Утверждено на заседании

Академического Совета программы от 11.05.2018 протокол № 1

**Концепция образовательной программы Системы больших данных**

**по направлению подготовки 38.04.05 «Бизнес-информатика»**

**(квалификация (степень) магистр)**

Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики» (далее – НИУ ВШЭ) имеет Лицензию Федеральной службы по надзору в сфере образования и науки на право ведения образовательной деятельности рег. № 2593 от 24 мая 2017 года.

Магистерская программа «Системы больших данных» по направлению 38.04.05 разработана в соответствии с «Положением о магистерской подготовке (магистратуре) в системе многоуровневого высшего образования Российской Федерации» и «Образовательным стандартом федерального государственного автономного образовательного учреждения высшего профессионального образования “Национальный исследовательский университет “Высшая школа экономики”», утвержденным учёным советом Национального исследовательского университета “Высшая школа экономики” (Протокол от 02.02.2018 г. № 1)

**1. Общая характеристика образовательной программы**

Основная идея, заложенная в концепцию магистерской программы, - влияние технологий больших данных на деятельность крупных предприятий, возхможность повысить эффективность управления за счет новых технологий.

Термин **«**BigData», возникший в последнее десятилетие, непосредственно связан с появлением проблемы стремительного увеличения объемов данных, возникающих, прежде всего, в бизнес-структурах вследствие полнофункциональной информатизации бизнес-процессов; в научных организациях - вследствие появления новых возможностей для измерений; в государственных организациях - при увеличении объема и функционала сервисов, а также в общественных коммуникационных сетях.

Проблема больших данных становится очевидной из попыток их определения. В качестве определяющих характеристик для больших данных (с точки зрения проблем их обработки) отмечают *«три V»*: объём (*volume*), скорость (*velocity,* в смысле как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов), многообразие (*variety*, в смысле возможности одновременной обработки различных типов структурированных и полуструктурированных данных). Очень часто к ним добавляют другие “V”, например, veracity (в смысле проблемы недостоверности, непроверенности данных) и визуализацию (visualization, в смысле необходимости новых подходов к представлению результатов анализа данных, в том числе, предсказательного и моментального). Сейчас этот термин понимается не только как, собственно, данные, но и как ряд технологий и методов их практического использования. Например, Forrester определяет понятие Big Data как технологию в области аппаратного и программного обеспечения, которая объединяет, организует, управляет и анализирует данные, характеризующиеся «четырьмя V»: объемом (Volume), разнообразием (Variety), изменчивостью (Variability) и скоростью (Velocity).

Проблемы обработки больших данных особенно остро поставили проблему объединения усилий математиков, программистов, специалистов в области инженерии бизнеса и бизнес-анализа, а также, как считают многие авторы, дизайнеров и психологов для разработки принципов визуализации.

Очевидно, что поиск решения проблемы больших данных также связан с интересом бизнеса получать доход от использования огромных объемов сохраняемой информации, компенсирующий расходы на ее хранение. Сегодня эти вопросы также требуют проработки специалистами в области экономики.

Национальный научный фонд США приравнял тематику Big Data к научной сфере, анонсировав новые области финансирования междисцисплинарных исследований по Большим данным.

В обращении «Core Techniques and Technologies for Advancing Big Data Science & Engineering» говорится о развитии ключевых научно-технических подходов к управлению, анализу, визуализации и сбору информации из объемных, распределенных и гетерогенных источников, что, в частности, позволит решить задачи, с которыми ранее не удавалось справиться. Новые методы, инструментарий и инфраструктура обеспечат инновационное развитие науки, техники, медицины, торговли, образования и национальной безопасности, что, как рассчитывает NSF, обеспечит конкурентоспособность США на десятилетия вперед.

Более того, данная научная инициатива входит в состав государственной долгосрочной стратегии для работы с Big Data в национальном масштабе, включающей согласованные действия в крупных государственных, инженерных, биомедицинских, образовательных и трудовых проектах.

