

## Catégorisation et IA Vision du monde et action sur le monde Entretiens de l'AFSCET

Institut des Systèmes Complexes de Paris-Ile de France

113 rue Nationale, 75013 Paris, salle 1 au premier étage

vendredi 25 novembre 2022 de 14h à 18h

### Programme

14h - 14h45: **Jean-Marc Brossier (Université de Grenoble)**

Inférer sans loi. Importance des représentations, des règles et de l'*a priori* en apprentissage statistique.

Résumé. L'intelligence artificielle, terme consacré et peu pertinent, la machine pouvant à la fois plus et infiniment moins que l'homme, a évolué dans bien des directions. L'une des tendances a été le glissement de systèmes qui n'utilisent pas de données, mais seulement des règles explicites fournies par des experts humains, vers des systèmes qui produisent des décisions en inférant des règles implicites à partir de données. Dans un cadre probabiliste, une caractéristique des problèmes de classification réside dans le fait que la loi des données est inconnue contrairement à des contextes plus classiques comme l'inférence Bayésienne. L'approche "PAC" (Probablement Approximativement Correct) permet d'appréhender ce contexte et d'établir qu'il n'existe pas une méthode universellement la meilleure (théorèmes de type "no free lunch") : une connaissance *a priori* du problème est nécessaire. Inévitable, cet *a priori* est le fondement d'une régularisation qui permet de dégager un ordre dans le chaos des données ; il est aussi à l'origine de biais dans la prise de décision.

14h45 - 15h : discussion.

15h - 15h45: **Pierre Collet (Université de Strasbourg)**

Vision du monde et action sur le monde par l'IA via la théorie des catégories et des topos de Grothendieck

Résumé. L'objectif de notre travail de recherche est de nous interroger sur les différentes manières possibles de construire artificiellement des connaissances ; de nous intéresser, donc, à une épistémologie de l'intelligence artificielle (IA). Nous souhaitons nous interroger sur les énoncés produits par des intelligences artificielles et, un peu plus largement, sur les impacts qu'ont de tels énoncés d'un point de vue éthique sur celles et ceux qui les utilisent. Cet exposé présentera donc la manière dont les IA, par l'intermédiaire d'outils tels que les catégories et les Topoi, amènera à mieux voir le monde et agir sur lui de manière éthique.

15h45 - 16h30 : discussion, puis pause

16h30 - 18h : Table ronde animée par **Danièle Bourcier (CNRS, CERSA, Paris)**

Chaque intervenant pourra présenter sa vision du sujet en dix à quinze minutes au maximum pour laisser la place à la discussion.