

# Éthique et épistémologie des données



***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!** nouvelles vulnérabilités liées aux données
3. **Éthiques de principes – HLEG-AI & al.**
4. **Épistémologie de l'apprentissage:**
  1. Induction
  2. Causalité
  3. Masses de données
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

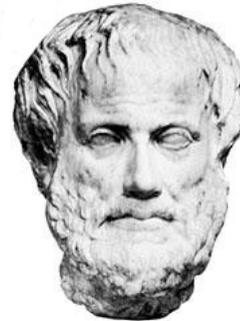


- *êthos* en grec
  - Mœurs
  - Etable: là où l'on rentre le soir, ce que l'on habite
- Fondement originel de l'éthique:
  - Traditions, usages

# Comment se mettre d'accord?

- **Préceptes**

- Religion
- Valeurs/vertus



Immanuel Kant  
Zeichnung im Schauffnermanner von Puttrich um 1799  
Archiv für Kunst und Geschichte

- **Principes**

- Utilitarisme / conséquentialisme
- Déontisme / impératif catégorique

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

- Savant, connaissance

- **Délibération *J. Habermas***

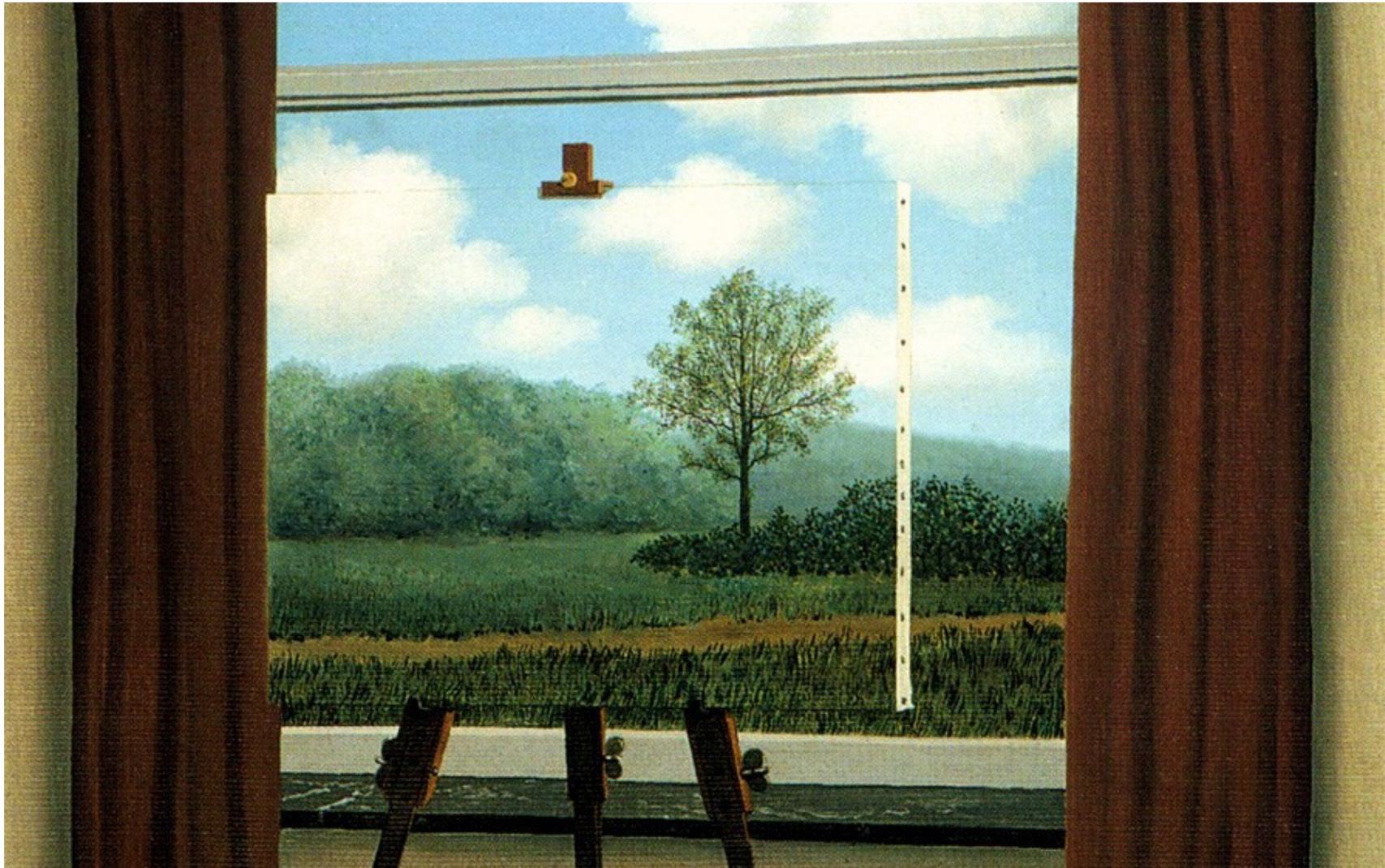
- Complexité du savoir – communauté de savants

