

Министерство образования и науки Республики Казахстан
Карагандинский государственный технический университет

«Утверждаю»
Председатель Ученого совета,
ректор, академик НАН РК
Газалиев А.М.

« ____ » _____ 2015г.

**ПРОГРАММА ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ ДЛЯ СТУДЕНТА
(SYLLABUS)**

Дисциплина OLAPT 3306 «OLAP-технологии»

Модуль ВHD 30 «Модуль Базы и хранилища данных»

Специальность 5В070400 « Вычислительная техника и программное
обеспечение»

Факультет информационных технологий

Кафедра Информационно-вычислительных систем

Предисловие

Программа обучения по дисциплине для студента (syllabus) разработана ст.преподавателем Садановой Б.М.

Обсуждена на заседании кафедры Информационно-вычислительных систем

Протокол № _____ от « ____ » _____ 2015 г.

Зав. кафедрой _____ « ____ » _____ 2015 г.

Одобрена учебно-методическим советом факультета информационных-технологий

Протокол № _____ от « ____ » _____ 2015 г.

Председатель _____ « ____ » _____ 2015 г.

Сведения о преподавателе и контактная информация

Саданова Бакытгуль Маратовна, старший преподаватель

Кафедра ВТ и ПО находится в главном корпусе КарГТУ (Б.Мира, 56), аудитория 301, контактный телефон 56-59-35(254), факс(-) электронный адрес sadanova_b@mail.ru

Трудоемкость дисциплины

Вид обучения	Семестр	Количество кредитов	Вид занятий					Количество часов СРС	Общее количество часов	Форма контроля
			количество контактных часов			количество часов СРСП	всего часов			
			лекции	практические занятия	лабораторные занятия					
Очн.	6	3	15	15	15	45	90	45	135	Экзамен Курсовая работа
Очн. сокр.	4	3	15	15	15	45	90	45	135	Экзамен Курсовая работа

Характеристика дисциплины

Дисциплина «OLAP-технологии» относится к циклу профильных дисциплин (компонент по выбору).

Цель дисциплины

Целью дисциплины «OLAP-технологии» является изучение теоретических основ представления и обработки знаний в информационных системах, а также получение студентами практических навыков проектирования систем, основанных на знаниях.

Задачи дисциплины

Задачи дисциплины следующие: обучение студентов систематизированным представлениям о прикладных системах искусственного интеллекта, моделях знаний и методах обработки знаний; приобретение практических навыков разработки баз знаний для различных моделей, необходимых при проектировании программных систем.

В результате изучения данной дисциплины студенты должны:

– иметь представление о структуре, характеристиках и разновидностях систем, основанных на знаниях;

– знать базовые модели представления знаний в информационных системах и уметь их анализировать; способы представления и обработки неточных и нечетких знаний; архитектуру баз знаний и различные подходы к их организации;

– уметь ставить и решать конкретные задачи по применению методов обработки знаний в прикладных системах, основные алгоритмы и стратегии ло-

гического вывода при реализации компьютерных систем обработки информации и управления;

–приобрести практические навыки по использованию основных компонентов языков инженерии знаний и инструментальных средств построения систем, основанных на знаниях.

Пререквизиты

Для изучения данной дисциплины необходимо усвоение следующих дисциплин (с указанием разделов (тем)):

Дисциплина	Наименование разделов (тем)
1 Алгоритмизация и основы программирования	Структуры данных. Методы и алгоритмы доступа к данным.
2 Технологии программирования	Методика проектирования программной системы
3 Инструментальные средства разработки программ	Принципы организации разработки программ
4 Базы данных	Функции СУБД. Виды моделей данных. Способы организации индексов баз данных.

Постреквизиты

Знания, полученные при изучении дисциплины «OLAP-технологии», используются при освоении следующих дисциплин:

- 1 Проектирование интеллектуальных систем
- 2 Компьютерное моделирование систем
- 3 Современные интернет-технологии

Тематический план дисциплины

Наименование раздела, (темы)	Трудоемкость по видам занятий, ч.				
	лекции	практические	лабораторные	СРСП	СРС
1 Модели представления знаний в информационных системах					
1.1 Определение знаний. Классификация знаний.	1				3
1.2 Логическая модель представления знаний	1			3	3
1.3 Продукционная модель представления знаний и правила их обработки	1			3	3
1.4 Реляционные модели представления знаний	1			3	

