

---

# L'éthique de la robotique

Les avancées combinées de l'ingénierie robotique et des technologies d'intelligence artificielle permettent une diversification des domaines où sont utilisés des systèmes robotiques. Cette factsheet discute les principaux enjeux éthiques et présente un outil de navigation pour aborder les questions pertinentes.

## Qu'est-ce qu'un robot?

Les robots sont partout, mais nous ne les voyons pas toujours à cause des lunettes que nous avons chaussées. Films, dessins animés, livres de science-fiction nourrissent notre imaginaire de robots très semblables aux humains. À l'inverse, le robot industriel qui assemble des pièces métalliques sur une chaîne de montage n'a aucune présence culturelle. Pour les spécialistes d'ingénierie robotique, cette omniprésence d'un imaginaire robotique est à la fois chance et malheur. Chance, car tout le monde s'intéresse à la robotique, malheur car les attentes sont démesurées. Le robot réel est presque condamné à décevoir.

Mais au juste, de quoi parlons-nous exactement? Le roboticien américain Georges Bekey a proposé l'un des définitions les plus connues du robot: une machine située dans le monde qui sent, pense et agit (*a machine, situated in the world, which senses, thinks and acts*).

- **Machine:** tous les artefacts réalisant une tâche, créés par les humains, qu'importe la matériau, la forme ou le type de mécanique à l'oeuvre
- **Situé dans le monde:** le robot peut se déplacer, saisir, pousser, tirer ou toutes autres actions exerçant une force sur le monde physique qui l'entoure. Cette capacité à agir directement sur le monde – et donc sur les humains – les qualifie comme un type d'entité particulière. Cet élément exclue les outils exclusivement numériques.
- **Qui sent, pense et agit:** le robot est créé afin de remplir un objectif. Pour l'atteindre, il est capable de percevoir l'environnement qui l'entoure. Il traite ensuite ces informations en utilisant certaines fonctions d'habitude associées à la réflexion et à l'intelligence. Il choisit ensuite des actions qu'il réalise sur son environnement. Les machines entièrement télécommandées et les automates exécutant un mouvement simple ne sont pas donc pas traités comme des robots.

## Canevas d'évaluation

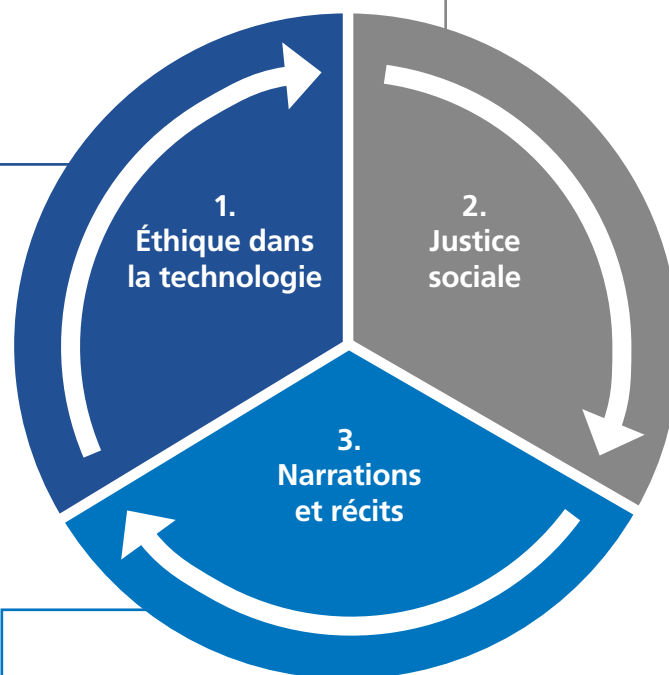
Ce canevas peut être utilisé pour structurer les questions éthiques principales à traiter dans un projet de robotique. Il s'adresse aux équipes de développements, aux responsables politiques ou aux citoyennes et citoyens intéressés par les défis technologiques.

### Éthique dans la technologie

- Le robot est-il autonome et, si oui, dans quel sens? Qui prend la responsabilité en cas de problèmes?
- Le robot agit-il dans un contexte où il doit respecter certaines règles morales? Qui a déterminé quelles sont ces règles?
- Si vous deviez imaginer un scénario catastrophe où l'utilisation du robot est détournée, à quoi ressemblerait-il?
- L'apparence extérieure du robot et son système d'interactions pourraient-ils tromper l'utilisateur-trice ou renforcer certains clichés?

