

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

«Утверждаю»
Заведующий кафедрой ТО и ЗИ



А.А. Сирота

05.07.2018 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
Б1.В.ОД.4Семантический WEB

- 1. Шифр и наименование направления подготовки/специальности:**
45.03.03 Фундаментальная и прикладная лингвистика
- 2. Профиль подготовки/специализации:-**
- 3. Квалификация (степень) 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 кафедры Технологий обработки и защиты информации
- 7. Рекомендована:** Научно-методическим советом ФКН, протокол № 6 от 25.06.2018 г.
- 8. Учебный год:** 2020/2021
- Семестр(-ы):** 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.

10. Место учебной дисциплины в структуре ООП:дисциплина Б1.Б.28 Технологии обработки текста и звучащей речи входит в базовую часть ООП. Для изучения дисциплины необходимы знания, умения и компетенции, сформированные дисциплин: Б1.Б.15 Введение в теорию языка, Б1.Б.27 Общая и компьютерная лексикография, Б1.Б.26 Технологии корпусной лингвистики, Б1.Б.14 Информатика и основы программирования Б1.В.Од.2 Проектирование баз данных,

11. Планируемые результаты обучения по дисциплине (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников):

Компетенция		Планируемые результаты обучения
Код	Название	
ПК-5	Владение основными способами описания и формальной репрезентации денотативной, концептуальной, коммуникативной и прагматической информации, содержащейся в тексте на естественном языке	<p>Знать: методы описания денотативной, концептуальной, коммуникативной и прагматической информации</p> <p>Уметь: использовать лингвистически-ориентированные программные системы.</p> <p>Владеть: основами дисциплин, необходимых для формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур.</p>

12. Объем дисциплины в зачетных единицах/часах в соответствии с учебным планом — 23ЕТ / 72час.

Форма промежуточной аттестациизачет.

13. Виды учебной работы:

Вид учебной работы	Трудоемкость (часы)			
	Всего	По семестрам		
		№ сем.5	№ сем.

Аудиторные занятия	38	39		
в том числе: лекции				
практические	0	0		
лабораторные	38	38		
Самостоятельная работа	34	34		
Итого:	72	72		
Форма промежуточной аттестации (зачет)				

13.1 Содержание дисциплины:

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
1. Лекции		
Не предусмотрены учебным планом		
2. Практические занятия		
Не предусмотрены учебным планом		
3. Лабораторные занятия		
3.1	Основы организации вычислительных сетей. Сетевые стандарты и протоколы. Сетевые сервисы.	Лабораторная работа 1. 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.	Лабораторная работа 2. 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.	Лабораторная работа 3. Язык разметки HTML. Структура документа. Основные теги языка. Обязательные метки . Теги <html> и </html>,<head> и </head>, <title> и </title>, <body> и </body>, <H1> и </H1> — <H6> и </H6>,<P> и </P> Стилиевое оформление текста. Гиперссылки. Форматирование текста с помощью тега . Специальные символы,<HR> и & -последовательности. Горизонтальные линейки. Разделители. Непарные метки. Конец строки . Лабораторная работа 4. Списки – упорядоченные, неупорядоченные, вложенные. Нумерованные списки: и Нумерованные

		<p>списки: и Списки определений: <DL> и </DL>Списки, формы. Сложные формы. Изображения в документе. Фреймы. Тэги <META>.</p> <p>Лабораторная работа 5</p> <p>JavaScript/Jscript. Управляющие структуры – структура выбора if , структура выбора if/else, Структуры повторения while, for, do /while . Операторы break , continue. Операции присваивания, операции инкремента и декремента. Типы данных. Структура с множественным выбором switch. Логические операции.</p>
3.4	Язык гипертекстовой разметки XML для задач хранения и обработки электронных хранилищ документов.	<p>Лабораторная работа 6.</p> <p>Структура XML документа. Пролог и корневой элемент. Комментарии. Специальные символы. Пример XML файла.XML процессор. Использование инструкций по обратотке. Диррективы 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 *.css. Основные стандартные элементы для отображения шрифта и форматирования текста. Форма записи раздела CDATA. Понятие валидного XML документа.</p> <p>Лабораторная работа 7.</p> <p>DocumentsTypeDefinitions (DTD). Внутреннее и внешнеописаниеDTD.Схемы данных для построения XML документов. Область схемы данных. Описание элементов. Атрибуты элементов в схемах данных. Модель содержимого элемента. Группировка элементов. Закрытая и открытая модели описания содержимого элемента.</p>
3.5	Языки разметки на базе XML Язык разметки AIML	<p>Лабораторная работа 8.</p> <p>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.</p>
3.6	Концепции WEB2.0, WEB3.0, Семантический WEB	<p>Лабораторная работа 9.</p> <p>Семь технологических уровней, на которых базируется Семантическая сеть. ResourceDescriptionFramework (RDF) и RDF Schema, –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. DublinCoreMetadata. Язык описания семантики ресурсов - OWL, язык запросов к знаниям SPERQL.</p>
3.7	Онтологии и тезаурусы	<p>Лабораторная работа 10.</p> <p>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. Инструментальные средства проектирования онтологий. Примеры крупных</p>

