

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет имени В. Н. Татищева»
(Астраханский государственный университет им. В. Н. Татищева)

СОГЛАСОВАНО
Руководитель ОПОП

О.Н. Выборнова

«05» мая 2025 г.

УТВЕРЖДАЮ
И.о. заведующего кафедрой
информационных технологий
О.Н. Выборнова

«05» мая 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

**«Процесс организации и использования данных внутри компании
на платформах Universe и Плюс 7»**

Составитель(и)

**Лулева Е.А., ст. преподаватель кафедры ИТ;
Выборнова О.Н., доцент, к.т.н., доцент кафедры
ИТ;**

Согласовано с работодателями:

**Лазарев Н.В., инженер 2 категории группы
контроля безопасности объектов критической
информационной инфраструктуры отдела
информационной безопасности управления
корпоративной защиты, ООО «Газпром добыча
Астрахань»;**

**Горбатенко С.Ю., заместитель директора,
ГБУ АО «Инфраструктурный центр
электронного правительства;**

Направление подготовки /
специальность

**10.03.01 ИНФОРМАЦИОННАЯ
БЕЗОПАСНОСТЬ**

Направленность (профиль) /
специализация ОПОП

**Организация и технологии защиты информации
(в сфере информационных и коммуникационных
технологий)**

Квалификация (степень)

бакалавр

Форма обучения

очная, очно-заочная

Год приёма

2025

Курс

**2 (по очной форме) /
2 (по очно-заочной форме)**

Семестр(ы)

**4 (по очной форме) /
4 (по очно-заочной форме) /**

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1. Целями освоения дисциплины (модуля) «Процесс организации и использования данных внутри компании на платформах Universe и Плюс 7» являются: изучить особенности организации и использования данных внутри компании на платформах «Юниверс» и «Плюс 7»

1.2. Задачи освоения дисциплины (модуля):

- сформировать у студентов представления об управлении корпоративными данными;
- ознакомить студентов с концепцией Управления данными и управления НСИ;
- обучить студентов основам операций с данными на платформах «Юниверс» и «Плюс 7».

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОПОП

2.1. Учебная дисциплина (модуль) «Процесс организации и использования данных внутри компании на платформах Universe и Плюс 7» относится к части, формируемой участниками образовательных отношений, и осваивается в 4 семестре.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы следующие знания, умения, навыки, формируемые предшествующими учебными дисциплинами (модулями):

- Информатика
- Основы программирования
- Аппаратные средства вычислительной техники

Знания: основных понятий информатики; принципов построения информационных систем; реляционной модели баз данных и базовое знание языка SQL (синтаксис языка); программирования и базовых знаний по системному администрированию;

Умения: использовать программные и аппаратные средства персонального компьютера;

Навыки: поиска информации в глобальной сети Интернет и работы с офисными приложениями (текстовыми процессорами, электронными таблицами, средствами подготовки презентаций и т.п.); работы с различными Case-средствами.

2.3. Последующие учебные дисциплины (модули) и (или) практики, для которых необходимы знания, умения, навыки, формируемые данной учебной дисциплиной (модулем):

- Базы данных
- Аттестация объектов информатизации
- Технологии облачных вычислений и виртуализации
- Комплексное обеспечение защиты информации объекта информатизации

Также дисциплина «Процесс организации и использования данных внутри компании на платформах Universe и Плюс 7» поможет студентам при реализации задач преддипломной практики и написании бакалаврской работы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Процесс освоения дисциплины (модуля) направлен на формирование элементов следующей(их) компетенции(ий) в соответствии с ФГОС ВО и ОПОП ВО по данному направлению подготовки / специальности:

- а) профессиональной (ПК): ПК-3. Способен разрабатывать программное обеспечение,

включая проектирование, отладку, проверку работоспособности и модификацию ПО

Таблица 1. Декомпозиция результатов обучения

Код компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине (модулю)		
		Знать (1)	Уметь (2)	Владеть (3)
ПК-3	ПК-3.1. Знать: современные информационные технологии разработки, отладки, проверки работоспособности, модификации программного обеспечения	методики использования программных средств для решения практических задач	использовать программные средства для решения практических задач	использования программных средств для решения практических задач
	ПК-3.2. Уметь: осуществлять выбор информационных технологий для решения задач по разработке, отладке, проверке работоспособности, модификации программного обеспечения	принципы и виды отладки программного обеспечения; методы оценки качества программ;	применять принципы и виды отладки программного обеспечения;	навыками разработки и отладки программ
	ПК-3.3. Владеть: навыками разработки, отладки, проверки работоспособности, модификации программного обеспечения с использованием современных информационных технологий	методики сопряжения компонентов информационных и автоматизированных систем	разрабатывать программное обеспечение с использованием баз данных	методами описания схем баз данных;

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость дисциплины в соответствии с учебным планом составляет 2 зачетные единицы (72 часа).

Трудоемкость отдельных видов учебной работы студентов очной и очно-заочной форм обучения приведена в таблице 2.1.

Таблица 2.1. Трудоемкость отдельных видов учебной работы по формам обучения

Вид учебной и внеучебной работы	для очной формы обучения	для очно-заочной формы обучения
Объем дисциплины в зачетных единицах	2	2
Объем дисциплины в академических часах	72	72
Контактная работа обучающихся с преподавателем (всего), в том числе (час.):	36	36
- занятия лекционного типа, в том числе:	0	0
- практическая подготовка (если предусмотрена)		
- занятия семинарского типа (семинары, практические, лабораторные), в том числе:	36	36
- практическая подготовка (если предусмотрена)		
Самостоятельная работа обучающихся (час.)	36	36
Форма промежуточной аттестации обучающегося (зачет/экзамен), семестр (ы)	зачет – 4 семестр;	зачет – 4 семестр

Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий и самостоятельной работы, для каждой формы обучения представлено в таблице 2.2.

