

# Éthique et numérique

M2 Innovation, Digital, Conseil

Serge Pajak • Maître de conférence en Économie

[serge.pajak@universite-paris-saclay.fr](mailto:serge.pajak@universite-paris-saclay.fr)

## Cours de 5 séances

Illustrations des situations nouvelles issues de l'économie numérique, et des questions éthiques qui sont associées

**Pré-requis** : aucun. Ni en économie, ni en éthique, ni en numérique

# Déroulement du cours

## Dossier – étude de cas.

Objectif : présentation de la situation, explicitation des enjeux, référence à la réglementation et aux principes éthiques applicables, propositions d'alternatives

Consignes précises disponibles sur eCampus

# Évaluation

# Pourquoi s'intéresser à l'éthique du numérique ?

***“And, so what?”***

---

1. Toute innovation vient aussi avec un aspect négatif ou critiquable

Leur mise en œuvre fera ressortir des questions éthiques

---

2. L'éthique est un élément de la stratégie et de l'organisation des entreprises et des acteurs publics

Les cours d'éthique anticipent souvent la réglementation... (ex. : RGPD)

**Les cadres du secteur comme vous doivent anticiper ces questions**

# ***Facebook, Citing Societal Concerns, Plans to Shut Down Facial Recognition System***

Saying it wants “to find the right balance” with the technology, the social network will delete the face scan data of more than one billion users.

*The New York Times, Nov. 2, 2021*

Business | Rift valley

# The Sam Altman drama points to a deeper split in the tech world

Doomers and boomers are fighting for AI dominance



*The Economist, Nov. 19th, 2023*

# Plan du cours

1. Théories de l'éthique et de la décision

---

2. Publicité et donnée personnelle

---

3. Illusion de contrôle, surcharge informationnelle et bad nudge

---

4. Présentation de l'information, mise en page et discrimination

---

5. Éthique de l'expérimentation de terrain

---

6. Tarification personnalisée : la fin de l'affichage des prix ?

---

7. Décision par algorithmes

---

8. Open data

---

9. L'Intelligence Artificielle : principes et applications

---

# Théories de l'éthique et de la décision

# Ne pas confondre

**L**oi, réglementation : obligations qui s'imposent par la force publique à l'ensemble de la société

**D**éontologie : règles professionnelles régissant l'ensemble des situations possibles

- Il n'est pas nécessaire de les comprendre ou de les partager pour les respecter
- Les conséquences n'entrent pas en ligne de compte

---

Pour ce cours, déontologie = loi = pas notre sujet

---

**É**thique : réflexion de l'acteur concerné sur ses pratiques

- Éthique de responsabilité (prise en compte des conséquences dans le contexte donné) vs. éthique de conviction (respecter des principes)
- Une réflexion éthique peut influencer la réglementation (ex. RGPD)

# Quelques principes

## Éthique de conviction (Weber) ou impératif catégorique (Kant)

- l'action est guidée par des principes généraux, basés sur l'intention et les motivations de l'acteur
- Le contexte ou la conséquence n'est pas pris en compte

## Three Laws of Robotics

---

From Wikipedia, the free encyclopedia

The **Three Laws of Robotics** (often shortened to **The Three Laws** or known as **Asimov's Laws**) are a set of rules devised by [science fiction](#) author [Isaac Asimov](#). The rules were introduced in his 1942 short story "[Runaround](#)" (included in the 1950 collection *I, Robot*), although they had been foreshadowed in some earlier stories. The Three Laws, quoted from the "Handbook of Robotics, 56th Edition, 2058 A.D.", are:<sup>[1]</sup>

### First Law

A robot may not injure a human being or, through inaction, allow a human being to come to harm.

### Second Law

A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

### Third Law

A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

# Quelques principes

**É**thique de responsabilité (Weber, 1919), conséquentialisme : on doit répondre des conséquences de ses actions

**U**tilitarisme : Bentham (1789), Stuart Mill (1861) : la société est la somme des intérêts et des états de chacun de ses membres.

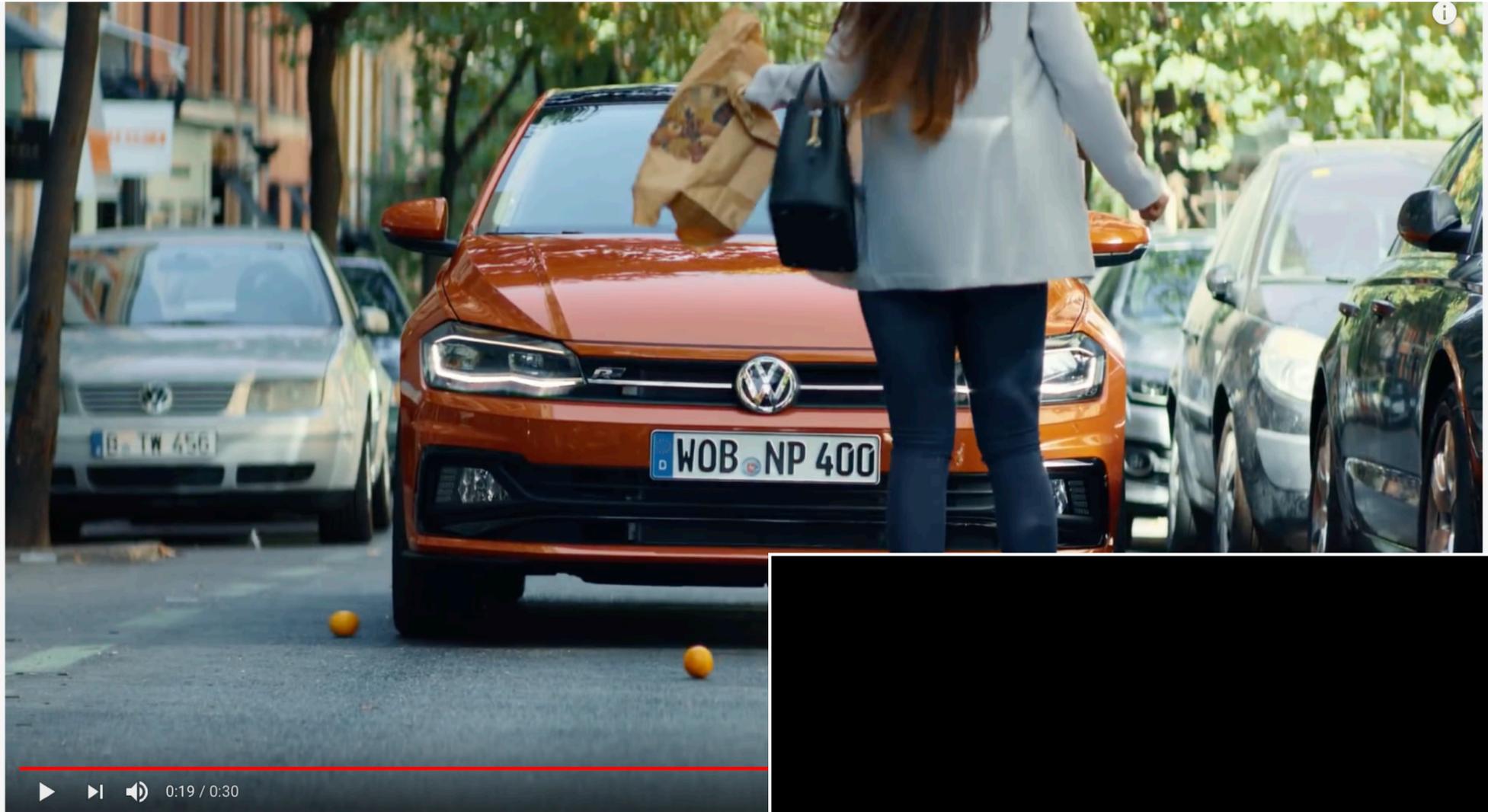
- Pas d'intérêt général, pas d'autorité supérieure
- Introduit l'analyse coûts-bénéfices face à des modes de décision basés sur des principes intangibles

# Quelques références & concepts

**Luddisme** Conflit industriel en Angleterre au 19<sup>e</sup> siècle autour des conséquences sur l'emploi de l'arrivée des machines à tisser

**Solutionnisme technologique** Courant de pensée originaire de la Silicon Valley qui souligne la capacité des nouvelles technologies à résoudre les grands problèmes du monde, comme la maladie, la pollution, la faim ou la criminalité

- Réduction des questions sociales à un concours Lépine de réponses
- Déresponsabilisation des acteurs



Volkswagen | La nouvelle Polo avec reconnaissance de piétons

437,979 views • Oct 10, 2017



Volkswagen | La nouvelle Polo avec reconnaissance de piétons

437,979 views • Oct 10, 2017

8 9 SHARE SAVE ...

# Quelques références & concepts

**M**elvin Kranzberg, Jacques Ellul

« La technologie n'est ni bonne, ni mauvaise, ni neutre »

**S**cientisme/rationalisme Croyance selon laquelle la science est la seule source de savoir sur le monde, sa méthode devrait être appliquée à tous les problèmes sociaux

**A**ceptation sociale Le public a des critères pour accepter ou rejeter une technologie qui ne se limitent pas aux caractéristiques techniques

