МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Магнитогорский государственный технический университет им. Г.И. Носова»

УТВЕРЖДАЮ
Директор ИЭиАС
_____ В.Р. Храмшин

13.02.2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

МЕТОДЫ УПРАВЛЕНИЯ ЗНАНИЯМИ

Направление подготовки (специальность)
09.03.01 Информатика и вычислительная техника

Направленность (профиль/специализация) программы
Программное обеспечение средств вычислительной техники и автоматизированных систем

Уровень высшего образования - бакалавриат

Форма обучения
очная

Институт/ факультет	Институт энергетики и автоматизированных систем
Кафедра	Вычислительной техники и программирования
Курс	4
Семестр	7

Магнитогорск
2024 год

Рабочая программа составлена на основе ФГОС ВО - бакалавриат по направлению подготовки 09.03.01 Информатика и вычислительная техника (приказ Минобрнауки России от 19.09.2017 г. № 929)

Рабочая программа рассмотрена и одобрена на заседании кафедры Вычислительной техники и программирования
25.01.2024, протокол № 5

Зав. кафедрой _____ О.С. Логунова

Рабочая программа одобрена методической комиссией ИЭиАС
13.02.2024 г. протокол № 4

Председатель _____ В.Р. Храмшин

Рабочая программа составлена:
ст. преподаватель кафедры ВТиП, _____ М.В.Зарецкий

Рецензент:
Директор НИИ «Промбезопасность» , д-р техн. наук _____ М.Ю. Наркевич

Лист актуализации рабочей программы

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2025 - 2026 учебном году на заседании кафедры Вычислительной техники и программирования

Протокол от __ _____ 20__ г. № __
Зав. кафедрой _____ О.С. Логунова

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2026 - 2027 учебном году на заседании кафедры Вычислительной техники и программирования

Протокол от __ _____ 20__ г. № __
Зав. кафедрой _____ О.С. Логунова

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2027 - 2028 учебном году на заседании кафедры Вычислительной техники и программирования

Протокол от __ _____ 20__ г. № __
Зав. кафедрой _____ О.С. Логунова

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2028 - 2029 учебном году на заседании кафедры Вычислительной техники и программирования

Протокол от __ _____ 20__ г. № __
Зав. кафедрой _____ О.С. Логунова

1 Цели освоения дисциплины (модуля)

Целями освоения дисциплины (модуля) «Методы управления знаниями» являются: ознакомление студентов с базовыми понятиями теории управления знаниями, формирование представлений об алгоритмах создания, поиска, хранения, распределения, актуализации знаний.

Для достижения поставленной цели в курсе «Методы управления знаниями» решаются задачи:

- изучение жизненного цикла знаний;
- изучение методологии управления процессом работы со знаниями;
- изучение онтологических моделей представления знаний;
- изучение языков описания знаний;
- освоение инструментальных сред для работы со знаниями;
- изучение систем управления знаниями.

2 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Методы управления знаниями входит в часть учебного плана формируемую участниками образовательных отношений образовательной программы.

Для изучения дисциплины необходимы знания (умения, владения), сформированные в результате изучения дисциплин/ практик:

Базы и хранилища данных

Базы данных OLTP-систем

Логическое программирование

Методологии и инструментальные средства моделирования и анализа бизнес-процессов

Алгоритмы и теория сложности

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

Технологии Data Mining и Big Data

Подготовка к сдаче и сдача государственного экзамена

Выполнение и защита выпускной квалификационной работы

3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения

В результате освоения дисциплины (модуля) «Методы управления знаниями» обучающийся должен обладать следующими компетенциями:

Код индикатора	Индикатор достижения компетенции
ПК-6 Способность к формализации и алгоритмизации поставленных задач, к написанию программного кода с использованием языков программирования, определения и манипулирования данными и оформлению программного кода в соответствии установленными требованиями	
ПК-6.1	Оценивает качество математической модели при формализации задачи предметной области
ПК-6.2	Оценивает качество разработанных алгоритмов для последующего кодирования
ПК-6.3	Оценивает выбор программных средств для программирования и манипулирования данными в соответствии установленными требованиями

4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 3 зачетных единиц 108 акад. часов, в том числе:

– контактная работа – 55 акад. часов:

– аудиторная – 54 акад. часов;

– внеаудиторная – 1 акад. часов;

– самостоятельная работа – 53 акад. часов;

– в форме практической подготовки – 0 акад. час;

Форма аттестации - зачет

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в акад. часах)			Самостоятельная работа студента	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		Лек.	лаб. зан.	практ. зан.				
1. Основные понятия теории управления знаниями								
1.1 Данные, информация, знания, классификация знаний, свойства знаний	7	2	4		3	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.1
1.2 Управление процессом работы со знаниями. Решающие деревья. Метод ближайших соседей.		2	4		3	1. Подготовка к лабораторным занятиям 2. Выполнение лабораторных работ 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.1, ПК-6.2
1.3 Жизненный цикл знаний		2	4		3	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.1, ПК-6.2
Итого по разделу		6	12		9			

2. Онтологические модели представления знаний и языки описания знаний								
2.1 Онтологии предметных областей, формальная модель онтологии, типы онтологий, метаописания	7	2	4		4	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.1, ПК-6.2
2.2 Дескриптивные логики		2	4		8	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.1, ПК-6.2
2.3 Языки описания онтологий		2	4		8	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.2, ПК-6.3
Итого по разделу		6	12		20			
3. Программные средства работы со знаниями								
3.1 Программные инструменты для работы с онтологиями	7	2	4		8	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.3
3.2 Системы управления знаниями		2	4		8	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.3

3.3 Семантический WEB		2	4		8	1. Подготовка к лабораторному занятию 2. Выполнение лабораторной работы 3. Самостоятельное изучение учебной и научной литературы	1. Беседа - обсуждение 2. Проверка индивидуальных заданий 3. Устный опрос.	ПК-6.3
Итого по разделу		6	12		24			
Итого за семестр		18	36		53		зачёт	
Итого по дисциплине		18	36		53		зачет	

5 Образовательные технологии

1. Традиционные образовательные технологии, ориентированные на организацию образовательного процесса и предполагающую прямую трансляцию знаний от преподавателя к аспиранту.

