

Apprentissage machine, masses de données et éthique du numérique



Jean-Gabriel Ganascia

Sorbonne Université – Labex OBVIL

Institut Universitaire de France

Président du comité d'éthique du CNRS

4, place Jussieu, 75252 Paris Cedex 05, FRANCE

Jean-Gabriel.Ganascia@lip6.fr

Synoptique

1. **Éthique**: quelques rappels
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes** – HLEG-AI & al.
4. **Épistémologie de l'apprentissage**:
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



Éthique

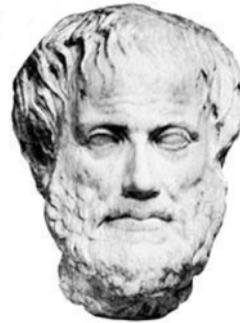


- Étymologie:
 - bas lat. *ethica*;
 - grec *êthikos*, *êthikê*,
de *êthos* « mœurs »
- Comportement de l'homme en société
- Art de diriger la conduite
- Science de la *morale*
 - lat. *moralis* (relatif aux mœurs),
de *mores* → Mœurs
- **Mœurs**: usages des hommes en société — traditions
- *êthos* en grec
 - Mœurs
 - Etable: là où l'on rentre le soir, ce que l'on habite
- Fondement originel de l'éthique:
 - Traditions, usages

Comment se mettre d'accord?

- **Préceptes**

- Religion
- Valeurs/vertus



- **Principes**

- Utilitarisme / conséquentialisme
- Déontisme / impératif catégorique

- **Conviction/Responsabilité *M. Weber***

- Savant, connaissance

- **Délibération *J. Habermas***

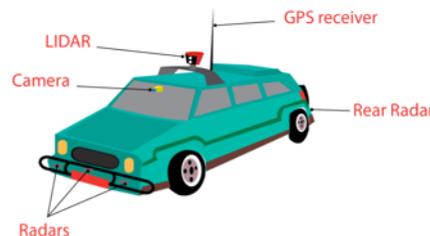
- Complexité du savoir – communauté de savants

Nouvelles questions à l'ère numérique



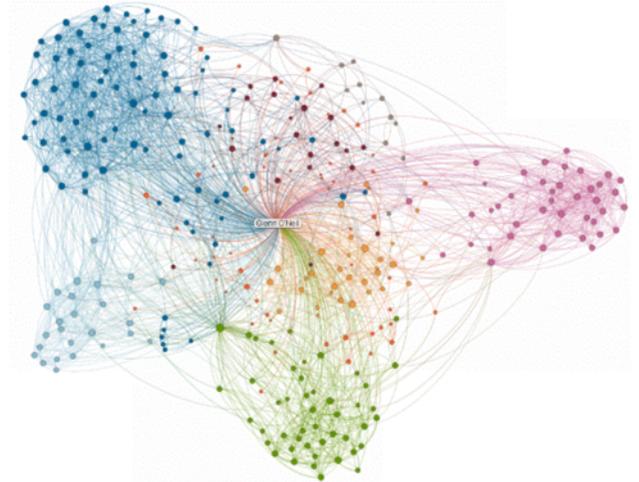
SALA – Systèmes d'Armes Létales Autonomes

« Golems » et systèmes autonomes



Données

- Vie privée,
- Exactitude,
- Biais,
- Propriété,
- Accessibilité



Décision

- Responsabilité,
- Transparence,
- Équité



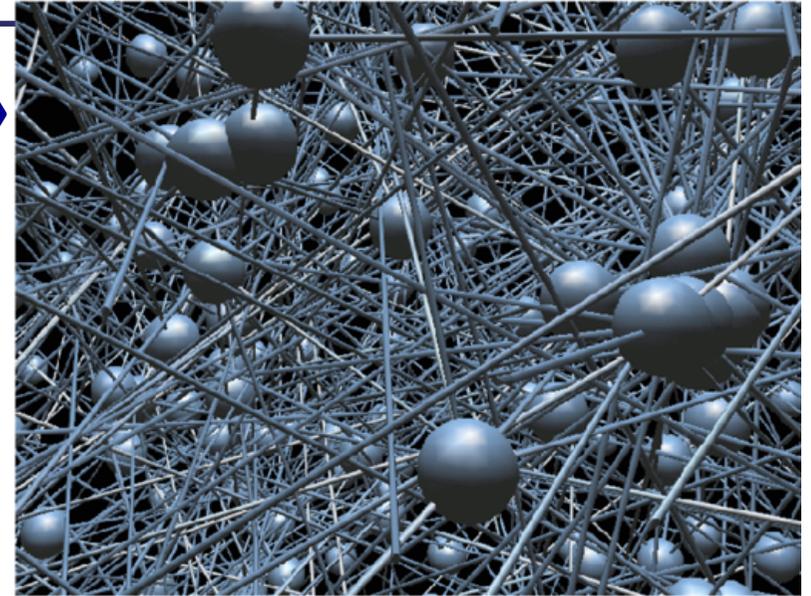
Synoptique

1. **Éthique**: quelques rappels
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes** – HLEG-AI & al.
4. **Épistémologie de l'apprentissage**:
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



