

Использование моделей инженерии знаний для
подготовки специалистов в области
информационных технологий

Т. А. Гаврилова И. А. Лещева
gavrilova@gsom.pu.ru leshcheva@gsom.pu.ru
Высшая Школа Менеджмента,
Санкт-петербургский государственный университет

Д. В. Кудрявцев
dmitry.ku@gmail.com
Санкт-петербургский государственный
политехнический университет

В статье рассматривается проблема применения инструментов визуального структурирования знаний для подготовки программистов. Основной акцент сделан на интеллект-картах, концептуальных картах и анализе типичных ошибок при их составлении.

Ключевые слова: интеллект-карты, концептуальные графы, концептуальные модели, инженерия знаний, mind maps, concept maps.

Введение

Идеи и парадигмы искусственного интеллекта (ИИ) часто служат «колыбелью» для развития новых теорий в смежных науках. Так фреймы Минского породили объектно-ориентированный подход. Сейчас наблюдается перенос идей инженерии знаний — науки, возникшей в рамках разработки интеллектуальных систем [1, 9], — в области обучения и педагогики [4, 5, 7, 18].

Особенно актуальным это оказалось при обучении программистов, системных аналитиков и архитекторов. Именно для них исключительно важны навыки извлечения и структурирования знаний, систематизации и категоризации понятий. Не секрет, что подготовка в области программной инженерии сегодня ориентирована в основном на технологические навыки программирования и страдает от фрагментарности. Системный подход, коммуникативные тренинги, основы когнитивной психологии могли бы существенно обогатить и расширить кругозор и набор компетенций будущих специалистов по информационным технологиям. Кроме того, ведущие преподаватели высшей школы также отмечают [10, 14], что программирование как промышленная дисциплина значительно шире кодирования и в него входит большое количество и управленческих (менеджмент), и психологических аспектов (работа в команде, общение с заказчиком). Общение с заказчиком вообще играет порой ключевую роль в успешности программных проектов — правильные, но ненужные людям системы бесполезны. Но при этом существует со стороны заказчиков трудность в восприятии ИТ-реальности, а со стороны программистов — пренебрежение (скрытое или даже явное), а также объективные трудности в понимании специфики предметной области заказчика.

Именно такая практика привела к ситуации, когда Россия безусловно порождает ярких и талантливых одиночных ИТ-гениев, и не может противостоять азиатской экспансии на ИТ-рынке.

К сожалению, в российской практике высшей школы также почти отсутствует подготовка специалистов в области information science (информационных наук), и единственным местом, где изучают методы систематизации, классификации и кодирования информации, являются библиотечные факультеты.

Кроме того, следует отметить, что описанные далее методы существенно влияют на развитие лингвистических способностей студентов, традиционно недооценивающих гуманитарные составляющие работы аналитика. Таким образом, основной акцент в статье делается на важности использования результатов междисциплинарных исследований для подготовки квалифицированных специалистов информационных технологий, нехватка которых существенно ощущается сейчас в различных секторах российской экономики.

В статье описывается опыт авторов в развитии навыков структурирования и концептуализации знаний при подготовке программистов в Санкт-Петербургском государственном политехническом университете. Методы инженерии знаний частично изучаются в курсах «Основы искусственного интеллекта», «Интеллектуальные системы», «Проектирование экспертных систем» и «Введение в специальность» на факультете технической кибернетики, факультете управления и информационных технологий и в институте международных образовательных программ.

Примерная структура курса может выглядеть, например, так:

- краткая история искусственного интеллекта;
- данные и знания, модели представления знаний;
- методология разработки экспертных систем;
- инженерия знаний (извлечение и структурирование);
- системы управления знаниями.

При этом ядром всех курсов являются концептуальные модели знаний, применимые не только для формализации знаний для машинной обработки, но и для передачи знаний в системе «человек-человек».

В этих курсах концептуальные модели используются двояко:

- для развития компетенций по визуальному структурированию;
- как дидактический инструмент.

