

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

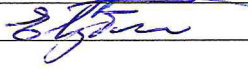
Директор ИШИТР

Сонькин Д.М.

«26» июль 2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИЕМ 2019 г.
ФОРМА ОБУЧЕНИЯ очно-заочная

Инструментарий для работы с большими данными			
Направление подготовки/специальность	09.04.03 Прикладная информатика		
Образовательная программа (направленность (профиль))	Информационные технологии в электроэнергетике		
Специализация	Информационные технологии в электроэнергетике		
Уровень образования	высшее образование - магистратура		
Курс	2	семестр	4
Трудоемкость в кредитах (зачетных единицах)	3		
Виды учебной деятельности	Временной ресурс		
Контактная (аудиторная) работа, ч	Лекции	16	
	Практические занятия	-	
	Лабораторные занятия	32	
	ВСЕГО	48	
Самостоятельная работа, ч		60	
ИТОГО, ч		108	

Вид промежуточной аттестации	Экзамен	Обеспечивающее подразделение	ОИТ
Заведующий кафедрой - руководитель ОИТ на правах кафедры			Шерстнёв В.С.
Руководитель ООП			Прохоров А.В.
Преподаватель			Губин Е. И.

2020 г.

1. Цели освоения дисциплины

Целями освоения дисциплины является формирование у обучающихся определенного ООП (п. 5.4 Общей характеристики ООП) состава компетенций для подготовки к профессиональной деятельности.

Код компетенции	Наименование компетенции	Индикаторы достижения компетенций		Составляющие результатов освоения (дескрипторы компетенции)	
		Код индикатора	Наименование индикатора достижения	Код	Наименование
ОПК(У)-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	И.ОПК(У)-2.3	Применяет интеллектуальные технологии при разработке программных средств решения профессиональных задач	ОПК(У)-2.3В1	Владеет: опытом применения современных информационных средств, технологий, инструментария для работы с большими объемами данных
				ОПК(У)-2.3У1	Умеет: проектировать многомерные кубы данных
				ОПК(У)-2.3З1	Знает: технологии манипуляции и анализа больших объемов данных
ПК(У)-2 2	Способен самостоятельно осваивать и применять информационные технологии для автоматизации и бизнес-процессов в электроэнергетике	И.ПК(У)-2.2	Проектирует и разрабатывает программное обеспечение информационных систем для автоматизации бизнес-процессов в электроэнергетике	ПК(У)-2.2З3	Знает: инструментальные средства реализации информационных систем на основе современных технологий разработки программного обеспечения и применения СУБД
				ПК(У)-2.2У3	Умеет: выполнять проект концептуальной модели базы данных информационной системы
				ПК(У)-2.2В2	Владеет: методами интеграции информационных систем и их компонентов на уровне данных
				ПК(У)-2.2У6	Умеет: применять методы интеграции информационных систем и их компонентов на уровне данных

2. Место дисциплины (модуля) в структуре ООП

Дисциплина относится к вариативной части Блока 1 учебного плана образовательной программы.

3. Планируемые результаты обучения по дисциплине

После успешного освоения дисциплины будут сформированы результаты обучения:

Планируемые результаты обучения по дисциплине		Индикатор достижения компетенции
Код	Наименование	
РД 1	Применять интеллектуальные технологии при разработке аналитических систем	И.ОПК(У)-2.3
РД 2	Планировать конфигурацию, осуществлять развертывание, настройку и администрирование хранилищ Hadoop	И.ПК(У)-2.2
РД 3	Реализовывать обмен данными между кластерами Hadoop и различными источниками и потребителями данных.	И.ОПК(У)-2.3
РД 4	Применять технологии обработки больших объемов данных	И.ОПК(У)-2.3
РД 5	Осуществлять интеграцию информационных систем и их компонентов на уровне данных	И.ПК(У)-2.2

Оценочные мероприятия текущего контроля и промежуточной аттестации представлены в календарном рейтинг-плане дисциплины.

4. Структура и содержание дисциплины

Основные виды учебной деятельности

Разделы дисциплины	Формируемый результат обучения по дисциплине	Виды учебной деятельности	Объем времени, ч.
Раздел (модуль) 1. Введение в обработку больших данных	РД1 РД2	Лекции	8
		Практические занятия	-
		Лабораторные занятия	16
		Самостоятельная работа	30
Раздел (модуль) 2. Hadoop для обработки больших данных	РД1 РД3 РД4 РД5	Лекции	8
		Практические занятия	-
		Лабораторные занятия	16
		Самостоятельная работа	30

Содержание разделов дисциплины:

Раздел 1. Введение в обработку больших данных

Введение, обзор имеющихся технологий для обработки больших данных, излагаются особенности работы с большими данными, возникающие проблемы и их решения, рассматривается установка и настройка кластера Hadoop.

