Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет» (Новосибирский государственный университет, НГУ)

Экономический факультет

Согласовано
Декан ЭФ
_____ Богомолова Т.Ю.
подпись
«19» октября 2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Цифровая трансформация и аналитика больших данных

Направление подготовки: 38.04.02 Менеджмент
Направленность (профиль): Стратегическое управление

Форма обучения: очная

Разработчики:

к.ф.-м.н. Павловский Е.Н. _____

Директор ММЦ НГУ
Д.ф.-м.н., Марчук И.В. _____

Новосибирск
2020

Содержание

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы 3

2. Место дисциплины в структуре образовательной программы 4

3. Трудоемкость дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающегося с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающегося 4

4. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий 4

5. Перечень учебной литературы 5

6. Перечень учебно-методических материалов по самостоятельной работе обучающихся 6

7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины 6

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине 7

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине 7

10. Оценочные средства для проведения текущего контроля и промежуточной аттестации по дисциплине 7

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Цель дисциплины «Цифровая трансформация и аналитика больших данных»: формирование у студентов профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных. Данная цель соотносится с целью образовательной программы в части формировании стратегии компании с использованием преимуществ технологий больших данных.

Основные задачи дисциплины:

- постановка задачи анализа данных;
- планирование проекта анализа данных;
- применение методов интеллектуального анализа данных для принятия решений;
- представление результатов работы.

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

Результаты освоения образовательной программы (компетенции)	В результате изучения дисциплины обучающиеся должны:		
	знать	уметь	владеть
ПК-2. Способностью разрабатывать корпоративную стратегию, программы организационного развития и изменений и обеспечивать их реализацию	- важность феномена больших данных для развития общества и науки;	- увязывать возможности технологий анализа больших данных с развитием организации; - убеждать заинтересованных лиц в необходимости цифровых решений;	
ПК-4. Способностью использовать количественные и качественные методы для проведения прикладных исследований и управления бизнес-процессами, готовить аналитические материалы по результатам их применения	- причины возникновения тренда больших данных; - важность применения научных методов для извлечения пользы из больших массивов данных	- формулировать бизнес-задачи в терминах анализа данных;	
ПК-11. Способность определять стратегию работы со структурированными и неструктурированными данными	- проблемы и возможности, связанные с появлением больших данных.	- выбрать подходящую технологию хранения и обработки больших данных; - использовать современные высоконагруженные системы хранения и обра-	

Результаты освоения образовательной программы (компетенции)	В результате изучения дисциплины обучающиеся должны:		
	знать	уметь	владеть
		ботки больших данных	

2. Место дисциплины в структуре образовательной программы

Дисциплина «Цифровая трансформация и аналитика больших данных» является элективной, преподается в 4 семестре.

Дисциплины (практики), изучение которых необходимо для освоения дисциплины «Цифровая трансформация и аналитика больших данных»: «Проектный анализ», «Инвестиционный менеджмент», «Современные отраслевые рынки», «Организационный дизайн и развитие организации», «Современный стратегический анализ», «Корпоративные финансы», «Управленческая экономика».

Дисциплины (практики), для изучения которых необходимо освоение дисциплины «Цифровая трансформация и аналитика больших данных»: выполнение ВКР.

3. Трудоемкость дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающегося с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающегося

Трудоемкость дисциплины – 3 зачетных единицы, 108 часов.

Форма промежуточной аттестации: дифференцированный зачет.

Вид деятельности	Семестр
	4
Контактная работа, часов, в том числе:	**38**
Лекции	-
Практические занятия	32
Групповая работа с преподавателем	4
Контактная работа при аттестации	2
Консультации перед экзаменом/зачетом	-
Самостоятельная работа, часов, в том числе:	**70**
Самостоятельная работа во время занятий	64
Самостоятельная работа во время промежуточной аттестации	6
Всего, часов	**108**

4. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4 семестр

Содержание дисциплины «Цифровая трансформация и аналитика больших данных»:

	Содержание разделов
1	Введение в большие данные. Цифровая трансформация экономики. Философия цифровизации.
2	Работа с заинтересованными лицами: идентификация, выявление предпочтение, матрица власть-интерес.

