

Документ подписан простой электронной подписью
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной работе

Дата подписания: 13.09.2023 10:25:51

Уникальный идентификатор документа:

d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное образовательное учреждение
высшего образования**

«Национальный исследовательский технологический университет «МИСИС»

Рабочая программа дисциплины (модуля)

ETL: автоматизация подготовки данных

Закреплена за подразделением

Кафедра АСУ

Направление подготовки

09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Профиль

Инженерия данных

Квалификация

Магистр информационных систем

Форма обучения

очная

Общая трудоемкость

3 ЗЕТ

Часов по учебному плану

108

Формы контроля в семестрах:

в том числе:

зачет с оценкой 2

аудиторные занятия

51

самостоятельная работа

57

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	2 (1.2)		Итого	
	11			
Неделя	УП	РП	УП	РП
Вид занятий	УП	РП	УП	РП
Лекции	17	17	17	17
Практические	34	34	34	34
Итого ауд.	51	51	51	51
Контактная работа	51	51	51	51
Сам. работа	57	57	57	57
Итого	108	108	108	108

Программу составил(и):

-, *ст.преп., Агабубаев А.Т.*

Рабочая программа

ЕТЛ: автоматизация подготовки данных

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - магистратура Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСиС» по направлению подготовки 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА (приказ от 28.06.2023 г. № 292 о.в.)

Составлена на основании учебного плана:

09.04.01 Информатика и вычислительная техника, 09.04.01-МИВТ-23-9-ПП.plx Инженерия данных, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

09.04.01 Информатика и вычислительная техника, Инженерия данных, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

Кафедра АСУ

Протокол от 10.04.2023 г., №5

Руководитель подразделения Темкин Игорь Олегович, д.т.н., доцент

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Дисциплина нацелена на освоение навыков работы с инструментальными средствами сбора, хранения и обработки данных
-----	--

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		Б1.В.ДВ.01
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Базы данных и SQL	
2.1.2	Системы хранения и обработки данных	
2.1.3	Современные технологии и инструменты разработки	
2.1.4	Языки программирования для работы с данными	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ПК-1: Способен управлять хранилищами данных на всех этапах ETL-процесса	
Знать:	
ПК-1-31 практики, процессы и технологии, которые сочетают, интегрированный и ориентированный на процесс работы с данными, посредством инструментария автоматизации и методами гибкой инженерии программного обеспечения	
Уметь:	
ПК-1-У1 применять современные инструментальные средства разработки программных компонентов сбора, хранения и обработки данных	
Владеть:	
ПК-1-В1 современными инструментальными средствами работы с данными	

4. СТРУКТУРА И СОДЕРЖАНИЕ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
Раздел 1. Лекции								
1.1	Введение в ETL и Big Data /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5			
1.2	Инструменты извлечения данных (Extraction) /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
1.3	Инструменты преобразования данных (Transformation) /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
1.4	Инструменты загрузки данных (Loading) /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
1.5	Инструменты мониторинга и управления процессом ETL /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
1.6	Оптимизация процесса ETL /Лек/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
1.7	Работа с реальными данными /Лек/	2	5	ПК-1-31 ПК-1-У1 ПК-1-В1				
Раздел 2. Практикум								
2.1	Обзор основных инструментов ETL в Big Data /Пр/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.2	Способы извлечения данных из источников /Пр/	2	2	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.3	Инструменты извлечения данных Sqoop, Flume и Kafka /Пр/	2	6	ПК-1-31 ПК-1-У1 ПК-1-В1				

