



Contribution ID: **10**

Type: **not specified**

Keynote - Savannah Thais (Columbia): Trustworthy and Robust AI for Science and Society

Monday 19 August 2024 13:00 (55 minutes)

Artificial intelligence (AI) and machine learning (ML) have become critical tools in many scientific research domains. In many areas of physics ML tools are essential to meeting the computing needs of current and future experiments and to ensuring robust data reconstruction and interpretation. In addition to being powerful tools for scientific research, ML and AI are now ubiquitous in nearly all facets of society, from healthcare to criminal justice, from education to transportation. These applications have the potential to address critical community needs and improve educational, health, financial, and safety outcomes; however, they also have the potential to exacerbate existing inequalities and raise concerns about privacy, surveillance, and data ownership.

In this talk I will explore critical ethical considerations that arise when designing, developing, and deploying data science and ML/AI systems when modeling both scientific and social systems, including bias and fairness, task design, equitable data practices, model evaluation, and more. These topics are essential to ensuring we can trust the results of our AI models in our scientific research and to helping us build a more complete understanding of how AI models behave and when we can appropriately use them to make or support decisions. Although these are not purely mathematical problems that can be fully solved with technological approaches, I will emphasize quantitative approaches and techniques to address these issues and discuss what unique roles physicists might play in co-creating a more just and responsible future of technology. I will show examples of recent work from particle physics and astronomy that exhibit how physics data can be used to develop AI models that both enhance the scientific capacity of our experiments and help us gain foundational insight into AI.

Session Classification: Plenaries & Keynotes