



## CHAIRE TRANSFORMATIONS DE L'ACTION PUBLIQUE

SÉMINAIRE « ALGORITHMES, INTELLIGENCE ARTIFICIELLE ET MONDE PUBLIC »

### Séance 3 - Impacts à long terme sur la société et l'action publique : apports des SHS, de la prospective et de l'imaginaire (2 décembre 2020)

*Exposants invités : Joffrey Becker, anthropologue au Collège de France. Daniel Kaplan, cofondateur de l'Université de la Pluralité.*

#### Points clefs de la présentation de Joffrey Becker

##### ***Vivre avec des robots, vivre dans des robots***

Dans la continuité d'un travail de thèse en anthropologie sociale, qui avait pour objet l'étude **des relations qu'entretiennent les humains avec les machines anthropomorphes et les robots humanoïdes**, les recherches de Joffrey Becker s'étendent aujourd'hui à des systèmes qui ne présentent pas forcément de ressemblances avec le corps humain, mais qui en intègrent toutefois les activités. D'une manière générale, les différents cas et études empiriques qu'il a analysés et menés dans le cadre de son travail illustrent **deux approches** principales en matière d'ingénierie sociale, qui peuvent être traduites par le biais de deux questions : à quelles conditions peut-on vivre **avec** des machines ? Et comment vivre **dans** des machines ? Plus précisément, l'enjeu est d'étudier la manière dont se construisent, en robotique et en IA, **des relations d'un nouveau genre**, ainsi que la manière dont ces relations nous placent devant **des choix de société nouveaux et importants**.

##### ***Le développement de la robotique sociale***

Comment la robotique et l'IA se sont intéressées aux relations sociales ? Alors que l'histoire de la robotique a souvent été présentée par les roboticiens en tant qu'aboutissement d'une longue et linéaire évolution des techniques, celle - plus spécifique - de la **robotique sociale** s'est quant à elle construite sur plusieurs éléments. Sur **des mythes** tout d'abord, à savoir la volonté de construire des machines vivantes ou encore celle de totalement domestiquer la vie. Mais aussi sur des principes davantage réalistes tels que **l'imitation de la nature** ou sur **des arguments démographiques** à l'image du vieillissement de la population.

##### ***Avec des robots : à quoi a-t-on affaire ?***

Quelle est la vocation du robot dit social ? **Remplacer l'humain** ou plutôt lui offrir de **nouvelles possibilités d'action** ? À travers des cas issus de ses recherches, Joffrey Becker s'est notamment intéressé aux agents robotiques autonomes et à leur analyse à l'aune de la « **suppléance du lien social** », un concept qui explore les façons dont on peut continuer à avoir des relations sociales ordinaires en s'appuyant sur des objets qui sont capables de les mimer. Qu'ils soient anthropomorphes ou non, on retrouve parmi ces agents robotiques

interactionnels certains conçus pour **accompagner les personnes dépendantes**, tels que la peluche Paro fréquemment utilisée en maison de retraite, d'autres pour être mis en relation avec des humains **dans l'espace public**, à l'instar du projet *Heart Robot*. De par leur nature, ces systèmes illustrent divers enjeux et posent différentes questions propres à la robotique sociale, allant de celle de **l'acceptation ou du rejet des machines anthropomorphes** à celle de **l'expression émotionnelle**, en passant par celle de **la communication sensorielle**.

Un aspect intéressant de ces recherches sur les mécaniques relationnelles est que malgré le fait que l'on puisse imputer la qualité d'« interactant » à des robots, ceux-ci demeurent cependant de **piètres partenaires de communication**. En effet, des problèmes assez évidents surgissent dans la plupart des situations d'interaction. En l'occurrence, il est surtout difficile, en interagissant avec une machine de ce genre, **de savoir exactement à quoi on a affaire**. Loin d'une interaction naturelle, les participants sont dans une sorte de test quasi-permanent où ils essaient de chercher les limites de l'interaction. La question du sens de l'interaction est également importante : bien souvent, on impute à la machine des capacités de perception ou des émotions qu'elle n'a pas.

