Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 12.09.2026 15:16:02
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742f9[illegible]

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

Факультет издательского дела и журналистики

УТВЕРЖДАЮ

Декан

_____/Е.Л. Хохлогорская/

«_26_» _февраля_2026 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ОСНОВЫ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ

Направление подготовки
42.03.02 Журналистика

Профиль
Мультимедийная журналистика и цифровое продюсирование

Квалификация
Бакалавр

Форма обучения
Очная

Москва, 2026 г.

Разработчик:

К.полит.н., доцент кафедры «Журналистика и массовые коммуникации имени М.Ф. Ненашева» /С.Р. Мухамадуллин/

Согласовано:
Заведующий кафедрой «Журналистика и массовые коммуникации имени М.Ф. Ненашева»,
к.с.н., доцент /Л.Г. Толчинский/

Содержание

1. Цели, задачи и планируемые результаты обучения по дисциплине 4
2. Место дисциплины в структуре образовательной программы 5
3. Структура и содержание дисциплины 5
3.1 Виды учебной работы и трудоемкость 5
3.2 Тематический план изучения дисциплины 5
3.3 Содержание дисциплины 6
3.4 Тематика семинарских/практических и лабораторных занятий 7
3.4.2. Лабораторные занятия 9
3.5 Тематика курсовых проектов (курсовых работ) 9
4. Учебно-методическое и информационное обеспечение 9
4.1 Нормативные документы и ГОСТы 9
4.2 Основная литература 9
4.3 Дополнительная литература 9
4.4 Электронные образовательные ресурсы 10
4.5 Лицензионное и свободно распространяемое программное обеспечение 10
4.6 Современные профессиональные базы данных и информационные справочные системы 10
5. Материально-техническое обеспечение 10
6. Методические рекомендации 10
6.1 Методические рекомендации для преподавателя по организации обучения 10
6.2 Методические указания для обучающихся по освоению дисциплины 11
7. Фонд оценочных средств 11
7.1 Методы контроля и оценивания результатов обучения 11
7.2 Шкала и критерии оценивания результатов обучения 12
7.3 Оценочные средства 12

1. Цели, задачи и планируемые результаты обучения по дисциплине

Цель изучения дисциплины «Основы работы с большими данными» – сформировать у студентов систему знаний и практических навыков в области сбора, обработки, анализа и визуализации больших массивов данных для решения профессиональных задач журналиста: создания достоверных, доказательных и наглядных материалов, проведения медиаисследований, работы с аудиторными метриками и реализации проектов в области дата-журналистики.

К числу основных **задач** освоения дисциплины относятся:

- сформировать представление о природе, свойствах и классификации больших данных, их источниках и областях применения в медиасфере;
- обучить основам работы с открытыми данными (государственные порталы, социальные сети, дата-сеты) и методам их верификации;
- сформировать навыки первичной обработки, очистки и агрегации данных с использованием электронных таблиц и специализированных сервисов;
- научить применять базовые статистические методы для выявления трендов, корреляций и аномалий в массивах данных;
- развить умение создавать наглядные визуализации (диаграммы, карты, инфографику, дашборды) и интегрировать их в журналистские форматы (лонгриды, спецпроекты);
- освоить принципы дата-сторителлинга и перевода сложных числовых данных в доступный нарратив для массовой аудитории;
- привить этические и правовые нормы работы с персональными данными, соблюдение законодательства в информационной сфере.

