

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего  
образования

«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»

---

Кафедра №63

«УТВЕРЖДАЮ»

Руководитель направления

к.ф.н., доц.

(должность, уч. степень, звание)

М.А. Чиханова

(подпись)

«28» мая 2020 г

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Информационный поиск и извлечение информации»

(Название дисциплины)

Код направления	45.03.02
Наименование направления/ специальности	Лингвистика
Наименование направленности	Теоретическая и прикладная лингвистика
Форма обучения	очная

Санкт-Петербург 2020 г.

Лист согласования рабочей программы дисциплины

Программу составил(а)

доц., к. филол. наук, доц.

должность, уч. степень, звание



подпись, 06. 05.2020 г.

М. А. Чиханова

инициалы, фамилия

Программа одобрена на заседании кафедры № 63

«06» мая 2020 г, протокол № 9

Заведующий кафедрой № 63

доц., к. филол. н., доц. «06» мая 2020 г

должность, уч. степень, звание

подпись, дата



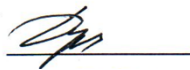
М.А. Чиханова

инициалы, фамилия

Ответственный за ОП 45.03.02(02)

доц., к. ф. н.

должность, уч. степень, звание



подпись, 06. 05.2020 г.

Е.Ю. Дубинина

инициалы, фамилия

Заместитель директора института (декана факультета) № 6 по методической работе



доц., к. п. н., доц.

должность, уч. степень, звание

подпись, 06. 05.2020 г.

И.М. Евдокимов

инициалы, фамилия

## Аннотация

Дисциплина «Информационный поиск и извлечение информации» входит в вариативную часть образовательной программы подготовки обучающихся по направлению 45.03.02 «Лингвистика» направленность «Теоретическая и прикладная лингвистика». Дисциплина реализуется кафедрой №63.

Дисциплина нацелена на формирование у выпускника

общефессиональных компетенций:

ОПК-11 «владение навыками работы с компьютером как средством получения, обработки и управления информацией»,

ОПК-20 «способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности»;

профессиональных компетенций:

ПК-19 «способность работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности»,

ПК-20 «владение методами формального и когнитивного моделирования естественного языка и методами создания метаязыков»,

ПК-21 «владение основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов»,

ПК-22 «владение стандартными способами решения основных типов задач в области лингвистического обеспечения информационных и других прикладных систем»,

ПК-25 «владение основами современных методов научного исследования, информационной и библиографической культурой»,

ПК-26 «владение стандартными методиками поиска, анализа и обработки материала исследования».

Содержание дисциплины охватывает круг вопросов, связанных с базовыми понятиями теории информационного поиска, основными принципами организации информационно-поисковых систем и алгоритмами аналитико-синтетической переработки документов.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лабораторные работы, самостоятельная работа обучающегося, консультации.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме экзамена.

Общая трудоемкость освоения дисциплины составляет 3 зачетных единицы, 108 часов.

Язык обучения по дисциплине русский.

## **1. Перечень планируемых результатов обучения по дисциплине**

### **1.1. Цели преподавания дисциплины**

Содержание дисциплины охватывает круг вопросов, связанных с базовыми понятиями теории информационного поиска, основными принципами организации информационно-поисковых систем и алгоритмами аналитико-синтетической переработки документов.

### **1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП**

В результате освоения дисциплины обучающийся должен обладать следующими компетенциями:

ОПК-11 «владение навыками работы с компьютером как средством получения, обработки и управления информацией»:

знать - основы теории информационного поиска

уметь - осуществлять обоснованный выбор технологии построения информационно-поисковой системы;

владеть навыками - основными технологиями построения информационно-поисковых систем;

иметь опыт деятельности - основные этапы аналитико-синтетической переработки документов.

ОПК-20 «способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности»:

знать - основные принципы организации информационно-поисковых систем;

уметь - практически реализовывать основные алгоритмы информационного поиска

владеть навыками - технологиями построения тезаурусов и онтологий;

иметь опыт деятельности - основными алгоритмами аналитико-синтетической переработки документов.