		онтологических проектов.
3.8	Многоагентные интеллектуальные системы для задач семантического поиска	Лабораторная работа 11. 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)				
		Лекции	Практические	Лабораторные	Самостоятельная 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 HTML для задач хранения и обработки электронных хранилищ документов.			6	6	12
5	Языки разметки на базе XML Язык разметки AIML			6	4	10
6	Концепции WEB2.0, WEB3.0, Семантический WEB			6	4	10
7	Онтологии и тезаурусы			6	6	12
8	Многоагентные интеллектуальные системы для задач семантического поиска			2	2	4
				38	34	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 рассылаются по электронной почте).

2) Для максимального усвоения дисциплины рекомендуется проведение письменного опроса (тестирование, решение задач) студентов по материалам лекций и лабораторных 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.

13. Перечень основной и дополнительной литературы, ресурсов Интернет, необходимых для освоения дисциплины:

а) основная литература:

№ п/п	Источник
1	Маннинг, Кристофер Д. Введение в информационный поиск = Introduction to Information retrieval / Кристофер Д. Маннинг, Прабхакар Рагхаван, Хайнрих Шютце ; [пер. с англ. Д.А. Ключина] .— М. ; СПб. ; Киев : Вильямс, 2011 .— 520 с.
2	Онтологии и тезаурусы: модели, инструменты, приложения : учебное пособие / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. - Москва : Интернет-Университет Информационных Технологий, 2009. - 173 с. : ил., табл., схем. - (Основы информационных технологий). - ISBN 978-5-9963-0007-5 ; То же [Электронный ресурс]. - URL: http://biblioclub.ru/index.php?page=book&id=233056
3	Информационные технологии. HTML и XHTML : учебное пособие / А.И. Костюк, С.М. Гушанский, М.Ю. Поленов, Б.В. Катаев ; Министерство образования и науки Российской Федерации, Южный федеральный университет, Инженерно-технологическая академия. - Таганрог : Издательство Южного федерального университета, 2015. - 131 с. : ил. - Библиогр. в кн. - ISBN 978-5-9275-1329-1 ; То же [Электронный ресурс]. - URL: http://biblioclub.ru/index.php?page=book&id=461923
4	Жданов, Александр Аркадьевич. Автономный искусственный интеллект / А.А. Жданов .— 2-е изд. — М. : БИНОМ. Лабораторий знаний, 2009 .— 359 с.

б) дополнительная литература:

№ п/п	Источник
1	Towards the Semantic Web: Ontology-driven Knowledge Management. — John Wiley & Sons, 2003.
2	Toby Segaran, Colin Evans, Jamie Taylor Programming the Semantic Web. — 2009.
3	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)», David Amerland, 240 pages, 10 July 2013
4	Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph Foundations of Semantic Web Technologies Chapman & Hall/CRC, 2009, 455 pages
5	Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL» by Dean Allemang and James Hendler, Publisher: Morgan Kaufmann, 352 pages, 2008.
6	Google Analytics для профессионалов», Брайан Клифтон, 3-е издание. — М.: «Диалектика», 2013, 608 стр.
7	Programming the Semantic Web» by Toby Segaran, Colin Evans and Jamie Taylor, - O'Reilly Media , 2009, 298 pages.
8	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.
9	Леонтьева Н.Н. Автоматическое понимание текстов. М., 2006

в) базы данных, информационно-справочные и поисковые системы:

№ п/п	Источник
-------	----------

1	Электронный каталог Научной библиотеки Воронежского государственного университета. – (http // www.lib.vsu.ru/).
2	Образовательный портал «Электронный университет ВГУ». – (https://edu.vsu.ru/)
3	ЭБС «Издательства «Лань», Договор №3010-06/71-14 от 25.11.2014, ЭБС «Университетская библиотека online», Договор №3010-06/70-14 от 25.11.14, Национальный цифровой ресурс «РУКОНТ», Договор №ДС-208 от 01.02.2012
4	Международная конференция по компьютерной лингвистике. http://www.dialog-21.ru/
5	Лаборатория компьютерной лингвистики Института проблем передачи информации РАН. http://proling.iitp.ru/
6	Лаборатория общей компьютерной лексикологии и лексикографии МГУ. http://www.philol.msu.ru/~lex/library.htm

16. Перечень учебно-методического обеспечения для самостоятельной работы:

№ п/п	Источник
1	Боярский К. К. Введение в компьютерную лингвистику. Учебное пособие. – СПб: НИУ ИТМО, 2013. – 72 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 подписок «Imagine», ежегодные сублицензионные договоры № 56035/ВРН3739 и № 56036/ВРН3739 от 07.10.2016.

3) Персер русского языка ТОМИТА (Свободно-распространяемое ПО)

4) Язык программирования Python, IDE Pycharm.

4) ПО Редактор онтологий и фреймворк для построения баз знаний Protege. Свободно-распространяемое ПО.

18. Материально-техническое обеспечение дисциплины

/ауд. 12/ - компьютерный класс: Компьютер ArbyteTempo/АОС (12 шт.), Проектор Benq MW523 (1 шт.), Сканер CanonCanoscanLiDE 120 (5 шт.) Экран проекционный (1 шт.) /ауд. 14/Проектор Benq MW523 (1 шт.) Экран проекционный (1 шт.) Компьютер Asus H81m-Plus (11 шт.)	г. Воронеж, пл. Ленина 10, ауд. 12, 14
--	--

19. Фонд оценочных средств:

19.1 Перечень компетенций с указанием этапов формирования и планируемых результатов обучения

Код и содержание компетенции (или ее части)	Планируемые результаты обучения (показатели достижения заданного уровня освоения компетенции посредством формирования знаний, умений, навыков)	Этапы формирования компетенции (разделы (темы) дисциплины или модуля и их наименование)	ФОС* (средства оценивания)

ПК-5 Владение основными способами описания и формальной репрезентации денотативной, концептуальной, коммуникативной и прагматической информации, содержащейся в тексте на естественном языке	знать: математические и алгоритмические методы обработки текстовой информации, принципы организации и этапы разработки лингвистически ориентированных программных продуктов.	Разделы 1,2,10,11	Контрольная работа по соответствующим разделам.
	уметь: прилагать полученные теоретические и практические знания к задачам обработки текста и речи	Разделы 1,2,10,11	Лабораторные работы 1,2,10,11
	владеть: навыками использования, интеграции и разработки программных систем обработки лингвистической информации.	Разделы 1,2,10,11	Лабораторные работы 1,2,10,11
Промежуточная аттестация			Комплект КИМ

19.2. Описание критериев и шкалы оценивания компетенций (результатов обучения) при промежуточной аттестации

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

знать: математические и алгоритмические методы обработки текстовой информации, принципы организации и этапы разработки лингвистически ориентированных программных продуктов.

уметь: прилагать полученные теоретические и практические знания к задачам обработки текста и речи

владеть: навыками использования, интеграции и разработки программных систем обработки лингвистической информации.

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, критериев 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

19.3 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения образовательной программы

19.3.1 Примерный перечень применяемых оценочных средств

№ п/п	Наименование оценочного средства	Представление оценочного средства в фонде	Критерии оценки
1	2	3	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 работа	Содержит 9 лабораторных заданий, предусматривающих освоение программных систем обработки текста и речи.	При успешно выполнении 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 задания вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции.	Шкалы оценивания приведены в разделе 19.2

19.3.2. 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 обратотке. Диррективы 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. DocumentsTypeDefinitions (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 – GAITOБОТ.
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, Sowaontology.

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 - Textmining.
37. Направления применения онтологического моделирования при проектировании информационных систем. Примеры проектов, использующих интеграцию с онтологиями и обработкой семантики.

19.3.3. Пример задания для выполнения лабораторной работы

Лабораторная работа №

Цель лабораторной работы.

Создание классов онтологической модели в Protégé.

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 (<http://protege.stanford.edu>).

