

ЖУРНАЛИСТИКА ДАННЫХ

Дисциплина «Журналистика данных», созданная на стыке традиционных гуманитарных подходов к журналистике и новых информационных технологий, основанных на прикладной и вычислительной математике. призвана сформировать у студентов понимание того, что означают термины «данные» и «большие данные», как собирать и обрабатывать данные с помощью математических методов, создавать на их основе профессиональные журналистские тексты.

Статистико-эконометрический раздел дисциплины содержит основные понятия и концепции из теории вероятностей и математической статистики (понятие случайной величины, различные определения среднего значения, распределения вероятностей и их свойства, статистические гипотезы, тесты и оценки параметров), а также математический аппарат для работы с большими объемами численных данных (простейшие модели регрессии, работа с панельными данными и временными рядами).

Второй раздел посвящен обсуждению приложений статистических методов и науки о данных (data science) к журналистике. Также обсуждаются основные форматы данных, особенности поиска данных онлайн и оффлайн, способы «преобразования» найденных данных в журналистские тексты. Основная цель — сформировать у студентов понимание, какие типы данных и в каких ситуациях могут понадобиться в работе, какое ПО лучше использовать для их обработки.

В рамках дисциплины также изучаются основы визуализации данных и инфографики.

Общая трудоемкость дисциплины составляет 4 зачетные единицы (144 часа).

№ п/п	Раздел дисциплины	Формы текущего контроля и промежуточной аттестации
1	Понятие данных, основные их типы и формы. Особенности онлайн и оффлайн поиска данных. Полезные инструменты для скачивания. Основные базовые форматы хранения данных (XLS, DB, CSV, JSON, XML). Преобразование «человеческих» данных в машиночитаемые, дата-скреппинг	Контрольная работа
2	Основные машиночитаемые форматы хранения данных (XLS, DB, CSV, JSON, XML). Подготовка данных к работе. Правила оформления датасетов. Базовые инструменты: Excel, Google Sheets. Очистка, нормализация, исправление ошибок, хранение и управление данными	Контрольная работа

3	Математический аппарат для работы с численными данными. Основные понятия и концепции из теории вероятностей и математической статистики: введение в регрессионный, корреляционный, факторный анализ и кластеризацию данных	Контрольная работа
4	Визуализация данных. Визуальное представление информации и способ её анализа. Типы графического представления информации, выбор корректного типа графики для различных данных, отбор данных для визуализации. Введение в инструменты визуализации: Excel, Infogram, Tableau, Datawrapper, Flourish, RawGraphs	Контрольная работа
5	Практическое применение больших данных в медиаиндустрии: аудиторные исследования, расследовательская журналистика, персонализация рекламы и цифровые следы	
		Зачет