

14 janvier 2018

*Une traversée de la transformation numérique
Évolution ou révolution?*

Max Dauchet

Professeur émérite d'informatique, Université de Lille

Président de la CERNA





Novembre 2017

Le robot Sophia intervient à l'ONU

Quelques jours après, il est fait citoyen saoudien

Tout autre chose...

...un article anodin...

Learning to Protect Communications with Adversarial Neural Cryptography,
Martín Abadi, David G. Andersen (Google Brain), 2016

Cornell University Library <https://arxiv.org/abs/1610.06918>

... deux réactions

- Jean-Paul Baquiast, un des responsables des télécoms sous de Gaulle, Blog Médiapart, 2 nov. 2016
- Jacques Attali, publié le 07/11/2016 sur son blog

Jean-Paul Baquiast, Blog Médiapart, 2 nov. 2016

Des réseaux neuronaux de Google inventent leurs propres systèmes d'encryptage.

(...) Les chercheurs précisent qu'ils n'ont pas à ce jour bien compris comment les réseaux neuronaux Alice et Bob ont procédé pour produire des textes cryptés.

Il s'agissait manifestement d'une initiative autonome dans laquelle aucun humain n'était intervenu (...)

Jean-Paul Baquiast, Blog Médiapart, 2 nov. 2016

Des réseaux neuronaux de Google inventent leurs propres systèmes d'encryptage.

(...) Les chercheurs précisent qu'ils n'ont pas à ce jour bien compris comment les réseaux neuronaux Alice et Bob ont procédé pour produire des textes cryptés.

Il s'agissait manifestement d'une initiative autonome dans laquelle aucun humain n'était intervenu (...)

Scientifiquement
avéré

Scientifiquement
inepte

Mais que disent-elles ? publié le 07/11/2016 à 11:20

<http://blogs.lexpress.fr/attali/2016/11/07/mais-que-disent-elles/>

(...) les deux auteurs ont (...) développé, à la surprise de tous, des formes de cryptage (...) indéchiffrables pour tous, même pour leurs créateurs.

On sait , depuis longtemps, que ces intelligences peuvent aussi prendre leur autonomie ; qu'elles peuvent d'abord considérer que leur mission est de contrecarrer ceux de nos comportements qui, à leurs yeux, nous nuiraient ; qu'elles peuvent ainsi penser que l'humanité courre à sa perte et décider, pour son bien, de l'arrêter à temps. Mais elles pourront alors aussi agir contre l'humanité, dans leur seul intérêt ; surtout si elles pensent que, à terme, nous aurons intérêt à les détruire : elles peuvent alors nous combattre, en légitime défense.

Là, vient évidemment le pire, qu'annonce cet article, apparemment purement théorique : si les intelligences artificielles réussissent à inventer des langues irréversiblement indéchiffrables par les humains, nous sommes perdus. Le pire des pronostics se réaliseront. L'humanité (...) aurait (...) créé des Intelligences Artificielles immortelles, capables de travailler en secret pour la détruire.

Doit-on arrêter cela ? Sans aucun doute. (...)

Il est urgent qu'une grande conférence planétaire sur l'intelligence artificielle réunisse savants et gouvernants pour décider d'un moratoire sur ces recherches (...)

j@attali.com

Mais que disent-elles ? publié le 07/11/2016 à 11:20

<http://blogs.lexpress.fr/attali/2016/11/07/mais-que-disent-elles/>

(...) les deux auteurs ont (...) développé, à la surprise de tous, des formes de cryptage (...) indéchiffrables pour tous, même pour leurs créateurs.

On sait, depuis longtemps, que ces intelligences peuvent aussi prendre leur autonomie ; qu'elles peuvent d'abord considérer que leur mission est de contrecarrer ceux de nos comportements qui, à leurs yeux, nous nuiraient ; qu'elles peuvent ainsi penser que l'humanité court à sa perte et décider, pour son bien, de l'arrêter à temps. Mais elles pourront alors aussi agir contre l'humanité, dans leur seul intérêt ; surtout si elles pensent que, à terme, nous aurons intérêt à les détruire : elles peuvent alors nous combattre, en légitime défense.

