

The Seminars on “Information Technology Outlook” – PhD Program in Computer Science and Mathematics



Chiara Di Bonaventura

PhD student in
Safe and Trusted Artificial Intelligence
at King's College London and Imperial College London

Wednesday January 17, 2024
15:00



Aula Hume
Dept. of Computer Science

Temporal dynamics in abusive language detection

Abusive language, a pervasive issue in online communication, poses significant challenges to maintaining a safe and inclusive digital environment. Among these challenges, this seminar explores the dynamic nature of abusive language, highlighting the shifts in language patterns and their temporal biases that make traditional detection models less effective over time. Temporality introduces a layer of complexity, requiring innovative approaches for effective detection mechanisms. The integration of external resources, particularly knowledge graphs, has emerged as a promising solution to enhance abusive language detection systems in several scenarios: from stereotypes and implicit abusive language to languages other than English. The seminar will discuss the potential synergies between language models and knowledge graphs in addressing the challenge of temporality in abusive language detection.

Chiara Di Bonaventura is a second-year PhD student in Safe and Trusted Artificial Intelligence at King's College London and Imperial College London. Currently, she works on leveraging symbolic knowledge for abusive language detection systems under the supervision of Albert Meroño-Peñuela and Barbara McGillivray. Before moving to London, she graduated in the MSc in Data Science at Bocconi University under the supervision of Dirk Hovy. Her research interests include Natural Language Processing, Online Abusive Language, AI Ethics and Knowledge Graphs.