Формы учебных занятий с использованием традиционных технологий:

Информационная лекция – последовательное изложение материала в дисциплинарной логике, осуществляемое преимущественно вербальными средствами (монолог преподавателя).

Лабораторная работа – организация учебной работы с реальными материальными и информационными объектами, экспериментальная работа с аналоговыми моделями реальных объектов.

2. Технологии проблемного обучения – организация образовательного процесса, которая предполагает постановку проблемных вопросов, создание учебных проблемных ситуаций для стимулирования активной познавательной деятельности аспирантов.

3. Интерактивные технологии – организация образовательного процесса, которая предполагает активное и нелинейное взаимодействие всех участников, достижение на этой основе личностно значимого для них образовательного результата.

Формы учебных занятий с использованием специализированных интерактивных технологий:

Лекция «обратной связи» – лекция–провокация (изложение материала с заранее запланированными ошибками), лекция-беседа, лекция-дискуссия, лекция-конференция.

4. Информационно-коммуникационные образовательные технологии – организация образовательного процесса, основанная на применении программных сред и технических средств работы со знаниями в различных предметных областях.

6 Учебно-методическое обеспечение самостоятельной работы обучающихся

Представлено в приложении 1.

7 Оценочные средства для проведения промежуточной аттестации

Представлены в приложении 2.

8 Учебно-методическое и информационное обеспечение дисциплины (модуля)

а) Основная литература:

1. Управление знаниями. Теория и практика : учебник для бакалавриата и магистратуры / А. И. Уринцов [и др.] ; ответственный редактор А. И. Уринцов. — Москва : Издательство Юрайт, 2022. — 255 с. — (Бакалавр и магистр. Академический курс). — ISBN 978-5-9916-3754-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/508130 (дата обращения: 03.05.2023).

б) Дополнительная литература:

1. Фролов, Ю. В. Управление знаниями : учебник для вузов / Ю. В. Фролов. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2023. — 324 с. — (Высшее образование). — ISBN 978-5-534-05521-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/515648 (дата обращения: 03.05.2023).

2. Паникарова, С. В. Управление знаниями и интеллектуальным капиталом : учебное пособие для вузов / С. В. Паникарова, М. В. Власов. — Москва : Издательство Юрайт, 2022. — 142 с. — (Высшее образование). — ISBN 978-5-534-10125-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/493564 (дата обращения: 03.05.2023).

в) Методические указания:

1. Управление знаниями в организации : учебник и практикум для вузов / А. И. Уринцов [и др.] ; под редакцией А. И. Уринцова. — Москва : Издательство Юрайт, 2023. — 254 с. — (Высшее образование). — ISBN 978-5-9916-9039-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/530666 (дата обращения: 03.05.2023).

г) Программное обеспечение и Интернет-ресурсы:

https://new.znanium.com/read?id=303165 Селетков, С. Н. Управление информацией и знаниями в компании : учебник / С.Н. Селетков, Н.В. Днепровская. — Москва : ИНФРА-М, 2018. — 208 с.

https://new.znanium.com/read?id=289990 Салихова, И. С. Управление качеством интеллектуального капитала самообучающейся организации в экономике знаний : монография / И. С. Салихова. - Москва : Дашков и К, 2015. - 147 с.

Программное обеспечение

Наименование ПО	№ договора	Срок действия лицензии
MS Office 2007 Professional	№ 135 от 17.09.2007	бессрочно
PostgreSQL	свободно распространяемое ПО	бессрочно
MS Visual Studio 2017 Community Edition	свободно распространяемое ПО	бессрочно
Oracle SQL Developer	свободно распространяемое ПО	бессрочно
Oracle SQL Developer Data Modeler	свободно распространяемое ПО	бессрочно

Профессиональные базы данных и информационные справочные системы

Название курса	Ссылка
Федеральное государственное бюджетное учреждение «Федеральный институт промышленной собственности»	URL: http://www1.fips.ru/

9 Материально-техническое обеспечение дисциплины (модуля)

Материально-техническое обеспечение дисциплины включает:

1. Лекционная аудитория ауд. 282. Мультимедийные средства хранения, передачи и представления информации.

2. Компьютерные классы Центра информационных технологий ФГБОУ ВО «МГТУ». Персональные компьютеры, объединенные в локальные сети с выходом в Internet, оснащенные современными программно-методическими комплексами для решения задач в области информатики и вычислительной техники.

3. Аудитории для самостоятельной работы: компьютерные классы; читальные залы библиотеки. Все классы УИТ и АСУ с персональными компьютерами, выходом в Интернет и с доступом в электронную информационно-образовательную среду университета.

4. Аудиторий для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Ауд. 282 и классы УИТ и АСУ.

5. Помещения для самостоятельной работы обучающихся, оснащенных компьютерной техникой с возможностью подключения к сети «Интернет» и наличием доступа в электронную информационно-образовательную среду организации. Классы УИТ и АСУ.

6. Помещения для хранения и профилактического обслуживания учебного оборудования. Центр информационных технологий – ауд. 372

Приложение 1. Учебно-методическое обеспечение самостоятельной работы обучающихся

Лабораторная работа

Данные, информация, знания, классификация знаний, свойства знаний

Для того чтобы начать работать с данными, необходимо сначала загрузить их из файла. В данном задании данные в формате CSV, предназначенном для хранения табличных данных: столбцы разделяются запятой, первая строка содержит имена столбцов. Пример загрузки данных в Pandas:

```
import pandas

data = pandas.read_csv ( 'titanic.csv' , index_col = 'PassengerId')
```

Данные будут загружены в виде DataFrame, с помощью которого можно удобно работать с ними. В данном случае параметр

```
index_col = 'PassengerId'
```

означает, что колонка PassengerId задает нумерацию строк данного датафрейма.