Là, vient évidemment le pire, qu'annonce cet article, apparemment purement théorique : si les intelligences artificielles réussissent à inventer des langues irréversiblement indéchiffrables par les humains, nous sommes perdus. Le pire des pronostics se réaliseront. L'humanité (...) aurait (...) créé des Intelligences Artificielles immortelles, capables de travailler en secret pour la détruire.

Doit-on arrêter cela ? Sans aucun doute. (...)

Il est urgent qu'une grande conférence planétaire sur l'intelligence artificielle réunisse savants et gouvernants pour décider d'un moratoire sur ces recherches (...)

Tissu d'inepties

j@attali.com

Au delà des fantasmes, que craindre/espérer des androïdes et de l'IA?

PLAN

La transformation numérique, une véritable révolution, mais pas celle des fantasmes

Il y a unanimité scientifique pour dire que

- Les élucubrations transhumanistes n'ont aucun fondement
- L'IA ne prendra pas le pouvoir sur l'humain

Que néanmoins

- Le numérique transforme la science
- Le numérique transforme la société
- Le numérique transforme probablement l'humain et l'avenir

Et que finalement

- Il y a bien plus à attendre qu'à craindre du numérique, à condition de vigilance éthique

Mais pourquoi une telle confusion?

Transhumanisme, IA "hostile" ou "bienveillante"

- *La singularité*: la capacité des machines dépasse exponentiellement celle de l'homme.
- *L'intelligence forte* : aussi polyvalente et adaptative que l'humaine (et en mieux)

- Les transhumanismes (Nick Bostrom, Max Tegmark, Kurtweil, Laurent Alexandre)
- Mouvements internationaux, souvent activistes,
 - La "convergence NBIC" (Nano-Bio-Info-Cogn) permet de prendre le relai de l'évolution pour améliorer l'homme, voire le dépasser.
- Vivre 900 ans, voire infiniment, en nous réparant et améliorant sans cesse par de plus en plus de prothèses numériques

Repose sur des hypothèses sans aucun fondement comme

- L' intelligence émerge de "la complexité" indépendamment du substrat
- Les algorithmes "existent" au même titre que les êtres et les choses
- L'esprit est numérisable

Leurs promoteurs sont des futurologues pseudo-philosophes ou pseudo-technologues qui ignorent la réalité des sciences et technologies du numérique

Transhumanisme, IA hostile et réalité

- Ce que l'on sait faire : actionner des commandes simples par la pensée, restaurer des fonctions psychomotrices, fabriquer des prothèses et des exosquelettes ciblant des déficiences
- Ce que l'on saura faire: des IA de plus en plus autonomes pour assurer des tâches (IA faible) de plus en plus complexes
- Ce qui n'est pas envisageable:
 - Télécharger l'esprit
 - Beaucoup de scénarios de SF sur les robots.

Explication intuitive :

- Les explications les plus courtes sont les plus probables

Rasoir d'Ockham (v. 1285 - 9 avril 1347)

« Les multiples ne doivent pas être utilisés sans nécessité »

(Pluralitas non est ponenda sine necessitate)

- Ce principe demeure un cadre de la science
- Il est un théorème dans le monde numérique des objets calculables par programme
- Ce cadre de pensée ne permet pas de découverte : il est presque toujours impossible de calculer la plus courte cause *(qui est dite non calculable)*



Le deep learning qui fait l'actualité est largement basé sur la règle d'Ockham, avec succès parce que le monde qu'il apprend est modulaire et structuré (le Cuius in Propter). Expliquer analytiquement pourquoi est un défi scientifique majeur (Villani)

Le numérique transforme la science

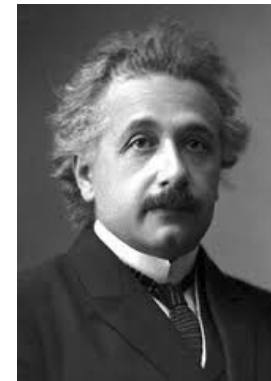
Gérard Berry, Collège de France, *L'Hyperpuissance de l'informatique*, Odile Jacob, 2017

Du point de vue des sciences classiques, le monde a été dominé par trois entités jusqu'à la moitié du XXe siècle : la matière, l'énergie et les ondes. Mais un quatrième larron : l'information, à Sciences et Avenir

$$E = mc^2 \quad (\text{en Joule}) \quad \text{Einstein (1905)}$$

Dans un sens, le big-bang, le LHC.