### Justice sociale

- Le robot modifie-t-il les interactions entre humains, par exemple une dynamique d'équipe?
- Le robot déploie-t-il un impact dans un contexte où l'interaction humaine est essentielle?
- Les humains qui vont interagir avec le robot ont-ils été adéquatement consultés et formés? Le cas échéant, un cahier des charges alternatifs a-t-il été développé?

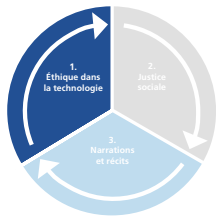


### Narrations et récits

- L'utilisation du robot et la communication déployée renforcent-elles un récit de compétition entre robots et humains?
- Le robot est-il présenté comme un presque-humain ou au contraire comme un outil au service de l'humain?

## Un outil de navigation pour l'éthique de la robotique

Les robots soulèvent de nombreuses questions éthiques. L'éthique est ici définie de manière large comme une réflexion systématique sur les valeurs et sur leur implémentation dans des systèmes socio-techniques, couplant à la fois une composante technique et une composante contextuelle. En mettant en évidence trois grandes thématiques, l'outil de navigation va permettre d'identifier les principales questions éthiques autour des robots. Il doit être utilisé comme une visite guidée pour qui veut faire découvrir le champ de l'éthique de la robotique. Dans une perspective d'innovation responsable, cet outil est particulièrement pertinent pour les équipes impliquées dans le design et la création d'outils technologiques. Les trois thématiques forment un tout et l'outil devrait s'utiliser de manière itérative.



### Éthique dans la technologie

La première thématique traite des questions éthiques dans la technologie. Durant le processus de création, de design et de fabrication du robot, de nombreuses questions s'imposent à ceux qui créent le robot. Trois grandes questions sortent du lot: l'autonomie du robot, le robot moral, et le design du robot.

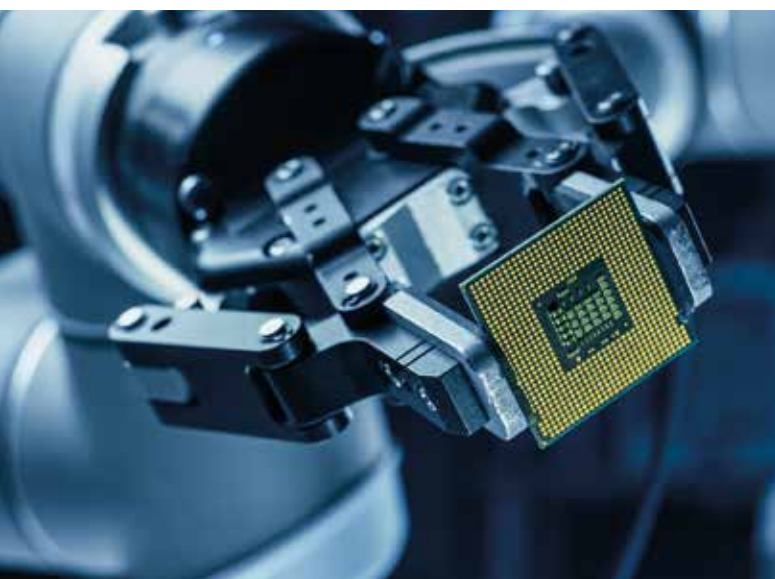
**Première question:** un robot est-il autonome? L'autonomie du robot n'est pas définie de manière similaire par les éthicien-ne-s et les ingénieur-e-s. Les normes ISO liées à la robotique nous fournissent un bon exemple de l'usage «ingénieur» du concept d'autonomie. L'autonomie y est définie comme «la capacité d'exécuter des tâches prévues à partir de l'état courant et des détections, sans intervention humaine». L'autonomie se définit donc comme la capacité à déterminer les moyens pour réaliser un objectif fixé d'avance. Les améliorations en IA vont augmenter cette autonomie car le robot sera toujours meilleur dans le choix et la mise en

œuvre des moyens utilisés. Un défi majeur consiste à cadrer le robot en lui donnant des informations claires sur les règles de comportement, limitant ainsi sa capacité à choisir n'importe quelle stratégie pour atteindre un but.