1.5 Фреймы и семантические сети.	1			3	
1.6 Представление знаний в виде нечетких высказываний	1			3	
2 Экспертные системы как инструмент автоматизированных обучающих систем					
2.1 Введение в экспертные системы	1			3	
2.2 Искусственный интеллект. Инженерия знаний.	1			3	
3 Онтологии в корпоративных системах					
3.1 Принципы построения онтологических систем	1			3	3
3.2 Онтологический анализ	1			3	3
3.3 Языки описания онтологий. Проект Protégé	1			3	3
4 Интеллектуальный анализ данных					
4.1 Технология исследования данных Data Mining. Задачи Data Mining	1		2	3	3
4.2 Процесс исследования данных	1			3	3
4.3 Технология многомерного оперативного анализа данных OLAP (On-Line Analytical Processing)	1			3	3
4.3.1 Построение витрин данных Создание куба.		3			
4.3.2 Заполнение производственной витрины данных.		4			
4.3.3 Особые элементы OLAP-кубов		4			
4.3.4 Управление измерениями OLAP-кубов		4			
4.3.5 Создание и заполнение хранилищ данных с помощью службы Data Transformation Services OLAP-сервера Microsoft Analysis Services			2		3
4.3.6 Создание многомерных баз данных и описание источников данных			2		3
4.3.7 Применение компонента Pivot-Table List для отображения OLAP-			3		2

данных					
4.3.8 Обзор языка формирования запросов к многомерной базе данных MDX			3		2
4.3.9 Создание Mining Model на основе реляционной таблицы. Использование алгоритма Microsoft Decision Tree			3		3
5 Извлечение знаний из INTERNET (WEB MINING)					
5.1 Web Usage, Content Mining				3	3
5.2 Semantic Web	1			3	2
Итого	15	15	15	45	45

Перечень практических (семинарских) занятий

- 1 Построение витрин данных. Создание куба.
- 2 Заполнение производственной витрины данных.
- 3 Особые элементы OLAP-кубов
- 4 Управление измерениями OLAP-кубов

Перечень лабораторных занятий

- 1 Введение в технологию Data Mining
- 2 Создание и заполнение хранилищ данных с помощью службы Data Transformation Services OLAP-сервера Microsoft Analysis Services
- 3 Создание многомерных баз данных и описание источников данных
- 4 Применение компонента PivotTable List для отображения OLAP-данных
- 5 Обзор языка формирования запросов к многомерной базе данных MDX
- 6 Создание Mining Model на основе реляционной таблицы. Использование алгоритма Microsoft Decision Tree

Тематика курсовой работы

Ознакомление с терминами и понятиями технологии многомерного анализа данных OLAP.

Тематический план самостоятельной работы студента с преподавателем

Наименование темы СРСП	Цель занятия	Форма проведения занятия	Содержание задания	Рекомендуемая литература
1.2 Логическая модель			Построение логиче-	[2,3]

представления знаний	Углубление знаний по данной теме	Выполнение упражнений	ской модели представления знаний	
1.3 Продукционная модель представления знаний и правила их обработки	Углубление знаний по данной теме	Выполнение упражнений	Построение продукционной модели представления знаний	[2,3]
1.4 Реляционные модели представления знаний	Углубление знаний по данной теме	Выполнение упражнений	Построение реляционной модели представления знаний	[2,3]
1.5 Фреймы и семантические сети.	Углубление знаний	Выполнение упражнений	Упражнения № 16,17,18	[4 стр. 158]
1.6 Представление знаний в виде нечетких высказываний	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 19,20	[4 стр.185]
2.1 Введение в экспертные системы	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 5,6	[2 стр. 221]
2.2 Искусственный интеллект. Инженерия знаний.	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 13,15	[6 стр. 230]
3.1 Принципы построения онтологических систем	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 1-3	[5 стр.185]
3.2 Онтологический анализ	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 4-5	[5 стр. 221]
3.3 Языки описания онтологий. Проект Protégé	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 6-10	[5 стр. 230]
4.1 Технология исследования данных Data Mining. Задачи Data Mining	Углубление знаний по данной теме	Выполнение упражнений	Упражнения № 17,18,22,23	[1]
4.2 Процесс исследования данных	Углубление знаний по данной теме	Выполнение упражнений	Упражнения [10 стр. 254, 257,265,269]	[10,23]
4.3 Технология многомерного оперативного анализа данных OLAP (On-Line Analytical Processing)	Углубление знаний по данной теме	Выполнение упражнений	Упражнения 1,2,3,5,9 [1 стр.365]	[2,3,4]
5.1 Web Usage, Content Mining	Углубление знаний по данной теме	Выполнение упражнений	Упражнения [6 стр.549, 559,566, 576]	[5,6,7]
5.2 Semantic Web	Углубление знаний по данной теме	Выполнение упражнений	Упражнения [6 стр. 584,658,668]	[9,11]

Темы контрольных заданий для СРС

1. Понятие инженерии знаний.
2. Обобщенная структура экспертных систем
3. Логические и эвристические методы представления знаний.
4. Теория нечетких множеств
5. Структура правил-продукций
6. Методы логического вывода
7. Семантические сети
8. Фреймы и объекты
9. Онтологический инжиниринг
10. Процесс построения онтологии.
11. Наиболее известные инструменты инженерии онтологий.