Обучение по дисциплине «Основы работы с большими данными» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	Индикаторы достижения компетенции
ОПК-5. Способен использовать современные информационные технологии и программные средства при решении профессиональных задач	ИОПК-5.1. Использует цифровые инструменты для сбора, обработки и визуализации данных (электронные таблицы, BI-системы, сервисы дата-журналистики). ИОПК-5.2. Применяет программные средства для анализа больших массивов данных в профессиональных целях. ИОПК-5.3. Оценивает возможности и ограничения различных технологий работы с данными применительно к журналистским задачам.
ПК-2. Способен производить сбор, обработку и анализ информации, проверку фактов, принятие решений в процессе создания медиатекстов и (или) медиапродуктов, с учетом требований к достоверности, объективности, оперативности, полноте, этичности	**ИПК-2.1.** Осуществляет поиск и сбор данных из открытых источников, включая государственные порталы, социальные сети. **ИПК-2.2.** Проводит первичную обработку и очистку данных, выявляет аномалии и пропуски. **ИПК-2.3.** Интерпретирует статистические

	показатели и визуализирует данные для аудитории, формулирует выводы и гипотезы. **ИПК-2.4.** Соблюдает этические и правовые нормы при работе с данными, включая требования к защите персональной информации.

2. Место дисциплины в структуре образовательной программы

Дисциплина относится к элективной части блока Б1 «Дисциплины (модули)».

Дисциплина взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ОП: «Информационная безопасность СМИ», «Медиаметрия и аналитика для журналистов».

3. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 2 зачетные единицы (72 часа).

3.1 Виды учебной работы и трудоемкость

(по формам обучения)

3.1.1. Очная форма обучения

п/п	Вид учебной работы	Количество часов	Семестры	
			7	-
	Аудиторные занятия		54	-
	В том числе:			-
1	Лекции		18	-
2	Семинарские/практические занятия		36	-
3	Лабораторные занятия		-	-
	Самостоятельная работа		18	-
	В том числе:			-
1	Проведение исследования, подготовка презентации, докладов, сообщений		9	-
2	Подготовка к практическим занятиям (изучение лекционного материала)		9	-
	Промежуточная аттестация			-
	Зачет		-	-
	Итого		72	-

3.2 Тематический план изучения дисциплины

(по формам обучения)

3.2.1. Очная форма обучения

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/ практические занятия	Лабораторные занятия	Практическая подготовка	
1	Тема 1. Введение в большие данные: природа, свойства, источники и роль в журналистике	16	4	8	-	-	4
2	Тема 2. Сбор, очистка и подготовка данных: открытые данные, табличные процессоры, методы очистки	16	4	8	-	-	4
3	Тема 3. Основы статистического анализа для журналиста: средние, корреляции, тренды, аномалии	14	4	6	-	-	4
4	Тема 4. Визуализация данных: от статической инфографики до интерактивных дашбордов	12	2	8	-	-	2
5	Тема 5. Дата-сторителлинг и этика работы с данными: нарратив, фактчекинг, правовые нормы	14	4	6	-	-	4
Итого		**72**	**18**	**36**	**-**	**-**	**18**

3.3 Содержание дисциплины

Тема 1. Введение в большие данные: природа, свойства, источники и роль в журналистике

Определение больших данных, ключевые характеристики (объём, скорость, многообразие, достоверность, ценность). Эволюция данных как медиаресурса: от статистических справочников к большим массивам. Роль данных в современной журналистике: фактологическая основа, расследования, прогнозирование, персонализация контента. Классификация источников данных: открытые государственные реестры, ведомственная статистика, данные социальных сетей, веб-аналитика, результаты опросов. Обзор российских дата-журналистских проектов и практик. Тенденции развития дата-журналистики в России. Практическое задание: анализ выбранного дата-проекта российского медиа, выявление использованных источников и методов.

Тема 2. Сбор, очистка и подготовка данных: открытые данные, табличные процессоры, методы очистки

Понятие открытых данных, принципы их публикации. Государственный портал открытых данных Российской Федерации, порталы открытых данных федеральных органов исполнительной власти и субъектов Российской Федерации. Реестры наборов открытых данных по категориям: «Безопасность», «Картография», «Спорт» и др. Работа с российскими социальными сетями и платформами. Основы веб-парсинга с использованием готовых конструкторов. Импорт данных в табличные процессоры из разных форматов. Первичная обработка: сортировка, фильтрация, условное форматирование, базовые функции. Сводные таблицы для агрегации. Типичные проблемы «грязных» данных: пропуски, дубликаты, выбросы, ошибки формата, несоответствие типов. Методы очистки: удаление дубликатов, интерполяция пропусков, стандартизация,

нормализация. Инструменты очистки. Практические кейсы: подготовка бюджетных данных, результатов голосований, данных переписи для дальнейшего анализа.