Для того, чтобы посмотреть, что представляют из себя данные, можно воспользоваться несколькими способами: более привычным с точки зрения Python (если индекс указывается только один, то производится выбор строк):

```
data [:10]
```

или же воспользоваться методом датафрейма:

```
data.head( )
```

Один из способов доступа к столбцам датафрейма — использовать квадратные скобки и название столбца:

data[’Pclass’]

Для подсчета некоторых статистик (количества, среднее, максимум, минимум) можно также использовать методы датафрейма:

data[’Pclass’].value_counts( )

Более подробно со списком методов датафрейма можно познакомиться в документации.

Инструкция по выполнению

Загрузите датасет *titanic.csv* (https://www.kaggle.com/c/titanic/data) и, используя описанные выше способы работы с данными, напишите программный код, который находит:

1. Какое количество мужчин и женщин ехало на корабле?

2. Какой части пассажиров удалось выжить? Посчитайте долю выживших пассажиров. Ответ приведите в процентах.

3. Какую долю пассажиры первого класса составляли среди всех пассажиров? Ответ приведите в процентах.

4. Какого возраста были пассажиры? Посчитайте среднее и медиану возраста пассажиров.

5. Коррелируют ли число братьев/сестер с числом родителей/детей? Посчитайте корреляцию Пирсона между признаками SibSp и Parch.

6. Какое самое популярное женское имя на корабле? Извлеките из полного имени пассажира (колонка Name) его личное имя (First Name). Это задание — типичный пример того, с чем сталкивается специалист по анализу данных. Данные очень разнородные и шумные, но из них требуется извлечь необходимую информацию. Попробуйте вручную разобрать несколько значений столбца Name и выработать правило для извлечения имен, а также разделения их на женские и мужские. При необходимости округляйте ответ до двух знаков после запятой.

Лабораторная работа

Управление процессом работы со знаниями. Решающие деревья.

Решающие деревья относятся к классу логических методов. Их основная идея состоит в объединении определенного количества простых решающих правил, благодаря чему итоговый алгоритм является интерпретируемым. Как следует из названия, решающее дерево представляет собой бинарное дерево, в котором каждой вершине сопоставлено некоторое правило вида "j-й признак имеет значение меньше b". В листьях этого дерева записаны числа-предсказания. Чтобы получить ответ, нужно стартовать из корня и делать переходы либо в левое, либо в правое поддерево в зависимости от того, выполняется правило из текущей вершины или нет. Одна из особенностей решающих деревьев заключается в том, что они позволяют получать важности всех

используемых признаков. Важность признака можно оценить на основе того, как сильно улучшился критерий качества благодаря использованию этого признака в вершинах дерева.

Данные

В этом задании мы вновь рассмотрим данные о пассажирах Титаника. Будем решать на них задачу классификации, в которой по различным характеристикам пассажиров требуется предсказать, кто из них выжил после крушения корабля.

Реализация в Scikit-Learn

В библиотеке scikit-learn решающие деревья реализованы в классах sklearn.tree.DecisionTreeClassifier (для классификации) и sklearn.tree.DecisionTreeRegressor (для регрессии). Обучение модели производится с помощью функции fit. Пример использования:

```
import numpy as np

from sklearn.tree import DecisionTreeClassifier

X = np.array([[1, 2], [3, 4], [5, 6]])

y = np.array([0, 1, 0])

clf = DecisionTreeClassifier()

clf.fit(X, y)
```

В этом задании потребуется находить важность признаков. Это можно сделать, имея уже обученный классификатор:

```
importances = clf.feature_importances_
```

Переменная importances будет содержать массив "важностей" признаков. Индекс в этом массиве соответствует индексу признака в данных. Стоит обратить внимание, что данные могут содержать пропуски. Pandas хранит такие значения как nan (not a number). Для того, чтобы проверить, является ли число nan’ом, можно воспользоваться функцией np.isnan. Пример использования:

```
np.isnan(X)
```

Инструкция по выполнению

1. Загрузите выборку из файла titanic.csv с помощью пакета Pandas.

2. Оставьте в выборке четыре признака: класс пассажира (Pclass), цену билета (Fare), возраст пассажира (Age) и его пол (Sex).

3. Обратите внимание, что признак Sex имеет строковые значения.

4. Выделите целевую переменную – она записана в столбце Survived.

5. В данных есть пропущенные значения – например, для некоторых пассажиров неизвестен их возраст. Такие записи при чтении их в pandas принимают значение nan. Найдите все объекты, у которых есть пропущенные признаки, и удалите их из выборки.

6. Обучите решающее дерево с параметром random_state=241 и остальными параметрами по умолчанию.

7. Вычислите важности признаков и найдите два признака с наибольшей важностью.

Лабораторная работа

Управление процессом работы со знаниями. Метод ближайших соседей.

Метод *k* ближайших соседей реализован в классе sklearn.neighbors.KNeighborsClassifier. Основным параметром является n_neighbors, который задает число соседей для построения прогноза. Вам понадобится производить кроссвалидацию по блокам. Кроссвалидация заключается в разделении выборки на *m* непересекающихся блоков примерно одинакового размера, после чего выполняется *m* шагов. На *i*-м шаге *i*-й блок выступает в качестве тестовой выборки, объединение всех остальных блоков – в качестве обучающей выборки. Соответственно, на каждом шаге алгоритм обучается на некоторой обучающей выборке, после чего вычисляется его качество на тестовой выборке. После выполнение *m* шагов мы получаем *m* показателей качества, усреднение которых и дает оценку кроссвалидации (подробнее в документации scikit-learn). В задании необходимо будет произвести кроссвалидацию по блокам. Это делается в два этапа:

1. Создается генератор разбиений sklearn.cross_validation.KFold, который задает набор разбиений на обучение и валидацию. Число блоков в кроссвалидации определяется параметром n_folds. Порядок следования объектов в выборке может быть неслучайным, это может привести к смещенности кроссвалидационной оценки. Чтобы устранить такой эффект, объекты выборки случайно перемешивают перед разбиением на блоки. Для перемешивания достаточно передать генератору KFold параметр shuffle=True.