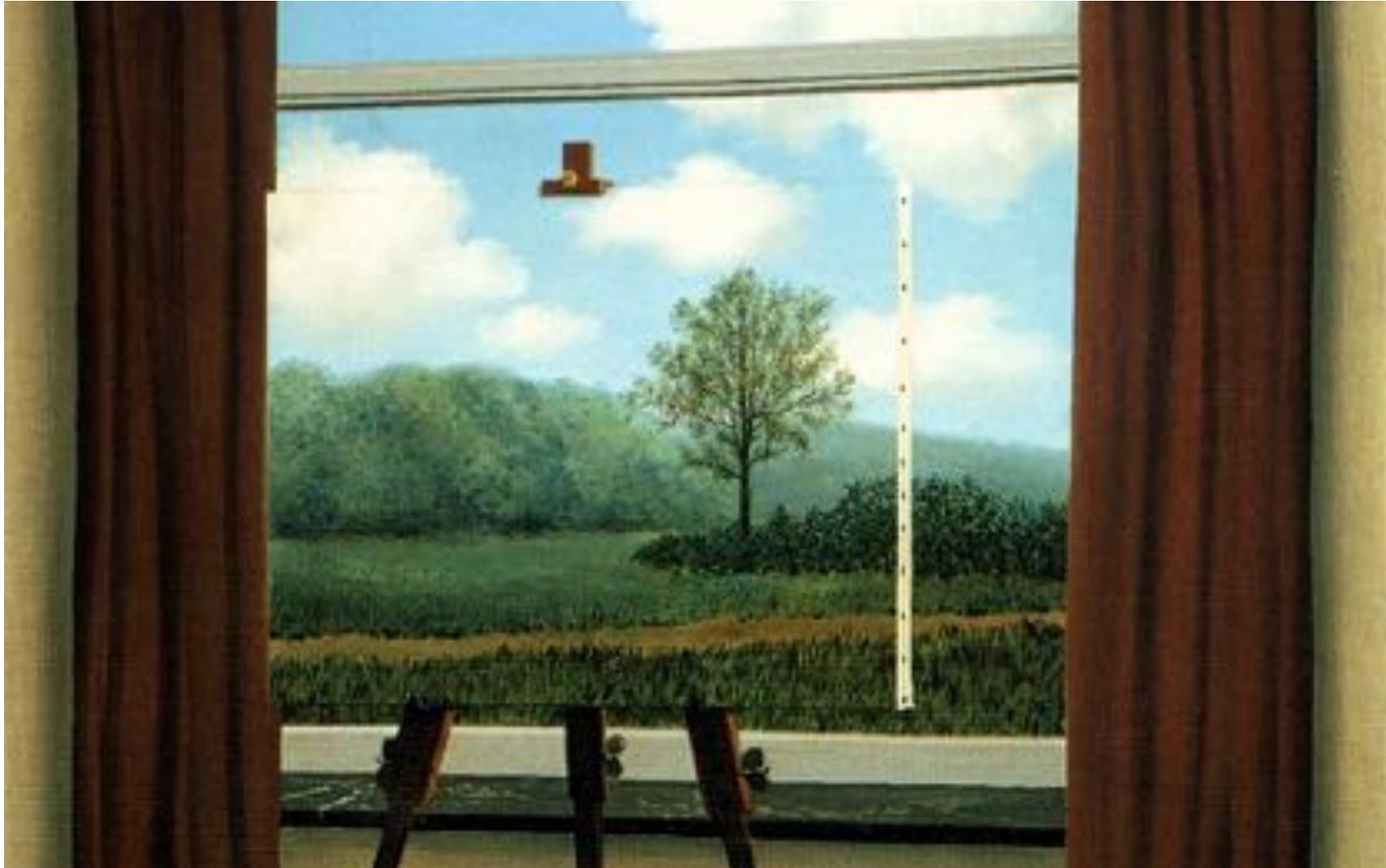
Le monde change! Plus d'étable...

Société « hyperconnecté »

