

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования «Новосибирский национальный исследовательский
государственный университет» (Новосибирский государственный университет, НГУ)

Факультет информационных технологий

СОГЛАСОВАНО

Декан ФИТ НГУ



М.М. Лаврентьев

«03» июля 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Теоретические основы обработки информации

Направление подготовки: 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
Направленность (профиль): Программная инженерия и компьютерные науки

Форма обучения: очная

Год обучения: 3, семестр: 6

№	Вид деятельности	Семестр
		6
1	Лекции, час.	32
2	Практические занятия, час.	32
3	Лабораторные занятия, час.	
4	Занятий в контактной форме без учета промежуточной аттестации, час, из них	66
5	в электронной форме, час.	
6	из них аудиторных занятий, час.	64
7	из них в активной и интерактивной форме, час.	64
8	консультаций, час.	2
9	Самостоятельная работа, час.	76
10	в том числе на выполнение письменных работ, час	36
11	Форма аттестации (экзамен, зачет, дифференцированный зачет), час	Э 2
12	Всего зачетных единиц ¹	4

Новосибирск 2019

¹ С учетом выделенных часов на промежуточную аттестацию

Рабочая программа дисциплины составлена на основании федерального государственного образовательного стандарта (ФГОС) высшего образования - бакалавриат по направлению подготовки 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА.

Федеральный государственный образовательный стандарт (ФГОС) высшего образования - бакалавриат по направлению подготовки 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА введен в действие приказом Минобрнауки от 19.09.2017 № 929.

Место дисциплины в структуре учебного плана: Блок 1 Дисциплины (модули); часть, формируемая участниками образовательных отношений, дисциплина по выбору

Рабочая программа дисциплины утверждена решением Ученого совета факультета информационных технологий от 02.07.2019, протокол № 75.

Программу разработал:

доцент кафедры систем информатики ФИТ,
кандидат физико-математических наук

Т.В. Батура

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук

М.М. Лаврентьев

Ответственный за образовательную программу:

доцент кафедры систем информатики ФИТ,
кандидат технических наук

А.А. Романенко

Аннотация к рабочей программе дисциплины «Теоретические основы обработки информации»

Дисциплина «Теоретические основы обработки информации» реализуется в рамках образовательной программы высшего образования – программы бакалавриата 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА, направленность (профиль): ПРОГРАММНАЯ ИНЖЕНЕРИЯ И КОМПЬЮТЕРНЫЕ НАУКИ по очной форме обучения на русском языке.

Место в образовательной программе: Дисциплина «Теоретические основы обработки информации» развивает знания, умения и навыки, сформированные у обучающихся по результатам изучения следующих дисциплин: «Информатика», «Основы объектно-ориентированного программирования».

Данный курс является базовым для освоения дисциплины «Интеллектуальный анализ данных», «Языки и системы искусственного интеллекта».

Дисциплина «Теоретические основы обработки информации» реализуется в 6 семестре в рамках части, формируемой участниками образовательных отношений, дисциплин (модулей) Блока 1 и является дисциплиной по выбору

Дисциплина «Теоретические основы обработки информации» направлена на формирование компетенций:

Способен разрабатывать компоненты системных программных продуктов (ПКС-2), в части следующих индикаторов достижения компетенции:

ПКС-2.6 Знать: алгоритмы, модели и объекты исследования в предметной области.

Перечень основных разделов дисциплины:

При освоении дисциплины студенты выполняют следующие виды учебной работы: лекции, практические занятия, консультации, самостоятельная работа.

Дисциплина включает в себя элементы теории кодирования и теории сложности алгоритмов, элементы обработки изображений и сигналов, алгоритмы поиска объектов на изображениях, компрессии видео, распознавание речи, классические ортогональные и современные вейвлет-преобразования. Дисциплина также охватывает подходы автоматизированной обработки текстов: формальные методы анализа текстов, алгоритмы семантического поиска и извлечения информации, особенности построения тезаурусов, знакомство с основами корпусной лингвистики.

Основные темы:

Раздел 1. Базовые понятия и основные теоремы теории информации

Раздел 2. Обработка изображений и сигналов

Раздел 3. Общие принципы и задачи построения систем автоматической обработки текстовой информации

Раздел 4. Семантический анализ текстов

При проведении лекций и практических занятий могут применяться дистанционные образовательные технологии с использованием платформы Google meet.

Общий объем дисциплины – 4 зачетные единицы (144 часа).

Правила аттестации по дисциплине. Текущий контроль по дисциплине осуществляется на практических занятиях на основании оценки за портфолио (защита результатов проведенной работы по основным разделам дисциплины). По результатам защиты портфолио выставляется оценка «зачтено» или «не зачтено».

Промежуточная аттестация по дисциплине «Теоретические основы обработки информации» проводится по завершению периода ее освоения (семестра). Промежуточная аттестация по дисциплине включает 2 этапа:

- 1) портфолио (контрольные работы, реферат);
- 2) экзамен.

Оценка «зачтено» за портфолио является необходимым условием для прохождения промежуточной аттестации. Результаты промежуточной (итоговой по дисциплине) аттестации определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Решение об окончательной оценке принимается по результатам 2 этапа (экзамен).

