

Ethics of New Technologies

Michal Masny

Description This course examines key technological developments and challenges of the 21st century from an ethical perspective. We will discuss: self-driving cars and autonomous weapons systems; surveillance and the right to privacy; algorithmic discrimination in the criminal justice system; the future of work and the promise of unconditional basic income; human enhancement; and the risk of human extinction.

Level Advanced undergraduate course.

Prerequisites No previous work in philosophy or STEM is required or assumed.

Requirements This course will be held in a seminar format. You will review the assigned materials in advance and then we will discuss them together during the seminar.

Student grades are determined by five components:

- First paper (1000 words): 15%
- Second paper (1500 words): 15%
- Third paper (2000 words): 30%
- Participation: 30%
- Oral presentation: 10%

Class participation is very important in this course. I am asking you do two things:

- Post a reaction to at least one of the assigned materials on the course discussion board by midnight on the day before the seminar. You can clarify a tricky aspect of the paper, raise an objection, ask a pertinent question, bring attention to an interesting case study, or simply identify an issue that you would like us to talk about during the seminar. I will read your comments and do my best take them up during the seminar.
- Actively participate in the discussions during the seminar. Please bear in mind that the topics covered in this course can be controversial. At the same time, do not hesitate to present arguments in support of unpopular positions or objections to those that are widely held.

You are required to do one oral presentation on a topic of your choosing. A presentation should be no more than 8-10 minutes. It can consist of a philosophical commentary on one of the readings, an explanation of the technical aspects of one of the discussed technologies, or a case study applying philosophical concepts to a specific practical issue. You must meet with me a few days before your presentation to go over an outline.

Blind grading Please submit your work with no identifying information other than your student number.

Office hours I hold weekly office hours. I welcome you to use this opportunity. You don't need to come with a specific question about the readings or an assignment. We can just grab a cup of coffee and talk about philosophy.

Resources Don't be discouraged if you have difficulty following an argument in an assigned reading or if you misunderstand something. Philosophy is challenging and many philosophers are not gifted stylists. I often read philosophical articles twice and encourage you to do the same. You can find many helpful tips about reading philosophical texts, participating in seminars, and writing essays in this [Pink Guide to Philosophy](#) by Helena de Bres. Another excellent resource, regularly used by students and professors alike, is the [Stanford Encyclopedia of Philosophy](#).

Week 1: Introduction to Ethical Theory

Shafer-Landau explains the ambitions, the limits, and the tools of ethics. Hooker outlines different forms that a consequentialist theory of ethics can take.

- Russ Shafer-Landau. (2012). Introduction. In his *Fundamentals of Ethics*. (17 pages)
- Brad Hooker. (2010). Consequentialism. In *The Routledge Companion to Ethics*. (12 pages)

Week 2: Self-Driving Cars

Thomson's article is a classic statement of the famous 'Trolley Problem'. Nyholm and Smith examine the relevance of this problem for the design of autonomous vehicles.

- Judith Thomson. (1986). Killing, Letting Die, and the Trolley Problem. In her *Rights, Restitution, and Risk: Essays in Moral Theory*. (16 pages)
- Sven Nyholm. & Jilles Smids. (2016). The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem? *Ethical Theory and Moral Practice*. (15 pages)
- *Bonus*: Wes Bard. (2020). [Tesla Autopilot is better than you think! Here's why. \(part 1\)](#). (7 minutes)

Week 3: Technology and War

McMahan's piece is an accessible introduction to contemporary Just War Theory. Shane makes the case for the use of drones as more precise and humane alternatives. Kaag and Kreps raise the worry that drones may lower the threshold for resorting to force. Sparrow considers who should be held responsible when an autonomous weapon is involved in an atrocity that would normally be described as a war crime.

- Jeff McMahan. (2012) Rethinking the 'Just War,' [Part 1](#), and [Part 2](#). *NYT*. (8 pages)
- Scott Shane. (2012). [The Moral Case for Drones](#). *NYT*. (4 pages)
- John Kaag and Sarah Kreps. (2012). [The Moral Hazard of Drones](#). *NYT*. (4 pages)
- Robert Sparrow. (2007). Killer Robots. *Journal of Applied Philosophy*. (16 pages)
- *Bonus*: Gavin Wood's film *Eye in the Sky* (2016), starring Helen Mirren and Alan Rickman.

Week 4: Privacy and Surveillance I

Thomson argues that putative violations of the right to privacy are really just violations of other rights, such as the right to control how other people use our body and property. Marmor argues that there *is* a distinctive right to privacy which is grounded in people's interest in having control over the ways in which they can present themselves to others.

