

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИИ



Федеральное государственное бюджетное образовательное учреждение высшего образования

«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГУМАНИТАРНЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «РГГУ»)

Институт лингвистики

УНЦ компьютерной лингвистики

Рабочая программа дисциплины

«Автоматический семантический анализ»

Направление подготовки 45.04.03 Фундаментальная и прикладная лингвистика

Магистерская программа: Фундаментальная и компьютерная лингвистика

Квалификация выпускника: магистр

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2019

Автоматический семантический анализ

Рабочая программа дисциплины

Составители:

к.ф.н Л.Л.Иомдин, д.ф.н. И.М.Богуславский

Ответственный редактор:

д. филол. н., профессор В.И.Подлеская

УТВЕРЖДЕНО

Протокол заседания УНЦ компьютерной
лингвистики

№ 1 от «28» августа 2019г.

Оглавление

1. Пояснительная записка

- 1.1. Предмет
- 1.2. Цель и задачи дисциплины
- 1.3. Формируемые компетенции и результаты освоения дисциплины
- 1.4. Место дисциплины в структуре образовательной программы

2. Структура дисциплины

3. Содержание дисциплины

4. Образовательные технологии

5. Оценка планируемых результатов обучения

- 5.1. Система оценивания
- 5.2. Критерии выставления оценок
- 5.3. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

6. Учебно-методическое и информационное обеспечение дисциплины

- 6.1. Список литературы

7. Материально-техническое обеспечение дисциплины

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

9. Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений

1. Пояснительная записка

1.1 Предмет

Предметом данной дисциплины является один из важнейших разделов компьютерной лингвистики – автоматический анализ текста. Семантический анализ текста – это самый глубокий уровень анализа, при котором для текста строится его смысловое представление. Если системы, осуществляющие другие, более поверхностные, уровни анализа текста (в частности, морфологический и синтаксический анализ) достигли высокой степени развития и единообразия (в конечном счете для представления синтаксической структуры предложения используются всего два основных стандарта), то семантический анализ все еще находится в стадии становления, и стандартов семантического представления значительно больше.

Данный курс даст магистрантам общее представление о наиболее распространенных типах представления семантической структуры в автоматического семантическом анализе и об алгоритмах семантического анализа. Среди них такие системы семантического анализа, как StanfordCoreNLP, Boxer, WASP, KRISP, Compreno, OntoEtap.

Курс призван дать магистрантам представление о принципах и методах автоматизации семантического анализа текста в современной компьютерной лингвистике. Курс позволит студентам ознакомиться с основными достижениями компьютерной лингвистики в области автоматизации построения семантической структуры текста для целей информационного поиска, в частности, для целей создания вопросно-ответных систем общения человека с компьютером на естественном языке.

Данный курс содержательно и системно связан с курсом «Лингвистическая и онтологическая семантика».

1.2 Цель и задачи курса

Курс направлен на решение следующих задач:

1. ознакомить магистрантов с важнейшими принципами семантического анализа текста;
2. дать представление об основных системах автоматического семантического анализа и их месте в современной компьютерной лингвистике;
3. дать представление о связи автоматического построения семантической структуры с информационными технологиями;

1.3 Компетенции обучающегося, формируемые в результате освоения дисциплины

способностью свободно вести коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке для решения задач профессиональной деятельности (ОПК-1);

готовностью руководить коллективом в сфере своей профессиональной деятельности, толерантно воспринимая социальные, этнические, конфессиональные и культурные различия (ОПК-2);

владением знаниями об истории лингвистических учений и современного состояния лингвистической науки (ОПК-3);

способностью анализировать, сопоставлять и критически оценивать различные лингвистические направления, теории и гипотезы (ОПК-5);

способностью проводить самостоятельные исследования и получать новые научные

результаты в области теории языка, лингвистики конкретных языков, прикладной и компьютерной лингвистики (ПК-1);

владением навыками перевода научной литературы по лингвистике и смежным дисциплинам с иностранных языков на государственный язык Российской Федерации и с государственного языка Российской Федерации на иностранный язык со снабжением ее необходимым редакторским и издательским комментарием и научным аппаратом (ПК-11).