Критерии оценки знаний студентов

Экзаменационная оценка по дисциплине определяется как сумма максимальных показателей успеваемости по рубежным контролям (до 60%) и итоговой аттестации (экзамен) (до 40%) и составляет значение до 100% в соответствии с таблицей.

Оценка по буквенной системе	Цифровые эквиваленты буквенной оценки	Процентное содержание усвоенных знаний	Оценка по традиционной системе
A	4,0	95-100	Отлично
A-	3,67	90-94	
B+	3,33	85-89	Хорошо
B	3,0	80-84	
B-	2,67	75-79	
C+	2,33	70-74	Удовлетворительно
C	2,0	65-69	
C-	1,67	60-64	
D+	1,33	55-59	
D	1,0	50-54	
F	0	0-49	Неудовлетворительно

Оценка «А» (отлично) выставляется в том случае, если студент в течение семестра показал отличные знания по всем программным вопросам дисциплины, а также по темам самостоятельной работы, регулярно сдавал рубежные задания, проявлял самостоятельность в изучении теоретических и прикладных вопросов по основной программе изучаемой дисциплины, а также по внепрограммным вопросам.

Оценка «А-» (отлично) предполагает отличное знание основных законов и процессов, понятий, способность к обобщению теоретических вопросов дисциплины, регулярную сдачу рубежных заданий по аудиторной и самостоятельной работе.

Оценка «В+» (хорошо) выставляется в том случае, если студент показал хо-

Защита лаб. работ	1,25	*	*	*	*	*	*	*	*	*	*	*	7,5	
Теорет. модуль	7,5					*						*	15	
Курсовая работа	30												30	
Всего по аттестации		30						30						60
Экзамен														40
Итого														100

Политика и процедуры

При изучении дисциплины «OLAP-технологии» прошу соблюдать следующие правила:

1 Не опаздывать на занятия.

2 Не пропускать занятия без уважительной причины, в случае болезни прошу представить справку, в других случаях – объяснительную записку.

3 В обязанности студента входит посещение всех видов занятий.

4 Согласно календарному графику учебного процесса сдавать все виды контроля.

5 Пропущенные занятия отрабатывать в указанное преподавателем время.

Учебно-методическая обеспеченность дисциплины

Автордың аты-жөні	Оқу-әдістемелік әдебиеттің атауы	Баспа, басылып шығатын күні	Даналар саны	
			Кітап-ханада	кафедрада
1	2	3	4	5
Негізгі әдебиеттер				
1. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.	Методы и модели анализа данных: OLAP и Data Mining	СПб.: БХВ – Петербург, 2004	5	1
2. Гаврилова Т.А., Хорошевский В.Ф.	Базы знаний интеллектуальных систем.	Питер, 2010	2	2
3. Питер Джексон	Введение в экспертные системы (3-е издание).	М.: Диалектика, 2004.	8	--
4. Романов В. П.	Интеллектуальные информационные системы в экономике.	М.: Экзамен, 2003.	10	--
5. Тузовский А.Ф., Чириков С.В., Ямпольский В.З.	Системы управления знаниями (методы и технологии)/ Под общ. ред. В.З. Ямпольского	Томск: НТЛ, 2005.	3	3
6. Башмаков А.И., Башмаков И.А.	Интеллектуальные информационные технологии: Учеб. пособие	М.: Изд-во МГТУ им. Н.Э. Баумана, 2005.	2	3