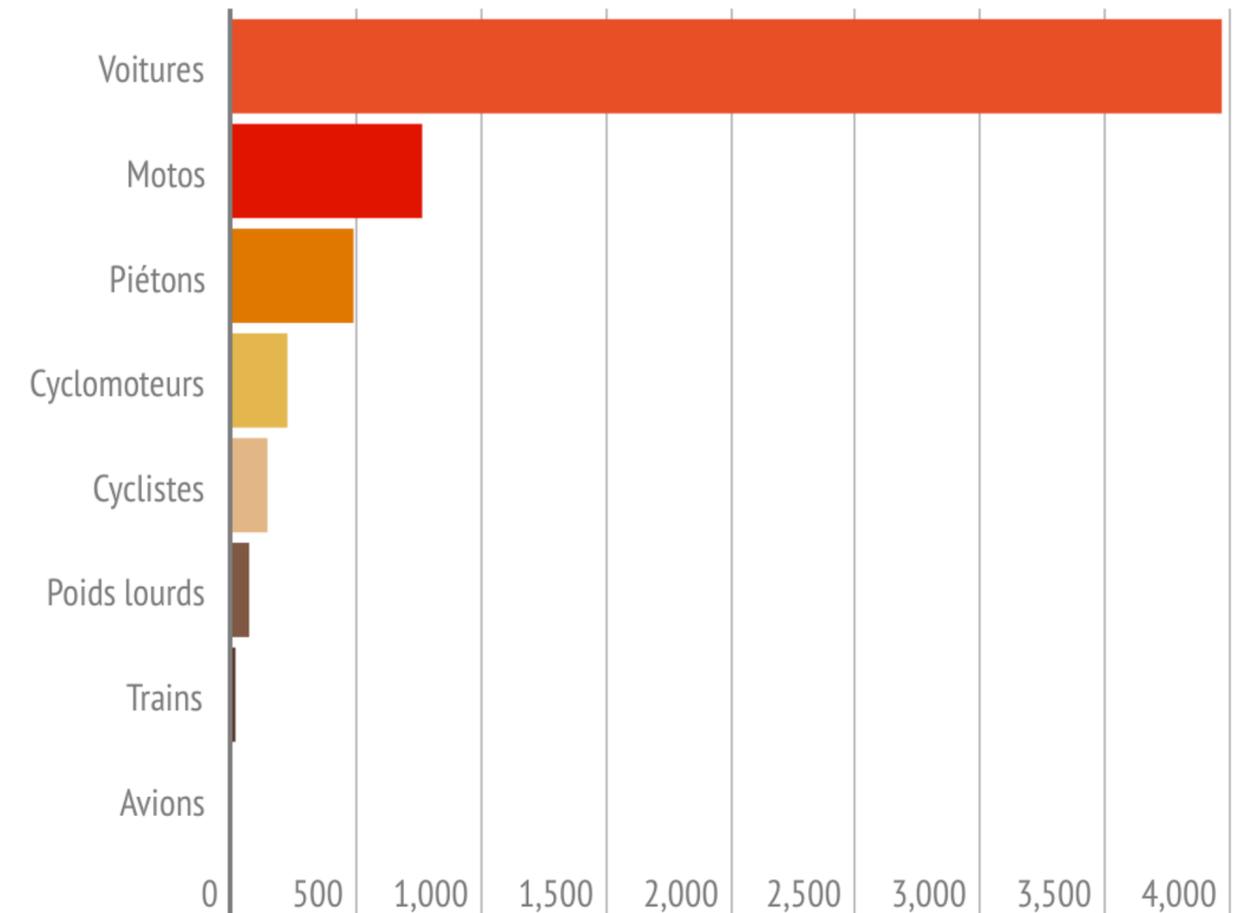
Ex. : la transparence; ou bien solliciter l'avis du public, peut changer l'acceptabilité

Ex. : le risque de dérive dans l'usage d'une technologie est faible, mais est-il accepté ?

# QUELS SONT LES RISQUES LIÉS AUX TRANSPORTS ?

## Le nombre de morts en France (2011)

Le nombre de tués en voitures (3963) est bien supérieur au nombre de morts en train (12) et en avion (0). Concernant les compagnies aériennes, il n'y a eu aucun mort en France sur des avions de ligne depuis le crash du Concorde en 2000.



Comparez la **réalité** de la dangerosité des transports avec l'**acceptation** des accidents

Et la voiture autonome ?

# La responsabilité de l'inventeur

« *Move fast & break things* »

**I**nventeur / découvreur : celui qui développe une technologie est-il responsable de ses applications ?

☰ **Bloomberg** [Subscribe](#)

---

Technology

## U.K. Warns Facebook to Keep Users Safe or Risk Criminal Charges

By [Nate Lanxon](#) [+Follow](#)  
4 November 2021, 12:30 CET

---

- ▶ Mark Zuckerberg, Nick Clegg named by lawmaker to make changes
- ▶ Culture Secretary Nadine Dorries criticizes Facebook rebrand



# Quelques motivations, au-delà de l'altruisme

## Les motivation intrinsèques

### Desire to contribute, pure altruism

- sequential reciprocity
- sympathy for others, hedonist altruism

### Identification to the community

- « OpenAI is nothing without its people »

# Quelques motivations, au-delà de l'altruisme

## Les motivations extrinsèques

### Warm glow effect or impure altruism

Reputation- and peer recognition-seeking (open source)

Improving one's human capital learning by contributing (open source, Wikipedia)

### Monetary

### Signal of quality in the labor market

- ▶ OSS contributor picked by companies as employees
- ▶ Differed gains

Publicité et données personnelles

# Le modèle économique de la donnée personnelle

Le modèle économique d'abonnement ou de micro-paiements pour le contenu en ligne existe mais reste faible

- La création et la diffusion de contenu sont principalement financées par la publicité

Les annonceurs définissent une audience et chargent les régies publicitaires (Google, Facebook, TikTok, etc.) de cibler cette audience

**Les régies se basent pour cela sur le profile des visiteurs établis par la collecte de données personnelles au fil de la navigation**

# Le modèle économique de la donnée personnelle

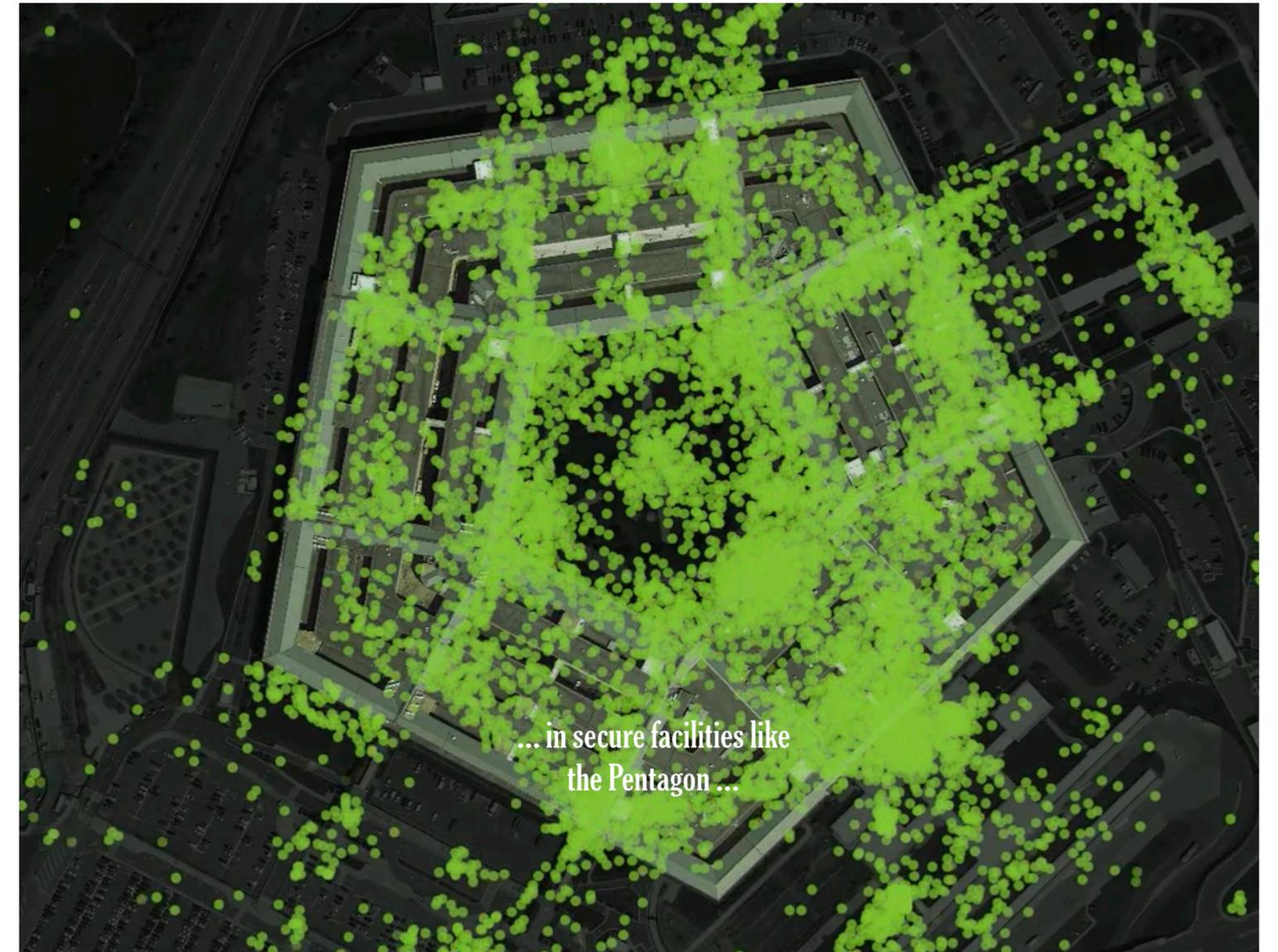
"The best minds of my generation are thinking about how to make people click ads."

--Jeff Hammerbacher, a 28-year-old Silicon Valley tech whiz who went from being an early employee at Facebook to co-founding the data analysis start-up Cloudera, in Ashlee Vance's *BusinessWeek* [story](#) about the advertising and social-media driven bubble in Silicon Valley.

# Données personnelles et surveillance

Enquête One Nation, Tracked du NY Times (2019)