Оценка «отлично» соответствует продвинутому уровню сформированности компетенции.

Оценка «хорошо» соответствует базовому уровню сформированности компетенции.

Оценка «удовлетворительно» соответствует пороговому уровню сформированности компетенции.

Оценка «неудовлетворительно» выставляется при неудовлетворительном прохождении одного или двух этапов промежуточной аттестации.

Учебно-методическое обеспечение дисциплины.

<https://sites.google.com/site/tatianavbatura/lectures>

1. Внешние требования к дисциплине

Таблица 1.1

Компетенция ПКС-2 Способен разрабатывать компоненты системных программных продуктов, в части следующих индикаторов достижения компетенции:	
ПКС-2.6	Знать: алгоритмы, модели и объекты исследования в предметной области.

2. Требования к результатам освоения дисциплины

Таблица 2.1

Результаты изучения дисциплины по уровням освоения (иметь представление, знать, уметь, владеть)	Формы организации занятий		
	Лекции	Практики / семинары	Самостояте льная работа
ПКС-2.6 Знать: алгоритмы, модели и объекты исследования в предметной области.			
1. Знать основные методы и алгоритмы классификации и кластеризации данных	+	+	+
2. Знать основную идею организации памяти с параллельным доступом к информации; знать формулировку базовой задачи обнаружения объектов на изображении	+	+	+
3. Знать общие принципы построения автоматизированных систем обработки текстов; основные подходы к исследованию семантики текстов	+	+	+
4. Знать методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов	+	+	+

3. Содержание и структура учебной дисциплины

Таблица 3.1

Темы лекций	Активные формы, час.	Часы	Ссылки на результаты обучения
Семестр: 6			
1. Базовые понятия и основные теоремы теории информации – Основные понятия и определения теории информации Шеннона. – Признак Крафта. – Теорема о нижней оценке. – Связь исправления ошибок с избыточностью и энтропией. – Модель Бернулли и теорема о верхней оценке. – Статистические модели источников сообщений. – Код Хемминга. – Элементы криптографии. (См. [1], [8])	6	6	2, 4

<p>2. Обработка изображений и сигналов. Математические основы алгоритмов обработки изображений и сигналов</p> <ul style="list-style-type: none"> – Системы символьных преобразований, алгоритмические основы. – Классификация типов изображений и алгоритмов обработки. – Меры близости изображений. – Цветовое пространство и цветовая константа. – Подавление шумов с помощью масок. – Выделение контуров. – Градиентный метод. – Комбинаторный метод. – Метод Слободы. – Двумерные унитарные преобразования. – Косинусные и синусные преобразования. – Преобразование на основе матриц Адамара. – Преобразование Хаара. – Дискретное вейвлет-преобразование. – Приложения к кодированию звуков и видео. <p>(См. [7], [9])</p>	5	5	1, 2, 4
<p>3. Организация памяти с параллельным доступом к информации</p> <ul style="list-style-type: none"> – Параллельная обработка информации. – Загруженность и асимптотическая загруженность устройств. – Формулы для загруженности и ускорения для системы конвейерных устройств, работающих с векторами в режиме зацепления; условие на длины векторов. – Общие принципы организации памяти с параллельным доступом к информации, некоторые перестановки и их свойства, методы адресации. – Нестандартные методы обработки информации. – Применение размытой логики в распознавании и системах управления. <p>(См. [9])</p>	6	6	2, 4
<p>4. Общие принципы и задачи построения систем автоматической обработки текстовой информации</p> <ul style="list-style-type: none"> – Основные понятия математической лингвистики. – Общие принципы построения систем автоматизированной обработки текстов. – Деревья составляющих. – Деревья зависимостей. – Проблемы синтаксического анализа. – Синтаксические анализаторы. – Проекты АОТ и Link Grammar Parser. – Устройство поисковых систем. – Порождающие грамматики. – Распознающие грамматики. <p>(См. [2], [3], [5])</p>	8	8	3, 4

<p>5. Семантический анализ текстов. Подходы и методы анализа текстов на семантическом уровне</p> <ul style="list-style-type: none"> – Подход И. Мельчука. – Теоретико-модельный подход. – Теоретико-множественный подход С. Маркуса. – Тезаурусы. – Семантические сети. – Онтологии. – Базы знаний. – Методы классификации текстов. – Методы кластеризации текстов. <p>(См. [4], [6])</p>	3	3	1, 3, 4
<p>6. Практическое применение алгоритмов обработки текстов в различных информационных системах</p> <ul style="list-style-type: none"> – Анализ тональности и тематическая индексация текстов. – Методы определения авторства текстов. – Методы энтропийного кодирования, алгоритм Хаффмана. – Статистические методы атрибуции. – Деревья решений. – Скрытые Марковские модели. – Алгоритм Витерби. – Методы анализа данных из социальных сетей. <p>(См. [5])</p>	4	4	1, 3, 4
Итого	32	32	