Les éthicien-ne-s ajoutent une deuxième dimension au terme d'autonomie: la capacité de fixer ses objectifs et de les réviser. Le robot n'a pas cette capacité, il réalise ce que d'autres ont choisi pour lui. Cette autonomie augmentée est souvent définie en lien avec la conscience de soi. La conscience que nous avons de notre existence, de nos désirs et de nos valeurs rend possible l'autonomie. Le robot n'a pas conscience d'exister en tant que robot.

Cette distinction entre deux niveaux d'autonomie joue un rôle crucial pour la question de la responsabilité. Si le robot n'est pas autonome, il ne peut pas être responsable. Le robot est un agent, il peut être causalement responsable de certains événements, mais il n'est pas un agent moral. Une personne humaine doit toujours être responsable. Les différents acteurs qui inventent, conçoivent, produisent et finalement utilisent un robot peuvent avoir différentes responsabilités. Le concepteur / La conceptrice fait un certain choix de base sur ce que le robot peut réaliser, le producteur du robot peut commettre des erreurs de fabrication, et l'utilisateur-trice peut faire un usage erroné de son achat. Les robots dotés de systèmes d'IA avancés capables d'améliorer leur comportement selon leurs expériences passées représentent un défi majeur pour établir ces responsabilités. Il est difficile de reconstruire l'évolution d'un de ces robots pour expliquer son comportement actuel. Néanmoins, le débat sur la responsabilité peut être mené avec les catégories usuelles. Tant que les choix des concepteurs-trices prédéterminent les décisions du robot – même les décisions autonomes – le robot ne peut se qualifier comme un agent moral et il ne peut pas être responsable.

**Deuxième question:** un robot peut-il être agir de manière morale? La question de l'autonomie revient ici sous une forme différente. Le robot peut agir de manière morale si on le programme de cette manière. Mais il ne comprend pas ce qu'il fait au sens d'un humain. Il met en œuvre le code de programmation utilisé. Le défi est donc triple pour les équipes de programmation: quelle



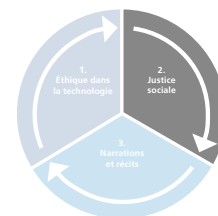


grande approche de l'éthique choisir, par exemple entre une éthique des conséquences et une éthique des devoirs? Sur cette base, comment définir les règles à programmer qui font consensus? Cette discussion ne peut être que contextuelle, distinguant par exemple entre les robots de soins et les armes autonomes. Et finalement, comment assurer que le robot soit techniquement à même de respecter ces règles? Si le robot doit identifier les conséquences pertinentes, les évaluer puis agir en conséquence, comment être sûr qu'il prenne en compte les bonnes informations? Ce défi est crucial et pour l'instant irrésolu.

**Troisième question:** comment faut-il designer un robot? Concevoir un robot, c'est lui choisir une forme et un habillage extérieur, mais également une voix et une interface avec ses utilisateurs-trices. Cette nécessité de choisir l'enveloppe physique du robot distingue la machine des entités numériques. Ce défi de design touche à la confiance que les humains vont témoigner au robot. Les humains souhaitent interagir avec des entités qu'ils (re)connaissent et qu'ils estiment cohérente dans un contexte spécifique. Un sentiment d'étrangeté se développe parfois face à une combinaison inhabituelle d'éléments perçus comme étant humains (formes, textures, comportements) et d'éléments associés au monde des robots. Dans ce contexte, le design d'un robot pose une question de manipulation. Est-il souhaitable de produire un robot qui ressemble à un humain ou à un animal dans le but de tromper ses utilisateurs-trices? Comme dans nos interactions numériques – par exemple avec un chatbot – il semble nécessaire d'avoir une information claire sur qui participe à une interaction. Outre ce défi de la manipulation, le design d'un robot soulève également des questions de représentation. Le choix de certaines caractéristiques physiques du robot projette une certaine image dans le monde. Ainsi la plupart des robots sexuels exhiberont une forte poitrine, des hanches étroites et un visage enfantin. Ces choix de design peuvent renforcer certains clichés. La robotique peut néanmoins nous permettre d'aborder certaines questions inaccessibles aux humains: un robot doit-il avoir un genre reconnaissable ou devrait-on privi-

légier un robot aussi neutre que possible? Pour certaines voix critiques, la robotique pourrait être une opportunité d'expérimenter un rapport plus fluide aux genres et moins dépendante de la binarité masculin-féminin qui imprègne la vie sociale.