<https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html>

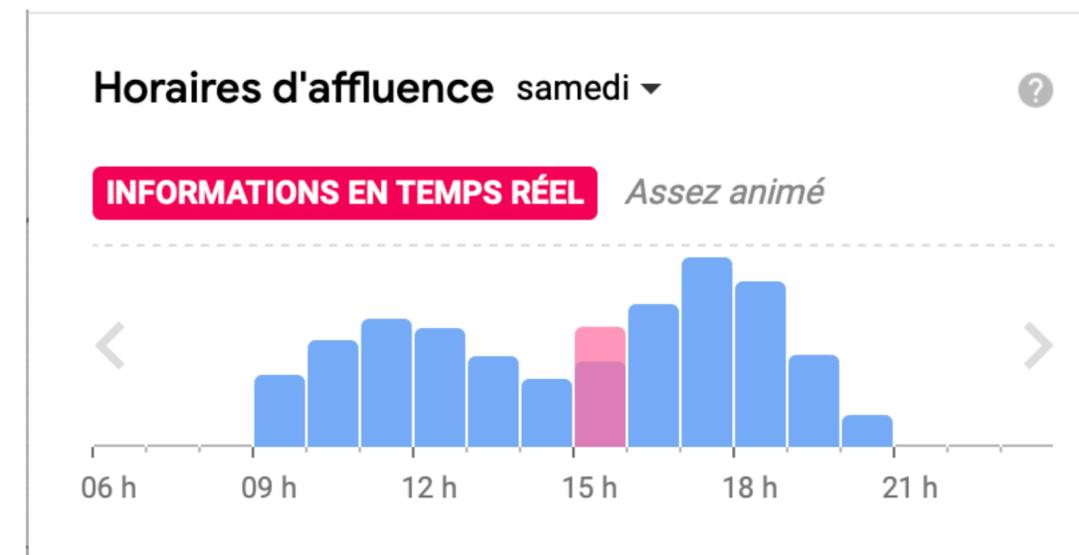
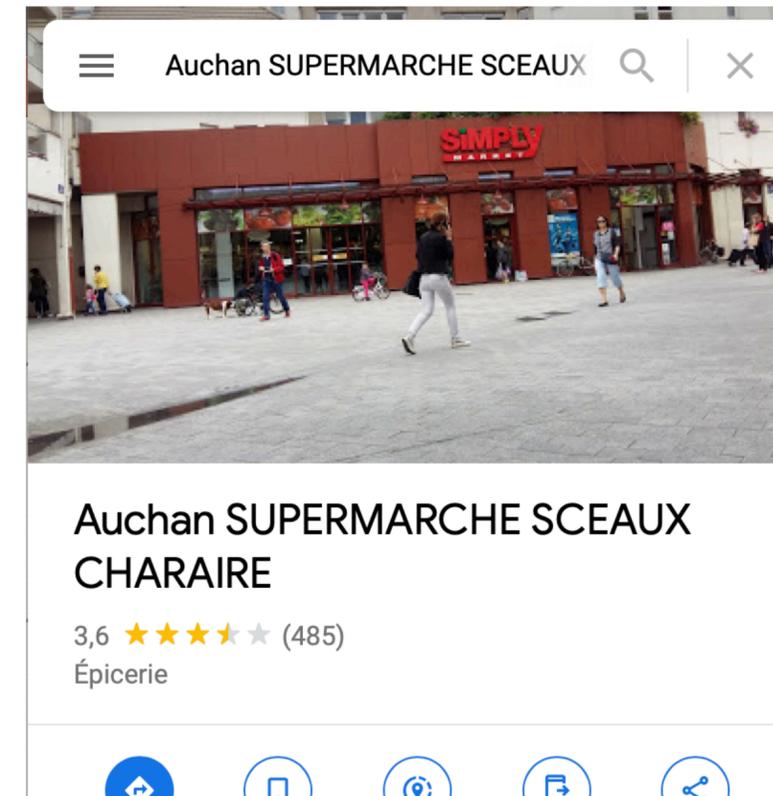


# Données personnelles et géo-localisation

Statistiques de fréquentation à base de données de géo-localisation

Pas d'identification individuelle, utilisation purement commerciale par Google Maps

Utilité indéniable du service... mais qui a consenti à fournir les données utilisées ?



# Publicité et trackers

## Ghostery : un détecteur de trackers

Très peu de sites sans aucun trackers

9 sur [www.universite-paris-saclay.fr](http://www.universite-paris-saclay.fr)

4 sur [lesechos.fr](http://lesechos.fr)

28 sur [cnn.com](http://cnn.com)

Plus nombreux sur les sites US (pas de RGPD !)

The screenshot displays the Ghostery browser extension interface. The top bar is blue with the Ghostery logo and name. The main content area is divided into two panels. The left panel shows the results for the website [www.mynews.com](http://www.mynews.com), indicating that 22 trackers were found. A donut chart shows the distribution of these trackers, with 4 of them blocked. Below the chart are three buttons: 'Trust Site' (with a shield icon), 'Restrict Site' (with a lock icon), and 'Pause Ghostery'. A link 'Map these trackers' is also present. The right panel is titled 'TRACKERS' and features a 'Block All' checkbox. It lists several trackers with their respective status (checked for blocked, unchecked for not blocked): Advertising (10 TRACKERS, 3 Blocked), Advertising.com (checked), DoubleClick (unchecked), Google AdSense (unchecked), Korrelate (unchecked), Moat (checked), NetRatings-Site-Center (checked), Polar-Mobile (checked), ScoreCard Research Beacon (unchecked), and Tacoda (unchecked).

# Les ambiguïtés du gratuit

« Si c'est gratuit, c'est vous le produit » dit l'adage

Quelle conséquence lorsqu'une offre est gratuite ?

Difficile de revenir en arrière : le consommateur a été habitué aux services en ligne gratuits

Quelle relation entre un utilisateur d'un produit gratuit et l'entreprise qui le fournit ?

Peut-il être exigeant, peut-il se plaindre du *tracking* comme contrepartie ?

# Acceptation du modèle actuel

## Les utilisateurs sont-ils responsables de cette situation ?

- Sont-ils d'accord ? Sont-ils informés ?

En déclaratif, ni l'un ni l'autre : « moi je n'ai rien demandé... »

- Comment expliquer que la situation actuelle existe dans la durée ?

On constate qu'ils refusent avant tout les alternatives (l'abonnement pour consulter un site)

---

**Les mécanismes de marché révèlent les préférences**

---

PIXELS · FACEBOOK

### **La nouvelle option d'abonnement payant à Facebook et Instagram suscite la confusion**

Le message des deux réseaux sociaux offrant le choix entre publicité et abonnement, diffusé depuis novembre, a provoqué l'incompréhension. Car loin d'être une nouveauté, la collecte de données personnelles est effective depuis les débuts de ces plates-formes.

Le Monde

Publié le 13 novembre 2023 à 12h01, modifié le 13 novembre 2023 à 13h39 · 🕒 Lecture 2 min.

# Plateformes et fonctions de recommandations

# Les fonctions de recommandations

Une part importante du trafic sur la plateforme provient des recommandations, construites à partir de l'historique et des interactions précédentes

La fonction objectif : favoriser le clic, favoriser l'engagement, le temps de visionnage, la durée de la session...

Risque de « chambre d'écho » : on n'obtient que les informations qui correspondent à ses préjugés

Théories du complot, comportements dépressifs, addictions

**YouTube / Facebook / TikTok etc.  
est-il responsable des recommandations de  
son algorithme ?**

**Oui : L'interface utilisateur des plateformes est  
soigneusement travaillée et optimisée**

# Instagram Will Test Hiding 'Likes' in the US Starting Next Week

Hiding like counts is just the latest step in Instagram's quest to become the safest place on the internet.



Menu +

TWITTER - X / TECH / CREATORS

**Twitter is now showing everyone how many views your tweets get** / Elon promised the feature earlier this month, and now they've started rolling out.

By [Mitchell Clark](#)

Updated Dec 23, 2022 at 1:25 AM GMT+1 |

## Psychologist, Research/Experimental

In order to create exceptional products that people will use and appreciate, we need to know about human behavior and about the motivations and influences underlying how and why people do what they do. We believe that all game designers are, in a sense, experimental psychologists. That is why we're looking for an experimental psychologist with superior research skills to apply knowledge and methodologies from psychology to game design and all aspects of Valve's operations. We want to leverage your experience with experimental design, research methods, statistics, and human behavior to help craft even more compelling gameplay experiences for future Valve titles. We'd also expect you to research and weigh in on any and all topics that are relevant to improving the experiences of our customers, partners, and employees.

### Responsibilities

- Providing relevant insight into human behavior in order to shape gameplay and customer experience
- Performing statistical analyses on all aspects of Valve's operations: gameplay, financial, and company data

Exemple de valvesoftware.com

# Des fonctions de recommandations responsables

## Peu de transparence ou d'explication

Une formulation pédagogique qui favorise la prise de recul est pourtant facile à intégrer : « *Les clients qui ont acheté cet article ont également acheté...* » (Amazon)

## Peu de sens des responsabilités des acteurs concernés

Neutralité de la technique, « *Guns don't kill people; people kill people* »

Or, une addiction est stimulée (ou pas) par l'environnement

Grande difficulté à auditer les systèmes de recommandations du fait de l'individualisation du service rendu

# Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads

– Anja Lambrecht (London Business School), Catherine E. Tucker (Massachusetts Institute of Technology MIT)

Une école d'ingénieur diffuse une annonce sur Facebook sans spécifier de cible, et constate qu'elle est plus diffusée auprès des femmes que des hommes

Est-ce parce que l'algorithme maximise la diffusion et que les hommes cliquent plus ?

Non car les femmes cliquent plus sur cette publicité... quand elles la voit.

# Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads

– Anja Lambrecht (London Business School), Catherine E. Tucker (Massachusetts Institute of Technology MIT)

Les publicités en ligne sont attribuées par enchères.

Les annonceurs proposent un prix pour un mot-clé et une cible et lorsque les critères sont vérifiés la plateforme (Facebook, Google, etc.) affiche la publicité de l'annonceur au prix le plus élevé.

Les annonceurs sont plus nombreux à cibler des femmes que des hommes - les femmes sont prescriptrices. *In advertising, men are cheap.*

# Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads

– Anja Lambrecht (London Business School), Catherine E. Tucker (Massachusetts Institute of Technology MIT)

Dans ce cas, il n'y a aucune intention malveillante : ni de l'annonceur, ni du diffuseur

Seulement l'agrégation des préférences des agents combinée à des mécanismes d'enchères

## Bonne pratiques pour éviter cette situation

- ▶ Savoir qu'une intention neutre ne suffit pas; anticiper
- ▶ Diffuser la publicité sur chaque public séparément
- ▶ Audit des plateformes

# For Sale: SAT-Takers' Names. Colleges Buy Student Data and Boost Exclusivity

Score SAT : *Scholastic Assessment Test*, épreuve standardisée à la fin du lycée dont le score sert de base à la décision d'admission par les universités

- The makers of the SAT are licensing personal data about test-takers to schools for 47 cents each, and schools are using it to inflate their acceptance rates, **according to the Wall Street Journal**.
- The College Board uses data from test-takers, like their names, ethnicity, and approximate PSAT and SAT scores, and creates personalized lists that they then sell to schools for recruitment.
- Schools use these lists to recruit students they know will apply but are unlikely to be admitted. By rejecting them, the school can claim a lower acceptance rate, which many see as a signifier of a school's academic prestige.

