

Maria Vanina Martinez, Artificial Intelligence Res. Institute (IIIA-CSIC), Barcelona, Spain

AI and cybersecurity: a delicate balance between data and knowledge

The effective and efficient management of data and expert knowledge lies—to a greater or lesser extent—at the heart of any successful application of Artificial Intelligence in any of its variants, ranging from purely perceptual (subsymbolic AI) to approaches that employ machinery based on computational logic (symbolic AI). Hybrid or Neurosymbolic AI (NeSy) emerged decades ago as a promising way to develop intelligent tools that could embrace the best of both worlds. It is now gaining greater relevance in the face of the weaknesses exhibited by production software based on subsymbolic methods. This talk will first provide a brief introduction to NeSy AI and Cybersecurity, a discipline currently undergoing a revolution due to the emergence of both defensive and offensive AI. Next, we will present some of the key threats observed in the modern socio-technical ecosystem, and present preliminary results on the use of NeSy AI to address them.

Ethical and Legal Approaches to the Autonomy of AI Systems (*panel discussion*)

Salient properties required for autonomous decision-making will be examined, exploring how far we are in the process of ensuring these properties in current AI systems. Moreover, the legality of Autonomous Weapons Systems (AWS) will be examined, focusing on whether AWS may align with International Humanitarian Law (IHL) and its principles of distinction, proportionality, and precaution. Finally, the two-tiered approach to prohibiting and regulating AWS, advanced by the International Committee of the Red Cross (ICRC), will be examined. The moral foundations of the ICRC distinction between regulated versus prohibited systems will be explored, as well as the moral foundations for the specific requirements for human control in regulated systems.