

**PHIL 3350: Robot Ethics**  
**(a.k.a. Philosophy, Ethics and Public Affairs)**

**Course Outline: Fall 2012**

**Prof: Jason Millar**

**Department of Philosophy – Carleton University**

**Logistics**

Class Times: TBD

Location: TBD

Instructor's Office: TBD

Instructor's Office hours: TBD

Email: jlmillar@connect.carleton.ca

- 1 *A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
- 2 *A robot must obey orders given to it by human beings except where such orders would conflict with the first law.*
- 3 *A robot must protect its own existence as long as such protection does not conflict with the first or second law.*

Isaac Asimov (*Runaround*, 1941)

**Course Description**

*Context:*

Robots are becoming ubiquitous in society. They are being used in every branch of the military to patrol streets, disarm bombs, gather intelligence, and kill from a distance. Software “bots” are populating the Internet, “chatting” autonomously with unsuspecting individuals on Facebook and Twitter, and building complex social networks, often to profile “friends” and “followers”. According to statistics released by the International Federation of Robotics, sales of service robots for personal use will top 14.4 million units through 2011-2014. Microsoft founder Bill Gates recently argued that we are at the point now with personal robots that we were with personal computers in the 1970’s.

The desire to expand and study the social implications of robots is widespread. Worldwide, research in artificial intelligence (AI) and robotics has acquired a new sense of promise over the past decade and a half. That is due in part to recent synergies between mathematics, neuroscience, computer science and psychology, and also to recent field “successes” with military robots, autonomous online agents, swarm robots and service bots. Consider Watson, IBM’s artificial Jeopardy! contestant. Watson recently beat two of Jeopardy!’s most winning human players in a match they called “Man vs. Machine”. IBM is currently deploying “Dr.” Watson in major health care centers in the US in order to perform complex system analyses and to aid in diagnosing disease. Indeed, robots are beginning to assume humanlike roles in society: establishing personal relationships with humans in order to assist in tasks; teaching students; patrolling warzones; and caring for the elderly. The pace of innovation suggests that future generations of robots will be quite unlike their “simple-minded” ancestors. Robots are increasingly autonomous by design and, according to many scientists and philosophers, they will surpass human intelligence within the next half-century. Eventually it might be extremely difficult to distinguish robots from humans.

Robots are challenging many of our philosophical and ethical notions while suggesting the need for new public policy. The European Commission has sponsored research into “roboethics” as a means of anticipating the ethical challenges associated with pursuing various robot technologies and deploying them in, and across, societies. Recently, academic institutions in North America have held major workshops (e.g. see <http://robots.law.miami.edu>) and are publishing more and more on the field of the social implications of robotics. It is an exciting time for the academic study of robotics!

What is a robot? Can robots be moral agents? What would that agency look like? How will robots impact privacy? Should robots be used to wage war? What are the current directions in roboethics, and robot design, and are they adequate/justifiable? What kinds of relationships ought we to encourage between robots and humans, or robots and robots? To what extent are people to be held morally or legally liable for the actions of robots? What norms, if any, ought to be considered with respect to governing the design of, and interaction with, robots? What might we owe robots in the future?

#### *Course Objectives:*

The goal of this course is to investigate the above questions through an examination of classic and contemporary texts in ethics, philosophy of mind, science and technology studies, public policy, and a newly emerging area called “robot ethics”. Students will be challenged to consider how philosophical concepts including identity, consciousness, moral agency, personhood, artificial intelligence, and rationality apply to roboethics. As this course focuses on ethics and policy issues, an emphasis will be placed on examining and developing normative claims from the literature. By engaging public policy documents students will also have an opportunity to examine and critique the application of philosophical concepts in the context of interdisciplinary science debates.

#### *Pedagogy:*

This course does not require any previous training in technology or robotics. The course will consist of a combination of lectures and regular class discussions. In addition to the readings a variety of sources will be engaged throughout the course including sci-fi literature, movies, print media and television as a means of providing thought experiments for framing discussions. Students will be encouraged to draw on all sources to participate in regular class discussions and are expected to have completed all readings prior to class.

#### **Course Readings:**

Required readings for the course will be drawn primarily from *Robot Ethics* and from *Can Animals and Machines be Persons?*, which are available at the Campus Bookstore. There are also required electronic resources available either through the Carleton Library, or online via the provided links.

#### **Required Texts:**

1. Robot Ethics: An Anthology, MIT Press (Campus Bookstore)
2. Can Animals and Machines be Persons?, Justin Leiber (Campus bookstore)
3. Online articles specified in the reading list.

#### ***Weekly Reading Schedule:***

Week 1: Setting the Scene: An Intro to Roboethics

Week 2: Overview of the Problems of Roboethics

Week 3: The Imitation Game I

Week 4: The Imitation Game II

Week 5: Robots and Personhood I

Week 6: Robots and Personhood II

Week 7: Military Robots

Week 8: Privacy & Law

Week 9: Social Robots and Human-Robot Interaction

Week 10: Sex and Robots

Week 11: Expert Robots

Week 12: Engineering Ethics and the Sociology of Technology