

**EthicalDataScience.**  
**New York University**  
**Professor Anne L. Washington, PhD**  
**APSTA-GE 2062 Ethics of Data Science**  
***DRAFT SHARE* Syllabus Spring 2020**

Topics:

1. Human Rights / Dignity and Autonomy
2. Human Subjects Research / Informed consent / Accountability / Validity
3. Classification, Fairness, Interpretation
4. Moral Philosophy of Kant, Bentham, Aristotle, Gilliam, Rawls, Rousseau
5. Case studies: Google Flu Trends, Facebook Emotional Contagion, Microsoft Tay, COMPAS risk assessment, Cambridge Analytica, Twitter content moderation

Learning Outcomes:

- Critically evaluate projects involving digital representations of human behavior
- Identify competing interests in the data supply chain
- Analyze the interdependence of values, data, & operational processes
- Demonstrate familiarity with theoretical models of ethics
- Give compelling oral presentations

**About EthicalDataScience. :**

**Ethics of Data Science** explores the societal consequences of extracting inferences about human behavior from digital abstractions. Historical and theoretical perspectives will serve as the foundation for exploring contemporary concerns about data science, predictive analytics, artificial intelligence, and other data-driven tools. The course is designed to build students' ethical imaginations and data literacy using both inductive and deductive reasoning. The course provides practical guidance on how to uncover ethical weaknesses as well as construct principled data-driven projects.

Data technology weaves together our social fabric by comparing digital information about people across time and populations. Comparing an individual to a larger population historically has always had moral, legal, and social implications. What is different today is that digital comparisons immediately cycle back to impact our lived experience of friendship, navigation, government, health care, community, politics, employment, transportation, or commerce. Because data are rarely created for one-time use, downstream ethical implications may accumulate, especially when comparing across jurisdictional or cultural contexts. Data science that shapes society requires expert technical abilities as well as careful critical thinking skills in quantitative reasoning. Together we will ask core questions about what it means to be human within predictive data technology.

**Assessment:**Attendance

You are expected to attend class regularly with the exception of rare work-related commitments. Important preparation for assignments is conducted during class time. Be aware that missing class might make it difficult to complete assignments appropriately and result in a lower grade. Please get the number of another student if you anticipate missing a session.

Teaching Philosophy

This class emphasizes experiential learning. In this model, students and instructors share a responsibility for learning. Reading, thinking-aloud, activities, practice, and revision are essential elements. Experiential learning is an ideal form of learning for adults because it assumes that everyone is interested in their own growth and development.

**Discussions.** Ethics requires conversations. Excellent class discussion participants are able to build on the contributions of others, focus on the substance of the course, ask reflective questions, and respect everyone in the room.

Assignments heighten an awareness of competing interests in ethical dilemmas. Students will explore ethical topics through their own cultural values as well as learning to reflect on the logic of alternative perspectives. By completing case studies, discussions, problem sets, and essays, students will have an opportunity to consider both the technical and theoretical aspects of ethical data science. Students will be encouraged to connect readings from in-depth reporting, technology ethnography, and primary sources to relevant current events. Students in this seminar will learn how technical choices in statistics and algorithms can impact society.

(c)

**EthicalDataScience. Books:**

Books are available through independent bookstores and at the university bookstore. Page numbers will be from the 2018 5<sup>th</sup> edition.

- Baase, S., & Henry, T. (2018). A gift of fire social, legal, and ethical issues for computing technology (5th ed.). ISBN: 978-0-13-461527-1  
*Baase presents a balanced exploration of the social, legal, philosophical, ethical, political, constitutional, and economic implications of computing and the controversies they raise. With an objective computer scientist's perspective, and with historical context for many issues, Baase covers the issues students will face both as members of a technological society and as professionals in computer-related fields.*

## EthicalDataScience. Assignments:

Complete details about each assignment will be available on the classroom technology platform.

**Online Forum :** Each student will post to the online forum 24 hours before class on scheduled weeks. Students will submit a short quote from the week's reading along with a question, comment, or image relevant to the topic. (30%)

**Case Study Leaders :** In pairs, students will serve as case study discussion leaders. Students will present an ethics case-study for 10-15 minutes and stimulate the class with a vigorous 15-20 minute discussion or exercise. The goal is to connect readings to real professional, personal, or societal situations. (10%)

**Homework :** Students will gather data on themselves and reflect on what they discover. Some information gathered for homework will be the basis for exercises during subsequent classes. (15%)

**Debate :** In groups, students will hold a debate about the societal consequences of abstracting human behavior into digital representations. Groups representing competing interests in a specific situation will argue their perspective and answer questions from other groups. An entire class period is devoted to the debate. (15%)

**Classroom Labs :** Together we will complete short thought experiments in our weekly ethics lab during class. Each student will submit documentary evidence of their thinking each week. (30%)

## EthicalDataScience. Schedule

The semester is organized into two-week segments where we will rotate between discussing concepts and analyzing case studies. The structure of the class follows the sequence of a data science project to expose points where technical choices have social ramifications. This practical format strengthens the student's ability to articulate specific points for improvement.

### 1 Introduction

On being human. On being digital. Defining data.

### 2 Human Rights

**EthicsLab:** Humanitarian Data Exchange

### 3 Dignity

Case Study (A): Vulnerable populations & mental health

**EthicsLab:** Future of Privacy Forum & De-identification

**4 Human Subjects Research****EthicsLab:** Quantified Self & Self-tracking**5 Informed Consent**Case Study **(B)**: Facebook Emotional Contagion Study**EthicsLab:** Terms of Service. Didn't Read ToS;DR**6 Modeling Reality****EthicsLab:** AOIR code of ethics**7 External Validity**Case Study **(C)**: Google Flu Trends**EthicsLab:** Search Trends**8 Social Classification****EthicsLab:** Consumer Scoring & Ranking**9 Individual Autonomy**Case Study **(D)**: Gender Binary**EthicsLab:** Facial Recognition**10 Calculated Intention**Case Study **(E)**: Social Prediction & Criminal Justice**EthicsLab:** ProPublica COMPAS Data <https://github.com/propublica/compas-analysis>**11 Fairness & Applied Ethics**Case Study **(F)**: Government algorithms & Due process**12 Data Science Reasoning**

Debate Case: Predicting social problems using data technology

**13 Debate****14 Conclusion****EthicsLab:** Everyday Algorithms

anne.washington@nyu.edu 0801-1215