Таблица 3.2

Темы практических занятий	Активные формы, час.	Часы	Ссылки на результаты обучения	Учебная деятельность
Семестр: 6				
<p>1. Базовые понятия и основные теоремы теории информации</p> <ul style="list-style-type: none"> – Кодирование, асимптотические формулы. – Спектры, префиксные коды. – Энтропия. – Оптимальное кодирование. – Приложения статистических моделей в лингвистике, генетике, при исследовании музыки. – Коды открытого ключа. <p>(См. [1], [8])</p>	6	6	2, 4	Разбор теоретической темы, представленной на лекции, решение задач
<p>2. Обработка изображений и сигналов. Математические основы алгоритмов обработки изображений и сигналов</p>	5	5	1, 2, 4	Разбор теоретической темы, представленной на лекции, решение задач

<ul style="list-style-type: none"> – Обработка изображений и области ее применения. – Дискретизация и квантование функции яркости. – Предварительная обработка изображений. – Коррекция яркости и контрастности. – Обнаружение объектов на изображении. – Системы поиска и сопровождения целей. (См. [7], [9])				
<p>3. Организация памяти с параллельным доступом к информации</p> <ul style="list-style-type: none"> – Распараллеливание, как метод повышения эффективности обработки информации. – Общие математические соотношения, характеризующие вычисления на векторно-конвейерных ЭВМ. – Приложения принципов организации памяти с параллельным доступом к численным методам и в обработке изображений. – Нейрокомпьютерный подход. (См. [9])	6	6	2, 4	Разбор теоретической темы, представленной на лекции, решение задач
<p>4. Общие принципы и задачи построения систем автоматической обработки текстовой информации</p> <ul style="list-style-type: none"> – Системы автоматизированной обработки текстов. – Методы задания синтаксической структуры предложений. – Проект Universal Networking Language. – Морфологический и синтаксический анализ в поисковых системах. – Регулярные выражения и конечные автоматы. (См. [2], [3], [5])	8	8	3, 4	Разбор теоретической темы, представленной на лекции, решение задач
<p>5. Семантический анализ</p>	3	3	1, 3, 4	Разбор теоретической темы,

<p>текстов</p> <p>Подходы и методы анализа текстов на семантическом уровне</p> <ul style="list-style-type: none"> – Методы теоретического исследования семантики текстов. – Представление знаний для компьютерной обработки. – Методы измерения семантического расстояния. – Машинное обучение с учителем. – Машинное обучение без учителя. <p>(См. [4], [6])</p>				представленной на лекции, решение задач
<p>6. Практическое применение алгоритмов обработки текстов в различных информационных системах</p> <ul style="list-style-type: none"> – Графовые модели социальных сетей. – Анализ отношения предпочтения. – Моделирование психологических операций. – Системы извлечения информации. – Распознавание именованных элементов. – Разрешение кореференций. <p>(См. [5])</p>	4	4	1, 3, 4	Разбор теоретической темы, представленной на лекции, решение задач
Итого	32	32		

4. Самостоятельная работа студентов

Таблица 4.1

№	Виды самостоятельной работы	Ссылки на результаты обучения	Часы на выполнении	Часы на консультации
Семестр: 6				
1	Изучение разделов дисциплины по учебной литературе, в том числе вопросов, не освещаемых на лекциях	1, 2, 3, 4	8	
	Изучение предлагаемых теоретических разделов в соответствии с настоящей Программой. Учебно-методические материалы по дисциплине «Теоретические основы обработки информации» выложены на странице курса в сети Интернет			
2	Подготовка к контрольной работе	1, 2, 3, 4	14	

	Методические рекомендации по подготовке к контрольной работе представлены в приложении к рабочей программе дисциплины.		
3	Подготовка реферата	1, 2, 3, 4	30
	Обучающийся готовит реферат по теме из списка, указанного в рабочей программе дисциплины. Методические рекомендации по подготовке реферата представлены в приложении к рабочей программе дисциплины.		
4	Подготовка к экзамену	1, 2, 3, 4	24
	Подготовка к экзамену по вопросам, представленным в фонде оценочных средств, являющихся приложением к рабочей программе дисциплины.		
	Итого		76
			2

5. Образовательные технологии

В ходе реализации учебного процесса по дисциплине проводятся лекции и практические занятия.

При проведении лекций и практических занятий могут применяться дистанционные образовательные технологии с использованием платформы Google meet. При проведении практических занятий студенты подключаются к онлайн сессии. На занятиях разбираются теоретические темы, представленные на лекции, и формулировки практических заданий. Для сдачи выполненного задания студент включает демонстрацию экрана, показывает результаты, обосновывает решение, отвечает на вопросы преподавателя.

Применение дистанционных образовательных технологий позволяет обеспечить эффективное взаимодействие преподавателя со студентами в различных ситуациях, когда преподаватель или студенты не могут территориально присутствовать в помещениях НГУ в силу объективных и уважительных субъективных факторов. Дистанционные технологии могут применяться, в том числе и в «смешанном формате»: при проведении лекционных и практических занятий преподаватель дополнительно подключает дистанционный сервис, позволяющий участвовать в занятии студентам, которые по каким-либо причинам не могут присутствовать в классе.

В ходе реализации учебного процесса по дисциплине применяются следующие интерактивные формы обучения (таблица 5.1).