Концептуальные модели (например, онтологии) — это инструмент структурирования и представления знаний, позволяющий описывать предметную область в виде понятий и отношений между ними. Концептуальные модели достаточно давно используются людьми для обучения, передачи знаний, проектирования информационных систем, интеграции информации и приложений, обработки естественного языка, а также для автоматизации рассуждений [25, 28]. Они позволяют создавать визуальные модели, удобные как для формирования первичных знаний, так и для восприятия и запоминания.

1. Основы инженерии знаний

Инженерия знаний — это ветвь информатики, изучающая модели и методы извлечения, структурирования и формализации (представления) знаний для их обработки в интеллектуальных и информационных системах.

Традиционно при разработке систем, основанных на знаниях (knowledge-based systems), выделяют три следующих фазы домашинной обработки (рис. 1):

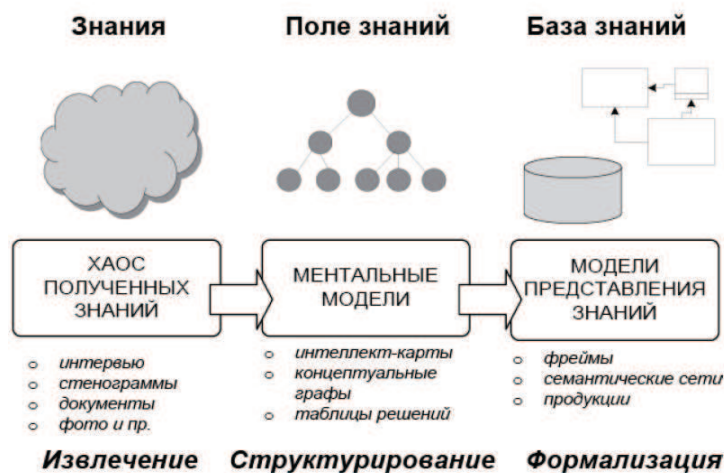


Рис. 1. Основные фазы обработки знаний

1. Извлечение (получение, приобретение) знаний из различных источников. Результат — большое количество разрозненных гетерогенных (разнохарактерных) противоречивых фрагментов и обрывков знаний из Интернета и специальной литературы.
2. Структурирование (концептуализация) разрозненных фрагментов в единую модель. Результат — некоторое наглядное условное, часто слабоформализованное представление, называемое *полем знаний* [2] или *ментальными моделями*. Эти модели включают основные объекты (концепты) предметной

области и связи между ними. Поле знаний обычно представляется в виде некоторой диаграммы, графа, сети, таблицы или схемы. Очевидно, что поле знаний — это один из способов «компрессии» знаний и их наглядного представления, необходимого для лучшего понимания.

3. Формализация поля знаний при помощи специализированных программных языков представления знаний (ЯПЗ). Результат — база знаний, описанная на ЯПЗ.

Эти фазы обработки знаний постоянны и не зависят от предметной области (ПО). Для преподавания можно ограничиться и первыми двумя фазами, работая с полем знаний без компьютерной поддержки, используя его как когнитивно-дидактический инструмент. Это существенное расширение диапазона применимости методологий ИИ.

Поле знаний формируется при структурировании разрозненных фрагментов знаний и представляет некоторое «умственное» упрощение или ментальную модель знаний о предметной области в том виде, в каком она существует в представлении преподавателя, структура которого формируется как базовый «скелет», необходимый для понимания.

Обычно формирование поля знаний начинается с выявления понятийной или концептуальной структуры (ЧТО-знания) предметной области или структурированного словаря. Визуализация всегда считалась мощным когнитивным инструментом понимания (mind tool), т. е. средством, предназначенным для организации и облегчения процесса познания [5].

В качестве инструмента визуального формирования такой структуры удобно использовать наглядность интеллект-карт и концептуальных карт.

2. Модели визуального структурирования знаний

2.1. Интеллект-карты (mind maps)

Методы визуализации идей, процессов, проектов, текстов в форме сетевых графов достаточно традиционны. Считается, что они являются инструментом, позволяющим сделать видимыми понятийные

(или семантические) сети памяти человека. В когнитивной психологии имеется целый набор теорий, основанных на предположении, что человеческая память наиболее адекватно представляется именно сетевой структурой.