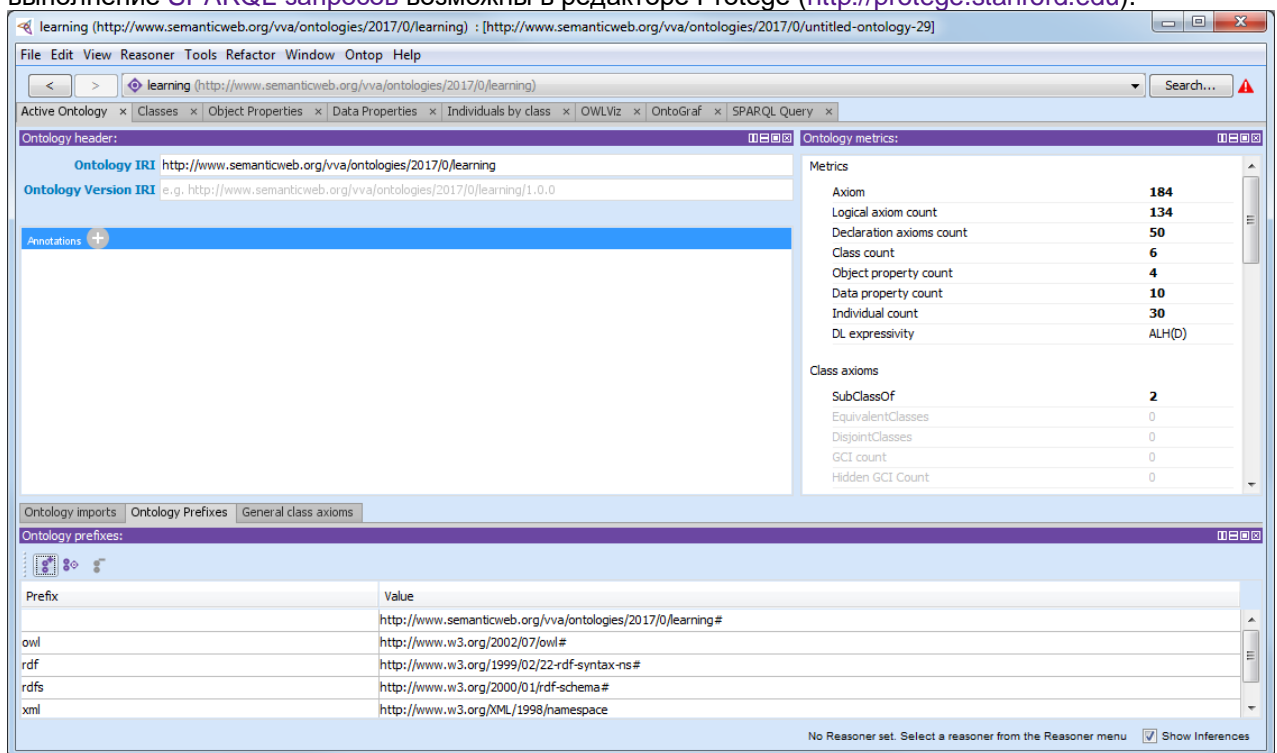


Рис.1. Редактор онтологий Protégé

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

Назначение

вкладок:

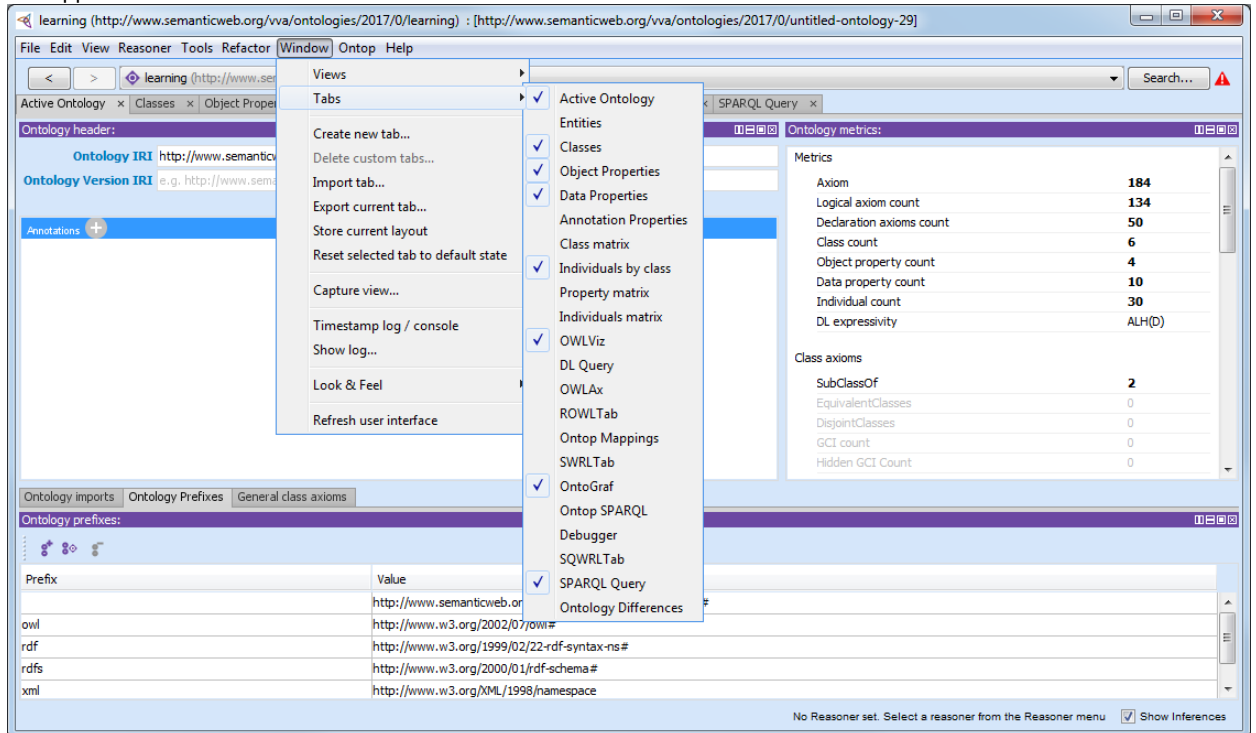


Рис.2. Выбор отображаемых вкладок

- ActiveOntology – отображение IRI, общих характеристик (количества аксиом, классов, свойств, объектов и т.п.) и аннотации онтологии;
- Classes – просмотр и редактирование классов;
- ObjectProperties – просмотр и редактирование свойств-отношений между индивидами;
- DataProperties – просмотр и редактирование свойств-данных индивидов;
- Individualsbyclass – просмотр и редактирование экземпляров классов (индивидов);
- 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.