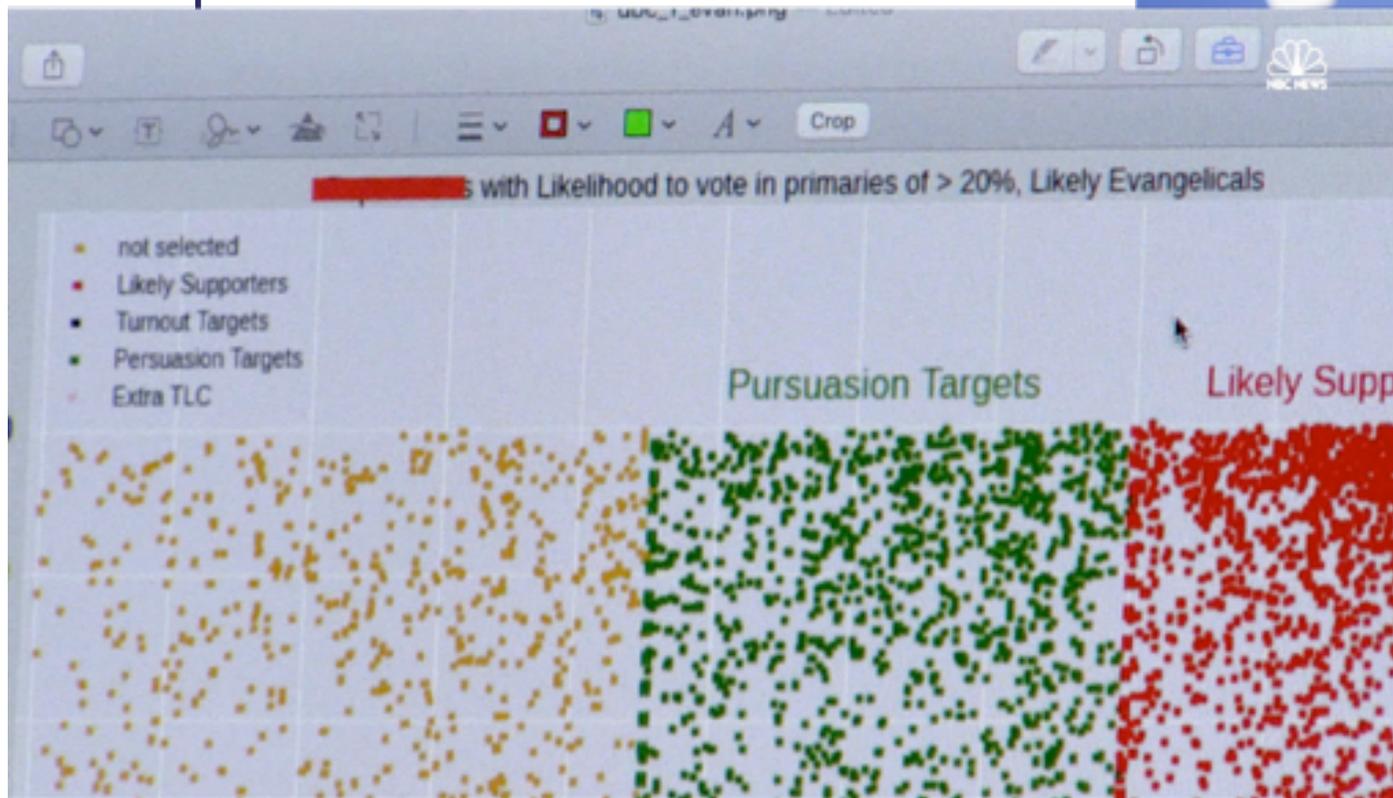


- **Tradition?**
- **Principes?**
 - comment les fonder?
- **Conviction/Responsabilité?**
 - comment faire face à la complexité?
- **Délibération**
 - multitude de cultures, de compétences, ...

La nouvelle condition humaine



Nouvelles vulnérabilités



Nouvelles vulnérabilités: deep fake!



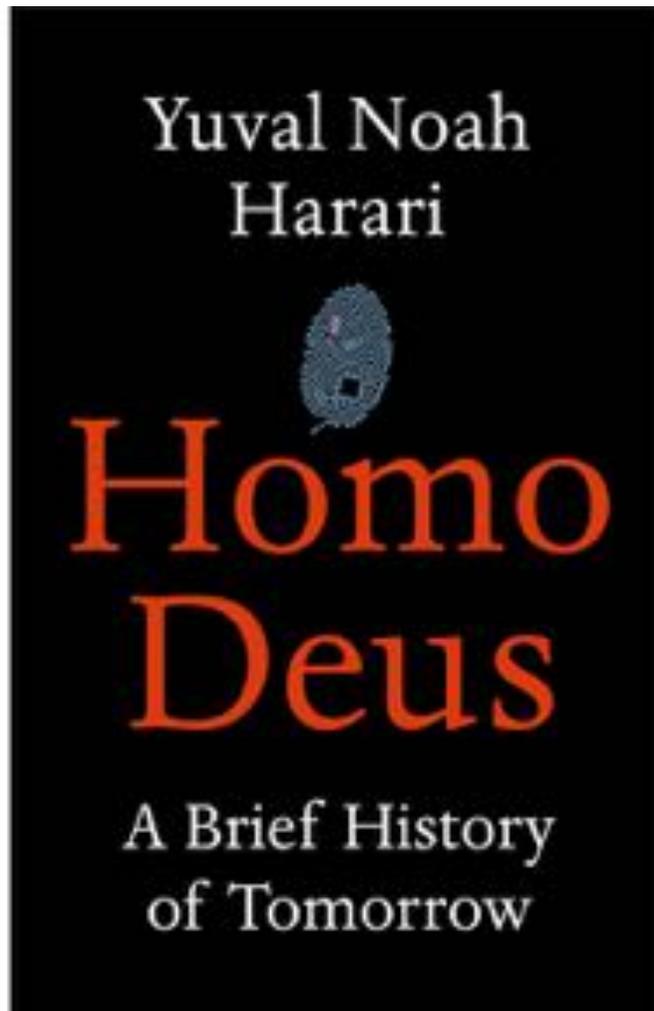
Manipulations
de paroles et
d'images...