<https://www.wsj.com/articles/for-sale-sat-takers-names-colleges-buy-student-data-and-boost-exclusivity-11572976621>

# Management par tableau de bord

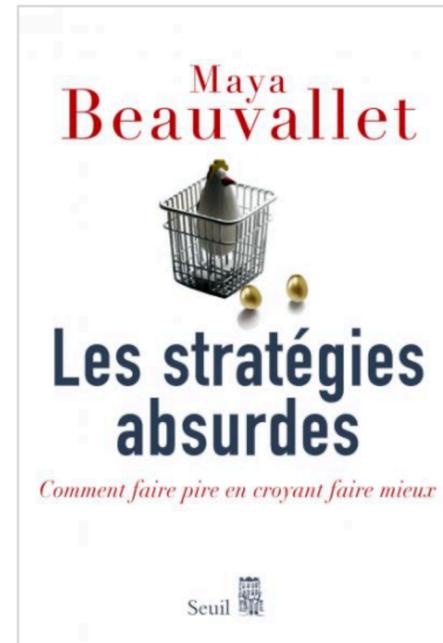
Management par tableau de bord, public comme privé

des standards explicites de mesures de la performance ; les objectifs et les indicateurs de performance sont exprimés quantitativement

Les agents réagissent aux incitations, et se focalisent sur l'objectif

Loi de Goodhart : « lorsqu'une mesure devient un objectif, elle cesse d'être une bonne mesure »

# Management par tableau de bord



VISUEL HD

## Les Stratégies absurdes

Comment faire pire en croyant faire mieux  
Maya Beauvallet

Un club de football met à l'amende un de ses joueurs au motif qu'il rend trop souvent la balle à l'adversaire. Résultat : il ne la passe plus à personne. Un patron décide d'organiser une compétition permanente entre ses salariés. Résultat : une partie d'entre eux commencent à saboter le travail de leurs collègues. Constatant que certains patients victimes de graves complications cardiaques décèdent régulièrement au bloc opératoire, une clinique fixe un quota maximal de « pertes » à ses chirurgiens. Résultat : lorsqu'ils approchent du chiffre fatidique, les chirurgiens refusent d'opérer. Une école décide de sanctionner financièrement les parents dont les enfants arrivent en retard le matin. Résultat : le nombre des retardataires se multiplie...

Le point de départ de ces histoires est presque toujours le même : la nouvelle idéologie managériale et ses méthodes, ses indicateurs de performance, ses dispositifs d'incitation et de sanction. Maya Beauvallet en propose ici à la fois le bêtisier le plus insolite et l'analyse la plus sérieuse.

# Plateforme et *user-generated content*

## Les responsabilités

Dans la presse traditionnelle, pour toute forme de publication : un directeur de la publication est co-responsable avec l'auteur

**Pour les sites web 1.0** : le site dispose d'un directeur de la publication connu de l'hébergeur

- L'hébergeur a le statut d'intermédiaire technique, il ne peut pas intervenir sur le contenu sauf décision de justice
- Il a obligation de préserver les éléments d'identification

**Pour les plateformes** : le titulaire du compte est directeur de la publication et Youtube/Instagram/etc revendique le statut d'intermédiaire technique... donc s'interdit d'intervenir sur le contenu

# Quand les CGU d'une plateforme ont plus d'impact que la loi ...sans se soucier des différences culturelles

**Facebook supprime** des contenus culturels jugés contraires aux bonnes mœurs, comme la Liberté guidant le Peuple Delacroix (1830)



**Mais autorise** la diffusion d'une fusillade de masse en Facebook Live

Justifie le live comme un témoignage en direct et une contribution au débat public (fusillades, violences policières etc.)

Biais Silicon Valley : **toute information améliore le fonctionnement du marché**

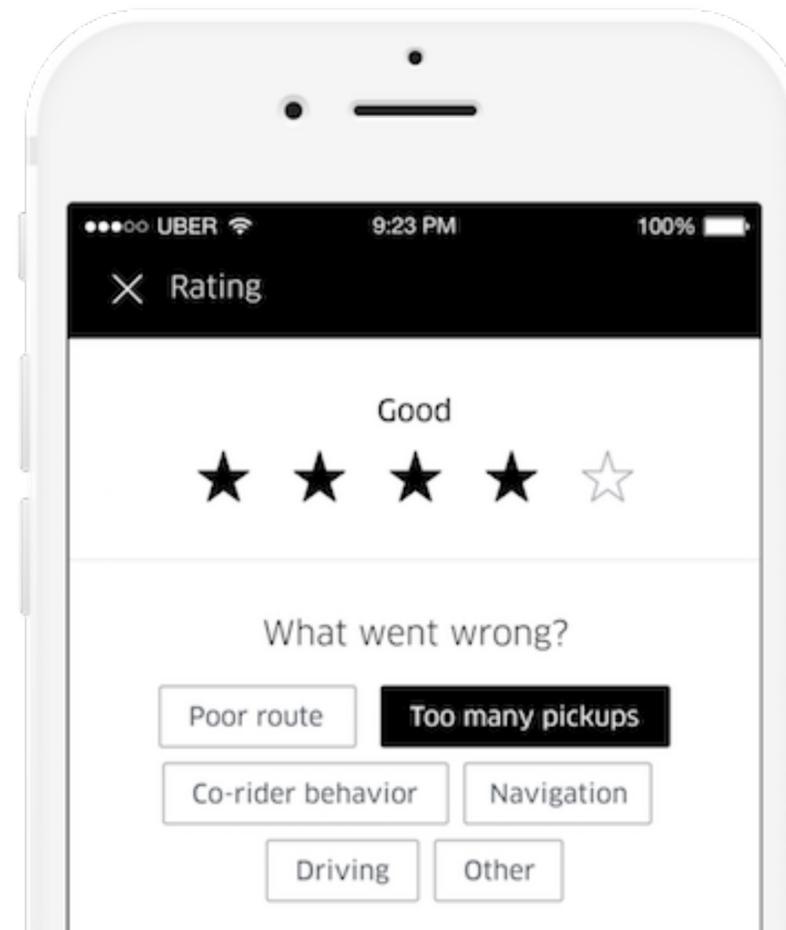


# Notations réciproques sur les plateformes

## Des différences culturelles importantes

Les plateformes utilisent presque systématiquement une notation réciproque entre les participants

Réduction de l'asymétrie d'information pour permettre l'existence du marché - Akerlof (1970)



Ranking system that evenly distributes grades

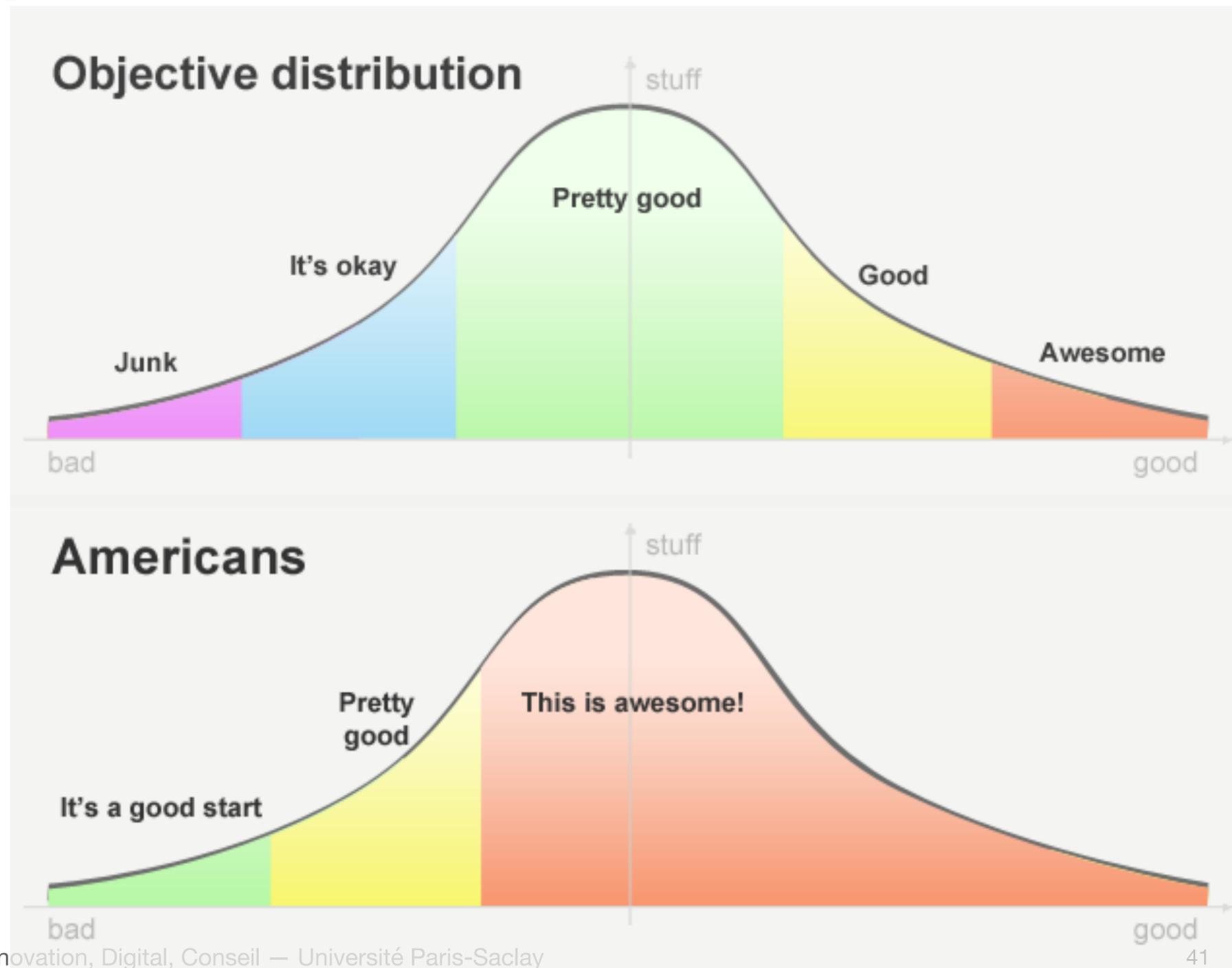
Grade	percentage of students receiving grade
A	top 20%
B	next 20%
C	middle 20%
D	next 20%
F	bottom 20%

# Notations réciproques sur les plateformes

## Des différences culturelles importantes

In America, 99% of people raised here will never ever say “I don’t like what you’ve done”. Instead, their definitions are something like this:

<https://medium.com/octavians-thoughts/what-does-its-a-good-start-really-mean-ff8c1a2495c4>



