Course Outline
Ethics of Artificial Intelligence

Course Description:
The ethics of artificial intelligence in a relatively new, but burgeoning, area within applied ethics that deals with moral questions related to artificial intelligence and its use. The first section of this course will familiarize students with philosophical methodology, provide a history of artificial intelligence, and address basic issues in describing what artificial intelligence is and is not. The second section of this course will cover concepts in normative ethics, including moral theories, trust, accountability, and risk. The last section of this course will survey applied ethics issues related to artificial intelligence, including relationships with artificial intelligence, education, autonomous vehicles, and military uses of artificial intelligence.

Course Texts:
I am aware of the financial impediments posed by acquiring materials for class. To remove this potential barrier to your academic success, all readings are free-to-use and made available electronically.

Course Reading Schedule:
Day 1: Introduction
Day 2: Reading and Writing Philosophy; University Resources
Reading: 1,000 Word Philosophy, "How to Read Philosophy"
Reading: Harvard Writing Center, "A Brief Guide to Writing the Philosophy Paper"

Day 3: The History of Artificial Intelligence
Reading: Stanford Encyclopedia of Philosophy, "Artificial Intelligence" (Section 1)
Day 4: What is Artificial Intelligence?
Reading: Stanford Encyclopedia of Philosophy, "Artificial Intelligence" (Section 2)

Day 5: The Turing Test
Reading: Alan Turing, "Computing Machinery and Intelligence"
Day 6: Weak vs. Strong Artificial Intelligence
Reading: John Searle, "The Chinese Room"

Day 7: Descriptive and Normative Ethics
Reading: Christoph Bartneck, et al., "What is Ethics?" (pp. 17-20)
Day 8: Ethics and Machines
Reading: Christoph Bartneck, et al., "What is Ethics?" (pp. 20-26)

Day 9: Trust and Fairness: Non-Maleficence and Beneficence
Reading: Christoph Bartneck, et al., "Trust and Fairness in AI Systems" (pp. 27-32)
Day 10: Trust and Fairness: Justice and Explicability
Reading: Christoph Bartneck, et al., "Trust and Fairness in AI Systems" (pp. 33-37)

Day 11: Responsibility and Liability
Reading: Christoph Bartneck, et al., "Responsibility and Liability in the Case of AI Systems"
Day 12: Accountability
Reading: Reuben Binns, "Algorithmic Accountability and Public Reason"

Day 13: Risk
Reading: Christoph Bartneck, et al., "Risks in the Business of AI" (pp. 45-49)
Day 14: Risk
Reading: Christoph Bartneck, et al., "Risks in the Business of AI" (pp. 49-53)

Day 15: Anthropomorphism and Psychological Problems
Reading: Christoph Bartneck, et al., "Psychological Aspects of AI"
Day 16: Sex Robots
Reading: Christoph Bartneck, et al., "Application Areas of AI" (pp. 78-80)

Day 17: Privacy
Reading: Christoph Bartneck, et al., "Privacy Issues of AI" (pp. 61-67)
Day 18: Privacy
Reading: Christoph Bartneck, et al., "Privacy Issues of AI" (pp. 67-70)

Day 19: Enhancement
Reading: Christoph Bartneck, et al., "Application Areas of AI" (pp. 71-73)
Day 20: Healthcare
Reading: Christoph Bartneck, et al., "Application Areas of AI" (pp. 73-76)

Day 21: Education
Reading: Christoph Bartneck, et al., "Application Areas of AI" (pp. 76-78)
Day 22: Education
Reading: Joshua Schulz, "Machine Grading and Moral Learning"

Day 23: Autonomous Vehicles
Reading: John Gogoll and Julian Muller, "Autonomous Cars"
Day 24: Autonomous Vehicles
Reading: Fiona Woollard, "The New Trolley Problem"

Day 25: Military Uses
Reading: Christoph Bartneck, et al., "Military Uses of AI" (pp. 93-96)
Day 26: Military Uses
Reading: Christoph Bartneck, et al., "Military Uses of AI" (pp. 97-99)

Day 27: Politics
Reading: John Danaher, "The Threat of AIgocracy"
Day 28: Politics
Reading: Bartek Chomanski, "Legitimacy and Automated Decisions"

Day 29: Artificial Intelligence and the Future
Reading: Bing Song, "Applying Ancient Chinese Philosophy to Artificial Intelligence"
Day 30: Artificial Intelligence and the Future
Reading: Stanford Encyclopedia of Philosophy, "Artificial Intelligence" (Section 9)