Таблица 5.1

1	Лекция в форме дискуссии	ПКС-2.6
Формируемые умения:		
Знать основные методы и алгоритмы классификации и кластеризации данных.		
Знать основную идею организации памяти с параллельным доступом к информации; знать формулировку базовой задачи обнаружения объектов на изображении.		
Знать общие принципы построения автоматизированных систем обработки текстов; основные подходы к исследованию семантики текстов.		
Знать методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов.		
Краткое описание применения: Представляется теория, обсуждаются идеи и способы решения задач, рекомендованных для практических занятий		
2	Портфолио	ПКС-2.6
Формируемые умения:		
Знать основные методы и алгоритмы классификации и кластеризации данных.		
Знать основную идею организации памяти с параллельным доступом к информации; знать формулировку базовой задачи обнаружения объектов на изображении.		
Знать общие принципы построения автоматизированных систем обработки текстов; основные		

подходы к исследованию семантики текстов.

Знать методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов.

Краткое описание применения: бакалавры ведут портфолио (оценки за реферат, за контрольную работу), которое является основой для проведения аттестации по дисциплине

Для организации и контроля самостоятельной работы бакалавров, а также проведения консультаций применяются информационно-коммуникационные технологии (таблица 5.2).

Таблица 5.2

Информирование	Адрес почты – сообщается бакалаврам на первом занятии.
Консультирование	Адрес почты – сообщается бакалаврам на первом занятии.
Контроль	Адрес почты – сообщается бакалаврам на первом занятии.
Размещение учебных материалов	https://sites.google.com/site/tatianavbatura/lectures

6. Правила аттестации студентов по учебной дисциплине

По дисциплине «Теоретические основы обработки информации» проводится текущая и промежуточная аттестация (итоговая по дисциплине).

Текущий контроль по дисциплине осуществляется на практических занятиях на основании оценки за портфолио (защита результатов проведенной работы по основным разделам дисциплины). По результатам защиты портфолио выставляется оценка «зачтено» или «не зачтено».

Промежуточная аттестация по дисциплине «Теоретические основы обработки информации» проводится по завершению периода ее освоения (семестра). Промежуточная аттестация по дисциплине включает 2 этапа:

- 1) портфолио (контрольные работы, реферат);
- 2) экзамен.

Необходимым условием для прохождения промежуточной аттестации является оценка «зачтено» по результатам выполненного портфолио. Для оценивания портфолио студенту необходимо сдать все работы, входящие в структуру портфолио.

Результаты промежуточной (итоговой по дисциплине) аттестации определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Решение об окончательной оценке принимается по результатам 2 этапа (экзамен).

В таблице 6.1 представлено соответствие форм аттестации заявляемым требованиям к результатам освоения дисциплины.

Таблица 6.1

Коды компетенций ФГОС	Результаты обучения	Формы аттестации	
		1 этап - портфолио	2 этап - экзамен
ПКС-2	ПКС-2.6 Знать: алгоритмы, модели и объекты	+	+

Требования к структуре и содержанию портфолио, оценочные средства, а также критерии оценки сформированности компетенций и освоения дисциплины в целом, представлены в Фонде оценочных средств, являющемся приложением 1 к настоящей рабочей программе дисциплины.

7. Перечень учебной литературы

1. Батура Т.В. Математическая лингвистика и автоматическая обработка текстов на естественном языке: учеб. пособие / Новосибирский государственный университет. Новосибирск: РИЦ НГУ, 2016. ISBN 978-5-4437-0548-4. 166 с. [Электронный ресурс]. URL: <http://e-lib.nsu.ru/dsweb/Get/Resource-1583/page001.pdf>

8. Перечень ресурсов информационно-коммуникационной сети «Интернет», необходимых для освоения дисциплины

При освоении дисциплины используются следующие ресурсы:

- электронная информационно-образовательная среда НГУ (ЭИОС);
- образовательные интернет-порталы;
- информационно-телекоммуникационная сеть Интернет.

Взаимодействие обучающегося с преподавателем (синхронное и (или) асинхронное) осуществляется через личный кабинет студента в ЭИОС.

Таблица 8.1

№ п/п	Наименование Интернет-ресурса	Краткое описание
1.	The Stanford Natural Language Processing Group. [Electronic resource]. URL: http://nlp.stanford.edu/	Каталог публикаций группы автоматической обработки текстов Стэнфордского университета
2.	ACL Anthology – A Digital Archive of Research Papers in Computational Linguistics. [Electronic resource]. URL: http://anthology.aclweb.org/	Электронный каталог научных статей на английском языке
3.	Международная конференция по компьютерной лингвистике "Диалог". [Электронный ресурс]. URL: http://www.dialog-21.ru	Сборники трудов ежегодной международной конференции по компьютерной лингвистике
4.	Ruscorpora.ru – Национальный корпус русского языка. [Электронный ресурс]. URL: http://www.ruscorpora.ru/	Информационно-справочная система, основанная на собрании текстов на русском языке в электронной форме
5.	AOT.ru – Проект АОТ (Автоматическая Обработка Текста). [Электронный ресурс]. URL: http://aot.ru/	Данные проекта АОТ
6.	Link Grammar Parser. [Electronic resource]. URL: http://www.link.cs.cmu.edu/link/	Документация по грамматике связей, разрабатываемой в университете Карнеги — Меллона
7.	NLPub Каталог ресурсов и инструментов для обработки текстов. [Электронный	Электронный каталог ресурсов и инструментов для обработки текстов

