

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**«Российский государственный гуманитарный университет»  
(ФГБОУ ВО «РГГУ»)**

*ОТДЕЛЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ В ГУМАНИТАРНОЙ СФЕРЕ  
Кафедра математики, логики и интеллектуальных систем в гуманитарной сфере*

***ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ТЕКСТА***

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

45.03.04 Интеллектуальные системы в гуманитарной среде

Разработка и программирование интеллектуальных систем в гуманитарной сфере  
**Уровень квалификации выпускника: бакалавр**

**Форма обучения: очная**

РПД адаптирована для лиц  
с ограниченными возможностями  
здоровья и инвалидов

Москва 2021

Интеллектуальный анализ текста

Рабочая программа дисциплины

Составитель:

Кандидат физико-математических наук, кандидат исторических наук,  
доцент С.В. Шпирко

.....

.....

УТВЕРЖДЕНО

Протокол заседания кафедры МЛиИС

№ 5 от 24.03.2021

## **ОГЛАВЛЕНИЕ**

### **1. Пояснительная записка**

1.1 Цель и задачи дисциплины

1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине

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»

6.3 Профессиональные базы данных и информационно-справочные системы

### **7. Материально-техническое обеспечение дисциплины**

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

### **9. Методические материалы**

9.1. Планы практических (семинарских, лабораторных) занятий

9.2. Методические рекомендации по подготовке письменных работ

9.3. Иные материалы

## 1. Пояснительная записка

### 1.1. Цель и задачи дисциплины

Цель дисциплины —

Формирование компетенций в области разработки, применения и анализа эффективности использования Интеллектуальных Программных Технологий (ИПТ) для (распознавания) текстов

Задачи дисциплины: теоретический анализ и сравнение реализаций различных версий изучаемых моделей для распознавания текстов и определение круга задач, решаемых с их использованием.

### 1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине:

ПК-8:

Способность разрабатывать техническую документацию и использовать средства автоматизации при проектировании информационных систем и систем, основанных на знаниях

<b>Коды компетенции</b>	<b>Содержание компетенций</b>	<b>Перечень планируемых результатов обучения по дисциплине</b>
ПК-8 Способен разрабатывать техническую документацию и использовать средства автоматизации и при проектировании информационных систем и систем, основанных на знаниях	ПК-8.1 Знает стандарты на техническую документацию	Знать: основные методы искусственного интеллекта и особенности их применения в текстологии. Уметь: применять модели, методы и интеллектуальные программные средства для текстологического анализа. Владеть: навыками разработки и применения интеллектуальных программных средств в текстологии (оценке и анализе текстов)
ПК-8 Способен разрабатывать техническую документацию и использовать средства автоматизации и при проектировании	ПК-8.2 Умеет применять CASE-технологии при проектировании информационных систем и систем, основанных на знаниях и отображать результаты проектирования в технической документации	Знать: основные методы искусственного интеллекта и особенности их применения в текстологии. Уметь: применять модели, методы и интеллектуальные программные средства для текстологического анализа. Владеть:

ии информацион ных систем и систем, основанных на знаниях		навыками разработки и применения интеллектуальных программных средств в текстологии (оценке и анализе текстов).
ПК-8 Способен разрабатывать техническую документацию и использовать средства автоматизации и при проектировании информацион ных систем и систем, основанных на знаниях	ПК-8.3 Имеет практический опыт участия в разработке технической документации и проектировании информационных систем и систем, основанных на знаниях	Знать: основные методы искусственного интеллекта и особенности их применения в текстологии. Уметь: применять модели, методы и интеллектуальные программные средства для текстологического анализа. Владеть: навыками разработки и применения интеллектуальных программных средств в текстологии (оценке и анализе текстов).

### 1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Интеллектуальный анализ текстов» относится к части блока дисциплин Б1 учебного плана, формируемой участниками образовательных отношений.