Nouvelles vulnérabilités: infox



Homo Deus — Pop Science



- Trois étapes dans l’histoire de l’humanité
 - Religions
 - Humanisme
 - “**Dataisme**”: fin de l’humanisme
- Dépossession du privilège de l’Homme dans la nature par les machines et, surtout, les **données** (*data*)

Synoptique

1. **Éthique: quelques rappels**
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes – HLEG-AI & al.**
4. **Épistémologie de l'apprentissage:**
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



Masses de données - « Big Data »

Data

- du latin *data*, données, observations; faits connus

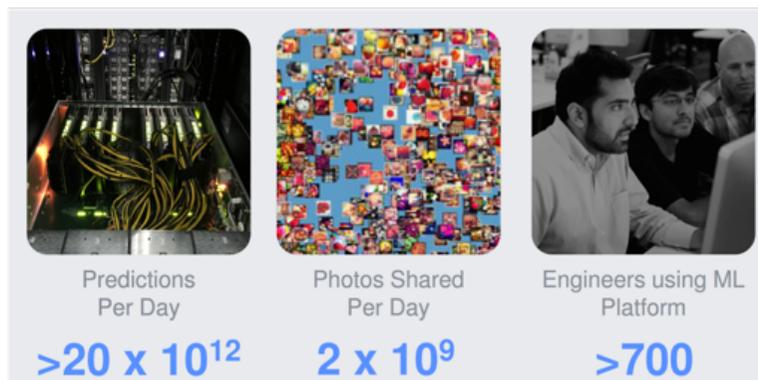
Quantité d'information

- 1 livre, 1 million de caractères = 1Mo
- Catalogue des livres et imprimés de la BNF:
14 millions d'ouvrages = **14To**
1 Go = 10^3 Mo = 10^9 o (1 milliard)
1 To = 10^{12} o (1000 milliards)
- Twitter produit 70 Go/jour = 2 BNF/an!
- Volume total du web en 2015:
7 Zetta-octets = $7 \cdot 10^{21}$ o = ½ milliard de BNF

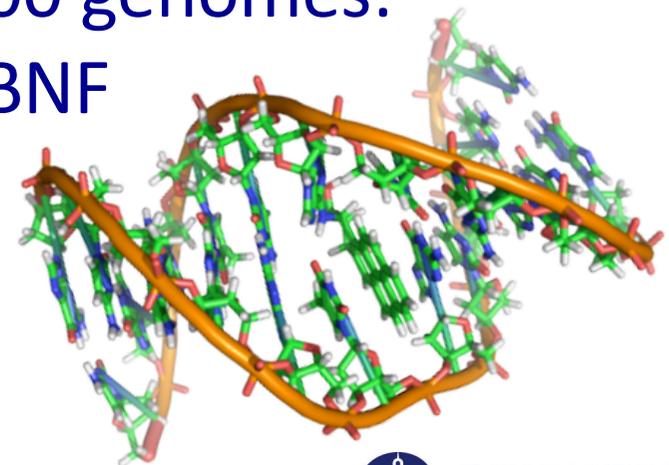


How Big are Big Data?

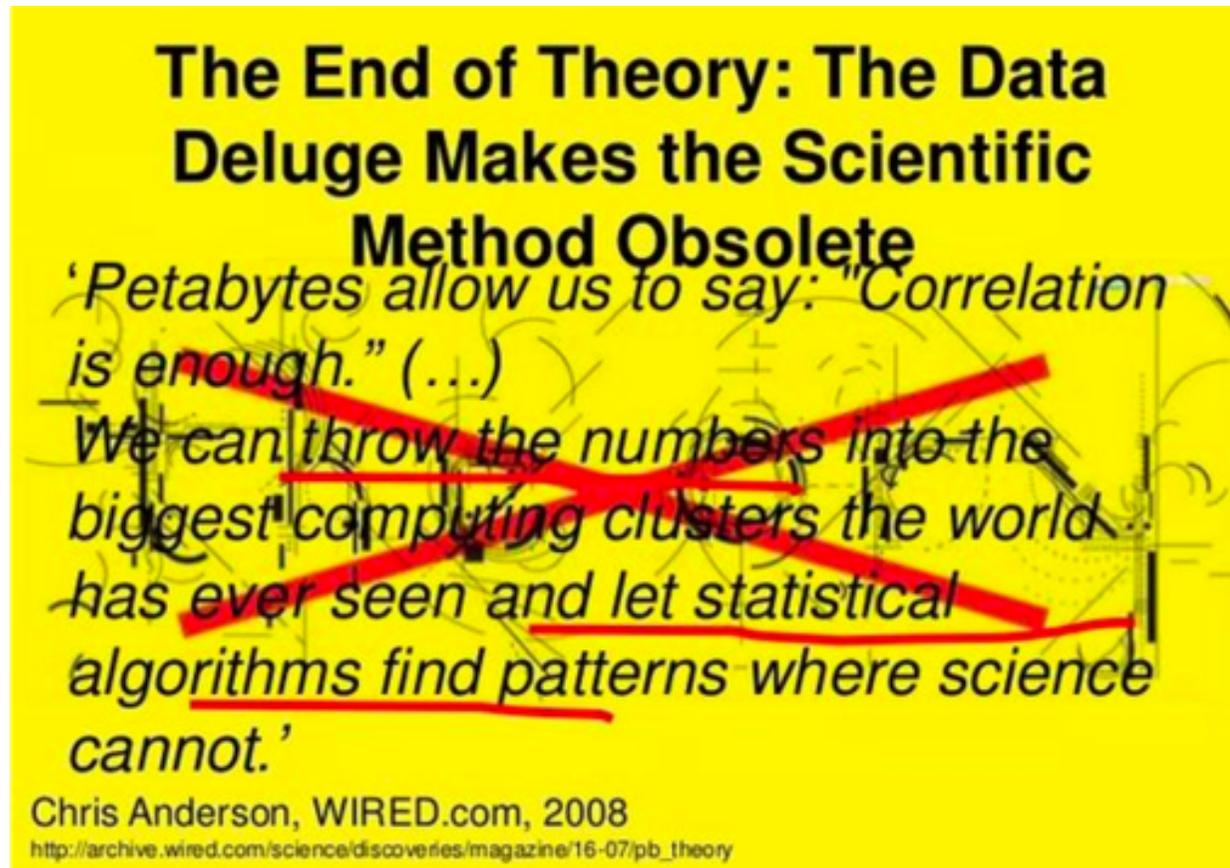
- Twitter produit
70 Go/jour = 2 BNF/an!
- *Volume total du web en 2015:*
7 Zetta-octets = $7 \cdot 10^{21} \text{o}$
= $\frac{1}{2}$ milliard de BNF