2. Вычислить ошибку на всех разбиениях можно при помощи функции sklearn.cross_validation.cross_val_score. В качестве параметра estimators передается классификатор, в качестве параметра cv – генератор разбиений с предыдущего шага. С помощью параметра scoring можно задавать меру качества, по умолчанию в задачах классификации используется доля верных ответов (accuracy). Результатом является массив, значения которого нужно усреднить. Приведение признаков к одному масштабу можно делать с помощью функции sklearn.preprocessing.scale, которой на вход необходимо подать матрицу признаков и получить масштабированную матрицу, в которой каждый столбец имеет нулевое среднее значение и единичное стандартное отклонение.

Инструкция по выполнению

В этом задании нужно подобрать оптимальное значение *k* для алгоритма kNN. Для выполнения задания предлагается набор данных Wine, где требуется предсказать сорт винограда, из которого изготовлено вино, используя результаты химических анализов. Выполните следующие шаги:

1. Загрузите выборку Wine по адресу https://archive.ics.uci.edu/ml/datasets/Wine (файл: wine.data).

2. Извлеките из данных признаки и классы. Класс записан в первом столбце (три варианта), признаки – в столбцах со второго по последний.

3. Оценку качества необходимо провести методом кроссвалидации по 5 блокам (5-fold). Создайте генератор разбиений, который перемешивает выборку перед формированием блоков (shuffle=True). Для воспроизводимости результата, создавайте генератор KFold с фиксированным параметром random_state=42. В качестве меры качества используйте долю верных ответов (accuracy).

4. Найдите точность классификации на кроссвалидации для метода k ближайших соседей (sklearn.neighbors.KNeighborsClassifier), при k от 1 до 50. При каком k получилось оптимальное качество? Чему оно равно (число в интервале от 0 до 1)?

5. Произведите масштабирование признаков с помощью функции sklearn.preprocessing.scale. Снова найдите оптимальное k на кроссвалидации.

6. Какое значение k получилось оптимальным после приведения признаков к одному масштабу? Как изменилось значение качества?

Лабораторная работа

Онтологии предметных областей, формальная модель онтологии, типы онтологий, метаописания

Описать предметную область с помощью онтологии по заданным темам:

Вариант 1

Предметная область ИС: Библиотека
Минимальный список характеристик:

Автор книги, название, год издания, цена, является ли новым изданием, краткая аннотация; номер читательского билета, ФИО, адрес и телефон читателя, дата выдачи книги читателю и дата сдачи книги читателем.

Вариант 2

Предметная область ИС: Университет
Минимальный список характеристик:

Номер, ФИО, адрес и должность преподавателя; код, название, количество часов, тип контроля и раздел предмета; код, название, номер заведующего кафедрой; номер аудитории, где преподаватель читает свой предмет.

Вариант 3

Предметная область ИС: Оптовая база
Минимальный список характеристик:

Код товара, название товара, количество на складе, стоимость единицы товара, примечания - описание товара; номер и ФИО поставщика товара, срок поставки и количество товаров в поставке.

Вариант 4

Предметная область ИС: Производство
Минимальный список характеристик:

Код изделия, название изделия, является ли типовым, примечание - для каких целей предназначено; код, название, адрес и телефон предприятий, выпускающих изделия; год выпуска и объем выпуска данного изделия предприятием.

Вариант 5

Предметная область ИС: Сеть магазинов
Минимальный список характеристик:

Номер, ФИО, адрес, телефон и капитал владельцев магазинов; номер, название, адрес и телефон магазина; номер, ФИО, адрес, телефон поставщика, а также стоимость поставки данного поставщика в данный магазин.

Вариант 6

Предметная область ИС: Авторемонтные мастерские
Минимальный список характеристик:

Номер водительских прав, ФИО, адрес и телефон владельца автомобиля; номер, ФИО, адрес, телефон и квалификация механика; номер, марка, мощность и цвет автомобиля; номер, название, адрес и телефон ремонтной мастерской.

Вариант 7

Предметная область ИС: Сессия
Минимальный список характеристик:

Наименование специальности, код группы, ФИО, дата рождения, домашний адрес, телефон слушателя, примечания - автобиография слушателя; код, название, количество часов и вид контроля предметов, код сессии и оценки каждого слушателя каждому предмету в каждую сессию.

Вариант 8

Предметная область ИС: Договорная деятельность организации
Минимальный список характеристик:

Шифр договора, наименование организации, сроки выполнения, сумма договора, примечания вид договора номер, ФИО, адрес, телефон, должность, оклад сотрудников, сроки работы данного сотрудника по данному договору.

Вариант 9

Предметная область ИС: Поликлиника
Минимальный список характеристик:

Номер, фамилия, имя, отчество, дата рождения пациента; ФИО, должность и специализация лечащего врача, диагноз, поставленный данным врачом данному пациенту, необходимо ли амбулаторное лечение, срок потери трудоспособности, состоит ли на диспансерном учете, примечание

Вариант 10

Предметная область ИС: Телефонная станция
Минимальный список характеристик:

Номер абонента, фамилия абонента, адрес, дата установки, наличие блокиратора, задолженность, примечание Код АТС, код района, количество номеров.

Вариант 11

Предметная область ИС: Спорт
Минимальный список характеристик:

Фамилия спортсмена, дата рождения, вид спорта, команда, спортивный разряд Название соревнования, дата проведения, место проведения, показанный результат Название вида спорта, мировой рекорд и его дата, единица измерения.