	ресурс]. URL: http://nlpub.ru/	
8.	Каталог продуктов по извлечению информации. [Электронный ресурс]. URL: http://pullenti.ru/CompetitorPage.aspx	Интернет-каталог продуктов по извлечению информации
9.	Каталог лингвистических программ и ресурсов. [Электронный ресурс]. URL: http://www.rvb.ru/soft/catalogue/catalogue.html	Каталог лингвистических программ и ресурсов
10.	Hanneman R., Riddle M. Introduction to social network methods. Riverside, CA: University of California, Riverside, 2005. [Electronic resource]. Published in digital format. URL: http://faculty.ucr.edu/~hanneman/nettext	Книга по методам анализа социальных сетей на английском языке
11.	Российский семинар по оценке методов информационного поиска (РОМИП). [Электронный ресурс]. URL: http://romip.ru/	Материалы семинара по оценке методов информационного поиска
12.	Сайт о сжатии. [Электронный ресурс]. URL: http://www.compression.ru	Электронный архив статей о сжатии данных, изображений и видео
13.	Impuls Imaging GmbH. [Electronic resource]. URL: http://www.impuls-imaging.com	Электронная подборка научных статей по обработке изображений
14.	Steven Bird, Ewan Klein, and Edward Loper. Natural Language Processing with Python. [Electronic resource]. URL: http://www.nltk.org/book/	Книга по автоматической обработке текстов

9. Учебно-методическое и программное обеспечение дисциплины

9.1. Перечень учебно-методических материалов по самостоятельной работе обучающихся

Для обеспечения самостоятельной работы обучающихся при изучении дисциплины используются следующие учебно-методические материалы:

1. Настоящая рабочая программа дисциплины, соответствующие разделы.
2. Учебники, учебные пособия и дополнительные материалы.
3. Перечень ресурсов информационно-коммуникационной сети «Интернет».
4. Методические указания для обучающихся по освоению дисциплины, обеспечивающие самостоятельную работу студента при подготовке к учебным занятиям, выполнении домашних работ, подготовке к контрольным мероприятиям и аттестациям, приведенные в Приложении к настоящей рабочей программе дисциплины.
5. <https://sites.google.com/site/tatianavbatura/lectures>

9.2. Программное обеспечение

Для обеспечения реализации дисциплины используется стандартный комплект программного обеспечения (ПО), включающий регулярно обновляемое лицензионное ПО Windows и MS Office.

Специализированное ПО не требуется.

10. Профессиональные базы данных и информационные справочные системы

1. Полнотекстовые журналы Springer Journals за 1997-2015 г., электронные книги (2005-2016 гг.), коллекция научных биомедицинских и биологических протоколов SpringerProtocols, коллекция научных материалов в области физических наук и инжиниринга SpringerMaterials, реферативная БД по чистой и прикладной математике zbMATH.

2. Электронная библиотека диссертаций Российской государственной библиотеки (ЭБД РГБ)

3. Электронные ресурсы Web of Science Core Collection (Thomson Reuters Scientific LLC.), Journal Citation Reports + ESI

4. БД Scopus (Elsevier)

11. Материально-техническое обеспечение

Для реализации дисциплины используются специальные помещения:

1. Учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля, промежуточной и итоговой аттестации;

2. Помещения для самостоятельной работы обучающихся;

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду НГУ.

Для проведения занятий лекционного типа предлагаются следующие наборы демонстрационного оборудования и учебно-наглядных пособий:

- комплект лекций-презентаций по темам дисциплины;

Таблица 11.1

№	Наименование	Назначение
1	Презентационное оборудование (мультимедиа-проектор, экран, компьютер для управления)	Для проведения лекционных и практических занятий
2	Компьютерный класс (с выходом в Internet)	Для организации самостоятельной работы обучающихся

Материально-техническое обеспечение образовательного процесса по дисциплине для обучающихся из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете».

**Лист актуализации рабочей программы дисциплины
«Теоретические основы обработки информации»**

№	Характеристика внесенных изменений (с указанием пунктов документа)	Дата и № протокола Ученого совета ФИТ	Подпись ответственного
1.	Актуализирован на 2020- 2021 уч.год	22.07.2020 №77	
2.	Актуализирован на 2021 - 2022 уч. год	26.04.2021 №80	
3.	Актуализирован на 2022 - 2023 уч. год	31.08.2022 №84	
4.	Дополнено приложение дистанционных образовательных технологий п. 5	30.01.2023 №90	

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования «Новосибирский национальный исследовательский
государственный университет» (Новосибирский государственный университет, НГУ)

Факультет информационных технологий

СОГЛАСОВАНО

Декан ФИТ НГУ

М.М. Лаврентьев

«03» июля 2019 г.