2.4	инструменты преобразования данных Pig и Spark /Пр/	2	4	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.5	Целевые хранилища, реляционные базы данных, NoSQL; хранилища и Hadoop Distributed File System (HDFS) /Пр/	2	8	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.6	Инструменты мониторинга и управления Oozie и Apache NiFi /Пр/	2	4	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.7	Использование индексов и кэширования для ускорения процесса ETL /Пр/	2	4	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.8	Инструментов оптимизации процесса ETL Apache Spark и Apache Flink /Пр/	2	4	ПК-1-31 ПК-1-У1 ПК-1-В1				
2.9	Примеры реальных случаев использования инструментов ETL в Big Data (анализ данных в социальных сетях и анализ данных в медицине) /Ср/	2	57	ПК-1-31 ПК-1-У1 ПК-1-В1			КМ1	Р1

5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Вопросы к зачету	ПК-1-31	<p>Что такое ETL и как он связан с Big Data?</p> <p>Какие проблемы могут возникнуть при обработке больших объемов данных?</p> <p>Какие инструменты ETL используются для обработки Big Data?</p> <p>Какие источники данных могут быть использованы в ETL?</p> <p>Какие методы извлечения данных могут быть применены в ETL?</p> <p>Какие методы преобразования данных могут быть применены в ETL?</p> <p>Какие методы загрузки данных могут быть применены в ETL?</p> <p>Какие инструменты мониторинга и управления процессом ETL могут быть использованы?</p> <p>Какие инструменты оптимизации процесса ETL могут быть использованы?</p> <p>Какие языки запросов могут быть использованы для ETL в Big Data?</p> <p>Какие преимущества может предоставить использование инструментов ETL в Big Data?</p> <p>Какие методы обработки реальных данных могут быть применены в ETL?</p> <p>Какие проблемы могут возникнуть при обработке данных из социальных сетей?</p> <p>Какие проблемы могут возникнуть при обработке медицинских данных?</p> <p>Какие методы анализа данных могут быть применены в ETL?</p> <p>Какие методы хранения данных могут быть использованы в ETL?</p> <p>Какие методы обеспечения безопасности данных могут быть использованы в ETL?</p> <p>Какие методы обеспечения качества данных могут быть использованы в ETL?</p> <p>Какие инструменты могут использоваться для визуализации данных, полученных из ETL?</p> <p>Какие будущие тенденции можно ожидать в области ETL в Big Data?</p>

5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)			
Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
P1	Индивидуальное задание	ПК-1-У1;ПК-1-В1;ПК-1-З1	<p>Цель задания: изучить реальные случаи использования инструментов ETL в Big Data для анализа данных в социальных сетях и медицине.</p> <p>Задание:</p> <p>Изучите примеры реальных случаев использования инструментов ETL в Big Data для анализа данных в социальных сетях и медицине.</p> <p>Сформулируйте свои выводы о том, как инструменты ETL могут быть полезны в анализе данных в социальных сетях и медицине. Примеры реальных случаев использования инструментов ETL в Big Data для анализа данных в социальных сетях и медицине могут включать в себя:</p> <p>Анализ социальных сетей для выявления тенденций и предсказания поведения пользователей Анализ медицинских данных для выявления связей между различными заболеваниями и лечением Анализ генетических данных для выявления связей между генами и заболеваниями Анализ данных о клиентах для повышения качества сервиса и улучшения маркетинговых стратегий.</p>

5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

Экзамен не предусмотрен

5.4. Методика оценки освоения дисциплины (модуля, практики. НИР)

Промежуточная аттестация в форме коллоквиума позволяет оценить уровень сформированности компетенций в целом по дисциплине и может осуществляться, как в письменной так и в устной форме.

По окончании изучения дисциплины в системе оценки знаний и умений используются следующие критерии:

«Отлично» – за глубокое и полное овладение содержанием учебного материала, в котором студент легко ориентируется, владение понятийным аппаратом за умение связывать теорию с практикой, решать практические задачи, высказывать и обосновывать свои суждения. Отличная отметка предполагает грамотное, логичное изложение ответа (как в устной, так и в письменной форме), качественное внешнее оформление;

«Хорошо» – если студент полно освоил учебный материал, владеет понятийным аппаратом, ориентируется в изученном материале, осознанно применяет знания для решения практических задач, грамотно излагает ответ, но содержание и форма ответа имеют некоторые неточности;