# Synoptique

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



# La nouvelle condition humaine



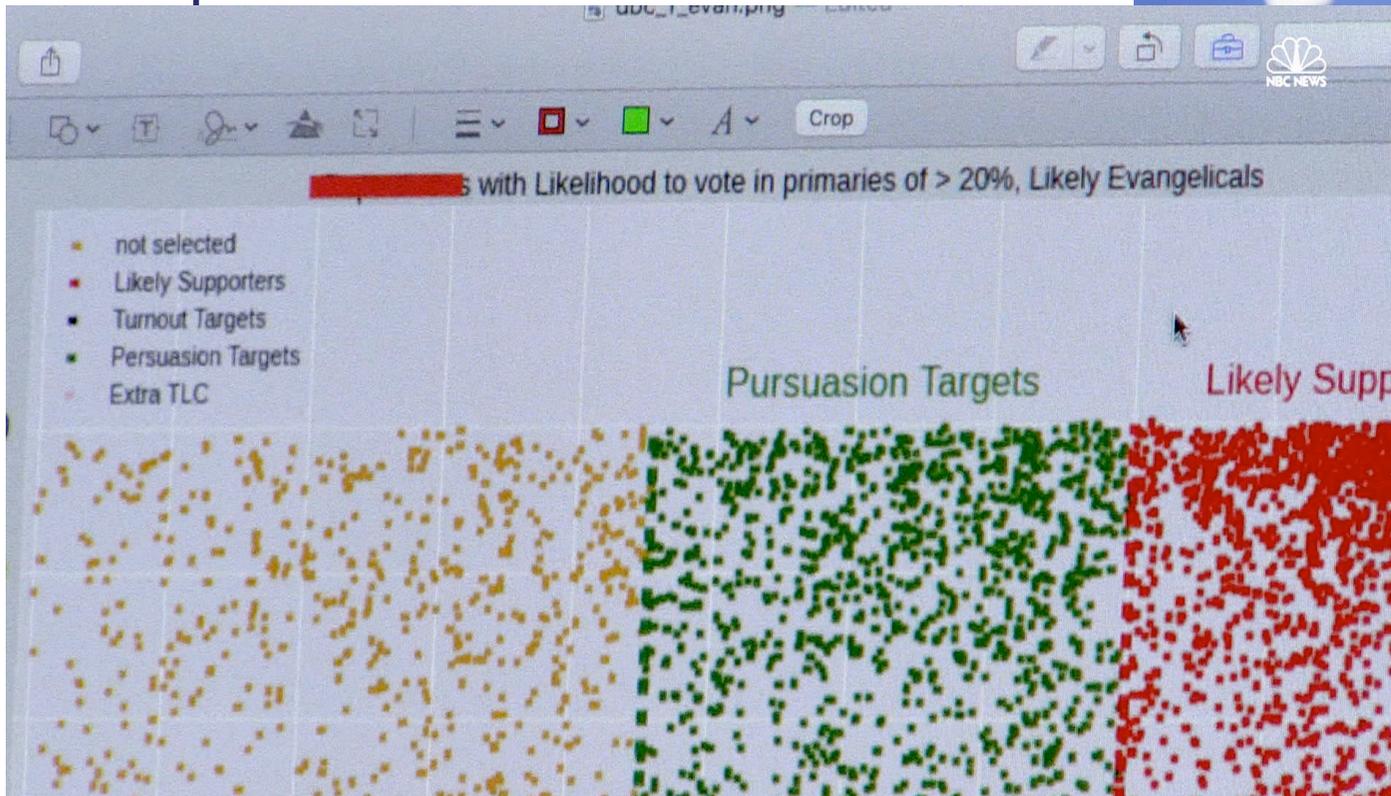
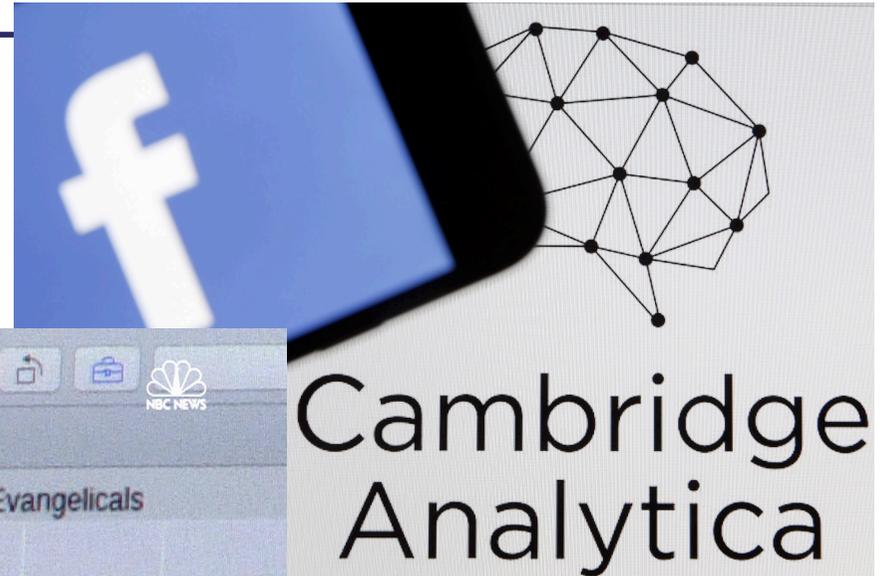
# Réontologisation des sociétés