Темы лекций:

1. Концепции данных
2. Центральные технологии экосистемы Hadoop
3. Планирование, установка и конфигурирование кластера Hadoop
4. Администрирование и управление кластером Hadoop

Названия лабораторных работ:

1. Введение в обработку больших данных

2. Введение в Hadoop
3. Установка и конфигурирование кластера Hadoop
4. Администрирование кластера Hadoop

Раздел 2. Hadoop для обработки больших данных

Описываются различные методы работы с данными в экосистеме Hadoop: с использованием технологий Pig, Spark и Hive. Рассматриваются преимущества, недостатки и границы применимости технологий.

Темы лекций:

1. Hadoop Distributed File System (HDFS). Обмен данных с кластером Hadoop
2. Apache Pig
3. Apache Hive
4. Apache Spark

Названия лабораторных работ:

1. HDFS. Импорт и экспорт данных в кластер
2. Обработка данных с помощью Apache Pig
3. Обработка данных с помощью Apache Hive
4. Обработка данных и машинное обучение с помощью Apache Spark

5. Организация самостоятельной работы студентов

Самостоятельная работа студентов при изучении дисциплины (модуля) предусмотрена в следующих видах и форм:

- Работа с лекционным материалом, поиск и обзор литературы и электронных источников информации по индивидуально заданной проблеме курса (10ч.);
- Изучение тем, вынесенных на самостоятельную проработку (20ч.);
- Подготовка к лабораторным работам (14ч.);
- Подготовка к экзамену (16ч.).

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Учебно-методическое обеспечение

Основная литература:

1. Чак, Л. Hadoop в действии / Л. Чак. — Москва : ДМК Пресс, 2012. — 424 с. — ISBN 978-5-94074-785-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/39997> (дата обращения: 26.05.2020). — Режим доступа: для авториз. пользователей.
2. Осипов, Д. Л. Технологии проектирования баз данных / Д. Л. Осипов. — Москва : ДМК Пресс, 2019. — 498 с. — ISBN 978-5-97060-737-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://ezproxy.ha.tpu.ru:2225/book/131692> (дата обращения: 26.05.2020). — Режим доступа: для авториз. пользователей.

6.2. Информационное и программное обеспечение

Internet-ресурсы (в т.ч. в среде LMS MOODLE и др. образовательные и библиотечные ресурсы):

1. Перминов Г.И. Видеокурс «Хранилища данных» в Интернет университете информационных технологий <https://intuit.ru/studies/courses/1168/314/info>
2. Профессиональные базы данных и информационно-справочные системы доступны по ссылке: <https://www.lib.tpu.ru/html/irs-and-pdb>

3. Полнотекстовые и реферативные базы данных для студентов и сотрудников ТПУ: <https://www.lib.tpu.ru/html/full-text-db>
4. Apache Hadoop: <https://hadoop.apache.org/>

Лицензионное программное обеспечение (в соответствии с **Перечнем лицензионного программного обеспечения ТПУ**):

1. Microsoft Office 2007 Standard Russian Academic
2. Document Foundation Libre Office
3. Microsoft Visual Studio 2019 Community

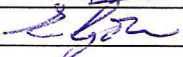
7. Особые требования к материально-техническому обеспечению дисциплины

В учебном процессе используется следующее лабораторное оборудование для практических и лабораторных занятий:

№	Наименование специальных помещений	Наименование оборудования
1.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (компьютерный класс) 634034, Томская область, г. Томск, Усова улица, 7 221	Компьютер – 20 шт., видеопроектор - 1 шт., звуковая система - 1 шт. Доска аудиторная настенная - 1 шт., тумба подкатная - 3 шт., комплект учебной мебели на 15 посадочных мест.

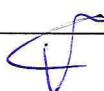
Рабочая программа составлена на основе Общей характеристики основной образовательной программы по направлению 09.04.03 Прикладная информатика, специализация «Информационные технологии в электроэнергетике» (приема 2019 г., очная очно-заочная форма обучения).

Разработчик(и):

Должность	Подпись	ФИО
Доцент ОИТ ИШИТР		Губин Е. И.

Программа одобрена на заседании Отделения электроэнергетики и электротехники Инженерной школы энергетики (протокол от «27» июня 2019 г. № 6).

И.о. заведующего кафедрой –
руководителя ОЭЭ на правах кафедры,
к. т. н., доцент

 А.С. Ивашутенко

Лист изменений рабочей программы дисциплины:

Учебный год	Содержание /изменение	Обсуждено на заседании Отделения электроэнергетики и электротехники (протокол)
2020/2021	1. Обновлено программное обеспечение 2. Обновлен состав профессиональных баз данных и информационно-справочных систем 3. Обновлено содержание разделов дисциплины 4. Обновлен список литературы, в том числе ссылок ЭБС	протокол от «25» июня 2020 г. № 6