- Génome
 $3,4 \cdot 10^9 \text{o} = 0,00024 \text{ BNF}$
- Humanités numériques
130000 livres numérisés
 $130 \cdot 10^9 \text{o} = 0,009 \text{ BNF}$
- 10000 génomes:
2,4 BNF



The End of Theory!



- Succès de l'apprentissage:
 - publicité ciblée,
 - traduction automatique, etc.

Synoptique

1. **Éthique:** quelques rappels
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes** – HLEG-AI & al.
4. **Épistémologie de l'apprentissage:**
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



Dignité Autonomie

Bienfaisance **Human-centered**

Justice

Non malfaisance...

Équité

Liberté Absence biais

Vie privée



Déclaration de Montréal

1- WELL-BEING PRINCIPLE

The development and use of artificial intelligence systems (AIS) must permit the growth of the well-being of all sentient beings.

2- RESPECT FOR AUTONOMY PRINCIPLE

AIS must be developed and used while respecting people's autonomy, and with the goal of increasing people's control over their lives and their surroundings.

3- PROTECTION OF PRIVACY AND INTIMACY PRINCIPLE

- + Top of the page
- + Reading the Declaration
- Preamble
- Well-being
- Respect for autonomy
- Privacy and intimacy
- Solidarity
- Democratic participation
- Equity
- Diversity inclusion
- Prudence
- Responsability
- + Sustainable development
- Glossary
- Credits

Comités d'éthique

Plus de 67 Principes & recommandations

- The **Asilomar AI Principles**, developed under the auspices of the Future of Life Institute, 2017
- The **Montreal Declaration for Responsible AI**, developed under the auspices of the University of Montreal, 2017
- The **General Principles of Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems**, IEEE, 2017
- The **Ethical Principles offered in the Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems**, European Commission's European Group on Ethics in Science and New Technologies, 2018
- The **"five overarching principles for an AI code"** offered in paragraph 417 of the UK House of Lords Artificial Intelligence Committee's report, *AI in the UK: ready, willing and able?*, 2018
- The **Tenets of the Partnership on AI**, a multi-stakeholder organization consisting of academics, researchers, civil society organizations, companies building and utilizing AI technology, 2018
- **20 recommendations** of the AI4People, an Atomium—EISMD initiative designed to lay the foundations for a "Good AI Society", 2018
- **Ethical Guidance for a Trustworthy AI**, High Level Expert Group on Artificial Intelligence, European Commission, 2019

Ces principes sont non seulement nombreux, mais contradictoires!

- Vie privée vs. Transparence
- Vie privée vs. Sécurité
- Absence de discrimination vs. Politique inclusive
- ...

Encore un autre guide d'éthique de l'IA!