## Évolution des relations humaines

- Amitié (*réseaux sociaux*)
- Réputation (*score...*)
- Confiance (*blockchain*)
- Monnaie (*paypal, bitcoin, ...*)
- Travail
- Souveraineté
- ...



# Nouvelles vulnérabilités



# Nouvelles vulnérabilités: deep fake!



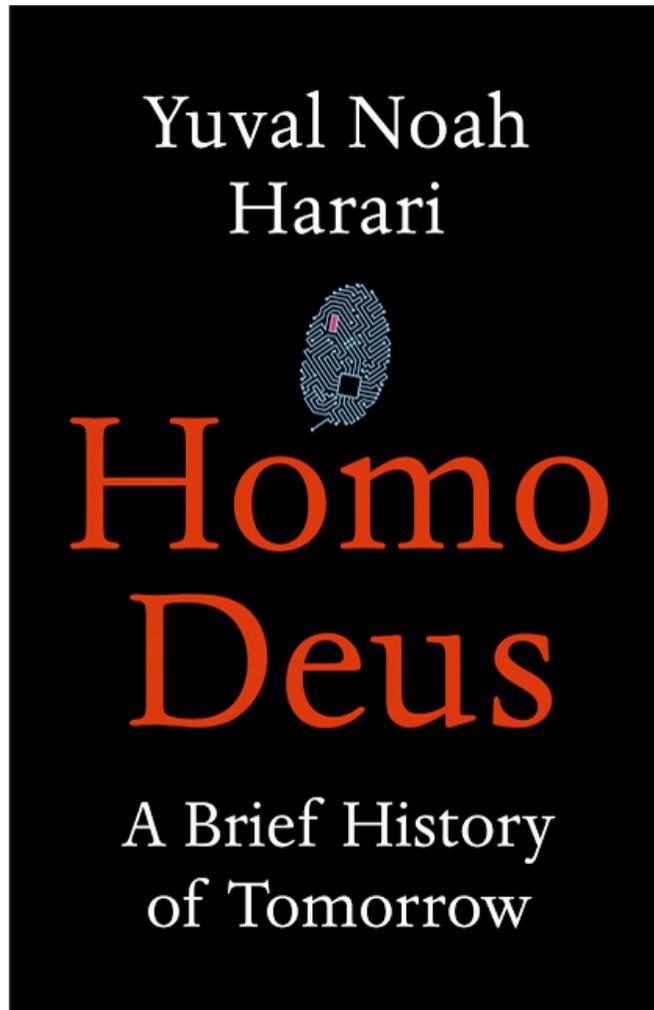
Manipulations  
de paroles et  
d'images...



