

федеральное государственное автономное образовательное учреждение высшего образования
Первый Московский государственный медицинский университет имени И.М. Сеченова
Министерства здравоохранения Российской Федерации
(Сеченовский Университет)

**Дополнительная профессиональная программа
(программа профессиональной переподготовки)**

Специалист по большим данным

(наименование программы)

дополнительное профессиональное образование

(подвид дополнительного образования)

Утверждена Центральной методической комиссией 12.07.2022

Проректор по учебно-воспитательной работе,
Председатель Центральной методической комиссии
ФГАОУ ВО Первый МГМУ им. И.М. Сеченова
Минздрава России (Сеченовский Университет)

Т.М.Литвинова

I. Общие положения

1. Дополнительная профессиональная программа (программа профессиональной переподготовки) ИТ-профиля «Специалист по большим данным» (далее – Программа, ДПП ПП) разработана в соответствии с нормами Федерального закона РФ от 29 декабря 2012 года № 273-ФЗ «Об образовании в Российской Федерации», с учетом требований приказа Минобрнауки России от 1 июля 2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам», с изменениями, внесенными приказом Минобрнауки России от 15 ноября 2013 г. № 1244 «О внесении изменений в Порядок организации и осуществления образовательной деятельности по дополнительным профессиональным программам, утвержденный приказом Министерства образования и науки Российской Федерации от 1 июля 2013 г. № 499», приказа Министерства образования и науки РФ от 23 августа 2017 г. N 816 «Об утверждении Порядка применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ»; паспорта федерального проекта «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации»; постановления Правительства Российской Федерации от 13 мая 2021 г. № 729 «О мерах по реализации программы стратегического лидерства «Приоритет-2030» (в редакции постановления Правительства Российской Федерации от 14 марта 2022 г. № 357 «О внесении изменений в постановление Правительства Российской Федерации от 13 мая 2021 г. № 729»); приказа Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации от 28 февраля 2022 г. № 143 «Об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации» и признании утратившими силу некоторых приказов Министерства цифрового

развития, связи и массовых коммуникаций Российской Федерации об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации» (далее – приказ Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации № 143); федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.03.02 Информационные системы и технологии (уровень бакалавриата), утвержденного приказом Минобрнауки России от 19 сентября 2017 г. (С изменениями и дополнениями от: 26 ноября 2020 г., 8 февраля 2021 г.), (далее вместе – ФГОС ВО)), а также профессионального стандарта 06.042 «Специалист по большим данным», утвержденного приказом Министерства труда и социальной защиты РФ от 6 июля 2020 г. № 405н.

2. Профессиональная переподготовка заинтересованных лиц (далее – Слушатели), осуществляемая в соответствии с Программой (далее – Подготовка), имеющей отраслевую направленность¹ «Информационно-коммуникационные технологии», проводится в Федеральном государственном автономном образовательном учреждении высшего образования Первый Московский государственный медицинский университет имени И.М. Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет) (далее – Университет) в соответствии с учебным планом в очной/заочной форме обучения².

3. Разделы, включенные в учебный план Программы, используются для последующей разработки календарного учебного графика, учебно-тематического плана, рабочей программы, оценочных и методических материалов. Перечисленные документы разрабатываются Университетом самостоятельно, с учетом актуальных положений законодательства об образовании, законодательства в области информационных технологий и

¹ Варианты отраслевой направленности: «Городское хозяйство»; «Финансовые услуги»; «Строительство»; «Добывающая промышленность»; «Обрабатывающая промышленность»; «Транспортная инфраструктура»; «Здравоохранение»; «Энергетическая инфраструктура»; «Образование»; «Сельское хозяйство и агропромышленный комплекс»; «Информационно-коммуникационные технологии»; «Искусство и культура»

² При реализации Программы допускается использовать сетевую форму обучения с организациями реального сектора экономики субъекта Российской Федерации

смежных областей знаний ФГОС ВО и профессионального стандарта 06.042«Специалист по большим данным».

4. Программа регламентирует требования к профессиональной переподготовке в области создание информационных технологий нового поколения, обеспечивающих экономически эффективное извлечение полезной информации из больших объемов разнообразных данных путем высокой скорости их сбора, обработки и анализа, и применение этих технологий в информационно-аналитической деятельности, в системах управления и принятия решений, а также для разработки на их основе новых продуктов и услуг.

Срок освоения Программы составляет 252 часа.

К освоению Программы в рамках проекта допускаются лица:

- получающие высшее образование по очной (очно-заочной) форме, лица, освоившие основную профессиональную образовательную программу (далее – ОПОП ВО) бакалавриата – в объеме не менее первого курса (бакалавры 2-го курса), ОПОП ВО специалитета – не менее первого и второго курсов (специалисты 3-го курса), а также магистратуры, обучающиеся по специальностям и направлениям подготовки ИТ-сферы.

5. Область профессиональной деятельности: Создание и применение технологий больших данных (Профстандарт 06.042).

II. Цель

6. Целью подготовки слушателей по Программе является получение компетенции³, необходимой для выполнения нового вида профессиональной деятельности в области информационных технологий: Большие данные. Сбор, обработка и анализ больших данных в организации, анализ больших данных с использованием существующей в организации методологической и

³Указать целевые группы обучающихся, определенные паспортом Федерального проекта: – обучающиеся по специальностям и направлениям подготовки, не отнесенным к ИТ-сфере, – обучающиеся по специальностям и направлениям подготовки ИТ-сферы (выбрать нужное)

технологической инфраструктуры, анализ больших данных в проектах под контролем опытных специалистов, приобретение новой квалификации «Специалист по большим данным».

III. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации

7. Виды профессиональной деятельности, трудовая функция, указанные в профессиональном стандарте по соответствующей должности «Аналитик», «Исследователь данных», «Руководитель (специалист) отдела по информационным технологиям», представлены в таблице 1:

Таблица 1

Характеристика новой квалификации, связанной с видом профессиональной деятельности и трудовыми функциями в соответствии с профессиональным стандартом «06.042 «Специалист по большим данным».

Область профессиональной деятельности	Тип задач профессиональной деятельности	Код и наименование профессиональной компетенции	Трудовые действия	Трудовая функция	Обобщенная трудовая функция	Вид профессиональной деятельности
06 Связь, информационные и коммуникационные технологии (в сфере исследования, разработки, внедрения и сопровождения информационных технологий и систем); 40 Сквозные виды	Создание информационных технологий нового поколения, обеспечивающих экономически эффективное извлечение полезной информации из больших объемов разнообразных данных путем высокой скорости их сбора, обработки и анализа, и применение этих	ОПК-1 (2) Способность понимать принципы работы современных информационных технологий ОПК-2 (3) Способность решать стандартные задачи профессиональной деятельности с применением информационно-коммуникационных технологий ОПК-3 (7) Способность осуществлять выбор платформ и инструментальных программно-аппаратных средств для реализации информационных систем ПК -1 Способность определение возможностей	Выявление требований заказчика к результатам анализа, определение возможностей применения анализа больших данных в предметной области и конкретных задачах заказчика. Консультирование заказчика по возможностям имеющейся методологической и технологической инфраструктуры анализа больших данных и результатам	А/01.6 Выявление, формирование и согласование требований к результатам аналитических работ с применением технологий больших данных. А/02.6 Планирование и организация аналитических работ с использованием технологий больших данных. А/03.6 Подготовка данных для проведения аналитических работ по исследованию больших данных. А/04.6 Проведение аналитического	Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры	<i>Специалист по большим данным</i>

<p>профессиональной деятельности в промышленности (в сфере организации и проведения научно-исследовательских и опытно-конструкторских работ в области информатики и вычислительной техники).</p>	<p>технологий в информационно-аналитической деятельности, в системах управления и принятия решений, а также для разработки на их основе новых продуктов и услуг</p>	<p>применения анализа больших данных</p> <p>ПК-2 Способность консультировать заказчика по возможностям имеющейся методологической и технологической инфраструктуры для анализа больших данных</p> <p>ПК-3 Способность согласовывать с заказчиком и утверждение требований к результатам аналитического исследования</p> <p>ПК-4 Выполняет проекты по анализу больших данных. Создает и поддерживает отказоустойчивые решения в ИТ-системах</p>	<p>применения технологий больших данных к аналогичным задачам.</p> <p>Согласование с заказчиком и утверждение требований к результатам аналитического исследования</p>	<p>исследования с применением технологий больших данных в соответствии с требованиями заказчика</p>		
--	---	---	--	---	--	--

Таблица 2

Характеристика новой и развиваемой цифровой компетенции в ИТ-сфере, связанной с уровнем формирования и развития в результате освоения Программы⁴ «Специалист по большим данным»

Наименование сферы	Код и наименование профессиональной компетенции	Пример инструментов	0 — способность не проявляется/ проявляется в степени, недостаточной для отнесения к 1 уровню сформированности компетенции	1 — способность проявляется под внешним контролем / при внешней постановке задачи/ обучающийся пользуется готовыми, рекомендованными продуктами	2 — способность проявляется, но обучающийся эпизодически прибегает к экспертной консультации/ самостоятельно подбирает и пользуется готовыми продуктами
Большие данные	<p>ОПК-1 (2) Способность понимать принципы работы современных информационных технологий</p> <p>ОПК-2 (3) Способность решать стандартные задачи профессиональной деятельности с применением информационно-коммуникационных технологий</p> <p>ОПК-3 (7) Способность осуществлять выбор платформ и инструментальных программно-аппаратных</p>	<p>Python, Pandas, SciPy, scikit-learn</p> <p>Big Data tools (Hadoop)</p>	<p>Не использует большие данные</p>	<p>Анализирует большие данные в проектах под контролем опытных специалистов</p>	<p>Выполняет проекты по анализу больших данных. Создает и поддерживает отказоустойчивые решения в ИТ-системах</p>

⁴ На основании Модели цифровых компетенций, указанной в Приложении 2

	<p>средств для реализации информационных систем</p> <p>ПК -1 Способность определение возможностей применения анализа больших данных</p> <p>ПК-2 Способность консультировать заказчика по возможностям имеющейся методологической и технологической инфраструктуры для анализа больших данных</p> <p>ПК-3 Способность согласовывать с заказчиком и утверждение требований к результатам аналитического исследования</p> <p>ПК-4 Выполняет проекты по анализу больших данных. Создает и поддерживает отказоустойчивые решения в ИТ-системах</p>				
--	---	--	--	--	--

IV. Характеристика новых и развиваемых цифровых компетенций, формирующихся в результате освоения программы

8. В ходе освоения Программы Слушателем приобретаются следующие профессиональные компетенции:

- Анализирует большие данные в проектах под контролем опытных специалистов;
- Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика;
- Выявление, формирование и согласование требований к результатам аналитических работ с применением технологий больших данных.