NSF выделил 10 млн долл Калифорнийскому университету Беркли на проект Expeditions in Computing по разработке новых алгоритмов машинного обучения, способных обрабатывать масштабные гетерогенные массивы данных.

Кроме того, в рамках программы Cyberinfrastructure Framework for 21st Century Science and Engineering (CIF210) осуществляется прием заявок на реализацию EarthCube – организацию инфраструктуры Больших данных реального времени для наук о Земле.

1,4 млн долл уйдет группе, в которой биологи вместе с математиками продолжат создание цифровых моделей работы белков на базе накопленных массивов данных, и 2 млн долл - на развитие учебных инфраструктур по тематике Big Data. Новая профессия – исследователь данных (DataScientist), в компетенции которого входит работа с большими данными, широко обсуждается образовательным и профессиональным сообществом, с каждым годом появляется все больше соответствующих образовательных программ.

Важными предпосылками реализации программы является успешная реализация программ, связанных с использованием информационных технологий в области бизнес-аналитики, моделирования бизнес-процессов предприятия, проектирования корпоративных информационных систем.

**Целью магистерской программы** является подготовка специалистов, способных оценить влияние технологий больших данных на деятельность крупных предприятий и предложить варианты эффективного использования этих технологий в управлении предприятиями.

Основными потребителями выпускников программы являются крупные компании, имеющие практику хранения больших массивов данных, в том числе и во внешних дата-центрах и заинтересованные в использовании огромных накопленных массивов информации. В выпускниках, обладающих компетенциями в области больших данных, заинтересованы также ИТ-компании и исследовательские организации. Особая сфера деятельности выпускников – образовательные организации, предполагающие развивать современные модели обучения на основе образовательной и академической аналитики.

**Основные задачи реализации программы:**

1. Создание новой, конкурентоспособной на мировом уровне образовательной программы, способной обеспечить позиционирование НИУ ВШЭ среди лучших университетов мира. С этой целью предусматривается активное позиционирование программы, как за счет привлечения в совет программы зарубежных ученых, так и за счет содействия компаний-партнеров программы.

2. Подготовка специалистов в области разработки и внедрения систем больших данных, востребованность которых на отечественном и мировом рынке сегодня очень высока. Появление таких специалистов в отечественных организациях и предприятиях обеспечит возможности внедрения новых, более эффективных подходов к управлению.

3. Создание вместе с компаниями-партнерами облачной инфраструктуры для обеспечения доступа обучающихся к программному инструментарию технологий больших данных.

4. Развитие новых аналитических компетенций у преподавателей и студентов в ходе реализации программы, что будет способствовать развитию навыков использования в учебном процессе образовательной аналитики.

**Важной особенностью и преимуществом** предлагаемой программы является ее ориентация именно на потребности бизнеса в новой технологии. Программа концентрирует свое внимание на предприятии, как на системе с развитой информационной инфраструктурой, обеспечивающей автоматизацию решения управленческих задач. В этом аспекте появляется возможность выделить классы управленческих функций, где внедрение технологий больших данных позволит получить новые характеристики деятельности для принятия решений, прогноза и формирования управляющих воздействий. Подход «от задач к технологии» позволит развить компетенции специалистов, которые ускорят внедрение технологии больших данных в практику, сделают и саму технологию, и предлагаемый инструментарий более востребованными.

**2. Целевая аудитория МП: критерии набора студентов**

Целевая аудитория магистерской программы:

* Выпускники бакалавриата и специалитета в области информационных технологий, менеджмента, экономики, а также других направлений, заинтересованные в использовании информационных технологий в сфере управления предприятиями, как с практической точки зрения, так и с точки зрения базы для научных исследований.
* Абитуриенты, имеющие опыт практической работы в ИТ-подразделениях компаний, заинтересованные в развитии новых перспективных компетенций для дальнейшего карьерного или академического роста

Программа рассчитана на российских и иностранных абитуриентов. Технологии больших данных развиваются и продвигаются, в первую очередь, компаниями – мировыми лидерами в ИТ, которые привлекаются для поддержки практической и научной составляющей образовательного процесса, поэтому для российских абитуриентов обучение на англоязычной программе привлекательно, так как позволяет получить опыт профессионального общения на иностранном языке, что важно для последующего трудоустройства.

