

МИНОБРНАУКИ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ  
Заведующий кафедрой ТО и ЗИ



А.А. Сирота  
21.06.2022 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**

**Б1.В.09 Семантический WEB**

**1. Код и наименование направления подготовки/специальности:**

45.03.03 Фундаментальная и прикладная лингвистика

**2. Профиль подготовки/специализация:**

Экспертно-аналитическая деятельность

**3. Квалификация выпускника:** бакалавр

**4. Форма обучения:** очная

**5. Кафедра, отвечающая за реализацию дисциплины:** кафедра Технологий обработки и защиты информации

**6. Составители программы:** Гаршина Вероника Викторовна, канд.тех.наук, доцент, Степанцов Вячеслав Алексеевич, доцент кафедры Технологий обработки и защиты информации

**7. Рекомендована:** Научно-методическим советом факультета ФКН, протокол № 10 от 10.06.2022 г.

**8. Учебный год:** 2024/2025

**Семестр:** 5

## **9. Цели и задачи учебной дисциплины**

Изучение принципов организации лингвистической информации в WEB и современных методов организации семантического поиска.

Основные задачи дисциплины:

- Изучить основные стандарты описания разметки текстовых документов (W3C)
- Знакомство с принципами моделирования семантики для обработки текстов (онтологическое моделирование)
- Принципы построения выводов заключений на знаниях, представленных онтологией.

Дисциплина реализуется частично в форме практической подготовки (ПП).

**10. Место учебной дисциплины в структуре ООП:** дисциплина Семантический WEB входит в вариативную часть ООП. Для изучения дисциплины необходимы знания, умения и компетенции, сформированные дисциплинами: Б1.О.25 Информатика и основы программирования, Б1.В.03 Введение в прикладную лингвистику, Б1.В.01 Проектирование баз данных.

**11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:**

Код	Название компетенции	Коды	Индикаторы	Планируемые результаты обучения
ПК-14	Владеет принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умеет пользоваться такими ресурсами.	ПК-14.1 ПК-14.2 ПК-14.3	Разрабатывает и документирует программные интерфейсы Пользуется электронными языковыми ресурсами для решения прикладных задач Анализирует требования к программному обеспечению	<b>Знать:</b> принципы создания и работы электронных языковых ресурсов <b>Уметь:</b> – ориентироваться в области профессионально-ориентированных электронных ресурсов; – пользоваться электронными языковыми ресурсами при решении профессиональных задач; – применять на практике знание компьютерных программ для создания электронных языковых ресурсов. <b>Владеть:</b> – организационно-методическими правилами доступа к языковым ресурсам; – основными инструментами работы в них; – навыками поиска информации в сети Интернет, работы с базами

				данных, а также прочими приложениями, необходимыми для программирования в лингвистике.
--	--	--	--	--

**12. Объем дисциплины в зачетных единицах/час.** — 2 з.е. /72 ч.

**Форма промежуточной аттестации:** зачет.

### **13. Трудоемкость по видам учебной работы**

Вид учебной работы	Трудоемкость	
	Всего	По семестрам
		5 семестр
Аудиторные занятия	<b>36</b>	<b>36</b>
в том числе:	лекции	
	практические	
	лабораторные	36 36
Самостоятельная работа	<b>36</b>	<b>36</b>
в том числе: курсовая работа (проект)		
Форма промежуточной аттестации (зачет)		
<b>Итого:</b>	<b>72</b>	<b>72</b>

#### **13.1. Содержание дисциплины**

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
<b>1. Лекции</b>		
Не предусмотрены учебным планом		
<b>2. Практические занятия</b>		
Не предусмотрены учебным планом		
<b>3. Лабораторные занятия</b>		
3.1	Основы организации вычислительных сетей. Сетевые стандарты и протоколы. Сетевые сервисы.	<b>Лабораторная работа 1.</b> Адресация в сети Интернет. Понятие IP адреса, маска подсети. Служба DNS, URL –адреса. Провайдеры интернет-услуг. Типы. Понятие хостинга. Виды хостинга. Способы доступа, подключения к Интернет коммерческих организаций. Функциональные службы в Интернет: E-mail, WWW, FTP, Archie. Функциональные службы в Интернет: USENET, News, WAIS, Whois, Telnet. Мессенджеры-программы или сервисы для мгновенного обмена сообщениями, голосовой связи и видеосвязи в сети Интернет. IP-телефония.
3.2	Организация информационного поиска в Internet.	<b>Лабораторная работа 2.</b> Особенности структуры WEB. Объем. Поверхностная и скрытая части Web. Графовая структура, сообщества, NK-кланы, ссылки. Особенности информационного поиска в WEB по сравнению с информационно-справочными системами. Модель поведения типичного пользователя. Архитектура поисковой системы для WEB. Понятия: хранилище документов, модуль индексирования, индекс,

		сетевой робот, поисковая машина, формат запроса. Стратегии сканирования пространства WEB сетевыми роботами. Особенности сканирования скрытого Web. Понятие релевантного и нерелевантного документа. Методы ранжирования результатов поиска. Модели PageRank, “голосования”, HITS. Архитектуры Информационно-поисковых систем (ИПС): распределенные, метапоисковые.
3.3	Язык гипертекстовой разметки HTML для задач хранения и обработки электронных хранилищ документов.	<p><b>Лабораторная работа 3.</b></p> <p>Язык разметки HTML. Структура документа. Основные тэги языка. Обязательные метки . Теги <code>&lt;html&gt;</code> и <code>&lt;/html&gt;</code>, <code>&lt;head&gt;</code> и <code>&lt;/head&gt;</code>, <code>&lt;title&gt;</code> и <code>&lt;/title&gt;</code>, <code>&lt;body&gt;</code> и <code>&lt;/body&gt;</code>, <code>&lt;H1&gt;</code> и <code>&lt;/H1&gt;</code> — <code>&lt;H6&gt;</code> и <code>&lt;/H6&gt;</code>, <code>&lt;P&gt;</code> и <code>&lt;/P&gt;</code> Стилевое оформление текста. Гиперссылки. Форматирование текста с помощью тега <code>&lt;FONT&gt;</code>. Специальные символы, <code>&lt;HR&gt;</code> и <code>&amp;</code> - последовательности. Горизонтальные линейки. Разделители. Непарные метки. Конец строки <code>&lt;BR&gt;</code>.</p> <p><b>Лабораторная работа 4.</b></p> <p>Списки – упорядоченные, неупорядоченные, вложенные. Ненумерованные списки: <code>&lt;UL&gt;</code> и <code>&lt;/UL&gt;</code> Нумерованные списки: <code>&lt;OL&gt;</code> и <code>&lt;/OL&gt;</code> Списки определений: <code>&lt;DL&gt;</code> и <code>&lt;/DL&gt;</code> Списки, формы. Сложные формы. Изображения в документе. Фреймы. Тэги <code>&lt;META&gt;</code>.</p> <p><b>Лабораторная работа 5</b></p> <p>JavaScript/Jscript. Управляющие структуры – структура выбора <code>if</code> , структура выбора <code>if/else</code>, Структуры повторения <code>while</code>, <code>for</code>, <code>do /while</code> . Операторы <code>break</code> , <code>continue</code>. Операции присваивания, операции инкремента и декремента. Типы данных. Структура с множественным выбором <code>switch</code>. Логические операции.</p>
3.4	Язык гипертекстовой разметки XML для задач хранения и обработки электронных хранилищ документов.	<p><b>Лабораторная работа 6.</b></p> <p>Структура XML документа. Пролог и корневой элемент. Комментарии. Специальные символы. Пример XML файла. XML процессор. Использование инструкций по обратотке. Директивы анализатора. Правила создания XML документа. Необходимое программное обеспечение для работы с XML документом. Конструкции языка XML. Элементы данных. Объявление типов элементов. Описание содержимого элемента. Основные модели задания содержимого элемента: последовательная, выборочная, смешанные. Иерархическая организация элементов. Элементы данных и их атрибуты. Типы атрибутов. Правила создания атрибутов. Типы атрибутов: Маркерный тип, нумерованный тип, значения по умолчанию. Отображение XML документа: с использованием таблиц стилей и без них. Таблицы стилей <code>*.css</code>. Основные стандартные элементы для отображения шрифта и форматирования текста. Форма записи раздела CDATA. Понятие валидного XML документа.</p> <p><b>Лабораторная работа 7.</b></p> <p>Documents Type Definitions (DTD). Внутреннее и внешнеописание DTD. Схемы данных для построения XML документов. Область схемы данных. Описание элементов.</p>