(Код и наименование профессиональной компетенции Таблица 1)

ПК-1 Определение возможностей применения анализа больших данных

ПК-2 Знание инфраструктуры анализа больших данных

ПК-3 Утверждение требований к результатам исследования

В ходе освоения Программы Слушателем совершенствуются следующие профессиональные компетенции:

ОПК-1 (2) Способность понимать принципы работы современных информационных технологий

ОПК-2 (3) Способность решать стандартные задачи профессиональной деятельности с применением информационно-коммуникационных технологий

ОПК-3 (7) Способность осуществлять выбор платформ и инструментальных программно-аппаратных средств для реализации информационных систем

(Код и наименование профессиональной компетенции Таблица 2)

ПК -формирования цифровых компетенций в области создания алгоритмов и компьютерных программ, пригодных для практического применения:

ПК-1 применения анализа больших данных в предметной области и конкретных задачах заказчика.

ПК-2 способность использовать методологическую и технологическую инфраструктуру анализа больших данных.

ПК-3 проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика

V. Планируемые результаты обучения по ДПП ПП

10. Результатами подготовки слушателей по Программе является получение компетенции, необходимой для выполнения нового вида профессиональной деятельности в области информационных технологий Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры; приобретение новой квалификации «**Специалист по большим данным**».

11. В результате освоения Программы слушатель должен:

Знать:

1. Инструменты и методы согласования с заказчиками требований к результатам аналитических исследований с использованием технологий больших данных.
2. Регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных.
3. Технологии межличностной и групповой коммуникации в деловом взаимодействии, основы конфликтологии.
4. Технологии подготовки и проведения презентаций.
5. Предметную область анализа больших данных в соответствии с требованиями заказчика.
6. Возможности имеющейся у исполнителя методологической и технологической инфраструктуры анализа больших данных.
7. Современный опыт использования анализа больших данных.
8. Теоретические и прикладные основы анализа данных.
9. Типы анализа больших данных, виды аналитики.
10. Современные методы и инструментальные средства анализа больших данных.
11. Стандарты проведения анализа данных.
12. Источники информации, в том числе информации, необходимой для

обеспечения деятельности в предметной области заказчика исследования.

13. Методы интерпретации и визуализации больших данных.

14. Правила деловой переписки

Уметь:

1. Проводить переговоры с целью выявления требований заказчика к результатам анализа, формировать и согласовывать требования к результатам аналитических работ с использованием технологий больших данных.
2. Проводить презентации при консультировании заказчика, согласовании и утверждении требований к результатам аналитических работ с использованием технологий больших данных
3. Подготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.
4. Использовать имеющуюся у исполнителя методологическую и технологическую инфраструктуру анализа больших данных для выполнения аналитических работ.
5. Проводить сравнительный анализ методов и инструментальных средств анализа больших данных.
6. Проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.

Иметь навыки:⁵

1. Выявления требований заказчика к результатам анализа, определение возможностей применения анализа больших данных в предметной области и конкретных задачах заказчика.
2. Консультирования заказчика по возможностям имеющейся методологической и технологической инфраструктуры анализа больших данных и результатам применения технологий больших

⁵ Выделяются знания и умения в соответствии с профстандартом, связанные с результатами освоения Программы

данных к аналогичным задачам.

3. Согласование с заказчиком и утверждение требований к результатам аналитического исследования.

VI. Организационно-педагогические условия реализации ДПП ПП

12. Реализация Программы должна обеспечить получение компетенции, необходимой для выполнения нового вида профессиональной деятельности в области информационных технологий Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры; приобретение новой квалификации **«Специалист по большим данным»**.

13. Учебный процесс организуется с применением⁶электронного обучения, дистанционных образовательных технологий, инновационных технологий и методик обучения, способных обеспечить получение слушателями знаний, умений и навыков в области⁷Создания и применения технологий больших данных (Код 06.042).

14. Реализация Программы обеспечивается научно-педагогическими кадрами Университета, допустимо привлечение к образовательному процессу высококвалифицированных специалистов ИТ-сферы и/или дополнительного профессионального образования в части, касающейся профессиональных компетенций в области создания алгоритмов и программ, пригодных для практического применения, с обязательным участием представителей профильных организаций-работодателей. Возможно привлечение региональных руководителей цифровой трансформации (отраслевых ведомственных и/или корпоративных) к проведению итоговой аттестации, привлечение работников организаций реального сектора экономики субъектов Российской Федерации.

⁶ При необходимости указать нужное — электронного обучения, дистанционных образовательных технологий

⁷ Разрабатывается на основе ФГОС ВО (3++), соответствует разделу 1.11 ФГОС ВО и конкретному профстандарту

VII. Учебный план ДПП

15. Объем Программы составляет 252 часа.

16. Учебный план Программы определяет перечень, последовательность, общую трудоемкость разделов и формы контроля знаний.

Учебный план программы профессиональной переподготовки

«Специалист по большим данным»

№ п/п	Наименование раздела (модуля)	Общая трудоемкость (252 часа)	Форма контроля
1.	Введение в анализ данных. Типы медицинских данных	30	тест
2.	Бизнес аналитика и визуализация медицинских данных	30	тест
3.	Исследовательская статистика медицинских данных	30	тест
4	Прикладная статистика в здравоохранении	30	тест
5	Работа с данными. Основы работы с БД	30	тест
6	Языки программирования для работы с данными. Библиотеки Python	30	тест
7	Основы машинного обучения	30	тест
8	Прикладной анализ данных на примере отрасли	30	тест
	Промежуточная аттестация	По 2 часа в каждом модуле	Тесты по модулям 1-8
	Итоговая аттестация	12	Тест и выпускная квалификационная работа
	Итого:	252	

VIII. Календарный учебный график

18. Календарный учебный график представляет собой график учебного процесса, устанавливающий последовательность и продолжительность обучения и итоговой аттестации по учебным дням.