Таблица 2.2. Структура и содержание дисциплины (модуля)**для очной формы обучения**

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости и, форма промежуточной аттестации [по семестрам]
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Семестр 4.										
<i>Тема 1. Введение в управление корпоративными данными</i>					2			2	4	<i>Лабораторная работа, устный опрос</i>
<i>Тема 2. Интеграция данных</i>					10			8	18	<i>Лабораторная работа, устный опрос</i>
<i>Тема 3. Обеспечение качества данных</i>					8			6	14	<i>Лабораторная работа, устный опрос</i>
<i>Тема 4. Управление НСИ и мастер-данными</i>					6			6	12	<i>Лабораторная работа, устный опрос</i>
<i>Тема 5. Основы DG – руководства на основе данных</i>					6			6	12	<i>Лабораторная работа, устный опрос</i>

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости и, форма промежуточ ной аттестации [по семестрам] <i>опрос</i>
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
<i>Тема 6. Основы BI</i>					4			8	12	<i>Лабораторная работа, устный опрос</i>
Консультации										
Контроль промежуточной аттестации										Зачёт
ИТОГО за семестр:					36			36	72	
Итого за весь период										

для очно-заочной формы обучения

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости и, форма промежуточ ной аттестации [по семестрам] <i>опрос</i>
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Семестр 4.										
<i>Тема 1. Введение в управление корпоративными данными</i>					2			2	4	<i>Лабораторная работа, устный опрос</i>
<i>Тема 2. Интеграция данных</i>					10			8	18	<i>Лабораторная работа, устный опрос</i>
<i>Тема 3. Обеспечение качества данных</i>					8			6	14	<i>Лабораторная работа, устный опрос</i>
<i>Тема 4. Управление НСИ и мастер-данными</i>					6			6	12	<i>Лабораторная работа, устный опрос</i>
<i>Тема 5. Основы DG – руководства на основе данных</i>					6			6	12	<i>Лабораторная работа, устный опрос</i>
<i>Тема 6. Основы BI</i>					4			8	12	<i>Лабораторная работа, устный</i>

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости и, форма промежуточ ной аттестации [по семестрам] <i>опрос</i>
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Консультации										
Контроль промежуточной аттестации										Зачёт
ИТОГО за семестр:					36			36	72	
Итого за весь период										

Примечание: Л – лекция; ПЗ – практическое занятие, семинар; ЛР – лабораторная работа; ПП – практическая подготовка; КР / КП – курсовая работа / курсовой проект; СР – самостоятельная работа

Таблица 3. Матрица соотнесения разделов, тем учебной дисциплины (модуля) и формируемых компетенций

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции	Общее количество компетенций
		ПК-3	
<i>Тема 1. Введение в управление корпоративными данными</i>	4	+	1
<i>Тема 2. Интеграция данных</i>	18	+	1
<i>Тема 3. Обеспечение качества данных</i>	14	+	1
<i>Тема 4. Управление НСИ и мастер-данными</i>	12	+	1
<i>Тема 5. Основы DG – руководства на основе данных</i>	12	+	1
<i>Тема 6. Основы BI</i>	12	+	1
Итого	72		

Краткое содержание каждой темы дисциплины (модуля)

Тема 1 «Введение в управление корпоративными данными»

Что такое данные. Управление данными. Данные и информация. Данные как актив организации. Принципы управления данными. Ценность данных. Управление данными как кросс-функциональный процесс. Целостный взгляд на функционирование организации. Жизненный цикл данных. Различные виды данных. Данные и риск. Управление данными и информационные технологии. Стратегия управления данными

Тема 2 «Интеграция данных»

Цели, задачи, проблемы. Подходы к интеграции. Сравнение ETL, EAI, EII. Инструменты. Базовые понятия ETL: источники и приемники, маппинги, Workflow. Преобразования данных. Инструмент отладки. Работа со строками в БД. Параметры и переменные. Задачи процесса и многократное использование задач.

Тема 3 «Обеспечение качества данных»

Понятие качества данных, метрики качества данных, проблемы качества. Циклический процесс обеспечения качества данных, роли участников в процессе обеспечения качества данных. Основные функции, используемые при управлении качеством данных. Анализ данных путём профилирования. Очистка данных: стандартизация, нормализация, обогащение данных. Поиск и обработка дублированных данных: алгоритмы поиска дубликатов, применение различных подходов в зависимости от типа данных. Применение решения по управлению качеством в различных проектах: построение хранилищ данных, построение справочников, миграция и т.д.

Тема 4 «Управление НСИ и мастер-данными»

Определение основных данных, отличие мастер-данных и справочников от транзакционных данных. Виды мастер-данных и справочников. Типовые структуры справочников: списки, таксономия, фасетная классификация Актуальность и цели управления мастер-данными. Источники данных, понятие «золотой записи», подход к её формированию. Типовые варианты архитектуры MDM: централизованная, гармонизированная, аналитическая. Ключевые элементы модель данных: реестры, справочники, атрибуты, связи. Использование классификаторов в модели данных. Подходы к выявлению дубликатов записей.