- Judith Thomson. (1975). The Right to Privacy. *Philosophy & Public Affairs*. (20 pages)
- Andrei Marmor. (2015). What is the Right to Privacy? *Philosophy & Public Affairs*. (24 pages)

Week 5: Privacy and Surveillance II

Veliz describes the extent to which corporations and governments collect data on us. The Guardian provides a brief overview of government surveillance programmes revealed by Edward Snowden. Snowden explains how our actions are being tracked through our phones. Kosinski and colleagues demonstrate how sensitive information about people can be reliably inferred from what might seem like innocuous online activity.

- Carissa Veliz. (2020). Introduction and Chapter 1: Data Vultures and Chapter 2: How Did We Get here?. In her *Privacy is Power: Why and How You Should Take Back Control of Your Data*. (24 pages)
- The Guardian. (2013). [The NSA and surveillance ... made simple - animation](#). (3 minutes)
- Edward Snowden. [How Your Cell Phone Spies on You](#). JRE Podcast. (24 minutes)
- Michal Kosinski, et al. (2013). [Private traits and attributes are predictable from digital records of human behavior](#). *Proceedings of the National Academy of Sciences*. (4 pages)

Week 6: Algorithmic Discrimination I

Angwin and colleagues argue that an algorithm used widely in the criminal justice system is biased against Black people. Zimmerman and colleagues argue that problems arising from the use of algorithms cannot be solved simply by designing better algorithms. Alexander explains huge racial disparities at every stage of the criminal justice process.

- Julia Angwin, et al. (2016). [Machine Bias](#). *ProPublica*. (~6 pages)
- Annette Zimmerman, Elena di Rosa, and Hochan Kim. (2020). [Technology Can't Fix Algorithmic Injustice](#). *Boston Review*. (~10 pages)
- Michelle Alexander. (2010). Chapter 3: The Color of Justice. In her *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. (42 pages)

Week 7: Algorithmic Discrimination II

Long responds to the ProPublica investigation and argues against one measure of algorithmic fairness.

- Robert Long. (2021). Fairness in machine learning: Against false positive rate equality as a measure of fairness. *Journal of Moral Philosophy* (30 pages)

Week 8: The Future of Work I

Vox considers whether technological developments will lead to mass unemployment. Gheaus and Herzog argue that work is important because it gives us an opportunity to attain excellence at something, make a social contribution, experience community, and gain social recognition. Danaher argues that work is bad and keeps getting worse, and that its technological elimination would be desirable.

- Vox. (2017). [The big debate about the future of work, explained](#). Video. (9 minutes).
- Anca Gheaus & Lisa Herzog. (2016). The Goods of Work (Other Than Money!). *Journal of Social Philosophy*. (20 pages)
- John Danaher. (2019). Chapter 3: Why you should hate your job. In his *Automation and Utopia: Human Flourishing in a World without Work*. (34 pages)

Week 9: The Future of Work II

Bregman outlines the basic case for Universal Basic Income. Van Parijs and Vanderborght consider the objection that UBI would be unfair. Gourevitch and Stanczyk critique UBI from the Left.

- Rutger Bregman. (2014). [Why we should give everyone a basic income](#). *TedX Talk*. (17 minutes)
- Philippe van Parijs and Yannick Vanderborght. (2017). Chapter 5: Ethically Justifiable? Free Riding Versus Fair Shares. In their *Basic Income: A Radical Proposal for a Free Society and a Sane Economy*. (34 pages)
- Alex Gourevitch and Lucas Stanczyk. (2018). The Basic Income Illusion. *Catalyst Journal*. (22 pages)

Week 10: Enhancement

Buchanan and colleagues argue that parents should be free to use genetic intervention techniques to produce the best offspring. Sandel argues that the pursuit of perfection will undermine our appreciation for the 'gifted' character of human life and parental relations.

- Allen Buchanan et al. (2001). Chapter 5: Why not the best? In their *From Chance to Choice: Genetics and Justice*. (47 pages)
- Michael Sandel. (2004). [The Case Against Perfection](#). *The Atlantic Monthly*. (~8 pages)

Week 11: Existential Risk I

Ord explains why safeguarding humanity's future is the key challenge of our time.

- Toby Ord. (2020). [The Precipice: existential risk and the future of humanity](#). Narrated presentation summarizing his new book by the same title. (20 min)

Week 12: Existential Risk II

Beckstead argues that achieving even small reductions in existential risk is more important than greatly improving the lives of currently existing people. Scheffler suggests that many things that we care deeply about would lose their value if human extinction was imminent

- Nick Beckstead. (2019). A brief argument for the overwhelming importance of shaping the far future. In *Effective Altruism: Philosophical Issues*. Edited by Hilary Greaves and Theron Pummer. Oxford University Press. (19 pages).
- Sam Scheffler. (2018). Chapter 2: Reasons to Worry – Interests and Love. In his *Why Worry about Future Generations?* (28 pages).