Аудитория иностранных абитуриентов – это, в первую очередь, студенты из стран бывшего Советского Союза, Восточной Европы, Ближнего Востока, Южной, Юго-Восточной и Восточной Азии, стран БРИКС, заинтересованных в изучении технологий и, одновременно, в знакомстве с российским бизнесом. Здесь важно отметить, что для большинства ИТ- компаний, например, при формировании академических альянсов, Россия входит в один регион со странами Восточной, Юго-Восточной и Южной Азии.

Многие из этих студентов рассматривают перспективы дальнейшей работы в России. В ряде стран имеются государственные программы поддержки обучения за рубежом, причем именно в ИТ.

Прием на магистерскую программу осуществляется по портфолио. При оценке портфолио подтверждение входных компетенций в области программирования, разработки баз данных, применения или разработки ИТ, анализа и проектирования бизнес-процессов, бизнес-аналитики, экономико-математического моделирования, а также опыт практической работы являются важными, но не обязательными для поступления на программу.

**3. Характеристика сегмента рынка образовательных услуг, основные конкуренты, сравнительные преимущества МП**

Первоначально рынок образовательных услуг слабо реагировал на появление новой технологии. Эксперты полагали, что специалисты области больших данных могут быть обеспечены существующими программами в области анализа данных. статистики и традиционной бизнес-аналитики. Поэтому вопросы решались за счет отдельных концентраций в существующих программах на технологиях больших данных.

Первые программы с названием Big Data, Data Science появились пару лет назад, что связано с пришедшим пониманием трансформирующего влияния этой технологии на бизнес, а также существенным отличием этой технологии от традиционных методов анализа.

Бизнес-анализ является процессом анализа результатов, достигнутых бизнесом в определенный период времени, между тем как скорость обработки больших данных позволяет сделать анализ предсказательным, способным предлагать бизнесу рекомендации на будущее. Технологии больших данных позволяют также анализировать больше типов данных в сравнении с инструментами бизнес-аналитики, что дает возможность фокусироваться не только на структурированных хранилищах.

В белой книге `Информационная архитектура Oracle: руководство архитектора по большим данным` (Oracle Information Architecture: An Architect's Guide to Big Data), указывается, что при работе с большими данными мы подходим к информации иначе, чем при проведении бизнес-анализа.

При сходстве целей, отличие технологии больших данных и бизнес-анализа состоит в следующем (Мэтт Слокум, O'Reilly Radar):

* Технологии больших данных предназначены для обработки более значительных объемов информации.
* Большие данные предназначены для обработки более быстро получаемых и меняющихся сведений, что означает глубокое исследование и интерактивность. В некоторых случаях результаты формируются быстрее, чем загружается веб-страница.
* Большие данные предназначены для обработки неструктурированных данных, требуются алгоритмы и возможность диалога для облегчения поиска тенденций, содержащихся внутри этих массивов

Осмысление новой технологии, большое количество высокоуровневых научных конференций, посвященных этой тематике, а также ряд успешных внедрений этой технологии привели к резкому увеличению числа соответствующих образовательных программ. Большая часть программ имеет технологический уклон, программ с уклоном во внедрение систем, управление информационными структурами, использующими такие технологии, пока мало. Одним из хороших примеров является программа университета Carnegie Mellon.

Это делает предлагаемую программу крайне перспективной, особенно, если учесть, что российские ИТ-компании, ориентированы, прежде всего, на внедрение, интеграцию и консалтинг. Для иностранных студентов программа будет привлекательна именно за счет ее ориентации на применение технологий больших данных в российском бизнесе.