Тема 5 «Основы DG – руководства на основе данных»

Цели, задачи, решаемые проблемы. Принципы управления данными, роль команды CDO и её взаимодействие с другими подразделениями в организации. Понятие мета-данных, бизнес-гlossарий, концептуальный и логический уровни модели данных. Каталог данных, объекты данных физического слоя, назначение сканера мета-данных. Связи моделей данных, понятия data lineage и impact analysis. Средства проверки и обеспечения качества данных. Средства выявления критичных данных и их защиты

Тема 6 «Основы BI»

Цели, задачи, проблемы. Визуализация данных. Понятие OLAP. Подходы к формированию OLAP кубов. Виды отчётов: плоские отчёты и кубы. Майнинг данных, применение ML в майнинге данных.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

5.1. Указания для преподавателей по организации и проведению учебных занятий по дисциплине (модулю)

При подготовке к лабораторным занятиям необходимо воспользоваться учебно-методической литературой из п.8.

5.2. Указания для обучающихся по освоению дисциплины (модулю)

В рамках дисциплины предполагается организация следующих видов самостоятельной работы студентов (таблица 4): работа с теоретическим материалом, учебно-методическим информационным обеспечением; подготовка к лабораторно-практическим работам, подготовка отчетов к защите отчетов; подготовка к контрольным работам в форме компьютерного тестирования, текущей и промежуточной аттестации (зачету). В качестве форм и методов контроля внеаудиторной самостоятельной работы используются: электронные отчеты по выполнению лабораторных работ; устный опрос, протоколы компьютерного тестирования и др.

Таблица 4. Содержание самостоятельной работы обучающихся

для очной формы обучения

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
<i>Тема 1. Введение в управление корпоративными данными</i> Общее представления по управлению данными и знаниями Знакомство с Плюс 7 ФормИТ	2	Изучение учебно-методического материала по теме, подготовка к устной сдаче лабораторных работ
<i>Тема 2. Интеграция данных</i> Папки, Источники и Приемники Трансформация Lookup для файлового источника Объединения таблиц баз данных с помощью Плюс 7 ФормИТ Логика Workflow	8	
<i>Тема 3. Обеспечение качества данных</i> Слияние, Роутинг, и сортировка данных	6	
<i>Тема 4. Управление НСИ и мастер-данными</i> Command Task. Запуск команд командной строки из workflow Отладчик	6	
<i>Тема 5. Основы DG – руководства на основе данных</i> Обновление таблиц баз данных Маплеты	6	
<i>Тема 6. Основы VI</i> Разработка маппинга Автоматическое объединение портов	8	

для очно-заочной формы обучения

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
<i>Тема 1. Введение в управление корпоративными данными</i> Общее представления по управлению данными и знаниями Знакомство с Плюс 7 ФормИТ	2	Изучение учебно-методического материала по теме, подготовка к устной сдаче лабораторных работ
<i>Тема 2. Интеграция данных</i> Папки, Источники и Приемники Трансформация Lookup для файлового источника Объединения таблиц баз данных с помощью Плюс 7 ФормИТ Логика Workflow	8	
<i>Тема 3. Обеспечение качества данных</i> Слияние, Роутинг, и сортировка данных	6	
<i>Тема 4. Управление НСИ и мастер-данными</i> Command Task. Запуск команд командной строки из workflow Отладчик	6	
<i>Тема 5. Основы DG – руководства на основе данных</i> Обновление таблиц баз данных Маплеты	6	
<i>Тема 6. Основы VI</i> Разработка маппинга Автоматическое объединение портов	8	

5.3. Виды и формы письменных работ, предусмотренных при освоении дисциплины (модуля), выполняемые обучающимися самостоятельно

Отчет по лабораторной работе

Оформляется в электронном виде: формат листа А4, книжная ориентация страницы. Отчеты по всем лабораторным работам имеют титульный лист, на котором указывается наименование дисциплины, ФИО и группа исполнителя, ФИО преподавателя, принимающего отчеты, тема лабораторной работы. В отчете по каждой лабораторной работе должны быть представлены цель, ход выполнения работы (скриншоты, краткое текстовое описание), выводы по результатам работы. Лабораторная работа в обязательном порядке отчитывается.

6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

При реализации различных видов учебной работы по дисциплине могут использоваться электронное обучение и дистанционные образовательные технологии.

6.1. Образовательные технологии

Название образовательной технологии	Темы, разделы дисциплины	Краткое описание применяемой технологии
Консультации по электронной почте	1 – 6	Подготовка к защите отчетов по лабораторным работам
Проектное обучение	1 – 6	Подготовка отчетов по лабораторным работам и обобщение их в единый проект
Анализ ситуаций	1 – 6	Лабораторные работы
Работа в малых группах	1 – 6	Выполнение лабораторной работы в малых группах (до 5 человек)

Учебные занятия по дисциплине могут проводиться с применением информационно-телекоммуникационных сетей при опосредованном (на расстоянии) взаимодействии обучающихся и преподавателя в режимах on-line в формах: видеолекций, лекций-презентаций, видеоконференции, собеседования в режиме чат, форума, чата, выполнения виртуальных практических и/или лабораторных работ и др.