### **Dans des robots : sommes-nous vraiment autonomes ?**

Ici, d'autres exemples nous viennent notamment de **l'automatisation de la surveillance et de l'assistance des personnes âgées à leur domicile** (e.g. projet *Personally Assisted Living*), ou encore **du milieu urbain** (e.g. ligne 1 du métro parisien). Le terrain n'est alors plus un environnement « naturel », mais plutôt un environnement **augmenté** par l'emploi de systèmes électroniques (e.g. sols intelligents, caméras, interfaces de communication, etc.), censés offrir aux personnes âgées la possibilité de retrouver un peu de leur autonomie perdue. L'idée principale est d'adapter l'environnement des personnes afin d'assurer le bon fonctionnement du dispositif robotique. Cette approche pose la question **des limites de l'espace privé ainsi que celles de l'autonomie elle-même**. Sommes-nous vraiment autonomes lorsque nous sommes enregistrés en permanence ? Rapportées dans la sphère du travail, ces systèmes robotiques en modifient sensiblement la nature.

### **Opacité, substitution et travail**

La robotique sociale, tout comme la communication avec des systèmes dits intelligents, illustre la thématique de **l'opacité**. Le fonctionnement en « boîte noire » de ces systèmes ainsi que la fascination pour ce fonctionnement rendent invisible le travail des concepteurs et surtout leurs représentations, qui sont socialement et culturellement situées. La finalité exacte des systèmes robotiques est ainsi **invisibilisée**. La thématique de **l'apprentissage** est également illustrée, au niveau de l'apprentissage de la machine mais aussi de l'humain qui cherche à apprendre à parler avec elle et à s'y adapter, par exemple en « **robotisant** » son langage. La question - ancienne - de la **substitution**, dans le travail notamment, des hommes par les robots surgit à son tour. L'automatisation, bien qu'elle implique quelques fois une sorte de « simplification » des chaînes humaines de médiation, a une grande tendance à remplacer celles-ci par une **médiation spécifique, machinique**, qui crée un effet **d'interface** et **de la distance** entre les humains qui sont pris dans le système.

### Points clefs de la présentation de Daniel Kaplan

#### **Automatiser et optimiser, un point c'est tout ?**

La réflexion de Daniel Kaplan part d'une question un peu plus restreinte que les discussions généralistes sur la singularité des IA : pourquoi est-ce qu'on pense aussi **majoritairement** - et **presque uniquement** - l'IA comme un moyen **d'automatiser et d'optimiser** ce qu'on fait déjà ? La question se pose car dans les faits, on constate l'invention continue de produits et

de services qui se révèle **inexplicable** par des objectifs d'optimisation et d'automatisation. Certaines IA concernent ainsi la création de désordre, par exemple dans le domaine artistique, et ne perdent pas en complexité pour autant. Des pistes de réponse et d'analyse nous sont données par **la fiction** ainsi que par **les arts**.

### ***L'IA est conçue comme une organisation***

Dans un article paru dans *The Economist* en 2018, le chercheur Jonnie Penn souligne que l'une des choses qui pourrait nous troubler dans l'IA, c'est que celle-ci **pense comme les grandes organisations**. De ce point de vue, le développement actuel de l'IA correspondrait largement à un vieux désir des organisations : **se débarrasser du facteur** imparfait et imprévisible qu'est **l'humain**, qui de surcroît nuit à la bonne marche des affaires. Penn va ainsi chercher les racines de la conception dominante de l'IA dans l'histoire des sciences de l'administration et de la gestion, qui fleurissent dans les années 30, notamment au sein des pays capitalistes. C'est à ce moment-là que l'économiste Herbert Simon se saisit de l'informatique dans le but **d'optimiser et d'automatiser** les décisions dans les organisations. Et c'est ainsi que sont retenues des propriétés particulières de l'IA telles que **l'exactitude, la vitesse ou encore l'efficacité**, tandis que d'autres, à l'instar de l'originalité ou de la fantaisie **sont mises de côté car dérangeantes**. Ce n'est donc pas par hasard que les craintes que l'on exprime face aux IA ressemblent à celles que l'on exprime face aux grandes organisations (e.g. opacité, déshumanisation, expression sans recours du pouvoir).

