

НОЧУ ДПО УЦ «Сетевая Академия»

УТВЕРЖДАЮ
Директор НОЧУ ДПО УЦ «Сетевая Академия»
 /Макарова М.М./



Образовательная программа
дополнительного профессионального образования
(повышения квалификации)
«Реализация решений данных в Microsoft Azure»
(DP-203T00 Data Engineering on Microsoft Azure)

Содержание

Описание образовательной программы	2
Цели программы	3
Планируемые результаты обучения	4
Учебный план	8
Календарный учебный график	10
Рабочая программа	11
Организационно-педагогические условия реализации Программы	11
Формы аттестации и оценочные материалы	15

Описание образовательной программы

Настоящая образовательная программа повышения квалификации (далее – Программа) разработана в соответствии с:

1. Федеральным законом от 29 декабря 2012 г. N 273-ФЗ «Об образовании в Российской Федерации»
2. Приказом Минобрнауки России от 1 июля 2013 г. N 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»
3. Уставом НОЧУ ДПО УЦ «Сетевая Академия»

Структура Программы включает цели, планируемые результаты обучения, учебный план, календарный учебный график, рабочую программу, организационно-педагогические условия, формы аттестации и оценочные материалы.

Цели Программы содержат описание целевой аудитории, целей обучения и необходимых начальных знаний и навыков слушателей.

Планируемые результаты обучения представлены в виде перечня профессиональных компетенций в рамках имеющейся квалификации (с отсылкой к профессиональному стандарту), качественное изменение которых осуществляется в результате обучения.

Учебный план определяет перечень, трудоемкость, последовательность и распределение модулей, иных видов учебной деятельности обучающихся и формы аттестации.

Календарный учебный график определяет основные параметры учебного процесса при организации занятий по освоению настоящей Программы, включая формы обучения, расписание занятий очных групп и т.п.

Рабочая программа раскрывает рекомендуемую последовательность изучения разделов (модулей).

Описание организационно-педагогических условий реализации Программы определяет организационные и методические требования НОЧУ ДПО УЦ «Сетевая Академия» к организации и проведению обучения по Программе.

Формы аттестации и оценочные материалы определяют формы проведения промежуточной и итоговой аттестации по Программе и форму учебно-методических материалов, необходимых для проведения указанных видов аттестации.

Цели программы

Данная Программа предназначена для:

- профессионалов работы с данными;
- архитекторов данных;
- профессионалов в области бизнес-аналитики, которые хотят узнать об инженерии данных и создании аналитических решений с использованием технологий платформы данных, существующих в Microsoft Azure;
- подготовки к сертификационному экзамену DP-203: «Data Engineering on Microsoft Azure» сертификации Microsoft - **Microsoft Certified: Azure Data Engineer Associate**.

Целью обучения является формирование у слушателей знаний и навыков в области технологий вычислений и хранения, которые используются для создания аналитического решения, а также использования методов проектирования данных, связанных с работой с пакетными аналитическими решениями и решениями в реальном времени с использованием технологий платформы данных Azure.

Для изучения данной Программы рекомендуется обладать следующими знаниями и навыками:

- Базовые знания по основам Azure в объеме курса AZ-900T00 «Microsoft Azure Fundamentals» (*рекомендуется пройти обучение*)
- Базовые знания по основам данных Azure в объеме курса DP-900T00 «Microsoft Azure Data Fundamentals» (*рекомендуется пройти обучение*).

Планируемые результаты обучения

Реализация Программы направлена на повышение профессионального уровня в рамках имеющейся квалификации, определяемой профессиональными стандартами «06.026 Системный администратор информационно-коммуникационных систем», утвержденным Приказом Минтруда России от 29.09.2020 №680н «Об утверждении профессионального стандарта «Системный администратор информационно-коммуникационных систем»; «06.001 Программист», утвержденным Приказом Минтруда России от 18.11.2013 №679н (ред. от 12.12.2016) «Об утверждении профессионального стандарта «Программист»» и «06.011 Администратор баз данных», утвержденным приказом Министерства труда и социальной защиты РФ от 17 сентября 2014 г. N 647н «Об утверждении профессионального стандарта «Администратор баз данных».

Результатами обучения по Программе станут знания и умения, соответствующие следующим обобщенным трудовым функциям указанных профессиональных стандартов:

- Обслуживание информационно-коммуникационной системы.
- Обслуживание сетевых устройств информационно-коммуникационной системы.
- Обслуживание серверных операционных систем информационно-коммуникационной системы.
- Проверка работоспособности и рефакторинг кода программного обеспечения.
- Интеграция программных модулей и компонент и верификация выпусков программного продукта.
- Разработка требований и проектирование программного обеспечения.
- Обеспечение функционирования БД.
- Оптимизация функционирования БД.