Таблица 5. Образовательные технологии, используемые при реализации учебных занятий

Раздел, тема дисциплины (модуля)	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
<i>Тема 1. Введение в управление корпоративными данными</i>	Не предусмотрено	Не предусмотрено	Выполнение лабораторной работы, Фронтальный опрос, Тематические дискуссии
<i>Тема 2. Интеграция данных</i>			
<i>Тема 3. Обеспечение качества данных</i>			
<i>Тема 4. Управление НСИ и мастер-данными</i>			
<i>Тема 5. Основы DG – руководства на основе данных</i>			
<i>Тема 6. Основы VI</i>			

6.2. Информационные технологии

При реализации различных видов учебной и внеучебной работы используются:

- возможности интернета в учебном процессе;

- использование электронных учебников и различных сайтов как источников информации;
- использование возможностей электронной почты преподавателя;
- использование средств представления учебной информации;
- использование виртуальной обучающей среды (LMS Moodle «Электронное образование») или иных информационных систем, сервисов и мессенджеров.

6.3. Программное обеспечение, современные профессиональные базы данных и информационные справочные системы

6.3.1. Программное обеспечение

Наименование программного обеспечения	Назначение
Adobe Reader	Программа для просмотра электронных документов
Платформа дистанционного обучения LMS Moodle	Виртуальная обучающая среда
Google Chrome	Браузер
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Офисная программа
7-zip	Архиватор
Microsoft Windows 10 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты

6.3.2. Современные профессиональные базы данных и информационные справочные системы

1. Универсальная справочно-информационная полнотекстовая база данных периодических изданий ООО «ИВИС» <http://dlib.eastview.com>
2. Электронные версии периодических изданий, размещенные на сайте информационных ресурсов www.polpred.com
3. Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: <https://library.asu-edu.ru/catalog/>.
4. Электронный каталог «Научные журналы АГУ»: <http://journal.asu.edu.ru/issledovaniya-i-innovacii/11745-nauchnye-jurnaly-agu.html>.
5. Корпоративный проект Ассоциации региональных библиотечных консорциумов (АРБИКОН) «Межрегиональная аналитическая роспись статей» (МАРС) <http://mars.arbicon.ru>

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

7.1. Паспорт фонда оценочных средств

При проведении текущего контроля и промежуточной аттестации по дисциплине (модулю) «Процесс организации и использования данных внутри компании на платформах Universe и Плюс 7» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин (модулей) и прохождением практик, а в процессе освоения дисциплины (модуля) – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

Таблица 6. Соответствие разделов, тем дисциплины (модуля), результатов обучения по дисциплине (модулю) и оценочных средств

Контролируемый раздел, тема дисциплины (модуля)	Код контролируемой компетенции	Наименование оценочного средства
Тема 1. Введение в управление корпоративными данными	ПК-3	Лабораторная работа, контрольные вопросы
Тема 2. Интеграция данных	ПК-3	Лабораторная работа, контрольные вопросы
Тема 3. Обеспечение качества данных	ПК-3	Лабораторная работа, контрольные вопросы
Тема 4. Управление НСИ и мастер-данными	ПК-3	Лабораторная работа, контрольные вопросы
Тема 5. Основы DG – руководства на основе данных	ПК-3	Лабораторная работа, контрольные вопросы
Тема 6. Основы VI	ПК-3	Лабораторная работа, контрольные вопросы

7.2. Описание показателей и критериев оценивания компетенций, описание шкал оценивания**Таблица 7. Показатели оценивания результатов обучения в виде знаний**

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует глубокое знание теоретического материала, умение обоснованно излагать свои мысли по обсуждаемым вопросам, способность полно, правильно и аргументированно отвечать на вопросы, приводить примеры
4 «хорошо»	демонстрирует знание теоретического материала, его последовательное изложение, способность приводить примеры, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует неполное, фрагментарное знание теоретического материала, требующее наводящих вопросов преподавателя, допускает существенные ошибки в его изложении, затрудняется в приведении примеров и формулировке выводов
2 «неудовлетворительно»	демонстрирует существенные пробелы в знании теоретического материала, не способен его изложить и ответить на наводящие вопросы преподавателя, не может привести примеры

Таблица 8. Показатели оценивания результатов обучения в виде умений и владений

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы
4 «хорошо»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует отдельные, несистематизированные навыки, испытывает затруднения и допускает ошибки при выполнении заданий, выполняет

Шкала оценивания	Критерии оценивания
«1»	задание по подсказке преподавателя, затрудняется в формулировке выводов
2 «неудовлетворительно»	не способен правильно выполнить задания

7.3. Контрольные задания и иные материалы, необходимые для оценки результатов обучения по дисциплине (модулю)

В данном разделе приведено общее описание лабораторных работ. Более подробное описание хода работы, и также контрольные вопросы, соответствующие ходу выполнения работы, размещаются в курсе на платформе «Электронное образование».

Тема 1 «Введение в управление корпоративными данными»

Лабораторная работа № 1 «Использование конструктора и диспетчера рабочих процессов»

Бизнес-Цель:

Нужно научиться использовать Плюс 7 ФормИТ, чтобы более эффективно выполнять задачи ETL и автоматизировать процесс разработки. Поскольку студенты имели ограниченный опыт работы с этим программным обеспечением или не имели его вообще, эта лабораторная поможет сориентироваться в основных интерфейсах разработки.

Техническое описание:

Плюс 7 ФормИТ включает в себя два приложения для разработки: Designer, который используется для создания мапинга, и Workflow Manager, который используется для создания и запуска рабочих процессов. Эта лабораторная предназначена для первого практического опыта работы с Плюс 7 ФормИТ. Студент импортирует исходные и целевые определения из папки быстрого доступа в свою собственную папку.