Illusion de contrôle  
surcharge informationnelle et *bad  
nudge*

# Nudge for good, nudge for bad

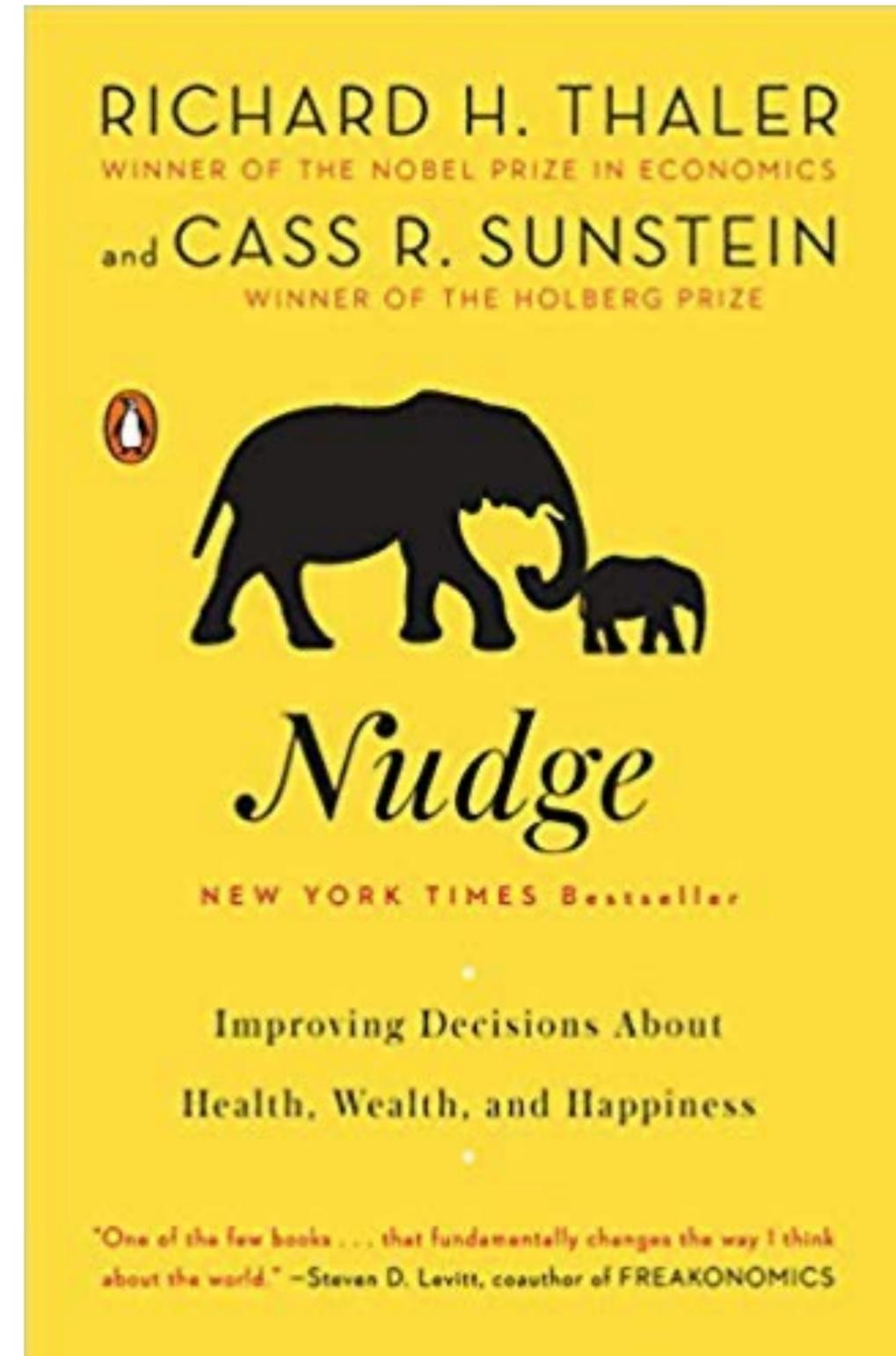
Nudge : modification de l'architecture du choix d'un individu

✓ qui exploite un biais cognitif identifié de façon à lui faire choisir l'issue préférée par les pouvoirs publics

≠ incitation monétaire ≠ réglementation

Ouvrage **Nudge, Improving Decisions About Health, Wealth and Happiness** en 2008 par R. Thaler et C. Sunstein.

Richard Thaler, Nobel d'économie en 2017.



# Nudge for good, nudge for bad

Les comportements ne sont pas toujours rationnels, mais la non-rationalité est prévisible : ce sont les **biais cognitifs**

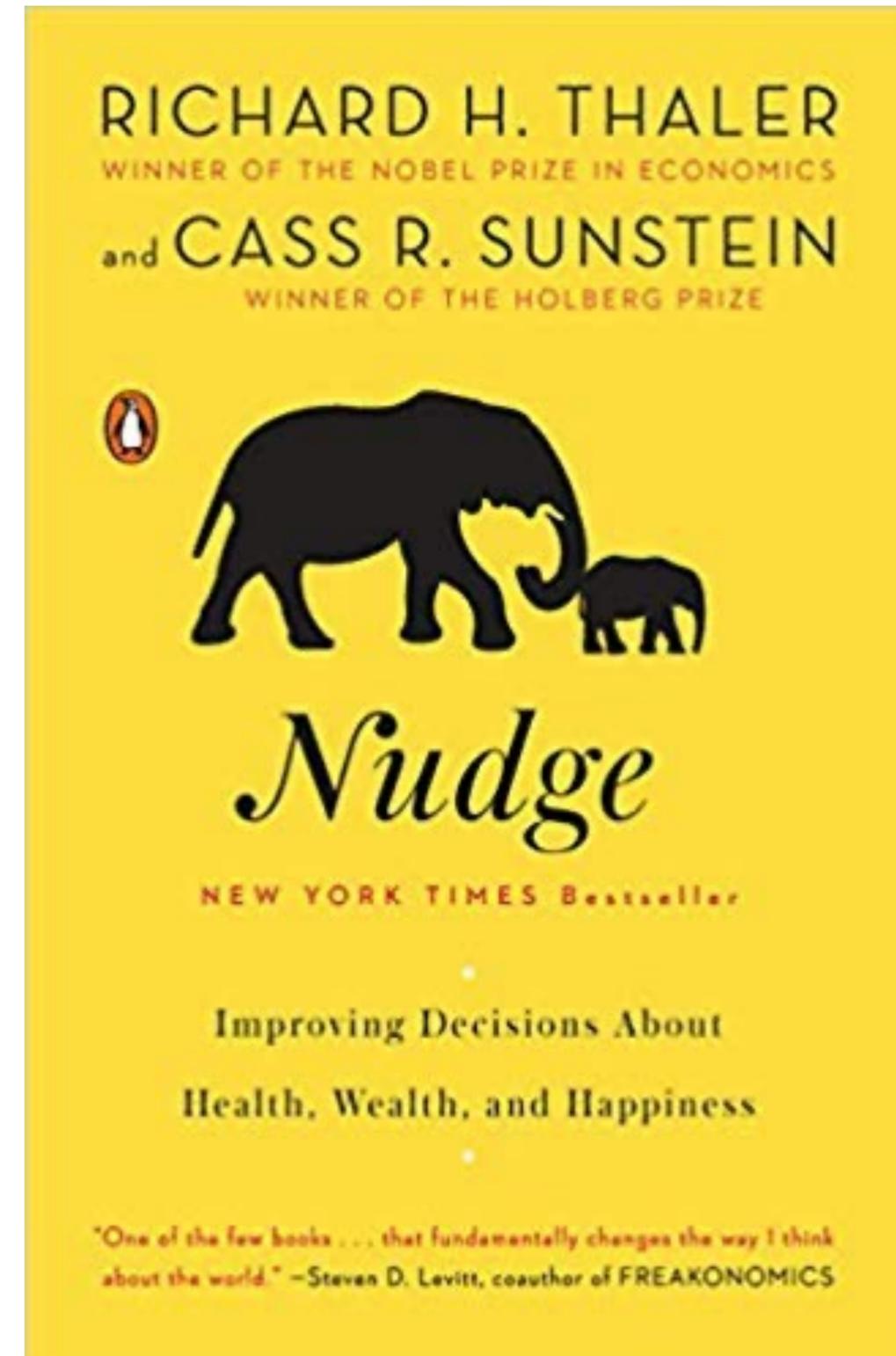
[https://fr.wikipedia.org/wiki/Biais\\_cognitif](https://fr.wikipedia.org/wiki/Biais_cognitif)