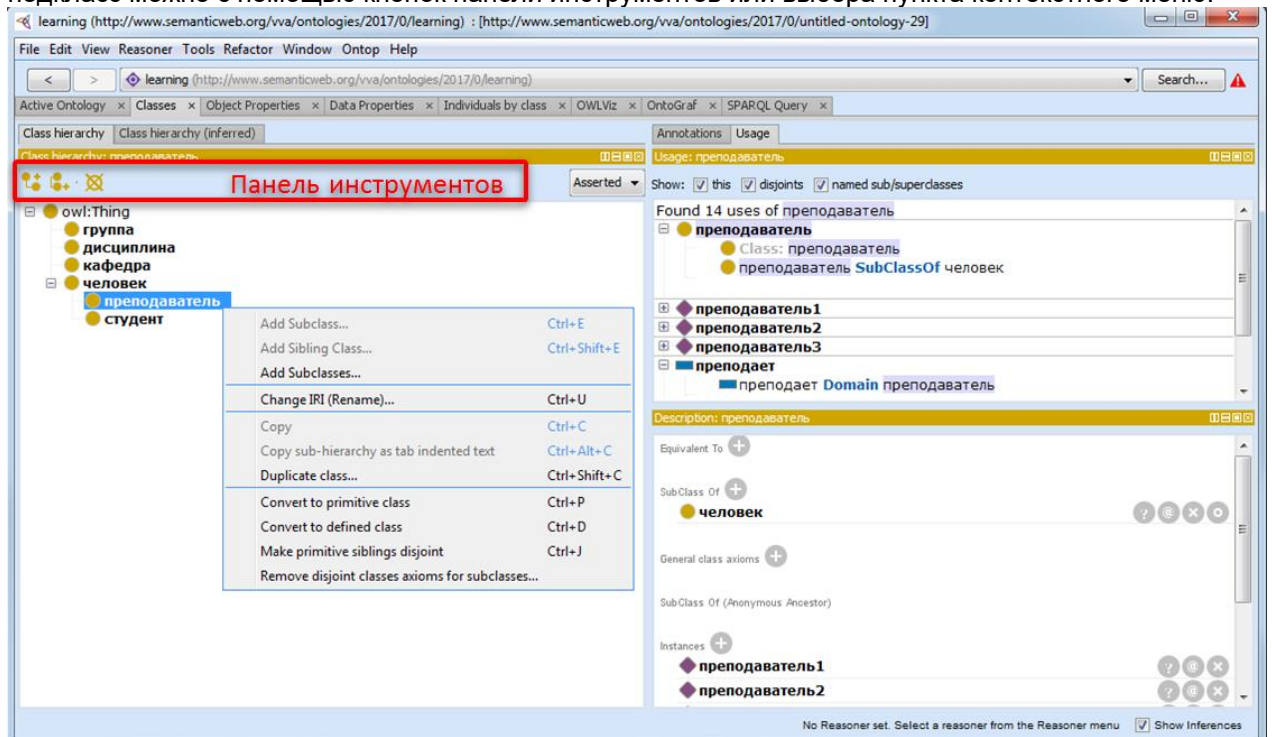


Рис.3. Вкладка «Classes»

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. SiblingClass).

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. EquivalentTo) или невозможность принадлежности экземпляров класса другим классам (англ. DisjointWith) – запрет множественного наследования.

Свойства классов и их экземпляров (предикаты RDF-троек) делятся на два вида:

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

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

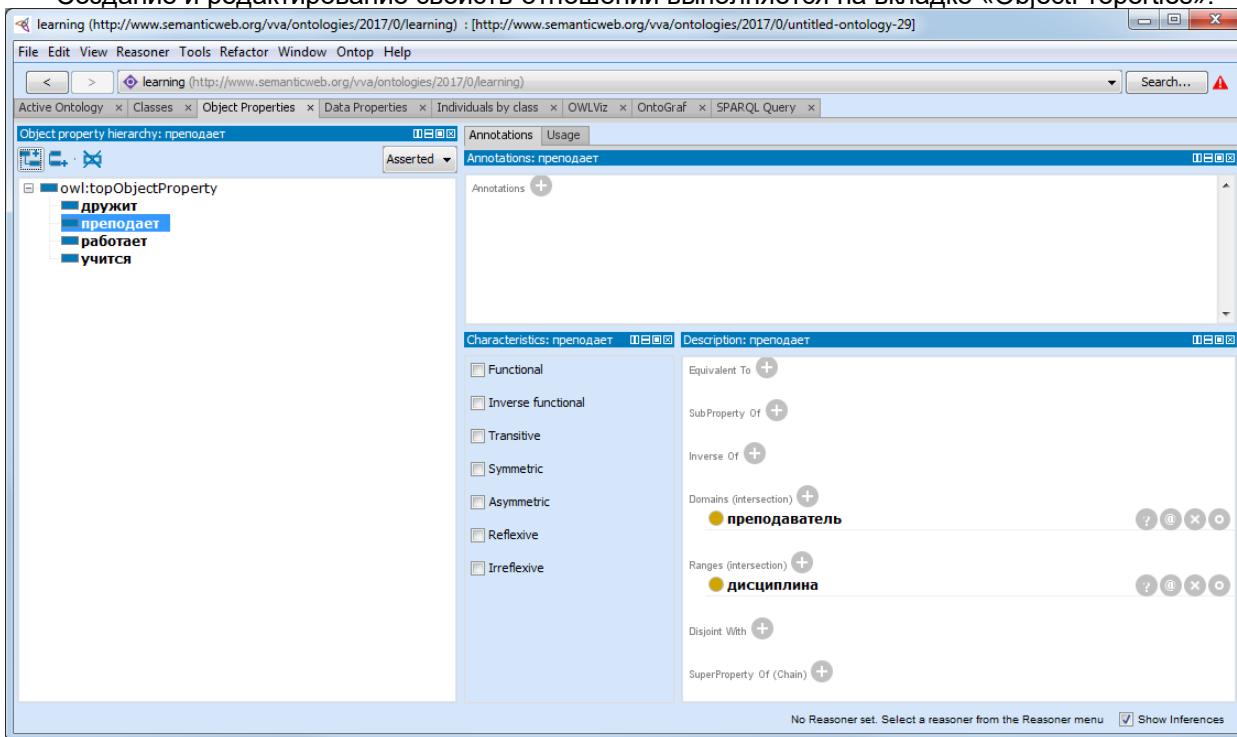


Рис.4. Вкладка «ObjectProperties»

Домен (англ. Domain) указывает, экземпляры каких классов в RDF-тройке 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 «DataProperties».