Вариант 12

Предметная область ИС: Сельскохозяйственные поставки
Минимальный список характеристик:

Наименование с/х предприятия, дата регистрации, вид собственности, число работников Название продукции, единица измерения, закупочная цена Название поставщика, дата поставки, объем, себестоимость поставщика

Вариант 13

Предметная область ИС: Городской транспорт
Минимальный список характеристик:

Вид транспорта, средняя скорость движения, стоимость проезда, количество машин в парке номер маршрута, количество пассажиров в день, количество остановок в пути, количество машин на маршруте начальный пункт, конечный пункт, расстояние.

Вариант 14

Предметная область ИС: География
Минимальный список характеристик:

Название страны, столица, площадь территории, является ли страна развитой в экономическом отношении, количество населения национальность, язык, общая численность

Вариант 15

Предметная область ИС: Домоуправление
Минимальный список характеристик:

Номер подъезда, номер квартиры, общая площадь, полезная площадь, количество комнат, фамилия квартиросъемщика, дата прописки, количество членов семьи, количество детей в семье, есть ли задолженность по квартплате, примечание

Вариант 16

Предметная область ИС: Аэропорт
Минимальный список характеристик:

Номер рейса, пункт назначения, дата рейса, тип самолета, время вылета, время в пути, является ли маршрут международным, сведения о пассажире, примечание

Вариант 17

Предметная область ИС: Рынок ПК
Минимальный список характеристик:

Фирма-изготовитель, тип процессора, тактовая частота, объем RAM, объем жесткого диска, дата выпуска, Сведения о фирмах - реализаторах: Наименование, адрес, телефон, примечание

Вариант 18

Предметная область ИС: Деканат.
Минимальный список характеристик:

Фамилия и инициалы студента, дата рождения, группа, сумма стипендии, год зачисления факультет, специальность, число мест

Вариант 19

Предметная область ИС: Микросхемы памяти
Минимальный список характеристик:

Обозначение, разрядность, емкость, дата начала выпуска, время доступа, является ли широко используемой, стоимость, примечание Дата продажи, количество, цена

Вариант 20

Предметная область ИС: Шахматы
Минимальный список характеристик:

Фамилия спортсмена, страна, рейтинг Город, страна, дата проведения турнира, квалификационный уровень Стартовый номер, место участников.

Вариант 21

Предметная область ИС: Судоходство
Минимальный список характеристик:

Название корабля, водоизмещение, порт прописки, капитан Номер причала, дата прибытия, дата убытия, цель посещения Порт, страна, категория.

Вариант 22

Предметная область ИС: Автотранспортное предприятие

Минимальный список характеристик:

Марка автомобиля, номер, пробег, грузоподъемность Номер рейса, пункт назначения, расстояние, дата выезда, дата возвращения ФИО водителя, его табельный номер и категория.

Вариант 23

Предметная область ИС: Научные конференции
Минимальный список характеристик:

ФИО ученого, ученая степень, организация, страна Тема доклада, тип участия, наличие публикаций, Название конференции, место проведения, дата

Вариант 24

Предметная область ИС: Программные продукты
Минимальный список характеристик:

Название, класс продукта, фирма, версия, дата выпуска, стоимость одной лицензии Дата инсталляции, дата деинсталляции, стоимость инсталляции, количество Пользователь, регион, сфера применения.

Вариант 25

Предметная область ИС: Операционная система
Минимальный список характеристик:

Идентификатор процесса, идентификатор владельца, приоритет, время выполнения Название ресурса, запрошенное количество, выделенное количество

Вариант 26

Предметная область ИС: Автотранспортное предприятие
Минимальный список характеристик:

номерной знак автомобиля, марка автомобиля, его техническое состояние, местонахождение автомобиля, средняя скорость, грузоподъемность, расход топлива, табельный номер водителя, фамилия, имя, отчество, дата рождения, стаж работы, оклад, дата выезда, дата прибытия, место назначения, расстояние, расход горючего, масса груза

Вариант 27

Предметная область ИС: Каталог видео клипов
Минимальный список характеристик:

Код видеоленты, дата записи, длительность, тема, стоимость, Сведения об актере: Фамилия, имя отчество, амплуа.

Вариант 28

Предметная область ИС: Заказы
Минимальный список характеристик:

Фамилия, имя, отчество клиента, номер счета, адрес, телефон, номер заказа, дата исполнения, стоимость заказа, название товара, его цена и количество

Вариант 29

Предметная область ИС: Рецепты
Минимальный список характеристик:

Наименование блюда, описание рецепта, категория, вес порции перечень ингредиентов, их количество, очередность добавления, предварительная обработка, количество порций название продукта, калорийность, единица измерения, цена за единицу.

Вариант 30

Предметная область ИС: Отдел кадров
Минимальный список характеристик:

Фамилия, имя, отчество, домашний адрес, телефон, дата рождения, образование, табельный номер, должность, дата зачисления, оклад наименование подразделения, количество штатных единиц, фонд заработной платы за месяц и за год.