Совершенствуемые компетенции в соответствии с трудовыми функциями профессионального стандарта:

Компетенция	Содержание компетенции Трудовые функции	Код
Обслуживание информационно-коммуникационной системы	Выполнение работ по выявлению и устранению инцидентов в информационно-коммуникационных системах	В/01.5
	Обеспечение работы технических и программных средств информационно-коммуникационных систем	В/02.5
	Реализация схемы резервного копирования, архивирования и восстановления конфигураций технических и программных средств информационно-коммуникационных систем по утвержденным планам	В/03.5
	Внесение изменений в технические и программные средства информационно-коммуникационных систем по утвержденному плану работ	В/04.5
	Проведение обновления программного обеспечения технических средств информационно-коммуникационных систем по инструкциям производителей	В/05.5
	Диагностика исчерпания типовых ресурсов информационно-коммуникационных систем с	В/06.5

	использованием прикладных программных средств и средств контроля	
Обслуживание сетевых устройств информационно-коммуникационной системы	Выполнение работ по выявлению и устранению сложных инцидентов, возникающих на сетевых устройствах информационно-коммуникационных систем	C/01.6
	Проведение анализа и выявление основных причин сложных проблем, возникающих на сетевых устройствах информационно-коммуникационных систем	C/02.6
	Разработка планов резервного копирования, архивирования и восстановления конфигураций сетевых устройств информационно-коммуникационных систем	C/03.6
	Планирование изменений сетевых устройств информационно-коммуникационных систем предметными специалистами из других областей	C/04.6
	Прогнозирование влияния внешних и внутренних воздействий на поведение сетевых устройств информационно-коммуникационной системы	C/06.6
	Прогнозирование потребности в изменении объемов ресурсов, необходимых для обеспечения бесперебойной работы сетевых устройств информационно-коммуникационных систем	C/07.6
Обслуживание серверных операционных систем информационно-коммуникационной системы	Выполнение работ по выявлению и устранению нетипичных инцидентов, возникающих в серверных операционных системах информационно-коммуникационной системы	D/01.6
	Проведение анализа и определение основных причин сложных проблем, возникающих на серверах и в серверных операционных системах	D/02.6
	Выполнение планирования резервного копирования, архивирования и восстановления конфигурации серверов и серверных операционных систем	D/03.6
	Планирование изменений параметров работы серверов и серверных операционных систем	D/04.6
	Выполнение обновления программного обеспечения серверных операционных систем	D/05.6
	Прогнозирование влияния внешних и внутренних воздействий на поведение серверных операционных систем	D/06.6
	Прогнозирование потребности в изменении объемов необходимых ресурсов для обеспечения бесперебойной работы серверов и серверных операционных систем	D/07.6
	Планирование и проведение работ по распределению нагрузки между имеющимися ресурсами, снятию нагрузки на серверы и серверные операционные системы перед проведением регламентных работ, восстановлению штатной схемы работы в случае сбоев	D/08.6
Разработка и отладка программного кода	Проверка и отладка программного кода	A/05.3

Проверка работоспособности и рефакторинг кода программного обеспечения	Разработка процедур проверки работоспособности и измерения характеристик программного обеспечения	B/01.4
	Разработка тестовых наборов данных	B/02.4
	Проверка работоспособности программного обеспечения	B/03.4
Разработка требований и проектирование программного обеспечения	Анализ требований к программному обеспечению	D/01.6
	Проектирование программного обеспечения	D/03.6
Обеспечение функционирования БД	Резервное копирование БД	A/01.4
	Восстановление БД	A/02.4
	Управление доступом к БД	A/03.4
	Установка и настройка программного обеспечения (ПО) для обеспечения работы пользователей с БД	A/04.4
	Установка и настройка ПО для администрирования БД	A/05.4
	Мониторинг событий, возникающих в процессе работы БД	A/06.4
	Протоколирование событий, возникающих в процессе работы БД	A/07.4
Оптимизация функционирования БД	Мониторинг работы БД, сбор статистической информации о работе БД	B/01.5
	Оптимизация распределения вычислительных ресурсов, взаимодействующих с БД	B/02.5
	Оптимизация производительности БД	B/03.5
	Оптимизация компонентов вычислительной сети, взаимодействующих с БД	B/04.5
	Оптимизация выполнения запросов к БД	B/05.5
	Оптимизация управления жизненным циклом данных, хранящихся в БД	B/06.5