«Удовлетворительно» – если студент обнаруживает знание и понимание основных положений учебного материала, но излагает его неполно, непоследовательно, допускает неточности в определении понятий, в применении знаний для решения практических задач, не умеет доказательно обосновать свои суждения;

«Неудовлетворительно» – если студент имеет разрозненные, бессистемные знания, не умеет выделять главное и второстепенное, допускает ошибки в определении понятий, искажает их смысл, беспорядочно и неуверенно излагает материал, не может применять знания для решения практических задач; за полное незнание и непонимание учебного материала или отказ отвечать.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Труб И. И.	СУБД Cache: работа с объектами: учебное пособие	Электронная библиотека	Москва: Диалог-МИФИ, 2006
Л1.2	Гудов А. М., Завозкин С. Ю., Рейн Т. С.	Базы данных и системы управления базами данных. Программирование на языке PL/SQL: учебное пособие	Электронная библиотека	Кемерово: Кемеровский государственный университет, 2010

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.3	Щелоков С. А.	Разработка и создание баз данных средствами СУБД Access и SQL Server: учебно-методическое пособие	Электронная библиотека	Оренбург: Оренбургский государственный университет, 2014
Л1.4	Васюткина И. А., Трошина Г. В., Бычков М. И., Менжулин С. А.	Разработка приложений на С с использованием СУБД PostgreSQL: учебное пособие	Электронная библиотека	Новосибирск: Новосибирский государственный технический университет, 2015
Л1.5		Администрирование MySQL: курс: учебное пособие	Электронная библиотека	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), 2007

6.3 Перечень программного обеспечения

П.1	Microsoft SQL server 2016
П.2	Oracle DB Express
П.3	Python
П.4	Qlik Sense Desktop 2019

6.4. Перечень информационных справочных систем и профессиональных баз данных

И.1	https://www.scopus.com
И.2	https://habr.com

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
Любой корпус Компьютерный класс	Учебная аудитория для проведения практических занятий:	экран, проектор, доска, комплект учебной мебели на 30 посадочных мест, персональные компьютеры, доступ к ЭИОС университета LMS Canvas, лицензионные программы MS Teams, MS Office
Любой корпус Мультимедийная	Учебная аудитория для проведения занятий лекционного типа и/или для проведения практических занятий:	комплект учебной мебели до 36 мест для обучающихся, мультимедийное оборудование, магнитно-маркерная доска, рабочее место преподавателя, ПКс доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Дисциплина требует значительного объема самостоятельной работы. Отдельные учебные вопросы выносятся на самостоятельную проработку и контролируются посредством текущей аттестации.

При этом организуются групповые и индивидуальные консультации. Качественное освоение дисциплины возможно только при систематической самостоятельной работе, что поддерживается системой текущей и рубежной аттестации. Практикум проводится с широким использованием компьютерных программ, как для выполнения, так и для оформления работы.

В процессе выполнения практических работ необходимо показать умелое применение полученных в процессе обучения знаний и навыков при решении задач.

С целью формирования практических навыков, знаний, полученных при изучении дисциплин Основы дискретной математики, Комбинаторика и теория графов, Алгоритмы дискретной математики и Высшая математика.

При выполнении самостоятельных работ акцент делается на формирование навыков работы студентов с научно-технической литературой; работы с документацией библиотек Python; на систематизацию материала для решения поставленных задач;

на формирование навыков оформления результатов выполненных работ (пояснительной записки, ссылок на литературные источники, выводов по работе).

Индивидуальные задания на самостоятельную работу студент получает у преподавателя в соответствии с прилагаемым перечнем их тематик.

Рекомендуемая форма их оформления – отчеты с приложением программного файла ЭИОР «Canvas». Защита работы проводится индивидуально каждым студентом. Студенты делают сообщения и отвечают на вопросы преподавателя.