ПК-19 «способность работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности»:

знать – основные информационно-поисковые и экспертные системы

уметь – работать с формальными системами обработки естественного языка

владеть навыками – работы с системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, автоматизированными системами идентификации и верификации личности»

иметь опыт деятельности – в области обработки лексикографической информации и автоматизированного перевода;

ПК-20 «владение методами формального и когнитивного моделирования естественного языка и методами создания метаязыков»:

знать – методы формального и когнитивного моделирования естественного языка и методы создания метаязыков

уметь – пользоваться указанными методами

владеть навыками – применения данных методов для анализа языкового материала

иметь опыт деятельности – построения текста при помощи метаязыка;

ПК-21 «владение основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов»:

знать – принципы автоматической обработки лингвистических корпусов

уметь - работать с электронными словарями и другими электронными ресурсами для решения лингвистических задач

владеть навыками – математико-статистической обработки лингвистической информации

иметь опыт деятельности - работы с компьютером как средством получения, обработки и управления информацией.

ПК-22 «владение стандартными способами решения основных типов задач в области лингвистического обеспечения информационных и других прикладных систем»:

знать - терминологию, алгоритмы и схемы, приемы и законы создания и чтения чертежей и документации по аппаратным и программным компонентам информационных систем;

уметь - определять назначение технического описания ИС, нормативной документации и методических указаний к выполнению работы;

владеть навыками - демонстрации результатов анализа и расчётов, выполненных согласно изученному техническому описанию, ПО, нормативной документации и методическим указаниям;

иметь опыт деятельности - в анализе потребности в ходе выполнения задания и необходимости в использовании дополнительных источников: ПО, учебных пособий и интернет-сайтов, демонстрирует использование новых материалов и дополнительных опций в работе

ПК-25 «владение основами современных методов научного исследования, информационной и библиографической культурой»:

знать – основы современных методов научного исследования, информационной и библиографической культуры;

уметь – пользоваться библиотеками разного типа;

владеть навыками - основными способами выражения семантической, коммуникативной и структурной преемственности между частями высказывания - композиционными элементами текста (введение, основная часть, заключение), сверхфразовыми единствами, предложениями;

иметь опыт деятельности - использования основных способов выражения семантической, коммуникативной и структурной преемственности между частями высказывания - композиционными элементами текста, сверхфразовыми единствами, предложениями;

ПК-26 «владение стандартными методиками поиска, анализа и обработки материала исследования»:

знать – основные механизмы поиска, анализа и обработки материала исследования;

уметь - применять теоретические знания в решении практических задач по теоретической и прикладной лингвистике;

владеть навыками - реализации коммуникативных целей высказывания применительно к особенностям текущего коммуникативного контекста иметь опыт деятельности – проведения лингвистического анализа;

## 2. Место дисциплины в структуре ОП

Дисциплина базируется на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- Компьютерные программы в лингвистических исследованиях

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и используются при изучении других дисциплин:

- Обработка текстовых массивов
- Базы данных
- Производственная преддипломная практика

## 3. Объем дисциплины в ЗЕ/академ. час

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 1

Таблица 1 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№5
1	2	3
<b>Общая трудоемкость дисциплины, ЗЕ/(час)</b>	3/ 108	3/ 108
<i>Аудиторные занятия</i> , всего час., <i>В том числе</i>	34	34
лекции (Л), (час)		
Практические/семинарские занятия (ПЗ), (час)		
лабораторные работы (ЛР), (час)	34	34
курсовой проект (работа) (КП, КР), (час)		
Экзамен, (час)	54	54
<i>Самостоятельная работа</i> , всего	20	20
<b>Вид промежуточного контроля:</b> зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.)	Экз.	Экз.

#### 4. Содержание дисциплины

##### 4.1. Распределение трудоемкости дисциплины по разделам и видам занятий

Разделы и темы дисциплины и их трудоемкость приведены в таблице 2.