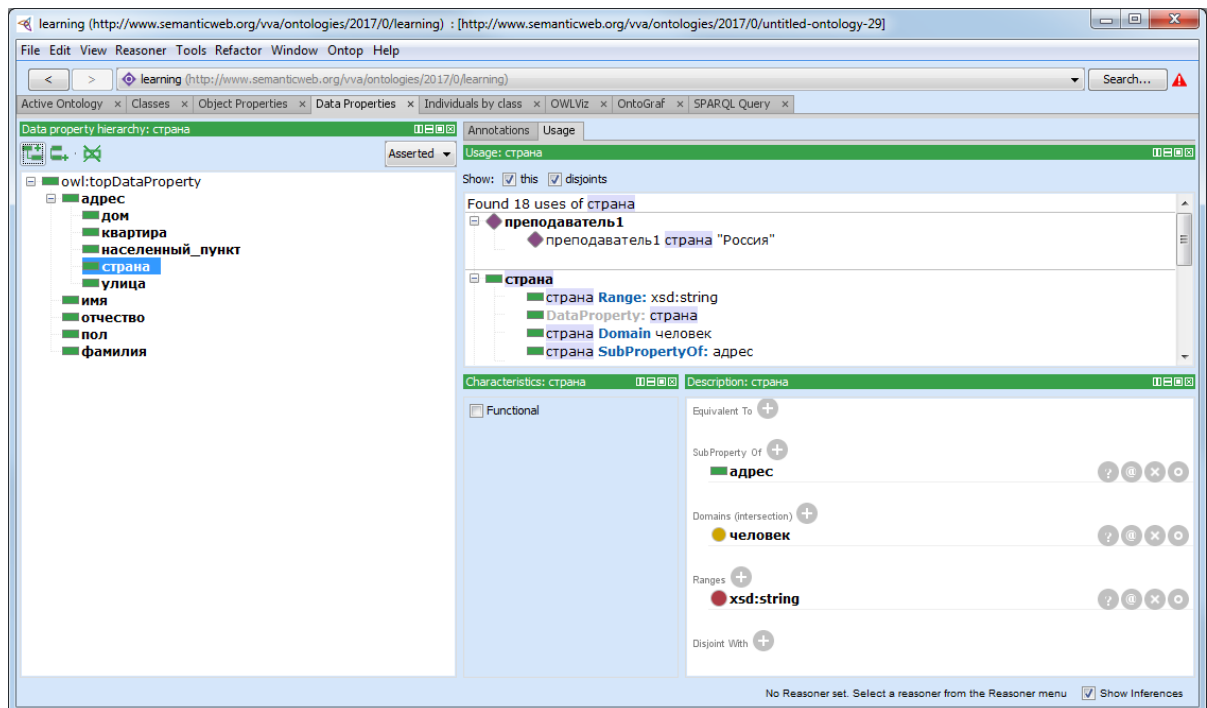


Рис.5. Вкладка «DataProperties»

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.