7.Н.Елманова, А.Федоров	Введение в OLAP-технологии Microsoft	Изд: Диалог- МИФИ, 2008 г. - 272 стр.	10	--
8.Б.Ларсон	Разработка бизнес-аналитики в Microsoft SQL Server 2005	Изд: Питер, 2008 г. - 688 стр	8	--
9.Бергер А. и др.	Microsoft SQL Server 2005 Anal- ysis Services. OLAP и многомер- ный анализ данных.	Изд:БХВ- Петербург, 2007г. - 928 стр.	15	1
Қосымша әдебиеттер				
1.Чубукова И.А	Data Mining	<a href="http://www.in
tuit.ru/">http://www.in tuit.ru/	2	1
2.Гладун А.Я., Рогошина Ю.В	Онтологии в корпоративных системах	“Корпора- тивные си- стемы ” №1,2006	1	1
3. Ландэ Д	Семантический Веб: От идеи – к технологии	Телеком- №6,2005	12	--
4.Яворский В.В.	Проектирование интеллектуаль- ных систем.	Караганда: КарГТУ, 2006.	15	10
5.Комов С.	Корпоративные знания: как ими управлять?	Корпоратив- ные системы. № 3, 2005.	2	
6.Шпак Н.	Управление корпоративными знаниями: это уже важно!!!	«PR в Рос- сии».№ 10, 2004	--	1
7.Колесов А	Интеллектуальные анализ дан- ных и прогнозирование в «1С:Предприятия 8.0»//	PC Magazine/ RE. 6/2006.	2	1
8.О’Лири Д.	Управление корпоративными знаниями//	Открытые системы – 2008. – № 4.	3	1
9.Клещев А.С., Артемьева И.Л.	Математические модели онто- логий предметных областей. Часть 1. Существующие подхо- ды к определению понятия "он- тология":	Научно– техническая информация. Серия.2.– 2008. – № 2.	4	1
10.Клещев А.С., Артемьева И.Л.	Математические модели онто- логий предметных областей. Часть 2. Компоненты модели:	Научно- техническая инф. Сер.2. 2008 № 3.С. 19-28	4	1
11.Смирнов А.В., Пашкин М.П., Шилов Н.Г., Ле- вашова Т.В	Онтологии в системах искус- ственного интеллекта: способы построения и организации. Ч. 1	Новости ис- кусственного интеллекта №1. 2010.	2	1

График выполнения и сдачи заданий по дисциплине

Вид кон-	Цель и содержание	Рекомендуемая	Продолжитель-	Форма	Срок
----------	-------------------	---------------	---------------	-------	------

троля	задания	литература	ность выполне- ния	контроля	сдачи
1	2	3	4	5	6
Посещаемость лекций	Усвоение материала по темам, изложенным в п.п.1-5	[1-9]	15ч	Текущий	На каждой лекции
Сдача лабораторных работ №№1-6	Усвоение материала по темам, изложенным в п.4	[1,7,8,9]	15ч	Текущий	На 2,4,6,9, 12,15 неделях
Рубежный контроль-теоретический модуль	Проверка знаний по темам разделов №№1-5	[1-9, 12,14,15]	0,5ч	Рубежный контроль	На 7,14 неделях
Курсовая работа	Проверка знаний по темам разделов №№1-5	[1-9, 12,14,15]	0,5ч	Итоговый	В период сессии
Экзамен	Проверка знаний по темам разделов №№1-5	[1-9, 12,14,15]	0,5ч	Итоговый	В период сессии

Вопросы для самоконтроля

Заполните пропуски в следующих высказываниях:

1. Раздел интеллектуального анализа данных, прикладной областью которого является Internet, называется
2. . . . подразумевает автоматический поиск и извлечение качественной информации из разнообразных источников Интернета, перегруженных "информационным шумом".
3. . . . , т.е. устойчивость к каким-либо нарушениям исходных предпосылок, означает возможность работы с зашумленными данными и пропущенными значениями в данных.
4. Синонимами термина . . . являются "автоматическая классификация", "обучение без учителя" и "таксономия".
5. . . . – целостная структурная спецификация некоторой предметной области, ее формализованное представление, которое включает словарь (или имена) указателей на термины предметной области и логические выражения, описывающие, как они соотносятся друг с другом.
6. Онтология предметной области есть соглашений; она определяет термины предметной области, задает их толкование, содержит утверждения, которые ограничивают смысл этих терминов, а также дает толкование этих утверждений.
7. Основная цель СУЗ – сделать знания . . . и повторно используемыми на уровне всей корпорации, ведь знание, которое не описано и не тиражировано, в конечном счете становится устаревшим и бесполезным.

8. . . . – это необработанный материал, предоставляемый поставщиками данных и используемый потребителями для формирования информации на основе данных.

9. . . . могут являться числовыми данными либо символьными.

10. . . . – это особым образом организованные и хранимые в электронном виде данные, т.е. данные организованы неким конкретным способом, способным облегчить их поиск и доступ к ним для одного или нескольких приложений.

11. . . . буквально переводится как «данные о данных», информация о другом наборе данных.

12. онтологий – наиболее важное приложение сервера Ontolingua является Web-приложением на основе форм HTML.