**4. «Портрет выпускника», анализ востребованности и возможностей трудоустройства выпускников**

«Портрет выпускника» магистерской программы в значительной мере определяется устойчивой тенденцией развития в ИТ-отрасли: информация, сервисы более важны для потребителя, чем ПО.

Основное отличие специалистов по изучению данных от, например, аналитиков, - это умение видеть логические связи в системе собранной информации, и на основании этого разрабатывать те или иные бизнес-решения. Специалисты по изучению данных собирают информацию, строят модели на ее основании и активно применяют количественный анализ.

Мнение Майкла Раппа (Michael Rappa), директора Института аналитики в Университете Северной Каролины: "Эти специалисты (по изучению данных) должны уметь извлекать нужную информацию из всевозможных источников, включая информационные потоки в режиме реального времени, и анализировать ее для дальнейшего принятия бизнес-решений. Дело не только в объеме обрабатываемой информации, но также в ее разнородности и скорости обновления".

100% выпускников разработанного Институтом аналитики курса для специалистов по изучению данных получили предложения о работе еще до того, как они завершили обучение.

Программа нацелена на формирование междисциплинарных компетенций:

* Математические и технологические знания и навыки для выбора, оценки, анализа и использования инструментария и технологий больших данных;
* Компетенции, обеспечивающие понимание архитектуры бизнеса, влияния внедрения новых ИТ-технологий, в том числе технологий больших данных, на эффективность управления предприятием, изменение цепочки добавленной стоимости;
* Управленческие компетенции в области внедрения систем больших данных и сервисов на основе технологий больших данных;
* Исследовательские компетенции в области аналитики больших данных, стохастической оптимизации, предсказательного моделирования, прогнозирования, управления данными предприятия, бизнес-анализа, экономико-математического моделирования.

Программа ориентирована на подготовку выпускников, способных:

* проводить работу по внедрению и оценке эффективности технологий и инструментария больших данных на предприятии
* управлять данными предприятия (Data Management);
* внедрять и применять основанный на технологиях больших данных инструментарий аналитики и поддержки принятия решений, осуществлять управление решениями (Decision Management);
* разрабатывать новые модели информационной инфраструктуры предприятия с учетом возможностей технологий больших данных (Model Management).

Основные компетенции выпускника соответствуют компетенциям стандарта направления «Бизнес-информатика», существенно расширяя их за счет новых компетенций в области аналитики больших данных. Согласно стандарту выпускник магистратуры по направлению «Бизнес-информатика» способен:

* Применять методы системного анализа и моделирования для анализа архитектуры предприятий;
* Разрабатывать стратегию развития архитектуры предприятия;
* Разрабатывать и внедрять компоненты архитектуры предприятия;
* Проводить исследования и поиск новых моделей и методов совершенствования архитектуры предприятия;
* Консультировать по вопросам развития ИТ-инфраструктуры предприятия;
* Управлять внедрением инноваций для развития архитектуры предприятия.

**Возможные места работы выпускников магистратуры**

* ИТ-подразделения коммерческих предприятий и организаций
* ИТ-компании
* Научные организации
* Образовательные организации

**5. Структура учебного плана, как она обеспечивает реализацию целей и задач МП**

Программа состоит из обязательных дисциплин, вариативных дисциплин, курсовой работы (первый год обучения), научного семинара и магистерской диссертации (второй год обучения).