В результате освоения дисциплины обучающийся должен демонстрировать следующие результаты образования:

знать место семантики в общей системе лингвистики и место семантического анализа в общем подходе к автоматической обработке текстов;

уметь совершенствовать и развивать свой интеллектуальный уровень, адаптироваться к изменению профиля деятельности; использовать фундаментальные знания по компьютерной лингвистике в целом и компьютерной семантике в частности в сфере профессиональной деятельности;

владеть основами методологии научного познания при изучении различных аспектов человеческого языка.

1.4. Место дисциплины в структуре образовательной программы

Дисциплина (модуль) «Автоматический семантический анализ» является дисциплиной по выбору *вариативной части* цикла дисциплин ООП ВПО (магистратуры) по направлению подготовки «Фундаментальная и прикладная лингвистика. Фундаментальная и компьютерная лингвистика» и адресована студентам 2 курса (3 семестр). Дисциплина (модуль) реализуется УНЦ компьютерной лингвистики Института Лингвистики.

Программой дисциплины (модуля) предусмотрены следующие виды контроля: текущий контроль успеваемости в форме: *подготовка к семинарским занятиям, реферат, коллоквиум по контрольным вопросам с выполнением практического задания*; промежуточная аттестация в форме: *зачет*.

Общая трудоемкость освоения дисциплины (модуля) составляет 2 зачетные единицы, 72 часа.

Программой дисциплины (модуля) предусмотрены: практические занятия – 20 часов; самостоятельная работа студента – 52 часа.

2. Структура дисциплины

Общая трудоемкость дисциплины составляет 2 зачетные единицы, 72 часа (6 лекционных часов, 14 семинарских часов, 52 часа самостоятельной работы).

№ п/п	Раздел Дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				лекции	семинары	самостоятельная работа		
1	Введение. Компьютерная лингвистика и автоматический анализ	3	1	1	1	2		

	текста. Уровни анализа текста. Семантическая структура предложения и семантическое представление. Поверхностная и глубинная семантика						
2.	Варианты семантической структуры предложения. Семантическое представление модели Смысл-Текст.	3	1-2		2	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
3.	Варианты семантической структуры предложения. Структура RDF.	3	2		1	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
4	Лингвистические ресурсы, используемые в построении семантических анализаторов. Семантический словарь. Структура аналитического толкования.	3	3		2	6	Выполнение практического задания: составление образцов словарных статей предикатных слов с упором на семантику
5	Лингвистические ресурсы, используемые в построении семантических анализаторов. Онтология. База знаний и фактов.	3	4		2	6	
6	Система семантического анализа StanfordCoreNLP	3	5		2	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
7.	Система семантического анализа Boxer.	3	6		1	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
8.	Системы семантического анализа Wasp и Crisp.	3	7		2	2	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
9.	Система семантического анализа Compreno.	3	8		2	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
10.	Система семантического анализа OntoEtap. Построение	3	9		3	6	Выполнение практического задания:

	семантической структуры предложения						построение несложных правил семантического анализа в формализме и алгоритме ЭТАП
11.	Система семантического анализа OntoEtap. Правила разрешения анафоры и установления кореферентности.	3	10		2	4	Выполнение практического задания: построение несложных правил разрешения анафоры и установления кореферентности в формализме и алгоритме ЭТАП
12	Зачет	3				6	Реферат; коллоквиум по контрольным вопросам с выполнением практического задания
	Итого:				20	52	

3. Содержание дисциплины

1. Введение. Компьютерная лингвистика и автоматический анализ текста. Уровни анализа текста. Семантическая структура предложения и семантическое представление. Поверхностная и глубинная семантика.
2. Варианты семантической структуры предложения. Семантическое представление модели Смысл-Текст.
3. Варианты семантической структуры предложения. Структура RDF.
4. Лингвистические ресурсы, используемые в построении семантических анализаторов. Семантический словарь. Структура аналитического толкования.
5. Лингвистические ресурсы, используемые в построении семантических анализаторов. Онтология. База знаний и фактов.
6. Система семантического анализа StanfordCoreNLP.
7. Система семантического анализа Boxer.
8. Системы семантического анализа WASP и CRISP.
9. Система семантического анализа Compreno.
10. Система семантического анализа OntoETAP. Создание семантической структуры. Разрешение анафоры.

