

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



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИЕМ 2019 г.
ФОРМА ОБУЧЕНИЯ очная

Интеллектуальный анализ web-данных			
Направление подготовки/специальность	09.04.04 Программная инженерия		
Образовательная программа (направленность (профиль))	«Технологии больших данных»		
Специализация	«Технологии больших данных»		
Уровень образования	высшее образование - магистратура		
Курс	2	семестр	3
Трудоемкость в кредитах (зачетных единицах)	6		
Виды учебной деятельности	Временной ресурс		
Контактная (аудиторная) работа, ч	Лекции	24	
	Практические занятия		
	Лабораторные занятия	24	
	ВСЕГО	48	
Самостоятельная работа, ч		168	
ИТОГО, ч		216	

Вид промежуточной аттестации	Экзамен	Обеспечивающее подразделение	ОИТ ИШИТР
Заведующий кафедрой-руководитель ОИТ на правах кафедры			Шерстнев В.С.
Руководитель ООП			Савельев А.О.
Преподаватель			Савельев А.О.

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)-2.1	Разрабатывает оригинальные алгоритмы и программные средства для решения профессиональных задач	ОПК(У)-2.1В1	Владеет опытом программной реализации алгоритмов обработки пространственной информации
				ОПК(У)-2.1У1	Умеет применять алгоритмы обработки пространственных данных для решения профессиональных задач
				ОПК(У)-2.1З1	Знает технологии обработки пространственной и картографической информации
ОПК(У)-6	Способен использовать методы и средства системной инженерии в области получения, передачи, хранения, переработки и предоставления информации посредством информационных технологий	И.ОПК (У)-6.1	Применяет знания технологий управления и хранения данных для решения профессиональных задач	ОПК(У)-6.1В1	Владеет опытом использования системной инженерии для проектирования и реализации информационных технологий передачи хранения и обработки данных
				ОПК(У)-6.1З1	Знает методы оценки качества программных продуктов

2. Место дисциплины (модуля) в структуре ООП

Дисциплина относится к вариативной части Блока М1.ВМ3 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	Проводить научные исследования, связанные с объектами профессиональной деятельности	ОПК(У)- 2
РД 2	Исследовать и анализировать большие данные, создавать их модели и интерпретировать структуры данных в таких моделях	ОПК(У)-6

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. <i>Основы интеллектуального анализа web-данных</i>	РД1	Лекции	6
		Практические занятия	
		Лабораторные занятия	
		Самостоятельная работа	18
Раздел (модуль) 2. <i>Математическое и алгоритмическое обеспечения анализа web-данных</i>	РД2	Лекции	18
		Практические занятия	
		Лабораторные занятия	24
		Самостоятельная работа	150

Содержание разделов дисциплины:

Раздел 1. Основы интеллектуального анализа web-данных

Раздел формирует общие представления о целях, задачах, особенностях и ограничениях интеллектуального анализа web-данных. Рассматриваются перспективные технологии и направления исследований. Приводится характеристика 3 основных дисциплин указанного научного направления: Web Usage Mining, Web Structure Mining и Web Content Mining.

Темы лекций:

- 1. Web Data Mining: введение в дисциплину.
- 2. Анализ данных социальных медиа.
- 3. Основы интеллектуального анализа данных.

Раздел 2. Математическое и алгоритмическое обеспечения анализа web-данных

Раздел формирует навыки организации работ по анализу web-данных. Рассматриваются вопросы построения и анализа социального графа, моделирования распространения информации, формирования рекомендаций и анализа пользовательского поведения.

Темы лекций:

- 1. Анализ сетевой топологии: узлы, связи и влияние.
- 2. Поиск и анализ сообществ.
- 3. Влияние и гомофилия.
- 4. Моделирование информационной диффузии.
- 5. Рекомендательные алгоритмы.
- 6. Анализ тональности текста.
- 7. Анализ поведения.
- 8. Извлечение знаний из лог-файлов веб-серверов.
- 9. Web Crawling: обзор возможностей и ограничений.