Для освоения дисциплины необходимы знания, сформированные в ходе изучения следующих дисциплин:

логика (логика высказываний, исчисление эквивалентных формул),

алгебра (элементы теории множеств),

программирование на C++, Delphi

интеллектуальные системы (элементы нечеткой логики, системы нечеткого вывода, нечеткие множества)

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик:

Интеллектуальный анализ текстов

## 2. Структура дисциплины

### Структура дисциплины для очной формы обучения

Общая трудоёмкость дисциплины составляет 6 з.е., 216 академических часов.

Объём дисциплины в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Семестр	Тип учебных занятий	Количество часов
7	Лекции	10
7	Практические занятия	32
8	Лекции	10
8	Практические занятия	32
Всего:		84

Объём дисциплины в форме самостоятельной работы обучающихся составляет 132 академических часа.

№ п/п	Раздел дисциплины/темы	Се м е ст р	Виды учебной работы (в часах)					Пром ежуто чная аттест ация	Са мо сто яте льная раб ота	Формы текущего контроля успеваемости, форма промежуточной аттестации ( <i>по семестрам</i> )
			контактная				Лабо ратор ные занят ия			
			Лек ции	Се ми нар	Пр ак тиче ск ие за ня тия					
1	Введение в компьютерную текстологию	7	2					16	Устный опрос	
2	Методы и технологии автоматического сличения текстов	7	2			12		16	Устный опрос Оценка выполнения практических заданий Домашнее задание	
3	Формализованная оценка значимости разночтений	7	2			6		16	Оценка выполнения практических заданий Домашнее задание	
4	Формализованные методы	7	2			10		16	Оценка выполнения	

	текстологического анализа (классификации)							практических заданий Письменная контрольная работа Доклад. Презентация	
<b>5</b>	Неформальное введение в теорию нечетких множеств	<b>7</b>	4			10		16	Устный опрос Домашнее задание Оценка выполнения практических заданий
<b>6</b>	Нечеткая генеалогическая классификация с заданным уровнем нечеткости	<b>8</b>	4			12		20	Устный опрос Домашнее задание Оценка выполнения практических заданий Письменная контрольная работа
<b>7</b>	Анализ текста в рамках логики распознавания А.Д. Закревского	<b>8</b>	2			8		16	Устный опрос Домашнее задание Оценка выполнения практических заданий
<b>8</b>	Закономерности и их выявление в рамках логики распознавания А.Д. Закревского	<b>8</b>	2			6		16	Устный опрос Домашнее задание Оценка выполнения практических заданий Письменная контрольная работа
	Зачет с оценкой								зачет
	итоги:		<b>20</b>			<b>64</b>		<b>132</b>	

### 3. Содержание дисциплины

№	Наименование раздела дисциплины	Содержание
<i>Первый семестр</i>		
1	Введение в компьютерную текстологию	Цель и задачи компьютерной текстологии Рукописная традиция. Примеры Контроль над стабильностью текста. Разночтения и узлы разночтений. Коллация (сличение) списков текста. Характер текстовых групп. Компьютерная классификация списков. Матрица расстояний как способ оценки близости текстов
2	Методы и технологии автоматического сличения текстов	Сличение списков:нахождение наибольшей общей подпоследовательности пары последовательностей Предварительный этап сличения списков Парное сравнение списков (Задача нахождения наибольшей общей подпоследовательности) Diff алгоритм Ю.Майерса (простая схема) Оценка сложности Diff алгоритма Сравнение текстов списков: проблемы и подходы на примере программы CollateX Цели создания CollateX; Этапы работы; Используемые методы и алгоритмы; Выявление перестановок Применение динамического программирования для сравнения пары списков (алгоритм Нидлмана-Вунша) Обобщение алгоритма на случай сравнения множества списков
3	Формализованная оценка значимости разночтений	Лингвистические типологизации разночтений Типологизация разночтений при классификации списков “Русской Правды Текстолого-лингвистическая типологизация разночтений на материале “Закона Судного людем” и “Предсловия покаянию” Метод косвенной экспертизы для оценки значимости разночтений
4	Формализованные методы текстологического анализа (классификации)	Кластерная классификация Группировка по некоторым признакам сходства или различия Метод Дееса-Ваттеля (агломеративно-иерархический метод) Кладистические методы Принцип экономии: Нахождение дерева с наибольшей длиной Метод групп Дом Фроже-Бородкина Метод Крускала