4. Образовательные технологии

В соответствии с требованиями ФГОС по направлению 45.04.03 «Фундаментальная и прикладная лингвистика» и с учетом специфики магистерская программы «Фундаментальная и компьютерная лингвистика» занятия лекционного типа составляют не более 20% аудиторных занятий, а удельный вес занятий, проводимых в интерактивных формах, составляют не менее 40% аудиторных занятий. Интерактивные формы обучения в данном курсе предполагают:

- систематическое использование компьютерных презентаций;
- онлайн демонстрации работы с лингвистическими базами данных и энциклопедическими интернет-ресурсами;
- использование лингвистических корпусов в режиме онлайн.

5. Оценка планируемых результатов обучения

5.1. Система оценивания

При выставлении оценки в ведомость и в зачетную книжку преподаватель должен указать результат в соответствии с традиционной шкалой оценок и со шкалой оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

Распределение баллов по видам учебной деятельности таково:

- посещение семинарских занятий – до 8 баллов,
- уровень активности студента при подготовке к занятиям (конспектирование специальной литературы, готовность отвечать на вопросы по анализу кейсов, активное участие в дискуссиях, коллоквиумах и мозговом штурме и проч.) и во время проведения занятий (участие в обсуждениях и выполнении коллективных заданий) – всего до 32 баллов,
- качество выполнения контрольной работы (текущая аттестация) – до 20 баллов,
- успешность выполнения итогового творческого задания – до 40 баллов.

Оценка «зачтено» выставляется, если студент набрал в сумме не менее 50 баллов. Магистрант, не набравший в сумме 50 баллов, сдает зачет по всему курсу и предъявляет преподавателю собственноручно написанные конспекты специальной литературы и выполненные домашние задания ко всем семинарам.

5.2. Критерии выставления оценок

При выставлении оценки преподаватель ориентируется на следующие содержательные критерии.

Количество баллов	Критерии оценки
95–100 (A)	Оценка выставляется с учетом текущей и промежуточной аттестации. Теоретическое содержание дисциплины освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные рабочей программой дисциплины учебные задания выполнены, качество

Количество баллов	Критерии оценки
	<p>их выполнения оценено числом баллов, близким к максимальному.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне «высокий».</p>
83–94 (B)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, почти все задания, предусмотренные рабочей программой дисциплины, выполнены, качество выполнения большинства из них оценено числом баллов, близким к максимальному.</p> <p>Обучающийся адекватно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Достаточно свободно ориентируется в учебной и профессиональной литературе.</p> <p>Почти все компетенции, закреплённые за дисциплиной, сформированы на уровне «высокий».</p>
68–82 (C)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные рабочей программой дисциплины учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне «хороший».</p>
56–67 (D)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных рабочей программой дисциплины учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки.</p>

Количество баллов	Критерии оценки
	<p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
50–55 (E)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины освоено частично, некоторые практические навыки работы не сформированы, многие предусмотренные рабочей программой дисциплины учебные задания не выполнены, либо качество выполнения некоторых из них оценено числом баллов, близким к минимальному.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне «достаточный».</p>
21–49 (FX)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины освоено частично, необходимые практические навыки работы не сформированы, большинство предусмотренных рабочей программой дисциплины учебных заданий не выполнено, либо качество их выполнения оценено числом баллов, близким к минимальному; при дополнительной самостоятельной работе над материалом курса возможно повышение качества выполнения учебных заданий.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p> <p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.</p>
0–20 (F)	<p>Оценка выставляется с учетом текущей и промежуточной аттестации.</p> <p>Теоретическое содержание дисциплины не освоено. Необходимые практические навыки работы не сформированы, все предусмотренные рабочей программой дисциплины учебные задания выполнены с грубыми ошибками. Дополнительная самостоятельная работа над материалом дисциплины не приведет</p>

Количество баллов	Критерии оценки
	<p>к какому-либо значимому повышению качества выполнения учебных заданий.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p> <p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.</p>

5.3. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Ниже приводятся контрольные вопросы по данному курсу.