# 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 machine

# Synoptique

1. **Éthique: quelques rappels**
2. **Le monde change!** nouvelles vulnérabilités liées aux données
3. **Éthiques de principes – HLEG-AI & al.**
4. **Épistémologie de l'apprentissage:**
  1. Induction
  2. Causalité
  3. Masses de données
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

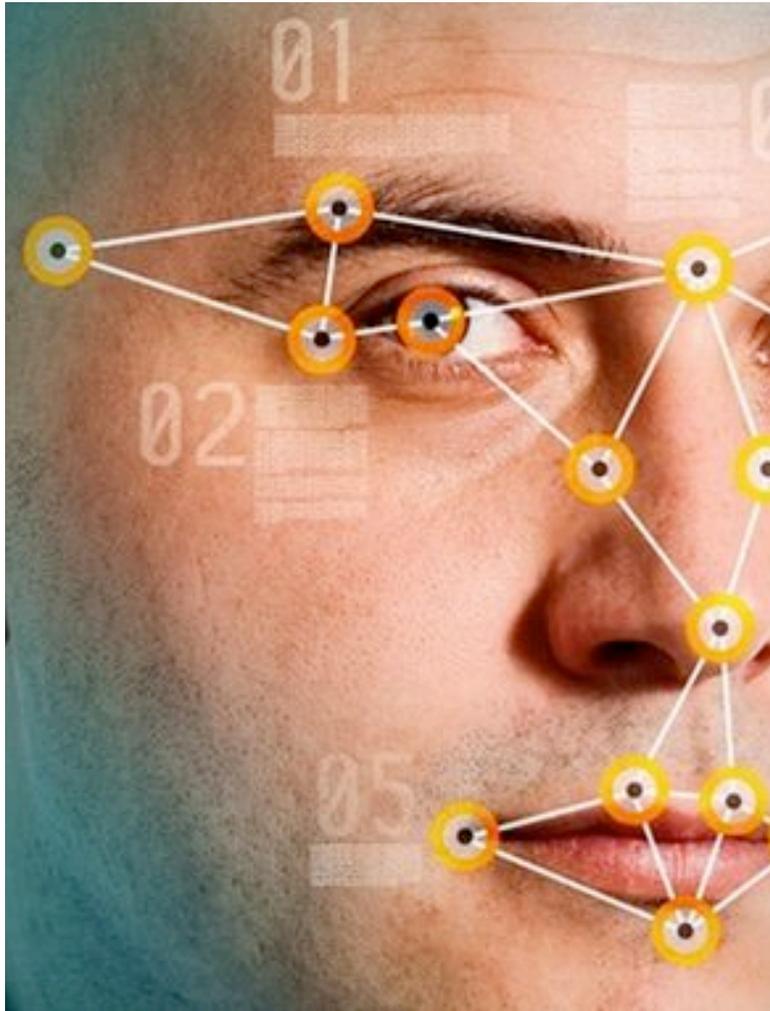
## *Règlement Général sur la Protection des Données*