Названия лабораторных работ:

- 1. Основы работы с VK API.
- 2. Автоматический анализ сообщества социальной сети.
- 3. Анализ сетевой топологии: значимость узлов.
- 4. Основы моделирования информационной диффузии.
- 5. Web-scraping: автоматизация извлечения данных открытых веб-страниц.

5. Организация самостоятельной работы студентов

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

- Работа с лекционным материалом, поиск и обзор литературы и электронных источников информации по индивидуально заданной проблеме курса;
- Изучение тем, вынесенных на самостоятельную проработку;
- Подготовка к лабораторным работам, к практическим и семинарским занятиям;
- Подготовка к оценивающим мероприятиям;

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Учебно-методическое обеспечение

- Степанов, П. Е. Планирование эксперимента: учебно-методическое пособие / П. Е. Степанов. — Москва: МИСИС, 2017. — 22 с. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/108113> (дата обращения: 03.03.2021). — Режим доступа: для авториз. пользователей.
2. Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва: ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст: электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131721> (дата обращения: 03.03.2021). — Режим доступа: для авториз. пользователей.
3. Макшанов, А. В. Технологии интеллектуального анализа данных: учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург: Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/120063> (дата обращения: 03.03.2021). — 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. Python v.3.0.1 documentation, <https://docs.python.org/3.0/>
2. A Survey of Data Mining Techniques for Social Media Analysis, <https://arxiv.org/vc/arxiv/papers/1312/1312.4617v1.pdf>

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

1. Microsoft Office Standart 2016

7. Особые требования к материально-техническому обеспечению дисциплины

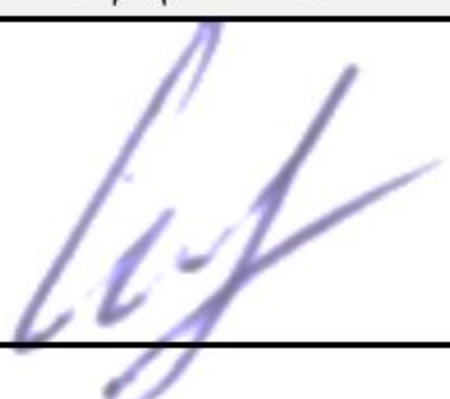
В учебном процессе используется следующее лабораторное оборудование для практических и лабораторных занятий:

№	Наименование специальных помещений	Наименование оборудования
1.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (лекционная аудитория)	Комплект оборудования для проведения лекционных занятий: – ПК, 1 шт. – Проектор, 1 шт.

	634034 Томская область, Томск, Советская улица, д. 84/3, учебный корпус «Кибернетический центр», аудитория 313.	
2.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (учебная лаборатория) 634034 Томская область, Томск, Советская улица, д. 84/3, учебный корпус «Кибернетический центр», аудитория 204.	Комплект оборудования для проведения лабораторных работ по основным разделам Информационных технологий: – ПК, 10 шт.

Рабочая программа составлена на основе Общей характеристики образовательной программы по направлению 09.04.04 Программная инженерия/ профиль Big Data Solutions / Технологии больших данных /специализации Big Data Solutions/Технологии больших данных (приема 2019 г., очная форма обучения).

Разработчик(и):

Должность	Подпись	ФИО
Доцент ОИТ		Савельев А.О.

Программа одобрена на заседании Отделения информационных технологий ИШИТР (протокол №13 от 19.06.2019 г.).

Заведующий кафедрой –
руководитель ОИТ на правах кафедры  /Шерстнев В.С

Лист изменений рабочей программы дисциплины:

Учебный год	Содержание /изменение	Обсуждено на заседании ИШИТР (протокол)
2020/2021	1. Обновлено программное обеспечение 2. Обновлен список литературы, в том числе ссылок ЭБС	от 01.09.2020г. № 19