Тема 2 «Интеграция данных»

Лабораторная работа №2 «Source Qualifier»

Бизнес-Цель:

Промежуточная область хранилища данных Mersche Motors содержит таблицу, которая назначает описания типов платежей для каждого идентификатора платежа. Поскольку эти описания могут меняться, таблицу необходимо ежедневно синхронизировать с соответствующими данными, находящимися в операционной системе. Администратор операционной системы использует простой плоский файл для записи и редактирования этих описаний.

Техническое описание:

Плюс 7 ФормИТ будет извлекать данные из плоского файла с разделителями и вставлять данные в таблицу базы данных без выполнения преобразований данных. Чтобы избежать дублирования записей при последующих загрузках, мы настроим Плюс 7 ФормИТ на усечение целевой таблицы перед каждой загрузкой.

Лабораторная работа №3 «Загрузите промежуточную таблицу клиентов»

Бизнес-Цель:

Промежуточная область для хранилища данных Mersche Motors имеет таблицу контактов с клиентами. Mersche Motors ежедневно получает новые данные из своих региональных офисов продаж в виде трех текстовых файлов. Текстовые файлы идентичны. Для

простоты обработки Mersche Motors будет использовать функцию Плюс 7 ФормИТ для чтения списка файлов из одного источника. Отображение, которое делает это, будет выполняться по ночному расписанию в полночь.

Техническое описание:

Плюс 7 ФормИТ будет исходить из списка файлов. Этот список файлов содержит имена трех плоских файлов с разделителями из региональных офисов продаж. Все строки с номером клиента 99999 необходимо будет отфильтровать. Существует ряд столбцов, в которых необходимо переформатировать данные, включая подстроки, конкатенацию и декодирование. Целевой объект будет усечен до тех пор, пока сопоставление не будет полностью протестировано.

Лабораторная работа №5 «Загрузить промежуточную таблицу транзакций продаж»

Бизнес-Цель:

Mersche Motors получает данные о транзакциях продаж из своих региональных офисов продаж в виде текстового файла. Данные о транзакциях продаж необходимо ежедневно загружать в промежуточную таблицу.

Техническое описание:

Плюс 7 ФормИТ будет использовать простой файл и реляционную таблицу. Преобразование Joiner используется для создания одного потока данных, который затем записывается в реляционную цель. В плоском файле отсутствует одно поле, необходимое для промежуточной таблицы — стоимость каждого продукта. Это значение можно прочитать из таблицы STG_PRODUCT. Каждая строка исходного файла содержит значение с именем Product. Это значение имеет идентичное соответствующее значение в столбце PRODUCT_ID таблицы STG_PRODUCT. Используйте преобразование Joiner для присоединения плоского файла к реляционной таблице (гетерогенное присоединение), а затем запишите результаты в таблицу STG_TRANSACTION.

Лабораторная работа № 6 «Загрузить промежуточную таблицу сотрудников»

Бизнес-Цель:

Информация о сотрудниках Mersche Motors сохраняется в трех текстовых файлах каждый день. Мы должны прочитать каждый из этих файлов по отдельности и загрузить их в промежуточную область. Файлы не содержат информации о зарплате сотрудников, поэтому мы должны найти зарплату каждого сотрудника. Мы также должны переформатировать некоторые другие поля.

Техническое описание:

У нас ежедневно поступают три текстовых файла с информацией о сотрудниках, которые мы хотели бы поместить в список файлов. Нам нужно найти зарплату для каждого сотрудника, соединить имя и фамилию, изменить формат возраста и номера телефона и добавить дату загрузки.

Тема 3 «Обеспечение качества данных»

Лабораторная работа № 7 «Отладчик»

Бизнес-Цель:

Мэппинг m_STG_DATES_DEBUG содержит по крайней мере одну ошибку, которая приводит к загрузке неверных данных в целевую таблицу. Эту ошибку необходимо найти и исправить, чтобы проект хранилища данных был успешным.

Техническое описание

Отладчик будет использоваться для отслеживания причины ошибки или ошибок.

Лабораторная работа № 8 «Таблица размеров даты загрузки»

Бизнес-Цель:

В хранилище данных Mersche Motors есть таблица измерений дат, которую необходимо загрузить. Измерение даты необходимо загрузить перед любыми другими таблицами измерений.

Техническое описание:

Плюс 7 ФормИТ извлечет даты из общей реляционной таблицы и загрузит их в общую реляционную таблицу. Все столбцы в исходной таблице имеют соответствующие столбцы в целевой таблице. Первичный ключ для целевой таблицы будет назначен с помощью преобразования Sequence Generator.

Тема 4 «Управление НСИ и мастер-данными»

Лабораторная работа № 9 «Загрузить таблицу измерений рекламных акций (Lookup and Persistent Cache)»

Бизнес-Цель:

Mersche Motors проводит ряд рекламных акций, которые начинаются и заканчиваются в определенные даты. Рекламные акции хранятся в таблице измерений рекламных акций. В этой таблице хранятся даты начала и окончания срока действия в виде ключей дат, которые ссылаются на таблицу измерения даты.

Техническое описание:

Для таблицы DIM_PROMOTIONS требуются ключи даты начала и окончания срока действия. Они существуют в таблице DIM_DATES, которая была заполнена в предыдущей лабораторной работе. Чтобы получить эти ключи дат, созданные sequence generator, необходимо выполнить поиск в таблице DIM_DATES в базе данных EDW. Таблица DIM_DATES изменяется нечасто, поэтому она будет загружаться в кэш в постоянном состоянии. Кэш поиска будет часто использоваться другими сопоставлениями, которые загружают таблицы измерений.