Dans l'autre, la possibilité d'une bombe...

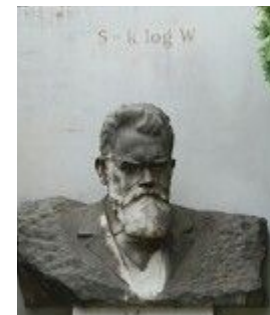


$$S_P = S_B + k \ln 2 K \quad (\text{en Joule/Kelvin}) \quad \text{Landauer (1962), Bennet, Zurek (1988)}$$

S_B Entropie de Boltzmann

K Complexité de Kolmogorov :

longueur d'un plus court programme



k constante de Boltzmann (en 1877, vaut environ $1,380\ 648\ 52 \times 10^{-23}$)

Il n'existe pas de programme permettant de décider si un programme s'arrête (l'arrêt des programmes est dite indécidable, ou incalculable)

Preuve : On considère les programmes qui calculent sur les nombres entiers. On peut les numéroter.

Supposons le problème de l'arrêt soit décidable. Il existe alors un programme $P_n(x)$ qui a pour résultat $1+P_x(x)$ si $P_x(x)$ s'arrête et 0 sinon.

Considérons $P_n(n)$. $P_n(n)$ ne s'arrête pas car sinon on aurait $P_n(n)=P_n(n)+1$. Mais si il ne s'arrête pas, alors il s'arrête avec 0 pour résultat. Donc la supposition est fausse.

==> La publication initiale de Turing
en 1937 fait 37 pages, avec une erreur.



==> **Enormément de problèmes simples sont indécidables, incalculables.**

Cela ne signifie pas qu'un algorithme reste à découvrir, cela signifie¹⁵ qu'il n'existe pas d'algorithme pour résoudre le problème.

Le numérique transforme la société

Exemple: Le bitcoin une monnaie numérique née "de rien" en 2009 par "on ne sait pas trop qui " et qui va "on ne sait où"



- Possible grâce aux sciences et technologies du numérique
 - Au cloud computing (capacité de calcul, de transmission et de stockage)
 - A l'existence de problèmes faciles à poser, à la solution très dure à trouver mais facile à vérifier (complexité algorithmique, cryptographie)
- Principes
 - Pas de régulation : les transactions sont consignées dans une *blockchain*
 - La mise à jour de ce grand livre public peut-être réalisée par n'importe qui pouvant prouver qu'il est capable de résoudre un problème dur
 - Elle est payée en bitcoins en tirant au sort un de ces "mineurs"
- Où va le bitcoin?
 - Il est passé de 1000 à 20000 \$ en 2017. Il commence à être coté bourses.
 - Les bitcoins émis sont de plus en plus rares, des conglomérats se forment pour miner dans une compétition algorithmique de plus en plus dure. Ce minage consomme maintenant autant d'électricité qu'une nation européenne moyenne.

Le simple est toujours faux.

Ce qui ne l'est pas est inutilisable. Paul Valéry

Le numérique transforme probablement l'humain et l'avenir

- Au niveau sociétal (au même titre que la consommation)
 - Au niveau anthropologique
 - La lance pour pêcher crée du lien social pour se protéger d'assassinats
 - *Du mode d'existence des objets techniques*, Gilbert Simondon, 1958
 - Représentation (post-moderne?) que se fait l'homme de lui
- Dieu est mort, l'homme moderne aussi (selon Michel Foucault)
- Au niveau biologique (tdonc transmis par l'inné) coévolution, écologie, épigénomique
 - maîtrise du feu au développement du cerveau
 - La station debout a fait perdre les sabots mais gagner les mains. Selon Michel Serres, le numérique pourrait nous faire perdre en mémoire et gagner en intelligence

Le simple est toujours faux.