Лабораторная работа

Дескриптивные логики

Пример терминологии TBox (описаний онтологии) семейных отношений показан на рис. 1, а описание конкретных экземпляров предметной области (ABox) показан на рис. 2.

$$
\begin{aligned}
\text{Женщина} &\equiv \text{Человек} \sqcap \text{женскийРод} \\
\text{Мужчина} &\equiv \text{Человек} \sqcap \neg\text{женскийРод} \\
\text{Мать} &\equiv \text{Женщина} \sqcap \exists\ \text{иметьДетей.Человек} \\
\text{Отец} &\equiv \text{Мужчина} \sqcap \exists\ \text{иметьДетей.Человек} \\
\text{Родитель} &\equiv \text{Мать} \sqcup \text{Отец} \\
\text{Бабушка} &\equiv \text{Мать} \sqcap \exists\ \text{иметьДетей.Родитель} \\
\text{МногодетнаяМать} &\equiv \text{Мать} \sqcap \geq 3\text{иметьДетей} \\
\text{МатьБезДочери} &\equiv \text{Мать} \sqcap \forall\ \text{иметьДетей.}\neg\text{Женщина} \\
\text{Жена} &\equiv \text{Женщина} \sqcap \exists\ \text{иметьМужа.Мужчина}
\end{aligned}
$$

Рисунок 1 - Терминология понятия семейных отношений TBox

МатьБезДочери (МАРИЯ)
иметьДетей (МАРИЯ, ПЕТР)
иметьДетей (МАРИЯ, СЕМЕН)

Отец (ПЕТР)
иметьДетей (ПЕТР, ИВАН)

Рисунок 2 - Описание мира ABox

Описать предметную область по заданным темам с помощью терминологий TBox и ABox.

Лабораторная работа

Языки описания онтологий

Описать предметную область по заданным темам с помощью языка описания онтологий.

Ниже приводится пример двух способов представления RDF графов: в форме XML-документа (часто более удобной для автоматической обработки) и в форме последовательностей троек - так называемый N Triple или N3 синтаксис (удобный для восприятия человеком).

XML-синтаксис

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:exterms="http://www.example.org/terms/">

<rdf:Description rdf:about="http://www.example.org/index.html">
<exterms:creation-date>August 16, 1999</exterms:creation-date>
</rdf:Description>

<rdf:Description rdf:about="http://www.example.org/index.html">
<dc:language>en</dc:language>
</rdf:Description>

</rdf:RDF>
```

N3-синтаксис *(удобный для чтения человеком и расширяющий исходную модель данных RDF)*