**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ
по дисциплине Теоретические основы обработки информации**

Направление подготовки: 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Направленность (профиль): Программная инженерия и компьютерные науки

Квалификация: бакалавр

Форма обучения: очная

Год обучения: 3, семестр 6

Форма аттестации	Семестр
Экзамен	6

Новосибирск 2019

Фонд оценочных средств промежуточной аттестации по дисциплине является **Приложением 1** к рабочей программе дисциплины «Теоретические основы обработки информации», реализуемой в рамках образовательной программы высшего образования – программы бакалавриата 09.03.01 Информатика и вычислительная техника, направленность (профиль): Программная инженерия и компьютерные науки

Фонд оценочных средств промежуточной аттестации по дисциплине утвержден решением ученого совета факультета информационных технологий, протокол № 75 от 02.07.2019.

Разработчики:

доцент кафедры систем информатики ФИТ,
кандидат физико-математических наук



Т.В.Батура

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук



М.М. Лаврентьев

Ответственный за образовательную программу:
доцент кафедры систем информатики ФИТ,
кандидат технических наук



А.А. Романенко

1. Содержание и порядок проведения промежуточной аттестации по дисциплине

1.1. Общая характеристика содержания промежуточной аттестации

Промежуточная аттестация по дисциплине «Теоретические основы обработки информации» проводится по завершению периода освоения образовательной программы (семестра) для оценки сформированности компетенций в части следующих индикаторов достижения компетенции (таблица П1.1).

Таблица П1.1

Код	Компетенции, формируемые в рамках дисциплины «Теоретические основы обработки информации»	Семестр 6	
		Портфолио	Экзамен
	ПКС-2 Способен разрабатывать компоненты системных программных продуктов		
ПКС-2.6	Знать: алгоритмы, модели и объекты исследования в предметной области	+	+

Промежуточная аттестация включает 2 этапа:

1. Портфолио.
2. Устный экзамен.

Все компетенции, формируемые в рамках дисциплины, оцениваются как через портфолио, так и на устном экзамене.

Тематика вопросов к экзамену соответствует избранным разделам (темам) дисциплины «Теоретические основы обработки информации»:

Базовые понятия и основные теоремы теории информации

Обработка изображений и сигналов

Математические основы алгоритмов обработки изображений и сигналов

Организация памяти с параллельным доступом к информации

Общие принципы и задачи построения систем автоматической обработки текстовой информации

Семантический анализ текстов

Подходы и методы анализа текстов на семантическом уровне

Практическое применение алгоритмов обработки текстов в различных информационных системах.

1.2. Порядок проведения промежуточной аттестации по дисциплине

Экзамен проводится в устной форме, в аудитории, студентам разрешено пользоваться бумагой для записей и авторучкой. Справочной, учебной и другой литературой пользоваться не разрешается. Использование электронных устройств (телефоны, любые виды компьютеров, т.д.) запрещено.

2. Требования к структуре и содержанию фонда оценочных средств промежуточной аттестации по дисциплине

Перечень оценочных средств, применяемых на каждом этапе проведения промежуточной аттестации по дисциплине, представлен в таблице П1.2.

Таблица П1.2

№ п/п	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
Этап 1 - портфолио			
1.	Портфолио	Целевая подборка работ студента, раскрывающая его индивидуальные образовательные достижения в одной или нескольких учебных дисциплинах.	Требования к структуре и содержанию портфолио
Этап 2 – Экзамен			
2	Экзаменационный билет	Комплекс вопросов	Список теоретических вопросов

2.1. Требования к структуре и содержанию оценочных средств аттестации в 6 семестре

Текущая аттестация по дисциплине «Теоретические основы обработки информации» проводится в форме портфолио.

Состав портфолио: задания, контрольная работа, реферат.

Реферат

В течение семестра обучающийся должен выполнить реферат по выбранной теме.

Для подготовки к сдаче реферата необходимо получить тему из Перечня тем для рефератов у преподавателя.

Работа над рефератом начинается с выбора исходного материала, в качестве которого могут быть печатные издания, источники из интернета. После анализа материала составляется краткое оглавление по теме. Затем следует последовательно скомпоновать содержание реферата в соответствии с оглавлением. Помимо текстовой части реферат может включать табличный материал, рисунки, если это улучшает качество изложения. В конце изложения приводится список использованной литературы и ссылки на материалы из интернета, если это имеет место. Объем реферата должен быть в пределах от 3 до 8 листов при междустрочном интервале 1,25. Причем в указанный объем не входят титульный лист, оглавление, список использованной литературы.

К реферату предъявляются следующие требования:

- содержание реферата должно соответствовать теме;
- объем реферата должен быть в пределах от 3 до 8 листов, причем в указанный объем не входят титульный лист, оглавление, список использованной литературы;
- реферат должен быть выполнен в печатном виде в формате "DOC";
- реферат в печатном оформлении должен иметь шрифт Times New Roman 12, междустрочный интервал 1,25;
- титульный лист для рефератов выполняется стандартным способом, т.е. должен содержать наименование учебного заведения, факультета, темы реферата, Ф.И.О. исполнителя, Ф.И.О. преподавателя, год

Качество выполнения оценивается по степени соответствия содержания реферата теме, полноте и глубине охвата, четкости и ясности изложения материала.

Реферат оформляют печатным способом, с оглавлением и титульным листом.

Сдача реферата на проверку не позднее 15-ой недели учебного семестра и возможна в трех вариантах: в печатном виде, в рукописном виде и в виде вложения в формате «DOC» по e-mail.