### 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



# Synoptique

1. **Éthique: quelques rappels**
2. **Le monde change!** nouvelles vulnérabilités liées aux données
3. **Éthiques de principes – HLEG-AI & al.**
4. **Épistémologie de l'apprentissage:**
  1. Induction
  2. Causalité
  3. Masses de données
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

# Caractère conjecturale de l'induction ?

## Ne pas confondre *induction* et *abduction*

Vision Aristotélicienne de l'induction (inversion de déduction)

Syllogisme:

Fumer donne le cancer

Majeure

Jean est un fumeur

Mineure

Jean risque d'avoir le cancer

Conclusion

**Induction**

=

**Déduction<sup>-1</sup>**

L'induction passe de la Mineure et de la Conclusion à la Majeure

L'abduction passe de la majeure et de la conclusion à la mineure

- Il existe des inductions certaines
  - Induction mathématique
  - Induction entière (induction logique pour Aristote; toutes les éventualité sont envisagées)

# 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**

# Jean-Nicod (1893-1924)

## Logique de l'induction:

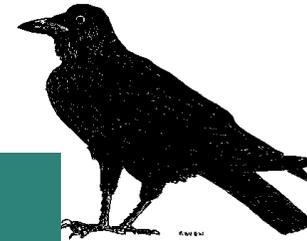
- But: Fonder l'induction sur des connaissances certaines
  - Confirmation: degré de plausibilité
  - Infirmité: exclusion
- 
- Tentative de fonder une logique de l'induction sur
    1. la confirmation uniquement: échec
    2. l'infirmité uniquement: échec
    3. L'infirmité et la confirmation: les exemples ont une force corroborante, mais on ne saurait atteindre la certitude absolue -> logique probabiliste de l'induction



# Logique de la confirmation



## Paradoxe de Hempel



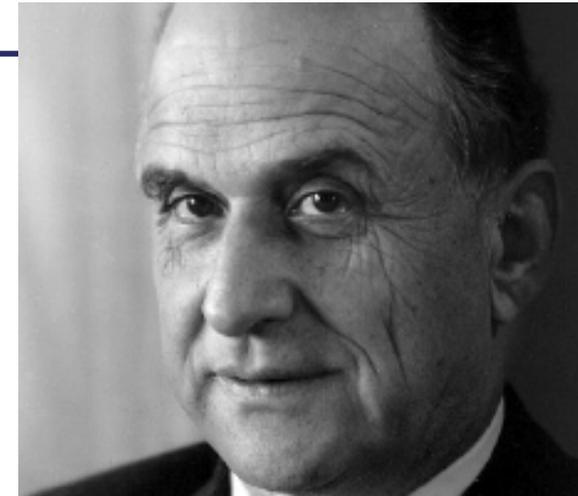
Tous les corbeaux sont noirs  
 $\equiv$   
Tous les non-noirs sont non-corbeaux

- La présence d' émeraudes vertes confirme L...



$\forall x (\text{corbeau}(x) \wedge \neg \text{noir}(x)) \Rightarrow (\text{corbeau}(x) \wedge \neg \text{corbeau}(x))$  est équivalent à S1 et S2... Mais aucun exemple ne le confirme...

# Nelson Goodman *“facts, fiction and forecast”*



Toutes les émeraudes ne sont peut-être pas vertes !

Imaginons un prédicat “mal élevé... appelé “Vleu”

“Vleu” = vert avant le 1er janvier 2020, bleu après

Pourquoi ne pas induire “**toutes les émeraudes sont vleues**” ?

