



Contribution ID: 59

Type: **not specified**

## **Towards Compliance of AI Algorithms : Bias Analysis and Robustness.**

*Thursday, 16 May 2024 09:00 (1 hour)*

AI is about to revolutionize many sectors of society and will soon have significant economic, legal, social and regulatory consequences.

In the world of production, transport, human resource management, and health, to name but a few, a growing share of diagnostic and planning processes is operated by AI-based systems.

Controlling the risks of deploying these AI for high risk systems is therefore becoming a considerable challenge for a variety of actors in society. The authorities in charge of regulation, in the first place, but also the producers of these systems themselves as well as the companies that buy and operate them.

While their performance gains are no longer in doubt, their interpretability and the transparency of their decisions is now considered crucial for their wider deployment. Auditing these processes based on artificial intelligence, if we refer to the DSA, DMA and AI texts, will become a necessity in the same way as the quality standards that govern the production of goods and market services. However, auditing a complex algorithm, based on AI-based components –to the point of identifying and measuring its risks and biases– is a scientifically and technically delicate task.

We will explain how mathematical methods based on the theory of optimal transport can provide a natural framework able to handle algorithmic biases, from their quantification and measures, to the mitigation process. Audit methods based on such tools can be used to certify the presence of disloyal behavior that may lead to possible discrimination in numerous fields but in an industrial context can also be used to identify deviations in the data that may lead to a loss of performance.

### **Type of presentation**

Invited Talk

**Primary author:** LOUBES, Jean-Michel (Université Toulouse)

**Presenter:** LOUBES, Jean-Michel (Université Toulouse)

**Session Classification:** Keynote

**Track Classification:** Spring Meeting