Лабораторная работа № 10 «Перезагрузите промежуточную таблицу сотрудников»

Бизнес-Цель:

Данные о сотрудниках Mersche Motors были загружены в таблицу STG_EMPLOYEES, но после проверки данных было установлено, что данные отсутствуют. Хотя поиск в файле salaries.txt и рабочем процессе прошел успешно, разработчик заметил, что в столбце DEALERSHIP_MANAGER целевой таблицы нет данных. Используя предыдущее маппинг, которое изначально загружало данные о сотруднике, разработчик должен указать полное имя менеджера по продажам в столбце DEALERSHIP_MANAGER.

Техническое описание:

Мы скопируем маппинг m_STG_EMPLOYEES_xx, созданное в предыдущей лабораторной работе, и изменим его, чтобы получить имя менеджера и загрузить его в столбец DEALERSHIP_MANAGER таблицы STG_EMPLOYEES. Для этого нам придется разделить данные на два потока. В одном потоке будут все записи о сотрудниках, а в другом — только записи менеджеров, которые нужно будет снова соединить вместе, используя записи менеджеров в качестве основных. В потоке менеджера мы будем фильтровать в столбце POSITION_TYPE записи MANAGER и связывать их обратно с записями SALESREP, используя DEALERSHIP_ID. Это необходимо, потому что в каждом дилерском центре есть только один менеджер. Нам также нужно будет поддерживать поиск в отношении файла salaries.txt, чтобы гарантировать, что данные о зарплате по-прежнему заполнены.

Лабораторная работа № 11 «Загрузить таблицу измерений сотрудников»

Бизнес-Цель:

Таблица сотрудников хранилища данных Mersche Motors обновляется ежедневно. Исходные строки из промежуточной области необходимо проверить, чтобы увидеть, существует ли уже строка в таблице измерений. Строки должны быть помечены для

обновления или вставки соответственно. Любые строки, содержащие неверные данные, должны быть записаны в файл ошибок.

Техническое описание:

Строки из таблицы STG_EMPLOYEES необходимо загрузить в таблицу DIM_EMPLOYEES. Перед загрузкой строк необходимо проверить EMPLOYEE_ID на наличие значений NULL. Недопустимые строки должны быть записаны в файл ошибок. Действительные строки необходимо проверить, чтобы убедиться, что они уже существуют в DIM_EMPLOYEES, и пометить их для INSERT или UPDATE соответственно. Наконец, любые строки, отправляемые в таблицу DIM_EMPLOYEE, должны получать действительные даты из DIM_DATES.

Тема 5 «Основы DG – руководства на основе данных»

Лабораторная работа № 12 «Загрузить таблицу измерений клиентов»

Бизнес-Цель:

В хранилище данных Mersche Motors есть таблица клиентов, которая загружается ежедневно. Многие клиенты посещают дилерские центры более одного раза в день, поэтому логика склада должна иметь возможность отслеживать несколько посещений в один и тот же день. Эта логика должна быть в состоянии проверить, была ли уже загружена запись о покупателе в текущем прогоне, и если да, то что изменилось в покупателе, если что-то изменилось.

Техническое описание:

Плюс 7 ФормИТ будет использовать промежуточную таблицу STG_CUSTOMERS и загрузит таблицу измерений DIM_CUSTOMERS. Данные клиента могут иметь более одного вхождения в источнике. Данные должны быть проверены на наличие новых строк, существующих строк и недопустимых строк. Необходимо будет использовать динамический поиск, поскольку строка клиента может встречаться в источнике более одного раза. В некоторых строках будут нулевые данные, поэтому для их записи будет использоваться журналирование ошибок в неструктурированных файлах.

Лабораторная работа № 13 «Загрузить таблицу фактов о продажах»

Бизнес-Цель:

Дилерские центры Mersche Motors иногда предоставляют агрессивные скидки, выходящие за рамки разрешенного диапазона. Скидки этого типа составляют небольшой процент по сравнению с количеством обрабатываемых строк, но информация должна обрабатываться соответствующим образом. Кроме того, несмотря на то что исходная промежуточная таблица содержит данные за 7 дней, этот производственный запуск загрузит всю промежуточную таблицу в таблицу SALES_FACT в рамках одного рабочего процесса.

Техническое описание:

Необходимая информация находится в двух отдельных промежуточных таблицах. Чтобы усугубить это, связь между двумя таблицами не существует в базе данных. Ссылочная целостность должна быть создана в Плюс 7 ФормИТ. Для обработки скидок вне допустимого диапазона необходимы специальные формулы. Чтобы сделать это более эффективным, будет использоваться использование параметров маппинга и переменных.

Лабораторная работа № 14 «Создание Mapplet»

Бизнес-Цель:

Руководитель группы заметил, что есть и другие ситуации, когда мы можем повторно использовать некоторые преобразования, разработанные в сопоставлении нагрузки FACT_SALES.

Техническое описание:

Чтобы воспользоваться преимуществами ранее созданных объектов, мы создадим мапплет из существующих объектов, использованных в предыдущем сопоставлении. Затем этот мапплет можно использовать в других сопоставлениях.

