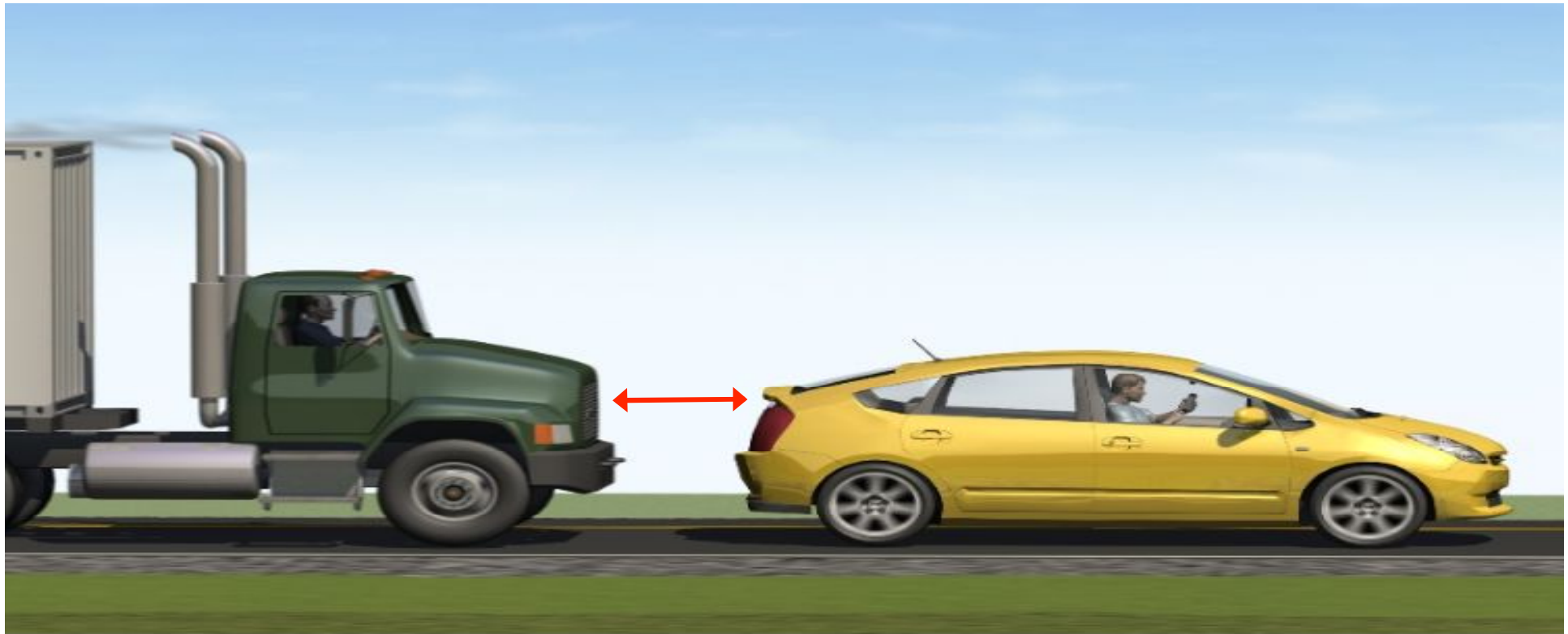


Dilemmes éthiques



Situation prévisible



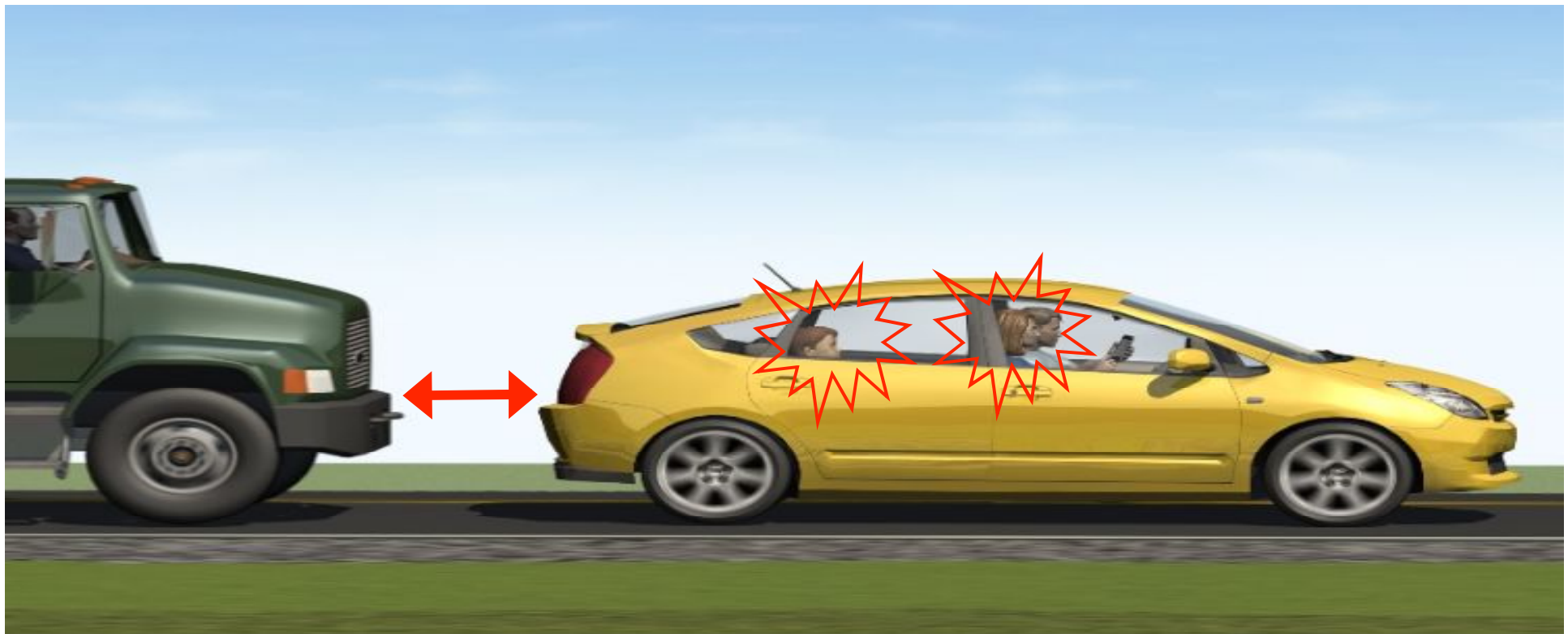
Obstacle imprévu



Événement imprévu



Influence du contexte



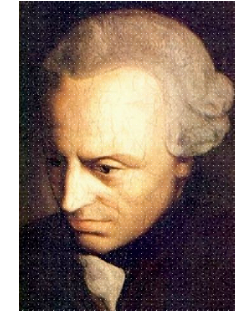
Ethique de la machine

- Le système respecte des valeurs morales dans ses décisions
- Les théories éthiques adoptées déterminent les valeurs respectées
- Difficulté: Evaluation de situation

Théories éthiques

- Théorie Déontique: respect en toute circonstance d'un impératif moral
- Théorie utilitariste: choix minimisant les conséquences négatives
- Théorie casuistique: détermination au cas par cas
- Théorie Rawlesienne: la justice comme fondement de l'éthique

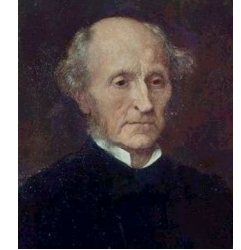
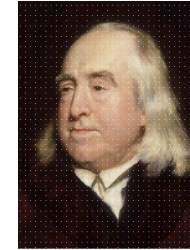
Ethique Déontique



Immanuel Kant (1797)

- Définir un **impératif moral**, valide en toutes circonstances et situations
- Le non-respect de l'impératif moral est non éthique et rend l'agent de toute conséquence
- Le "devoir" est plus important, plus moral, que le "bien"
- Exemples:
 - "Ne jamais tuer"
 - "Toujours dire la vérité "
 - "Tous les humains sont égaux"
 - "Toujours protéger les passagers de la voiture"
 - "Toujours sacrifier les passagers de la voiture"