		Атрибуты элементов в схемах данных. Модель содержимого элемента. Группировка элементов. Закрытая и открытая модели описания содержимого элемента.
3.5	Языки разметки на базе XML Язык разметки AIML	<b>Лабораторная работа 8.</b> Структура AIML документа, основные тэги языка. Примеры работы с редактором AIML – GAITOBOT. Программный чат-бот, его программирование, регистрация в WWW, сообщества чат-ботов. Области применения.
3.6	Концепции WEB2.0, WEB3.0, Семантический WEB	<b>Лабораторная работа 9.</b> Семь технологических уровней, на которых базируется Семантическая сеть. Resource Description Framework (RDF) и <a href="#">RDF Schema</a> , –синтаксическая модель для описания ресурсов, Базовая модель RDF. Синтаксис RDF. Синтаксис сериализации. Схемы и пространства имен. Контейнеры. Формальная модель RDF. Формальная грамматика RDF. Агрегирование. N-арные отношения. Dublin Core Metadata. Язык описания семантики ресурсов - OWL, язык запросов к знаниям SPARQL.
3.7	Онтологии и тезаурусы	<b>Лабораторная работа 10.</b> Понятие онтологии. Задачи, решаемые с помощью онтологий. Элементы онтологии: экземпляры, концепты, атрибуты, отношения. Картирование онтологий. Эквивалентность между классами и свойствами. Сложные классы. Типы онтологий: верхнего уровня, предметных областей, прикладные онтологии. Языки описания онтологий. Стандарт OWL. Виды OWL. Инstrumentальные средства проектирования онтологий. Примеры крупных онтологических проектов.
3.8	Многоагентные интеллектуальные системы для задач семантического поиска	<b>Лабораторная работа 11.</b> Агенты, Сообщества агентов. Мультиагентные системы. Основные типы агентных моделей и архитектур. Делиберативные, реактивные, гибридные. Восприятие. Агенты с состояниями. Методологии построения агентно-ориентированных систем. MAS DK, Gaia, Tropos. Сообщества агентов и протоколы взаимодействия: KQML. KIF. Языки программирования агентов. Программные платформы разработки агентно-ориентированных систем JADE.

### 13.2. Темы (разделы) дисциплины и виды занятий

№ п/ п	Наименование темы (раздела) дисциплины	Виды занятий (часов)				
		Лекци и	Практ . .	Лаб.	Сам. работа	Всего
1	Основы организации вычислительных сетей. Сетевые стандарты и протоколы. Сетевые сервисы.			4	4	8
2	Концепции WEB2.0, WEB3.0, Семантический WEB			4	4	8
3	Организация информационного поиска в Internet.			4	4	8
4	Язык гипертекстовой разметки			6	6	12

	HTML для задач хранения и обработки электронных хранилищ документов.					
5	Языки разметки на базе XML Язык разметки AIML		5	5	10	
6	Концепции WEB2.0, WEB3.0, Семантический WEB		5	5	10	
7	Онтологии и тезаурусы		5	5	10	
8	Многоагентные интеллектуальные системы для задач семантического поиска		3	3	6	
			36	36	72	

#### 14. Методические указания для обучающихся по освоению дисциплины

1) При изучении дисциплины рекомендуется использовать следующие средства:

- рекомендуемую основную и дополнительную литературу;
- методические указания и пособия;
- контрольные задания для закрепления теоретического материала;

электронные версии учебников и методических указаний для выполнения лабораторно-практических работ (при необходимости материалы рассыпаются по электронной почте).

2) Для максимального усвоения дисциплины рекомендуется проведение письменного опроса (тестирование, решение задач) студентов по материалам лекций и практических работ. Подборка вопросов для тестирования осуществляется на основе изученного теоретического материала. Такой подход позволяет повысить мотивацию студентов при конспектировании лекционного материала.

3) При проведении лабораторных занятий обеспечивается максимальная степень соответствия с материалом лекционных занятий и осуществляется экспериментальная проверка методов, алгоритмов и технологий обработки информации, излагаемых в рамках лекций.

4) При переходе на дистанционный режим обучения для создания электронных курсов, чтения лекций онлайн и проведения лабораторно-практических занятий используется информационные ресурсы Образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете.

5) При использовании дистанционных образовательных технологий и электронного обучения обучающиеся должны выполнять все указания преподавателей, вовремя подключаться к онлайн - занятиям, ответственно подходить к заданиям для самостоятельной работы. В ходе самостоятельной работы необходимо уделить основное внимание работе с текстом конспекта лекции, изучению рекомендованной литературы, изучению нормативных документов по информационной безопасности.