После обучения слушатель сможет:

- изучить варианты вычислений и хранения для проектирования данных рабочих нагрузок в Azure;
- понимать особенности проектирования данных;
- выполнять интерактивные запросы с использованием бессерверных пулов SQL;
- изучать, преобразовывать и загружать данные в хранилище данных с помощью Apache Spark;
- выполнять исследования и преобразования данных в Azure Databricks;
- получать и загружать данные в хранилище данных;
- преобразовывать данные с помощью Azure Data Factory или Azure Synapse Pipelines;
- интегрировать данные из записных книжек с помощью Azure Data Factory или Azure Synapse Pipelines;

- оптимизировать производительность запросов с помощью выделенных пулов SQL в Azure Synapse;
- анализировать и оптимизировать хранилище данных;
- поддерживать гибридную транзакционную аналитическую обработку (Hybrid Transactional Analytical Processing, HTAP) с помощью Azure Synapse Link;
- обеспечивать комплексную безопасность с помощью Azure Synapse Analytics;
- выполнять потоковую обработку в реальном времени с помощью Stream Analytics;
- создавать решения для потоковой обработки с помощью Event Hubs и Azure Databricks.

Учебный план

Учебный план Программы определяет перечень, трудоемкость, последовательность и распределение модулей, иных видов учебной деятельности обучающихся и формы аттестации.

№ п/п	Наименование разделов (модулей)	Всего, час	В том числе		Форма аттестации
			Лекции	Практические занятия	
0.	Введение	1,8	0,5	1,3	Опрос, практические занятия
1.	Возможности вычислений и хранения для инженерии данных	2	0,5	1,5	Опрос, практические занятия
2.	Выполнение интерактивных запросов с использованием бессерверных пулов SQL в Azure Synapse Analytics	2,5	1	1,5	Опрос, практические занятия
3.	Исследование и преобразование данных в Azure Databricks	2	0,5	1,5	Опрос, практические занятия
4.	Изучение, преобразование и загрузка данных в хранилище данных с помощью Apache Spark	2,5	0,7	1,8	Опрос, практические занятия
5.	Получение и загрузка данных в хранилище данных	3	1	2	Опрос, практические занятия
6.	Преобразование данных с помощью фабрики данных Azure Data Factory или конвейера Azure Synapse Pipelines	2,5	0,7	1,8	Опрос, практические занятия
7.	Интеграция данных из Notebooks с помощью Azure Data Factory или Azure Synapse Pipelines	2,5	1	1,5	Опрос, практические занятия
8.	Комплексная безопасность с помощью Azure Synapse Analytics	2,7	0,7	2	Опрос, практические занятия
9.	Сопровождение гибридной аналитической обработки транзакций (HTAP) с помощью Azure Synapse Link	2,7	0,7	2	Опрос, практические занятия
10.	Обработка потоков в реальном времени с помощью Stream Analytics	3	1	2	Опрос, практические занятия

11.	Создание решения для потоковой обработки с концентраторами событий Event Hubs и Azure Databricks	3	1	2	Опрос, практические занятия
12.	Итоговая аттестация	1,8	-	1,8	Тестирование
	Итого:	32	8,3	23,7	

Допускается формирование индивидуального учебного плана для каждого слушателя в пределах осваиваемой Программы в порядке, установленном Положением об организации образовательного процесса в НОЧУ ДПО УЦ «Сетевая Академия».

Календарный учебный график

Учебный год: круглогодичное обучение.

Продолжительность Программы: 32 академических часа.

Форма организации образовательного процесса: очная, очно-заочная (вечерняя) и заочная формы обучения, в том числе, с применением дистанционных образовательных технологий и электронного обучения.

Сменность занятий (при очной форме обучения): I смена.

Количество учебных дней в неделю при очном обучении: 4 дня.

Начало учебных занятий: 9.30

Окончание учебных занятий: 17.00

Продолжительность урока: 45 минут (1 академический час).

Продолжительность перемен: 15 минут, перерыв на обед – 60 минут.