Au croisement de la psychologie, des sciences cognitives et de l'économie

Exploitation des biais cognitifs

pour la politique publique : nudge

pour l'influence du consommateur : marketing comportemental

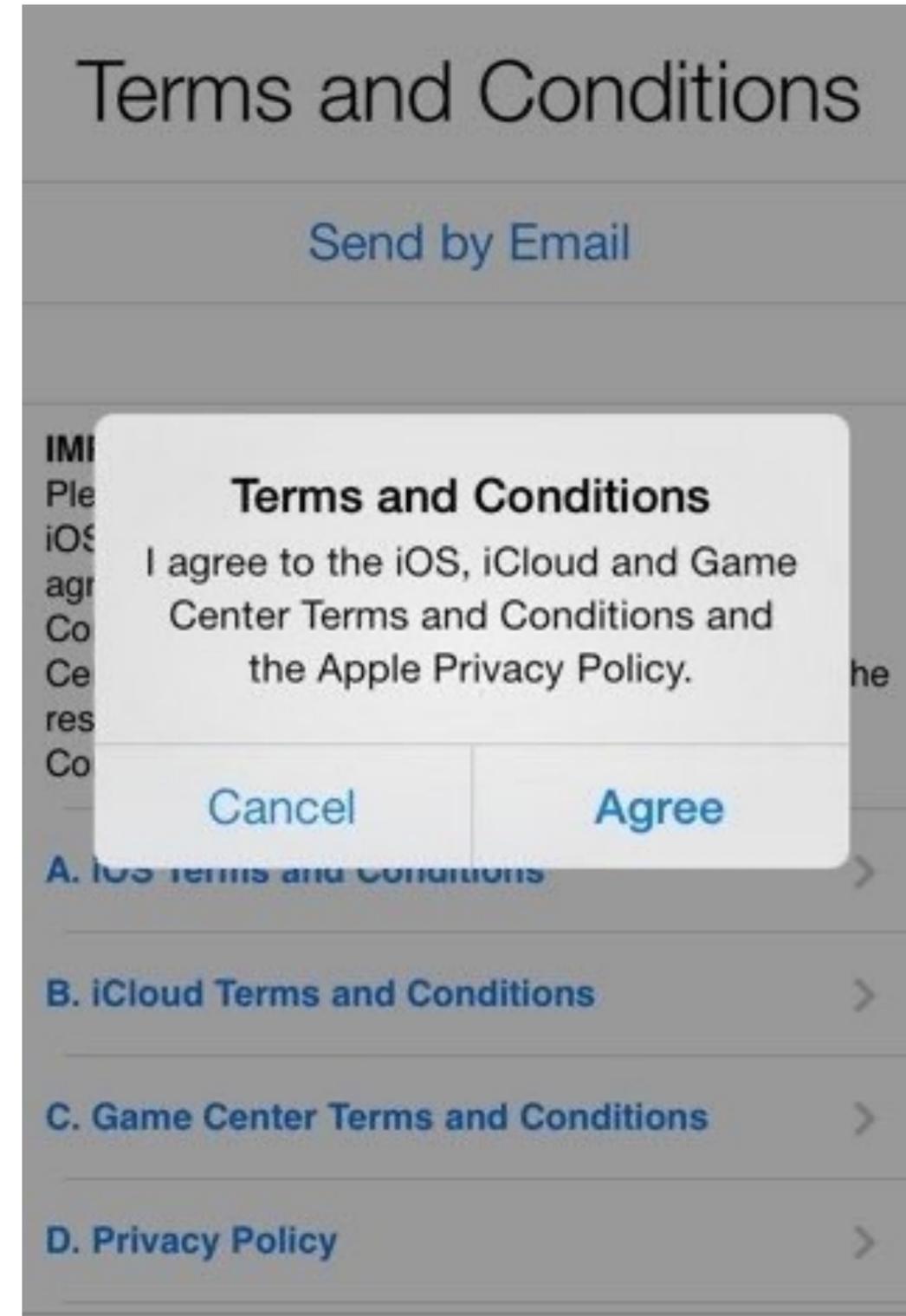


# bad nudges are among us

Boxes checked by default

Option & information preferred by the developer made more salient

Repeated asks

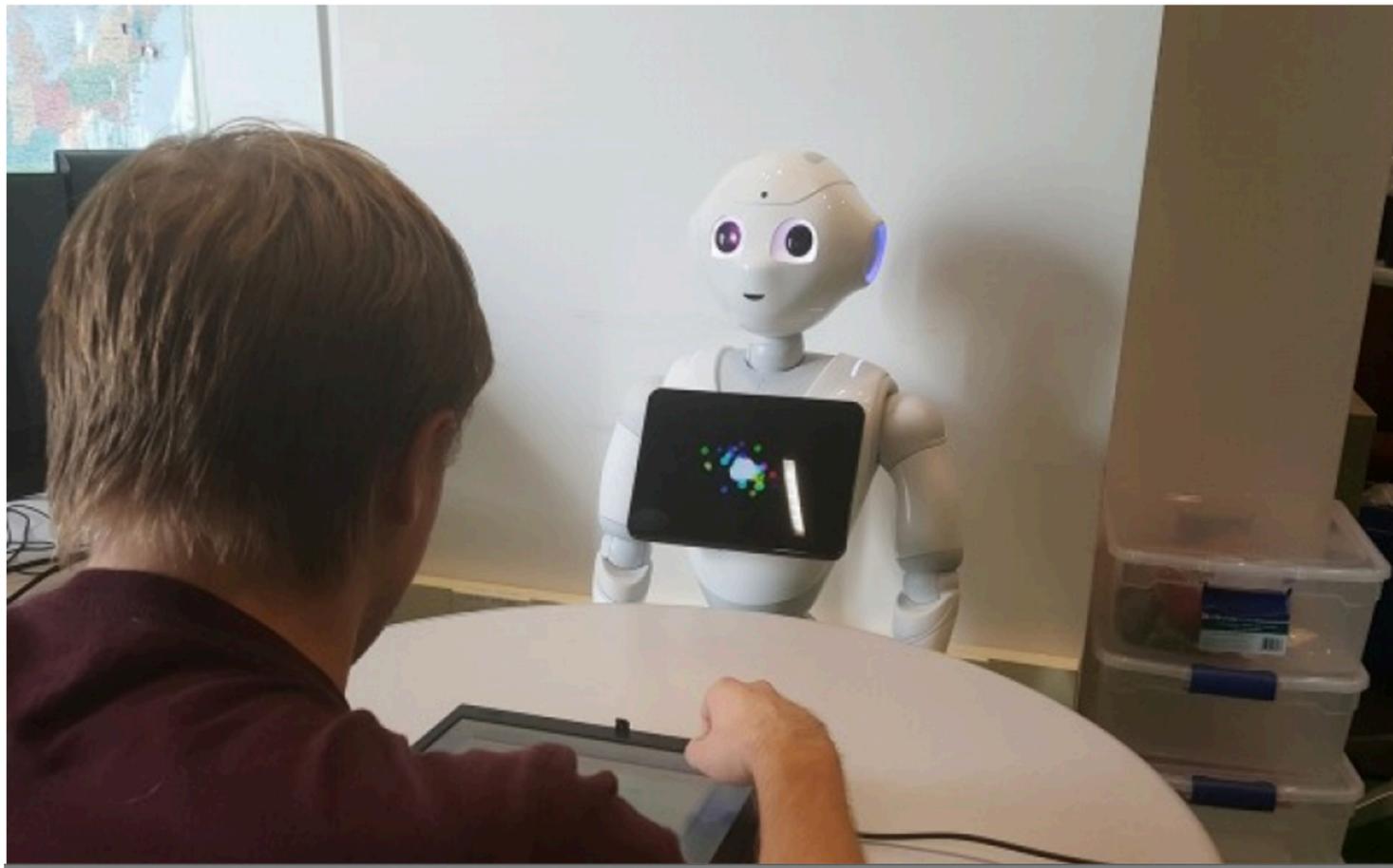




# Un robot parvient à déstabiliser des joueurs en les rabaisant

Lucile Meunier

Like 



En face à face avec un robot humanoïde doté d'IA et prénommé Pepper, quarante personnes ont été invitées à jouer à un jeu de stratégie, à trente-cinq reprises. Lors des parties, les joueurs étaient soit encouragés, soit découragés par la machine. Dans sa bouche, des propos démoralisants comme : « *Je dois te dire que tu es un très mauvais joueur* » ou « *au fur et à mesure de la partie, ton jeu est devenu brouillon* ». Résultat : les humains qui avaient été découragés par le robot ont moins bien joué que les autres.

<https://usbeketrica.com/article/robot-destabilise-rabaisse-joueurs>

# Que vaut le consentement face aux biais cognitifs ?

*Nudges for Privacy and Security:  
Understanding and Assisting Users'  
Choices Online*

– Alessandro Acquisti (Carnegie-Mellon University) and coauthors



## **Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online** – Alessandro Acquisti (Carnegie-Mellon University) and coauthors

1 formulaire sollicitant des informations personnelles, 2 groupes :

Consentement pour chaque question, une case par ligne.

Consentement global , une case pour toute la page.

**Illusion de contrôle** : je suis rassuré si je peux refuser... et j'autorise.

**Surcharge informationnelle** : je valide sans réfléchir.

The image shows a screenshot of the Facebook notification settings interface. On the left, there is a navigation menu with categories: Notifications (selected), Mobile, Public Posts, Apps, Ads, Payments, Support Inbox, and Videos. The main content area is titled 'WHAT YOU'LL RECEIVE' and has three sections: 'WHAT YOU'LL RECEIVE' with three radio button options, 'LIVE VIDEO SETTINGS' with two radio button options, and 'NOTIFICATIONS YOU'VE TURNED OFF' with a list of 18 notification categories, each with a 'Turn On' button.

**WHAT YOU'LL RECEIVE**

- All notifications, except the ones you unsubscribe from
- Important notifications about you or activity you've missed
- Only notifications about your account, security and privacy

**LIVE VIDEO SETTINGS**

- Turn off email notifications about comments added to your live video conversations
- Turn on email notifications about comments added to your live video conversations

**NOTIFICATIONS YOU'VE TURNED OFF**

Notes you're tagged in	Turn On
Comments on your photo albums	Turn On
Comments on your links	Turn On
Comments after you on a photo	Turn On
Comments on photos of you	Turn On
Comments after you on a note	Turn On
Comments after you on a link	Turn On
Comments after you on a photo album	Turn On
Upcoming birthdays	Turn On
You being tagged in a video	Turn On
Comments on your videos	Turn On
Comments on videos of you	Turn On
Likes on your posts	Turn On
Updates on your friends since you last logged in	Turn On
Tags of your videos	Turn On
Friend suggestions	Turn On
Friends added based on your suggestions	Turn On
Comments on stories on your timeline	Turn On
Comments after you on timeline stories	Turn On
Comments after you on a video	Turn On

# Your Roomba already maps your home. Now the CEO plans to sell that map.

Josh Hafner and Edward C. Baig | USA TODAY  
Published 9:06 PM EDT Jul 25, 2017



# Google says the built-in microphone it never told Nest users about was 'never supposed to be a secret'

Nick Bastone 20 Feb 2019, 02:04



- In early February, [Google announced](#) that Assistant would now work with its home security and alarm system Nest Secure.
- The problem — users didn't know a microphone even existed on their Nest security devices to begin with.
- On Tuesday, a Google spokesperson told Business Insider it had made an "error."
- "The on-device microphone was never intended to be a secret and should have been listed in the tech specs. That was an error on our part," the spokesperson said.

# Tile, AirTags

Un effort important est consacré à anticiper et limiter les dérives liées aux objets connectés

**AirTag is designed to discourage unwanted tracking.** If someone else's AirTag finds its way into your stuff, your iPhone will notice it's traveling with you and send you an alert. After a while, if you still haven't found it, the AirTag will start playing a sound to let you know it's there.

Of course, if you happen to be with a friend who has an AirTag, or on a train with a whole bunch of people with AirTag, don't worry. These alerts are triggered only when an AirTag is separated from its owner.

