

## Conférence « Les robots et le mal » d'Alexei Grinbaum

Dans le cadre de la Chaire Éthique & IA de l'Institut MIAI Grenoble Alpes,

le 22 janvier de 14 à 16 h,

à l'amphi de la Maison de la Création et de l'innovation (MaCi)

Des robots domestiques se font délateurs, des agents conversationnels injurient leurs interlocuteurs. Pire encore : des systèmes informatiques participent aux conflits humains et parfois même les provoquent. Le 18 mars 2018, un véhicule autonome de la société Uber a tué une femme qui traversait la rue dans une ville de l'Arizona. Ce fut la première mort d'un piéton provoquée par un algorithme.

Qui est responsable ? La réponse à cette question compte parmi les défis les plus urgents à relever dans notre rapport aux technologies numériques. Mais il ne s'agit pas de savoir comment rendre l'intelligence artificielle bienveillante. Il s'agit de faire en sorte qu'elle ne se substitue pas à l'homme en tant qu'agent moral. **Seul le recours au hasard, et ceci dès sa conception, peut libérer la machine de la responsabilité qu'on veut lui faire porter.**

*Deep learning techniques lead to fundamentally non-interpretable decisions made by the machine. Although such choices do not have an explanation, they impact the users in significant ways. If the ultimate innovator is a machine, what is the meaning of responsible conduct? I argue in a recent book that the capacity to extract an AI system from human judgment, by reducing transparency in favor of opacity, is an essential value in machine ethics. This can be achieved through the use of randomness, illustrated on several examples including the trolley dilemma. Methodologically, a comparison of common motives between technological setups and mythological narratives is used to achieve ethical insights.*

Alexei Grinbaum  
**LES ROBOTS ET LE MAL**



**Ne diabolisons pas  
les machines !**

DESCLÉE DE BROUWER



**Alexei Grinbaum** est philosophe et physicien. Chercheur au laboratoire Larsim du CEA-Saclay, il est spécialiste de l'information quantique. Depuis 2003, il s'intéresse aux questions éthiques liées aux nouvelles technologies, notamment aux nanotechnologies, à l'intelligence artificielle et à la robotique. Il a été coordinateur pour la France de l'Observatoire européen des nanotechnologies et partenaire du projet européen « Recherche et innovation responsables en pratique » (RRI-Practice). Membre du Comité national pilote d'éthique du numérique et de l'IA et de la Commission d'éthique de la recherche en numérique (Cerna), il a publié « Mécanique des étreintes » (Encre Marine, 2014) et « Les robots et le mal » (Desclée de Brouwer, 2019).

**Informations et inscriptions :** [tyler.reigeluth@univ-grenoble-alpes.fr](mailto:tyler.reigeluth@univ-grenoble-alpes.fr)

**Accès :** Maison de la Création et de l'Innovation, Université Grenoble Alpes, 339 avenue Centrale.