Тема 3. Основы статистического анализа для журналиста: средние, корреляции, тренды, аномалии

Базовые статистические показатели: среднее арифметическое, медиана, мода, дисперсия, стандартное отклонение. Интерпретация этих величин в журналистском контексте: когда использовать среднее, а когда медиану. Понятие корреляции и причинно-следственных связей; типичные ошибки при трактовке. Выявление трендов: скользящее среднее, сезонные колебания. Обнаружение аномалий (выбросов) и их возможное значение для расследования. Использование простых статистических тестов для проверки гипотез (на интуитивном уровне). Применение статистики для анализа общественного мнения, социально-экономических данных, эффективности государственных программ. Предостережения: манипуляции статистикой, ошибки выборки, недостоверность данных. Практические упражнения: расчёт показателей на реальных данных, визуализация распределений.

Тема 4. Визуализация данных: от статической инфографики до интерактивных дашбордов

Принципы эффективной визуализации: выбор типа диаграммы в зависимости от цели (сравнение, структура, распределение, связь, география). Создание столбчатых, линейных, круговых диаграмм, точечных графиков, тепловых карт, географических карт. Особенности восприятия графической информации: цвет, композиция, иерархия, подписи. Инструменты для статической визуализации. Интерактивные визуализации и дашборды: обзор российских BI-систем. Создание дашбордов с фильтрами и параметрами. Интеграция визуализаций в веб-страницы, лонгриды и спецпроекты. Анализ удачных примеров интерактивных историй от российских изданий. Практика: создание интерактивной карты или временной шкалы на основе реальных данных.

Тема 5. Дата-сторителлинг и этика работы с данными: нарратив, фактчекинг, правовые нормы

Концепция дата-сторителлинга: искусство рассказывать истории с помощью данных. Этапы построения истории: поиск идеи, формулировка вопроса, сбор и анализ данных, выбор формата, визуализация, написание текста. Сочетание числовых выкладок с эмоциональным повествованием, примерами из жизни, интервью. Адаптация контента для разных платформ (социальные сети, веб, мобильные устройства). Правовые аспекты: законодательство о персональных данных, ответственность за разглашение конфиденциальной информации. Этические принципы: достоверность, объективность, недопустимость манипуляций через подбор статистики, уважение к субъектам данных. Методы фактчекинга числовых данных: проверка источников, перекрёстное сравнение, выявление скрытых допущений. Как избежать распространённых ошибок и представить данные честно. Разбор громких случаев журналистских ошибок и скандалов, связанных с неверной интерпретацией данных. Итоговая практическая работа: разработка концепции и прототипа дата-истории на выбранную тему с использованием открытых данных.

3.4 Тематика семинарских/практических и лабораторных занятий

3.4.1. Семинарские/практические занятия для очной формы обучения

Семинар 1. Большие данные в современной журналистике: обзор российских дата-проектов

Анализ успешных примеров использования данных в российских СМИ (ТАСС, РИА Новости, «Коммерсантъ», «Ведомости»). Определение источников данных, методов их обработки и визуализации. Характеристика «5 V» на конкретных кейсах.

Семинар 2. Поиск и сбор данных на государственных порталах открытых данных Практическая работа с порталом data.gov.ru и региональными порталами

(data.mos.ru). Освоение навигации, фильтрации наборов данных, выгрузка в форматах CSV, JSON, XML.

Семинар 3. Работа с открытыми данными федеральных органов исполнительной власти. Обзор реестров наборов открытых данных по категориям: «Безопасность», «Картография», «Спорт», «Образование», «Здравоохранение». Выбор и загрузка данных для учебного проекта.