Таблица 2. – Разделы, темы дисциплины и их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
<b>Семестр 5</b>					
Раздел 1. Понятие «информация». Многоуровневая модель информации. Уточнение используемой терминологии.			4		2
Раздел 2. Информатика и семиотика.			4		2
Раздел 3. Общие принципы организации информационно-поисковых систем.			4		2
Раздел 4. Метаданные и обработка электронных ресурсов					2
Раздел 5. Модель информационно-поисковой системы. Структура логических компонентов информационно-поисковой системы.			4		4
Раздел 6. Построение тезаурусов и онтологий информационно-поисковых систем.			4		4
Раздел 7. Извлечение метаданных из слабоструктурированных документов.			4		4
Итого в семестре:			34		20
Итого:	0	0	34	0	20

##### 4.2. Содержание разделов и тем лекционных занятий

Содержание разделов и тем лекционных занятий приведено в таблице 3.

Таблица 3 - Содержание разделов и тем лекционных занятий

Номер раздела	Название и содержание разделов и тем лекционных занятий
	<b>Учебным планом не предусмотрено</b>

##### 4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 4.

Таблица 4 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	№ раздела дисциплины
<b>Учебным планом не предусмотрено</b>				

#### 4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	№ раздела дисциплины
Семестр 5			
1	Построение многоуровневая модели информации.	4	1
2	Знакомство с поисковыми системами.	4	2
3	Общие принципы организации информационно-поисковых систем.	4	3
4	Метаданные и обработка электронных ресурсов.	4	4
5	Модель информационно-поисковой системы. Структура логических компонентов информационно-поисковой системы.	4	5
6	Построение тезаурусов и онтологий информационно-поисковых систем.	4	6
7	Извлечение метаданных из слабоструктурированных документов.	4	7
8	Кластеризация текстовых документов на основе меры сходства.	4	8
9	Анализ проведенной кластеризации	2	8
Всего:		34	

#### 4.5. Курсовое проектирование (работа)

Учебным планом не предусмотрено

#### 4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 6.

Таблица 6 Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 5, час
1	2	3
<b>Самостоятельная работа, всего</b>	20	20
изучение теоретического материала дисциплины (ТО)	10	10
курсовое проектирование (КП, КР)		



расчетно-графические задания (РГЗ)		
выполнение реферата (Р)		
Подготовка к текущему контролю (ТК)	3	3
домашнее задание (ДЗ)	7	7
контрольные работы заочников (КРЗ)		

## 5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 8-10.

## 6. Перечень основной и дополнительной литературы

### 6.1. Основная литература

Перечень основной литературы приведен в таблице 7.

Таблица 7 – Перечень основной литературы

Шифр	Библиографическая ссылка / URL адрес	Количество экземпляров в библиотеке (кроме электронных экземпляров)
	Большакова Е. И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Е. И Большакова., Э. С Клышинский., Д. В. Ландэ, А. А Носков., О. В. Пескова, Е. В. Ягунова. - М.:МИЭМ, 2011. – 272 с.	
	Лукашевич Н.В. Тезаурусы в задачах информационного поиска / Н. В. Лукашевич. - М.: Изд-во Московского университета, 2011	
	Назаров С.В. Основы информационных технологий / С.В. Назаров, Л.П. Гудыно и др.. - М.: ИНТУИТ 2012. - 596 с.	

### 6.2. Дополнительная литература

Перечень дополнительной литературы приведен в таблице 8.

Таблица 8 – Перечень дополнительной литературы

Шифр	Библиографическая ссылка/ URL адрес	Количество экземпляров в библиотеке (кроме электронных экземпляров)
	Сегаран. Т. Программируем коллективный разум. / Т. Сегаран. – Пер. с англ. – СПб: Символ-Плюс, 2008. – 368с.	

	Семенов А.Л. Современные информационные технологии и перевод: учеб. Пособие для студ. перевод. фак. высш. учеб. заведений / А.Л. Семенов. – М.: Издательский центр «Академия», 2008. - 224с.	
	Щипицина Л.Ю. Информационные технологии в лингвистике : учеб. пособие / Л.Ю. Щипицина. - М. : Флинта : Наука, 2013. - 128с.	

