AI Robots & Ethics 2015/16
Dr. David J. Gunkel
Week 4

Objective
This week we will continue our investigation of machine moral agency by looking at the way philosophers have defined moral agency in general and "machine moral agency" in particular. In reading this week's texts, we not only want to understand the different sets of criteria that have been advanced by this effort, but we also want to evaluate the philosophical assumptions and consequences of this strategy by asking whether such criteria can be (or have ever been) sufficient for defining moral agency.

Readings
Kenneth Einar Himma, Artificial Agency
John Sullins, When is a Robot a Moral Agent?

Questions
1) For both Himma and Sullins, moral agency is defined and characterized on the basis of a set of "qualifying criteria." What are the criteria offered by Himma? What are the criteria offered by Sullins? How are they similar? How are they different? In other words, do a comparison and contrast of these two accounts of the criteria for machine moral agency.

2) Why are the two accounts of criteria for moral agency different? What is the effect of this difference on our understanding of moral agency? And what complications does this difference introduce into the debate concerning machine moral agency? In other words, if we cannot be entirely certain about the exact ontological criteria that make someone (or something) a moral agent, can this effort be considered an effective means for deciding the question of machine moral agency? Why or why not?