#### 15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

(список литературы оформляется в соответствии с требованиями ГОСТ и используется общая сквозная нумерация для всех видов источников)

а) основная литература:

№ п/п	Источник
1	Семантический веб / Г. Антониоу, П. Грос, в. Ф. Хармлен, Р. Хоекстра ; перевод с английского Т. Шульга. — Москва : ДМК Пресс, 2016. — 240 с. — ISBN 978-5-97060-333-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/69963">https://e.lanbook.com/book/69963</a> (дата

	обращения: 31.08.2020). — Режим доступа: для авториз. пользователей.
2	Маннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. — Вильямс, 2011, 528 с.
3	Лукашевич Н. В. Тезаурусы в задачах информационного поиска. — Издательство МГУ имени М. В. Ломоносова, 2011
4	Онтологии и тезаурусы: модели, инструменты, приложения: учебное пособие / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. - М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2009. - 173 с.: ил.
5	Стивен Шафер. HTML, XHTML и CSS. Библия пользователя, 5-е издание = HTML, XHTML, and CSS Bible, 5th Edition. — М.: «Диалектика», 2010. — 656 с.
6	Дэвид Хантер, Джейфф Рафтер, Джо Фаусетт, Эрик ван дер Влист, и др. XML. Работа с XML. 4-е издание. — М.: «Диалектика», 2009.
7	Рассел С., Норвиг П. Искусственный интеллект: современный подход. — М. : Вильямс , 2006.

б) дополнительная литература:

№ п/п	Источник
8	Towards the Semantic Web: Ontology-driven Knowledge Management. — John Wiley & Sons, 2003.
9	Toby Segaran, Colin Evans, Jamie Taylor Programming the Semantic Web. — 2009.
10	<a href="#"><b>Google Semantic Search: Search Engine Optimization (SEO) Techniques That Get Your Company More Traffic, Increase Brand Impact, and Amplify Your Online Presence (Que Biz-Tech)</b></a> , David Amerland, 240 pages, 10 July 2013
11	<a href="#"><b>Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph Foundations of Semantic Web Technologies</b></a> Chapman & Hall/CRC, 2009, 455 pages
12	<a href="#"><b>Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL</b></a> by Dean Allemang and James Hendler, Publisher: Morgan Kaufmann, 352 pages, 2008.
13	<a href="#"><b>Google Analytics для профессионалов</b></a> , Брайан Клифтон, 3-е издание. — М.: «Диалектика», 2013, 608 стр.
14	<a href="#"><b>Programming the Semantic Web</b></a> by Toby Segaran, Colin Evans and Jamie Taylor, - O'Reilly Media , 2009, 298 pages.
15	Добров Б.В. Онтологии и тезаурусы: модели, инструменты, приложения: учебное пособие / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. / - М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2009.
16	Леонтьева Н.Н. Автоматическое понимание текстов. М., 2006

в) информационные электронно-образовательные ресурсы (официальные ресурсы интернет)\*:

№ п/п	Ресурсы Интернет
17	Электронный каталог Научной библиотеки Воронежского государственного университета. – ( <a href="http://www.lib.vsu.ru/">http://www.lib.vsu.ru/</a> ).
18	Образовательный портал «Электронный университет ВГУ». – ( <a href="https://edu.vsu.ru/">https://edu.vsu.ru/</a> ).
19	«Университетская библиотека online» - Контракт № 3010-07/33-19 от

	11.11.2019 «Консультант студента» - Контракт № 3010-07/34-19 от 11.11.2019 ЭБС «Лань» - Договор 3010-04/05-20 от 26.02.2020. «РУКОНТ» (ИТС Контекстум) - Договор ДС-208 от 01.02.2018 ЭБС «Юрайт» - Договор № 43/8 от 10.02.2020.
--	---

\* Вначале указываются ЭБС, с которыми имеются договора у ВГУ, затем открытые электронно-образовательные ресурсы

**16. Перечень учебно-методического обеспечения для самостоятельной работы (учебно-методические рекомендации, пособия, задачники, методические указания по выполнению практических (контрольных) работ и др.)**

№ п/п	Источник
1	С.В. Одиночкина Основы технологий XML - СПб: НИУ ИТМО, 2013. – 56 с.
2	Д.И. Муромцев. Онтологический инжиниринг знаний в системе Protégé. – СПб: СПб ГУ ИТМО, 2007. – 62 с.

**17. Информационные технологии, используемые для реализации учебной дисциплины, включая программное обеспечение и информационно-справочные системы (при необходимости)**

Для реализации учебного процесса используются:

1. ПО Microsoft в рамках подписки "Imagine/Azure Dev Tools for Teaching", договор №3010-16/96-18 от 29 декабря 2018г.
2. ПО Редактор онтологий и фреймворк для построения баз знаний Protege. Свободно-распространяемое ПО.
3. При проведении занятий в дистанционном режиме обучения используются технические и информационные ресурсы Образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете, а также другие доступные ресурсы сети Интернет.

**18. Материально-техническое обеспечение дисциплины:**

*(при использовании лабораторного оборудования указывать полный перечень, при большом количестве оборудования можно вынести данный раздел в приложение к рабочей программе)*

Мультимедийная лекционная аудитория, персональный компьютер (ПК) рабочее место преподавателя: проектор, видеокоммутатор, персональные компьютеры (ПК), наушники и микрофоны по числу студентов, специализированная мебель: доска меловая., столы, стулья; выход в Интернет, доступ к фондам учебно-методической документации и электронным изданиям

**19. Оценочные средства для проведения текущей и промежуточной аттестаций**

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенции	Индикаторы достижения компетенции	Оценочные средства
----------	---	-------------	---	-----------------------

№ п/п	Наименование раздела дисциплины (модуля)	Компетенции	Индикаторы достижения компетенции	Оценочные средства
1.	1. Основы организации вычислительных сетей. Сетевые стандарты и протоколы. Сетевые сервисы. 2. Концепции WEB2.0, WEB3.0, Семантический WEB 3. Организация информационного поиска в Internet. 4. Язык гипертекстовой разметки HTML для задач хранения и обработки электронных хранилищ документов. 5. Языки разметки на базе XML Язык разметки AIML 6. Концепции WEB2.0, WEB3.0, Семантический WEB 7. Онтологии и тезаурусы 8. Многоагентные интеллектуальные системы для задач семантического поиска	ПК-14	Разрабатывает и документирует программные интерфейсы (ПК-14.1) Пользуется электронными языковыми ресурсами для решения прикладных лингвистических задач (ПК-14.2) Анализирует требования программному обеспечению (ПК-14.3)	Лабораторные работы 3-11
Промежуточная аттестация форма контроля – зачет				КИМ

## 20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

### 20.1 Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств:

№ п/ п	Наименование оценочного средства	Представление оценочного средства в фонде	Критерии оценки
			4
1	Устный опрос	Вопросы по темам/разделам дисциплины	Правильный ответ – зачтено, неправильный или принципиально неточный ответ - не зачтено
2	Контрольная работа по разделам дисциплины	Теоретические вопросы по темам/разделам дисциплины	Шкала оценивания соответствует приведенной в разделе 19.2
3	Лабораторная	Содержит	При успешно выполнении