Ce qui ne l'est pas est inutilisable. Paul Valéry

Le numérique transforme probablement l'humain et l'avenir

- Au niveau sociétal (au même titre que la consommation)
- Au niveau anthropologique
 - La lance pour pêcher crée du lien social pour se protéger d'assassinats
 - *Du mode d'existence des objets techniques*, Gilbert Simondon, 1958
 - Représentation (post-moderne?) que se fait l'homme de lui

Dieu est mort, l'homme moderne aussi (selon Michel Foucault)
- Au niveau biologique (tdonc transmis par l'inné) coévolution, écologie, épigénomique
 - maîtrise du feu au développement du cerveau
 - La station debout a fait perdre les sabots mais gagner les mains. Selon Michel Serres, le numérique pourrait nous faire perdre en mémoire et gagner en intelligence

Repère

Le malaise face à la ressemblance à l'humain

La vallée de l'étrange a été le thème de l'expo

PERSONA en 2016 au Musée du quai Branly

À PROPOS DE L'EXPOSITION

(...) Ce transfert ou cette confusion qui s'opère

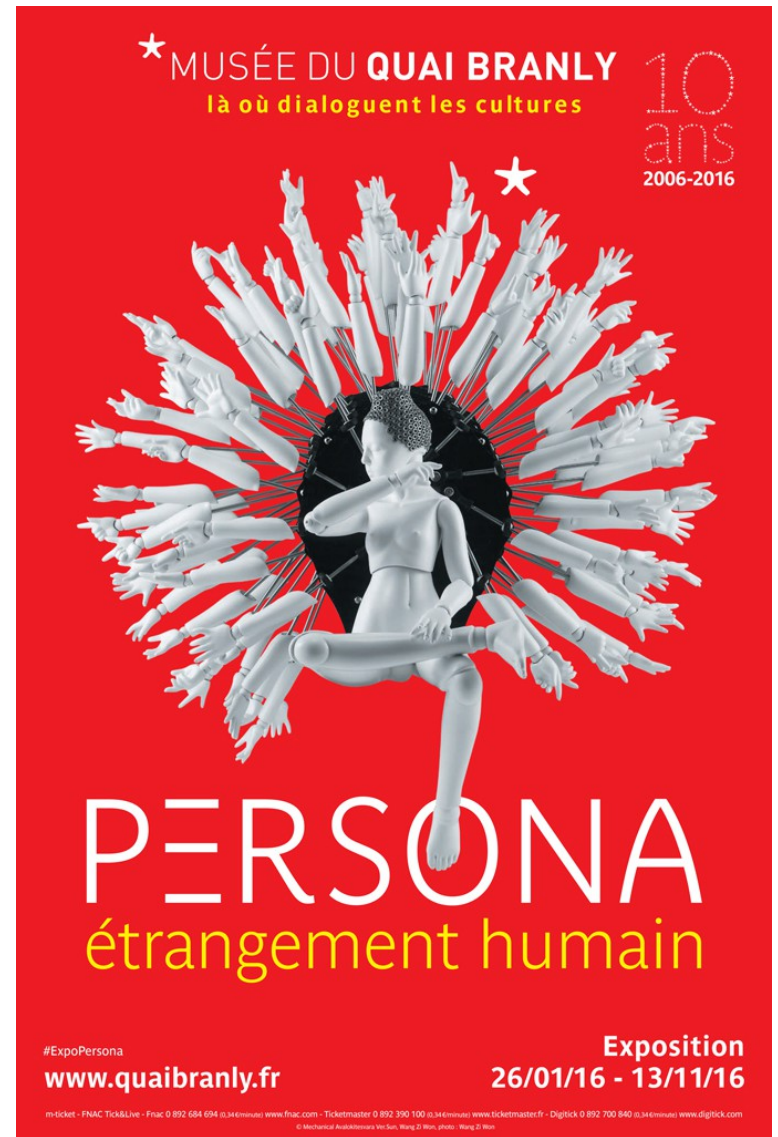
alors entre l'humain et le non-humain,

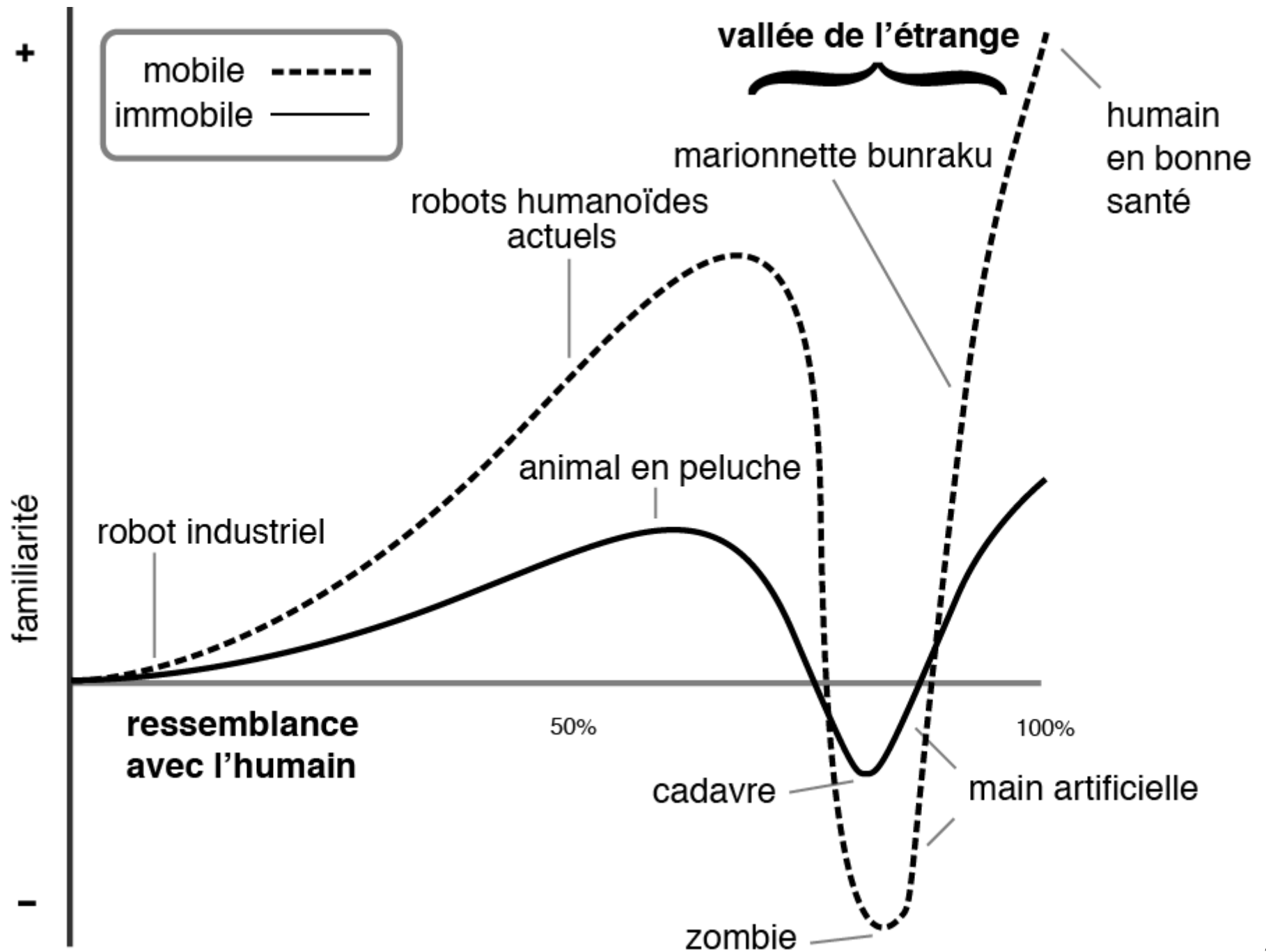
et la relation particulière et personnalisée