Семинар 4. Основы работы с API российских социальных сетей и платформ Знакомство с API ВКонтакте и Одноклассников. Получение данных о публикациях, комментариях, просмотрах. Использование готовых конструкторов запросов.

Семинар 5. Импорт и первичная обработка данных в электронных таблицах Импорт данных из различных форматов (CSV, JSON, XML). Сортировка, фильтрация, условное форматирование. Базовые функции (СУММ, СРЗНАЧ, СЧЁТ, ЕСЛИ, ВПР).

Семинар 6. Сводные таблицы и агрегация данных Создание сводных таблиц для группировки данных по категориям, временным периодам, географическим признакам. Анализ бюджетных данных, результатов голосований, данных переписи.

Семинар 7. Очистка данных: выявление и устранение пропусков, дубликатов, ошибок формата

Практические упражнения по обнаружению «грязных» данных. Методы интерполяции пропусков, удаления дубликатов, стандартизации и нормализации.

Семинар 8. Использование OpenRefine для очистки и преобразования данных Знакомство с интерфейсом и основными функциями OpenRefine. Кластеризация, разделение столбцов, преобразование форматов. Применение к реальным наборам данных.

Семинар 9. Базовые статистические показатели: среднее, медиана, мода Расчёт и анализ показателей на реальных данных (социально-экономические показатели регионов). Интерпретация результатов для журналистского текста. Когда использовать среднее, а когда медиану.

Семинар 10. Дисперсия, стандартное отклонение и оценка разброса данных Вычисление показателей, интерпретация результатов. Анализ данных опросов общественного мнения, выявление консенсусных и спорных позиций.

Семинар 11. Корреляционный анализ: поиск связей и типичные ошибки интерпретации

Выявление корреляций между различными показателями (на примере экономических и социальных данных). Разбор случаев ошибочного отождествления корреляции с причинно-следственной связью.

Семинар 12. Выявление трендов и аномалий в данных Построение временных рядов, расчёт скользящего среднего. Обнаружение аномальных значений и их значение для журналистского расследования.

Семинар 13. Принципы визуализации и создание статической инфографики Выбор типа диаграммы в зависимости от цели. Создание столбчатых, линейных, круговых диаграмм, тепловых карт. Оформление цветовой гаммы, подписей, легенд.

Семинар 14. Интерактивные визуализации. Подключение источников данных, создание дашбордов с фильтрами и параметрами. Настройка интерактивных элементов. Интеграция визуализаций в веб-страницы и лонгриды.

Семинар 15. Создание интерактивных карт и временных шкал Обзор российских сервисов для создания карт. Построение интерактивных таймлайнов с помощью доступных инструментов. Применение к журналистским задачам.

Семинар 16. Дата-сторителлинг: структура и композиция дата-истории Этапы создания дата-истории: поиск темы, формулировка вопроса, сбор и анализ данных, выбор формата. Сочетание числовых данных с нарративом, интервью, визуальными элементами.

Семинар 17. Адаптация дата-материалов для разных платформ. Правовые и этические аспекты. Особенности подачи данных в социальных сетях, на сайтах, в мобильных приложениях. Обзор законодательства о персональных данных. Этические принципы: достоверность, объективность, недопустимость манипуляций.

Семинар 18. Фактчекинг числовых данных. Защита дата-проектов Методы верификации данных: проверка источников, перекрёстное сравнение, выявление скрытых допущений. Разбор громких случаев журналистских ошибок. Презентация и обсуждение разработанных студентами дата-историй (итоговый проект).