3	Жизненный цикл анализа больших данных. Как управляются проекты больших данных
4	Документация по проектам анализа данных. Шаблоны технических заданий, отчётов о выполнении этапов работ.
5	Выявление требований и работа с заказчиками в проектах анализа данных.
6	Постановки задач анализа данных. Регрессионные задачи. Задачи классификации и кластеризации.
7	Язык Python, стек библиотек анализа данных.
8	Проблема переобучения. Регуляризация. Машина опорных векторов. Нейронные сети. Глубинные нейронные сети.
9	Научные проблемы в области больших данных

Практические занятия (32 ч)

Содержание практического занятия	Объем, час
Семинар по изучению основных понятий в сфере больших данных, цифровизации в рамках дискуссии со студентами по вопросам трансформации бизнес-процессов – о роли технологий искусственного интеллекта, о роли талантов.	2
Семинар по разбору аналитического цикла на примерах, с пониманием отличий анализа больших данных от небольших.	2
Семинар по изучению документации по проектам анализа данных на примере отчёта о понимании бизнеса для микрокредитной организации	2
Семинар по изучению технического задания на исследование по данным заказчика	2
Семинар по представлению результатов анализа заинтересованных лиц	2
Семинар по определению требований к проекту	4
Семинар по оцениванию состояния инфраструктуры данных	4
Семинар по разбору примеров задач машинного обучения на практике	2
Семинар по разбору примеров задач предобработки данных на языке Python	2
Семинар по разбору критериев выбора систем хранения больших данных	2
Семинар по изучению примеров работы с заказчиками и спонсорами, способов определения репрезентативных систем и нацеливания предложений	2
Семинар по изучению примеров метрик качества для решения задач анализа данных	2
Семинар по представлению и обсуждению результатов проекта	4
ИТОГО:	**32**

Самостоятельная работа студентов (70 ч)

Перечень занятий на СРС	Объем, час
Подготовка к практическим занятиям	22
Изучение теоретического материала, не освещаемого на лекциях	20
Подготовка итогового проекта	22
Подготовка к дифференцированному зачету	6

5. Перечень учебной литературы

5.1 Основная литература

1. Грибанов, Ю.И. Цифровая трансформация бизнеса : учебное пособие : [16+] / Ю.И. Грибанов, М.Н. Руденко ; Пермский государственный национальный исследовательский университет. – 2-е изд. – Москва : Дашков и К°, 2021. – 214 с. : ил., схем., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=600303 (дата обращения: 30.11.2020). – Библиогр. в кн. – ISBN 978-5-394-04192-1. – Текст : электронный.

2. Петров, А.В. Моделирование процессов и систем [Электронный ресурс] : учебное пособие / А.В. Петров. — Электрон. дан. — Санкт-Петербург : Лань, 2015. — 288 с. — Режим доступа: https://e.lanbook.com/book/68472. — Загл. с экрана.

5.2 Дополнительная литература

3. Келлехер, Д. Наука о данных: базовый курс : [16+] / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров ; пер. с англ. М. Белоголовского. – Москва : Альпина Паблишер, 2020. – 224 с. : схем., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=598235 (дата обращения: 30.11.2020). – ISBN 978-5-9614-3170-4. – Текст : электронный.
4. Лбов, Г. С. Анализ данных и знаний : учебное пособие / Г.С. Лбов ; Федер. агентство по образованию, Новосиб. гос. ун-т, Мех.-мат. фак .— Новосибирск : Новосибирский государственный университет, 2010 .— 107 с. (Абонемент учебной и научной литературы: НК-7840, 35 экз).

6. Перечень учебно-методических материалов по самостоятельной работе обучающихся

5. Павловский Е.Н. Бизнес-анализ [Электронный ресурс] : электронный учебно-методический комплекс / Е. Н. Павловский ; Новосиб. гос. ун-т. - Новосибирск, [2018]. - Режим доступа: https://bigdata.nsu.ru/BA/. - Загл. с экрана.
6. Система управления проектами Trello.com // Atlassian. www.trello.com. 2020.