В процессе изучения дисциплины предусмотрены следующие контрольные точки:

Контрольная точка	Срок сдачи (номер недели семестра)
Выполнение контрольной работы	10
Сдача реферата	8, 15

Контрольная работа (КР) выполняется письменно по билетам. Каждому обучающемуся рекомендуется ответить на 4 вопроса. Вопросы в билетах составлены по материалам лекций за предыдущие учебные недели. Для подготовки к КР обучающемуся необходимо повторить материалы лекций и рекомендованной литературы.

Система оценивания самостоятельной работы студентов основывается на следующих критериях:

- точность ответа на поставленный вопрос;
- логичность и последовательность изложения;
- полнота и глубина рассматриваемого вопроса, проблемы;
- способность к работе с литературными источниками, интернет-ресурсами;
- способность самостоятельно анализировать и обобщать информационный материал;
- умение формулировать цели и задачи работы;
- структурная упорядоченность оформления материала;
- соблюдение меры при оформлении материалов (объем, шрифты, интервалы, таблицы, рисунки, ссылки) на компьютере

Для сдачи экзамена необходимо повторить материалы, пройденные в течение семестра.

Для получения допуска к экзамену по дисциплине "Теоретические основы обработки информации" студенту необходимо, но не достаточно своевременно сдать реферат.

Экзамен студенты сдают устно по билетам. Ответ по билету оценивается по степени соответствия содержания ответа вопросу, четкости и ясности изложения материала.

Промежуточная аттестация по дисциплине производится: в 6 семестре в виде экзамена.

По результатам аттестации выставляется оценка по шкале «неудовлетворительно», «удовлетворительно», «хорошо», «отлично». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Задания и подробная инструкция по сдаче решений в систему выкладываются на странице курса <https://sites.google.com/site/tatianavbatura/lectures>

2.2.2 Форма и перечень вопросов экзаменационного билета бсеместра

Форма экзаменационного билета

Таблица П1.3

Новосибирский государственный университет Экзамен
Теоретические основы обработки информации наименование дисциплины

09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
Программная инженерия и компьютерные науки
наименование образовательной программы

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №

1. Вопрос из категории 1
2. Вопрос из категории 2

Составитель

Т.В.Батура

(подпись)

Ответственный за образовательную программу

А.А. Романенко

(подпись)

« ____ » _____ 20 г.

Перечень вопросов экзамена, структурированный по категориям, представлен в таблице П1.4

Таблица П1.4

Семестр 6	Формулировка вопроса
Категория 1 (ПКС-2.6)	1. Элементы теории информации Шеннона, некоторые определения и формулы. Основные понятия, кодирование, асимптотические формулы. Спектры, префиксные коды.
	2. Энтропия. Оптимальное кодирование. Признак Крафта. Теорема о нижней оценке.
	3. Связь исправления ошибок с избыточностью и энтропией. Модель Бернулли и теорема о верхней оценке. Статистические модели источников сообщений. Приложения в лингвистике, в генетике, при исследовании музыки.
	4. Код Хемминга. Элементы криптографии, коды открытого ключа.
	5. Системы символьных преобразований, алгоритмические основы.
	6. Обработка изображений и области ее применения. Классификация типов изображений и алгоритмов обработки. Дискретизация и квантование функции яркости.
	7. Меры близости изображений, цветовое пространство и цветовая константа. Стадия предварительной обработки изображений. Коррекция яркости и контрастности. Подавление шумов с помощью масок.
	8. Выделение контуров. Градиентный метод, комбинаторный метод, метод Слободы.
	9. Обнаружение объектов на изображении. Системы поиска и сопровождения целей.
	10. Двумерные унитарные преобразования. Косинусные и синусные

	<p>преобразования. Преобразование на основе матриц Адамара. Преобразование Хаара. Дискретное вейвлет-преобразование. Приложения к кодированию звуков и видео.</p>
	<p>11. Параллельная обработка информации. Распараллеливание, как метод повышения эффективности обработки информации. Загруженность и асимптотическая загруженность устройств.</p>
	<p>12. Общие математические соотношения, характеризующие вычисления на векторно-конвейерных ЭВМ: формулы для загруженности и ускорения для системы конвейерных устройств, работающих с векторами в режиме зацепления, условие на длины векторов.</p>
	<p>13. Организация памяти с параллельным доступом к информации: общие принципы, некоторые перестановки и их свойства, методы адресации. Приложения к численным методам и в обработке изображений.</p>
	<p>14. Нестандартные методы обработки информации. Нейрокомпьютерный подход. Применение размытой логики в распознавании и системах управления.</p>
	<p>15. Двумерные унитарные преобразования. Косинусные и синусные преобразования. Преобразование на основе матриц Адамара. Преобразование Хаара. Дискретное вейвлет-преобразование. Приложения к кодированию звуков и видео.</p>
Категория 2 (ПКС-2.6)	<p>16. Системы извлечения информации. Распознавание именованных элементов. Разрешение кореференций.</p>
	<p>17. Методы определения авторства текстов. Методы энтропийного кодирования, алгоритм Хаффмана. Статистические методы атрибуции. Деревья решений.</p>
	<p>18. Анализ тональности и тематическая индексация текстов. Нечёткая логика Заде.</p>
	<p>19. Методы классификации текстов. Машинное обучение с учителем.</p>
	<p>20. Скрытые Марковские модели. Алгоритм Витерби.</p>
	<p>21. Методы кластеризации текстов. Машинное обучение без учителя.</p>
	<p>22. Методы анализа данных из социальных сетей. Графовые модели социальных сетей. Анализ отношения предпочтения. Моделирование психологических операций.</p>
	<p>23. Представление знаний для компьютерной обработки. Тезаурусы, семантические сети, онтологии и базы знаний. Методы измерения семантического расстояния.</p>
	<p>24. Методы теоретического исследования семантики текстов. Подход Мельчука, теоретико-модельный подход.</p>
	<p>25. Порождающие грамматики. Распознающие грамматики. Регулярные выражения и конечные автоматы.</p>
	<p>26. Методы теоретического исследования семантики текстов. Теоретико-множественный подход Маркуса.</p>
	<p>27. Основные понятия математической лингвистики. Системы автоматизированной обработки текстов. Общие принципы построения систем автоматизированной обработки текстов.</p>
	<p>28. Синтаксические анализаторы. Проекты АОТ и Link Grammar Parser. Устройство поисковых систем. Морфологический и синтаксический анализ в поисковых системах.</p>
	<p>29. Методы задания синтаксической структуры предложений. Деревья составляющих и деревья зависимостей. Проблемы синтаксического</p>