Ces trois questions rappellent l'une des leçons les plus importantes de l'éthique des technologies: l'innovation technologique n'est jamais neutre. Elle est conçue, développée et produite avec l'ambition explicite ou implicite de répondre à des objectifs et des valeurs spécifiques. Tous ces choix sont considérés comme des questions d'éthique, car ils mettent en jeu certaines valeurs et engageant, de manière générale, la responsabilité des équipes de développement et de production.



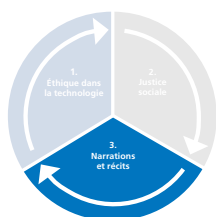
### Justice sociale

La deuxième thématique rassemble les questions de justice sociale liées à l'utilisation des robots. L'analyse porte alors sur l'impact du robot, ou de la robotique en général, sur la vie en société. Les questions générales portent sur l'impact des robots sur les relations humaines et sur le marché du travail. Une approche contextuelle s'intéresse plus précisément aux différents domaines d'utilisation des robots, notamment le soin, l'éducation, ou encore la sécurité.

L'impact sur le marché du travail est traité comme l'une des conséquences sociales les plus importantes de la robotique. Côté pile, les robots libèrent les humains de tâches dangereuses, répétitives ou stupides. Côté face, les mêmes robots sont dénoncés comme des destructeurs de places de travail. Les modèles développés par

les économistes pour chiffrer ce mouvement mettent en évidence plusieurs dimensions. D'une part, le type de domaine économique détermine la vulnérabilité des emplois face à la robotisation. Les secteurs qui reposent sur des processus planifiables, répétitifs et avec peu d'interactions humaines sont les plus à même d'être touchés. D'autre part, les impacts de la robotisation ne touchent pas tous les travailleurs-euses de la même manière. Les emplois les plus menacés par les robots – à la différence de l'IA – concernent les personnes les moins qualifiées travaillant dans des secteurs industriels. Les personnes qui vont devoir intégrer des robots collaboratifs dans leur quotidien vont, quant à elles, faire évoluer leur cahier des charges. L'impact du robot va plus loin que l'interaction humain-robot. L'intégration d'un robot va modifier les dynamiques des relations entre humains, par exemple au sein d'une équipe. Par les exigences de structuration et de mesure qu'il porte en lui, le robot contribue à normer les comportements. En effet, le robot ne sait pas improviser. Il faut donc le cadrer pour qu'il fonctionne et ce cadrage produit en retour un effet sur le personnel et sur l'institution.

Pour une réflexion de justice sociale, la grande question qui se joue en arrière-plan concerne la relation entre emploi, travail, moyens de subsistance et reconnaissance sociale. Trois éléments peuvent d'ores et déjà être mis en œuvre (et le sont déjà partiellement). Premièrement, il nous faut adapter la manière dont nous nous formons tout au long de notre vie. Deuxièmement, il importe de préparer une réponse de politique sociale aux changements en cours sur le marché du travail. Si un domaine économique est touché par un changement technologique rapide et profond, la collectivité doit pouvoir amortir le choc pour les personnes directement concernées. L'idée de transition responsable est cruciale: les changements technologiques doivent être abordés comme des changements permanents, et non dans une perspective de sauvetage une fois la catastrophe en cours. Troisièmement, nous devons mener une discussion sur les domaines qui devraient échapper à la robotisation. Cette approche vise à identifier certaines situations de vie où un contact avec un autre humain doit être sauvé.



## Narrations et récits

La troisième thématique traite des grands récits de société autour de la robotique. Ces grands récits représentent à la fois le cadre général de nos débats sur les robots et une histoire collective que nous entretenons. En matière de robotique, trois récits structurent une large partie des débats de société. L'objectif est ici de faire apparaître la dimension éthique de ces récits et de l'imaginaire qu'ils transportent. Le langage n'est pas neutre et il porte en lui une certaine vision de la réalité. De plus, certains acteurs et certaines actrices ont un intérêt économique ou politique à présenter la réalité dans une perspective spécifique.

Le premier récit porte sur la relation générale entre les robots et les humains. Celle-ci oscille entre une narration de compétition et une narration de coopération. Dans la science-fiction, la compétition finit tôt ou tard par une situation où les robots menacent de contrôler et soumettre les humains. À l'opposé, nous pouvons considérer les relations robots-humains comme des relations de coopération. Le robot est au service des humains, il est comparable à un outil perfectionné que les humains utilisent

pour accomplir leurs objectifs. Le choix d'un récit sur les relations humain-robot modifie en profondeur la manière dont le robot lui-même est appréhendé (comme chose plus ou moins complexe ou comme personne d'un nouveau type).