7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

При освоении дисциплины используются следующие ресурсы:

– образовательные интернет-порталы (Федеральный образовательный портал «Экономика. Социология. Менеджмент» http://ecsocman.hse.ru/):

– другие ресурсы информационно-телекоммуникационной сети «Интернет»:

- Введение в аналитику больших массивов данных [Электронный ресурс] : электронный учебно-методический комплекс / Е. Н. Павловский, Ю.А. Аникин, Н.Г. Загоруйко, И.А. Борисова, А.О. Зырянов; НОУ ИНТУИТ. - Москва, [2014]. - Режим доступа: http://www.intuit.ru/studies/courses/12385/1181/info. - Загл. с экрана.
- CRISP-DM 1.0 / Step-by-step data mining guide. Pete Charman and others. URL: https://www.the-modeling-agency.com/crisp-dm.pdf
- DJ Patil. Building Data Science Teams. O'Reilly. 2011. ISBN: 978-1-449-31623-5. URL: http://cdn.oreilly.com/radar/2011/09/Building-Data-Science-Teams.pdf

Взаимодействие обучающегося с преподавателем (синхронное и (или) асинхронное) осуществляется через электронную почту и социальные сети (Telegram, VK).

7.1 Современные профессиональные базы данных:

Не используются.

7.2. Информационные справочные системы

Не используются.

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Перечень программного обеспечения:

1) Windows (7, 8 или 10) или Ubuntu (14.04, 16.04 или 18.04).
2) Microsoft Office (версии выше 2010) или LibriOffice (версии от 5.0).

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Для реализации дисциплины «Цифровая трансформация и аналитика больших данных» используются специальные помещения:

1. Учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля, промежуточной и итоговой аттестации;

2. Помещения для самостоятельной работы обучающихся;

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду НГУ.

Материально-техническое обеспечение образовательного процесса по дисциплине для обучающихся из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете».

10. Оценочные средства для проведения текущего контроля и промежуточной аттестации по дисциплине

Перечень результатов обучения по дисциплине «Цифровая трансформация и аналитика больших данных» и индикаторов их достижения представлен в виде знаний, умений и владений в разделе 1.

10.1 Порядок проведения текущего контроля и промежуточной аттестации по дисциплине

Оценивание результатов обучения по дисциплине «Цифровая трансформация и аналитика больших данных» осуществляется по балльно-рейтинговой системе и включает следующие оценочные средства:

Текущий контроль успеваемости:

Текущий контроль осуществляется в виде приёма заданий из следующей таблицы:

Вид контроля/задания	Срок сдачи (нед.*)	Баллы
Задание 1 Прохождение опроса о целях и ожиданиях от курса: - почему выбрали именно этот курс? - чего хотели бы достичь по окончанию курса? - что готовы сделать для достижения этого? - чего хотели бы избежать? Регистрация в системе управление проектами Trello.	1	
Задание 2 Изучение вводного материала по онлайн лекциям	3	
Задание 3		

Выбрать проект из предложенных, сформировать команду.		
Задание 4 Составить список заинтересованных лиц. Идентифицировать всех, определить уровень власти/интереса.	4	
Задание 5 Представить результаты анализа заинтересованных лиц	4	20
Задание 6 Подготовка проекта: - свести документацию воедино; - выложить в проектное пространство.	10	20
Задание 7 Получить поддержку проекта у заинтересованных лиц	14	20
ИТОГО	**17**	**60**

* нед. – неделя от начала курса.

Баллы начисляются на каждого студента после командного выполнения задания, результат которого оценивается преподавателем. Сдача после указанного срока снижает количество баллов вдвое. В ходе выполнения заданий можно набрать до 60 баллов, остальные баллы распределяются на промежуточной аттестации. Не соблюдение сроков сдачи позволяет набрать не более 30 баллов в ходе текущего контроля. Таким образом, не позволяет получить оценку «отлично» за курс.

Итоговая оценка за семестр рассчитывается на основе полученной суммы баллов, и баллов, набранных при заключительном контроле знаний на зачете.

Промежуточная аттестация:

Промежуточная аттестация проводится в виде теста в устной или письменной форме в зависимости от возможности гарантировать самостоятельность ответа студентов.

В случае невозможности гарантировать самостоятельного выполнения студентами итоговых заданий (пандемия, карантин, вынужденное дистанционная форма контактной работы – онлайн), применяется устная форма тестирования. Устная форма тестирования состоит в выдаче индивидуальных вопросов в ходе индивидуальной беседы.