Интеллект-карты (mind maps) — один из наиболее привлекательных и простых способов отображения понятийных структур. Хотя автором идеи интеллект-карт (и-карт) считается Тони Бьюзен [43], в России похожий метод успешно использовался в педагогике уже в 1925 году [33].

Бьюзен сформулировал идею еще в 1970-е годы в качестве компактного средства организации конспектов. Позднее он понял, что метод гораздо шире и может использоваться как мощное орудие мышления применительно к педагогике, научной работе, инновациям, бизнес-идеям, политическим дискуссиям, мозговому штурму и пр.

Идея и-карт сильна еще и тем, что имеет серьезный нейрофизиологический базис. Человеческий мозг обладает выраженной сетевой структурой. Естественные нейронные сети мозга включают триллион нейронов, каждый из которых может связываться примерно с 10 000 ближайших соседей. И хотя природа их взаимодействия исследована далеко не полностью, так называемый радиантный характер передачи возбуждения от центра на периферию и обратно доказан.

Идея и-карт заключается в использовании и совмещении функций левого и правого полушарий для достижения целостного и наглядного представления идеи. Фактически, это переход от последовательного (текстового) изложения к сетевому (образному).

Можно предложить первичную классификацию использования таких карт в преподавании:

- организационно-методические и-карты;
- учебно-дидактические и-карты.

Первая группа предназначена для формирования целостного представления о самом курсе и включает структуру курса, дипломной работы, учебной программы. Вторая группа обширнее — к ней относятся карты основных понятий, научных школ, подходов, генеалогии исследователей и др.

Примером учебной и-карты может послужить рис. 2, отражающий представления преподавателя о различных типах знаний для курса «Интеллектуальные системы».

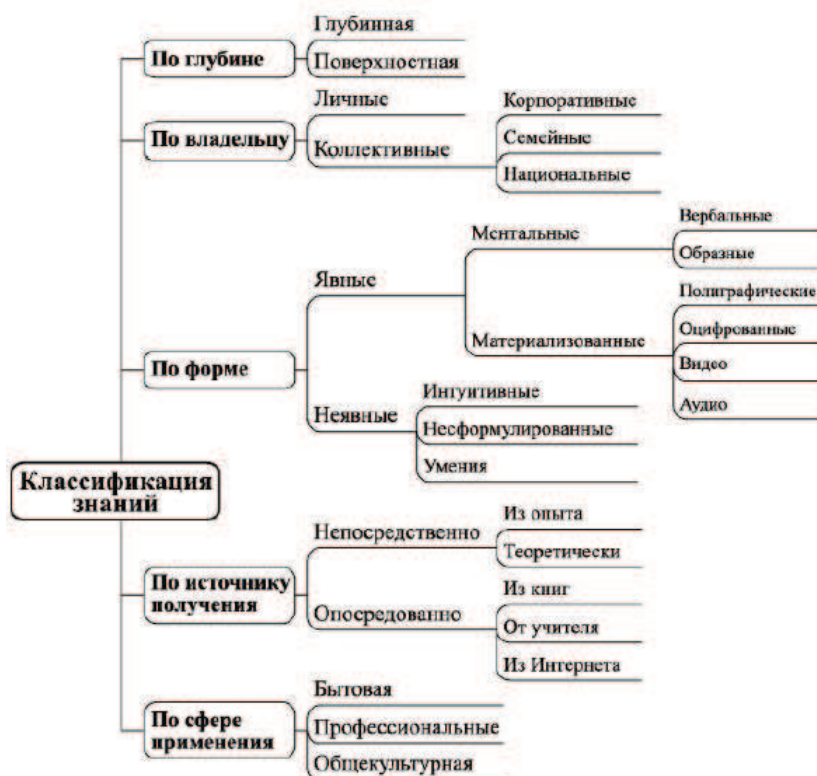


Рис. 2. Пример и-карты по классификации знаний для курса «Интеллектуальные системы»

Любые средства визуализации (от простейших, таких как Visio) и-карт применимы, однако все большую популярность приобретают специализированные методы и инструменты. За последние 5 лет на рынке программного обеспечения появились десятки средств для быстрого и красивого формирования и-карт (например, Comapping, Creately, Mindomo, Map it!, MindMapper Pro, MindGenius Business,

Visual Mind, Mindjet, Freemind, The Brain, Mind Meister, Conception и др.). Особенно интересен и эффективен инструмент Comapping [17] для групповой разработки и-карт.