3.4.2. Лабораторные занятия

Не предусмотрены

3.5 Тематика курсовых проектов (курсовых работ)

Не предусмотрены

4. Учебно-методическое и информационное обеспечение

4.1 Нормативные документы и ГОСТы

Закон РФ от 27.12.1991 N 2124-1 (ред. от 31.07.2025) «О средствах массовой информации». 27 декабря 1991 года N 2124-1 // https://www.consultant.ru/document/cons_doc_LAW_1511/ [Режим доступа - свободный]

Федеральный закон от 27.07.2006 N 149-ФЗ «Об информации, информационных технологиях и о защите информации» // https://www.consultant.ru/document/cons_doc_LAW_61798/ [Режим доступа - свободный]

Федеральный закон от 27.07.2006 N 152-ФЗ «О персональных данных» // https://www.consultant.ru/document/cons_doc_LAW_61801/ [Режим доступа - свободный]

Гражданский кодекс Российской Федерации (ГК РФ) 30 ноября 1994 года N 51-ФЗ // https://www.consultant.ru/document/cons_doc_LAW_5142/ [Режим доступа - свободный]

4.2 Основная литература

Макшанов, А. В. Большие данные. Big Data / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — 4-е изд., стер. — Санкт-Петербург : Лань, 2024. — 188 с. — ISBN 978-5-507-47346-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/362318 (дата обращения: 29.06.2026). — Режим доступа: для авториз. пользователей.

Баранова, Е. А. Конвергентная журналистика как новая модель для медиабизнеса / Е. А. Баранова. - Текст : электронный // Znanium.com. - 2017. - №1-12. - URL: https://znanium.com/catalog/product/506113 (дата обращения: 03.07.2026). – Режим доступа: по подписке.

4.3 Дополнительная литература

Ланских, Ю. В. Введение в большие данные : учебное пособие / Ю. В. Ланских, В. Г. Ланских, К. В. Родионов. — Киров : ВятГУ, 2023. — 172 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/408566 (дата обращения: 06.07.2026). — Режим доступа: для авториз. пользователей.

Лабковская, Р. Я. Анализ больших данных : учебное пособие / Р. Я. Лабковская, П. В. Косов. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2025. — 152 с. — ISBN 978-5-89160-366-0. — Текст : электронный // Лань : электронно-библиотечная система. —

URL: https://e.lanbook.com/book/508654 (дата обращения: 07.07.2026). — Режим доступа: для авториз. пользователей.

4.4 Электронные образовательные ресурсы

Сайт Союза журналистов России : нормативные документы, этические кодексы, профессиональные стандарты. – URL: https://ruj.ru/ (дата обращения: 25.06.2026).

4.5 Лицензионное и свободно распространяемое программное обеспечение

Браузеры – для доступа к образовательным ресурсам, онлайн-платформам и сервисам обработки больших данных (свободно распространяемое ПО)

4.6 Современные профессиональные базы данных и информационные справочные системы

1. Официальный интернет-портал правовой информации. Законодательная база, нормативно-правовое обеспечение, статьи. http://pravo.gov.ru/
2. Российская государственная библиотека. http://www.rsl.ru/
3. Российская национальная библиотека http://www.nlr.ru/
4. Научная электронная библиотека https://elibrary.ru/
5. Российский книжный союз https://bookunion.ru/
6. Гильдия издателей периодической печати (ГИПП) www.gipp.ru
7. Портал открытых данных Российской Федерации – https://data.gov.ru/
8. Портал открытых данных Правительства Москвы – https://data.mos.ru/

5. Материально-техническое обеспечение

Аудитории для лекционных и семинарских занятий из общего фонда. Необходима интерактивная доска/ноутбук, проектор для демонстрации презентаций.

Университет обеспечивает инвалидов и лиц с ограниченными возможностями здоровья специальными материально-техническими средствами обучения (включая специальное программное обеспечение) и электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья, при наличии их заявлений о необходимости предоставления специализированных электронных образовательных ресурсов.

6. Методические рекомендации

6.1 Методические рекомендации для преподавателя по организации обучения

Изучение учебной дисциплины осуществляется в форме лекционных и семинарских занятий. Задача лекционных занятий – изложение теоретического материала в рамках программы. Задача семинарских занятий – закрепление полученных теоретических знаний и формирование практических навыков работы с большими данными, их обработки и визуализации для журналистских задач.

