Course Title: The Philosophical Foundations of Artificial Intelligence  
Course Code: PHI 104  
Instructor: Forrest Hartman

**Course Summary:** The attempt to create artificial intelligence in machines necessarily presumes a certain understanding of philosophical concepts: What thinking is and what it means to be conscious are first of all questions about the philosophy of mind, the discussion of which may call into question prevailing paradigms in cognitive science and may even suggest alternative ways of thinking. Whether consciousness and intelligence are necessarily embodied, even extended beyond the body, are some of the questions we will consider in looking at the future prospects of the cognitivist model of symbolic representation and manipulation as the paradigm for the design of artificial intelligence.

**Grade Options and Requirements:**

- **No Grade Requested (NGR)**  
  - This is the default option. No work will be required; no credit shall be received; no proof of attendance can be provided.
- **Credit/No Credit (CR/NC)**  
  - Score will be determined by student attendance and participation. Whoever chooses this option needs to sign in on an attendance sheet each week.
- **Letter Grade (A, B, C, D, No Pass)**  
  - Written work, probably a 5 to 10 page paper, the topic of which is to be arranged with the instructor, will determine the grade.
- **Attendance Requirement:** Participants must attend at least six of the eight sessions in order to receive either a grade or credit for the course. Be sure to indicate your attendance at each session on the roster provided.

*Please Note: If you require proof that you completed a Continuing Studies course for any reason (for example, employer reimbursement), you must choose either the Letter Grade or Credit/No Credit option. Courses taken for NGR will not appear on official transcripts or grade reports.*

**Tentative Weekly Outline:**

**Week 1: 28 June**  
Overview and Background of the Artificial Intelligence Project: Definitions, Intellectual History, Current Developments, and Philosophical Orientation  
_Artificial Intelligence: What Everyone Needs to Know_, pp. 1 to 89.
**Week 2: 5 July**
The Cartesian background of the prevailing cognitive paradigm of artificial intelligence: Man, the programmable machine, and the mechanization of the modern world view in the 17th century, the Age of Enlightenment.
*Meditations, “Meditation 6,”* (online) and
*The Turing Test: Verbal Behavior as the Hallmark of Intelligence,* Chapters 1 through 3, pp. 17 to 58.

**Week 3: 12 July**
The Turing Test as the first approximation of a way to evaluate whether machines can think and why it may be an inadequate measure of intelligence:
*The Turning Test,* Chapters 4 through 7, pp. 61 to 140 and
*Philosophy for a New Century,* Chapter 3, pp. 53 to 67 (Canvas).

**Week 4: 19 July**
Could Machines Think?
*Philosophy for a New Century,* Chapter 5, pp. 86 to 107 and
*The Turing Test,* Chapter 16, pp. 267 to 293.

**Week 5: 26 July**
John Searle’s (in)famous Chinese Room Argument:
*The Turing Test,* Chapter 14, pp. 199 to 225 and
*Philosophy for a New Century,* Chapter 4, pp. 67 to 86.

**Week 6: 2 August**
The phenomenological critique of cognitivism: the frame problem and the embodied mind as being-in-the-world:
*What Computers Still Can’t Do* (Canvas) and
“Why Heideggerian AI Failed and How Fixing it would Require it to be More Heideggerian” (online) and
“Heideggerian AI and the Being of Robots” (Canvas)

**Week 7: 9 August**
The Extended Mind and Enactivism: You are not your merely your brain. Consciousness may be something we *do* and not something we *have.*
*Out of Our Heads,* Preface and Chapters 1 through 4, pp. 3 to 96.

**Week 8: 16 August**
Robot Ethics and Moral Machines: Can machines be ethical; will artificial *general* intelligence constitute an existential threat to our humanity; will superintelligent robots replace human beings altogether, or even domesticate us; should we be afraid of the rise of the robots; do we need to take precautions?
Nick Bostrom’s *Superintelligence* (selections) and
*Moral Machines* (selections) and other readings on machine ethics, as time permits.