2.2. Концептуальные графы (concept maps)

Если и-карты показывают связи и древовидную структуру произвольных фрагментов знаний, то концептуальные карты или графы (concept maps) позволяют глубже рассмотреть предметную область и включают отношения между понятиями или концептами. Такие концептуальные карты (к-карты) состоят из узлов и направленных поименованных отношений или связей, соединяющих эти узлы. Связи могут быть различного типа, например, «является», «имеет свойство» и т. п. Концепты и связи имеют универсальный характер для некоторого класса понятий предметной области. Поэтому любая разработка к-карты подразумевает глубокий анализ структурных взаимодействий между отдельными понятиями предметной области.

В процессе создания к-карты преподаватель анализирует структуру отношений предметной области, что помогает ему самому глубже понять ее природу. Зачастую приходится генерировать новые, ранее невербализованные связи.

Впервые к-карты были предложены в психологии в начале 1970-х годов при изучении детского мышления и формирования первых научных понятий. К-карты оказались эффективным инструментом отображения понятийной системы человека [28].

В простейшем случае построение к-карты сводится к ряду шагов:

- определение основного предмета путем задания конкретного фокусирующего вопроса (focus question), определяющего главную тему и границы к-карты;
- выделение концептов — базовых понятий данной предметной области (обычно не более 15–20 понятий);
- построение связей между концептами — определение отношений и взаимодействий базовых понятий;
- упорядочение графа — уточнение, удаление лишних связей, снятие противоречий.

«Хороший» граф обычно получается после 2–3 итераций. Стоит еще раз подчеркнуть, что к-карта — не только цель, но и средство. В процессе построения, т. е. при взаимодействии семантических связей нашей памяти с визуальной информацией, связи перестраиваются, порождая в свою очередь новые знания. Можно строить к-карты на бумаге, на доске, в любом графическом или текстовом редакторе, но удобнее всего использовать специальные программные продукты (например, Causa, Bubble.us, Yed, CoFFEE и др.). В нашей практике хорошо зарекомендовал себя инструмент ИМС SmartTools, свободно распространяемый на сайте <http://smar.ihmc.us>. Фрагмент к-карты для понятия «экспертная система» приведен на рис. 3.



Рис. 3. Пример к-графа

Онтологии [8, 18, 27] позволяют формализовать концептуальные модели с помощью понятного компьютеру языка представления знаний (например, Ontolingua, CycL, RDFS, OWL).

Концептуальные карты легки в использовании и понятны экспертам в различных предметных областях, поэтому их используют в качестве промежуточного этапа при разработке онтологий для информационных систем. Для перевода визуальных концептуальных карт в формализованные онтологии и обратно существуют специальные программные средства, например Concept-map Ontology Environment (COE) [24]. Проблема конвертации концептуальных

карт в онтологии не столько программно-техническая, сколько методическая. При этом на концептуальную карту налагаются дополнительные требования, отсутствующие в случае «традиционной» концептуальной карты.

Разработка таких карт требует развитых аналитических способностей и высоких показателей таких характеристик когнитивного стиля, как способность к обобщению и диапазон когнитивной эквивалентности.

Визуальная разработка концептуальных карт тренирует когнитивные способности будущих программистов и прививает навыки системного мышления. Карты можно активно использовать для контроля знаний студентов, для групповой работы над проектами, для планирования, для диагностики, для проведения мозговых штурмов, при подготовке презентаций и учебных материалов и т. д.

3. Анализ ошибок

Типовые ошибки в разработке интеллект-карт и формализуемых концептуальных карт были получены в результате анализа контрольных работ студентов, обучающихся по специальности программист-математик.

3.1. Интеллект-карты

Ошибки, которые наблюдаются в и-картах студентов 1-го курса, связаны с нарушением одной или нескольких рекомендаций по структурированию, что обычно свидетельствует о тех или иных пробелах в знаниях.

Далее представлены, следуя [26], правила структурирования и-карт.