	работа	лабораторные задания, предусматривающих освоение программных систем.	работы ставится оценка зачтено и осуществляется допуск к зачету, в противном случае ставится оценка не зачтено и обучающийся не допускается к зачету.
4	КИМ промежуточной аттестации	Каждый контрольно-измерительный материал для проведения промежуточной аттестации включает 2 заданий вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции.	Шкалы оценивания приведены в разделе 19.2

#### **Примерный перечень вопросов к зачету:**

1. XML как основа построения семантического информационного пространства в WWW сети.
2. Семь технологических уровней, на которых базируется Семантическая сеть, предложенные Тимом Бернерс-Ли.
3. Языки гипертекстовой разметки HTML и XML для задач хранения и обработки электронных хранилищ документов. Их достоинства и недостатки.
4. Автономные интеллектуальные агенты. Сообщества агентов для решения задач поиска в WEB.
5. Структура XML документа. Пролог и корневой элемент. Комментарии. Специальные символы. Пример XML файла.
6. XML процессор. Использование инструкций по обработке. Директивы анализатора. Примеры.
7. Правила создания XML документа. Необходимое программное обеспечение для работы с XML документом.
8. Конструкции языка XML. Элементы данных. Объявление типов элементов. Описание содержимого элемента.
9. Основные модели задания содержимого элемента: последовательная, выборочная, смешанные. Примеры.
10. Иерархическая организация элементов. Пример XML файла. Элементы данных и их атрибуты. Типы атрибутов. Правила создания атрибутов. Примеры.
11. Типы атрибутов: Маркерный тип, нумерованный тип, значения по умолчанию.
12. Отображение XML документа: с использованием таблиц стилей и без них. Пример.
13. Таблицы стилей \*.css. Основные стандартные элементы для отображения шрифта и форматирования текста.
14. Форма записи раздела CDATA. Примеры. Понятие валидного XML документа.
15. Documents Type Definitions (DTD). Внутреннее и внешнее описание DTD.
16. Схемы данных для построения XML документов. Область схемы данных. Описание элементов. Атрибуты элементов в схемах данных. Модель содержимого элемента. Примеры.
17. Группировка элементов. Закрытая и открытая модели описания содержимого элемента.
18. Структура AIML документа, основные тэги языка.

19. Примеры работы с редактором AIML – GAITOBOT.
20. Программный бот, его программирование, регистрация в WWW, сообщества ботов. Области использования ботов.
21. Адресация в сети Интернет. Понятие IP адреса, маска подсети. Служба DNS, URL –адреса.
22. Провайдеры интернет-услуг. Типы. Понятие хостинга. Виды хостинга. Способы доступа, подключения к Интернет коммерческих организаций.
23. Функциональные службы в Интернет: E-mail, WWW, FTP, Archie.
24. Функциональные службы в Интернет: USENET, News, WAIS, Whois, Telnet.
25. Мессенджеры-программы или сервисы для мгновенного обмена сообщениями, голосовой связи и видеосвязи в сети Интернет.
26. Особенности структуры WEB. Объем. Поверхностная и скрытая части Web. Графовая структура, сообщества, NK-кланы, ссылки.
27. Особенности информационного поиска в WEB по сравнению с информационно-справочными системами. Модель поведения типичного пользователя.
28. Архитектура поисковой системы для WEB. Понятия: хранилище документов, модуль индексирования, индекс, сетевой робот, поисковая машина, формат запроса.
29. Стратегии сканирования пространства WEB сетевыми роботами. Особенности сканирования скрытого Web.
30. Понятие релевантного и нерелевантного документа. Методы ранжирования результатов поиска. Модели PageRank, “голосования”, HITS.
31. Архитектуры Информационно-поисковых систем(ИПС): распределенные, метапоисковые.
32. Понятие онтологии, элементы онтологии: экземпляры (примеры), понятия (концепты), атрибуты, отношения. Назначение онтологий. Задачи, решаемые с помощью онтологий.
33. Классификации онтологий. Типы онтологий: верхнего уровня, предметных областей, прикладные онтологии. Примеры крупных онтологических проектов CYC, SUMO, Sowa ontology.
34. Математическая модель определения онтологии в предметной области.
35. Языки описания онтологий – XML, RDF, RDFS, OWL.
36. Возможности применения онтологий для семантического анализа предметной области. Семантическая аналитика - Text mining.
37. Направления применения онтологического моделирования при проектировании информационных систем. Примеры проектов, использующих интеграцию с онтологиями и обработкой семантики.

## Пример задания для выполнения лабораторной работы

### Лабораторная работа № 1

#### Цель лабораторной работы.

Создание классов [онтологической модели](#) в Protégé (<http://protege.stanford.edu>).

Рекомендации по выполнению лабораторной работы.

Создание и разработка онтологий с помощью языков [RDF/RDFS](#) и [OWL](#), а также выполнение SPARQL – запросов возможны в редакторе Protégé.

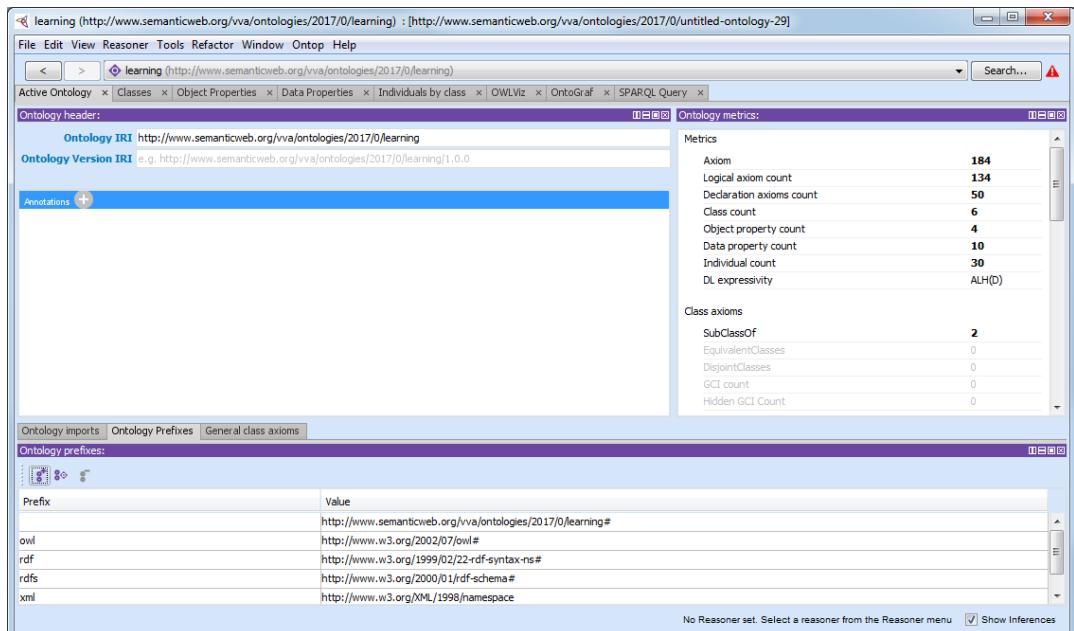


Рис.1. Редактор онтологий Protégé

Перед созданием онтологии необходимо настроить отображение в окне программы необходимых вкладок.