IA digne de confiance

1. Respect de la loi
2. Éthique
3. Robuste

Trois couches

1. Principes (inspirés de la bioéthique):
 - Respect de l'autonomie humaine
 - Prévention des dommages (harm)
 - Équité
 - Explicabilité et/ou transparence
2. Réalisation d'une IA digne de confiance
 - **Sept exigences:** respect l'initiative de l'action (*agency*), robustesse, vie privée, transparence, non-discrimination et équité, bien-être social en environnemental, responsabilité (*accountability*)
 - **Méthodes techniques et non techniques**
3. Évaluation d'une IA digne de confiance



RGPD

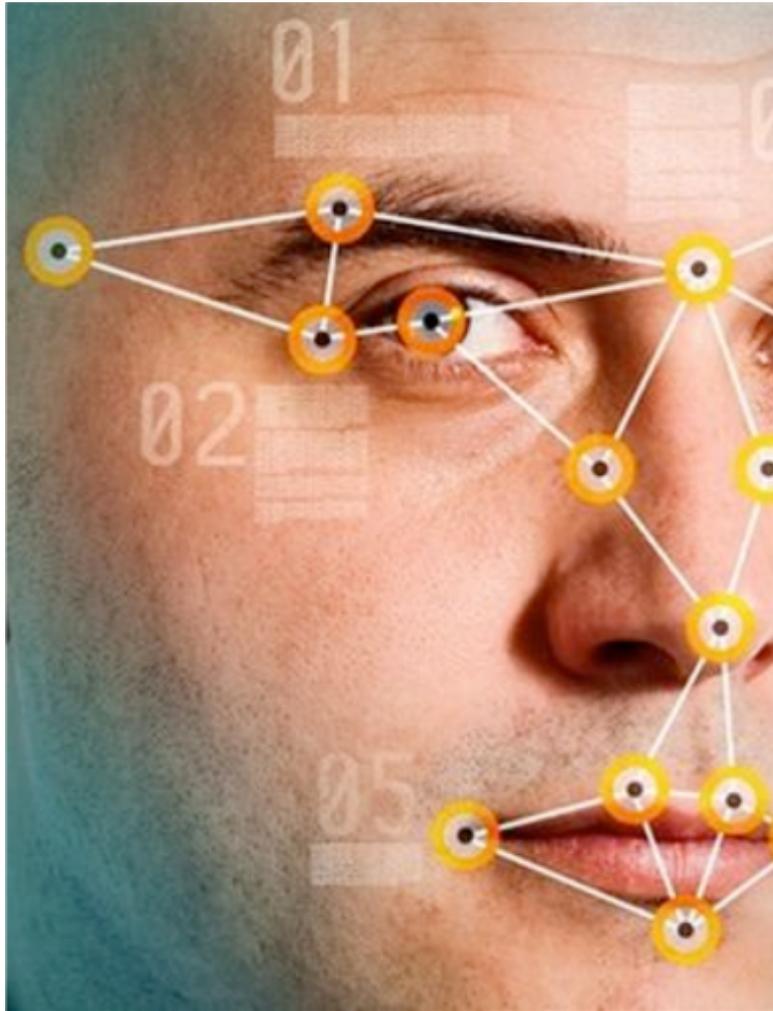
Règlement Général sur la Protection des Données



Principes

- **Finalité:** une organisation doit donner un objectif légitime avant de collecter des données personnelles
- **Transparence:** une organisation doit notifier les utilisateurs de la collecte et du partage de leur données personnelles.
- **Respects de droits de la personne:** les utilisateurs ont le droit d'accepter ou de rejeter la collecte de leurs données personnelles. Elles peuvent aussi en demander la correction ou l'effacement

Reconnaissance faciale et homosexualité



- Wang, Y., & Kosinski, M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*.
- Étude approuvée par le comité d'éthique de l'université de Stanford

Censure: *l'origine du monde* Courbet



Elon Musk, Neuralink et la ruche humaine

- “Over time I think we will probably see a **closer merger of biological intelligence and digital intelligence.**”
- “it's mostly about the bandwidth, the speed of the connection between your brain and **the digital version of yourself**, particularly output. ”



Hivemind: l'esprit de la ruche...
Collaboration des esprits qui partagent entre eux leurs connaissances et leurs opinions
Le « lacet neuronal »

Synoptique

1. **Éthique**: quelques rappels
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes** – HLEG-AI & al.
4. **Épistémologie de l'apprentissage**:
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



Induction

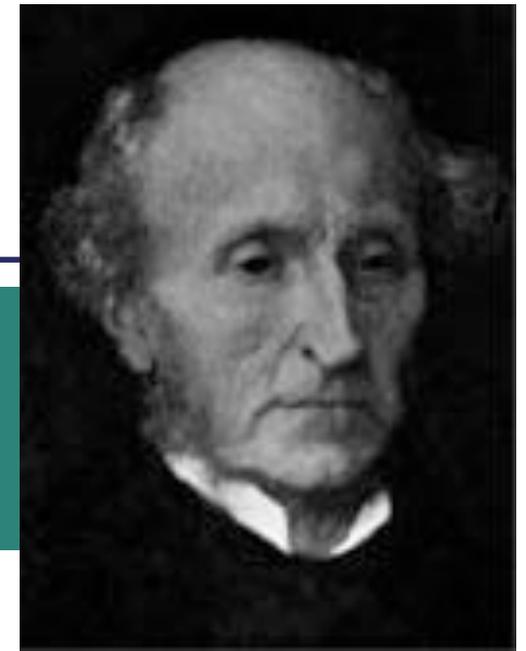
=

passage du particulier au général

"On appelle induction lorsque la recherche de plusieurs choses particulières nous mène à la connaissance d'une vérité générale. Ainsi, lorsqu'on a éprouvé sur beaucoup de mers que l'eau est salée, et sur beaucoup de rivières que l'eau en est douce, on conclut généralement que l'eau de la mer est salée et celle des rivières douce." [Arnauld & Nicole], p. 321