IX. Рабочая программа учебных предметов, курсов, дисциплин (модулей)

19. Рабочая программа содержит перечень разделов и тем, а также рассматриваемых в них вопросов с учетом их трудоемкости.

Рабочая программа разрабатывается Университетом с учетом профессионального стандарта 06.042 Специалист по большим данным.

№ п/п	Наименование и краткое содержание раздела(модуля)	Объем, часов
1.	<p>Введение в анализ данных. Типы медицинских данных</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • <i>Понятие информации</i> • <i>Предпосылки возникновения больших данных</i> • <i>Компоненты больших данных</i> • <i>Источники больших данных в биомедицине</i> • <i>Информационные технологии, основанные на больших данных</i> • <i>Уровни оценки в медицине и здравоохранении</i> • <i>Виды дизайнов исследований в биомедицине</i> • <i>Классификация научных исследований</i> <p><u>Краткое содержание:</u></p> <p><i>Каждый этап развития человечества и последующего научно-технического прогресса сопровождался информационными революциями: речь, письменность, книгопечатание, теле-радиовещание, персональный компьютер, интернет. Объединение всех компьютеров в глобальную сеть породило информационный взрыв, после которого поток информации только нарастает. Однако, природные когнитивные способности человека остаются на прежнем уровне и не позволяют обработать и проанализировать большие объемы данных даже с помощью персонального компьютера. Это послужило предпосылкой для создания специализированных средств по обработке и анализу первичных данных на удаленных серверах. Совмещение информационного потока и облачных вычислений стало называться “большими данными”.</i></p> <p><i>Источниками больших данных в биомедицине являются: ученые, клиники, лаборатории, пациенты, аптеки. Большие данные, вместе с технологиями машинного обучения, нейросетей, искусственного интеллекта и блокчейна, послужили развитию таких направлений, как: телемедицина, электронное здравоохранение, систем поддержки принятия врачебных решений. Это позволило повысить уровень оценки эффективности и качества отрасли здравоохранения: доказательная медицина, оценка технологий здравоохранения, доказательная организация здравоохранения.</i></p> <p><i>Современные методы обработки и анализа данных вместе с компьютерными технологиями объединились в новую науку о данных - Data Science. При этом эпидемиология, как наука еще со времен Гиппократ развивала свой математический аппарат на</i></p>	28+2(тест)

	<p>данных о распространении болезней среди населения. И сейчас, когда математические операции стало возможным автоматизировать в виде программных алгоритмов, эта наука о жизни, находясь на стыке медицины, математики, социологии, демографии и экономики, способна совершить прорыв, объединив естественные науки с гуманитарными.</p>	
2.	<p>Бизнес аналитика и визуализация медицинских данных</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • Круговая диаграмма • Столбиковая диаграмма • Линейная диаграмма • Гистограмма • Коробчатая диаграмма • Картографирование <p><u>Краткое содержание</u></p> <p>Результаты всех видов исследований в биомедицине можно визуализировать с помощью основных типов диаграмм: круговая, столбиковая, линейная, гистограмма, коробчатая диаграмма. Некоторые специфические исследовательские алгоритмы подразумевают построение более сложной инфографики, такой как картографирование.</p>	28+2(тест)
3	<p>Исследовательская статистика на основе медицинских данных</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • Обоснование актуальности исследования • Фрейминг проблемы исследования • Выявление исходных знаний • Идентификация теоретической базы • Постановка исследовательской гипотезы • Выбор дизайна исследования • Установление границ исследования • Получение и обработка данных • Анализ данных и интерпретация результатов • Передача и использование исследовательских знаний <p><u>Краткое содержание</u></p> <p>Любое исследование в области наук о жизни предназначено для того чтобы ответить на поставленный научный вопрос. Чтобы ответить на этот вопрос с высокой уверенностью, нужны точные результаты. Правильное определение основной гипотезы и плана исследования уменьшит количество ошибок при изучении явления. План исследования включает исследовательский вопрос, проверяемую гипотезу, план эксперимента, методы сбора данных, перспективы анализа данных и связанные с этим затраты. Крайне важно проводить исследование на основе трех основных принципов экспериментальной статистики: рандомизация, повторяемость и внутренний контроль.</p> <p>Исследовательский вопрос включает актуальность, объект, предмет и цель исследования. Чтобы определить способ постановки научного вопроса, требуется исчерпывающий обзор литературы.</p> <p>Как только цель исследования определена, необходимо предложить</p>	28+2(тест)