Тема 6 «Основы VI»

Лабораторная работа № 15 «Загрузить таблицу фактов инвентаризации»

Бизнес-Цель:

Загрузка таблицы фактов инвентаризации выполняется сразу после загрузки промежуточной таблицы инвентаризации. Иногда загрузка промежуточной таблицы инвентаризации выполняется дольше, чем допустимо, и задерживает остальную часть процесса. Электронное письмо необходимо отправить администратору, если промежуточная загрузка занимает больше часа.

Техническое описание:

Группа поддержки предложила создать рабочий процесс, ограничивающий время рабочего процесса. Этот рабочий файл будет содержать задачу «Таймер» для отслеживания времени выполнения, задачу «Электронная почта» для информирования администратора о том, что загрузка займет более 1 часа, и задачу «Управление» для остановки рабочего процесса, если он выполняется более одного часа. Будет создан рабочий процесс, который содержит рабочий лет, а затем выполняет промежуточную загрузку и загрузку фактов.

Перечень вопросов и заданий, вносимых на экзамен / зачёт / дифференцированный зачёт

1. Что такое интеграция данных?
2. Что такое хранилище данных?
3. Как работают хранилища данных?
4. Основные функции ETL-платформы?
5. Какие есть уровни данных?
6. Как лицензируется продукт Плюс 7 ФормИТ?
7. Какие у Плюс 7 ФормИТ есть преимущества продукта перед обычными используемыми средствами?
8. Когда следует отключать кэширование lookup?
9. Должны ли маппинги иметь более одного Pipeline?
10. Возможно ли повторное применение ранее разработанных решений в Плюс 7 ФормИТ?
11. Что такое репозиторий?
12. Что такое интеграционный сервис?
13. Для чего используется преобразование Sequence Generator?
14. Когда применяется Command Tasks?
15. Для чего применяется Propagate?
16. Когда используется «Префикс имени файла кэша»?
17. Трансформация Router. Для чего применяется?
18. Что делать, если в исходных таблицах нет первичного-внешнего ключа?
19. Как обновлять метаданные маппинга?
20. Как создается Lookup?
21. Что лучше применять Joiner или Lookup?
22. Что такое Source Definitions?
23. Что такое Target Definitions?
24. Какие есть преобразования в Плюс 7 ФормИТ?

25. В каком инструменте может создавать и редактировать Source Definitions?

Таблица 9. Примеры оценочных средств с ключами правильных ответов

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
ПК-3. Способен разрабатывать программное обеспечение, включая проектирование, отладку, проверку работоспособности и модификацию ПО				
1.	Задание закрытого типа	Виды преобразований: 1) Пассивные 2) Внутренние 3) Внешние 4) Активные 5) Смешанные	1, 4	1
2.		В каком клиентском приложении Plus 7 FormIT создается маппинг: 1) Designer 2) Workflow Manager 3) Workflow Monitor	1	1
3.		Логика передачи и обработки данных от источников до приемников, выраженная в виде документа или направленного графа в системе, выполняющей интеграцию это: 1) Mapping 2) Workflow 3) Source Qualifier 4) Maplets	1	1
4.		База данных или система файлов для хранения логики обработки данных средством интеграции^ 1) Маппинг 2) Источник 3) Репозиторий 4) Приемник	3	1
5.		Трансформация, которая может объединять реляционный источник и неструктурированный файл по общему значению, например по идентификатору транзакции: 1) Aggregator 2) Joiner 3) Expression 4) Filter	2	1
6.		Задание открытого типа	Базовые данные, которые определяют бизнес-сущности, с которыми имеет дело предприятие (клиенты,	Мастер-данные

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		поставщики, продукция, услуги, договора, счета, пациенты), называются?		
7.		Система, база данных, файл, сервис и т.д., предоставляющие данные для обработки?	Источник	1
8.		Informatica PowerCenter – это	ведущая промышленная платформа интеграции данных, позволяющая эффективно решать любые стоящие перед компанией задачи в области интеграции, синхронизации данных, построения хранилищ данных, MDM-решений, миграции данных в новые приложения, обмена информацией с контрагентами и т.д. для реализации бизнес-целей компании	3
9.		Инструмент, в котором можно создавать и редактировать источники	Source Analyzer	1
10.		Назовите типы портов в преобразованиях	Входной, Выходной	1
11.	Задание комбинированного типа	<p>Маппинги должны содержать:</p> <ol style="list-style-type: none"> 1) Хотя бы один источник 2) Хотя бы одну цель 3) Хотя бы одно преобразование 4) Все вышеуказанные ответы <p>Прокомментируйте свой ответ</p>	<p>4</p> <p>Каждый элемент дополняет друг друга и формирует целостный цикл передачи данных. Если хотя бы один компонент отсутствует, весь процесс становится неполноценным и рискует привести к некорректному результату</p>	3

Полный комплект оценочных материалов по дисциплине (модулю) (фонд оценочных средств) хранится в электронном виде на кафедре, утверждающей рабочую программу дисциплины (модуля).

7.4. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине (модулю)

Отчет по лабораторной работе представляется в электронном виде. Защита отчета проходит в форме доклада студентов по выполненной работе и ответов на вопросы преподавателя. В обсуждении результатов работы могут принимать участие также студенты, не входящие в группу отвечающих. В случае, если оформление отчета и поведение студента во время защиты соответствуют указанным требованиям, студент получает максимальное количество баллов.

По некоторым работам допускается выполнение в группах по 2-3 человека.

Основаниями для снижения количества баллов в диапазоне от max до min являются:

- небрежное выполнение,
- отсутствие выводов.

Отчет не может быть принят и подлежит доработке в случае:

- отсутствия необходимых разделов,
- отсутствия необходимого графического материала.