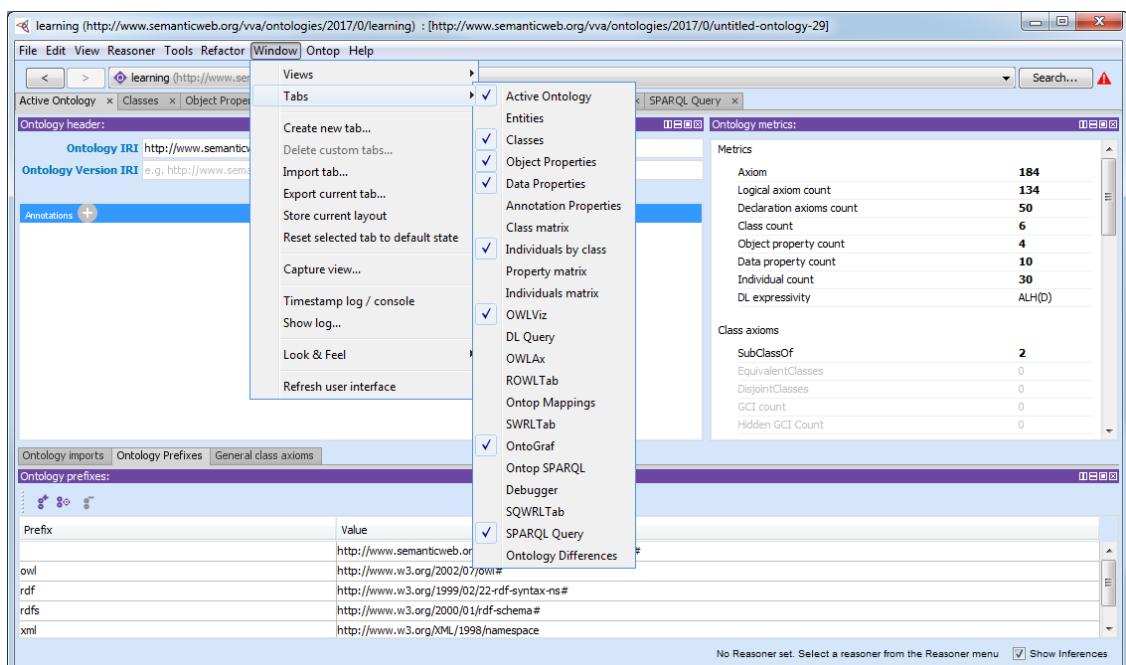


Рис.2. Выбор отображаемых вкладок

#### Назначение вкладок:

- Active Ontology – отображение [IRI](#), общих характеристик (количества аксиом, классов, свойств, объектов и т.п.) и аннотации онтологии;
- Classes – просмотр и редактирование классов;
- Object Properties – просмотр и редактирование свойств-отношений между индивидами;
- Data Properties – просмотр и редактирование свойств-данных индивидов;
- Individuals by class – просмотр и редактирование экземпляров классов (индивидуов);
- OWLViz и OntoGraf – визуализация онтологии в виде графа;
- SPARQL Query – выполнение запросов.

Классы создаются на вкладке «Classes». Добавить новый (удалить имеющийся) класс или подкласс можно с помощью кнопок панели инструментов или выбора пункта контекстного меню.

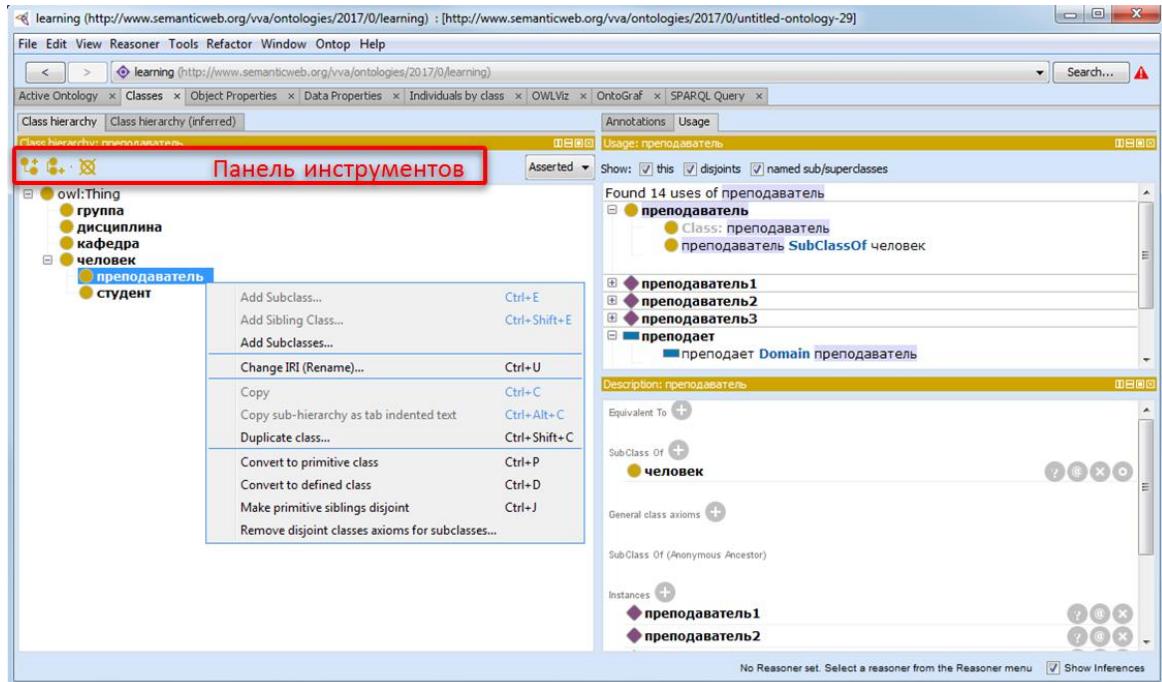


Рис.3. Вкладка «Classes»

В [OWL](#) базовым классом, на основе которого создаются классы онтологии, является класс «owl:Class». Остальные классы по отношению к нему являются дочерними подклассами (англ. Subclass). Классы одного уровня иерархии в Protégé называются родственными (англ. Sibling Class).

На вкладке «Применение» (англ. Usage) для класса, выделенного в иерархии, отображается его связь с родительскими и дочерними классами, его свойства, экземпляры и т.п.

В правой нижней панели «Описание» (англ. Description) можно указать дополнительные характеристики класса. Например, его эквивалентность другим классам (англ. Equivalent To) или невозможность принадлежности экземпляров класса другим классам (англ. Disjoint With) – запрет множественного наследования.