		<p>Попытки преодоления однозначности в формализованной текстологической классификации:</p> <p>Метод Дом Кантэна</p> <p>Метод Грега формализованного представления вариантов и генеалогической классификации текстов</p> <p>(The Calculus of Variants)</p>
5	Неформальное введение в теорию нечетких множеств	<p>Моделирование в условиях неопределенности с позиции нечетких множеств:</p> <p>“Лингвистические” переменные; Нечеткие высказывания; Нечеткие алгоритмы</p> <p>Нечеткие множества, основные понятия. Функция принадлежности. Операции над нечеткими множествами. Нечеткие числа. Сравнение нечетких чисел</p> <p>Нечеткие отношения</p> <p>Нечеткий вывод. Основные правила вывода в нечеткой логике.</p> <p>Нечеткое управление.</p>
6	Нечеткая генеалогическая классификация с заданным уровнем нечеткости	<p>Нечеткие отношения и их свойства</p> <p>Нечеткое отношение: Матричное и табличное представление</p> <p>Нечеткие отношения, их свойства.</p> <p>Транзитивное замыкание</p> <p><math>\alpha</math>-срез нечеткого множества</p> <p>Класс нечеткой эквивалентности</p> <p>Каноническое разбиение</p> <p>Нечеткие отношения порядка</p> <p>Линейное упорядочивание базового множества</p> <p>Построение матрицы нечеткого отношения</p> <p>Алгоритм нечеткой генеалогической классификации с заданным уровнем нечеткости</p>
7	Анализ текста в рамках логики распознавания А.Д. Закревского	<p>Трансформации (синтаксические) исходного текста</p> <p>Логика распознавания: Пространство многозначных признаков</p> <p>построение интервала запрета</p> <p>Исчисление предикатов как формальная модель рассуждений</p> <p>Представление данных и знаний в виде секционированного булева вектора</p> <p>Центральная задача дедуктивного вывода</p> <p>Метод резолюции</p> <p>Задача распознавания</p>
8	Закономерности и их выявление в рамках логики распознавания А.Д. Закревского	<p>Матрица запретов</p> <p>Проблема восполнения данных</p>

		Этапы нахождения кратчайшей формы матрицы запретов Индуктивный и дедуктивный этапы алгоритма восстановления пропусков

#### 4. Образовательные технологии

<b>№ п/п</b>	<b>Наименование раздела</b>	<b>Виды учебной работы</b>	<b>Информационные и образовательные технологии</b>
<b>1</b>	<b>2</b>	<b>3</b>	<b>5</b>
1	Введение в компьютерную текстологию	Лекция 1 Самостоятельная работа	Вводная лекция-беседа. Работа с интернет-ресурсами. Консультирование посредством электронной почты
2	Методы и технологии автоматического сличения текстов	Лекция 2 Лабораторное занятие 1-6 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
3	Формализованная оценка значимости различий	Лекция 3 Лабораторное занятие 7-9 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
4	Формализованные методы текстологического анализа (классификации)	Лекция 2 Лабораторное занятие 10-14 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Письменная контрольная работа Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты Доклад. Презентация
5	Неформальное введение в теорию нечетких множеств	Лекция 2 Лабораторное занятие 15-19 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
6	Нечеткая генеалогическая	Лекция 2	Теоретическая лекция. Практическая работа в компьютерном классе.

	классификация с заданным уровнем нечеткости	Лабораторное занятие 20-25 Самостоятельная работа	Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Письменная контрольная работа Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
7	Анализ текста в рамках логики распознавания А.Д. Закревского	Лекция 2 Лабораторное занятие 26-29 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
8	Закономерности и их выявление в рамках логики распознавания А.Д. Закревского	Лекция 2 Лабораторное занятие 30-32 Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Письменная контрольная работа Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
1	Мягкие вычисления	Лекция 1. Самостоятельная работа	Вводная лекция-беседа. Работа с интернет-ресурсами. Консультирование посредством электронной почты
2	Нечеткие множества	Лекция 2 Лабораторное занятие 1, 2 Самостоятельная работа	Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
3	Эволюционные алгоритмы	Лекция 3 Лабораторное занятие 3, 4  Самостоятельная работа	Теоретическая лекция. Практическая работа в компьютерном классе. Изучение инструментальных средств и реализация простейших алгоритмов. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
4	Искусственные нейронные сети	Лекция 3 Лабораторное занятие 5, 6 Лабораторное	Теоретическая лекция. Практическая работа в компьютерном классе. Выполнение практического задания. Практическая работа в компьютерном классе.