Le syllogisme inductif de Mill

System of logic ratiocinative and inductive, Vol 1, Book III, ch. 1



- Introduction d'un "principe d'uniformité"
- Construction d'un syllogisme inductif

Ce qui est vrai de A, B, C, D, ... est vrai de tous les cygnes
A, B, C, D, ... sont blancs

Tous les cygnes sont blancs

Induction de Mill

=

Déduction sous condition!



Causalité

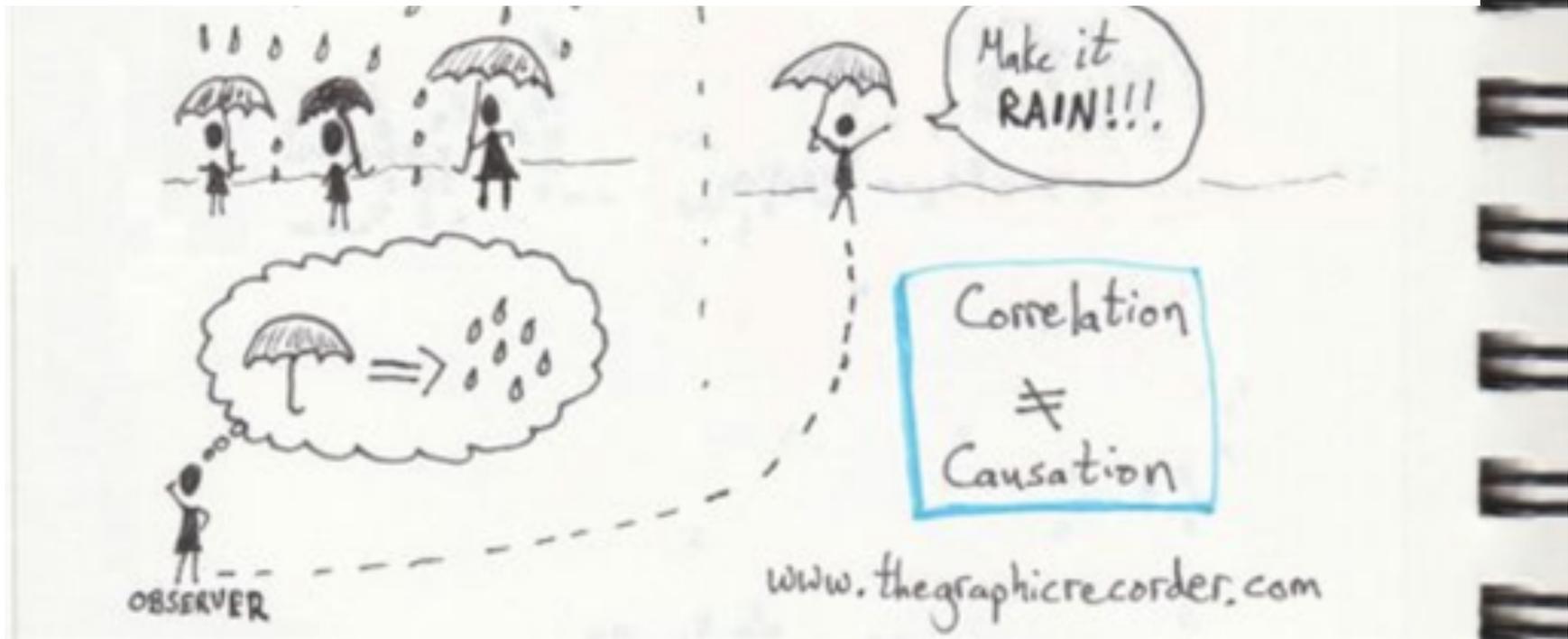
- Relation(s) Cause \rightarrow Effet
 - Si C alors D (relation logique: $C \Rightarrow D$; cela signifie aussi que $\neg D \Rightarrow \neg C$)
 - Si $C[t]$ alors $D[t + \delta_t]$ (relation physique)
 - Si C change alors D change (Interventionisme)



- Y-a-t-il une ou plusieurs relations de causalité?

Causalité \neq Corrélation

Il existe une corrélation entre l'utilisation des crèmes solaires et l'apparition de cancers de la peau... Qu'en conclut-on?



