

Тем не менее, такой способ кластеризации также не является универсальным. Очень часто все документы, хранящиеся в информационно-поисковой системе, имеют принадлежность к некоторой тематической категории, и при кластеризации результатов поиска эта одна из наиболее важных характеристик после релевантности документа. Проблема заключается в том, что данная характеристика не имеет измеримой величины. В некоторых случаях, правда, тематические категории могут иметь и числовые характеристики (например, шифры ББК или УДК), однако эти числовые характеристики не всегда отображают степень близости различных категорий между собой.

Одним из способов определения числовых характеристик тематических категорий можно предложить иерархическую структуризацию таких категорий. Если для каждой категории определить подчинённые ей подкатегории и предположить, что все подкатегории одной категории одинаково близки между собой, можно каждой категории назначить свой весовой коэффициент, исходя из её уровня. Фактически, это задача кластеризации самих категорий. Для решения этой задачи можно использовать экспертные системы (что является предпочтительным для специализированных предметных областей), или кластеризовать все документы, но не по релевантности некоторому запросу, а по семантической схожести полного содержания документов между собой. Последняя задача осложняется, во-первых, трудностью определения степени семантической схожести документов, а, во-вторых, значительными вычислительными и временными затратами.

Заключение. В данной работе рассмотрены некоторые аспекты кластеризации результатов поиска информации в информационно-поисковых системах. Продемонстрированы возможности кластерного анализа по решению различных задач, возникающих в процессе построения информационно-поисковых систем. Рассмотрены также и трудности в адаптации методов кластерного анализа к обработке текстовой информации, не имеющих чётких числовых характеристик.

Список литературы

1. Ландэ, Д. В. Интернетика. Навигация в сложных сетях: модели и алгоритмы / Д. В. Ландэ, А. А. Снарский, И. В. Безсуднов. – Москва: Научная и учебная литература, 2009. – 290 с.

ВИЗУАЛЬНОЕ МОДЕЛИРОВАНИЕ ПРИЛОЖЕНИЙ С ПОМОЩЬЮ UML И ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ CASE ПРИ РАЗРАБОТКЕ ПРИЛОЖЕНИЙ

*О.Г. Казанцева
Витебск, ВГУ имени П.М. Машерова*

Разработка программного продукта – это сложный и многоэтапный процесс, начинающийся с постановки задачи и разработки технического задания. Далее следует разработка архитектуры системы, написание кода, тестирование и частичное внедрение, устранение «багов» (ошибок в программном обеспечении), написание технической документации. Заканчивается процесс разработки передачей программного комплекса в эксплуатацию и дальнейшим сопровождением продукта.

При подготовке студентов специальности «Прикладная математика» в учебном плане предусмотрена дисциплина «Избранные главы информатики». В рамках данной дисциплины студенты изучают современные технологии разработки

программного обеспечения, на практике осваивают современные методы анализа, моделирования, проектирования и разработки программных продуктов.

Цель данной работы – проанализировать влияние знаний, умений и навыков, приобретаемых студентами при изучении UML (Unified Modeling Language) и технологий Caché, на формирование профессиональной компетентности и профессиональной культуры студента.

Материал и методы. В работе проводится анализ содержания учебной дисциплины на предмет соответствия современным тенденциям разработки программного обеспечения.

Результаты и их обсуждение. К моделированию относится разработка программных приложений вплоть до этапа кодирования. Моделирование является неотъемлемой частью крупных программных проектов и полезно для средних и, даже, небольших проектов. Модель играет аналогичную роль в разработке программного обеспечения, как чертежи играют в проектировании зданий. Используя модель, лица, ответственные за успех проекта разработки программного обеспечения, еще до кодирования приложения могут удостовериться в том, что функциональные требования являются полными и правильными, система поддерживает требования масштабируемости, надежности, безопасности, расширяемости и т.д. Если приходится вносить изменения в приложение, которое уже начали кодировать, то эти изменения окажутся более сложными и дорогостоящими. Моделирование является единственным способом для визуализации дизайна и архитектуры приложения и сравнения его с требованиями, прежде чем команда программистов приступит к кодированию [1].

В настоящее время для моделирования и проектирования объектно-ориентированных программных систем широкое применение получили язык визуального моделирования UML и средства автоматизации проектирования программного обеспечения (CASE-средства, Computer-Aided Software Engineering, набор инструментов и методов программной инженерии для проектирования программного обеспечения, который помогает обеспечить высокое качество программ, отсутствие ошибок и простоту в обслуживании программных продуктов).