**Обязательные дисциплины**

1. System Analysis&Organization Design (Системный анализ и организационный дизайн)
2. Economic and Mathematic Modeling (Экономико-математическое моделирование)
3. Enterprise Architecture Modeling (Моделирование архитектуры предприятия)
4. Advanced Data Analysis&Big Data for Business Intelligence (Перспективные методы анализа данных и Большие данные в бизнес-интеллекте)
5. Big Data Systems Development and Implementation (Pазработка и внедрение систем больших данных)

**Вариативные модули**

**Технологический блок**

1.Data Visualization (Визуализация данных)

2. Predictive Modeling (Предсказательное моделирование)

3. Natural Language Processing and Cognitive Modelling (Обработка естественного языка и когнитивное моделирование)

4. Cloud Computing (Облачные вычисления)

5. Big Data Collection, Storage&Processing in Heterogeneous Distributed Computer Networks (Сбор, хранение и обработка данных в гетерогенных распределенных компьютерных сетях)

6. Knowledge Discovery in Data at Scale Technologies (Технологии извлечения знаний из большого объема данных)

7. Applied Machine Learning (Прикладные аспекты машинного обучения)

**Управленческий блок**

1. Creating and Managing Enterprise Information Assets (Создание и управление информационными активами предприятия )

2. Advanced Data Management (Современный менеджмент данных)

3. Big Data Based Marketing Analytics (Маркетинговая аналитика на основе больших данных)

4. Big Data Based Risk Analytics (Анализ рисков на основе больших данных)

5. Process Mining and Big Data Driving Process Management **(**Анализ процессов и управление процессами на основе больших данных)

6. Big Data Analytics for Industrial Internet (Аналитика больших данных в индустриальном интернете)

Студент может выбирать дисциплины из обоих блоков в соответствии с индивидуальными входными компетенциями и ориентируясь на указанные в программе курса пререквизиты, а также курсы из других магистерских программ (на английском языке).

**Адаптационные курсы**

Enterprise Architecture (Архитектура предприятия)

Data Analysis (Анализ данных)

**Курсовая работа**

Выбор темы курсовой работы осуществляется в соответствии с примерной тематикой курсовых работ или выполняется по тематике, предложенной студентом по согласованию с преподавателем.

Примерная тематика курсовых работ формируется по следующим направлениям с учетом использования концепции, технологии или методов больших данных:

* Анализ информационных активов предприятия (по отраслям)
* Информационное моделирование архитектуры и бизнес-процессов предприятия
* Методы и технологии обработки больших объемов разнородных данных
* Применение технологий больших данных в управлении предприятием
* Экономические аспекты использования больших данных

Курсовая работа рассматривается как важный этап в процессе подготовки итоговой магистерской диссертации и, как правило, готовит студента к ее написанию.

**6. Организация проектной работы и концепция научно-исследовательского (проектного) семинара**

Научно-исследовательский семинар ориентирован на исследование областей применения технологий больших данных, изучение практической работы с инструментарием больших данных, а также на анализ развития этой технологии.

Основной **целью** научно-исследовательского семинара является формирование академических компетенций в области анализа и оценки влияния новых информационных технологий, в том числе технологий больших данных и связанных с ними технологий, на эффективность бизнеса и его архитектуру, а также практическое освоение технологий больших данных при проведении исследований.

Основными **задачами** семинара являются:

* обучение студентов навыкам академической работы, включая подготовку и проведение научных исследований, написание научных работ;
* обучение навыкам научного обсуждения и презентации идей, концепций, результатов исследований, проектов и исследовательских работ;
* обучение использованию технологий больших данных при проведении научных исследований;
* обучение методам и навыкам использования научного прогнозирования для определения технологических трендов в области информационных технологий

Конечная задача семинара – сделать научную работу студентов постоянным и систематическим элементом учебного процесса, включить их в жизнь научного сообщества, помочь освоить методологию, технологию и инструментарий научно-исследовательской деятельности.

Ряд занятий семинара проводятся в компаниях-партнерах программы, которые являются разработчиками продуктов для больших данных (IBM, SAP, EMC).

Программа работы с партнерами является важной частью обучения. Предварительные договоренности с партнерами предполагают не только обеспечение доступа к работе с новейшими программными продуктами в режиме облачных сервисов, что не требует создания собственной технологической инфраструктуры, но и стажировки, как в российских подразделениях, так и в зарубежных исследовательских лабораториях компаний. Часть научных и проектных семинаров проводится в режиме телеконференций. Это позволит привлечь к учебному процессу и проектной работе ведущих ученых, обеспечить участие студентов в реальных проектах.