	анализа. Проект Universal Networking Language.
	30. Представление знаний для компьютерной обработки. Тезаурусы, семантические сети, онтологии и базы знаний. Методы измерения семантического расстояния.

Набор вопросов для экзамена формируется и утверждается в установленном порядке в начале учебного года при наличии контингента обучающихся, осваивающих дисциплину «Теоретические основы обработки информации» в текущем учебном году.

3. Критерии оценки сформированности компетенций в рамках промежуточной аттестации по дисциплине

Таблица П1.5

Шифр компетенций	Структурные элементы оценочных средств	Показатель сформированности	Не сформирован	Пороговый уровень	Базовый уровень	Продвинутый уровень
ПКС-2	Портфолио (этап 1), Экзамен (этап 2)	ПКС-2.6 Знать: алгоритмы, модели объекты исследования в предметной области	Не знает основных принципов построения автоматизи рованных систем обработки текстов; не может спроектиро вать и разработать программн ые средства для решения практическ их задач в соответств ии с технически м заданием	Имеет фрагментарны е знания принципов построения автоматизиро ванных систем обработки текстов; формулировк и базовой задачи обнаружения объектов на формулировк и базовой задачи обнаружения объектов на изображении; основные подходы к исследованию семантики текстов	Допускает незначительные погрешности понимания в принципов построения автоматизированн ых систем обработки текстов; формулировки базовой задачи обнаружения объектов на изображении; основные подходы к исследованию семантики текстов; основных подходов к классификации и кластеризации данных; основной идеи организации памяти с параллельным доступом к информации	Демонстрирует четкое и целостное представление принципов построения автоматизированных систем обработки текстов; формулировки базовой задачи обнаружения объектов на изображении; основные подходы к исследованию семантики текстов; основных подходов к классификации и кластеризации данных; основной идеи организации памяти с параллельным доступом к информации

ШКС-2	Портфолио (этап 1), Экзамен (этап 2)	ПКС-2.6 Знать: алгоритмы, модели объекты исследования предметной области	Не умеет применять методы обработки изображен ий, сигналов и текстовой информации и при разработке и тестирован ии программн ых продуктов; применять методы теории информации и в различных областях	Допускает грубые ошибки, применя методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов; применять методы теории информации для учебных задач	Допускает незначительные ошибки, применя методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов; применять методы теории информации в различных областях	Уверенно применяет методы обработки изображений, сигналов и текстовой информации при разработке и тестировании программных продуктов; применять методы теории информации в различных областях для широкого класса задач
--------------	---	---	---	--	---	--

4. Критерии выставления оценок по результатам промежуточной аттестации по дисциплине

Результаты промежуточной аттестации в семестре определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Решение об окончательной оценке принимается по результатам 2 этапа (экзамен).

Оценка «отлично» соответствует продвинутому уровню сформированности компетенции.

Оценка «хорошо» соответствует базовому уровню сформированности компетенции.

Оценка «удовлетворительно» соответствует пороговому уровню сформированности компетенции.

Оценка «неудовлетворительно» выставляется при неудовлетворительном прохождении одного или двух этапов промежуточной аттестации.

**Лист актуализации фонда оценочных средств промежуточной аттестации
по дисциплине «Теоретические основы обработки информации»**

№	Характеристика внесенных изменений (с указанием пунктов документа)	Дата и № протокола Ученого совета ФИТ	Подпись ответственного
1.	Актуализирован на 2020- 2021 уч.год	22.07.2020 №77	
2.	Актуализирован на 2021 - 2022 уч. год	26.04.2021 №80	
3.	Актуализирован на 2022 - 2023 уч. год	31.08.2022 №87	
4.	Актуализирован на 2023 - 2024 уч. год	30.01.2023 №90	