Ce premier grand récit entre compétition et coopération amène ainsi naturellement vers le second grand récit sur les spécificités humaines. Le jeu d'échecs et le go ont longtemps été présentés comme l'apanage de l'humain, mêlant créativité, ingéniosité et sens tactique. Aujourd'hui, l'humain a perdu. Devons-nous continuer à chercher ce qui distingue l'humain ou au contraire accepter l'idée que l'humain est un être vivant comme les autres? La discussion sur les différences entre humains et animaux, ainsi que leur place générale dans la nature, revient ici par la bande.

Cette quête sur les spécificités humaines nous emmène vers un troisième grand récit portant sur le sens de la vie humaine. Les robots sont souvent présentés comme l'apogée d'une société travailleuse et tout entière orientée vers la productivité. Ils besognent sans relâche. La robotisation interroge profondément la place du travail rémunéré (l'emploi) comme vecteur d'identité personnelle. Si demain nos petits-enfants travaillent 4 heures par semaine dans une société largement robotisée, quel sera le sens de leur vie? Quels récits pouvons-nous développer pour accompagner cette société? La question du sens de la vie revient ici par la grande porte.



## L'Académie suisse des sciences techniques (SATW) soutient des projets dans le domaine de l'intelligence artificielle

Deux évolutions majeures modifient notre relation avec les robots: l'interaction croissante des robots avec le grand public et l'amélioration de l'intelligence artificielle intégrée dans les robots. Ces changements profonds nous confrontent à une série de défis éthiques que nous devons résoudre en tant qu'individus et en tant que société. Il s'agit de savoir avec quels types de robots nous voulons interagir, comment ils sont programmés et comment nous préparons la société aux interactions humain-robot afin d'obtenir une meilleure qualité de vie.

Ces questions sont d'une extrême importance. Pourtant, le débat n'a souvent pas lieu et lorsqu'il a lieu, il est mené avec des clichés et des scénarios extrêmes. La situation de la formation des futur·e·s ingénieur·e·s n'est pas non plus satisfaisante. Le projet «Éthique en robotique» vise à relever l'un de ces défis en fournissant au grand public les outils nécessaires à un débat de qualité.

La SATW a soutenu une équipe de chercheur·e·s de l'EPFL, dirigée par le professeur Francesco Mondada, membre de la SATW, dans la planification, la mise en œuvre et la promotion du projet pilote présenté dans cette fiche d'information.

Dans le domaine de l'intelligence artificielle (IA), la SATW dispose d'une plateforme thématique et d'un programme prioritaire, organise régulièrement des discussions et édite des publications – toujours en collaboration avec des expert·e·s et des organisations partenaires. La SATW est également très active dans le domaine des données, car l'IA ne fonctionne pas sans données. Au printemps 2021, la SATW a fondé, avec la Direction du droit international public du *DFAE*, l'*Office fédéral de la communication* et la *Swiss Data Alliance*, le «Réseau d'autodétermination numérique» national pour permettre aux citoyen·ne·s, aux entreprises et aux institutions publiques d'utiliser l'économie des données sur la base des valeurs fondamentales de la démocratie libérale.

### Impressum

**Auteur:** Johan Rochel (EPFL, UZH, Ethix)

**Chef de projet:** Francesco Mondada (EPFL)

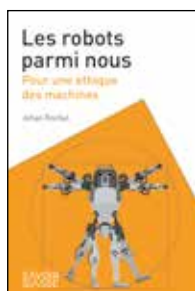
**Équipe de projet:** Johan Rochel (EPFL, UZH, Ethix), Alexandre Luyet (SATW), Adriana Cantaluppi (SATW)

**Rédaction:** Esther Lombardini

**Traduction:** weiss traductions genossenschaft

**Graphisme:** Andy Braun

Mai 2022



Disponible auprès de la maison EPFL Press:

*Les robots parmi nous: pour une éthique des machines*

**satw** it's all about  
technology

Académie suisse des sciences techniques SATW  
St. Annagasse 18 | 8001 Zürich | 044 226 50 11 | info@satw.ch | www.satw.ch

Photos: Adobe Stock