Тематически семинар разделен на 4 раздела:

1. Математические и технологические основы инструментария больших данных – исследование особенностей применения технологий на задачах, связанных с формированием информационной инфраструктуры предприятия, новые возможности для аналитики и принятия решений.

2. Архитектурные решения на основе больших данных для предприятия: трансформация системы управления данными, формирование информационных активов предприятия, системы сбора данных о бизнес-процессах, обработка внешней информации, новый дизайн взаимодействия, взаимодействие систем.

3. Экономика больших данных: оценка экономической эффективности решений для управления предприятием на основе технологии больших данных, возможности использования информационных активов предприятия, возможности использования в управлении предприятием решений на основе неструктурированной информации из разных источников

4. Перспективы развития функциональности и областей применения технологии

Основными формами проведения семинара являются семинарские занятия в аудитории и в компьютерных классах в НИУ ВШЭ, а также занятия на территории компаний-партнеров программы, которые являются разработчиками продуктов и технологий с использованием концепции больших данных IBM, SAP, ФОРС и других.

Важной частью семинара является подготовка студентами исследовательских материалов в формате статей или выступлений на конференциях, а также работа в реальных проектах компаний.

**Порядок организации семинара**

Научно-исследовательский семинар проводится в течение 1-4 модулей первого и второго года обучения. На протяжении всего этого периода он включает еженедельные аудиторные занятия по утвержденному расписанию и самостоятельную работу студентов. Работа в  научно-исследовательском семинаре является для студента обязательной. При этом он вправе посещать научно-исследовательские семинары других кафедр на добровольной основе без обязательств по сдаче зачетов.

**Формы работы на семинаре**

Работа НИС формируется на пересечении двух циклов деятельности:

* цикл подготовки магистерской диссертации,
* цикл организационных форм научно-исследовательского семинара.

С самого начала научно-исследовательский семинар ориентирован на подготовку магистерской диссертации. Промежуточной формой ее подготовки в рамках первого года обучения является написание курсовой работы. Цикл подготовки магистерской диссертации является основным, под него подстраивается цикл организационных форм научно-исследовательского семинара, который, по существу, является способом организации публичного обсуждения каждого этапа единого исследовательского цикла. Структура организационных форм НИС построена на основе двухфазного принципа, предполагающего постоянное чередование:

* фазы вложений, где главными выступающими являются преподаватели и приглашенные ими практики,
* фазы отдачи, где главными выступающими являются студенты магистерской программы.

На фазе вложений часть занятий семинара может быть заменена практической работой в компании (по согласованию с руководителем семинара)

**Схема организации семинара**

Общая схема организации НИС представлена в таблице. Реализация отдельных форм не обязательно полностью совпадает с границами соответствующих модулей, однако приоритет в каждом модуле отдается именно указанным формам.

Таблица. Циклы подготовки магистерской диссертации и форм научно-исследовательского семинара