Check out

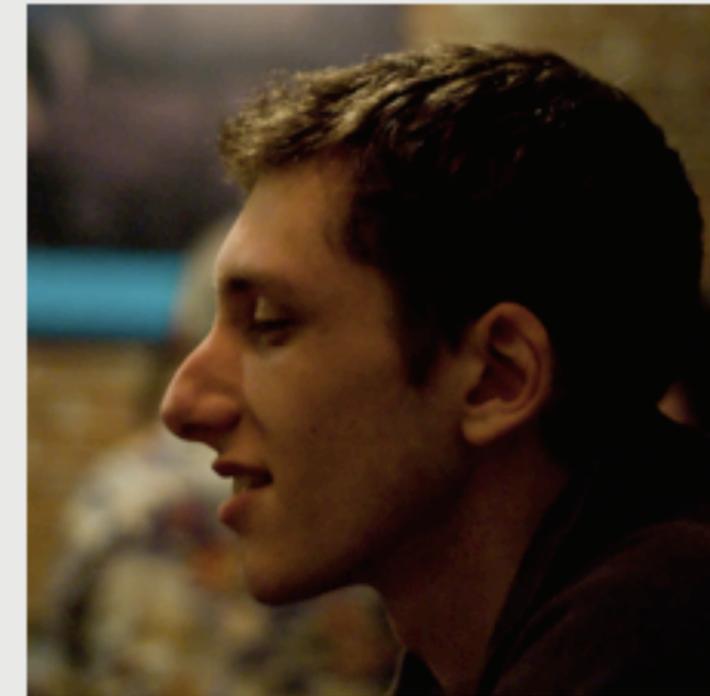
08/20/2008

Guests

1

Search

# Présentation de l'information, mise en page et discrimination



Modern - \$99



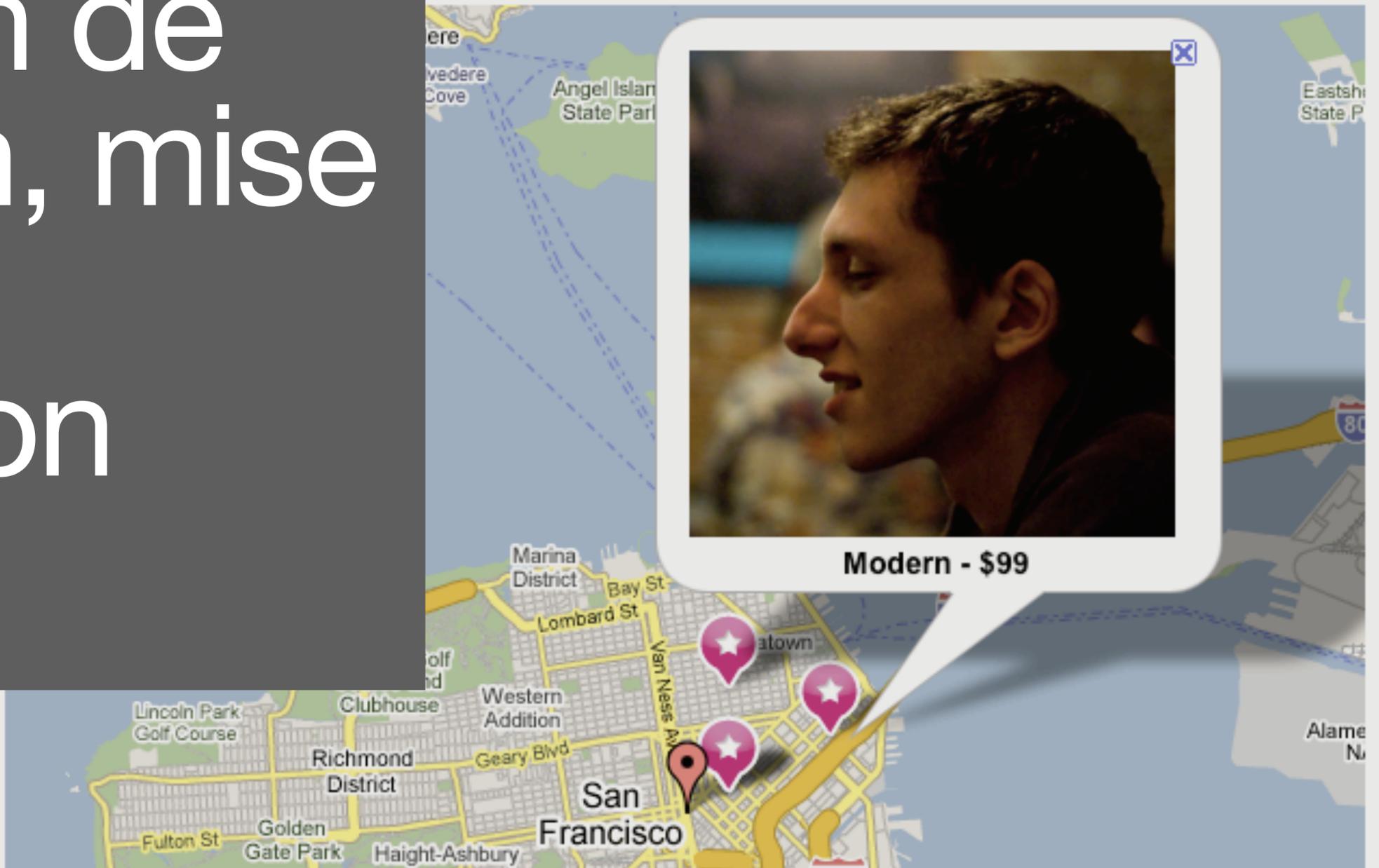
"Get Fit!"

\$80



"Sunny"

\$85



# Économie du partage et discrimination

Dans l'économie du partage, la discrimination à l'encontre des minorités est fréquente

Airbnb : Les propriétaires qui sont noirs doivent baisser leur prix de 14 % par rapport aux blancs afin d'obtenir la même demande. Ce résultat tient compte de la différence d'emplacement.

Blablcar : l'allocation est défavorable aux chauffeurs maghrébins

Résultats identiques sur les marchés de prêts en pairs-à-pairs

Airbnb users pay more to stay with attractive hosts, The Economist, 30 Dec 2019. <https://www.economist.com/graphic-detail/2019/12/30/airbnb-users-pay-more-to-stay-with-attractive-hosts>

Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. Benjamin Edelman, Michael Luca and Dan Svirsky . American Economic Journal: Applied Economics, 2017)

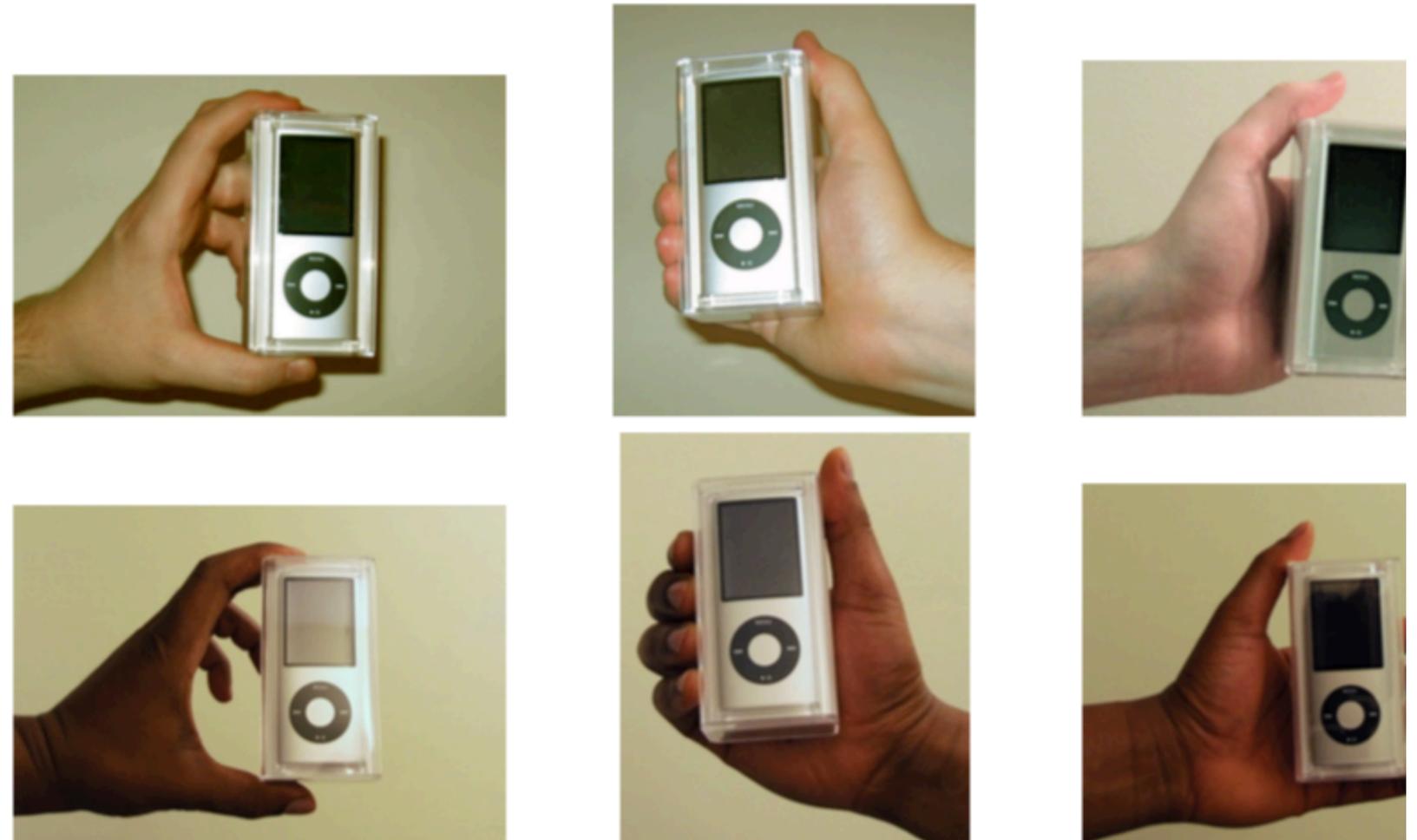
What drives pricing behavior in Peer-to-Peer markets? Evidence from the carsharing platform blablacar. Mehdi Farajallah, Robert G. Hammond, ThierryPénard. <https://doi.org/10.1016/j.infoecopol.2019.01.002>

# ***The Visible Hand: Race and Online Market Outcomes***

– Jennifer L. Doleac (University of Virginia) Luke C.D. Stein (Arizona State University),

The Economic Journal 123:572  
(November 2013)

**Figure 1:** Advertisement photographs



# Présentation de l'information, mise en page et discrimination

Dans l'économie du partage, la relation entre offreurs et acheteurs n'est pas (seulement) commerciale

- Les pratiques sont nouvelles, encore peu de routines
- Pas de droit de la consommation

➔ Il faut donc créer de la confiance

**Problème : les stéréotypes et la discrimination sont des moyens efficaces d'échanger entre interlocuteurs de confiance...**

# Présentation de l'information, mise en page et discrimination

L'offreur transmet des informations sur lui-même qui peuvent être utilisées par l'acheteur

La plateforme **reflète** la discrimination qui existe off-line...

- est-ce qu'elle la **facilite** ?
- est-ce qu'elle la **limite** ?