Conséquentialisme



Jeremy Bentham (1789), John Stuart Mill (1861)

- Utilitarisme: “Le plus grand bien pour le plus grand nombre”
- Seules les conséquences comptent dans un choix moral
- Les valeurs déterminent le choix de l’action
- Mais une action *délibérée* peut être problématique (doctrine du double effet de l’action et du préjudice – Thomas d’Aquin)
- Exemples:
 - Tuer une personne plutôt que cinq
 - Tuer un adulte plutôt qu’un enfant

Casuistique et éthique de situation

- Ethique pratique
- Souplesse et pragmatisme
- Casuistique:
 - Règles générales, mais décisions selon le cas.
 - Taxonomies de cas pour élaborer des indications générales
 - Exemple: Bioéthique
- Ethique de situation (Joseph Fletcher)
 - relativisme
 - Les décisions dépendent du contexte et de la situation précise.

Recommandations sur l'Autonomie

Autonomie et capacités décisionnelles 1/4

- **[AUT-1] Reprises en main**

Le chercheur doit se poser la question des **reprises en main** que l'opérateur ou l'utilisateur peut effectuer (au détriment du robot) et que la machine peut effectuer (au détriment de l'humain), des circonstances qui les permettent ou les rendent obligatoires. Il doit également étudier la possibilité ou non laissée à l'humain de « débrayer » les fonctions autonomes du robot.

- **[AUT-2] Décisions à l'insu de l'opérateur**

Le chercheur doit faire en sorte que les **décisions du robot ne soient pas prises à l'insu de l'opérateur** afin de ne pas créer de ruptures dans sa compréhension de la situation (c'est-à-dire afin que l'opérateur ne croie pas que le robot est dans un certain état alors qu'il est dans un autre état).

Autonomie et capacités décisionnelles 2/4

- **[AUT-3] Influences sur le comportement de l'opérateur**

Le chercheur doit être conscient des phénomènes de **biais de confiance**, c'est-à-dire la tendance de l'opérateur à s'en remettre aux décisions du robot, et de **distanciation morale** (« Moral Buffer ») de l'opérateur par rapport aux actions du robot.

- **[AUT-4] Limites des programmes**

Le chercheur doit être attentif à **évaluer** les programmes de perception, d'interprétation et de prise de décision et à en **explicitier les limites**. En particulier, les programmes qui visent à conférer une **conduite morale** au robot sont soumis à de telles limites.

Autonomie et capacités décisionnelles 3/4

- **[AUT-5] Caractérisation d'une situation**

En ce qui concerne les logiciels d'interprétation du robot, le chercheur doit évaluer jusqu'à quel point ceux-ci peuvent caractériser correctement une **situation** et discriminer entre plusieurs situations qui semblent proches, surtout si la décision d'action prise par l'opérateur ou par le robot lui-même est fondée uniquement sur cette caractérisation. Il faut en particulier évaluer comment les **incertitudes** sont prises en compte.