### ***La fiction - mettre en scène la démythification de l'IA***

Dans l'imaginaire collectif, la représentation la plus partagée de l'IA dans la fiction est celle d'une IA qui devient consciente, et qui, sans doute du fait de son anthropomorphisme, décide qu'il va falloir se débarrasser des humains. Dans plusieurs œuvres néanmoins, à l'instar du roman *Accelerendo* (Stross, 2015) ou du plus connu *2001, l'Odyssée de l'espace* (Kubrick, 1968), la dangerosité de l'IA se situe plutôt dans le fait qu'elle poursuive les objectifs qu'on lui a fixés **sans se soucier des conséquences**, ne tenant ainsi pas compte de tout ce qui ne leur a pas été dit. On peut relire de cette manière l'histoire de *l'Odyssée 2001* mettant en scène le robot HAL 9000, lequel n'a pas décidé que les humains en général le gênaient et qu'il fallait par conséquent s'en débarrasser. Si HAL veut se débarrasser des cosmonautes de son vaisseau, c'est **uniquement car ils compromettent la mission**.

### ***L'art - vers une imagination artificielle ?***

À la question de savoir si on peut attendre autre chose de l'IA, le champ artistique nous fournit une réponse **plutôt favorable**. On peut citer à cet égard les travaux de l'artiste franco-américain Grégory Chatonsky sur *l'imagination artificielle*, où les IA sont nourries de représentations humaines qu'elles recombinaient pour produire des cartes, des territoires ainsi que des objets **qui ne se produiraient pas autrement**, et qui viennent à leur tour nourrir l'imagination des humains. Ou encore le travail du musicien iranien Ash Koosha, qui est devenu le producteur de l'IA Yona en réintervenant sur les créations de cette dernière.

Une autre piste de réponse se situe cette fois-ci du côté du champ de **la prospective**, qui consiste à utiliser le futur comme **une source d'inspiration** pour les actions transformatrices dans le présent. Riel Miller, qui travaille sur la *littérature du futur*, distingue trois types de futurs : **contingents**, qui résultent de forces extérieures, **optimisés**, c'est-à-dire fabriqués à partir de nos connaissances et constitués principalement de morceaux de passé (IA classique), et enfin **nouveaux**, qui se caractérisent plutôt par leurs potentialités que par leurs résultats qui sont quant à eux indéterminables. Pour Daniel Kaplan, c'est justement de cette forme de futur dont se saisit une IA comme celle de Chatonsky et **laisse envisager autre chose que l'automatisation ou l'optimisation**.

Pour aller plus loin ...

- Becker, J. (2015). *Humanoïdes : Expérimentations croisées entre arts et sciences*. Presses universitaires de Paris Ouest.
- Kirman, B., Linehan, C., Lawson, S., & O'Hara, D. (2013). CHI and the future robot enslavement of humankind : A retrospective. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. <https://doi.org/10.1145/2468356.2468740>
- Koosha, A. (2018). *Yona—Art and Law (Written and Performed by Artificial Intelligence)*. <https://www.youtube.com/watch?v=ZzgGgFl9e4s>
- Kubrick, S. (2012). *Hal, open the pod bay doors*. Extrait accessible à partir du lien suivant : [https://www.youtube.com/watch?v=qDrDUmuUBTo&feature=emb\\_title](https://www.youtube.com/watch?v=qDrDUmuUBTo&feature=emb_title)
- McGoran, D. (2008). *Heart Robot Project*. Vidéo accessible à partir du lien suivant : <https://www.youtube.com/watch?v=C7W9wjhiv2o>
- Miller, R. (2018). *Transforming the Future: Anticipation in the 21st Century*. Routledge.
- Penn, J. (2018). AI thinks like a corporation—And that's worrying. *The Economist*. <https://www.economist.com/open-future/2018/11/26/ai-thinks-like-a-corporation-and-thats-worrying>
- Stross, C. (2015). *Accelerando* (J. Bonnefoy, Trad.). PIRANHA.