- *Однородность понятий* — на одном уровне должны находиться понятия одного уровня общности.
- *Однородность отношений* — понятия одного уровня иерархии связываются с родительским концептом одним и тем же типом отношений.
- *Равномерность* — разница в числе уровней в различных ветвях не должна быть больше 2.

- *Обобщение* — максимальное число дочерних концептов не должно превышать число Ингве-Миллера (7 ± 2) [36].
- *Детализация* — глубина ветви также не должна превышать 7 ± 2 .
- *Полнота* — и-карта должна быть полной в том смысле, что на каждом уровне присутствуют все основные концепты.

Рассмотрим подробнее некоторые типы ошибок на примере и-карты понятия «Персональный компьютер» (рис. 4). Самой распространенной ошибкой является нарушение однородности понятий. Например, «Модем» оказался на одном уровне с такими понятиями, как «устройства ввода/вывода», что неверно. В данном случае естественно выделить группу «Устройства связи» и поместить туда «Модем» (предполагая, что он внешний). Более того, «устройства ввода/вывода», хранения и связи можно объединить в понятие «Периферийные устройства» — ошибка отсутствия обобщения. Те же ошибки наблюдаются и в ветви «Приложения».

Примером нарушения однородности отношений является появление «MiniTower» в «Системном блоке», так как материнская плата, блок питания и т. д. находятся в системном блоке, а «MiniTower» является типом корпуса.

Условие полноты не соблюдено: например, среди «операционных систем» не хватает MacOS, которая более широко распространена, чем Linux, а ветка «Приложения» практически не раскрыта.

Несмотря на то, что на приведенной и-карте найдено немало ошибок, ее качество относительно высокое. На практике встречаются более грубые ошибки, которые не являются напрямую нарушением одной из рекомендаций структурирования, а говорят о непонимании значений понятий. Например, «Программное обеспечение» помещают на «Рабочий стол», что говорит о незнании основных принципов работы компьютера.

Представленная и-карта достаточно сбалансирована, хотя даже из нее очевидно, что студент значительно лучше разбирается в устройстве компьютера на аппаратном уровне, чем в программном обеспечении. Таким образом, и-карты могут служить эффективным инструментом для оценки знаний в предметной области.

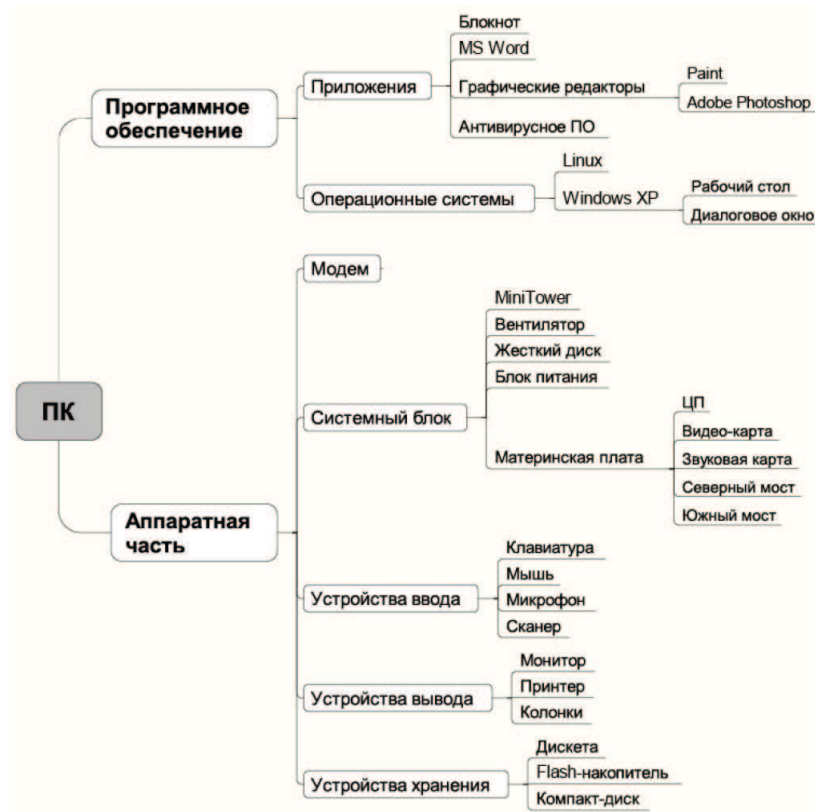


Рис. 4. Пример и-карты «Персональный компьютер», построенной студентом

3.2. Концептуальные графы

В работах студентов 4-х и 5-х курсов было выделено два класса ошибок в практике дизайна к-карт.