## **7. Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины**

Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины

URL адрес	Наименование
<a href="http://www.logrus.ru">http://www.logrus.ru</a>	Многоязыковые переводы и локализация программ, документации, игр и мультимедиа
<a href="http://www.ruscorpus.ru">http://www.ruscorpus.ru</a>	Собрание грамматически размеченных русских текстов XIX-XXI вв. в электронной форме
<a href="http://www.tra-service.ru">http://www.tra-service.ru</a>	Системы автоматизации перевода: решения SDL Trados, Plunet, WhiteSmoke

## **8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине**

### **8.1. Перечень программного обеспечения**

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10 – Перечень программного обеспечения

№ п/п	Наименование
1.	Операционная система: Microsoft® Windows® Vista Business Russian Тип лицензии: Academic Номер лицензии 44260430
2.	Офис: Microsoft® Office Professional Plus 2007 Russian Тип лицензии: Academic Номер лицензии 44260430

### **8.2. Перечень информационно-справочных систем**

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11 – Перечень информационно-справочных систем

№ п/п	Наименование
-------	--------------

1.	Электронно-библиотечная система Лань URL: <a href="https://e.lanbook.com/">https://e.lanbook.com/</a>
2.	Электронно-библиотечная система Znanium URL: <a href="https://znanium.com/">https://znanium.com/</a>
3.	Информационно-правовой портал ГАРАНТ.РУ URL: <a href="http://www.garant.ru/">http://www.garant.ru/</a>
4.	Правовая поддержка КОНСУЛЬТАНТПЛЮС URL: <a href="http://www.consultant.ru">http://www.consultant.ru</a>

## **9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине**

Состав материально-технической базы представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Аудитория для практических занятий Компьютерный класс	Аудитория укомплектована специализированной мебелью, техническими средствами обучения, служащими для представления учебной информации большой аудитории Ауд. 34-10
2	Аудитории общего пользования (для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации)	Аудитория укомплектована специализированной мебелью, техническими средствами обучения, служащими для предоставления учебной информации большой аудитории
3	Библиотека, Интернет-класс ГУАП (для самостоятельной работы)	Помещения укомплектованы специализированной мебелью, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечено доступом в электронную информационно-образовательную среду ГУАП

## **10. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине**

10.1. Состав фонда оценочных средств приведен в таблице 13

Таблица 13 - Состав фонда оценочных средств для промежуточной аттестации

Вид промежуточной аттестации	Примерный перечень оценочных средств
Экзамен	Список вопросов к экзамену.

10.2. Перечень компетенций, относящихся к дисциплине, и этапы их формирования в процессе освоения образовательной программы приведены в таблице 14.

Таблица 14 – Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Номер семестра	Этапы формирования компетенций по дисциплинам/практикам в процессе освоения ОП
----------------	--

ОПК-11 «владение навыками работы с компьютером как средством получения, обработки и управления информацией»	
1	Информатика
2	Учебная практика по получению первичных профессиональных умений и навыков
4	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности
5	Информационный поиск и извлечение информации
5	Автоматизированная обработка письменного языка
6	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности (педагогическая)
8	Производственная преддипломная практика
ОПК-20 «способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности»	
4	Компьютерные программы в лингвистических исследованиях
5	Информационный поиск и извлечение информации
5	Обработка текстовых массивов
5	Автоматизированная обработка письменного языка
6	Обработка текстовых массивов
6	Основы информационной безопасности
ПК-19 «способность работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности»	
2	Учебная практика по получению первичных профессиональных умений и навыков
3	Основы теоретической и прикладной лингвистики
4	Компьютерные программы в лингвистических исследованиях
4	Основы теоретической и прикладной лингвистики
4	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности
4	Информационные технологии в лингвистике
5	Обработка текстовых массивов
5	Технологии обработки текста и звучащей речи
5	Информационный поиск и извлечение информации
5	Базы данных
5	Автоматизированная обработка письменного языка

6	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности (педагогическая)
6	Обработка текстовых массивов
7	Информационные языки
7	Автоматизированный перевод
8	Информационные языки
8	Производственная преддипломная практика
ПК-20 «владение методами формального и когнитивного моделирования естественного языка и методами создания метаязыков»	
4	Компьютерные программы в лингвистических исследованиях
4	Информационные технологии в лингвистике
5	Информационный поиск и извлечение информации
5	Автоматизированная обработка письменного языка
6	Когнитивные технологии в прикладной лингвистике
8	Производственная преддипломная практика
ПК-21 «владение основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов»	
3	Математика. Математический анализ
4	Компьютерные программы в лингвистических исследованиях
4	Математика. Теория вероятностей и математическая статистика
4	Математические методы в лингвистике
4	Информационные технологии в лингвистике
4	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности
5	Автоматизированная обработка письменного языка
5	Информационный поиск и извлечение информации
6	Производственная практика по получению профессиональных умений и опыта профессиональной деятельности (педагогическая)
6	Машинное обучение
8	Производственная преддипломная практика
ПК-22 «владение стандартными способами решения основных типов задач в области лингвистического обеспечения информационных и других прикладных систем»	
4	Компьютерные программы в лингвистических исследованиях
4	Информационные технологии в лингвистике
5	Автоматизированная обработка письменного языка
5	Информационный поиск и извлечение информации
7	Автоматизированный перевод