Synoptique

1. **Éthique**: quelques rappels
2. **Le monde change!**
 1. Nouvelles vulnérabilités
 2. Apprentissage et masses de données
3. **Éthiques de principes** – HLEG-AI & al.
4. **Épistémologie de l'apprentissage**:
 1. Induction
 2. Causalité
5. **Éthique vs. épistémologie**



CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE

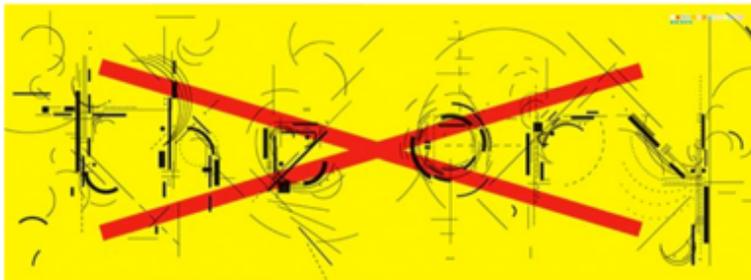


Illustration: Marian Bantjes "All models are wrong, but some are useful."

L'ère du péta-octet!

- Pas d'échantillonnage
- Recueillir tout, sans modèle *a priori*
- Ni sémantique, ni analyse causale requise
- "Correlation is enough"

Épistémologie vs. Éthique des Big Data



- Non conforme au RGPD
- Non conforme au rapport du HLEG-IA



L'âge du pétaoctet!

- Pas d'échantillonnage
- Collecte intégrale, sans modèle et objectif a priori
- Pas d'analyse sémantique ou causale requise.
- “Les corrélations suffisent”

Le mirage des masses de données



- Les « dogmes » des chercheurs et/ou des machines...
- Difficultés éthiques et épistémologiques:
 - Question de représentation
 - Biais des données (uniformité?)

Exemple de problème éthique

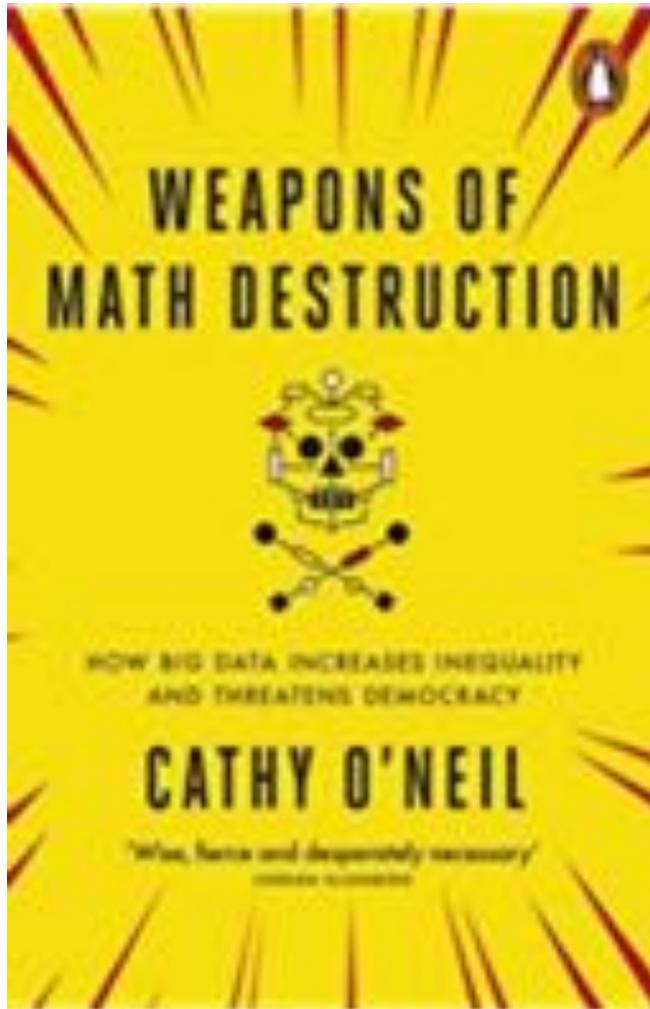
- Etude aux Etats-Unis en 2007 sur 15000 adolescents
 - Dérivation de la conclusion à partir de données d'observations
 - Article *Ohio State University*: l'âge de la première expérience sexuelle **cause** la délinquance
 - *pour réduire la délinquance, retarder l'âge de la première expérience sexuelle*
 - Étude *University of Virginia Charlottesville*: conclusions très différentes... facteurs génétiques et environnementaux déterminants

Systemes prédictifs

- Utilisation d'Apprentissage sur de grandes masses de données (*big data*)
- Applications:
 - Profilage individus (publicité ciblée)
 - Prédiction
 - Risques (assurances),
 - gains,
 - comportements des clients
 - établissement dynamique des prix,
 - ...



Caractère prédateur des systèmes prédictifs



- Non représentativité des exemples: l'axiome d'uniformité (Axiome de Mill) n'est pas vérifié! – c'est ce qu'on appelle le biais
- Objectif de l'induction: risque d'accident de voiture ou de changement d'assurance...
- Causalité \neq corrélation
- Autoconfirmation dans les applications à la société: ex. déploiement force de police ou justice prédictive

Conclusions

Faut-il adopter une éthique de principes?

- **Études de cas**
- **Nécessité de compromis:**
 - Vie privée
 - Transparence
 - Sécurité
 - Non discrimination
 - Inclusion
 - ...
- **Arbitrage:**
 - Dilemmes éthique à résoudre
 - Logique non monotones