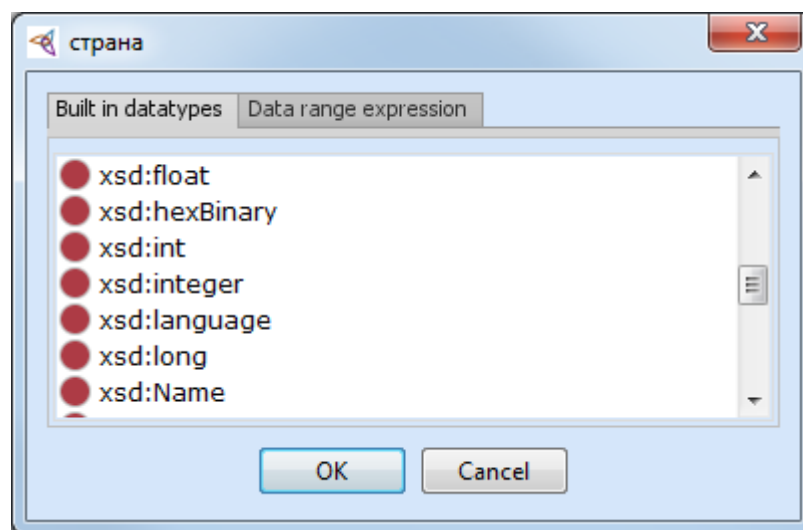
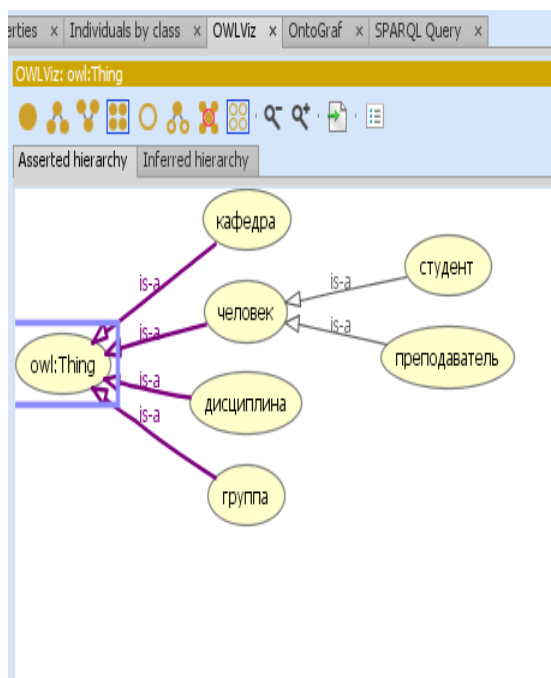
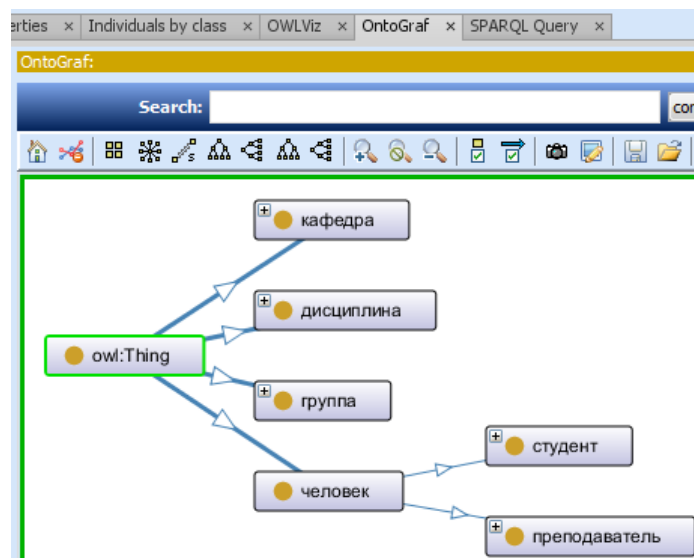


Рис.6. Окно выбора типа данных

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.

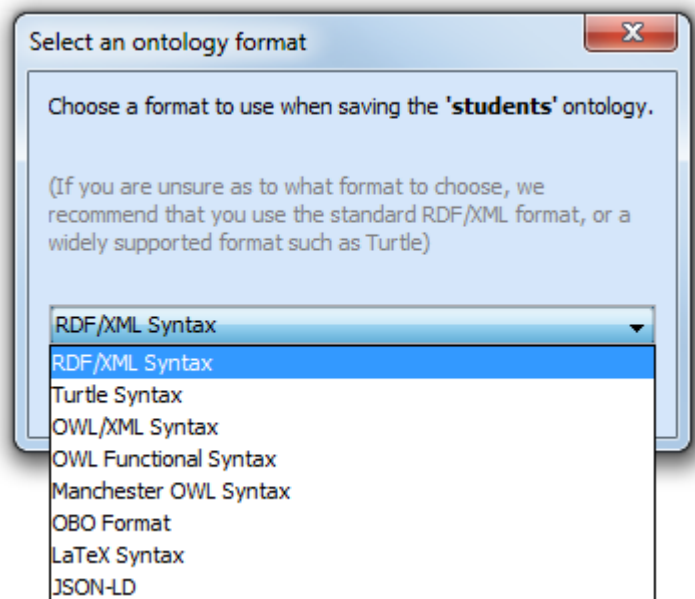


Рис.8. Выбор формата хранения онтологии

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 свойственных.

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 (OWLViz или OntoGraf);
- текст онтологии, сохраненной в формате «TurtleSyntax»;

19.3.4. Пример контрольно-измерительного материала

УТВЕРЖДАЮ

Заведующий кафедрой технологий обработки и защиты информации

_____ А.А. Сирота

__._.2017

Направление подготовки / специальность

45.03.03 Фундаментальная и прикладная лингвистика

Дисциплина Б1.В.ОД.4 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.В.Гаршина

19.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

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. Критерии оценивания приведены выше в таблице раздела 19.2.