7	Основы проектирования информационных систем
7	Информационные языки
8	Информационные языки
8	Производственная преддипломная практика
ПК-25 «владение основами современных методов научного исследования, информационной и библиографической культурой»	
3	Теоретические основы современного русского языка
4	Основы теории языка. Фонетика и фонология
5	Лингвистика текста
5	Основы теории языка. Морфология и синтаксис
5	Автоматизированная обработка письменного языка
5	Информационный поиск и извлечение информации
6	Основы теории языка. Семантика и прагматика
8	Производственная преддипломная практика
ПК-26 «владение стандартными методиками поиска, анализа и обработки материала исследования»	
3	Теоретические основы современного русского языка
4	Основы теории языка. Фонетика и фонология
5	Лингвистика текста
5	Основы теории языка. Морфология и синтаксис
5	Автоматизированная обработка письменного языка
5	Информационный поиск и извлечение информации
6	Основы теории языка. Семантика и прагматика
8	Производственная преддипломная практика

10.3. В качестве критериев оценки уровня сформированности (освоения) у обучающихся компетенций применяется шкала модульно–рейтинговой системы университета. В таблице 15 представлена 100–балльная и 4–балльная шкалы для оценки сформированности компетенций.

Таблица 15 –Критерии оценки уровня сформированности компетенций

Оценка компетенции		Характеристика сформированных компетенций
100-балльная шкала	4-балльная шкала	
$85 \leq K \leq 100$	«отлично» «зачтено»	<ul style="list-style-type: none"> <li>- обучающийся глубоко и всесторонне усвоил программный материал;</li> <li>- уверенно, логично, последовательно и грамотно его излагает;</li> <li>- опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления;</li> <li>- умело обосновывает и аргументирует выдвигаемые им идеи;</li> <li>- делает выводы и обобщения;</li> <li>- свободно владеет системой специализированных понятий.</li> </ul>
$70 \leq K \leq 84$	«хорошо» «зачтено»	<ul style="list-style-type: none"> <li>- обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы;</li> <li>- не допускает существенных неточностей;</li> <li>- увязывает усвоенные знания с практической деятельностью направления;</li> <li>- аргументирует научные положения;</li> </ul>

		<ul style="list-style-type: none"> <li>- делает выводы и обобщения;</li> <li>- владеет системой специализированных понятий.</li> </ul>
$55 \leq K \leq 69$	«удовлетворительно» «зачтено»	<ul style="list-style-type: none"> <li>- обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы;</li> <li>- допускает несущественные ошибки и неточности;</li> <li>- испытывает затруднения в практическом применении знаний направления;</li> <li>- слабо аргументирует научные положения;</li> <li>- затрудняется в формулировании выводов и обобщений;</li> <li>- частично владеет системой специализированных понятий.</li> </ul>
$K \leq 54$	«неудовлетворительно» «не зачтено»	<ul style="list-style-type: none"> <li>- обучающийся не усвоил значительной части программного материала;</li> <li>- допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении;</li> <li>- испытывает трудности в практическом применении знаний;</li> <li>- не может аргументировать научные положения;</li> <li>- не формулирует выводов и обобщений.</li> </ul>