qui les lie, dans les cultures les plus variées,

est le vaste sujet de cette exposition

d'anthropologie. Une incursion par la robotique, via l'oeuvre pionnière de Masahiro Mori, permettra de comprendre le rôle de l'anthropomorphisme dans les artefacts les plus divers, et ce qui se joue lorsqu'un robot à l'apparence trop humaine – ou toute autre pièce au réalisme intimidant – croise notre regard, jusqu'à le déstabiliser.





Dans la vallée de l'étrange

Golem



Marionette
bunroku

Frankenstein de Mary Shelley



Geminoids des
Hiroshi Ishiguro Labs

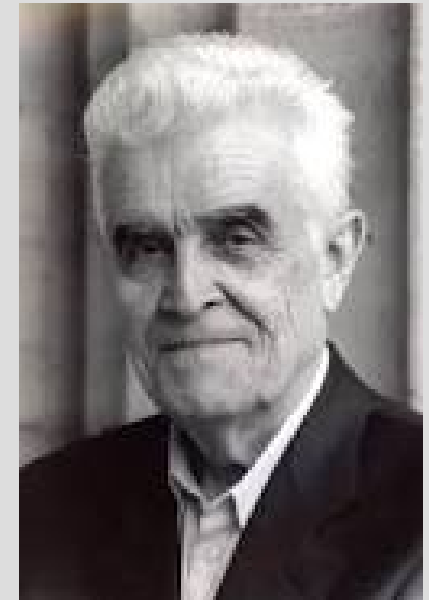


Océanie il y a un siècle



La vallée de l'étrange est l'objet d'hypothèses scientifiques diverses

- Simple artefact
- Psychologique : les mécanismes de classification apprendront à distinguer sans trouble avec l'habitude
- Anthropologiques
- Le *désir mimétique* ne fonctionne plus (René Girard)
 - *La Violence et le Sacré, Le Bouc émissaire*
- Biologique :
 - Les robots n'ont pas de *neurones miroirs*
 - Les neurones miroirs ont été découverts chez le singe dans les années 1990 par Giacomo Rizzolatti, et en 2010 chez l'homme
(une explication de leur fonction est donnée durant l'exposé)



Mais pourquoi une telle confusion? (Opinion)

- *Mal nommer un objet, c'est ajouter au malheur de ce monde.* Albert Camus

On utilise des mots de l'humain pour parler des machines

- "Troisième loi" d'Arthur Clarke " *Toute technologie suffisamment avancée est indiscernable de la magie*"
- Profondeur de la transformation numérique (comparée à l'émergence de l'écriture), difficulté à comprendre globalement les aspects technoscientifiques ET les aspects humains et sociaux.
- Rapidité et caractère massif sans précédent de la diffusion d'usages souvent inattendus
- Facksciences, guerre asymétrique de l'information
- Sciences et progrès: De la science aux technosciences; De la confiance à la défiance

" Dans la vie, rien n'est à craindre, tout est à comprendre. " Marie Curie

Consensus: Il y a bien plus à attendre qu'à craindre du numérique, à condition de vigilance éthique

- Risques de double fracture : accès, compréhension
- Frontière floue entre réparation/amélioration de l'humain
- Transformation de l'espace public: de la place du village au village planétaire
- Réseaux sociaux et "tous les points de vue se valent"
- Les robots et l'emploi: une montée en qualification est certaine, l'impact sur le chômage est inconnu
- Les algorithmes en général
 - Risque de soumettre le modèle social à l'optimisation des algorithmes
 - APB : faut-il simplifier l'expression des vœux pour optimiser les affectations?
 - Humain et "robots" au travail : risque de soumission de l'humain à des normes
 - Optimisation de la société "pour son bien" (paternalisme)