		занятие 7, 8 Лабораторное занятие 9, 10 Самостоятельная работа	Выполнение практического задания. Практическая работа в компьютерном классе. Выполнение практического задания. Работа с конспектом и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты
--	--	--	---

## 5. Оценка планируемых результатов обучения

### 5.1. Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль:		
<input type="checkbox"/> опрос	5 баллов	10 баллов
<input type="checkbox"/> дом. задание (темы 2-4)	5 баллов	15 баллов
<input type="checkbox"/> контр. работа (темы 2-3)	15 баллов	15 баллов
<input type="checkbox"/> выполнение практических заданий (темы 3-4)	5 баллов	20 баллов
Промежуточная аттестация (зачет с оценкой)		40 баллов
Итого за семестр (дисциплину)		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу 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

## 5.2. Критерии выставления оценки по дисциплине

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. 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	«удовлетворительно»/ «зачтено (удовлетворительно)»/ «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 необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p>

		Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	«неудовлетворительно»/ не зачтено	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.

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

#### 5.3.1. Образцы заданий для самостоятельного выполнения

1.  $U = \{0, 1, 2, \dots, 120\}$  — возраст человека. Постройте графики функций принадлежности следующих нечетких множеств: А – молодой, В — старый, С — нестарый, D — очень молодой. Постройте графики функций принадлежности объединения и пересечения множеств А и В, В и С.
2. Даны 5 списков (А, В, С, D, E), состоящие из заданной последовательности символов из алфавита:  $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, b_1, b_5, b_7, c_1\}$ . Построить стемму списков, используя модель Фроже, выбрав в качестве основного список А.
3. Даны 5 списков (А, В, С, D, E), состоящие из заданной последовательности символов из алфавита:  $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, b_1, b_5, b_7, c_1\}$ . Построить неориентированный граф, используя модель Кантэна, для четверки списков А, В, С, D
4. Даны 4 списка (А, В, С, D), состоящие из заданной последовательности символов из алфавита с указанием ошибочных чтений:  $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, b_1, b_5, b_7, c_1\}$ . Провести нечеткую классификацию с заданным уровнем нечеткости.
5. Объекты исследуемого мира характеризуются 4-мя признаками:  $x_1, x_2, x_3, x_4$ , принимающие заданные значения. Построить модель мира, используя заданные закономерности между признаками. Применяя принцип дедукции, выяснить истинность данного утверждения.

#### 5.3.2. Образцы заданий для контрольных работ

##### Контрольная работа

1. Два числа задаются нечеткой функцией принадлежности. Используя известные критерии (детерминистский, интегральный), сравнить эти числа.
2. Задана матрица типа объект-признак. Построить соответствующую матрицу запрета.

### 5.3.3 Список теоретических вопросов, выносимых на промежуточную аттестацию (зачет)

1. Цель и задачи текстологии
2. Матрица расстояний как способ оценки близости текстов
3. Сличение списков:нахождение наибольшей общей подпоследовательности пары последовательностей
4. Парное сравнение списков
- 5.Применение динамического программирования для сравнения пары списков
6. Метод косвенной экспертизы для оценки значимости различий
7. Кластерная классификация
8. Кладистические методы классификации
9. Метод групп Дом Фроже-Бородкина
10. Нечеткие множества
11. Нечеткие отношения и их отношения
12. Основные правила вывода в нечеткой логике
13. Каноническое разбиение по отношению нечеткой эквивалентности
14. Линейное упорядочивание для совершенного порядка
15. Этапы нечеткой генеалогической классификации с заданным порогом
16. Логика распознавания А.Д. Закревского: Пространство многозначных признаков
17. Исчисление предикатов как формальная модель рассуждений
- 18.Центральная задача дедуктивного вывода
19. Метод резолюции
20. Задача распознавания
21. Алгоритм восстановления пропусков

