области по сети и совместное решение задач с применением распределенных по сети знаний.

Список литературы

1. Буряк М.П. Німецько-українсько-російський словник з технології виробництва будівельних матеріалів та ніздрюватих бетонів. 5000 сл. / М.П. Буряк, В.О. Мартиненко, Г.П. Євсєєва. – Дніпропетровськ: ПДАБА, 2007. – 137 с.
2. Литвинов А.І. Архітектура, Будівництво: Короткий термінологічний словник / А.І. Литвинов, А.І Волощенко. – Вінниця: ВДТУ, 2003. – 120 с.
3. Петров Е.Г. Методи та засоби прийняття рішень у соціально-економічних системах / Е.Г. Петров, М.В. Новожилова, І.В. Гребеннік. – К.: Техніка, 2003. – 240 с.
4. Иванников А.Д. Основообразующие понятия информационных систем второго поколения в научных исследованиях и проектировании / А.Д. Иванников, В.А. Мордовinov. – М.: МИРЭА, 2004. – 73 с.
5. Вендров А.М. Проектирование программного обеспечения экономических информационных систем / А.М. Вендров. – М.: Финансы и статистика, 2000. – 352 с.
6. Романов В.П. Теоретические основы информатики. Информационные структуры и фактографический поиск информации / В.П. Романов. – М.: РЭА им. Г.В. Плеханова, 1996. – 190 с.
7. Оре О. Теория графов / О. Оре. – М.: Мир, 1996. – 240 с.
8. Международный стандарт ISO 12207 «Процессы жизненного цикла программных средств». [Электронный ресурс]. – Режим доступа к ресурсу: <http://www.klubok.net/pageid313.html>.
9. Международный стандарт ISO 9001 «Управление качеством». [Электронный ресурс]. – Режим доступа к ресурсу: <http://www.klubok.net/pageid132.html>.
10. Ушеров-Маршак А.В. Опыт разработки информационной системы «Лексикон «Бетонноедение» / А.В. Ушеров-Маршак, Е.В. Латорец, М.В. Новожилова, И.А. Михеев // Науковий вісник будівництва. – Х.: ХДТУБА, 2009. – Вип. 54. – С. 248-252.

Поступила в редколлегию 2.09.2010

Рецензент: д-р физ.-мат. наук, проф. Н.Д. Сизова, Харьковский государственный технический университет строительства и архитектуры, Харьков.

РОЗРОБКА ІНФОРМАЦІЙНОЇ ТЕРМІНОЛОГІЧНОЇ СИСТЕМИ В ГАЛУЗІ БЕТОНОВЕДЕННЯ

М.В. Новожилова, О.В. Ушеров-Маршак, К.В. Латорець, І.А. Міхєєв

Розглянуті концептуальні засади інформаційної термінологічної системи та наведено опис її практичної реалізації.

Ключові слова: інформаційна система, термінологічний словник.

DEVELOPMENT OF THE INFORMATICS TERMINOLOGY SYSTEM IN AREA OF CONCRETE SCIENCE

M.V. Novozhilova, A.V. Usherov-Marshak, E.V. Latorec, I.A. Mikheev

Conceptual bases of the informatics terminology system are covered and description of its program implementation is provided.

Keywords: informative system, terminology dictionary.