Le temps n’intervient pas dans “vleu”, mais dans vert:  
Vert = “Vleu” avant le 1er janvier 2020, “Blert” après

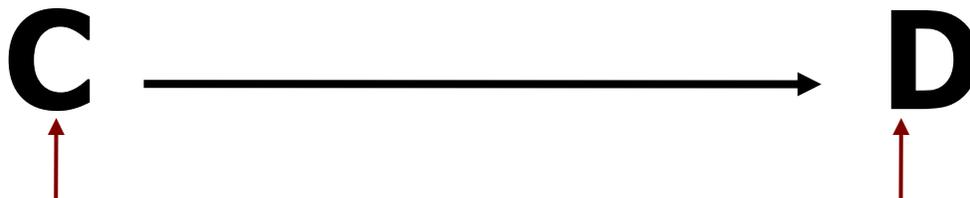
# Synoptique

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



# Causalité

- Relation 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)



# Questions philosophiques

---

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

Woodward: il n'existe qu'une seule relation à laquelle tout ce que l'on entend par causalité se réduit. Pour Woodward, c'est l'interventionisme qui est l'explication fondamentale.

Glymour: il peut y avoir plusieurs causalité...



# Science et causalité

---

- Connaissance causalité → décision
  - Médecine
  - Politique
  - Sciences sociales
  - Physique et ingénierie
  - ...
- Question: comment détecter les relations de causalité?

# Causalité $\neq$ Corrélation



# Causalité et corrélation statistique

- Si C cause D alors...  
... chaque fois que C intervient, D doit aussi intervenir
- But: détecter relations de causalité
  - Si C apparaît corrélé à D, autrement si chaque fois que C apparaît D apparaît, peut-on en déduire que C cause D?
- *Il existe une corrélation entre l'utilisation des crèmes solaires et l'apparition de cancers de la peau... Qu'en déduit-on?*

# Synoptique

1. **Éthique: quelques rappels**
2. **Le monde change!** nouvelles vulnérabilités liées aux données
3. **Éthiques de principes – HLEG-AI & al.**
4. **Épistémologie de l'apprentissage:**
  1. Induction
  2. Causalité
  3. Masses de données
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

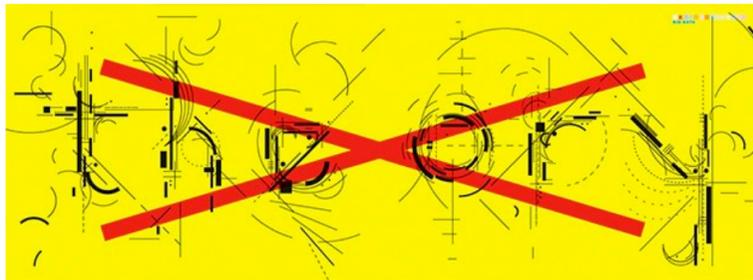


Illustration: Marian Bantjes "All models are wrong, but some are useful."

## L'ère du péta-octet!

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

# Le mirage des masses de données



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

# Synoptique

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



# 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*
  - Etude *University of Virginia Charlottesville*: conclusions très différentes... facteurs génétiques et environnementaux déterminants