Таблица 10. Технологическая карта рейтинговых баллов по дисциплине (модулю)

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
Основной блок				
1.	<i>Отчет по лабораторной работе</i>	15/3	45	По расписанию
2.	<i>Ответ на контрольные вопросы</i>	15/3	45	
Всего			90	-
Блок бонусов				
3.	<i>Посещение занятий без пропусков</i>		3	
4.	<i>Своевременное выполнение всех заданий</i>		3	
5.	<i>Успешное участие в профильных олимпиадах, конкурсах</i>		4	
Всего			10	-
ИТОГО			100	-

Таблица 11. Система штрафов (для одного занятия)

Показатель	Балл
<i>Опоздание на занятие</i>	-1
<i>Нарушение учебной дисциплины</i>	-1
<i>Неготовность к занятию</i>	-1
<i>Пропуск занятия без уважительной причины</i>	-1

Таблица 12. Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине (модулю)

Сумма баллов	
90–100	Зачтено
85–89	
75–84	
70–74	
65–69	

Сумма баллов	
60–64	
Ниже 60	Не зачтено

При реализации дисциплины (модуля) в зависимости от уровня подготовленности обучающихся могут быть использованы иные формы, методы контроля и оценочные средства, исходя из конкретной ситуации.

8. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Материал курса построен на основании уникальных материалов и методик разработанных DIS Group.

8.1. Основная литература

1. Тарасов, С. В. СУБД для программиста. Базы данных изнутри / С. В. Тарасов. - Москва : СОЛОН-ПРЕСС, 2020. - 320 с. - ISBN 978-2-7466-7383-0. - URL : <https://www.studentlibrary.ru/book/ISBN9782746673830.html> (ЭБС «Консультант студента»).

8.2. Дополнительная литература

1. Базы данных и системы управления базами данных : учеб. пособие / Е. А. Лазицкас, И. Н. Загумённикова, П. Г. Гилевский - Минск : РИПО, 2018. - 268 с. - ISBN 978-985-503-771-3. - URL : <https://www.studentlibrary.ru/book/ISBN9789855037713.html> (ЭБС «Консультант студента»).

8.3. Интернет-ресурсы, необходимые для освоения дисциплины (модуля)

1. **Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента».** Многопрофильный образовательный ресурс «Консультант студента» является электронной библиотечной системой, предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Каталог в настоящее время содержит около 15000 наименований. www.studentlibrary.ru.

2. <https://dis-group.ru/>

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для обеспечения учебного процесса по дисциплине используются компьютерные классы, библиотеки АГУ, мультимедийные аудитории.

Минимальные требования к ПК:

Компьютер PC или Mac:

не ниже CPU: 1.6 GHz

не менее 6 GB RAM

размер экрана не менее 1024x768 или выше

Операционная система:

Windows 7 и выше

Mac OS X

· Linux

Поддерживаемые браузеры: Firefox, Opera, Chrome, Microsoft EDGE

Доступ к интернет:

ширина интернет-канала на каждого участника не менее 300 кбит/с.

10. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ (МОДУЛЯ) ПРИ ОБУЧЕНИИ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Рабочая программа дисциплины (модуля) при необходимости может быть адаптирована для обучения (в том числе с применением дистанционных образовательных технологий) лиц с ограниченными возможностями здоровья, инвалидов. Для этого требуется заявление обучающихся, являющихся лицами с ограниченными возможностями здоровья, инвалидами, или их законных представителей и рекомендации психолого-медико-педагогической комиссии. При обучении лиц с ограниченными возможностями здоровья учитываются их индивидуальные психофизические особенности. Обучение инвалидов осуществляется также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).

Для лиц с нарушением слуха возможно предоставление учебной информации в визуальной форме (краткий конспект лекций; тексты заданий, напечатанные увеличенным шрифтом), на аудиторных занятиях допускается присутствие ассистента, а также сурдопереводчиков и тифлосурдопереводчиков. Текущий контроль успеваемости осуществляется в письменной форме: обучающийся письменно отвечает на вопросы, письменно выполняет практические задания. Доклад (реферат) также может быть представлен в письменной форме, при этом требования к содержанию остаются теми же, а требования к качеству изложения материала (понятность, качество речи, взаимодействие с аудиторией и т. д.) заменяются на соответствующие требования, предъявляемые к письменным работам (качество оформления текста и списка литературы, грамотность, наличие иллюстрационных материалов и т. д.). Промежуточная аттестация для лиц с нарушениями слуха проводится в письменной форме, при этом используются общие критерии оценивания. При необходимости время подготовки к ответу может быть увеличено.

Для лиц с нарушением зрения допускается аудиальное предоставление информации, а также использование на аудиторных занятиях звукозаписывающих устройств (диктофонов и т. д.). Допускается присутствие на занятиях ассистента (помощника), оказывающего обучающимся необходимую техническую помощь. Текущий контроль успеваемости осуществляется в устной форме. При проведении промежуточной аттестации для лиц с нарушением зрения тестирование может быть заменено на устное собеседование по вопросам.

Для лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата, на аудиторных занятиях, а также при проведении процедур текущего контроля успеваемости и промежуточной аттестации могут быть предоставлены необходимые технические средства (персональный компьютер, ноутбук или другой гаджет); допускается присутствие ассистента (ассистентов), оказывающего обучающимся необходимую техническую помощь (занять рабочее место, передвигаться по аудитории, прочитать задание, оформить ответ, общаться с преподавателем).