Необходимым условием допуска к зачёту является выполнение всех заданий текущего контроля.

Оценивание результатов обучения по дисциплине «Цифровая трансформация и аналитика больших данных» осуществляется по балльно-рейтинговой системе и включает следующие оценочные средства:

Оценочные средства	**Баллы (максимум)**
Текущий контроль	
Задание 5	20
Задание 6	20
Задание 7	20
Промежуточная аттестация	
Тест - дифференцированный зачёт	40
Итого	**100**

Описание критериев и шкал оценивания индикаторов достижения результатов обучения по дисциплине «Цифровая трансформация и аналитика больших данных»

Таблица 10.1.

Код компетенции	**Результат обучения по дисциплине**	**Оценочное средство**

ПК-2	Знание: - важности феномена больших данных для развития общества и науки;	Тест
	Умение: - увязывать возможности технологий анализа больших данных с развитием организации; - убеждать заинтересованных лиц в необходимости цифровых решений;	Задание 5 Задание 6 Задание 7
ПК-4	Знание: - причин возникновения тренда больших данных; - важности применения научных методов для извлечения пользы из больших массивов данных;	Тест
	Умение: - формулировать бизнес-задачи в терминах анализа данных;	Задание 6
ПК-11	Знание: - проблем и возможностей, связанных с появлением больших данных;	Тест
	Умение: - выбрать подходящую технологию хранения и обработки больших данных; - использовать современные высоконагруженные системы хранения и обработки больших данных	Задание 6

Таблица 10.2.

Критерии оценивания результатов обучения	**Шкала оценивания**
<u>Задание 5:</u> - успешно увязывает возможности технологий анализа больших данных с развитием организации в рамках раздела «заинтересованные лица» (выявлены все заинтересованные лица проекта); **<u>Задание 6:</u>** - успешно увязывает возможности технологий анализа больших данных с развитием организации: в рамках разделов «заинтересованные лица», «критерий успеха» (есть измеримый критерий успеха), «оценка ситуации» (есть описание реальных источников данных); - грамотно и верно формулирует бизнес-задачи в терминах анализа данных в рамках раздела «цели анализа данных»; - верно выбирает подходящую технологию хранения и обработки больших данных (требования к проекту полные, проверяемые, непротиворечивые и недвусмысленные); - использует современные высоконагруженные системы хранения и обработки данных; **<u>Задание 7:</u>** - демонстрирует умение убеждать заинтересованных лиц (заказчика/спонсора проекта) в необходимости цифровых решений;	*Отлично*

Тест: - демонстрирует всестороннее знание причин возникновения тренда больших данных; - демонстрирует отличные знания важности феномена больших данных для развития общества и науки; - обладает отличными знаниями важности применения научных методов для извлечения пользы из больших массивов данных; - демонстрирует исчерпывающие знания проблем и возможностей, связанных с появлением больших данных	
Задание 5: - увязывает основные возможности технологий анализа больших данных с развитием организации в рамках раздела «заинтересованные лица» (выявлены основные заинтересованные лица проекта); **Задание 6:** - увязывает основные возможности технологий анализа больших данных с развитием организации: в рамках разделов «заинтересованные лица», «критерий успеха» (есть измеримый критерий успеха), «оценка ситуации» (есть описание большинства реальных источников данных); - грамотно и верно формулирует бизнес-задачи в терминах анализа данных в рамках раздела «цели анализа данных»; - верно выбирает подходящую технологию хранения и обработки больших данных (требования к проекту полные, проверяемые, непротиворечивые и недвусмысленные); - использует высоконагруженные системы хранения и обработки данных; **Задание 7:** - демонстрирует умение убеждать основных заинтересованных лиц (заказчика/спонсора проекта) в необходимости цифровых решений; **Тест:** - демонстрирует знание основных причин возникновения тренда больших данных; - демонстрирует знания важности феномена больших данных для развития общества и науки; - обладает знаниями важности применения научных методов для извлечения пользы из больших массивов данных; - демонстрирует знания основных проблем и возможностей, связанных с появлением больших данных	*Хорошо*
Задание 5: - предпринимает попытку увязать возможности технологий анализа больших данных с развитием организации в рамках раздела «заинтересованные лица» (выявлены некоторые заинтересованные лица проекта); **Задание 6:** - предпринимает попытку увязать основные возможности технологий анализа больших данных с развитием организации: в рамках разделов «заинтересованные лица», «критерий успеха», «оценка ситуации» (есть описание отдельных источников данных); - предпринимает попытку сформулировать бизнес-задачи в терминах анализа данных в рамках раздела «цели анализа данных»; - предпринимает попытку выбирать подходящую технологию хранения и обработки больших данных (требования к проекту неполные, не	*Удовлетворительно*