Свойства классов и их экземпляров (предикаты [RDF-треек](#)) делятся на два вида:

- **свойства-отношения** задаются на вкладке «Object Properties» и определяют некоторые отношения между двумя индивидами (экземплярами классов), т.е. субъектом и объектом [RDF-трееки](#) будут индивиды;
- **свойства-данные** задаются на вкладке «Data Properties» и определяют некоторые фактические характеристики индивидов (экземпляров классов), т.е. субъектом [RDF-трееки](#) будет индивид, а объектом значение характеристики в виде строки, числа, даты и т.п.

Создание и редактирование свойств-отношений выполняется на вкладке «Object Properties».

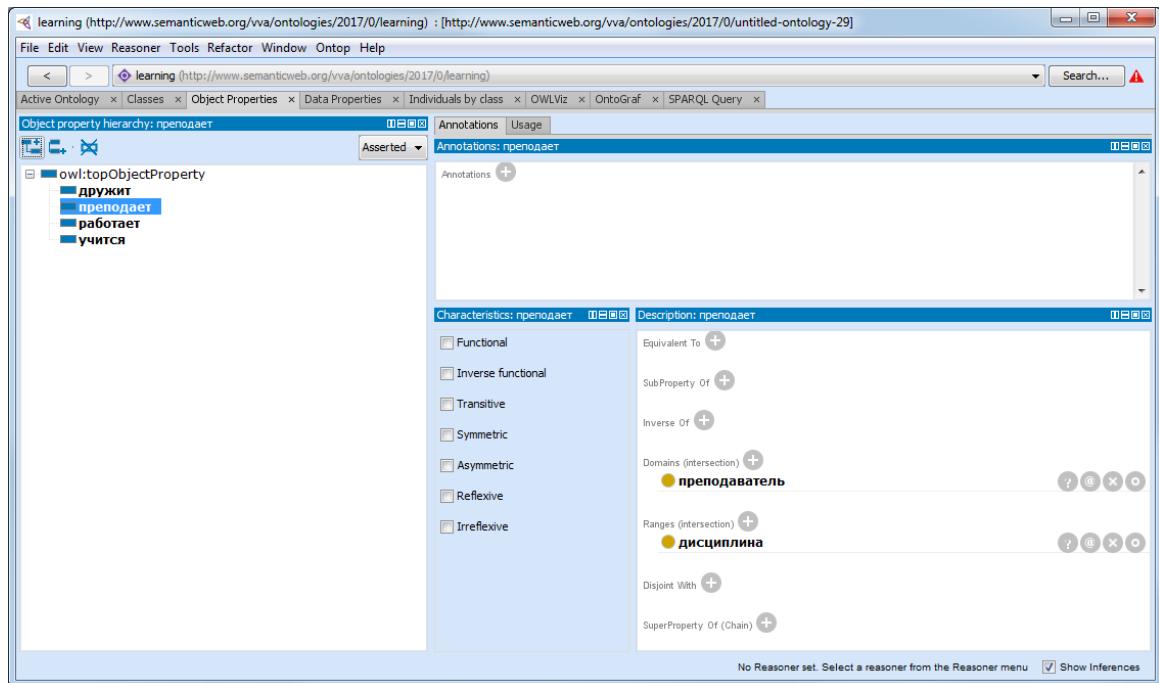


Рис.4. Вкладка «Object Properties»

Домен (англ. Domain) указывает, экземпляры каких классов в RDF-треюке при использовании данного свойства будут выступать в качестве субъектов, а диапазон (англ. Range) – в качестве объектов.

На панели «Характеристики» (англ. Characteristics) выбираются дополнительные характеристики свойства: транзитивность (англ. Transitive), симметричность (англ. Symmetric), рефлексивность (англ. Reflexive) и т.п.

Создание и редактирование свойств-данных выполняется на вкладке «Data Properties».

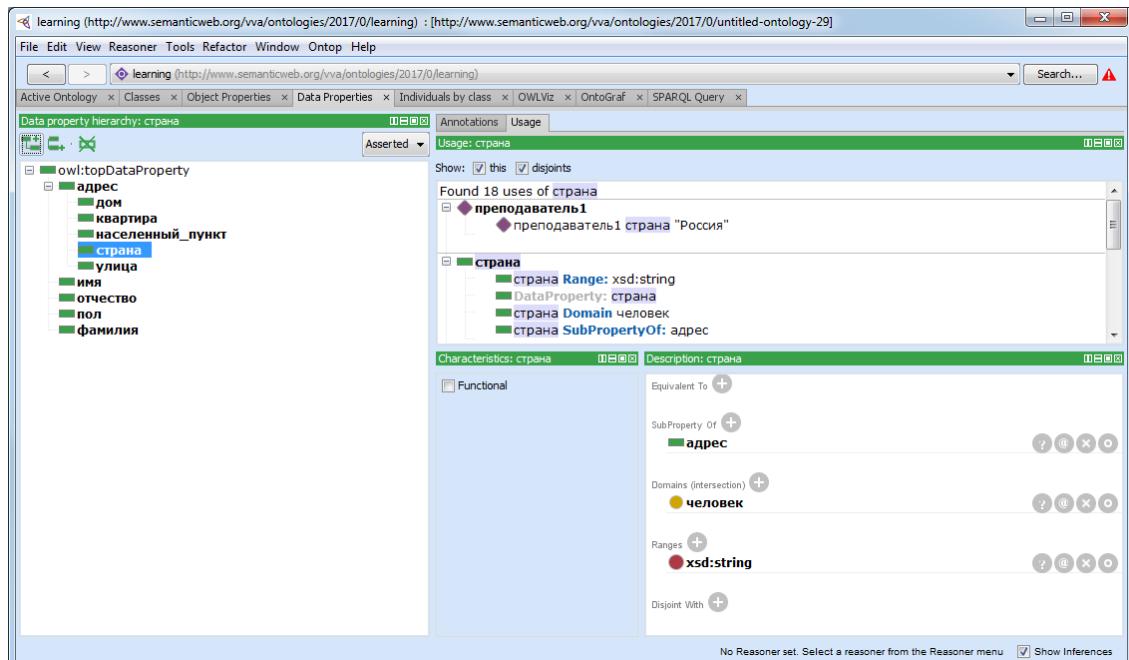


Рис.5. Вкладка «Data Properties»

Домен указывает, для экземпляров каких классов данное свойство может быть использовано. Диапазон задает область допустимых значений (тип данных и ограничения), которые можно указывать для данного свойства экземпляру класса. Тип данных выбирается из стандартных типов XML.

