

Filosofia dell'intelligenza artificiale

Prof. Dr. Mario De Caro

Sigla: FFSI

Semestre primaverile

Ects: 3

Obiettivo del corso: offrire i fondamenti concettuali delle discussioni epistemologiche e morali sull'intelligenza artificiale.

Descrizione del corso: HAL 9000, Terminator, i replicanti di Blade Runner's replicants, Eva di Ex machina: i film di fantascienza sono pieni di macchine intelligenti, che il più delle volte danneggiano gli esseri umani – se non vogliono distruggerli tutti. E, in effetti, sebbene nessuno possa seriamente dubitare che gli straordinari, e sempre più veloci, progressi tecnologici – e in particolare l'avvento dell'intelligenza artificiale – apportino notevoli benefici alle nostre vite, altrettanto indubitabile è che tali progressi generino nuove sfide e minacce gravi, se non terribili. Durante il corso discuteremo tre di queste minacce: la grave disoccupazione che l'A.I. può generare; gli usi potenzialmente distorti, dal punto di vista morale e sociale, delle macchine intelligenti; e la futuribile, ma non remota, possibilità che l'intelligenza artificiale possa autonomizzarsi in modo consapevole, al punto da provocare danni enormi al genere umano, indipendentemente dalla volontà dei suoi programmatori.

Modalità didattica: insegnamento in presenza.

Modalità di valutazione: compito scritto.

Bibliografia essenziale:

- J. SEARLE, "Minds, Brains, and Programs", <http://cogprints.org/7150/1/10.1.1.83.5248.pdf>;
- N. BOSTROM, "The Ethics of Artificial Intelligence", <https://www.nickbostrom.com/ethics/artificial-intelligence.pdf>;
- J. THOMAS, "In defense of philosophy: a review of Nick Bostrom, Superintelligence: Paths, Dangers, Strategies", <https://www.tandfonline.com/doi/full/10.1080/0952813X.2015.1055829>