

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 2021

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

№ 6 от «9» апреля 2021г.

ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

1.3. 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 (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

7. Материально-техническое обеспечение дисциплины

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений

1. Пояснительная записка

1.1 Цель и задачи дисциплины

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».

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

1.2 Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-2 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 умением пользоваться такими ресурсами	ПК-2.2	Умеет пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами

ПК-1 Владеет основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов	ПК-1.1	Знает основные понятия и категории современной лингвистики; основные методы научно-исследовательской деятельности в области фонологического, морфологического, синтаксического, дискурсивного и семантического анализа и правила их применения. Имеет представление об уровневой структуре естественного языка; основных параметрах разнообразия естественных языков; генетической, ареальной и типологической классификации языков; структуре лингвистической науки и ее основных направлениях; основных классических трудах по лингвистике
---	--------	--

1.4. Место дисциплины в структуре образовательной программы

Дисциплина «Автоматический семантический анализ» является дисциплиной по выбору цикла дисциплин ООП ВПО (магистратуры) и относится к части, формируемой участниками образовательных отношений.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: Введение в компьютерную лингвистику, Лингвистическая и онтологическая семантика.

2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 2 з.е., 76ч., в том числе контактная работа обучающихся с преподавателем 20ч., промежуточная аттестация 0ч., самостоятельная работа обучающихся 56ч.

№ 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 текста. Уровни анализа текста. Семантическая структура предложения и семантическое представление. Поверхностная и глубинная семантика	3	1	1	1	6	
2.	Варианты семантической	3	1-2		2	4	Реферирование научного

	структуры предложения. Семантическое представление модели Смысл-Текст.						сочинения из рекомендованного списка: хэндаут или презентация
3.	Варианты семантической структуры предложения. Структура RDF.	3	2		1	5	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	5	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	5	
6	Система семантического анализа StanfordCoreNLP	3	5		2	5	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
7.	Система семантического анализа Voxer.	3	6		1	4	Реферирование научного сочинения из рекомендованного списка: хэндаут или презентация
8.	Системы семантического анализа Wasp и Crisp.	3	7		2	5	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 семантической структуры предложения	3	9		3	5	Выполнение практического задания: построение несложных правил семантического анализа в формализме и алгоритме ЭТАП
11.	Система семантического анализа OntoEtap. Правила разрешения	3	10		2	4	Выполнение практического задания:

	анафоры и установления кореферентности.						построение несложных правил разрешения анафоры и установления кореферентности в формализме и алгоритме ЭТАП
12	Зачет	3				4	Реферат; коллоквиум по контрольным вопросам с выполнением практического задания
	Итого:				20	56	

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

Интерактивные формы обучения в данном курсе предполагают:

1. систематическое использование компьютерных презентаций (как преподавателем в установочной части занятия, так и студентом, выступающим с критическим разбором реферируемого научного сочинения);
2. он-лайн демонстрации работы с лингвистическими базами данных и энциклопедическими интернет-ресурсами (информационно-справочный ресурс по языкам мира Etnolog <http://www.ethnologue.com> (частично-платный ресурс, социолингвистическая информация, ареал распространения, карты); типологическая база данных WALS, The World Atlas of Language Structures <http://wals.info>; информационно-справочный ресурс по языкам мира <http://glottolog.org> (генеалогическая аффилиация и обширная библиография); многоязычная платформа для составления конкордансов и исследования грамматической и лексической дистрибуции WordSketchEngine <https://www.sketchengine.co.uk/>; лексикологическая база данных WordNet (<http://wordnet.princeton.edu/>; электронная

энциклопедия по лингвистике серии Oxford Research Encyclopedia

<http://linguistics.oxfordre.com/>; и др.;

3. использование открытых он-лайн аудио- и видео обучающих ресурсов (для академического английского: www.cambridge.org/elt/english-for-academics);
4. практическую работу с электронными анализаторами речи и платформами лингвистического документирования (платформы E-Language Archiving Technology, ELAN <http://www.lat-mpi.eu/tools/elan>; коллекция электронных ресурсов на сайте Summer Institute of Linguistics (www.sil.org), в том числе, программа для акустического анализа речи Speech analyzer, программа IPA Help для обучения и пользования Международной фонетической транскрипцией, комплекс программ для полевой работы; программа для акустического анализа речи PRAAT (<http://www.fon.hum.uva.nl/praat>); открытое он-лайн издание по документированию и консервации языков Language Documentation & Conservation, LD&C (<http://www.nflrc.hawaii.edu/lcdc>).
5. 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.

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, сдаёт зачёт по всему курсу и предъявляет преподавателю собственноручно написанные конспекты специальной литературы и выполненные домашние задания ко всем семинарам.

5.2. Критерии выставления оценок

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ А,В	«отлично»/ «зачтено (отлично)»/ «зачтено»	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</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>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</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>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,Е	«удовлетвори- тельно»/ «зачтено (удовлетвори- тельно)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>необходимыми для этого базовыми навыками и приёмами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F,FX	«неудовлетворительно»/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. 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. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. 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 источников и 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.

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

№ п/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

7. Материально-техническое обеспечение дисциплины

Занятия по курсу 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.

Состав программного обеспечения

№п /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

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.

АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина реализуется в Институте лингвистики УНЦ компьютерной лингвистики.

Цель дисциплины:

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».

Задачи дисциплины:

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;

Дисциплина направлена на формирование следующих компетенций:

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-2 Владеет принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических, грамматических и иных баз данных и баз	ПК-2.2	Умеет пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами

знаний) и умением пользоваться такими ресурсами		
ПК-1 Владеет основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов	ПК-1.1	Знает основные понятия и категории современной лингвистики; основные методы научно-исследовательской деятельности в области фонологического, морфологического, синтаксического, дискурсивного и семантического анализа и правила их применения. Имеет представление об уровневой структуре естественного языка; основных параметрах разнообразия естественных языков; генетической, ареальной и типологической классификации языков; структуре лингвистической науки и ее основных направлениях; основных классических трудах по лингвистике

По дисциплине предусмотрена промежуточная аттестация в форме зачета.
Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы.

ЛИСТ ИЗМЕНЕНИЙ

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
1	Приложение №		