

---

# Agents moraux artificiels : Approches, proxys et signaux

Dominic Martin\*<sup>1</sup>

<sup>1</sup>Université du Québec à Montréal = University of Québec in Montréal – Université du Québec à Montréal CP 8888, succursale Centre-ville Montréal (Québec) H3C 3P8, Canada

## Résumé

La possibilité et la manière de créer un agent moral de manière artificielle soulèvent des questions importantes et les développements récents en intelligence artificielle et dans les technologies du numérique ont mené à une explosion des contributions scientifiques sur le sujet. Mon but dans cette présentation sera de dresser un portrait de la littérature sur les agents moraux artificiels. J'essayerai ensuite de montrer quelles sont les suggestions les plus prometteuses pour développer ce type d'agent tout en évitant les écueils associés à une réflexion trop spéculative. Premièrement, il faut distinguer les agents moraux avec un haut ou bas niveau de compétence moral, car nous sommes déjà en mesure de développer des systèmes de bas niveau. Deuxièmement, il ne faut pas seulement réfléchir aux grandes approches pour développer un agent moral, mais aussi aux meilleurs proxys et signaux pour fournir une information morale.

**Mots-Clés:** Intelligence artificielle (IA), Éthique, Agents moraux artificiels

---

\*Intervenant