## 6. Учебно-методическое и информационное обеспечение дисциплины

### 6.1. Список источников и литературы

#### а) Основная литература

1. Бородкин Л.И., Милов Л.В. О некоторых аспектах автоматизации текстологического исследования (Закон Судный людям)// Математические методы в историко-экономических и историко-культурных исследованиях/ ред. И.Д. Ковальченко. М., 1977. С.235-280;
2. Закревский А.Д. Логика распознавания. М., 2003. 144 С.;
3. Мелихов А.Н., Бернштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечеткой логикой. М., 1990. 272 С.;
4. Шпирко С.В. Применение теории нечетких множеств к задаче генеалогической классификации в текстологическом исследовании// Историческая информатика: Информационные технологии и математические методы в исторических исследованиях и образовании. Барнаул, 2013. № 3. С. 39-51.

#### б) Дополнительная литература

1. Введение в математическое моделирование/ В.Н. Ашихмин, М.Б. Гитман и др. М., 2007. 440 С;
2. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. М., 1976. 165 С.;
3. Прикладная и компьютерная лингвистика/ ред. И.С. Николаева, О.В. Митренина, Т.М. Ландо. М., 2016. 320 С.;
4. Шпирко С.В. Логика распознавания А.Д. Закревского на примере задачи восстановления пропусков в историко-археологических исследованиях. 2014. № 1. С. 62–69
5. Шпирко С.В. О возможности логического анализа текста (на примере восточных средневековых сведений о народах Восточной Европы // Историческая информатика. 2014. №2-3. С. 80–93.

## 6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

<http://www.wolfram.com/>

<https://www.wolframcloud.com/>

## 6.3 Профессиональные базы данных и информационно-справочные системы

№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 Доступ к профессиональным базам данных: <a href="https://liber.rsuh.ru/ru/bases">https://liber.rsuh.ru/ru/bases</a>
4	Компьютерные справочные правовые системы Консультант Плюс, Гарант

## 7. Материально-техническое обеспечение дисциплины

Компьютерный класс, компьютер преподавателя, компьютеры студентов, проектор, экран, доступ в интернет.

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2013	Microsoft	лицензионное
3	Windows 10 Pro	Microsoft	лицензионное

4	Kaspersky Endpoint Security	Kaspersky	лицензионное
5	Zoom	Zoom	лицензионное

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

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, регулируемыми эргономическими партами СИ-1;*
- *компьютерной техникой со специальным программным обеспечением.*

## 9. Методические материалы

### 9.1 Планы лабораторных занятий

*Тема1 (16ч.) Введение* в компьютерную текстологию

*Цель занятия:* изучение современных формализованных математических методов текстологической классификации

Форма проведения –опрос, решение задач.

*Вопросы для обсуждения:*

Цель и задачи компьютерной текстологии

Рукописная традиция. Примеры

Контроль над стабильностью текста.

Разночтения и узлы разночтений.

Колляция (сличение) списков текста.

Характер тестовых групп.

Компьютерная классификация списков.

Матрица расстояний как способ оценки близости текстов

Лингвистические типологизации разночтений

Типологизация разночтений при классификации списков “Русской Правды

Текстолого-лингвистическая типологизация разночтений на материале “Закона Судного людем” и “Предсловия покаянию”

Метод косвенной экспертизы для оценки значимости разночтений

Кластерная классификация

Группировка по некоторым признакам сходства или различия

Метод Дееса-Ваттеля

(агломеративно-иерархический метод)

Кладистические методы

Принцип экономии: Нахождение дерева с наибольшей длиной

Метод групп Дом Фроже-Бородкина

Метод Крускала

Попытки преодоления однозначности в формализованной текстологической классификации:

Метод Дом Кантэна

Метод Грега формализованного представления вариантов и генеалогической классификации текстов

(The Calculus of Variants)

