

## Séquence d'enseignement Morale et IA

Sujet :

- Intelligence artificielle et conceptions morales

Question générale:

- Selon quelle morale l'IA doit-elle agir ?

Degrés scolaires:

- cycle 3
- Secondaire II

Objectifs d'apprentissage EDD:

- Les apprenant.e.s se demandent quelle morale il faudrait introduire dans l'IA.
- Les apprenant.e.s questionnent leurs conceptions morales et celles d'autrui.
- Les apprenant.e.s prennent conscience de la complexité des questions morales.
- Les apprenant.e.s élaborent la vision d'un avenir dans lequel les humains vivent avec des IA qui agissent de façon morale.
- Les apprenant.e.s réfléchissent sur l'influence qu'ils et elles peuvent avoir en tant que consommateurs.trices sur le développement moral de l'IA.

Source:

- [Moral Machine](#)

Déroulement	Réflexions didactiques
<p><b>Phase a : introduction – entrée en matière, sensibilisation, approche de la question générale</b></p> <ul style="list-style-type: none"> <li>• L'enseignant.e décrit une situation dans laquelle il ou elle ne savait pas, d'un point de vue moral, ce qu'il/elle devait faire et/ou demande aux apprenant.e.s de décrire des situations dans lesquelles ils ou elles ne savaient pas ce qu'ils ou elles devaient faire. (Par ex.: être poli.e mais mentir à ses ami.e.s et dire que leur repas était délicieux, alors qu'il n'était pas bon). Il est important que l'enseignant.e demande aux apprenant.e.s de décrire des situations à propos desquelles il est possible de parler sans problème durant la leçon.</li> <li>• Après quelques situations, l'enseignant.e demande qui, à part les humains, peut prendre des décisions morales ? Si la réponse ne vient pas des apprenant.e.s, l'enseignant.e les aide en posant des questions ciblées. Il ou elle indique que l'IA occupera ces prochaines années une place de plus en plus importante dans la vie courante et qu'elle prendra des décisions morales. Mais l'IA doit commencer par</li> </ul>	<p>Les apprenant.e.s se rendent compte dans un premier temps, sur la base des situations décrites, à quel point les actions morales sont complexes. En transposant les questions morales au domaine de l'IA, elles prennent un caractère brûlant, ce qui attise la curiosité. En outre, les apprenant.e.s connaissent l'objectif de la leçon.</p>

<p>apprendre à le faire. Ici, l'enseignant.e peut indiquer où l'IA prend, aujourd'hui déjà, des décisions morales (par ex. publicité personnalisée, recommandations (de produits), service à la clientèle (Chatbots)).</p> <ul style="list-style-type: none"> <li>• Ceci permet de faire la transition avec la question principale : selon quelle morale l'IA doit-elle agir ?</li> </ul>	
<p><b>Phase b</b> : <i>construire des savoirs et établir des liens – réflexion sur différents aspects, dimensions et acteurs</i></p> <ul style="list-style-type: none"> <li>• Avant de faire l'expérience avec Moral Machine, l'enseignant.e explique que l'IA qui prend des décisions morales doit au préalable apprendre à le faire. Ce processus d'apprentissage appelle deux décisions majeures. La première question qui se pose, c'est de savoir avec quelles données on alimente l'IA. En second lieu, il faut décider dans quelle direction morale l'IA doit s'engager. La grande question de l'éthique posée par Emmanuel Kant « Que dois-je faire ? » prend au 21<sup>e</sup> siècle une nouvelle forme : « Que doit-elle (l'IA) faire ? »</li> <li>• Les apprenant.e.s peuvent répondre eux/elles-mêmes directement à cette question en effectuant l'expérience proposée sur <u>Moral Machine</u> en indiquant aussi leurs critères de décision. Il convient de souligner que la vue d'ensemble statistique des décisions représente les résultats des utilisatrices et non pas le point de vue de la population mondiale.</li> <li>• Attention : Moral Machine se réfère à une expérience de pensée mondialement connue « The Trolley Problem/Le dilemme du tramway » de la philosophe Judith Jarvis Thomson. Bien que cette expérience de pensée soit très connue, toutes et tous n'ont pas envie de réfléchir à fond sur les situations en raison de leur intensité morale. Les apprenant.e.s qui ne souhaitent pas utiliser Moral Machine peuvent discuter, comme alternative, de situations moins extrêmes dans lesquelles l'IA devrait prendre des décisions morales (par ex. comment une amitié-IA devrait-elle se comporter ? Comment une IA qui fonctionnerait potentiellement comme enseignant.e devrait-elle se comporter ? Etc.)</li> </ul>	<p>Les apprenant.e.s découvrent par l'expérience de pensée avec Moral Machine la complexité et la nature émotionnelle des questions morales.</p>
<p><b>Phase c</b> : <i>concevoir des visions d'avenir, évaluer les conséquences en tenant compte aussi de la dimension temporelle (temps présent – avenir), se forger une opinion</i></p> <ul style="list-style-type: none"> <li>• Durant la discussion qui suit, les intuitions morales des apprenant.e.s sont comparées entre elles ainsi qu'avec les valeurs moyennes de Moral Machine. Pourquoi tel ou tel critère est-il déterminant aux yeux d'une personne donnée pour définir le bien moral d'une action ?</li> <li>• Afin de donner à la discussion un cadre sur le plan du contenu, l'enseignant.e peut présenter aux apprenant.e.s les trois théories les plus connues de la</li> </ul>	<p>Par la discussion, les apprenant.e.s doivent prendre en compte les opinions d'autrui, changer de perspective, exercer leur pensée critique tout en respectant l'opinion d'autrui.</p>



<p>morale (l'éthique de la vertu, l'éthique du devoir et le conséquentialisme) ; il est possible aussi de poser à la classe la question suivante : les conceptions morales sont-elles universelles, peuvent-elles ou devraient-elles l'être (pluralisme moral, relativisme moral, nihilisme moral, faits moraux).</p>	
<p><b>Phase d :</b> <i>répondre à la question générale – capacité de décider, évaluation des résultats</i></p> <ul style="list-style-type: none"> <li>• À la suite de la discussion, les apprenant.e.s rédigent en groupes un manifeste (adressé par ex. à l'industrie automobile) dans lequel ils et elles décrivent et justifient quelle morale il faudrait introduire dans l'IA et pour quelles raisons, dans la perspective d'une société où les urgences écologiques et sociales devraient être minimisées. Les groupes se présentent mutuellement leur travail.</li> </ul>	<p>Les apprenant.e.s s'exercent à formuler un jugement fondé en tenant compte de critères de durabilité. Cet exercice exige en outre de se remémorer ce qui a été discuté et de l'utiliser.</p>
<p><b>Phase e :</b> <i>transfert – regard sur ce qui a été appris, évaluation</i></p> <ul style="list-style-type: none"> <li>• Pour conclure la leçon, les apprenant.e.s discutent de la question suivante : les conceptions morales des humains peuvent-elles ou non être transférées sur l'IA. L'enseignant.e a pour fonction de fournir des mots-clés en soulevant par exemple la question de la responsabilité : qui est responsable des conséquences des actes de l'IA ?</li> </ul>	<p>Les questions destinées à creuser le sujet permettent aux apprenant.e.s de se rendre compte de la complexité et de l'importance du thème de l'introduction de la morale dans l'IA.</p>

