

Программа учебной дисциплины _____

Утверждена

Академическим советом ООП

Протокол № от «__» _____ 20__ г.

Автор	Критические исследования данных
Число кредитов	3
Контактная работа (час.)	12
Самостоятельная работа (час.)	12
Курс	1 Магистратура
Формат изучения дисциплины	с использованием элементов онлайн-курса

I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Актуальность

В течение последних пяти лет в НИУ ВШЭ, как и в других ведущих мировых университетах, появляется большое количество [образовательных](#) и [исследовательских](#) программ, посвящённых работе с данными.

В частности, появляется всё больше программ на ФКН, но и не только: многие студенты бакалавриата проходят обучение на программе [Data Culture](#).

В рамках программы работа с данными представляется в первую очередь как набор необходимых навыков, которые применяются в отношении конкретного направления в науках, например, в социологии или международных отношениях.

Однако, работа с алгоритмами и большими данными — это не только навык, но и новая веха в развитии исследовательских подходов разных наук. В ней есть немало сложных вопросов, этических конфликтов, эпистемологических проблем. Эти проблемы необходимо обсуждать: не существует готовых решений, которые могли бы уместиться в инструкцию. Но задача университета — это не только предложение конкретных решений, но и организация интеллектуально насыщенного диалога по проблемам, важным для науки и общества. Проблема данных — одна из них.

В течение тех же пяти лет в обществах разных стран возникают разные соображения по поводу того, как связаны развитие наук о данных и социальная жизнь. Нередко эти соображения становятся [опасениями](#) или превращаются в моральную панику (O’Neil, 2017). Так, можно вспомнить многочисленные скандалы, связанные с социальной сетью Facebook, потребностью европейских стран в принятии GDPR, участвовавшие разговоры о влиянии алгоритмов на избирательные кампании и ещё пример. Есть и немало сомнений со стороны разных представителей научного сообщества, в крупнейших университетах мира идут курсы, посвящённые этике работы с данными. Для примера — такие курсы в [Стэнфорде](#) или университете Висконсина.

Сами инициаторы работы с данными понимают, что этот инструмент не является полностью нейтральным и по крайней мере может быть связан с разнообразными и непредсказуемыми

последствиями при внедрении и участии в принятии решений. В индустрии также разрабатываются [этические элементы профессиональных кодексов](#), которые могут быть полезными для самих разработчиков. В Европе принят GDPR, ограничивающий возможности сбора пользовательских данных в интернете, в России также происходят обсуждения таких законодательных мер.

Основные вопросы и сферы применения

Ключевые вопросы, которые обсуждаются в связи с данными:

- Приватность пользователей и право на исключение себя из источников данных
- Коммерциализация данных и исключённость источников данных из этого процесса
- Алгоритмизация и возможное усиление неравенства и ограничение политических свобод
- Недостаточное понимание практик, в ходе которых пользователи городского пространства или онлайн-сред производят данные.

Эти проблемы возникают при использовании данных в разных отраслях: экономике, городском планировании, социальных исследованиях, разработке ПО и так далее.

Цель и содержание курса

Цель курса — разработать комплексное представление о рисках работы с данными, предусмотреть возможные механизмы решения возникающих проблем и наладить регулярную дискуссию, которая будет необходима для специалистов в разных областях.

В данном курсе будут проработаны основные риски работы с данными, этические аспекты и возможности решения проблем, возникающих при работе с данными.

II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Тема 1. Лекция. Как устроен процесс сбора и анализа данных с точки зрения разных участников этого процесса. Определение стейкхолдеров и участников процесса сбора данных. История статистических наблюдений и этических вызовов, связанных с работой с данными. Роли разработчика и аналитика при проектировании и выполнении сбора и обработки данных. Теории, разрабатывающие этические и политические проблемы при работе с данными.

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208.

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2), 2053951715594634.

Тема 2. Лекция. Критический анализ работы с данными в городских, индустриальных проектах, образовании и государственном управлении. Анализ и рассмотрение того, какие данные используются в современных бизнес-процессах, госуправлении, медицине, образовании. История внедрения данных в процессы, «датафикация»

boyd, d. Crawford, K. (2012). “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information, Communication, & Society* 15 (5): 662-679.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985-1002.

Тема 3. Лекция. Социальные последствия непродуманной работы с данными в разных сферах. Обзор основных кейсов, в которых проявляются социальные последствия непродуманной работы с данными. Выявление основных правил, которые могут предотвратить

подобные кейсы. Обзор инициатив, которые рассматривают эту проблему и предлагают различные решения.

