

Un'iniziativa di



In collaborazione con



LAKE COMO SCHOOL
OF ADVANCED STUDIES

LE SERATE DELLA LAKE COMO SCHOOL OF ADVANCED STUDIES

AI good vs. evil

14 maggio ore 18.00, Sala Recchi (Palazzo Lambertenghi)

Filippo Menczer

Luddy school of informatics, computing and engineering
Indiana University

Visita il sito alessandrovolta.it



Un'iniziativa di



In collaborazione con



LAKE COMO SCHOOL
OF ADVANCED STUDIES

LE SERATE DELLA LAKE COMO SCHOOL OF ADVANCED STUDIES

Filippo Menczer



"Luddy distinguished professor" di informatica e scienza dei computer, nonché direttore dell'Osservatorio sui Social Media presso l'Università dell'Indiana. Ha conseguito la laurea in Fisica alla Sapienza Università di Roma e un dottorato di ricerca (Ph.D.) in Informatica e Scienze Cognitive presso l'Università della California, San Diego. La sua ricerca recente si focalizza sulla vulnerabilità dei social network alla disinformazione e alla manipolazione. Il Dr. Menczer è, inoltre, un Fellow dell'ACM e dell'AAAS.

Abstract

La manipolazione online è un campo di battaglia in cui l'IA viene utilizzata come arma sia dagli attaccanti sia dai difensori. Questa presentazione offre una panoramica della ricerca che esplora entrambi i lati. Attori malintenzionati stanno sfruttando l'IA generativa per creare profili e contenuti falsi difficili da riconoscere, con facilità e a prezzi contenuti. Allo stesso tempo, l'IA fornisce strumenti per mitigare i danni online, come metodi per riconoscere account inautentici e operazioni di informazione coordinate, scalare il fact-checking e supportare la moderazione. Anche dal lato "positivo" del campo di battaglia, dobbiamo tenere conto dei limiti dell'IA, delle conseguenze indesiderate quando le persone interagiscono con essa, e delle vulnerabilità dell'IA a forme sottili di bias e abuso.