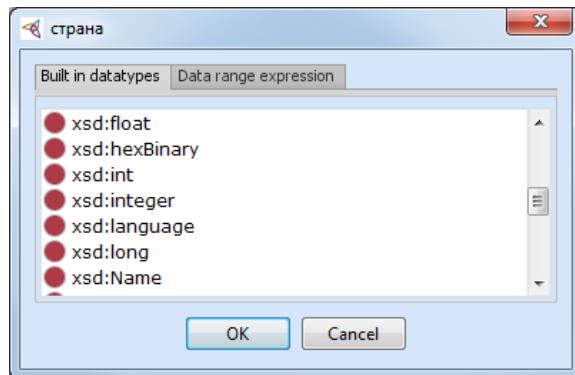
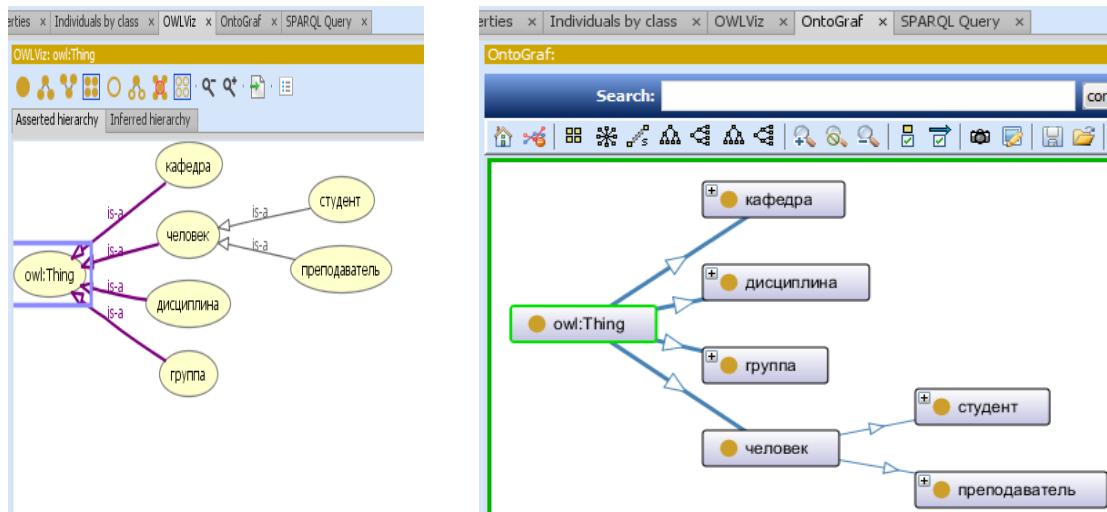


Рис.6. Окно выбора типа данных

Автоматическое построение и отображение графа онтологии и взаимосвязей между классами выполняется при выборе вкладок «OWLviz» и «OntoGraf».



а) OWLviz

б) OntoGraf

Рис.7. Варианты графического отображения онтологии

Перед записью на диск (пункт меню «File / Save as ...») можно выбрать формат (нотацию) хранения онтологии.

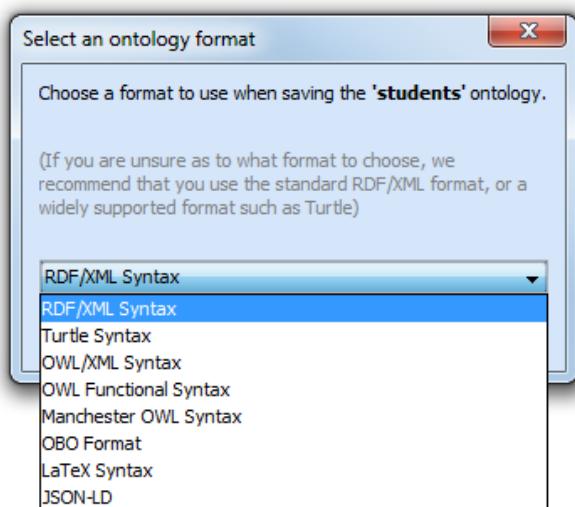


Рис.8. Выбор формата хранения онтологии

### Задание на выполнение лабораторной работы.

А) Построить онтологическую модель в редакторе Protégé, включающую не менее 10 классов и подклассов. Для каждого класса и подкласса определить 1-2 свойства- отношения и 2-10 свойств-данных.

**Б)** Предметная область онтологии выбирается по индивидуальному заданию:

1. автомобили;
2. самолеты;
3. железная дорога;
4. флора;
5. фауна;
6. искусственный интеллект;
7. информационные системы;
8. библиотека;
9. персонал компании;
10. медицина;
11. строительство;
12. астрономия;
13. физика;
14. путешествия;
15. населенные пункты;
16. география;
17. гаджеты;
18. вооруженные силы;
19. книжный магазин;
20. недвижимость;
21. печатное издательство;
22. гостиница;
23. авиакомпания;
24. рекламное агентство;
25. игровая индустрия.

**В)** Отчет должен содержать:

- титульный лист;
- описание задания;
- копии экранов программы (по одной с описанием класса, свойства-отношения и свойства-данных);
- граф онтологии (OWLVis или OntoGraf);
- текст онтологии, сохраненной в формате «Turtle Syntax»;

#### **Пример контрольно-измерительного материала**

УТВЕРЖДАЮ

Заведующий кафедрой технологий обработки и защиты информации

\_\_\_\_\_ А.А. Сирота

\_\_\_\_\_.\_\_\_\_\_.2020

Направление подготовки / специальность 45.03.03 Фундаментальная и прикладная лингвистика

Дисциплина Б1.В.09 Семантический WEB

Форма обучения Очное

Вид контроля Зачет

Вид аттестации Промежуточная

#### **Контрольно-измерительный материал № 1**

1. Конструкции языка XML. Элементы данных. Объявление типов элементов. Описание содержимого элемента.

2. Понятие релевантного и нерелевантного документа. Методы ранжирования результатов поиска. Модели PageRank, “голосования”, HITS.

Преподаватель \_\_\_\_\_ В.А.Степанцов

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенций в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в формах устного опроса (индивидуальный опрос, фронтальная беседа) и письменных работ (контрольные, лабораторные работы). При оценивании могут использоваться количественные или качественные шкалы оценок.

**Промежуточная аттестация может включать в себя теоретические вопросы, позволяющие оценить уровень полученных знаний и/или практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.**

При оценивании используется количественная шкала.

Для оценивания результатов обучения на зачете используются следующие содержательные показатели (формулируются с учетом конкретных требований дисциплины):

- 1) знание теоретических основ учебного материала, основных определений, понятий и используемой терминологии;
- 2) умение проводить обоснование и представление основных теоретических и практических результатов (теорем, алгоритмов, методик) с использованием математических выкладок, блок-схем, структурных схем и стандартных описаний к ним;
- 3) умение связывать теорию с практикой, иллюстрировать ответ примерами, в том числе, собственными, умение выявлять и анализировать основные закономерности, полученные, в том числе, в ходе выполнения лабораторно-практических заданий;
- 4) умение обосновывать свои суждения и профессиональную позицию по излагаемому вопросу;
- 5) владение навыками программирования и экспериментирования в рамках выполняемых лабораторных заданий;

Различные комбинации перечисленных показателей определяют критерии оценивания результатов обучения (сформированности компетенций) на зачете:

- высокий (углубленный) уровень сформированности компетенций;
- повышенный (продвинутый) уровень сформированности компетенций;
- пороговый (базовый) уровень сформированности компетенций.