#### 10.4. Типовые контрольные задания или иные материалы:

##### 1. Вопросы (задачи) для экзамена (таблица 16)

Таблица 16 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена
	<ol style="list-style-type: none"> <li>1. Обсуждение понятия «информация».</li> <li>2. Краткая история информатики как науки о семантической информации.</li> <li>3. Многоуровневая модель информации.</li> <li>4. Соотношение между понятиями «данные», «информация» и «знания».</li> <li>5. Семиотика и проблема восприятия информации.</li> <li>6. Знак, концепт, денотат – треугольник Фреге.</li> <li>7. Связь между многоуровневой моделью информации и семиотическими моделями.</li> <li>8. История возникновения тезаурусов и онтологий.</li> <li>9. Типы тезаурусов.</li> <li>10. Многообразие определения онтологий.</li> <li>11. Определение документа как носителя информации.</li> <li>12. Типы классификации документов.</li> <li>13. Форматы представления документов (текстовых, графических, аудио, видео...) и их особенности.</li> </ol>

	<p>14. Цели и задачи аналитико-синтетической переработки документов.</p> <p>15. Основные этапы аналитико-синтетической переработки документов.</p> <p>16. Индексация документов.</p> <p>17. Сравнение наиболее распространенных классификаторов (УДК, ГРНТИ, MSC).</p> <p>18. Понятие метаданных.</p> <p>19. Типы и иерархия метаданных.</p> <p>20. Стандарт Dublin Core.</p> <p>21. Особенности аналитико-синтетической переработки электронных документов.</p> <p>22. Библиометрика: исследование количественных характеристик документов.</p> <p>23. Понятие об импакт-факторе.</p> <p>24. Основные библиометрические индексы: индекс цитирования, индекс Хирша и его модификации.</p> <p>25. Вебометрика: особенности количественного анализа интернет-ресурсов. 26. Применение теории графов в вебометрических исследованиях.</p>
--	--

## 2. Вопросы (задачи) для зачета / дифференцированного зачета (таблица 17)

Таблица 17 – Вопросы (задачи) для зачета / дифф. зачета

№ п/п	Перечень вопросов (задач) для зачета / дифференцированного зачета
	Учебным планом не предусмотрено

## 3. Темы и задание для выполнения курсовой работы / выполнения курсового проекта (таблица 18)

Таблица 18 – Примерный перечень тем для выполнения курсовой работы / выполнения курсового проекта

№ п/п	Примерный перечень тем для выполнения курсовой работы / выполнения курсового проекта
	Учебным планом не предусмотрено

## 4. Вопросы для проведения промежуточной аттестации при тестировании (таблица 19)

Таблица 19 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов
-------	--



	Учебным планом не предусмотрено
--	---------------------------------

5. Контрольные и практические задачи / задания по дисциплине (таблица 20)

Таблица 20 – Примерный перечень контрольных и практических задач / заданий

№ п/п	Примерный перечень контрольных и практических задач / заданий
	<p>Темы рефератов</p> <ol style="list-style-type: none"> <li>1. Форматы представления текстовых документов.</li> <li>2. Форматы представления графических документов (векторная и растровая графика).</li> <li>3. Форматы представления звуковых документов.</li> <li>4. Форматы представления видеодокументов.</li> <li>5. Информационно-поисковые системы. Их виды и особенности.</li> <li>6. Сравнение поисковых систем сети Интернет.</li> <li>7. Системы научно-технической информации.</li> <li>8. Классификационные информационно-поисковые языки.</li> <li>9. Дескрипторные информационно-поисковые языки.</li> <li>10. УДК как информационно-поисковый язык.</li> <li>11. Классификатор ГРНТИ как информационно-поисковый язык.</li> <li>12. Стандарт метаданных Dublin Core.</li> <li>13. Сравнительный анализ понятий «релевантность» и «пертинентность».</li> <li>14. Особенности поиска документов «по аналогии».</li> <li>15. Лингвистическое обеспечение информационно-поисковых систем.</li> <li>16. Построение поисковых индексов.</li> <li>17. Эволюция понятий «тезаурус» и «онтология».</li> <li>18. Установление меры близости между объектами.</li> <li>19. Сравнительный обзор алгоритмов кластеризации текстовых документов.</li> <li>20. WordNet Wiki и автоматическая генерация тегов для интернет-контента.</li> <li>21. Семантическая поисковая система.</li> <li>22. Генерация текстов на основе семантического представления.</li> <li>23. SEO и методы борьбы.</li> </ol>

10.5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и / или опыта деятельности, характеризующих этапы формирования компетенций, содержатся в Положениях «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программам высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

## 11. Методические указания для обучающихся по освоению дисциплины

Целью дисциплины является ознакомление с базовыми понятиями теории информационного поиска, основными принципами организации информационно-поисковых систем и алгоритмами аналитико-синтетической переработки документов.