Gillespie, T. (2010). "The Politics of 'Platforms.'" *New Media & Society* 12 (3): 347–64.
Data Science professional Code <http://www.datascienceassn.org/code-of-conduct.html>

Тема 4. Семинар. Делая данные и алгоритмы подотчетными: изучение и обсуждение мировых практик контроля технических процессов обществом. Обсуждение собственных данных, собранных участниками и постановка проблем для практических занятий.

Тема 5. Семинар. Выявление того, как именно собираются данные и на каком этапе заложены возможные конфликты и непредсказуемые социальные последствия. Семинар-игра, на которой участники делятся на разные заинтересованные группы и представляют сценарии сбора и анализа данных и их последствия для этих групп.

Тема 6. Семинар. Предотвращение проблем с данными и алгоритмами. Заключительное занятие с презентацией карт возможностей и рисков работы с данными для разных стейкхолдеров.

Задания

1. Медиа-дневник, позволяющий понять, какие именно данные каждый человек оставляет как пользователь
2. Составить карту возможностей и рисков использования данных в своей области. Карта составляется по конкретной сфере: городское планирование, экономика, журналистика, международные отношения и т.д. Здесь указаны именно форматы и общие темы.

III. ОЦЕНИВАНИЕ

$$O_{\text{итог}} = O_{\text{п}} * 0,30 + O_1 * 0,3 + O_2 * 0,4, \text{ где}$$

$O_{\text{п}}$ — оценка за работу на семинарах,

O_1 — оценка за задание 1.

O_2 — оценка за задание 2.

Критерии задания 1. Медиа-дневник

- 4 балла Полнота и достаточная подробность описания
- 4 балла Наличие аналитической части, описывающей объект
- 2 балла Наличие выводов и структурность изложения

Критерии задания 2. Карта возможностей и рисков

4 балла Карта соответствует фактической ситуации на рынке и подкреплена ссылками и указанием источников

4 балла Карта представляет собой аналитическую работу, в которой представлены и описаны риски и возможности

2 балла Наличие выводов и структурность изложения

IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

Описываются примеры оценочных средств или ссылка на наличие оценочных материалов на сайте дисциплины в LMS.

V. РЕСУРСЫ

1. Основная литература

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208.

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2), 2053951715594634.

boyd, d. Crawford, K. (2012). "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication, & Society* 15 (5): 662-679.

Gillespie, T. (2010). "The Politics of 'Platforms.'" *New Media & Society* 12 (3): 347-64.

Data Science professional Code <http://www.datascienceassn.org/code-of-conduct.html>

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985-1002.

2. Дополнительная литература

O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Van Zoonen, L. (2016). Privacy concerns in smart cities. *Government Information Quarterly*, 33(3), 472-480.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *new media & society*, 14(7), 1164-1180.

Lupton, D. (2016). *The quantified self*. John Wiley & Sons.

Mai, J. E. (2016). Big data privacy: The datafication of personal information. *The Information Society*, 32(3), 192-199.

Winner, Langdon. (1980). "Do Artifacts Have Politics?" *Daedalus* 109(1): 121-136.

Gillespie, Tarleton. (2014). "The relevance of algorithms." Pp. 167-194 in T. Gillespie, P.

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (Eds.). (2014). Media technologies: Essays on communication, materiality, and society. MIT Press.

Fourcade, M, and Kieran H. (2016). "Seeing Like a Market." Socio-Economic Review 1-21.

<https://medium.com/@laura.noren/critical-data-studies-thinking-group-join-6e3340f66393>

3. Программное обеспечение

№ п/п	Наименование	Условия доступа
1.	Microsoft Windows 7 Professional RUS Microsoft Windows 10 Microsoft Windows 8.1 Professional RUS	<i>Из внутренней сети университета (договор)</i>
2.	Microsoft Office Professional Plus 2010	<i>Из внутренней сети университета (договор)</i>

4. Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

№ п/п	Наименование	Условия доступа
	<i>Интернет-ресурсы (электронные образовательные ресурсы)</i>	
1.	Открытое образование	URL: https://openedu.ru/

5. Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.

Учебные аудитории для лабораторных и самостоятельных занятий по дисциплине оснащены _____, с возможностью подключения к сети Интернет и доступом к

электронной информационно-образовательной среде НИУ ВШЭ.