Для оценивания результатов обучения на зачете используется – зачтено (выше порогового уровня), не зачтено (ниже порогового уровня) по результатам тестирования.

Соотношение показателей, критерии и шкалы оценивания результатов обучения на государственном экзамене представлено в следующей таблице.

**Критерии оценивания компетенций и шкала оценок**

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
---------------------------------	--------------------------------------	--------------

Обучающийся демонстрирует полное соответствие знаний, умений, навыков по приведенным критериям свободно оперирует понятийным аппаратом и приобретенными знаниями, умениями, применяет их при решении практических задач.	Повышенный уровень	Отлично
Ответ на контрольно-измерительный материал не полностью соответствует одному из перечисленных выше показателей, но обучающийся дает правильные ответы на дополнительные вопросы. При этом обучающийся демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателям, но допускает незначительные ошибки, неточности, испытывает затруднения при решении практических задач.	Базовый уровень	Хорошо
Обучающийся демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателям, допускает значительные ошибки при решении практических задач. При этом ответ на контрольно-измерительный материал не соответствует любым двум из перечисленных показателей, обучающийся дает неполные ответы на дополнительные вопросы.	Пороговый уровень	Удовлетворительно
Ответ на контрольно-измерительный материал не соответствует любым трем из перечисленных показателей. Обучающийся демонстрирует отрывочные, фрагментарные знания, допускает грубые ошибки	–	Неудовлетворительно

### 20.3 Материалы для диагностической работы

1. Необходимо корректно записать код гиперссылки на сайт [www.certifications.ru](http://www.certifications.ru), чтобы пользователь мог произвести по ней щелчок.  
Какой из представленных фрагментов кода позволяет решить поставленную задачу?

- 1) <A TARGET="http://www.certifications.ru">www.certifications.ru</A>
- 2) <A HREF="http://www.certifications.ru"></A>
- 3) <A HREF="http://www.certifications.ru" NAME="www.certifications.ru"></A>
- 4) <A HREF="www.certifications.ru">www.certifications.ru</A>
- 5) <A HREF="http://www.certifications.ru">www.certifications.ru</A>

**Ответ – 5**

2. Приведен код таблицы.

```
<table border>
<tr>
<td>&nbsp;</td>
<td>&nbsp;</td>
<td rowspan="2">&nbsp;</td>
<td>&nbsp;</td>
</tr>
<tr>
<td colspan="2">&nbsp;</td>
<td>&nbsp;</td>
</tr>
<tr>
<td>&nbsp;</td>
<td colspan="3">&nbsp;</td>
</tr>
</table>
```

Какой из вариантов таблиц соответствует структуре кода?

1)


2)


3) Ни один из вариантов не подходит


4)


**Ответ – 4**

3. Необходимо сделать текстовое содержимое первого ряда таблицы центрованным по горизонтали в ячейке и отформатированным полужирным стилем.

Какой из представленных фрагментов кода позволяет решить поставленную задачу?

- 1) <TR><TH>содержимое</TH></TR>
- 2) <TH><TD>содержимое</TD></TH>
- 3) <TR><TD>содержимое</TD></TR>
- 4) <TR><THEAD>содержимое</THEAD></TR>
- 5) <TR><TD><B>содержимое</B></TD></TR>

**Ответ – 5**

4. Какая ошибка, допущенная в следующем теге, приводит к неработоспособности внешней таблицы стилей?

<LINK REL="stylesheet" HREF="my.css" TYPE="text/html">

- 1) Необходим атрибут MEDIA
- 2) Значение атрибута TYPE должно быть "text/css"
- 3) Необходим закрывающий тег
- 4) Значения атрибутов следует записать без кавычек
- 5) Вместо атрибута REL следует задать атрибут REV

**Ответ – 2**

5. Требуется, чтобы при наведении курсора мыши на гиперссылку рядом с ней появлялся всплывающий текст.

Какой атрибут тега <A> необходимо использовать для решения поставленной задачи?

**Ответ – ALT**

6. Укажите какой из перечисленных атрибутов являются обязательными для вставки изображения тегом IMG? (BORDER, ALT, SRC, HSPACE, WIDTH)

**Ответ – SRC**

7. Необходимо в html документе создать заголовок четвёртого уровня. Напишите строку html кода для отображения в браузере – Заголовок-4 уровня

**Ответ – <H4>Заголовок-4 уровня</H4>**

8. Какой из вариантов вложения тегов является синтаксически верным?

- 1) <B><U><I></I></U></B>
- 2) <B><U><I></B></I></U>
- 3) <B><U><I></U></B></I>
- 4) <B><U><I></B></U></I>
- 5) <B><U><I></I></B></U>

**Ответ – 1**

9. Требуется создать упорядоченный (нумерованный) список. Какой из тегов позволяет решить поставленную задачу?

**Ответ – <OL>**

10. Чем является любая область данных, заключенная между последовательностями символов: <!---->

- 1) Комментариями
- 2) Инструкциями
- 3) Описанием

4) Сущностями

**Ответ – Комментариями**

11. Для описания данных XML использует DTD (Document Type Definition - Определение типа документа) или \_\_\_\_ документа

- 1) инструкции
- 2) структуру
- 3) оглавление
- 4) схему

**Ответ – схему**

12. Набор общепринятых правил, определяющих механизм взаимодействия компьютеров в сети, называют

- 1) языком
- 2) интерфейсом
- 3) протоколом
- 4) программой

**Ответ –3**

13. Чтобы задать область документа, которую при разборе анализатор будет рассматривать как простой текст, игнорируя любые инструкции и специальные символы, но, в отличие от комментариев, иметь возможность использовать их в приложении, необходимо использовать тэги

- 1) <! [cdata] и ] ] !>
- 2) <? и ?>
- 3) <! [cdata] и [cdata]!>
- 4) <! [cdata] и ] ] >

**Ответ –4**

14. Элемент в DTD определяется при помощи дескриптора \_\_\_, в котором указывается название элемента и структура его содержимого

- 1) <!ELEMENT element-name (element-content)>
- 2) <?ELEMENT element-name (element-content)?>
- 3) <!--ELEMENT element-name (element-content)-->
- 4) <!ELEMENT element-name (element-content)!>

**Ответ –1**

15. Как называется расширяемый язык разметки, разработанный консорциумом W3C для представления информационных ресурсов Web нового поколения

## **Ответ – XML**