Задание: провести классификацию набора списков из последовательности символов из алфавита, используя известные методы формализованной текстологической классификации:

*Список источников и литературы:*

1. Бородкин Л.И., Милов Л.В. О некоторых аспектах автоматизации текстологического исследования (Закон Судный людям)// Математические методы в историко-экономических и историко-культурных исследованиях/ ред. И.Д. Ковальченко. М., 1977. С.235-280;
2. Прикладная и компьютерная лингвистика/ ред. И.С. Николаева, О.В. Митренина, Т.М. Ландо. М., 2016. 320 С.;
3. Borsetta P. and Zarri G.P. An application of the Quentin/80 software to the study of the manuscript tradition of the Appendix Vergiliana (semi-automatic construction of the Stemmata Codicum)// Actes du Congres international informatique et sciences humaines. Liege, 1980. P. 73-92;
4. Colwell E.C. and Tune E.W. The quantitative relationship between MS text-types// Biblical and patristic studies in memory of R.P.Casey/ Ed. by J.N.Birdsall and P.W.Thomson. Freiburg, 1963. P.25-32;
5. Dees A. Sur une constellation de quatre manuscrits// Melanges de linguistique et de literature offertes a Lein Geschiere. Amsterdam, 1975. P.1-9;
6. Froger D.J. La critique des textes et son automatization. Paris, 1968. 280 P.;
7. Greg W. W. The Calculus of Variants. An Essay on Textual Criticism. Oxford, 1927. 63 P.;
8. Greg W. W Recent theories of textual criticism// Modern philology. Chicago, 1931. Vol. XXVIII. May. № 4. P. 401-404;
9. Quentin H. Dom. Memoire sur l'etablissement du texte de la Vulgate. Rome and Paris, 1922. 520 P.;
10. Quentin H. Dom. Essais de critique textuelle. Paris, 1926. 179 P.;
11. Saaty T.L. Measuring the fuzziness of sets// Journal of Cybernetics. V. 4. 1974. P.53-61;
12. Shepard William P. Recent theories of textual criticism// Modern philology. Chicago, 1930. Vol. XXVIII. November 2. P. 129-141;

*Тема2 (12ч.)* Методы и технологии автоматического сличения текстов

*Цель занятия:* изучение современных подходов и алгоритмов автоматического сличения текстов

Форма проведения –опрос, решение задач, контрольная работа, индивидуальные доклады с презентацией.

*Вопросы для обсуждения:*

Сличение списков:нахождение наибольшей общей подпоследовательности пары последовательностей

Предварительный этап сличения списков

Парное сравнение списков (Задача нахождения наибольшей общей подпоследовательности)

Diff алгоритм Ю.Майерса (простая схема)

Оценка сложности Diff алгоритма

Сравнение текстов списков: проблемы и подходы на примере программы CollateX

Цели создания CollateX; Этапы работы; Используемые методы и алгоритмы; Выявление перестановок

Применение динамического программирования для сравнения пары списков (алгоритм Нидлмана-Вунша)

Обобщение алгоритма на случай сравнения множества списков

Задание: найти общую подпоследовательность из пары текстов списков, используя известные методы выравнивания последовательностей.

*Список источников и литературы:*

1. Eugene W. Myers. An O(ND) Difference Algorithm and Its Variations// *Algorithmica* 1. p.251-266 (1986);
2. H. Carrillo & D. Lipman. The multiple sequence alignment problem in biology. *Siam J. Appl. Math.* Vol. 48. № 5 (1988);
3. R. Dekker etc. Computer-supported collation of modern manuscripts: CollateX and tge Beckett Digital Manuscript// *Digital Scholarship in the Humanities*. Vol. 30. № 3. Pp. 452-470 (2014).

*Тема 3 (22ч.)* Нечеткая генеалогическая классификация (текстов списков)

*Цель занятия:* изучение аппарата теории нечетких множеств для решения задачи нечеткой генеалогической классификации (текстов списков)

Форма проведения – опрос, решение задач, контрольная работа.