Активность на семинарских занятиях имеет приоритетное значение перед посещением лекций. Дискуссии на семинарах по заранее объявленным темам проходят в свободной форме и имеют целью выявление знания, понимания проблемы и способности критической оценки различных подходов к работе с данными, их статистическому анализу и интерпретации.

Формы текущего контроля: выступление на семинарах с докладами и сообщениями; участие в работе на семинарских занятиях, устные и письменные опросы; выполнение

практических заданий по работе с открытыми данными; подготовка и защита итогового дата-проекта.

Форма промежуточного контроля: зачёт. Принимается в форме собеседования, обязательны дополнительные вопросы по каждой пропущенной или плохо отработанной на семинаре/практическом занятии теме. Студент должен продемонстрировать знание основных понятий, методов работы с данными, умение применять статистические инструменты и создавать визуализации, а также понимание этических и правовых аспектов работы с данными.

Результирующая оценка складывается из следующих составляющих: посещение лекций; активное участие в семинарских занятиях; результаты письменных и устных опросов, выполнения практических заданий, качество итогового дата-проекта, ответ во время зачёта.

6.2 Методические указания для обучающихся по освоению дисциплины

Студенту, желающему углубить свои знания, предоставляется возможность расширить круг специализированного чтения за счёт обращения к дополнительным источникам, указанным в программе курса, а также самостоятельно изучать открытые наборы данных на государственных порталах. В процессе самоподготовки студенту следует: изучать и анализировать источники и научную литературу, рекомендованную преподавателем по программе курса; пользоваться ресурсами Интернета для самостоятельного поиска актуальных открытых данных и примеров их использования в российской журналистике; регулярно выполнять практические задания по работе с электронными таблицами, инструментами очистки данных и визуализации; фиксировать все этапы обработки данных и источники для обеспечения воспроизводимости результатов; при подготовке итогового дата-проекта особое внимание уделять достоверности исходных данных, корректности статистических расчётов и этичности интерпретации выводов; обращаться за разъяснениями к преподавателю по вопросам курса, которые вызывают затруднение. При подготовке к зачёту рекомендуется систематизировать материал по темам, повторить ключевые определения, методы анализа и визуализации, а также примеры из российской дата-журналистской практики, разобранные на семинарах.

7. Фонд оценочных средств

7.1 Методы контроля и оценивания результатов обучения

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций: подготовка, представление и обсуждение докладов и сообщений на семинарских занятиях; устные и письменные опросы; выполнение практических заданий по работе с открытыми данными; подготовка и защита итогового дата-проекта; круглый стол, дискуссия, коллоквиум.

Форма промежуточной аттестации – зачёт в форме собеседования.

Оценочные средства текущего контроля успеваемости включают тематику сообщений для подготовки к семинарским занятиям, тематику практических заданий; вопросы для проведения дискуссий, коллоквиума; контрольные вопросы для проведения зачёта.

Промежуточная аттестация обучающихся проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине, при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине проводится преподавателем, ведущим занятия по дисциплине методом

экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «зачтено» или «не зачтено».

7.2 Шкала и критерии оценивания результатов обучения

По результатам работы в семестре, включая итоговую защиту дата-проекта и собеседование по темам курса, студенту выставляется оценка по двухбалльной системе («зачтено», «не зачтено»).

Оценка «**зачтено**» выставляется в том случае, если обучающийся демонстрирует полное или частичное соответствие следующих знаний: определение и свойства больших данных, основные источники открытых данных, базовые статистические показатели, этапы подготовки данных к анализу и визуализации, правовые ограничения при работе с данными. Свободно или с незначительными ошибками оперирует приобретенными знаниями. Обучающийся демонстрирует полное или частичное соответствие следующих умений: формулировать запросы к источникам, проводить очистку и агрегацию данных в таблицах, строить визуализации, создавать простые дашборды, писать аналитические материалы на основе данных. Обучающийся демонстрирует полное или частичное владение технологиями работы с данными.