Задачами дисциплины являются:

- определить базовые понятия теории информационного поиска, показав их тесную связь с семиотикой;
- изучить основные технологии построения информационно-поисковых систем;
- ознакомиться с технологиями построения тезаурусов и онтологий; изучить основные алгоритмы аналитико-синтетической переработки документов.

**Методические указания для обучающихся по прохождению лабораторных работ (если предусмотрено учебным планом по данной дисциплине)**

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач у обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

### Задание и требования к проведению лабораторных работ

. Задание №1

- . Представьте как технологический процесс подготовку реферата, проверку электронной почты с ответами на письма, написание курсовой работы. Какие алгоритмические структуры при этом использовались? Приведите основные уровни информационной

технологии. Проиллюстрируйте их на примере из предыдущего пункта. Докажите, что приведенные примеры отвечают требованиям к технологии. Приведите инструментарий для рассматриваемой технологии.

. Задание №2

Проведите морфологический анализ заданного текста на естественном языке. Для всех слов в тексте укажите их лексико – грамматический класс. Отредактируйте текст, внося разметку о лексико – грамматических классах (используя аббревиатуры классов и/или выделив элементы разметки наклоном или иным образом). Рекомендуется использовать собственную разметку, расшифровку которой привести после текста. Составьте словарь на основании заданного текста. Составьте словарь словоформ для пяти существительных и трех глаголов заданного текста. Составьте словарь устойчивых словосочетаний для данного текста с указанием их значения (пара «словосочетание» - «значение»). Составьте биграмную модель встречаемости лексико-грамматических классов на рассматриваемом тексте. Проанализируйте текст на возможное наличие продукционных правил, исходя из предыдущего этапа и общих представлений о языке. Сделайте выводы.

. Задание № 3

- . Проведите автоматический перевод фрагмента текста, используя несколько доступных переводчиков. Переведите тот же фрагмент текста вручную. Укажите фрагменты, отличающиеся в переводе различными системами. Объясните, почему они возникли. Укажите, какие особенности систем автоматического перевода приводят к возникновению ошибок каждого рода. Приведите пример иерархической памяти переводов на выбранный вами язык для словаря, созданного в одном из предыдущих разделов.

. Задание №4

Создайте в MS Excel терминологический словарь. Конвертируйте созданный файл в формат SDL Trados. Изучите полученные в итоге файлы. Опишите структуру XML – файла и LOG – файла.

- . Создайте в MultiTerm Desktop базу терминов на основании сконвертированных данных. Загрузите термины в созданную базу. Создайте в Trados собственную пустую базу переводов, собственный проект с новой памятью переводов. Подключите в проект созданные в предыдущей работе терминологические базы. Сохраните проект на диск (flash – карту). Познакомьтесь с полученными файлами проекта и их назначением. Закройте проект, смените его расположение и откройте его заново. Проверьте настройки проекта. Установите точность совпадений сегментов – 70%.
- . Выполните перевод заданного текста. При выполнении перевода пополните словарь терминов. В режиме Reports просмотрите отчет по переводу (Pre-Translate Files Report и Analyze Files Report). Сделайте выводы. Создайте новый мультиязыковой проект в SDL Trados с новой памятью переводов. Выполните перевод на оба целевых языка. При выполнении перевода пополните словарь терминов. Откройте режим Translation Memory. Экспортируйте (аналогично импорту) из ранее созданных баз переводов данные в tmx файлы. Создайте общую новую базу перевода с полями для пометки проекта. Импортируйте в нее данные из двух предыдущих переводов, указав,

к какому переводу какой сегмент относится. «Почистите» полученную базу перевода от слишком длинных или некорректных сегментов. Экспортируйте полученную базу перевода в tmx файл. Ознакомьтесь с его структурой. Самостоятельно изучите тэги и разметку tmx файла для редактирования вручную. Создайте подпроект для перевода. Закройте проект. Откройте подпроект для редактирования. Откройте файл проекта для редактирования. Отредактируйте перевод, внося в него необходимую редакторскую правку. Создайте подпроект для переводчика. Закройте текущий подпроект. Откройте подпроект для перевода. Откройте файл проекта для переводчика. Подтвердите или отклоните изменения, внесенные редактором. Создайте файл результата. Подготовьте отчет Verify Files. Финализируйте проект.