	<p><i>возможные ответы на исследовательский вопрос, превратив этот вопрос в гипотезу. Основное предложение называется нулевой гипотезой (H0) и обычно основывается на постоянном знании темы или очевидных фактах явления, подкрепленных обзором литературы. С другой стороны, альтернативной гипотезой является отрицание нулевой гипотезы. Обычно исследование направлено на то, чтобы понять влияние какого-либо явления на популяцию.</i></p>	
4	<p>Прикладная статистика в здравоохранении <u>Основные темы:</u></p> <ul style="list-style-type: none"> • <i>Математический и вычислительный минимум</i> • <i>Теория вероятности</i> • <i>Численные методы оптимизации</i> <p><u>Краткое содержание</u> Понятие вектор, матрица, тензор и операции с ними. Единичная и обратная матрица. Нормы. Специальные матрицы и их свойства – диагональные, симметричные, ортогональные. Матричные разложения. Оператор следа. Определитель. Понятие вероятности. Случайные величины. Вероятностные распределения, непрерывные и дискретные. Условная вероятность, формула Байеса. Математическое ожидание, дисперсия и ковариация случайных величин. Основные распределения случайных величин, смеси распределений. Задача вычислительной оптимизации. Градиентные методы поиска экстремумов. Задачи оптимизации с ограничениями.</p>	28+2(тест)
5	<p>Работа с данными. Основы работы с БД <u>Основные темы:</u></p> <ul style="list-style-type: none"> • <i>Введение в Big Data (Большие данные)</i> • <i>Подготовка данных</i> <p><u>Краткое содержание</u> <i>Большие данные и цифровизация бизнеса</i> <i>Базовые принципы и отличия от классических подходов к работе с данными</i> <i>Обзор методологии CRISP DM – модели жизненного цикла аналитики данных: получение данных, подготовка данных, планирование модели, построение модели, проверка результатов, внедрение</i> <i>Отраслевая специфика аналитики больших данных. Общий разбор отраслевых сценариев (cases)</i> <i>Тенденции в подходах Big Data и что актуально на сегодняшний день</i> <i>Технологии Big Data в условиях импортозамещения</i> <i>Подготовка данных – подходы Data Science: нормализация, очистка, выборки, enrichment, форматирование данных</i> <i>Подготовка данных – как процесс формирования Data Pipeline:</i> <i>Процессы ETL и ELT</i> <i>Зонирование Data Lake и сегментирование данных.</i> <i>Назначение и сравнительная характеристика компонент экосистемы Apache Hadoop, NoSQL, DWH, платформ потоковой обработки для хранения и обработки Big Data на примерах (отраслевые сценарии)</i></p>	28+2(тест)

	<p><i>Инструменты оркестрирования и автоматизации (Data Flow)</i> <i>И снова о Метаданных: Data Lineage, Data Provenance, Data Governance, ...</i> <i>Безопасность больших данных</i> <i>Основные трудности и проблемы, критические факторы успеха</i> <i>Специалисты и их компетенции на данной стадии</i></p>	
6	<p>Языки программирования для работы с данными. Библиотеки Python</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • Введение в языки программирования Python • Управляющие конструкции • Функции и коллекции • Модули и пакеты • Регулярные выражения и основы синтаксического разбора • Работа с файловой системой <p><u>Краткое содержание</u> <i>Версии и реализации языка Python. Основные принципы работы интерпретатора Python. Установка интерпретатора и среды разработки используемых на курсе. Основные элементы программирования. Ветвление. Циклы. Простой ввод и простой вывод. Практикум. Словарь. Список. Кортеж. Последовательность.. Функция. Генератор-функция. Подсистема <i>pip</i>. Краткий обзор стандартной библиотеки. Установка стороннего модуля. Понятие о регулярном выражении. Синтаксис регулярных выражений. Применение регулярных выражений</i> <i>Работа с файлами. Работа с каталогами. Обработка параметров командной строки</i></p>	28+2(тест)
7	<p>Основы машинного обучения</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • классификация, • регрессия, • транскрипция, • машинный перевод, • структурный вывод, • обнаружение аномалий <p><u>Краткое содержание</u> <i>Задачи машинного обучения – классификация, регрессия, транскрипция, машинный перевод, структурный вывод, обнаружение аномалий, подстановка (прогнозирование) отсутствующих значений, преобразования и шумоподавления. Меры качества. Емкость обучения, пере- и недообучение. Регуляризация, функции штрафов. Оценки, смещение, дисперсия. Значения смещения и дисперсии для минимизации СКО. Оценка максимального правдоподобия. Байесовская статистика. Алгоритмы обучения с учителем. Алгоритмы обучения без учителя. Стохастический градиентный спуск. Проблемы глубокого обучения – проклятие размерности, регуляризация, обучение многообразий.</i></p>	28+2(тест)
8	<p>Прикладной анализ данных на примере отрасли</p> <p><u>Основные темы:</u></p> <ul style="list-style-type: none"> • Структура научной публикации IMRAD 	28+2(тест)

	<ul style="list-style-type: none"> • Система цифровых идентификаторов объекта DOI • Блокчейн • Анализ сетей, графов • Систематический обзор, мета-анализ • Парсинг, текст-майнинг • Машинное обучение, нейросети • Искусственный интеллект <p><u>Краткое содержание</u></p> <p>Глобальная стандартизация в области проведения научно-исследовательских работ, публикации научных статей и присвоения цифровых идентификаторов объектам делает возможным автоматизировать многие рутинные процессы. Так например, обнаруживать однотипные данные в ходе мета-анализа можно при помощи парсинга по регулярным выражениям и текст-майнинга. Выделять концептуальные смыслы в непрерывном потоке научных знаний становится возможным с применением анализа сетей (графов) и машинного обучения.</p> <p>Основопологающим фактором децентрализации информации обладает новая блокчейн-технология. В созданных с ее помощью распределенных сетях осуществляется разметка однотипных признаков (метаданных) в непрерывном информационном потоке, которые в последующем становятся материалом в процессе машинного обучения при разработке нейросетей и искусственного интеллекта.</p>	
	Завершение курса, итоговая аттестация, выпускная работа	12

20. Учебно-тематический план Программы определяет тематическое содержание, последовательность разделов и (или) тем и их трудоемкость.