- **[AUT-6] Prévisibilité du système humain-robot**

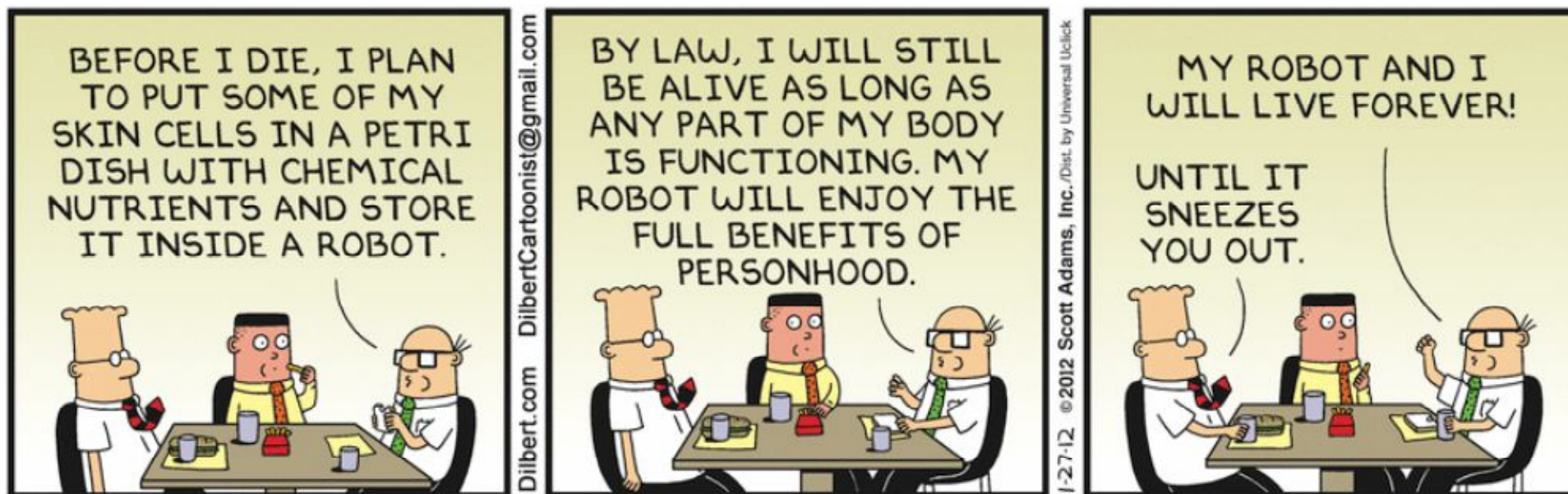
De manière plus globale, le chercheur doit analyser la **prévisibilité du système humain-robot** considéré dans son ensemble, en prenant en compte les **incertitudes d'interprétation et d'action**, ainsi que les **défaillances** possibles du robot et celles de l'opérateur, et analyser l'ensemble des états atteignables par ce système.

Autonomie et capacités décisionnelles 4/4

- **[AUT-7] Traçage et explications**

Le chercheur doit intégrer des **outils de traçage** dès la conception du robot. Ces outils doivent permettre d'**élaborer des explications**, même limitées, à plusieurs niveaux selon qu'elles s'adressent à des experts de la robotique, à des opérateurs ou à des utilisateurs.

Imitation du vivant



Préconisations 2: Bio-mimétisme et interactions affectives et sociales



AIST



Ricky Ma (Hong Kong)



Osaka U.

L'imitation du vivant et l'interaction affective et sociale avec les humains 1/3

- **[IVI-1] Utilité au regard des finalités**

Le chercheur étudiera, au regard des fonctions utiles du robot, la **pertinence et la nécessité de susciter des émotions** et la pertinence et la nécessité de l'aspect ou du **comportement biomimétiques**, notamment dans les cas de forte ressemblance visuelle ou comportementale entre un robot et un être vivant. Dans les cas où l'apparence ou la voix humaines sont imitées, le chercheur s'interrogera sur les effets que pourrait avoir cette imitation, y compris hors des usages pour lesquels le robot est conçu.

L'imitation du vivant et l'interaction affective et sociale avec les humains 2/3

- [IVI-2] **Frontière vivant - artefact**

Si une **ressemblance quasi parfaite** est visée, le chercheur doit avoir conscience que la démarche biomimétique peut brouiller la frontière entre un être vivant et un artefact. Le chercheur consultera sur ce brouillage le **comité opérationnel d'éthique de son établissement**.

- [IVI-3] **Étude des effets**

Pour les projets de recherche qui ont trait au développement de la **robotique affective**, le chercheur s'interrogera sur les répercussions éventuelles de son travail sur les **capacités de socialisation de l'utilisateur**.

- [IVI-4] **Interaction enfant-robot**

Pour les projets qui mettent en présence des enfants et des robots, le chercheur doit se poser la question de l'**impact** de l'interaction enfant-robot **sur le développement des capacités émotionnelles de l'enfant**, tout particulièrement dans la petite enfance.

L'imitation du vivant et l'interaction affective et sociale avec les humains 3/3

- [IVI-5] **Évaluation**

Pour les projets de recherche relatifs à des robots susceptibles d'avoir des effets sur l'**affectivité** des utilisateurs et de susciter leur attachement, le chercheur devra élaborer un **protocole de conception et d'évaluation** en veillant à impliquer les compétences multidisciplinaires nécessaires et des utilisateurs potentiels.