В качестве домашних заданий предлагаются задания следующих типов

Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация.

Выполнение практического задания: составление образцов словарных статей предикатных слов с упором на семантику.

Выполнение практического задания: построение несложных правил семантического анализа в формализме и алгоритме ЭТАП.

Выполнение практического задания: построение несложных правил разрешения анафоры и установления кореферентности в формализме и алгоритме ЭТАП.

Зачет ориентирован на следующие контрольные вопросы

1. Уровни представления языковых выражений.
2. Поверхностная и глубинная семантика.
3. Семантическая структура предложения в русле модели «Смысл – Текст».
4. RDF-структуры.
5. Лингвистические ресурсы, используемые в построении семантических анализаторов.
Семантический словарь. Структура аналитического толкования.
6. Лингвистические ресурсы, используемые в построении семантических анализаторов.
Онтология. База знаний и фактов.
7. Система семантического анализа StanfordCoreNLP.
8. Система семантического анализа Boxer.
9. Система семантического анализа WASP.
10. Система семантического анализа CRISP.
11. Система семантического анализа Compreno.
12. Система семантического анализа OntoETAP: семантический анализ текста с использованием онтологии.
13. Кореференция и анафора в автоматическом семантическом анализе.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список литературы

Основная литература

1. Гращенков Павел Валерьевич. Автоматический анализ русских сложных слов / П. В. Гращенков // Вестник Московского университета. Сер. 9, Филология. - 2016. - № 6. - С. 119-132. - Библиогр. в примеч.
2. Шелманов А.О. Исследование методов автоматического анализа текстов и разработка интегрированной системы семантико-синтаксического анализа. – Кандидатская диссертация. – Москва, 2015.
http://www.ipiran.ru/announce/avto_Shelmanov.pdf
3. Igor Boguslavsky. Semantic descriptions for a text understanding system. Dialog 2017.
4. Sergei Nirenburg and Marjorie McShane. Agents modeling agents: Incorporating ethics-related reasoning (2012). Proceedings of the symposium Moral Cognition and Theory of Mind at the AISB/IACAP World Congress 2012, Birmingham, UK.
5. Техническое описание семантического анализатора Abbyy Compreno
<https://www.abbyy.com/ru-ru/infoextractor/compreno/>

Рекомендованная литература

1. Анисимович К.В. и др. Синтаксический и семантический парсер, основанный на лингвистических технологиях Abbyy Compreno. Диалог 2012.
2. Богуславский И.М., Диконов В.Г., Иомдин Л.Л. и др. Семантический анализ и ответы на вопросы: система в стадии разработки) // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27–30 мая 2015 г.). М.: Изд-во РГГУ, 2015. Вып. 14(21). Т. 1. С. 62-79
3. О. Невзорова, В.Невзоров, Н. Пяткин. Методы решения лингвистических задач на основе онтологий. Institute of Information Theories and Applications FOI ITHEA, 2008.
4. Ф.И. Андон, И.Ю. Гришанова, В.А. Резниченко. Semantic Web как новая модель информационного пространства Интернет. Киев, 2009. (Интернет-версия: <http://shcherbak.net/semantic-web-kak-novaya-model-informacionnogo-prostranstva-internet/>)
5. E.Mueller. Common sense reasoning. Elsevier, Morgan Kaufmann Publishers. 2006.
6. Igor Boguslavsky. Semantic descriptions for a text understanding system. Dialog 2017.
7. Johan Bos. Open-Domain Semantic Parsing with Boxer. NODALIDA 2015 (20th Nordic Conference on Computational Linguistics): 301-304
8. Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60
9. Ontology and the Lexicon. A Natural Language Processing Perspective. Edited by: Churen Huang, N.Calzolari, A.Gangemi, A.Lenci, A.Oltramari, L.Prevot. Series: Studies in Natural Language Processing. Cambridge University Press, 2010.
10. Philipp Cimian, Christina Unger, John McCrae. Ontology-Based Interpretation of Natural Language (Synthesis Lectures on Human Language Technologies). 2014/ pp. 895-900, Vancouver, Canada, July 2007

11. Resource Description Framework. <https://www.w3.org/RDF/>
12. Rohit J. Kate, Raymond J. Mooney. Learning Language Semantics from Ambiguous Supervision. Proceedings of the 22nd Conference on Artificial Intelligence (AAAI-07).
13. Sergei Nirenburg and Marjorie McShane. Agents modeling agents: Incorporating ethics-related reasoning (2012). Proceedings of the symposium Moral Cognition and Theory of Mind at the AISB/IACAP World Congress 2012, Birmingham, UK.
14. Sergei Nirenburg, Victor Raskin. Ontological Semantics. MIT Press, 2004.
15. Yuk Wah Wong. Learning for semantic parsing and natural language generation using statistical machine translation techniques. 2007.