Расписание занятий для очных групп:

	№ урока	Время
Конкретный день недели согласовывается во время учебного процесса	1-2	09:30 - 11:00
	3-4	11:15 - 12:45
	5-6	13:45 - 15:15
	7-8	15:30 - 17:00

Модуль 0: Введение

- *Лабораторная работа: Настройка среды для выполнения лабораторных работ.*

Модуль 1: Возможности вычислений и хранения для инженерии данных

- Введение в Azure Synapse Analytics.
- Введение в Azure Databricks.
- Введение в архитектуру Delta Lake в Azure Databricks.
- Введение в хранилище озера данных Azure.
- Работа с потоками данных с помощью Azure Stream Analytics
- *Лабораторная работа: Изучение вариантов вычислений и хранения для рабочих нагрузок инженерии данных.*

Модуль 2: Выполнение интерактивных запросов с использованием бессерверных пулов SQL в Azure Synapse Analytics

- Изучение возможностей бессерверных пулов SQL в Azure Synapse.
- Выполнение запроса данных в озере с помощью бессерверных пулов SQL Azure Synapse.
- Создание объектов метаданных в бессерверных пулах SQL Azure Synapse.
- Защита данных и управление пользователями в бессерверных пулах SQL Azure Synapse.
- *Лабораторная работа: Выполнение интерактивных запросов с использованием бессерверных пулов SQL.*

Модуль 3: Исследование и преобразование данных в Azure Databricks

- Обзор Azure Databricks.
- Чтение и запись данных в Azure Databricks.
- Работа с DataFrames в Azure Databricks.
- Работа с расширенными методами DataFrames в Azure Databricks.
- *Лабораторная работа: Исследование и преобразование данных в Azure Databricks.*

Модуль 4: Изучение, преобразование и загрузка данных в хранилище данных с помощью Apache Spark

- Введение в проектирование больших данных с помощью Apache Spark в Azure Synapse Analytics.
- Получение данных с помощью записных книжек Apache Spark в Azure Synapse Analytics.
- Преобразование данных с помощью DataFrames в пулах Apache Spark в Azure Synapse Analytics.
- Интеграция пулов SQL и Apache Spark в Azure Synapse Analytics.
- Мониторинг управления рабочими загрузками с помощью Apache Spark в Azure Synapse Analytics.
- *Лабораторная работа: Изучение, преобразование и загрузка данных в хранилище данных с помощью Apache Spark.*

Модуль 5: Получение и загрузка данных в хранилище данных

- Использование лучших практик загрузки данных в Azure Synapse Analytics.
- Прием петабайтов с помощью фабрики данных Azure Data Factory.
- *Лабораторная работа: Получение и загрузка данных в хранилище данных.*

Модуль 6: Преобразование данных с помощью фабрики данных Azure Data Factory или конвейера Azure Synapse Pipelines

- Интеграция данных с фабрикой данных Azure Data Factory.
- Масштабируемое преобразование без кода с помощью Azure Data Factory.
- *Лабораторная работа: Преобразование данных с помощью Azure Data Factory.*

Модуль 7: Интеграция данных из Notebooks с помощью Azure Data Factory или Azure Synapse Pipelines

- Управление перемещением и преобразованием данных в Azure Data Factory.
- *Лабораторная работа: Управление перемещением и преобразованием данных в Azure Synapse Pipelines.*

Модуль 8: Комплексная безопасность с помощью Azure Synapse Analytics

- Защита хранилища данных в Azure Synapse Analytics.
- Настройка и управление секретами в Azure Key Vault.
- Внедрение контроля соответствия для конфиденциальных данных.
- *Лабораторная работа: Комплексная безопасность с помощью Azure Synapse Analytics.*

Модуль 9: Сопровождение гибридной аналитической обработки транзакций (HTAP) с помощью Azure Synapse Link

- Проектирование гибридной транзакционной и аналитической обработки с помощью Azure Synapse Analytics.
- Настройка ссылки Azure Synapse Link с помощью Azure Cosmos DB.
- Выполнение запросов к Azure Cosmos DB с пулами Apache Spark с помощью Azure Synapse Analytics.
- Выполнение запросов к Azure Cosmos DB с бессерверными пулами SQL с помощью Azure Synapse Analytics.
- *Лабораторная работа: Сопровождение гибридной аналитической обработки транзакций (HTAP) с помощью Azure Synapse Link.*

Модуль 10: Обработка потоков в реальном времени с помощью Stream Analytics

- Обеспечение надежного обмена сообщениями для приложений Big Data с помощью концентраторов событий Azure Event Hubs.
- Работа с потоками данных с помощью Azure Stream Analytics.
- Преобразование данных с помощью Azure Stream Analytics.
- *Лабораторная работа: Обработка потоков в реальном времени с помощью Stream Analytics.*

Модуль 11: Создание решения для потоковой обработки с концентраторами событий Event Hubs и Azure Databricks

- Обработка потоковых данных с помощью структурированной потоковой передачи Azure Databricks.
- *Лабораторная работа: Создание решения для потоковой обработки с концентраторами событий и Azure Databricks.*

Организационно-педагогические условия реализации Программы

При реализации Программы применяется форма организации образовательной деятельности, основанная на модульном принципе представления содержания образовательной программы и построения учебных планов, использовании различных образовательных технологий, в том числе дистанционных образовательных технологий и электронного обучения.

