DIPARTIMENTO DI
INFORMATICA, SISTEMISTICA E
COMUNICAZIONE
Viale Sarca, 336
20126 Milano

SEMINAR ANNOUNCEMENT

Wednesday, 4th June 2025
at 11:00 am
Room "Sala Seminari" - Abacus Building (U14)

Flushing out AI-Generated Content with Cryptography

Speaker
Dr Luca Mariot
University of Twente (The Netherlands)

Abstract
With the recent success of Large Language Models (LLMs), the web is increasingly being flooded by AI-generated content, a phenomenon also known as "AI slop". This poses the following question: how do we ascertain whether a particular piece of content (be it text, images, or other formats) has been produced by a generative AI model? This attribution problem has significant societal implications, for instance in education where academic cheating through ChatGPT or similar products is becoming more and more common. Moreover, training the next generation of AI models could become problematic if we are not able to distinguish AI-generated from human-made content, especially sizeable portion of the training data available on the web will be synthetic.

In this talk, we give an overview of the recent literature that addresses this identification problem using cryptography. Specifically, we review a few recent techniques based on pseudorandom functions and pseudorandom codes to apply a stealthy watermark on the output of LLMs, in such a way that only those who know the corresponding decryption key can later verify whether some piece of content has been indeed generated by the watermarked model.

Short bio
Luca Mariot is an assistant professor at the University of Twente, the Netherlands. His main research interests lie at the intersection of cryptography and artificial intelligence, focusing on natural computing models and techniques to design cryptographic primitives. Previously, Luca was a postdoc researcher at Radboud University and TU Delft, the Netherlands, and at the University of Milano-Bicocca, Italy. He received his PhD in Computer Science under a double degree agreement, from the University of Milano-Bicocca and the Université Côte d'Azur, France.