```
<ex:index.html> <dc:creator> exstaff:85740 .
<ex:index.html> <exterms:creation-date> "August 16, 1999" .
<ex:index.html> <dc:language> "en".
```

Лабораторная работа

Программные инструменты для работы с онтологиями

Изучить возможности инструменты Web 2.0.

Пользователи интернета сами становятся создателями информации - создаваемый в сети контент, его описание и структурирование, оценка и ранжирование становятся результатом деятельность рядовых пользователей. Технологически вовлечение пользователей обеспечивается засчет интерактивность веб-страниц и простоты их интерфейсов – для размещения информации в интернете пользователю не нужно никаких специальных знаний.

Список инструментов Веб 2.0:

1. интернет почта (mail.ru, yandex.ru, gmail.com),
2. средства обмена сообщениями (ICQ, google-чат)

3. блоги (например, Живой Журнал, корпоративные блоги Яндекса, Google-а) – он-лайн журналы и дневники, позволяющие обмениваться мнениями, размещать новости. Обновления блогов могут рассылаться с помощью RSS-потоков – см. далее,
4. социальные сети (например, vkontakte.ru, moikrug), позволяющие узнать о знаниях, опыте, интересах других людей и представить свои,
5. средства интеграции и чтения новостных RSS-потоков (Яндекс Лента, Google Reader). **RSS** (Really Simple Syndication) - средство подписки на обновления новостей сайтов, записей блогов, результатов поиска по определенному запросу и на любые другие обновляемые материалы – фотографии, видео, веб-закладки и т.д.
6. Peer-to-peer сети (иногда P2P-сети) – средство эффективного обмена файлами (музыкой, видео, текстами) либо через интернет, либо между ограниченным числом пользователей. Данное средство открывает доступ к содержимому персональных компьютеров пользователей и обеспечивает высокую скорость передачи информации, благодаря специальной технологии разделения файла на отдельные фрагменты.
7. веб-закладки (del.icio.us, БобрДобр), позволяющие пользователям не только сохранять ссылки на интересные для них ресурсы, но и сопровождать их ключевыми словами, обмениваться этими ссылками, анализировать популярность ресурсов (количество ссылок) и тенденции, находить других пользователей со схожими интересами.
8. средства для размещения фотографий (photosight.ru, flickr.com), презентаций (slideshare) и видео (youtube), с богатыми возможностями описания материалов (теги, категории), поиска, обсуждения.
9. wiki-средства (pbwiki.com, wikispaces.com) – средства создания и функционирования веб-сайтов по типу Википедии (см. ru.wikipedia.org). Страницы вики-сайта совместно редактируются сотрудниками организации и выступают альтернативой письмам и вложениям электронной почты.
10. программы для коллективного редактирования документов в онлайне (Google Docs),
11. гибридные приложения (mashups), объединяющие несколько инструментов Web 2.0 между собой или с другими приложениями (например, интеграция географических карт от Google, Yahoo со средствами размещения фотографий - Panoramio, flickr позволяет пользователям привязывать фотографии к местности)
12. средства планирования – календари (Google Календарь) и т.п.

К Web 2.0 можно также отнести «ПО как сервис» (Software as a Service, SaaS). Такое ПО работает через интернет, а компания ежемесячно оплачивает услуги сервис-провайдера. Это дешево и удобно. Не нужно покупать дорогие лицензии, устанавливать и настраивать сложный софт.

Более полумиллиона западных компаний успешно используют онлайновую платформу AppExchange от Salesforce.com, что позволяет подписаться на любое из нескольких сотен бизнес-приложений, доступных для подписки (CRM-системы, бухгалтерские программы и др.). Инструменты Web 2.0 подходят для поддержки совместной работы в системах управления знаниями: они просты, эффективны и уже знакомы некоторым пользователям.

Лабораторная работа

Программные средства работы со знаниями

Изучить возможности отечественных продуктов, позиционирующих себя как программные средства работы со знаниями:

1) Система "Айкумена Аналитик" компании IQMen (http://www.iqmen.ru/our_proposal/) обеспечивает информационно-аналитическую поддержку принятия решений с помощью таких современных технологий как:

1. Автоматический сбор информации из разнородных источников: электронные СМИ, аналитические агентства, специализированные базы данных.
2. Автоматическое составление и рассылка индивидуальных тематических отчетов экспертам и руководству.
3. Средства графического анализа событий по любой тематике, оперативное выявление в новостном потоке горячих тем.
4. Мгновенная подготовка досье по любому вопросу.
5. Выборка наиболее важной фактической информации - дат, финансовых данных, персон и организаций.
6. Автоматическое выявление физических и юридических лиц, имеющих отношение к любому вопросу и анализ взаимосвязей между ними.

2) Система «Галактика ZOOM» — инструмент для создания хранилища текстовой информации, который обладает возможностями для проведения поиска и аналитических исследований.

1. Поиск информации в больших информационных массивах.
2. Выявление значимых слов и словосочетаний документа, отражающих содержание текста.
3. Сравнение документов - обнаружение сходства, различия и аномалий изучаемых объектов.
4. Динамика во времени - отслеживание изменений во времени слов или словосочетаний характеризующих изучаемый объект.

3) Система DocsVision также предназначена для задач управления документами и процессами (http://www.docsvision.com/). Система DocsVision "Делопроизводство" для создает архивы документов, автоматизирует делопроизводство в организации.

4) Система ЕВФРАТ-Документооборот (http://www.evfrat.ru) - комплексное решение по организации электронного документооборота на предприятии. Внедрение системы обеспечивает:

1. систематизацию учета и хранения документов;
2. оперативный доступ к документам и отчетной информации;
3. управление процессами движения и обработки документов.

5) Яndex.Server — программа для поиска информации в интернете или в локальной сети с учетом морфологии русского языка.

Лабораторная работа

Семантический WEB

Описать предметную область по заданным темам с помощью Semantic Web.

Семантический веб (Semantic Web) привносит смысловую структуру в содержание веб-страниц. Семантический веб — это не какая-то отдельная сеть, а расширение уже существующей, такое, что в ней информация снабжена точно определённым смыслом, позволяющим человеку и машине успешно взаимодействовать.

Сценарий работы таких программных средств представлен в основополагающей статье идеологов семантического веба – Тима Бернерса-Ли и Джеймса Хендлера, опубликованной в 2001 году в журнале Scientific American [Berners-Lee, Hendler, 2001]:

«Пит поднял трубку, звонила его сестра Люси из кабинета врача: «Моей маме нужно попасть на приём к врачу, а потом ей требуется пройти несколько сеансов физиотерапии. Примерно два раза в неделю. Я сейчас поручу своему агенту записать нас на приём к врачу». Пит сразу же согласился подвезти её маму на своей машине. В кабинете врача Люси дала указания своему электронному агенту Семантической Сети через свой портативный веб-браузер. Агент сразу же получил информацию о *назначенном маме лечении* от агента её врача, просмотрел несколько перечней *медицинских учреждений*, оказывающих подобные услуги, проверил, какие из них *подходят* [in-plan] по её страховому полису и находятся в *окрестности 20 миль* от её *дома* и при этом имеют *рейтинг — «отличный»* либо *«очень хороший»* согласно достоверным рейтинговым источникам. Далее агент попытался сопоставить имеющиеся *часы приёма* врачей (предоставленные агентами соответствующих мед. учреждений на их веб-сайтах) с расписаниями Пита и Люси. (Слова, выделенные курсивом, указывают на термины, семантика, или смысл, которых была предоставлена агентам посредством Семантической Сети.) Уже через несколько минут агент представил им план. Питу он не понравился — Университетская Клиника расположена на другой стороне города, и ему пришлось бы возвращаться обратно как раз в час пик. Он попросил своего агента повторить поиск с более строгими предпочтениями относительно *места* и *времени*. Тем временем агент Люси, *полностью доверяя* агенту Пита в рамках данной задачи, автоматически помогал, предоставляя тому права доступа и ссылки на уже добытую им информацию. Практически мгновенно был предложен новый план: найденная клиника была уже гораздо ближе, и часы приёма более ранние».