3.2.1. Нарушения правил синтаксиса языка представления

онтологий, например, OWL [25, 27]. Такие ошибки приводят к потере информации при переводе к-карты в онтологию и непосред-

ственно связаны с составом элементов выбранного языка. Основными элементами одного из наиболее распространенных языков представления онтологий OWL являются классы, экземпляры, свойства. Сущности могут объединяться в группы, называемые классами. Члены класса называются экземплярами класса. На множестве классов определено родовидовое отношение АКО (A Kind Of). Свойства определяются как пары (домен, диапазон). При этом домен представляет некоторое множество классов, к которым данное свойство применимо, диапазон определяет допустимое множество сущностей — значений свойства. В OWL выделяют две категории свойств: свойства-объекты (или связи) и свойства-значения. Первые связывают между собой индивиды (экземпляры классов). Вторые связывают индивиды со значениями данных. При переводе узлы к-карты становятся классами, экземплярами классов, названиями или значениями свойств классов онтологии, а имена связей между узлами — связями в онтологии.

3.2.2. Нарушения эвристических правил и рекомендаций

по созданию онтологий [2], которые снижают качество получаемых онтологий.

1. Типовые нарушения правил синтаксиса языка представления онтологий:

- не указываются направления связей;
- значения свойства связаны с классами и/или экземплярами, а не только с самим свойством;
- неправильная формулировка связи между свойством и вариантами его значений (вместо связи «имеет значения» используется «включает» или «например»);
- свойство имеет составляющие компоненты;
- отсутствие имен связей между сущностями.

2. Типовые нарушения эвристических правил и рекомендаций по созданию онтологий:

- неправильные направления в родовидовых отношениях АКО;
- путаница в определении отношений «Класс-Подкласс» или «Класс-Экземпляр»;

- замена отношения АКО на part of;
- неправильные имена связей (например, именованье связи вопросом «В чем?», «Куда?» и т. п.).

Заключение

Применение инструментов инженерии знаний в подготовке специалистов в области информационных технологий делает первые шаги, и наш анализ носит предварительный характер. Следует подчеркнуть удобство и плодотворность использования визуальных моделей в преподавании различных дисциплин [4], а также для разработки онтологий, столь популярных сейчас в интеллектуальных приложениях [3, 9] и технологиях Semantic Web [6]. При этом данный подход может хорошо дополнить курсы по обучению средствам проектирования ПО (UML, предметно-ориентированное моделирование и т. д.) [8].

Безусловно, анализ ошибок требует обсуждения, выходящего за рамки данной статьи. Авторы планируют в дальнейшем углубить анализ ошибок, включив в него изучение особенностей когнитивных стилей студентов технических и математических специальностей.

Также важно, что все большее количество преподавателей понимает важность проблемы структурирования и наглядного представления знаний не только для традиционных форм обучения (лекции, семинары, коллоквиумы), но и для дистанционных форм обучения — в интеллектуальных обучающих системах, в системах e-learning и различных электронных учебниках.

Список литературы

- [1] *Башлыков А. А., Еремеев А. П.* Экспертные системы поддержки принятия решений в энергетике / под. ред. А. Ф. Дьяковой. М.: Изд-во МЭИ. 1994. 216 с.
- [2] *Бьюзен Т., Бьюзен Б.* Супермышление. Минск: Изд-во «Попурри». 2003. 320 с.
- [3] *Гаврилова Т. А.* Об одном подходе к онтологическому инжинирингу // *Новости искусственного интеллекта.* 2005. № 3. С. 25–31.
- [4] *Гаврилова Т. А.* Онтологический инжиниринг: от истории к формированию прикладных онтологий // *Когнитивные исследования*