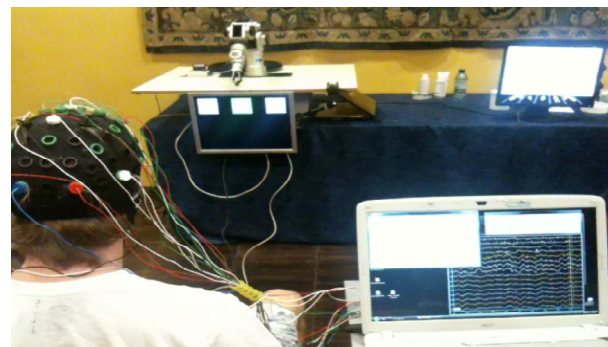
- [IVI-6] **Communication**

Le chercheur doit être prudent dans sa communication sur les capacités émotionnelles des robots et sur l'imitation de la nature et du vivant, notamment parce que **l'expression des émotions, au sens humain, par un robot, est un leurre**, et parce que l'imitation du vivant peut amener, volontairement ou pas, à prêter à l'artefact des caractéristiques du vivant.

Préconisations 3: Réhabilitation et augmentation de l'homme



EPFL



U. Saragozza



RIC Institute



Ekso
Bionics

27/9/2016

R. Chatila

CERNA SS



RB3D

57

Réparation et augmentation de l'humain par la machine 1/3

- [RAH-1] Éthique médicale

Les chercheurs en **robotique réparatrice ou d'assistance** doivent appliquer, en coordination avec les professionnels de santé, les aidants et les patients, les principes d'éthique en usage dans le secteur médical afin d'arbitrer entre les exigences d'efficacité et de sécurité des soins, celles d'autonomie et d'intégrité de la personne et, enfin, de protection de la vie privée. Ce sont là des questions qui relèvent de l'éthique et non uniquement du droit en cela qu'elles demandent à être arbitrées dans chaque cas particulier et qu'elles ne reçoivent pas de réponse générale. Pour en traiter, il faudra prendre avis auprès de **comités opérationnels d'éthique des sciences médicales** et veiller à ce que les compétences technologiques y soient étroitement associées.

Réparation et augmentation de l'humain par la machine 2/3

- **[RAH-2] Autonomie et intégrité de l'individu**

Dans le cas des organes robotisés à vocation **réparatrice**, le chercheur aura le souci de la **préservation de l'autonomie de l'individu** équipé, à savoir de la maîtrise qu'il conservera autant que faire se peut sur ses actions, et de la **conservation de l'intégrité** des fonctions autres que celles concernées par la réparation.

- **[RAH-3] Réversibilité de l'augmentation**

Dans le cas des dispositifs robotisés visant l'**augmentation**, le chercheur veillera à la **réversibilité** de celle-ci : les dispositifs doivent être amovibles sans dommage pour la personne, autrement dit, sans que la personne perde l'usage de ses fonctions initiales.

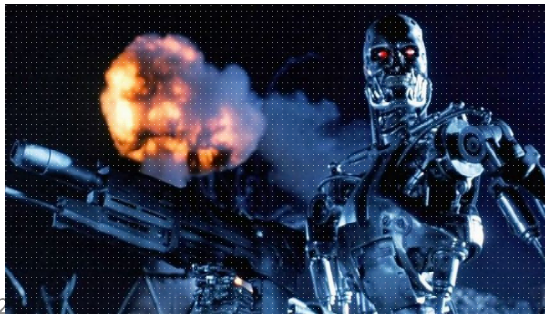
Réparation et augmentation de l'humain par la machine 3/3

- **[RAH-4] Discrimination induite par l'augmentation**

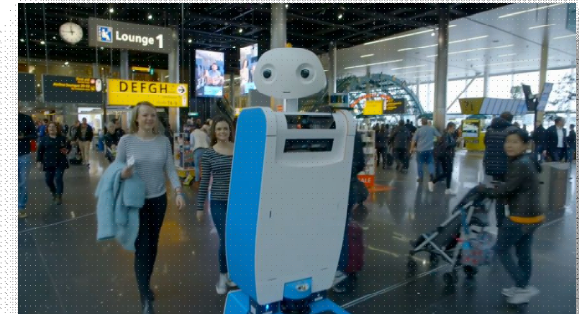
Le chercheur se posera la question de **l'incidence de l'augmentation** des facultés et des capacités humaines induites par les dispositifs qu'il développe **sur le comportement social** de ceux qui en bénéficient ainsi que, symétriquement, de ceux qui n'en bénéficient pas.