7. Материально-техническое обеспечение дисциплины

Занятия по курсу «Автоматический семантический анализ» можно проводить с максимальной эффективностью, если проводить их в компьютерном классе с доступом в Интернет, проектором и экраном для презентаций. Необходимо также наличие доски, чтобы преподаватель мог разбирать примеры по ходу объяснения и записывать задания.

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

При необходимости рабочая программа дисциплины может быть адаптирована для обеспечения образовательного процесса лицам с ограниченными возможностями здоровья, в том числе для дистанционного обучения. Для этого от студента требуется представить заключение психолого-медико-педагогической комиссии (ПМПК) и личное заявление (заявление законного представителя).

В заключении ПМПК должно быть прописано:

- рекомендуемая учебная нагрузка на обучающегося (количество дней в неделю, часов в день);
- оборудование технических условий (при необходимости);
- сопровождение и (или) присутствие родителей (законных представителей) во время учебного процесса (при необходимости);
- организация психолого-педагогического сопровождение обучающегося с указанием специалистов и допустимой нагрузки (количества часов в неделю).

Для осуществления процедур текущего контроля успеваемости и промежуточной аттестации обучающихся, при необходимости могут быть созданы фонды оценочных средств, адаптированные для лиц с ограниченными возможностями здоровья и позволяющие оценить достижение ими запланированных в основной образовательной программе результатов обучения и уровень сформированности всех компетенций, заявленных в образовательной программе.

Форма проведения текущей и итоговой аттестации для лиц с ограниченными возможностями здоровья устанавливается с учетом индивидуальных психофизических особенностей (устно, письменно (на бумаге, на компьютере), в форме тестирования и т.п.). При необходимости студенту предоставляется дополнительное время для подготовки ответа на зачете или экзамене.

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;

- письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
- обеспечивается индивидуальное равномерное освещение не менее 300 люкс;
- для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
- письменные задания оформляются увеличенным шрифтом;
- экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;
 - в форме аудиофайла.
- для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
 - устройством для сканирования и чтения с камерой SARA CE;
 - дисплеем Брайля PAC Mate 20;
 - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
 - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемые эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

9. Приложения

Приложение 1. Аннотация дисциплины

Дисциплина (модуль) «Автоматический семантический анализ» является дисциплиной по выбору *вариативной части* цикла дисциплин ООП ВПО (магистратуры) по направлению подготовки «Фундаментальная и прикладная лингвистика. Фундаментальная и компьютерная лингвистика» и адресована студентам 2 курса (3 семестр). Дисциплина (модуль) реализуется УНЦ компьютерной лингвистики Института Лингвистики.

Курс призван дать магистрантам представление о принципах и методах автоматизации семантического анализа текста в современной компьютерной лингвистике. Курс позволит студентам ознакомиться с основными достижениями компьютерной лингвистики в области автоматизации построения семантической структуры текста для целей информационного поиска, в частности, для целей создания вопросно-ответных систем общения человека с компьютером на естественном языке.

Данный курс содержательно и системно связан с курсом «Лингвистическая и онтологическая семантика».

Предметом данной дисциплины является один из важнейших разделов компьютерной лингвистики – автоматический анализ текста. Семантический анализ текста – это самый глубокий уровень анализа, при котором для текста строится его смысловое представление. Если системы, осуществляющие другие, более поверхностные, уровни анализа текста (в частности, морфологический и синтаксический анализ) достигли высокой ступени развития и единообразия (в конечном счете для представления синтаксической структуры предложения используются всего два основных стандарта), то семантический анализ все еще находится в стадии становления, и стандартов семантического представления значительно больше.