- / под ред. В. Д. Соловьева, Т. В. Черниговской. Вып. 2, Изд-во «Институт психологии РАН». 2008. С. 293–308.
- [5] *Гаврилова Т. А., Гулякина Н. А.* Визуальные методы работы со знаниями: попытка обзора // Искусственный интеллект и принятие решений. 2008. № 1. С. 15–21.
- [6] *Глаголева Н. М.* Техника рассказывания. Л.; М.: Гос. изд. 1925. 520 с.
- [7] *Добров Б. В., Иванов В. В., Лукашевич Н. В., Соловьев В. Д.* Онтологии и тезаурусы: модели, инструменты, приложения: учебное пособие. Изд. БИНОМ. 2011. 172 с.
- [8] *Загорюлько Ю. А.* Построение порталов научных знаний на основе онтологий // Вычислительные технологии. Т. 12, спецвыпуск. 2007. С. 169–177.
- [9] *Кознов Д. В.* Методика обучения программной инженерии на основе карт памяти // Системное программирование. СПб.: Изд. СПбГУ. 2008. Вып. 3. С. 121–140.
- [10] *Кознов Д. В.* О спецификации диаграммных преобразований в графических редакторах // Вестн. С.-Петерб. ун-та. Серия 10: Прикладная математика. Информатика. Процессы управления. 2011. № 3. С. 100–111.
- [11] *Ной Н., МакГиннесс Д.* Разработка онтологий 101: руководство по созданию Вашей первой онтологии. Перевод статьи Natalya F. Noy and Deborah L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. March 2001. <http://www.bibliofond.ru/view.aspx?id=101169>.
- [12] *Осипов Г. С.* Лекции по искусственному интеллекту. М.: КРАСАНД. 2009. 272 с.
- [13] *Терехов А. Н.* Что такое программная инженерия // Программная инженерия. 2010. № 1. С. 40–45.
- [14] *Терехов А. А., Терехов А. Н.* Computing Curricula: Software Engineering и Российское образование. Открытые системы. СУБД. 2006. № 8. С. 61–66.
- [15] *Allemang D., Hendler J.* Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL. Morgan Kaufmann. 2008. 352 с.
- [16] *Eskridge T., Hayes P., Hoffman R., Warren M.* Formalizing the Informal: A Confluence of Concept Mapping and the Semantic Web // A. J. Caas, J. D. Novak (Eds.), Concept Maps: Theory, Methodology, Technology. Proceedings of the

- Second International Conference on Concept Mapping. San Jose, Costa Rica: Universidad de Costa Rica. 2006. V.1. P.8. <http://coe.ihmc.us/groups/coe/wiki/f8c65/attachments/50914/cmc2006-p199.pdf>.
- [17] *Koznov D., Pliskin M.* Computer-Supported Collaborative Learning with Mind-Maps. *Communications in Computer and Information Science*. 2008. Vol.17 CCIS. P. 478–489.
- [18] *Koznov D. V.* Teaching To Write Software Engineering Documents With Focus on Document Design by Means of Mind Maps. *Proceedings of the IASTED International Conference on Computers and Advanced Technology in Education, CATE 2012*. P. 112–118.
- [19] *Gómez-Pérez, A., Fernández-López, M., Corcho, O.* *Ontological Engineering with Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer. 2004. 415 p.
- [20] *Gruber T. R.* A Translation Approach to Portable Ontologies // *Knowledge Acquisition*. V. 2. № 5. 1993. P. 199–220.
- [21] *Jonassen D. H.* Designing Constructivist Learning Environments // *C. Reigeluth (Ed.). Instructional Design Theories and Models: A New Paradigm of Instructional Theory*. Mahwah, NJ: Lawrence Erlbaum Associates. V. II. P. 215–239.
- [22] *Miller G.* The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information // *Psychological Rev.* 1956. V. 63. P. 81–97.
- [23] *Novak J.* Concept maps and diagrams: Two Metacognitive Tools for Science and Mathematics Education // *Instructional Science*. 19. 1990. P. 29–52.
- [24] *Novak J., Cacas A.* The Theory Underlying Concept Maps and How to Construct Them. Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition. 2006. 36 p. <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>
- [25] OWL Web Ontology Language Guide. <http://www.w3.org/TR/owl-guide/>