Для того чтобы Семантический Веб мог функционировать, компьютеры должны иметь доступ к структурированным хранилищам информации и множествам правил вывода, которые могли бы использоваться для проведения автоматических рассуждений. Однако два хранилища информации или базы данных могут использовать различные идентификаторы для обозначения одного и того же понятия, такого, например, как *почтовый индекс*. И программа, желающая сравнить или как-то скомбинировать информацию из этих баз данных, должна знать, что два конкретных термина используются ими для обозначения одного и того же. В идеале, у программы должен быть способ распознавать подобные термины с одинаковым смыслом, с какими бы базами данных ей не пришлось столкнуться в процессе своей работы.

Решение этой проблемы даётся третьим базовым компонентом Семантического Веба — совокупностью информации, которое специалисты именуют *онтологией* (см. статью про инженерию знаний). В философии онтологией называют некую теорию о природе бытия, ИТ-специалисты заимствовали этот термин, и для них уже онтология — это структура, задающая отношения между понятиями.

Онтология определяет классы объектов и отношения между ними. Например, понятие *адрес* может быть определено как разновидность понятия *местонахождение*[location], а *код города* можно задавать применительно лишь к *местонахождениям* и так далее. Задание классов, подклассов, а также отношений между индивидами [entities] является чрезвычайно мощным инструментом для использования в Вебе.

В простейшем случае, онтологии используют для увеличения точности поиска в Вебе — поисковая машина будет выдавать только такие сайты, где упоминается в точности искомое понятие, а не произвольные страницы, в тексте которых встретилось данное многозначное ключевое слово.

В полную силу Семантический Веб будет реализован тогда, когда люди создадут множество *программ - агентов,* которые, знакомясь с содержимым Веба из различных источников, обрабатывают полученную информацию и обмениваются результатами с другими программами. Семантический Веб стимулирует подобного рода синергию: даже те агенты, которые не были созданы специально для работы сообща, могут передавать информацию друг другу, коль скоро эта информация будет снабжена семантикой.

Полноценный Семантический Веб – это технология управления знаниями будущего, однако, уже сейчас можно пользоваться его отдельными технологиями и применять в ограниченных областях. Примером является Семантический веб организации — или реализация этой концепции в рамках отдельной организации [Cerebra, 2005].

Также и онтологии могут не только использоваться в Семантическом Вебе, но и применяться в системах управления знаниями предприятий. Онтологии задают единый язык, повышая тем самым эффективность коммуникаций сотрудников и обмен знаниями. Они могут использоваться для интеграции информации и выполнения простых автоматических рассуждений, повышая тем самым качество результатов поиска информации. Современные исследователи [Mika, Akkermans, 2004; Davies et al, 2005] считают онтологии основной парадигмой управлением знаниями предприятия.

Сегодняшние поисковые системы зачастую выдают бесчисленное множество совершенно не относящихся к запросу «хитов», обрекая пользователя на длительный ручной отбор материала. Например, если вы ввели для поиска слово «орган», то компьютеру совершенно непонятно, имеете ли вы в виду музыкальный инструмент, часть человеческого тела или орган исполнительной власти. Вся проблема в том, что для компьютера слово «орган» не имеет чёткого смысла, или другими словами, семантического содержания.

Онтологии дают возможность производить запросы на основе понятий, а не на основе совпадения строк. Например, если пользователь задаст вопрос *«Какие транспортные средства производятся в России?»*, то он получит из базы ответ, в который попадут автомобили (=подкласс транспортных средств) производимые во Всеволожске (=город, который находится в России).

Также онтологии дают возможность получения не заданных явно знаний из информационных хранилищ путем логического вывода - поиск «скрытой информации». Например, пользователь системы может задать вопрос: *Какие поставки продукции находятся сейчас в состоянии риска?* В ответ на такой вопрос система в одной онтологии тарифов определит, что с учетом текущих условий (например, географических или погодных) существуют риски связанные с перевозкой овощей и фруктов. А в другой базе или онтологии деклараций по перевозке груза определит, что в декларации №А345 указаны арбузы, которые являются подклассом «Овощей и фруктов» (см. рис. 3). В результате, система сможет дать конкретный ответ на поставленный вопрос: *Поставка COSCO #A345*



Рисунок 3 Интеллектуальный поиск на основе логического вывода

Темы докладов

1. Хранилище данных (модель данных типа «звезда»; свойства ХД; витрины данных; OLTP-системы; ETL-погружение данных; архитектура современной корпоративной информационно-аналитической системы)
2. OLAP (многомерное представление данных; операции над данными; требования к программному продукту класса OLAP; тест FASMI; классификация OLAP)
3. Системы управления корпоративной эффективностью (CPM-системы) (CPM-системы; ERP, Cognos)
4. Корпоративный портал.
5. Система поддержки принятия решения.
6. Ситуационные центры.
7. Интеллектуальный анализ данных (data mining)(понятие data mining; ассоциация, последовательность, классификация, кластеризация, временные закономерности; методы data mining; инструменты data mining; применение технологии data mining)
8. Интеллектуальный анализ текстов (text mining)
9. Анализ социальных сетей (SNA)(структурное, ресурсное, нормативное и динамическое направление анализа соц. сетей; методы анализа)
10. Системы управления взаимоотношениями с клиентами – CRM
11. Структура интеллектуального капитала.
12. Подходы к оценке интеллектуального капитала (BSC и EVA, коэффициент Тобина, методы DIC)
13. Базовые методы анализа данных (корреляционный, регрессионный, дисперсионный, факторный, кластерный, дискриминантный).

Приложение 2. Оценочные средства для проведения промежуточной аттестации.

Проверяемая компетенция ПК-6: Способность к формализации и алгоритмизации поставленных задач, к написанию программного кода с использованием языков программирования, определения и манипулирования данными и оформлению программного кода в соответствии установленными требованиями

ПК-6.1. Оценивает качество математической модели при формализации задачи предметной области

Задания:

1. Укажите, какое утверждение о понятиях «данные» и «информация» является верным:

а) данные извлекаются из информации;
б) информация извлекается из данных;
в) понятия «данные» и «информация» тождественны.

2. Укажите, какое утверждение об отношении понятия «средство обработки информации» к понятию «компьютер» является верным:

а) является интенсионалом;
б) является экстенсионалом;
в) является понятием одного уровня абстракции.

3. Укажите, какое утверждение об отношении понятия «накопитель на магнитном диске» к понятию «средство хранения информации» является верным:

а) нормальный закон распределения;
б) закон распределения Пуассона;
в) закон распределения Бернулли.

ПК-6.2. Оценивает качество разработанных алгоритмов для последующего кодирования

Задания:

1. Укажите, как называется процесс трансформации неявных знаний в явные:

а) социализация;
б) формализация;
в) комбинирование.

2. Укажите, как называется процесс трансформации неявных знаний в неявные:

а) социализация;
б) формализация;
в) комбинирование.

3. Укажите, как называется процесс трансформации явных знаний в явные:

а) социализация;
б) формализация;
в) комбинирование.

ПК-6.3. Оценивает выбор программных средств для программирования и манипулирования данными в соответствии установленными требованиями

Задания:

1. Укажите, какое утверждение **не соответствует** структуре треугольника Фреге (семантического треугольника):

а) знак замещает вещь;
б) понятие обозначает знак;
в) понятие ссылается на вещь.

2. Укажите, какое утверждение **не соответствует** структуре треугольника Фреге (семантического треугольника):

а) вещь замещает знак;
б) знак обозначает понятие;
в) понятие ссылается на вещь.

3. Укажите, какое утверждение **не соответствует** структуре треугольника Фреге (семантического треугольника):

а) знак замещает вещь;
б) знак обозначает понятие;
в) вещь ссылается на понятие.