*Вопросы для обсуждения:*

Моделирование в условиях неопределенности с позиции нечетких множеств:

“Лингвистические” переменные; Нечеткие высказывания; Нечеткие алгоритмы

Нечеткие множества, основные понятия. Функция принадлежности. Операции над нечеткими множествами. Нечеткие числа. Сравнение нечетких чисел

Нечеткие отношения

Нечеткий вывод. Основные правила вывода в нечеткой логике.

Нечеткое управление

Нечеткие отношения и их свойства

Нечеткое отношение: Матричное и табличное представление

Нечеткие отношения, их свойства.

Транзитивное замыкание

$\alpha$ -срез нечеткого множества

Класс нечеткой эквивалентности

Каноническое разбиение

Нечеткие отношения порядка

Линейное упорядочивание базового множества

Построение матрицы нечеткого отношения

Алгоритм нечеткой генеалогической классификации с заданным уровнем нечеткости

Задание: провести сравнение двух нечетких чисел; смоделировать высказывания на естественном языке в терминах нечеткой логики и нечеткого вывода, построить матрицу нечеткого отношения с помощью метода косвенной экспертизы, провести операцию транзитивного замыкания, выделить классы эквивалентности, провести операцию упорядочивания для нечеткого отношения порядка; провести нечеткую генеалогическую классификацию набора текстов списков с заданным уровнем нечеткости

*Список источников и литературы:*

1. Мелихов А.Н., Бернштейн Л.С., Коровин С.Я. Ситуационные советуемые системы с нечеткой логикой. М., 1990. 272 С.;

2. Шпирко С.В. Применение теории нечетких множеств к задаче генеалогической классификации в текстологическом исследовании// Историческая информатика: Информационные технологии и математические методы в исторических исследованиях и образовании. Барнаул, 2013. № 3. С. 39-51.;
3. Шпирко С.В., Баранкова Г.С. О некоторых аспектах построения формализованной генеалогической классификации текстов списков средневекового произведения с применением теории нечетких множеств (на материале "Закона Судного людем") // Исторический журнал: научные исследования. 2017. № 1. с. 56-64.

*Тема4 (22ч.)* Анализ текста на естественном языке и выявление закономерностей текста в рамках логики распознавания А.Д. Закревского.

*Цель занятия:* неформальное введение в логику распознавания А.Д. Закревского

Форма проведения –опрос, решение задач, контрольная работа.

*Вопросы для обсуждения:*

Трансформации (синтаксические) исходного текста

Логика распознавания: Пространство многозначных признаков  
построение интервала запрета  
Исчисление предикатов как формальная модель рассуждений  
Представление данных и знаний в виде секционированного булева вектора  
Центральная задача дедуктивного вывода  
Метод резолюции  
Задача распознавания  
Матрица запретов  
Проблема восполнения данных

Этапы нахождения кратчайшей формы матрицы запретов

Индуктивный и дедуктивный этапы алгоритма восстановления пропусков  
Задание: представить знания о мире в виде матрицы КНФ, проверить истинность элементарного дизъюнкта с помощью правила дедукции, провести операцию резольвирования, построение матрицы запретов, решить задачу восполнения пропусков (расознавания объекта)

*Список источников и литературы:*

1. Гуц А.К. Математическая логика и теория алгоритмов. М., 2009. 120 с.;
2. Закревский А.Д. Логика распознавания. М., 2003. 144 С.;
3. Закревский А.Д. Алгоритмы синтеза дискретных автоматов. М., 1971. 512 С.
4. Закревский А.Д. Представление знаний и логический вывод в пространстве многозначных признаков // Логика и компьютер. Логические языки, содержательные рассуждения и методы поиска доказательств. Вып. 2. М., 1995. С. 3–16;
5. Шпирко С.В. Логика распознавания А.Д. Закревского на примере задачи восстановления пропусков в историко-археологических исследованиях. 2014. № 1. С. 62–69;
6. Шпирко С.В. О возможности логического анализа текста (на примере восточных средневековых сведений о народах Восточной Европы // Историческая информатика. 2014. №2-3. С. 80–93;
7. Гуц А.К. Математическая логика и теория алгоритмов. М., 2009. 120 с.;