13. . . . онтологий ODE (Ontological Design Environment), который взаимодействует с пользователями на концептуальном уровне в отличие от инструментов, подобно OntoSaurus, общающихся на символьном уровне.

14. В центре большинства онтологий находятся описывают понятия предметной области.

15. Метаданные системы ХД содержатся в . . . – структурированной системе хранения и извлечения, реализованной на основе СУБД.

16. . . . – совокупность фактов, закономерностей и эвристических правил, с помощью которых решается поставленная задача.

17. Одно из ключевых требований к поисковым системам – . . . как по объему обрабатываемой информации, так и по числу пользователей.

18. Наличие . . . данных может действительно привести к финансовым потерям и юридической ответственности, если их присутствие не предотвращается или они не обнаруживаются и не очищаются.

19. . . . – резко отличающиеся объекты или наблюдения в наборе данных.

20. После окончания этапа подготовки данных можно переходить к . . . модели.

21. . . . – задача нахождения зависимостей в наборе данных

22. . . . – упорядоченное по некоторому принципу множество объектов, которые имеют сходные классификационные признаки (одно или несколько свойств), выбранных для определения сходства или различия между этими объектами.

23. . . . множество – множество, которое включает данные, использующиеся для обучения (конструирования) модели.

24. . . . множество также содержит входные и выходные значения примеров. Здесь выходные значения используются для проверки работоспособности модели.

25. В основе онтологического . . . лежит описание системы (например, корпорации) в терминах сущностей, отношений между ними и преобразование сущностей, которое выполняется в процессе решения определенной задачи

26. . . . – это интегрированное инструментальное программное средство, которое используется разработчиками систем и экспертами по предметным областям для разработки систем, основанных на знаниях.

27. Ключевым моментом в проектировании онтологии является выбор соответствующего . . . спецификации онтологий.

28. На сегодняшний день основными такими языками представления данных являются Существует также ряд других форматов, однако они предоставляют больше возможностей, потому и обладают статусом рекомендаций W3C.

29. . . . хорошо подходит для описания различных нетекстовых объектов, например, коллекций изображений, видео- или аудиофайлов

30. Основное назначение многомерных баз данных (МБД) – реализация систем, ориентированных на динамический, многомерный анализ исторических и текущих данных, . . . тенденций, моделирование и прогнозирование будущего

31. Системы, поддерживающие . . . модель, предполагают, что в МБД может быть определено несколько гиперкубов с различной размерностью и с различными измерениями в качестве их граней.

32. . . . (в буквальном переводе с английского – «добыча данных») – это направление в информационных технологиях, которое связано с автоматизированным извлечением знаний (неявным образом присутствующих в обрабатываемой информации) и базируется на интеллектуальном анализе данных.

33. . . . – свойство вычислительной системы, которое обеспечивает предсказуемый рост системных характеристик, например, быстроты реакции, общей производительности и пр., при добавлении к ней вычислительных ресурсов.

34. . . . предназначена для разбиения совокупности объектов на однородные группы (кластеры или классы).

35. . . . – это мысленно ограниченная область реальной действительности, подлежащая описанию или моделированию и исследованию

36. . . . – это циклическая управляющая структура, она содержит выбор между альтернативами и следование избранной.

37. . . . – это набор взаимодействующих друг с другом объектов встроенного языка, дающих возможность разработчику использовать его составные части в произвольной комбинации в любом прикладном решении.

38. В основе концепции . . . лежит принцип многомерного представления данных.

39. . . . OLAP-средства представляют собой приложения, осуществляющие вычисление агрегатных данных (сумм, средних величин, максимальных или минимальных значений) и их отображение, при этом сами агрегатные данные содержатся в кэше внутри адресного пространства такого OLAP-средства.

40. Идея . . . схемы реляционной базы данных заключается в том, что имеются таблицы для каждого измерения, а все факты помещаются в одну таблицу, индексируемую множественным ключом, составленным из ключей отдельных измерений

ПРОГРАММА ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ ДЛЯ СТУДЕНТА (SYLLABUS)

по дисциплине «OLAP-технологии»

для студентов специальности 5В070400 – «Вычислительная техника и
программное обеспечение»

институт компьютерных технологий и системотехники

кафедра вычислительной техники и программного обеспечения

Гос.изд.лиц. № 50 от 31.03.2004. Подписано в печать 30.12.08г. Формат
60x90/16 Усл.печ.л. 0,9 Тираж Цена договорная

Издательство Карагандинского государственного технического
университета
100027, Караганда, б.Мира, 56