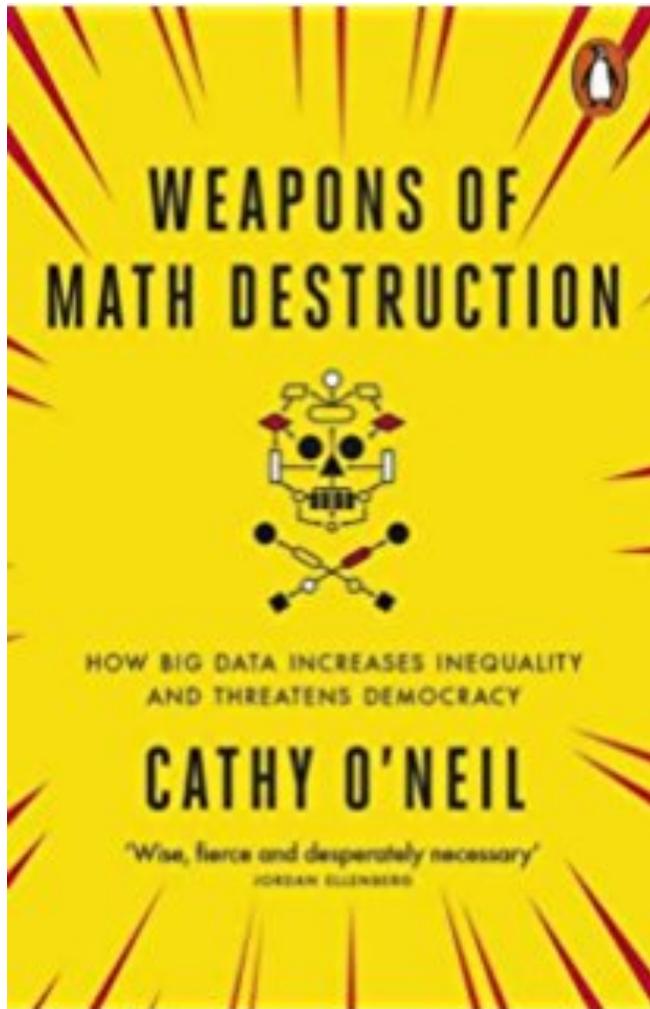
# Systemes prédictifs

- 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

## Ethique et épistémologie



- Non représentativité des exemples: l'axiome d'uniformité (Mill) n'est pas vérifié!
- Auto-confirmation: ex. déploiement force de police ou justice prédictive
- But de l'induction: calculer le risque d'accidents de voiture ou de changement d'assurance...
- Causalité  $\neq$  corrélation

# 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”

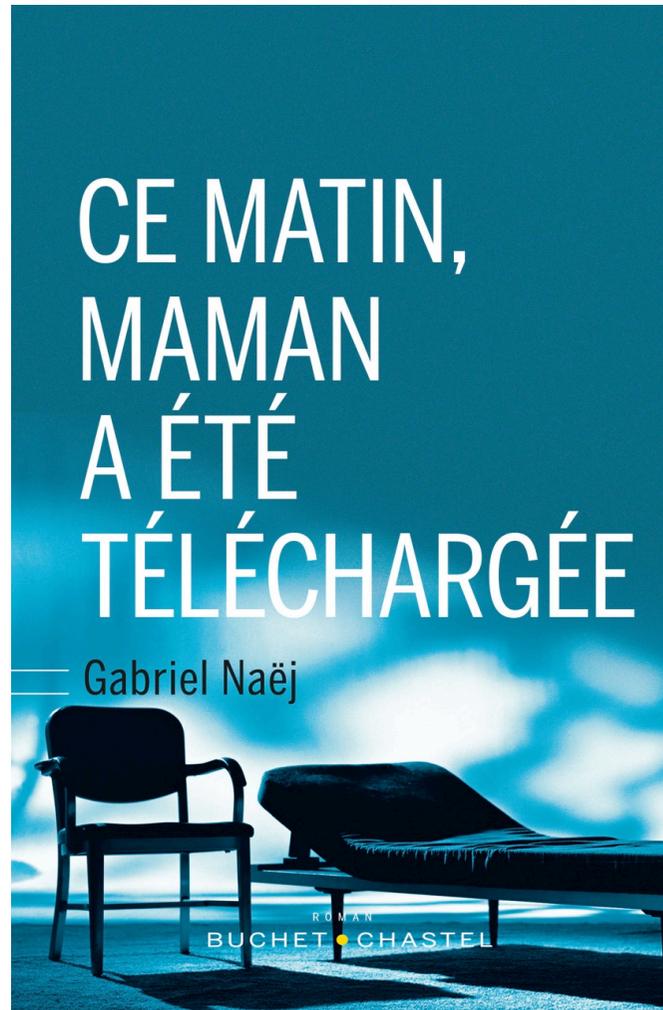
# Conclusions

**Faut-il adopter une éthique de principes?**

- **Études de cas**
- **Compromis:**
  - Vie privée
  - Transparence
  - Sécurité



# Post-conclusion



## Confession

Je me suis trompé!

La Singularité est advenue

Et, comble de malheur pour moi, maman,  
ma vraie mère s'est téléchargée.

Et en plus, elle s'est réincarnée dans un  
robot sexuel ce qui me trouble beaucoup

Et, elle fait irruption à tout moment dans ma  
vie...

Je vous demande d'être compatissant