все проверяемые, частично противоречат друг другу, но некоторые недвусмысленные); - предпринимает попытку использовать высоконагруженные системы хранения и обработки данных; **Задание 7:** - демонстрирует умение убеждать отдельных заинтересованных лиц в необходимости цифровых решений; **Тест:** - демонстрирует знание отдельных причин возникновения тренда больших данных; - демонстрирует отдельные знания важности феномена больших данных для развития общества и науки; - обладает знаниями некоторых аспектов важности применения научных методов для извлечения пользы из больших массивов данных; - демонстрирует знания отдельных проблем и возможностей, связанных с появлением больших данных	
Задание 5: - не предпринимает попытку увязать возможности технологий анализа больших данных с развитием организации в рамках раздела «заинтересованные лица» (не выявлены заинтересованные лица проекта); **Задание 6:** - не предпринимает попытку увязать возможности технологий анализа больших данных с развитием организации: в рамках разделов «заинтересованные лица», «критерий успеха», «оценка ситуации» (нет описания реальных источников данных); - не предпринимает попытку сформулировать бизнес-задачи в терминах анализа данных в рамках раздела «цели анализа данных»; - не предпринимает попытку выбирать подходящую технологию хранения и обработки больших данных (требования к проекту неполные, непроверяемые, противоречат друг другу, двусмысленные); - не предпринимает попытку использовать современные высоконагруженные системы хранения и обработки данных; **Задание 7:** - не демонстрирует умение убеждать заинтересованных лиц в необходимости цифровых решений; **Тест:** - демонстрирует отсутствие знания причин возникновения тренда больших данных; - демонстрирует отсутствие знания важности феномена больших данных для развития общества и науки; - не обладает знаниями важности применения научных методов для извлечения пользы из больших массивов данных; - не демонстрирует знания проблем и возможностей, связанных с появлением больших данных	*Неудовлетворительно*

Баллы, набранные за выполнение заданий текущего контроля и промежуточной аттестации, конвертируются в оценку по дисциплине следующим образом:

Итоговая сумма набранных баллов	**Оценка**
≤ 40	неудовлетворительно
от 40,1 до 60	удовлетворительно
от 60,1 до 80	хорошо

от 80,1 до 100	отлично

Типовые контрольные задания и иные материалы, необходимые для оценки результатов обучения

1. Типовая таблица для заинтересованных лиц (задание 5)

Заинтересованная сторона (лицо или группа)	Роль	Желаемый уровень участия (интерес)	Уровень влияния (власти)	Наиболее заинтересован в
	Бизнес-аналитик, Потребитель (Заказчик), Эксперт предметной области, Пользователь, Эксперт по реализации, Руководитель проекта, Тестировщик, Регулятор, Спонсор, Поставщик			

2. Типовая структура итогового документа (задание 6):