Оценка «**не зачтено**» выставляется в том случае, если обучающийся демонстрирует полное отсутствие или недостаточное соответствие следующих знаний: определения и свойства больших данных, основных источников открытых данных, базовых статистических показателей, этапов подготовки данных к анализу и визуализации, правовых ограничений при работе с данными. Обучающийся не умеет или в недостаточной степени умеет формулировать запросы к источникам, проводить очистку и агрегацию данных в таблицах, строить визуализации, создавать простые дашборды, писать аналитические материалы на основе данных. Обучающийся не владеет или в недостаточной степени владеет технологиями работы с данными.

7.3 Оценочные средства

7.3.1. Текущий контроль

Примерные вопросы для коллоквиумов, собеседований

1. Актуальная тема для дата-журналистского материала. Как её найти и обосновать?
2. Как сформулировать стержневой вопрос дата-истории?
3. Какие источники открытых данных можно использовать для выбранной темы?
4. Какие методы очистки и подготовки данных потребуются?
5. Какие статистические показатели и методы анализа необходимо применить?
6. Какой тип визуализации наиболее подходит для данной истории?
7. Как сочетать числовые данные с нарративом и эмоциональным повествованием?
8. Какие этические и правовые ограничения необходимо учитывать при работе с данными?

Примерные вопросы для контрольной работы

Контрольная работа позволяет выявить знания обучающихся по основным компетенциям работы с большими данными.

Вариант 1.

1. Охарактеризуйте ключевые характеристики больших данных («5 V») и приведите примеры их проявления в журналистской практике.
2. Опишите этапы сбора, очистки и подготовки данных для визуализации на примере любого открытого набора данных с портала data.gov.ru.

Вариант 2.

1. Объясните, что такое открытые данные и как они могут быть использованы журналистом. Приведите примеры российских порталов открытых данных.

2. Выберите набор открытых данных, проведите его первичный анализ с помощью Excel, постройте диаграмму и сформулируйте выводы.

Вариант 3.

1. Что такое корреляция? Чем она отличается от причинно-следственной связи? Приведите примеры типичных ошибок интерпретации корреляций в журналистике.
2. На основе предоставленного набора данных рассчитайте среднее, медиану и стандартное отклонение. Интерпретируйте результаты.

Вариант 4.

1. Охарактеризуйте принципы эффективной визуализации данных. Какие типы диаграмм используются для сравнения, структуры, распределения и связи?
2. Создайте интерактивную визуализацию на основе открытых данных. Опишите этапы работы.

Вариант 5.

1. Что такое дата-сторителлинг? Опишите этапы создания дата-истории от идеи до готового материала.
2. Разработайте концепцию дата-проекта на выбранную тему: сформулируйте вопрос, укажите источники данных, предложите тип визуализации и структуру нарратива.

Вариант 6.

1. Каковы основные правовые и этические ограничения при работе с данными в журналистике? Опишите требования Федерального закона № 152-ФЗ «О персональных данных».
2. Проведите фактчекинг предоставленного набора данных: проверьте источники, выявите возможные ошибки и искажения, предложите способы их устранения.