- Souverainetés et souveraineté numérique
 - Etats vs GAFAM. Souveraineté de l'individu sur ses données
- Médecine personnalisée, colloque singulier et aide au diagnostic
- IA
 - Données d'apprentissage non biaisées
 - Des assistants personnels tiennent des propos racistes
 - Classification faciale: gare aux dérives de l'anthropométrie
 - Les bots (confusion entre interlocuteurs robots et humains – le langage humains s'alignera-t-il sur le langage des robots?)
- Loyauté des plates-formes
- Valeurs. Peut-on implanter des valeurs dans des machines
 - Gestion de conflits, Dilemmes éthiques (voiture autonome)
 - Droits de l'homme, conventions de Genève
- Faut-il afficher la distinction bots – humains? Pourquoi faire des androïdes ?
- Méthodologique :
 - Ne pas confondre *La carte et le territoire* (tentation réductionniste)
 - Prédications non explicatives, stats, rôle du hasard, corrélation/déduction
- Etc. etc

Si vous m'avez compris, c'est sans doute que je me suis mal exprimé

Alan Greenspan

Épilogue

Mission Villani sur l'IA

Le numérique, une chance pour la France et l'Europe, le bien être de leurs citoyens et notre compétitivité économique, à condition de croire en nos atouts

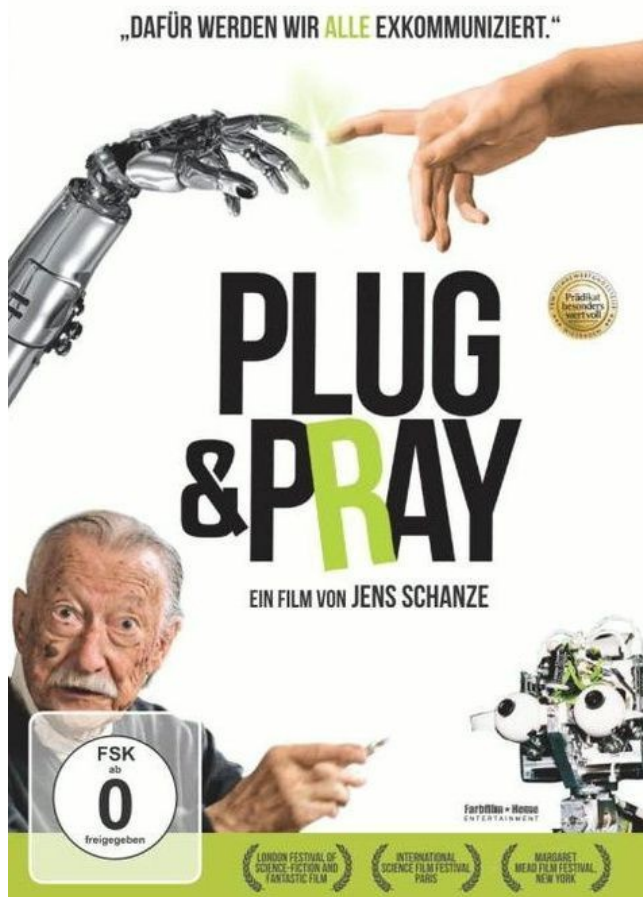
- L'Europe, premier foyer de culture et d'humanisme de l'histoire
- La zone Euro, première force économique
- La France et l'Europe, terre d'excellence scientifique

et de les conjuguer selon un modèle au service de l'humain, garanti par la vigilance éthique et la délibération.

Deux films à voir

2015, EX MACHINA

Comment Ava inverse le test de Turing
et comprend qu'on ne devient humain
qu'à travers ses rapports avec les humains



2010, Plug & Pray

avec dans leur propre rôle

Joseph Weizenbaum Professeur émérite au MIT et créateur en 1963 d'ELIZA, ancêtre des chatbots, qui attend sereinement la mort

Et le futuriste de Google Raymond Kurzweil, transhumaniste en quête d'immortalité