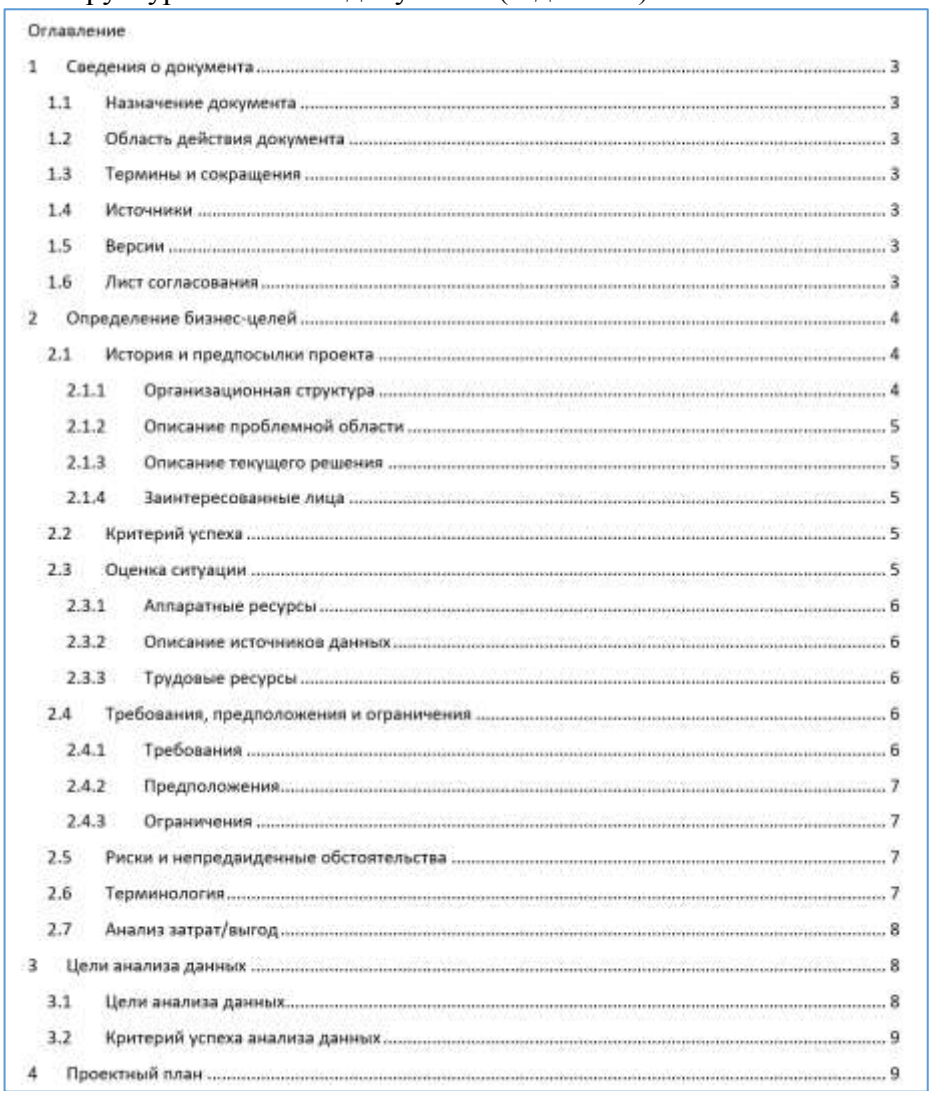

Рисунок 1.Типовая структура итогового документа

3. Типовые вопросы теста:

1. О соотношении аналоговой и цифровой информации:
    1. Большинство данных в мире в 2011 году содержалось:
        i. В цифровом виде
        ii. В аналоговом виде
2. История больших данных
    1. Укажите фактор, способствовавший появлению тренда больших данных
        i. Маркетинговые кампании крупных корпораций
        ii. Снижение издержек на хранение данных
        iii. Появление новых технологий обработки потоковых данных
        iv. Выпуск баз данных с обработкой данных в памяти
3. Определение больших данных:
    1. Выберите верный ответ
        i. Большие данные – это обработка или хранение более 1 Тб информации.
        ii. Проблема больших данных – это такая проблема, когда при существующих технологиях хранения и обработки сущностная обработка данных затруднена или невозможна.
        iii. Большие данные – это огромная PR-акция крупных вендоров и не более того.
        iv. Большие данные – это явление, когда цифровые данные наиболее полно представляют изучаемый объект.
4. Технологии хранения
    1. Какие из следующих технологий СУБД не используют принцип MapReduce
        i. Hadoop
        ii. Cassandra
        iii. HDInsight
        iv. Redis

Рисунок 2. Типовые вопросы теста

Оценочные материалы по текущему контролю и промежуточной аттестации, предназначенные для проверки соответствия уровня подготовки по дисциплине «Цифровая трансформация и аналитика больших данных» планируемым результатам освоения образовательной программы (в соответствии с образовательными стандартами), хранятся на кафедре-разработчике РПД в печатном и электронном виде.

Лист актуализации рабочей программы дисциплины
«Цифровая трансформация и аналитика больших данных»

№	Характеристика внесенных изменений (с указанием пунктов документа)	Дата и № протокола Ученого совета ЭФ	Подпись ответственного