Ethics and Societal Implications of Artificial Intelligence

Philosophy 5340
Fall 2020
Friday 8:30-11:20 EST
Dr. Devlin Russell

After a crash course in ethical theory, we'll look at the ethical consequences of A.I. now, near, and far. Right now, machine learning and big data are manipulating democracies, public opinion, preferences, education, and opportunity. We will analyze this manipulation through the lens of ethical theory. We will then look at issues involving autonomous cars, artificial caregivers, and artificial killers. And finally we will turn to the singularity: the point at which A.I. meets or exceeds human consciousness. When that happens, there will be some serious, brain-bending consequences for humanity.

Here's what you will *not* get out of this course: answers. So what will you get out of this course? The ability to formulate clear, precise, and fruitful questions about how and whether artificial intelligence *should* be implemented. And the ability to work out for yourself the answers to these questions. The ethical consequences of the choices you make will exist whether you look for them or not, and even the choice to ignore them is an ethical choice.

REMOTE LEARNING

This course will not meet in person.

Seminars will be conducted over Zoom at the scheduled time:
Friday 8:30-11:20 EST.
They will not be recorded, so you must attend at the scheduled time.

Course materials and assignments will be distributed and submitted through Canvas. Please review York University's guide to eLearning.

Office hours will be conducted over Zoom. They will be limited to 15 minutes per students, by appointment only, to a maximum of 2 hours per week. Contact me by email to arrange an appointment. First-come; first-serve.

e-mail: devlin@yorku.ca
E-mail correspondence will only be accepted from a YorkU registered e-mail address.
I will respond as soon as I can within two business days.

MATERIALS

The following are required materials for this course:

1. Shafer-Landau readings posted on Canvas.

2. O'Neil, C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Broadway Books (2017). ISBN: 978-0553418835

3. Lin, P., Jenkins, R., Abney, K. (Eds.). *Robot Ethics 2.0: From Autonomous Cars to Artificial intelligence*. Oxford University Press (2019). ISBN: 978-0197503584

4. Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press (2016). ISBN: 978-0198739838

REQUIREMENTS

The following are mandatory elements of taking this course:

1. Seminar attendance, engagement, and participation.
2. Thoroughly reading and thinking about the assigned readings before class.
3. Eleven conversation starters.
4. Six writing assignments (900-1200 words).

EVALUATION

	Due	Course Grade
Conversation Starters	Before Class	18%
Writing Assignment 1	October 4	12%
Writing Assignment 2	October 18	12%
Writing Assignment 3	November 1	12%
Writing Assignment 4	November 15	12%
Writing Assignment 5	November 29	12%
Writing Assignment 6	December 13	12%
Participation	Ongoing	10%

ABSOLUTELY NO LATE ASSIGNMENTS

No late assignments will be accepted. Late assignments will receive a zero. If you have a major emergency and expect that you cannot meet the deadline, you must contact me as soon as possible and <u>before</u> the deadline.

SCHEDULE OF READINGS

The following schedule is subject to change.

Week 1: Introduction
9/18   No reading

Week 2: Ethical Theory
9/25   Shafer-Landau - *The Fundamentals of Ethics* Ch. 19-21

Week 3: Ethical Theory
10/2   Shafer-Landau - *The Fundamentals of Ethics* Ch. 5-10

WA1 DUE OCTOBER 4

Week 4: Ethical Theory

10/9   Shafer-Landau - *The Fundamentals of Ethics* Ch. 11-16

Week 5: Now

10/16   O'Neil - *Weapons of Math Destruction* Intro.-Ch. 5

WA2 DUE OCTOBER 18

Week 6: Now

10/23   O'Neil - *Weapons of Math Destruction* Ch. 6-Concl.

WA3 DUE NOVEMBER 1   READING WEEK OCTOBER 27 - 30

Week 7: Near

11/6   Lin (et al.) - *Robot Ethics 2.0* Part I

Week 8: Near

11/13   Lin (et al.) - *Robot Ethics 2.0* Part II

WA4 DUE NOVEMBER 15

Week 9: Near

11/20   Lin (et al.) - *Robot Ethics 2.0* Part III

Week 10: Far

11/27   Bostrom - Superintelligence Ch. 1-5

WA5 DUE NOVEMBER 29

Week 11: Far

12/4   Bostrom - Superintelligence Ch. 6-11

Week 12: Far

12/11   Bostrom - Superintelligence Ch. 12-15

WA6 DUE DECEMBER 13

ACCOMMODATIONS

York University is committed to principles of respect, inclusion and equality of all persons with disabilities across campus. The University provides services for students with disabilities (including physical, medical, learning and psychiatric disabilities) needing accommodation related to teaching and evaluation methods/materials. These services are made available to students in all Faculties and programs at York University.

Students in need of these services are asked to register with disability services as early as possible to ensure that appropriate academic accommodation can be provided with advance notice. You are encouraged to schedule a time early in the term to meet with each professor to discuss your accommodation needs.

Please note that registering with disabilities services and discussing your needs with your professors is necessary to avoid any impediment to receiving the necessary academic accommodations to meet your needs. (accessibility.students.yorku.ca)

ACADEMIC INTEGRITY

In this course, we strive to maintain academic integrity to the highest extent possible. Please familiarize yourself with the meaning of academic integrity by completing SPARK's Academic Integrity module at the beginning of the course. Breaches of academic integrity range from cheating to plagiarism (i.e., the improper crediting of another's work, the representation of another's ideas as your own, etc.). All instances of academic dishonesty in this course will be reported to the appropriate university authorities, and can be punishable according to the Senate Policy on Academic Honesty.

Turnitin

To promote academic integrity in this course, students will be normally required to submit their written assignments to Turnitin (via the course

Moodle) for a review of textual similarity and the detection of possible plagiarism. In so doing, students will allow their material to be included as source documents in the Turnitin.com reference database, where they will be used only for the purpose of detecting plagiarism. The terms that apply to the University's use of the Turnitin service are described on the Turnitin.com website.

STUDENT RIGHTS & RESPONSIBILITIES