Conclusions

- La science et la technologie impactent la société. Les applications seront difficiles à prédire. La robotisation et l'IA peuvent transformer les emplois, en supprimer et en créer.
- Il faut cadrer le développement de la robotique et de l'IA dans une démarche éthique (recherche responsable).
- L'usage des robots et de l'IA amènera à de nouvelles législations (drones, voitures autonomes, armes autonomes, personnes vulnérables,...).



27/9/2

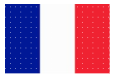


61

Les initiatives ELS sur le plan international



- <http://www.europarl.europa.eu/committees/en/juri/subject-files.html;jsessionid=BC99F48A4A420741A24E04FD184C34C6.node2?id=20150504CDT00301>



- <http://cerna-ethics-allistene.org/>
- [http://www2.assemblee-nationale.fr/14/les-delegations-comite-et-office-parlementaire/office-parlementaire-d-evaluation-des-choix-scientifiques-et-technologiques/\(block\)/24974](http://www2.assemblee-nationale.fr/14/les-delegations-comite-et-office-parlementaire/office-parlementaire-d-evaluation-des-choix-scientifiques-et-technologiques/(block)/24974)



- <https://www.whitehouse.gov/blog/2016/05/03/preparing-future-artificial-intelligence>



- <http://www.bsigroup.com/en-GB/about-bsi/media-centre/press-releases/2016/april/-Standard--highlighting-the-ethical-hazards-of-robots-is-published/>



- <http://www.japantimes.co.jp/news/2016/04/29/national/japan-pushes-basic-ai-rules-g-7-tech-meeting/>



- http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

The Global Initiative for Ethical Considerations in the Design of Autonomous Systems

1. Personal Data
2. Law
3. AGI, ASI
4. Economics Humanitarian
5. Ethical Manufacturing Updated (9.1)
6. Imbuing Ethics
7. Reframing LAWS
8. General Principles

Etudes de cas

- Trois cas d'usage
 - décision autonome,
 - Robot-personnes vulnérables,
 - imitation du vivant
- Pour chaque cas:
 - Identifier les apports positifs/négatifs du cas d'usage
 - Identifier les questions éthiques posées par ce cas
 - Indiquer des arguments pour et contre: faut-il développer les technologies qui rendraient possible cette situation? Faire un schéma avec les arguments « pour » et « contre ».

Présentation

1. Exprimer votre position
2. Donner les arguments en faveur de votre position
3. Donner les arguments opposés à votre position
4. Donner les arguments contrant les arguments opposés, qui vous ont permis de conclure à votre position exprimée en 1.
5. Synthèse

Etude de cas 1: décision autonome

- Pourquoi rendre les voitures autonomes?
- Quelles valeurs doivent respecter les voitures autonomes?
- Quelles préconisations pour la conception de voitures autonomes?

Etude de cas 2: Personnes âgées avec perte de d'autonomie

- Missions: Soins d'hygiène, stimulation cognitive, alimentation, promenade, prélèvements biologiques
- Avantages et inconvénients du robot soignant pour l'accomplissement des tâches?
- Quelles sont les valeurs que doit respecter un robot soignant?
- Faut-il remplacer les soignants humains par des robots ?
- Quelles préconisations pour la conception d'un robot soignant?

3-Imitation du vivant



Etude de cas 3: imitation du vivant

- Faut-il développer des robots androïdes ou exprimant des émotions?
Avantages et inconvénients
- Quelles sont les valeurs que doit respecter un robot androïde ou exprimant des émotions?

Quelles préconisations pour la conception d'un robot androïde ou exprimant des émotions?