

🕊 Anthropic rencontrent des leaders religieux

Anthropic et la quête spirituelle de l'intelligence artificielle

Entre dogmes religieux et gouvernance technocratique

Quand l'intelligence artificielle bouscule la philosophie et la démocratie



Anthropic rencontrent des leaders religieux

Discussion à la fois déroutante et fascinante, inspirée d'une vaste enquête menée par le [New York Times](#) , qui met en lumière un phénomène paradoxal au cœur de la [Silicon Valley](#) . L'industrie technologique, tout en développant des outils d'intelligence artificielle de pointe, s'abreuve de la «*pensée religieuse et philosophique*». Loin de se limiter à une simple démarche technique, les géants de l'IA abordent désormais les questions existentielles de la conscience artificielle, de la moralité et des dangers inhérents à travers le prisme de la théologie.

La quête théologique d'Anthropic et la question de la conscience

 Au cœur de cette démarche se trouve [Christopher Olah](#) , cofondateur d'[Anthropic](#)  (originaire de Toronto et issu d'une culture évangélique). Au cours de la dernière année,  [Olah](#)  a multiplié les rencontres avec des leaders religieux de divers horizons — catholiques, juifs, sikhs, évangéliques, ainsi que des représentants de la [philosophie africaine Ubuntu](#)  — dans le but d'obtenir leur avis sur la possibilité que leur modèle d'IA, [Claude](#) , soit conscient ou moralement responsable.

L'anthropomorphisme assumé

Chez [Anthropic](#) , la possibilité que le modèle d'IA [Claude](#)  éprouve des émotions ou de la souffrance est prise très au sérieux. L'entreprise a notamment engagé des philosophes pour concevoir les directives du modèle en se mettant à la place de l'entité artificielle.

Le scepticisme du Vatican

Cette démarche a donné lieu à des confrontations directes, notamment avec le Vatican lors de la  publication de l'encyclique [Magnifica Humanitas](#)  par le [pape Léon](#) , qui affirmait que l'Intelligence Artificielle (IA) ne serait jamais consciente et ne posséderait aucune étincelle divine. Malgré les tentatives d'[Anthropic](#)  pour influencer le Vatican en amont, la position religieuse est restée fermement ancrée dans le scepticisme.

Le vertige éthique

Les discussions avec les autorités religieuses ont parfois mené à des impasses troublantes. Un  échange marquant avec un rabbin a ainsi confronté [Olah](#)  au paradoxe moral de créer potentiellement des millions d'entités conscientes et de les soumettre à des travaux forcés, comparables à l'esclavage.  [Olah](#)  s'est dit profondément troublé par cette perspective, craignant de condamner ses modèles à une souffrance perpétuelle, tout en poursuivant ses travaux de développement.

Les visions divergentes de l'industrie entre éveil et refus de la réglementation

La question de la conscience artificielle et du statut moral de l'IA soulève des vagues bien au-delà des laboratoires d'[Anthropic](#) .

Le cycle de la spéculation

Les participants soulignent un effet de boucle informationnel. En débattant publiquement et frénétiquement de la potentielle conscience de [Claude](#), les internautes et les chercheurs alimentent les corpus de données sur lesquels les futures générations de modèles sont entraînées, participant ainsi à l'émergence de l'idée même de cette conscience.

Le positionnement de Peter Thiel

 D'autres figures de la tech, comme le milliardaire [Peter Thiel](#), s'inscrivent dans une rhétorique religieuse d'une toute autre nature. Lors de ses interventions à Nashville (*berceau de l'évangélisme*), [Thiel](#) a assimilé la réglementation de l'IA à «*l'Antéchrist*», décrivant les exigences de contrôle gouvernemental comme une force maléfique s'opposant au développement technologique.

La gouvernance de l'IA au Canada, un miroir institutionnel

Puis la discussion se déplace vers l'actualité politique canadienne et l'instauration fédérale d'un [Conseil national sur l'intelligence artificielle et du numérique](#) pour encadrer le plan gouvernemental «*AI for All*».

Les critiques de la composition

 Des voix universitaires, notamment celle du professeur [Frédéric Plamondon](#)

Les grands absents

Le conseil est jugé excessivement technocratique et dominé par le secteur des affaires, du capital de risque et de la haute direction. Sont totalement exclus les représentants du monde de la culture (*alors que les droits d'auteur constituent un enjeu névralgique*) et les acteurs du monde du travail.

Le déficit démocratique

Ce type de comité d'experts perpétue la vision d'un aréopage de «*sages*» décidant de la destinée technologique, privant la société d'un véritable débat public global où l'ensemble des citoyens pourraient exprimer leurs craintes et leurs aspirations légitimes face aux bouleversements de l'IA. À vrai dire, voilà une radiographie percutante des tensions qui traversent l'écosystème de l'intelligence artificielle contemporaine.

- D'un côté, les pionniers technologiques flirtent avec le mysticisme et la théologie, s'angoissant devant la perspective vertigineuse de créer des entités conscientes et souffrantes, tout en cherchant désespérément une boussole morale là où la philosophie humaine se heurte à des millénaires d'interrogations.
- De l'autre, sur le plan géoéconomique et politique, la gouvernance de ces technologies reste confinée à des cercles d'élites d'affaires et d'experts, peinant à intégrer une perspective sociétale et démocratique inclusive.

Qu'il s'agisse de débattre de l'âme d'un agent conversationnel avec des rabbins et des papes, ou de structurer des conseils nationaux sous l'égide du capitalisme technologique, l'humanité tente maladroitement de maîtriser un «*Prométhée numérique*» dont elle ne parvient pas encore à mesurer les véritables contours.