### **Структура и форма отчета о лабораторной работе**

Отчёт по лабораторной работе оформляется индивидуально каждым студентом, выполнившим необходимые (независимо от того, выполнялся ли эксперимент индивидуально или в составе группы студентов). Страницы отчёта следует пронумеровать (титульный лист не нумеруется, далее идет страница 2 и т.д.).

Титульный лист отчёта должен содержать фразу: «Отчёт по лабораторной работе «Название работы», чуть ниже: Выполнил студент группы (номер группы) (Фамилия, инициалы)». Внизу листа следует указать текущий год. Например, Отчёт по лабораторной работе № (номер работы) «Введение в спектральный анализ», Выполнил студент группы 4332 Иванов И.И.

Вторая страница текста, следующая за титульным листом, должна начинаться с пункта: Цель работы. Отчёт, как правило, должен содержать следующие основные разделы: 1. Цель работы; 2. Теоретическая часть; 3. Программное обеспечение, используемое в работе; 4. Результаты; 5. Выводы.

В случае необходимости в конце отчёта приводится перечень литературы.

### **Требования к оформлению отчета о лабораторной работе**

Теоретическая часть должна содержать минимум необходимых теоретических сведений о предметной области. Не следует копировать целиком или частично методическое пособие (описание) лабораторной работы или разделы учебника.

В разделе Программное обеспечение необходимо описать, с помощью каких инструментальных средств и каким образом были разработаны модели и получены результаты. Рисунки, блок-схемы, описание модели и её особенностей, необходимость отладки – все это должно быть представлено в указанном разделе.

Раздел Результаты включает в себя скриншоты программного приложения, полученные при выполнении лабораторной работы. Рисунки, графики и таблицы нумеруются и подписываются заголовками.

Выводы не должны быть простым перечислением того, что сделано. Здесь важно отметить, какие новые знания о предмете исследования были получены при выполнении работы, к чему привело обсуждение результатов, насколько выполнена заявленная цель работы.

Выводы по работе каждый студент делает самостоятельно. В случае необходимости в конце отчёта приводится Список литературы, использованной при подготовке к работе.

В тексте отчёта делаются краткие ссылки на литературу (учебники, справочники, иные источники...) номером в квадратных скобках, напр., [1]. Литературные источники нумеруются по мере их появления в тексте отчёта. В конце отчёта даётся их подробный список. На все источники списка литературы должны быть ссылки в тексте отчёта, там, где это необходимо.

При сдаче отчёта преподаватель может сделать устные и письменные замечания, задать дополнительные вопросы. Все ответы на дополнительные вопросы, обсуждения выполняются студентом на отдельных листах, включаемых в отчёт (при этом в тексте

основного отчёта делается сноска или другой значок, которому будет соответствовать новый материал). При этом письменные замечания преподавателя должны остаться в тексте для ясности динамики работы над отчётом.

Объём отчёта должен быть оптимальным для понимания того, что и как сделал студент, выполняя работу. Обязательные требования к отчёту включают общую и специальную грамотность изложения, а также аккуратность оформления.

После приёма преподавателем отчёт хранится на кафедре.

### **Методические указания для обучающихся по прохождению самостоятельной работы**

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Для обучающихся по заочной форме обучения, самостоятельная работа может включать в себя контрольную работу.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются:

- учебно-методический материал по дисциплине.

–

### **Методические указания для обучающихся по прохождению промежуточной аттестации**

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

- экзамен – форма оценки знаний, полученных обучающимся в процессе изучения всей дисциплины или ее части, навыков самостоятельной работы, способности применять их для решения практических задач. Экзамен, как правило, проводится в период экзаменационной сессии и завершается аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Система оценок при проведении промежуточной аттестации осуществляется в соответствии с требованиями Положений «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программам высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

## Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой