



Module Announcement

PhD in Information Technology and Electrical Engineering

Università degli Studi di Napoli Federico II

Module Title: *Intelligenza Artificiale ed Etica: La ricerca in IA alla prova delle sfide etiche (workshop)*

Lecturers:

- **Daniele Amoroso**, Dipartimento di Giurisprudenza, Università di Cagliari
- **Piero A. Bonatti**, DIETI, Università di Napoli Federico II
- **José M. Galvan**, Dipartimento di Teologia Morale, Pontificia Università della Santa Croce, Roma
- **Riccardo Guidotti**, KDD-Lab – ISTI-CNR, Pisa
- **Paola Inverardi**, DISIM, Università degli studi dell'Aquila
- **Roberto Prevete**, DIETI, Università di Napoli Federico II
- **Luciano Serafini**, ICT, Fondazione Bruno Kessler, Trento
- **Viola Schiaffonati**, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano

Dates and Locations: Ed.3, Aula seminari (ex Softel), I piano, via Claudio 21, Napoli

- 09.20 D. Amoroso: *Sistemi d'arma autonomi e (ir)responsabilità giuridica e morale*
- 10.00 R. Prevete: *Quali spiegazioni per il controllo etico dei sistemi di machine learning?*
- 10.40 V. Schiaffonati: *Intelligenza artificiale e metodo scientifico: dall'epistemologia all'etica*
- 12.00 L. Serafini: *Integration of Learning and Reasoning in Logic Tensor Networks*
- 12.40 R. Guidotti: *Explaining Explanation Methods*
- 14.30 J. Galvan: *Il dilemma morale della libertà di espressione nei networked publics*
- 15.10 P. Inverardi: *La Soft-Ethics ed il progetto EXOSOUL*
- 16.00 P. Bonatti: *L'Etica è sostenibile?*





Module Announcement

Motivations

Il workshop raccoglie ricercatori che operano in vari settori - scientifici, tecnologici e umanistici - intorno alle questioni etiche e giuridiche sollevate dall'Intelligenza Artificiale (IA) e alle prospettive che tali questioni aprono per la ricerca scientifica e tecnologica in IA.

L'analisi di questioni etiche sarà al centro di varie relazioni al workshop: la manipolazione dell'opinione pubblica e la creazione di "echo chambers" mediante l'IA, le decisioni delle armi autonome e il diritto internazionale, le responsabilità dell'uomo alla luce delle sue limitate capacità di interpretare, spiegare e prevedere i sistemi dell'IA. Al centro di altri contributi al workshop vi saranno i problemi di comprensione e controllo dei sistemi di IA. Si tratta di problemi per la ricerca in IA che hanno forti motivazioni di tipo etico-giuridico: trasparenza, interpretazione simbolica e spiegazione dei processi e delle decisioni dei sistemi dell'IA, conformità dei sistemi dell'IA alle preferenze o agli obblighi morali di gestori e utenti, la fiducia che l'uomo può in essi riporre.

Nel programma del workshop, i contributi dei due tipi (di analisi etico-giuridica dell'IA e di ricerca in IA con motivazioni etico-giuridiche) sono intercalati tra loro. Questa scelta riflette l'intenzione di favorire il dialogo e l'interazione produttiva tra ricercatori, diversi per formazione – umanistica, scientifica, tecnologica – ma accomunati dall'obiettivo di capire e affrontare le nuove sfide etiche per l'IA.

ECTS Credits: 0.2 per ora

Info: **Dott. Roberto Prevete** - tel. 081 679262 - roberto.prevete@unina.it