Организационные условия реализации программы в разных формах обучения регулируются следующими локальными нормативными актами:

- Положение об организации образовательного процесса в НОЧУ ДПО УЦ «Сетевая Академия».
- Положение о порядке применения электронного обучения, дистанционных образовательных технологий в НОЧУ ДПО УЦ «Сетевая Академия».

Учебные материалы по Программе включают: рабочую программу, раздаточные материалы по курсу, методические материалы по курсу, данные примеров по курсу. Учебное пособие по Программе выдается слушателям в бумажном или электронном виде в зависимости от формы обучения в порядке, установленном Положением о библиотеке в НОЧУ ДПО УЦ «Сетевая Академия».

Занятия по Программе проводятся преподавателями, предварительно подтвердившими свою квалификацию. В числе базовых требований ко всем преподавателям – требование обязательного прохождения программы «Андрагогика. Эффективное обучение взрослых» в форме учебного курса и пробной лекции, а также сдачи технических сертификационных тестов по продукту или технологии, рассматриваемым в курсе.

Освоение Программы сопровождается промежуточной аттестацией обучающихся в формах, определенных учебным планом, и в порядке, установленном Положением об организации образовательного процесса в НОЧУ ДПО УЦ «Сетевая Академия».

Освоение Программы завершается итоговой аттестацией обучающихся в форме, определенной учебным планом, и в порядке, установленном Положением об организации образовательного процесса в НОЧУ ДПО УЦ «Сетевая Академия».

Слушателям, успешно освоившим соответствующую Программу и прошедшим итоговую аттестацию, выдается удостоверение о повышении квалификации на бланке, образец которого самостоятельно устанавливается организацией.

Слушателям, не прошедшим итоговой аттестации или получившим на итоговой аттестации неудовлетворительные результаты, а также лицам, освоившим часть Программы и (или) отчисленным из организации, выдается справка об обучении или о периоде обучения по образцу, самостоятельно устанавливаемому организацией.

Оценочные материалы для промежуточной аттестации по Программе разрабатываются в форме лабораторных работ и/или контрольных вопросов после изучения каждого модуля.

Оценочные материалы для итоговой аттестации по Программе разрабатываются в форме теста.

Контрольные задания и вопросы для оценки знаний и навыков слушателей задаются и выполняются в следующих областях:

- Возможности вычислений и хранения для инженерии данных. Введение в Azure Synapse Analytics, Azure Databricks. Введение в архитектуру Delta Lake в Azure Databricks и хранилище озера данных Azure.
- Работа с потоками данных с помощью Azure Stream Analytics.
- Разработка многомерной схемы для оптимизации аналитических рабочих нагрузок.
- Выполнение масштабного преобразования без кода с помощью фабрики данных Azure.
- Заполнение медленно изменяющихся измерений в конвейерах Azure Synapse Analytics.
- Проектирование современного хранилища данных с помощью Azure Synapse Analytics.
- Защита хранилища данных в Azure Synapse Analytics.
- Выполнение интерактивных запросов с использованием бессерверных пулов SQL в Azure Synapse Analytics.
- Изучение, преобразование и загрузка данных в хранилище данных с помощью Apache Spark.
- Исследование и преобразование данных в Azure Databricks.
- Получение и загрузка данных в хранилище данных.
- Преобразование данных с помощью фабрики данных Azure Data Factory или конвейера Azure Synapse Pipelines.

- Управление перемещением и преобразованием данных в Azure Data Factory или Azure Synapse Pipelines.
- Оптимизация производительности запросов с помощью выделенных пулов SQL в Azure Synapse.
- Анализ и оптимизация хранилища данных.
- Сопровождение гибридной аналитической обработки транзакций (HTAP) с помощью Azure Synapse Link.
- Комплексная безопасность с помощью Azure Synapse Analytics.
- Обработка потоков в реальном времени с помощью Stream Analytics.
- Создание решения для потоковой обработки с концентраторами событий Event Hubs и Azure Databricks.
- Создание отчетов с использованием интеграции Power BI с Azure Synapse Analytics.
- Выполнение интегрированных процессов машинного обучения в Azure Synapse Analytics.