№ п/п	Наименование раздела(модуля)	Количество часов		
		аудиторных		самостоятельной работы (выполнение тестовых* заданий)
		Лекции	Семинары	
1.	Введение в анализ данных. Типы медицинских данных	28		2
2.	Бизнес аналитика и визуализация медицинских данных	28		2
3.	Исследовательская статистика медицинских данных	28		2

4.	Прикладная статистика в здравоохранении	28		2
5	Работа с данными. Основы работы с БД	28		2
6	Языки программирования для работы с данными. Библиотеки Python	28		2
7	Основы машинного обучения	28		2
8	Прикладной анализ данных на примере отрасли	28		2
9	Завершение курса, итоговая аттестация, выпускная работа			10 (подготовка и защита выпускной квалификационной работы; 2(итоговый тест)
1-8	Промежуточная аттестация	Тесты по модулям 1-8		
9	Итоговая аттестация	12 Тест и выпускная квалификационная работа		

**указать вид (-ы) запланированной самостоятельной работы*

Х. Формы аттестации

21. Слушатели, успешно выполнившие все элементы учебного плана, допускаются к итоговой аттестации.

Итоговая аттестация по Программе проводится в формате демонстрационного экзамена с участием представителей профильных организаций работодателей.

22. Лицам, успешно освоившим Программу (в области создания алгоритмов и программ, пригодных для практического применения, или навыков использования и освоения цифровых технологий, необходимых для выполнения нового вида профессиональной деятельности) и прошедшим итоговую аттестацию в рамках проекта «Цифровые кафедры», выдается документ о квалификации: диплом о профессиональной переподготовке.

При освоении ДПП ПП параллельно с получением высшего образования диплом о профессиональной переподготовке выдается не ранее получения соответствующего документа об образовании и о квалификации (за исключением лиц, имеющих среднее профессиональное или высшее образование).

23. Лицам, не прошедшим итоговую аттестацию или получившим на итоговой аттестации неудовлетворительные результаты, а также лицам, освоившим часть Программы и (или) отчисленным из Университета, выдается справка об обучении или о периоде обучения по образцу, самостоятельно устанавливаемому Университетом.

XI. Оценочные материалы

24. Контроль знаний, полученных слушателями при освоении разделов (модулей) Программы, осуществляется в следующих формах:

- входной контроль дает представление о базовых знаниях и проводится в форме комплексной оценки (ассесмента);

- текущий контроль успеваемости – обеспечивает оценивание хода освоения разделов Программы, проводится в форме теста после изучения каждой темы;

- промежуточная аттестация проводится с целью оценки динамики изменения уровня сформированности цифровых компетенций у обучающихся по завершению изучения отдельного модуля Программы, проводится в форме комплексной оценки (ассесмента);

- итоговая аттестация – проводится в форме комплексной оценки (ассесмента) и выполнения выпускной квалификационной работы завершает изучение всей программы.

25. В ходе освоения Программы каждый слушатель выполняет следующие отчетные работы:

№ п/п	Наименование раздела (модуля)	Задание	Критерии оценки
1.	Введение в анализ данных. Типы медицинских данных	тест	Не менее 70% правильных ответов
2.	Бизнес аналитика и визуализация медицинских данных	тест	Не менее 70% правильных ответов
3.	Исследовательская статистика медицинских данных	тест	Не менее 70% правильных ответов
4.	Прикладная статистика в здравоохранении	тест	Не менее 70% правильных ответов
5	Работа с данными. Основы работы с БД	тест	Не менее 70% правильных ответов
6	Языки программирования для работы с данными. Библиотеки Python	тест	Не менее 70% правильных ответов
7	Основы машинного обучения	тест	Не менее 70% правильных ответов
8	Прикладной анализ данных	тест	Не менее 70% правильных ответов

	на примере отрасли		
9	Завершение курса, итоговая аттестация, выпускная работа	Тест + выпускная квалификационная работа	Не менее 70% правильных ответов на вопросы теста и положительная оценка квалификационной работы
	Промежуточная аттестация	Тесты по темам 1-8	Не менее 70% правильных ответов
	Итоговая аттестация	Итоговый тест + выпускная квалификационная работа	Положительные оценки ассесмента и квалификационной работы

26/27. Текущий контроль. Перечень примерных 8 заданий

Модуль (тема модуля при наличии)

26. 1. «Введение в анализ данных»

26.2. Бизнес аналитика и визуализация

26.3. Исследовательская статистика

26.4. Прикладная статистика

26.5. Работа с данными. Основы работы с БД

26.6. Языки программирования для работы с данными. Базовый Python

26.7. Основы машинного обучения

26.8. Прикладной анализ данных на примере отрасли

27. Промежуточная аттестация. Перечень примерных 10 заданий

1 Информационная система-это

А Любая система обработки информации

Б Система обработки текстовой информации

В Система обработки графической информации

Г Система обработки табличных данных

Д Нет верного варианта

2 Разновидность информационной системы, в которой реализованы функции централизованного хранения и накопления обработанной информации организованной в одну или несколько баз данных это