Примерная тематика докладов, рефератов, сообщений

1. Большие данные в журналистике: определение, свойства и роль в современном медиапространстве.
2. Характеристики больших данных («5 V») и их проявление в журналистских задачах.
3. Открытые данные: понятие, принципы публикации и использование в российских СМИ.
4. Государственный портал открытых данных Российской Федерации: структура, навигация, возможности.
5. Региональные порталы открытых данных: сравнительный анализ (на примере Москвы и других регионов).
6. Работа с API российских социальных сетей (ВКонтакте, Одноклассники) для сбора данных.
7. Электронные таблицы как инструмент журналиста: основные функции и возможности.
8. Очистка данных: методы выявления и устранения ошибок, пропусков, дубликатов.
9. Статистические показатели в журналистике: среднее, медиана, мода – когда и как применять.
10. Корреляционный анализ: поиск связей и типичные ошибки интерпретации.
11. Тренды и аномалии в данных: методы выявления и значение для журналистского расследования.
12. Принципы эффективной визуализации данных для массовой аудитории.
13. Создание статической инфографики: инструменты и правила оформления.
14. Интерактивные карты и временные шкалы в журналистских проектах.
15. Дата-сторителлинг: искусство рассказывать истории с помощью данных.
16. Структура и композиция дата-сторителлинга.
17. Адаптация дата-материалов для разных платформ: социальные сети, веб, мобильные устройства.

18. Правовые аспекты работы с данными: законодательство о персональных данных.
19. Этические принципы работы с данными: достоверность, объективность, недопустимость манипуляций.
20. Фактчекинг числовых данных: методы проверки источников и перекрёстного сравнения.
21. Типичные ошибки журналистов при работе с данными и способы их предотвращения.
22. Будущее дата-журналистики в России: тренды и перспективы.

7.3.2. Промежуточная аттестация

Примерные вопросы к зачету

1. Определение больших данных, их ключевые характеристики.
2. Роль данных в современной журналистике: фактологическая основа, расследования, персонализация контента.
3. Классификация источников данных: открытые государственные реестры, ведомственная статистика, данные социальных сетей, веб-аналитика.
4. Обзор российских дата-журналистских проектов и практик.
5. Понятие открытых данных, принципы их публикации.
6. Государственный портал открытых данных Российской Федерации: структура, навигация, возможности.
7. Региональные порталы открытых данных (на примере Москвы).
8. Первичная обработка данных: сортировка, фильтрация, условное форматирование.
9. Базовые функции электронных таблиц: СУММ, СРЗНАЧ, СЧЁТ, ЕСЛИ, ВПР.
10. Создание сводных таблиц для агрегации и группировки данных.
11. Типичные проблемы «грязных» данных: пропуски, дубликаты, выбросы, ошибки формата.
12. Методы очистки данных: удаление дубликатов, интерполяция пропусков, стандартизация, нормализация.
13. Инструменты очистки данных: основные функции и возможности.
14. Базовые статистические показатели: среднее арифметическое, медиана, мода, дисперсия, стандартное отклонение.
15. Интерпретация статистических показателей в журналистском контексте: когда использовать среднее, а когда медиану.
16. Понятие корреляции и причинно-следственных связей; типичные ошибки при трактовке.
17. Выявление трендов: скользящее среднее, сезонные колебания.
18. Обнаружение аномалий (выбросов) и их значение для журналистского расследования.
19. Принципы эффективной визуализации данных: выбор типа диаграммы в зависимости от цели.
20. Создание столбчатых, линейных, круговых диаграмм, точечных графиков, тепловых карт.
21. Особенности восприятия графической информации: цвет, композиция, иерархия, подписи.
22. Создание интерактивных карт и временных шкал с использованием российских сервисов.
23. Интеграция визуализаций в веб-страницы, лонгриды и спецпроекты.
24. Концепция дата-сторителлинга: искусство рассказывать истории с помощью данных.
25. Этапы построения дата-истории: поиск идеи, формулировка вопроса, сбор и анализ данных, выбор формата, визуализация, написание текста.
26. Сочетание числовых данных с нарративом, интервью, визуальными элементами.

27. Адаптация контента для разных платформ (социальные сети, веб, мобильные устройства).
28. Правовые аспекты работы с данными: законодательство о персональных данных.
29. Этические принципы работы с данными: достоверность, объективность, недопустимость манипуляций.
30. Методы фактчекинга числовых данных: проверка источников, перекрёстное сравнение, выявление скрытых допущений.
31. Типичные ошибки журналистов при работе с данными и способы их предотвращения.
32. Перспективы развития дата-журналистики в России.