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| I год обучения | Модули | | | |
| I | II | III | IV |
| Подготовка курсовой работы | Выбор темы и построение общего плана курсовой работы | Подготовка научно-аналитического обзора по тематике курсовой работы.  Подготовка и обсуждение проекта курсовой работы | Выполнение курсовой работы Проведение исследований, связанных с оценкой полученных результатов | Апробация курсовой работы |
| Формы НИС | Профориентационные лекции преподавателей кафедры и экспертов по тематике больших данных, в том числе на предприятиях-партнерах | Семинары с участием преподавателей кафедры и экспертов по тематике больших данных Занятия по написанию научно-аналитических обзоров. | Занятия  по выполнению курсовой работы и документированию проекта. Занятия по оценке проектов | Семинары с участием преподавателей кафедры и экспертов по тематике больших данных |
| II год обучения | Модули | | | |
| I | II | III | IV |
| Подготовка магистерской диссертации | Выбор темы и построение общего плана магистерской диссертации | Обсуждение программы исследований и плана-проспекта магистерской диссертации | Выполнение магистерского исследования. Проведение исследований, связанных с оценкой полученных результатов | Выполнение магистерского исследования. Подготовка текста магистерской диссертации. Предзащита магистерской диссертации |
| Формы НИС | Профориентационные лекции преподавателей кафедры и экспертов по тематике больших данных | Семинар по написанию академических работ с участием преподавателей кафедры | Занятия по организации и проведению исследований | Семинары по написанию  и оформлению научно-исследовательских  работ Семинары с участием преподавателей кафедры и экспертов по тематике больших данных |

В **первый год обучения** семинар начинается с профориентационных лекций, проводимых преподавателями кафедры и экспертов по тематике больших данных, которые делятся опытом своей собственной исследовательской работы, знакомят студентов с процедурами организации исследовательских проектов и с частью полученных результатов, формулируют задачи, которые могут лечь в основу тематики курсовых работ. При этом часть занятий может быть проведена в компаниях, которые разрабатывают или внедряют решения на основе больших данных.

Эта форма призвана помочь студентам выбрать тему курсовой работы (с прицелом на будущую магистерскую диссертацию) и сформировать первоначальный план этой работы к концу *первого модуля*. *Второй модуль* посвящен подготовке студентами проектов курсовой работы, подготовке презентации проекта и обсуждению этих проектов на научно-исследовательском семинаре. На этом этапе студент должен окончательно определиться с тематикой и структурой курсовой работы. К концу 3-го модуля студент обязан выбрать тему и руководителя, а также представить проект курсовой работы, чтобы быть допущенным к дальнейшему участию в семинаре.

В ходе *третьего модуля* студенты приступают к написанию курсовой работы, включающей в обязательном порядке аналитическое исследование на основе обзора имеющейся специальной академической и аналитической литературы, электронных источников информации, в том числе имеющихся статистических и аналитических баз данных. При анализе имеющихся материалов особое внимание уделяется теоретическим подходам и методам исследования. Основным результатом должен стать аналитический обзор. Также проводятся занятия по документированию выполняемого проекта, исследованию и оценке полученных результатов. Результатом является разработанный бизнес-план реализации проекта.

Значительная часть семинара уделяется формированию навыков использования инструментария больших данных для обеспечения аналитической части работы, особенно при использовании информации Интернет. У студентов начинают формироваться компетенции исследователей данных.

В течение *четвертого модуля* студенты готовят презентации выполненных работ, проводится обсуждение, по результатам обсуждения студенты проводят коррекцию и завершают написание курсовой работы. В этот период семинар работает в режиме консультаций, основная цель – помочь автору доработать первоначальный текст и привести его к окончательному виду. Затем проводится апробация работ в формате их рецензирования экспертами в области больших данных.

Структура НИС **второго года обучения**  в значительной мере повторяет структуру первого года и ориентирована на написание магистерской диссертации.

В **первом модуле** проводится второй цикл профориентационных лекций преподавателей кафедры и экспертов по тематике больших данных, а также рассматриваются требования к магистерской диссертации и ее общая структура. Это позволяет студентам  определиться с темой и объемом исследований будущей магистерской диссертации. Выбор темы и руководителя должен быть сделан в течение первого месяца. В конце модуля проводится обсуждение представляемых студентами развернутых планов диссертационной работы. Если к концу модуля студент не представил развернутого плана диссертации, то он имеет 2 недели во втором модуле для того, чтобы его представить.

В ходе *второго модуля* происходит формулирование целей и задач исследования, подготовка программы исследований, определение требований к результатам исследований, формирование структуры магистерской диссертации, подготовка плана-проспекта магистерской диссертации общим объемом не более 10 стр. Результаты обсуждаются в форме презентации на научно-исследовательском семинаре.