А Банк данных

Б База данных

В Информационная система

Г Словарь данных

Д Вычислительная система

3 Совокупность специальным образом организованных данных, хранимых в памяти вычислительной системы и отображающих состояние объектов и их взаимосвязей в рассматриваемой предметной области - это

Ф База данных

Б СУБД

В Словарь данных

Г Информационная система

Д Вычислительная система

4 Комплекс языковых и программных средств, предназначенный для создания, ведения и совместного использования БД многими пользователями - это

А СУБД

Б База данных

В Словарь данных

Г Вычислительная система

Д Информационная система

- 5 Подсистема банка данных, предназначенная для централизованного хранения информации о структурах данных, взаимосвязях файлов БД друг с другом, типах данных и форматах их представления, принадлежности данных пользователям, кодах защиты и разграничения доступа и т.п. — это
- А Словарь данных
 - Б Информационная система
 - В Вычислительная система
 - Г СУБД
 - Д База данных.
- 6 Лицо или группа лиц, отвечающих за выработку требований к БД, ее проектирование, создание, эффективное использование и сопровождение - это
- А Администратор базы данных
 - Б Диспетчер базы данных
 - В Программист базы данных
 - Г Пользователь базы данных
 - Д Технический специалист
- 7 (1балл) Совокупность взаимосвязанных и согласованно действующих ЭВМ или процессов и других устройств, обеспечивающих автоматизацию процессов приема, обработки и выдачи информации потребителям - это
- А Вычислительная система
 - Б Информационная система
 - В Словарь данных
 - Г СУБД
 - Д База данных
- 8 Модель представления данных - это
- А Логическая структура данных, хранимых в базе данных
 - Б Физическая структура данных, хранимых в базе данных
 - В Иерархическая структура данных
 - Г Сетевая структура данных
 - Д Нет верного варианта
- 9 Наиболее используемая (в большинстве БД) модель данных
- А Реляционная модель
 - Б Сетевая модель данных
 - В Иерархическая модель данных
 - Г Системы инвертированных списков
 - Д Все вышеперечисленные варианты
- 10 Назовите вариант ответа, который не является уровнем архитектуры СУБД
- А Физический уровень
 - Б Внутренний уровень
 - В Внешний уровень
 - Г Концептуальный уровень
 - Д Все выше перечисленные варианты
- 12 Что из перечисленного является примером неструктурированных данных?
- А Данные семейного анамнеза
 - Б Рецепт
 - В Результат общего анализа крови
 - Г Результат УЗИ

- 13 Что из перечисленного является примером описания лекарственной формы?
А Таблетка
Б 10 мг
В 2 раза в день
Г Перорально
- 14 Какая переменная используется в базах данных для связи таблицы пациентов с таблицей приемов?
А Patient_ID
Б Patient_name
В Birth_date
Г Visit_date
- 15 Какое утверждение является правильным для структуры баз данных?
А Наименования переменных находятся в столбцах, записи находятся в строках
Б Наименования переменных находятся в строках, записи находятся в столбцах
В Наименования столбцы находятся в переменных, наименования строки находятся в записях
Г Наименования строк находятся в переменных, наименования столбцов находятся в записях
- 16 Что из перечисленного является особенностями реляционных баз данных, которых нет в электронных таблицах?
А Проверка типа данных
Б Имеется функция форматирования текста
В Данные записаны в таблицу, у которой есть один или несколько столбцов
Г Все данные записаны в одной таблице
- 17 Что из перечисленного является особенностями реляционных баз данных, которых нет в электронных таблицах?
А Данные записаны в нескольких связанных таблицах
Б Имеется функция форматирования текста
В Данные записаны в таблицу, у которой есть один или несколько столбцов
Г Все данные записаны в одной таблице
- 18 Какая из терминологических пар является логически верной
А Переменная - колонка
Б Переменная - строка

- В Переменная - таблица
- Г Переменная - наблюдение

19 Какая из терминологических пар является логически верной

- А Запись - строка
- Б Запись - колонка
- В Запись - таблица
- Г Запись - переменная

20 Реляционная база данных содержит:

- А Таблицы, связанные друг с другом одним или несколькими столбцами
- Б Одну таблицу
- В Список команд для запроса данных из таблиц
- Г Список переменных и типов данных

21 Какой из типов объединений таблиц используется, если вы хотите включить в объединение все строки из второй таблицы вместе с данными из первой таблицы, в которой есть совпадения?

- А Правое
- Б Левое
- В Полное
- Г Внутреннее

22 Какой из типов объединений таблиц используется, если вы хотите включить в объединение все строки из первой таблицы вместе с данными из второй таблицы, в которой есть совпадения?

- А Левое
- Б Правое
- В Полное
- Г Внутреннее

23 Какой из типов объединений таблиц используется, если вы хотите включить в объединение все данные первой и второй таблиц?