Данный курс даст магистрантам общее представление о наиболее распространенных типах представления семантической структуры в автоматического семантическом анализе и об алгоритмах семантического анализа. Среди них такие системы семантического анализа, как StanfordCoreNLP, Boxer, WASP, KRISP, Compreno, OntoEtap.

Курс направлен на решение *следующих задач*:

- ознакомить магистрантов с важнейшими принципами семантического анализа текста;
- дать представление об основных системах автоматического семантического анализа и их месте в современной компьютерной лингвистике;
- дать представление о связи автоматического построения семантической структуры с информационными технологиями.

Компетенции обучающегося, формируемые в результате освоения дисциплины:

способностью свободно вести коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке для решения задач профессиональной деятельности (ОПК-1);

готовностью руководить коллективом в сфере своей профессиональной деятельности, толерантно воспринимая социальные, этнические, конфессиональные и

культурные различия (ОПК-2);

владением знаниями об истории лингвистических учений и современного состояния лингвистической науки (ОПК-3);

способностью анализировать, сопоставлять и критически оценивать различные лингвистические направления, теории и гипотезы (ОПК-5);

способностью проводить самостоятельные исследования и получать новые научные результаты в области теории языка, лингвистики конкретных языков, прикладной и компьютерной лингвистики (ПК-1);

владением навыками перевода научной литературы по лингвистике и смежным дисциплинам с иностранных языков на государственный язык Российской Федерации и с государственного языка Российской Федерации на иностранный язык со снабжением ее необходимым редакторским и издательским комментарием и научным аппаратом (ПК-11).

В результате освоения дисциплины обучающийся должен демонстрировать следующие *результаты образования*:

знать место семантики в общей системе лингвистики и место семантического анализа в общем подходе к автоматической обработке текстов;

уметь совершенствовать и развивать свой интеллектуальный уровень, адаптироваться к изменению профиля деятельности; использовать фундаментальные знания по компьютерной лингвистике в целом и компьютерной семантике в частности в сфере профессиональной деятельности;

владеть основами методологии научного познания при изучении различных аспектов человеческого языка.

Дисциплина (модуль) «Автоматический семантический анализ» является дисциплиной по выбору *вариативной части* цикла дисциплин ООП ВПО (магистратуры) по направлению подготовки «Фундаментальная и прикладная лингвистика. Фундаментальная и компьютерная лингвистика» и адресована студентам 2 курса (3 семестр). Дисциплина (модуль) реализуется УНЦ компьютерной лингвистики Института Лингвистики.

Программой дисциплины (модуля) предусмотрены следующие виды контроля: текущий контроль успеваемости в форме: *подготовка к семинарским занятиям, реферат, коллоквиум по контрольным вопросам с выполнением практического задания*; промежуточная аттестация в форме: *зачет*.

Общая трудоемкость освоения дисциплины (модуля) составляет 2 зачетные единицы, 72 часа.

Программой дисциплины (модуля) предусмотрены: практические занятия – 20 часов; самостоятельная работа студента – 52 часа.

Приложение 2. Лист изменений

ЛИСТ ИЗМЕНЕНИЙ

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
1	Приложение к листу изменений №1	31.08.2020г	1

Приложение к листу изменений №1

1. Структура дисциплины (к п. 2 РПД на 2020)

Общая трудоёмкость дисциплины составляет 2 з.е., 76 ч., в том числе контактная работа обучающихся с преподавателем 20 ч., самостоятельная работа обучающихся 56 ч.

2. Образовательные технологии (к п.4 на 2020 г.)

В период временного приостановления посещения обучающимися помещений и территории РГГУ. для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

3. Перечень БД и ИСС (к п. 6 на 2020 г.)

№п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru

4. Состав программного обеспечения (ПО) (к п. 7 на 2020 г.)

№п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)

1	Microsoft Share Point 2010	Microsoft	лицензионное
2	Windows 10 Pro	Microsoft	лицензионное
3	Kaspersky Endpoint Security	Kaspersky	лицензионное
4	Microsoft Office 2016	Microsoft	лицензионное
5	Zoom	Zoom	лицензионное