Основной задачей работы студента в течение *третьего модуля* является работа над диссертацией, основную часть которой составляют формирование гипотезы и параметров исследования, сбор данных и проведение аналитических исследований. Важной часть работы является выбор, а некоторых случаях создание программного инструментария для проведения исследования данных. В это время преподавателями кафедры проводятся занятия и консультации по организации и проведению исследований, использованию существующих информационно-аналитических источников, использованию и созданию исследовательского программного инструментария на основе технологий больших данных.

В *четвертом модуле* студенты завершают работу над магистерской диссертацией, пишут первоначальный текст магистерской диссертации, готовят презентацию для предзащиты диссертации. Предзащита диссертации проводится на семинаре с участием преподавателей кафедры и экспертов-практиков.

Семинар предполагает активное взаимодействие преподавателей и докладчиков в ходе работы семинара.

**Руководство семинаром**

Общую координацию деятельности по проведению и научно-методическому обеспечению семинара осуществляет координатор магистерской программы. Руководство научным семинаром в разных модулях может осуществлять один из членов научного совета программы.

**7.Характеристика кадрового потенциала, собственного и привлекаемого со стороны, а также имеющихся и требуемых для реализации МП ресурсов (до 2 стр.).**

Школа бизнес-информатики имеет преподавательский состав, включающий квалифицированных преподавателей в области анализа и моделирования бизнес-процессов, проектирования архитектуры предприятия, информационных технологий бизнес-анализа, компьютерного моделирования.

Это позволяет большую часть программы реализовать преподавателями НИУ ВШЭ. Однако, учитывая то, что технологии больших данных являются новыми, имеют существенные отличия от применяемых сегодня в учебном процессе, на проведение мастер-классов, чтение отдельных курсов приглашаются как отечественные, так и зарубежные внешние преподаватели. Также идет подготовка собственных преподавателей для чтения курсов, предоставляемых нам как членам академических альянсов компаний. Для преподавания также привлекаются сотрудники базовых кафедр и компаний-партнеров Школы бизнес-информатики.

Факультет оснащен достаточным количеством компьютерной техники и современными программными инструментами ведущих производителей продуктов для анализа больших объемов данных, такими как SPSS, Cognos и т.п. Свои решения для больших данных и соответствующие курсы, доступные в рамках членства в академических альянсах по предварительной договоренности предоставляют компании EMC, SAP, IBM.

Научно-методическое управление программой осуществляет международный Академический **Совет программы**.

1. Алескеров Фуад Тагиевич, д.т.н., ординарный профессор НИУ ВШЭ, директор департамента математики факультета экономики, НИУ ВШЭ, Россия

2. Голосов Алексей Олегович,  к.т.н., заведующий базовой кафедрой компании «Форс», Россия

3. Громов Александр Иванович, к.х.н., профессор, заведующий кафедрой моделирования и оптимизации бизнес-процессов, НИУ ВШЭ, Россия

4. Дием Хо, PhD, руководитель по связям с университетами компании IBM в Европе, на Среднем Востоке и Африке (EMEA), США

5. Илия Полиан, профессор, заведующий кафедрой компьютерной инженерии Университета Пассау, Германия

6. Исаев Евгений Анатольевич, к.т.н., заведующий базовой кафедрой компании «Стэк», НИУ ВШЭ, Россия

7. Мальцева Светлана Валентиновна, д.т.н., профессор, заведующая кафедрой инноваций и бизнеса в сфере ИТ, академический руководитель программы, НИУ ВШЭ, Россия

8. Стефан Маршанд-Майллет,  ассоциированный профессор департамента компьютерных наук Университета Женевы, Швейцария

9. Хорошевский Владимир Федорович,  д.т.н., главный научный сотрудник Центра информационно-аналитических систем Института статистических исследований и экономики знаний, НИУ ВШЭ, Россия