Почти все современные приложения создаются с использованием объектно-ориентированной технологии, допускающей более быстрый и очевидный способ описания и использования информации. Эта технология позволяет ускорить разработку программных продуктов и увеличить их надежность. При этом разработчики сталкиваются с проблемой организации хранения объектов приложения в базе данных. В случае использования реляционной СУБД существует сложность отображения объектов сложной структуры. Для устранения этой проблемы существует два подхода: «ручное» отображение объектов в таблицы или промежуточный слой программного обеспечения, с неизбежными потерями производительности и усложнением семантики. Избежать вышеуказанных проблем позволяет использование объектных СУБД.

InterSystems Caché – это высокопроизводительная технология баз данных, которая сочетает в себе объектную базу данных, высокоскоростной SQL (для обеспечения реляционного доступа к данным), мощный многомерный доступ к данным при одновременном доступе тремя этими способами к одним и тем же данным. Сервер приложений Caché с высокопроизводительной средой выполнения и уникальной технологией кеширования данных поддерживает расширенные средства разработки и дает возможность легкой интеграции с различными технологиями. Caché включает в себя эффективную среду разработки веб-приложений. Для традиционных приложений без использования браузеров при

создании пользовательского интерфейса применяются популярные технологии: Java, .Net, Delphi, C#, C++ [2].

Тематический план курса «Избранные главы информатики» предполагает знакомство студентов с документированием требований к разрабатываемому программному продукту, с основными понятиями и конструкциями языка UML, с основными возможностями разработки моделей баз данных и архитектуры приложений с использованием CASE-средств, с основами разработки информационных систем на базе СУБД InterSystems Caché на языке Caché Object Script.

В результате изучения дисциплины студенты осваивают методы анализа и проектирования программных систем, методы визуализации и документирования объектно-ориентированных систем и бизнес-процессов, используя язык UML; выполняют кодогенерацию и реинжиниринг с помощью CASE-средств, разрабатывают веб-приложения средствами СУБД Caché.

В дальнейшем студенты при прохождении преддипломной практики закрепляют свои знания при выполнении заданий по разработке информационных систем на предприятии – базе практики.

В 2012 году ВГУ имени П.М.Машерова стал участником Академической программы InterSystems Campus, членами которой являются более 200 ВУЗов по всему миру. В рамках данной программы ВУЗ получает бесплатные лицензии на технологии InterSystems Caché, Ensemble, DeepSee; круглосуточную техническую поддержку и доступ к World-wide Response Center; бесплатное обучение преподавателей вуза; помощь в создании методических материалов; обмен международным опытом с другими вузами-участниками программы; участие в ежегодных конкурсах студенческих работ по технологиям InterSystems; гранты на студенческие проекты и исследования; стажировку талантливых студентов; помощь в трудоустройстве выпускников в крупные IT-компании в России и за рубежом.

Так в рамках данной программы в октябре 2012 года 12 студентов 4 курса специальности прикладная математика прошли обучение по курсу «СУБД Caché: разработка приложений. Базовый курс». В марте 2013 года планируется проведение курса «BI DeepSee», в рамках которого студенты познакомятся с современной платформой для бизнес-аналитики, позволяющей проводить анализ многомерных данных.

Заключение. Полученные в ходе изучения курса «Избранные главы информатики» практические знания и навыки могут быть успешно использованы в дальнейшем при выполнении проектов разработки информационных моделей и программных приложений. Знание изучаемых в курсе основ практической работы с современными CASE-средствами позволит студентам самостоятельно использовать эти средства с целью автоматизации выполнения всех этапов проектирования при разработке информационных систем и программных приложений.

Список литературы

1. Introduction to OMG's Unified Modeling Language (UML) [Электронный ресурс] / Object Management Group, Inc. – US, 1997-2013. – Режим доступа: http://www.omg.org/gettingstarted/what_is_uml.htm. Дата доступа: 1.02.2013.
2. Технологический справочник по Caché [Электронный ресурс] / InterSystems Corporation. – Москва, 1996-2013. – Режим доступа: <http://www.intersystems.ru/cache/technology/techguide/index.html>. Дата доступа: 02.02.2013.