**Artificial Intelligence, Ethics, and Politics 2020/2021**

Optional Course: MA in Governance, Leadership, and Democracy Studies, Instituto de Estudos Politicos, Universidade Catolica Portuguesa (IEP-UCP)

Professor: William Hasselberger

Hours: 9

ECTS: 3

Within a short span of time, Artificial Intelligence (A.I.) has moved from the margins of philosophical speculation and science fiction to an increasingly powerful and prevalent aspect of contemporary life.

Various forms of Artificial Intelligence are now used to analyze enormous troves of data and assist doctors to diagnose our medical conditions and prescribe treatments, to assist banks to more precisely assess our credit worthiness, to assist judges in determining the length of a criminal sentence and a prisoner’s likelihood of reoffending or suitability for parole, to assist us in deciding what to click on when online, what to watch, what to buy, whom to date, whom to vote for.

Artificial Intelligence is also remaking modern work, outperforming many human experts, and some argue that A.I. will soon enable the automation of vast numbers of white collar and service sector jobs in advanced economies—and, thereby, render many human trades, skills, and workers, obsolete.

Artificial Intelligence is also used by many countries’ police departments, by intelligence agencies around the world, by militaries, and by authoritarian governments to ever more carefully surveil their citizens in public and online and to stifle dissent and unrest.

This seminar will explore the vexing ethical and political questions raised by the rapid advances in machine intelligence, from the spheres of economics, domestic and international politics, and everyday life. Some examples include: Should A.I. be used to make morally significant decisions about human life and death? Are there forms of intelligence that are uniquely human? When a sophisticated A.I. delivers a different verdict than a human judge, who should have the final say? What is the spread of A.I. doing to human intelligence and skill in the material world? Are there human jobs that advanced A.I. can’t replace? Will further advances and applications of Artificial Intelligence serve to strengthen open societies and liberal democratic principles, or weaken them and strengthen high-tech authoritarianism? What are the appropriate ethical norms for A.I. development? What is a trustworthy, humane role for A.I. in society in the coming decades?

The course evaluation will be based on two factors: (i) regular attendance and active seminar discussion (20%) and (ii) a final essay, roughly 12-15 pages in length, including notes (80%).

**Part 1: What is Artificial Intelligence? (And what is *human* intelligence?)**

Readings (selections):

Joseph Weizenbaum, *Computer Power and Human Reason*

Meredith Broussard, *Artificial Unintelligence*

Brian Cantwell Smith, *The Promise of Artificial Intelligence*

**Part 2: The Economic and Social Impact of Artificial Intelligence**

Readings:

Kai Fu Lee, *A.I. Superpowers*

Martin Ford, *The Rise of the Robots*

Sean Kelly, “A.I. Can’t Be an Artist”

Nicholas Carr, *The Glass Cage: Automation and Us*

**Part 3: Emotional A.I. and Machine Morality**

Readings:

Gary Marcus, *Rebooting A.I., Building Artificial Intelligence We Can Trust*

Nick Bostrom, *Superintelligence*

Roberto Simonawski, *The Death Algorithm and Other Digital Dilemmas*

**Part 4: Surveillance Capitalism and Surveillance Authoritarianism**

Readings:

Shoshana Zouboff, *The Age of Surveillance Capitalism*

Tim Wu, “Bigger Brother”

Brookings Institution Policy Brief, “Exporting Digital Authoritarianism”