- А Полное
- Б Правое
- В Левое
- Г Внутреннее

28. Итоговая аттестация. Перечень тем выпускных квалификационных работ и тестовых заданий

Обоснование метода дистанционного мониторинга здоровья юных спортсменов на основе мобильных технологий

1. Основные принципы организации автоматизированного учета медицинской помощи
2. Динамика COVID-19: гетерогенная модель
3. Применение методов искусственного интеллекта для распознавания патологий на фотографиях глазного дна
4. Искусственные нейронные сети для мониторинга лежачих пациентов
5. Картирование моделей избыточного веса и истощения детей в странах с низким и средним уровнем дохода
6. Создание системы поддержки принятия медицинских решений с использованием доказательной медицины
7. Совершенствование форм статистической отчетности в научно-исследовательских учреждениях Минздрава России
8. Удаленное распознавание человеческих эмоций с помощью глубокого машинного обучения искусственных нейронных сетей
9. Научные подходы к цифровизации мониторинга ассортимента лекарственных средств с использованием искусственных нейронных сетей

ХII. Материально-техническое и учебно-методическое обеспечение Программы

- Персональный компьютер/ноутбук (не ниже не хуже Intel Core i5 или AMD Ryzen 5, 8 ГБ RAM, 256ГБ SSD) BYOD
- iPython Notebook
- Google Colaboratory
- IDE PyCharm

ХIII. Список литературы

- Федеральный закон № 273-ФЗ «Об образовании в Российской Федерации» от 29.12.2012г.
- Федеральный проект «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации», 2022 г.
- Постановления Правительства Российской Федерации от 13 мая 2021 г. № 729 «О мерах по реализации программы стратегического лидерства «Приоритет-2030» (в редакции постановления Правительства Российской Федерации от 14 марта 2022 г. № 357 «О внесении изменений в постановление Правительства Российской Федерации от 13 мая 2021 г. № 729»)
- Приказ Минтруда РФ № 148н от 12.04.2013 «Об утверждении уровней квалификации в целях разработки проектов профессиональных стандартов»

- Приказ Министерства труда и социальной защиты Российской Федерации от 6 июля 2020 года N 405н «Об утверждении профессионального стандарта «Специалист по большим данным»» (<http://profstandart.rosmintrud.ru/reestr-professionalnyh-standartov.>)
- Приказ Минобрнауки России от 1 июля 2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»
- Приказ Минобрнауки России от 23 августа 2017 г. № 816 «Об утверждении Порядка применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ»
- Приказ Минцифры России от 28 февраля 2022 г. № 143 «Об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации»
- Приказ от 19 сентября 2017 г. № 926 «Об утверждении Федерального государственного образовательного стандарта Высшего образования - бакалавриат по направлению подготовки 09.03.02 Информационные системы и технологии» (С изменениями и дополнениями от: 26 ноября 2020 г., 8 февраля 2021 г.)
- Протокол заседания комиссии Министерства науки и высшего образования Российской Федерации по проведению отбора российских образовательных организаций высшего образования в целях участия в программе стратегического академического лидерства «Приоритет-2030» № 1 от 26 сентября 2021 г.
- Методические рекомендации по разработке основных профессиональных программ и дополнительных профессиональных программ с учётом соответствующих профессиональных стандартов (утверждены Министром образования и науки РФ 22 января 2015 г. № ОЛ-1/05 вн)
- Концепция реализации результата «Обучающимся обеспечена возможность прохождения профессиональной переподготовки в рамках проекта «Цифровые кафедры» образовательной организации высшего образования – участника программы стратегического академического лидерства «Приоритет-2030» посредством получения дополнительной квалификации по ИТ-профилю» федерального проекта «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации», 2022 г.
- Рекомендации к дополнительным профессиональным программам (программам профессиональной переподготовки) ИТ-профиля, реализуемым в рамках проекта «Цифровые кафедры» образовательной организации высшего образования – участника программы стратегического академического лидерства «Приоритет-2030» посредством получения дополнительной квалификации по ИТ-профилю федерального проекта «Развитие кадрового потенциала ИТ-отрасли»

национальной программы «Цифровая экономика Российской Федерации», 2022 г.

- Журавлев А.Е., Макшанов А.В., ТындыкаръА.Н.. Большие данные. Big Data: учебник для вузов. – М.: Лань, 2021. –188с.
- Еременко К.А. Работа с данными в любой сфере: учеб.пособие для студ.вузов. – М:Альпина, 2021.–292с.
- НатанМарц, Большиеданные. Принципыипрактикапостроениямасштабируемыхсистемобработкиданных вреальномвремени / НатанМарц,ДжеймсУоррен,– М.: Вильямс. -2016. - 292 с.
- Дубров А.М., Многомерные статистические методы / А.М. Дубров,В.С. Мхитарян,Л.И.Трошин, – М.: Дрофа. 2003г., 352 с.
- Лебедев Г.С., Фомина И.В., Шадеркин И.А., Лисненко А.А., Рябков И.В., Качковский С.В., Мелаев Д.В. Основные направления развития интернет технологий в здравоохранении (систематический обзор). // Социальные аспекты здоровья населения. – 2017, №5 (56). URL: <http://vestnik.mednet.ru/content/view/923/30/lang.ru/> (Дата обращения 10.12.2017 г.)
- Клименко Г.С.. Лебедев Г.С., Стародубов В.И. Основные направления развития интернет-технологий в здравоохранении. // Федеральный справочник. Здравоохранение России. [информационно-аналитическое издание]; Т.18 / Центр стратегических программ. – М: Центр стратегических программ, 2018 – с.295-304.
- Лебедев Г.С., Невзорова Д.В., Клименко Г.С., Шадеркин И.А., Ким С.Р., Ашенбреннер И.В. Применение цифровых технологий в паллиативной медицинской помощи. // Pallium: паллиативная и хосписная помощь. Российский научно-практический журнал. №2(3), 2019
- Касимовская Н.А., Лебедев Г.С., Полещук И.А. Оценка готовности медицинских сестер к введению цифровых технологий и системы видеоконтроля при уходе за тяжелобольными (маломобильными) пациентами // Социальные аспекты здоровья населения. – 2020, №66 (5).