# Présentation de l'information, mise en page et discrimination

## Les suggestions de solutions

- Une fonction d'achat immédiat
  - Afin de se ramener le plus possible à une relation d'affaires classique
- Par la mise en page du site, mettre l'accent sur le bien lui-même plutôt que sur le propriétaire

# Airbnb en 2008

## Lors de son lancement



[Sign Up / Sign In](#) | [Post a Room](#)

san francisco

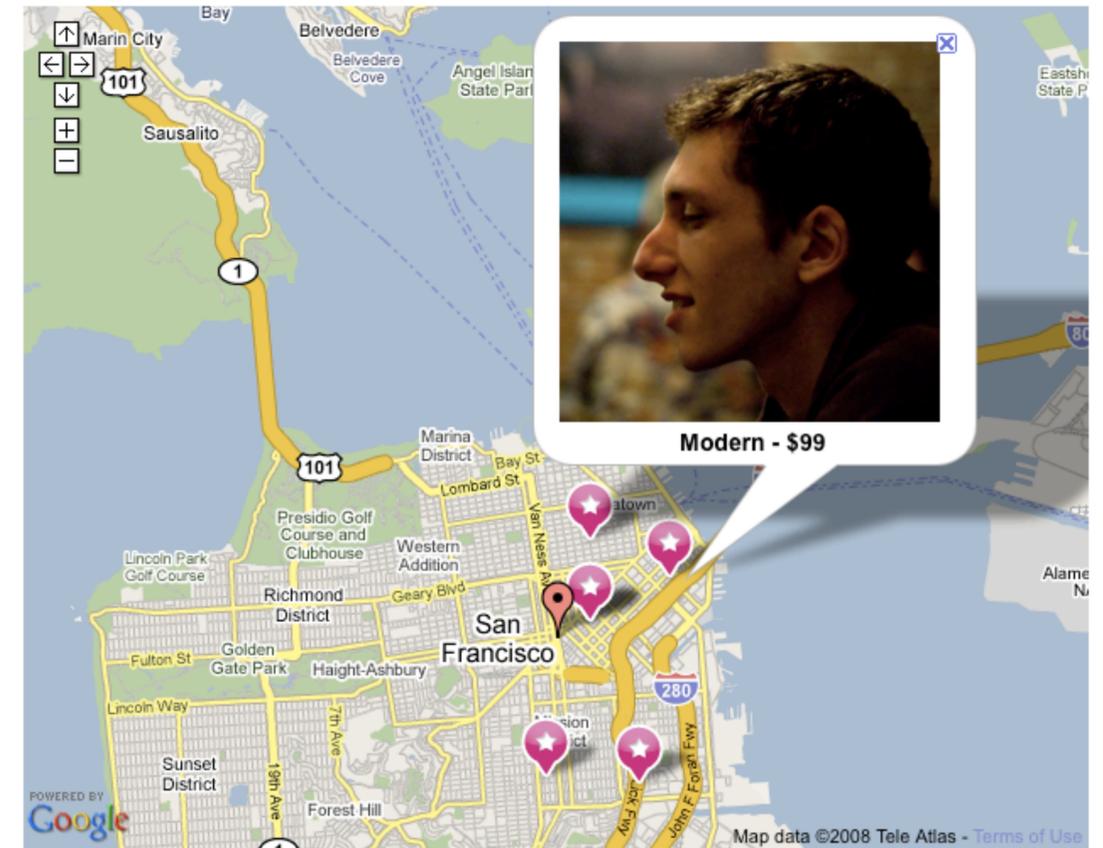
Check in: 08/18/2008    Check out: 08/20/2008    Guests: 1

[Sorting Options](#)    [Search](#)

1-10 out of 16 listings.

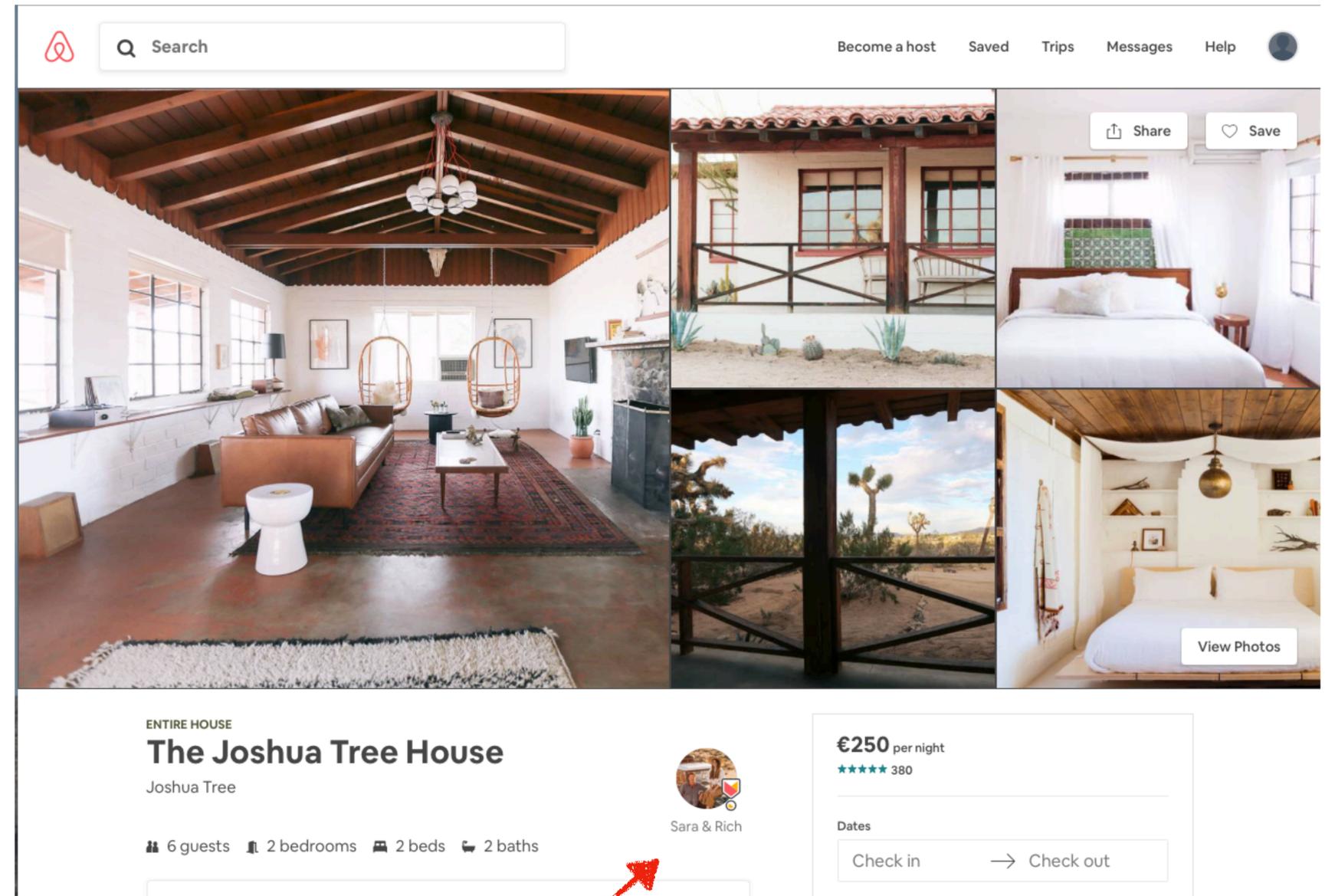
	"tikitastic"	\$10
	"Cute!"	\$25
	"Simple"	\$50
	"AwesomeRoom"	\$50
	"happy"	\$50
	"Victorian"	\$50
	"Get Fit!"	\$80
	"Sunny"	\$85
	"Cute"	\$90
	"Modern"	\$99

1 2 Next



# Airbnb en 2018

Après une réflexion sur le sujet,  
inhabituelle chez les GAFA



# Éthique de l'expérimentation de terrain

- ✓ L'innovation porte aussi sur l'organisation & les pratiques managériales
  - ✓ Son efficacité dépend des conditions de mise en œuvre (cas par cas)
- Besoin d'évaluer

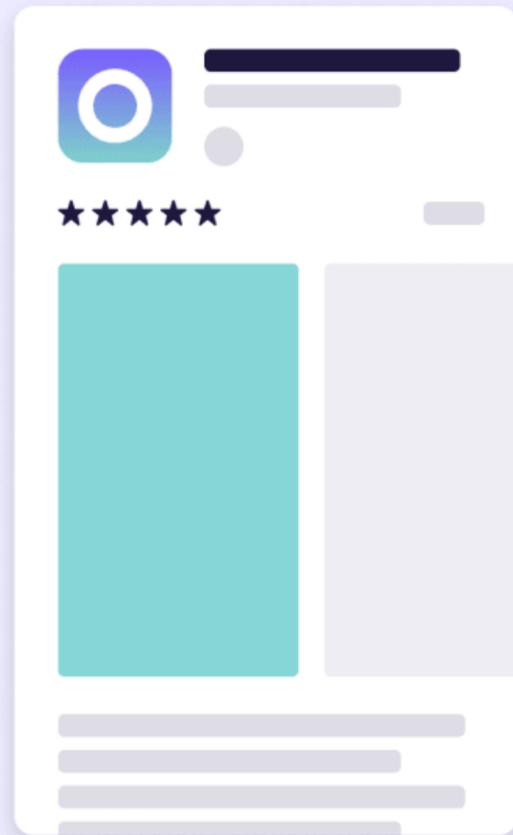
## Pourquoi faut-il expérimenter ?

# Mobile A/B Testing

Variation A

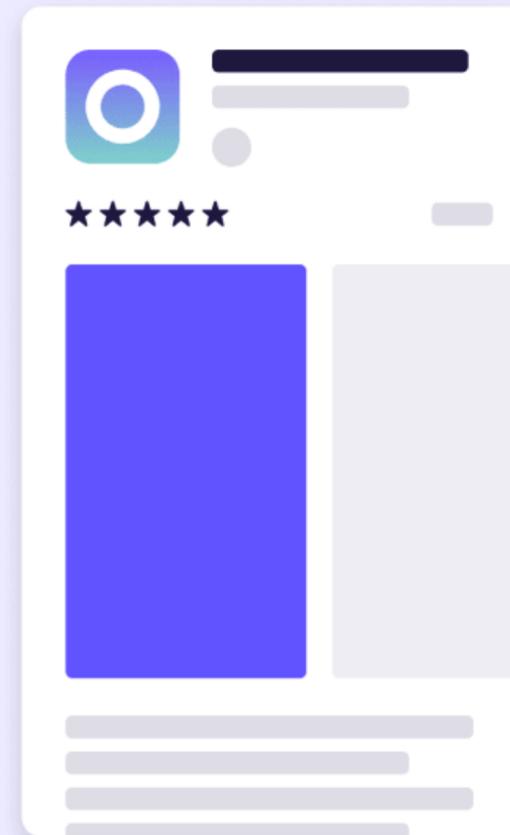
Variation B

 →  
50% visitors  
see variation A



↓  
**15% Conversions**

←   
50% visitors  
see variation B



↓  
**36% Conversions**



# La méthodologie de référence

La démarche expérimentale (test A/B) est issue de la médecine :

On constitue un **groupe de contrôle** puis un **groupe traité** qui subit une (et une seule) modification par rapport au contrôle.

Important : groupe traité et groupe de contrôle doivent être identiques en moyenne. Sinon, il existe un biais de sélection.

- D'où la nécessité d'une attribution **aléatoire** d'un individu donné à un groupe.

---

**De plus, afin d'éviter un effet placebo un individu ne doit pas savoir à quel groupe il appartient... voire ne pas savoir qu'il est observé**

---

# Tous des cobayes

Overview

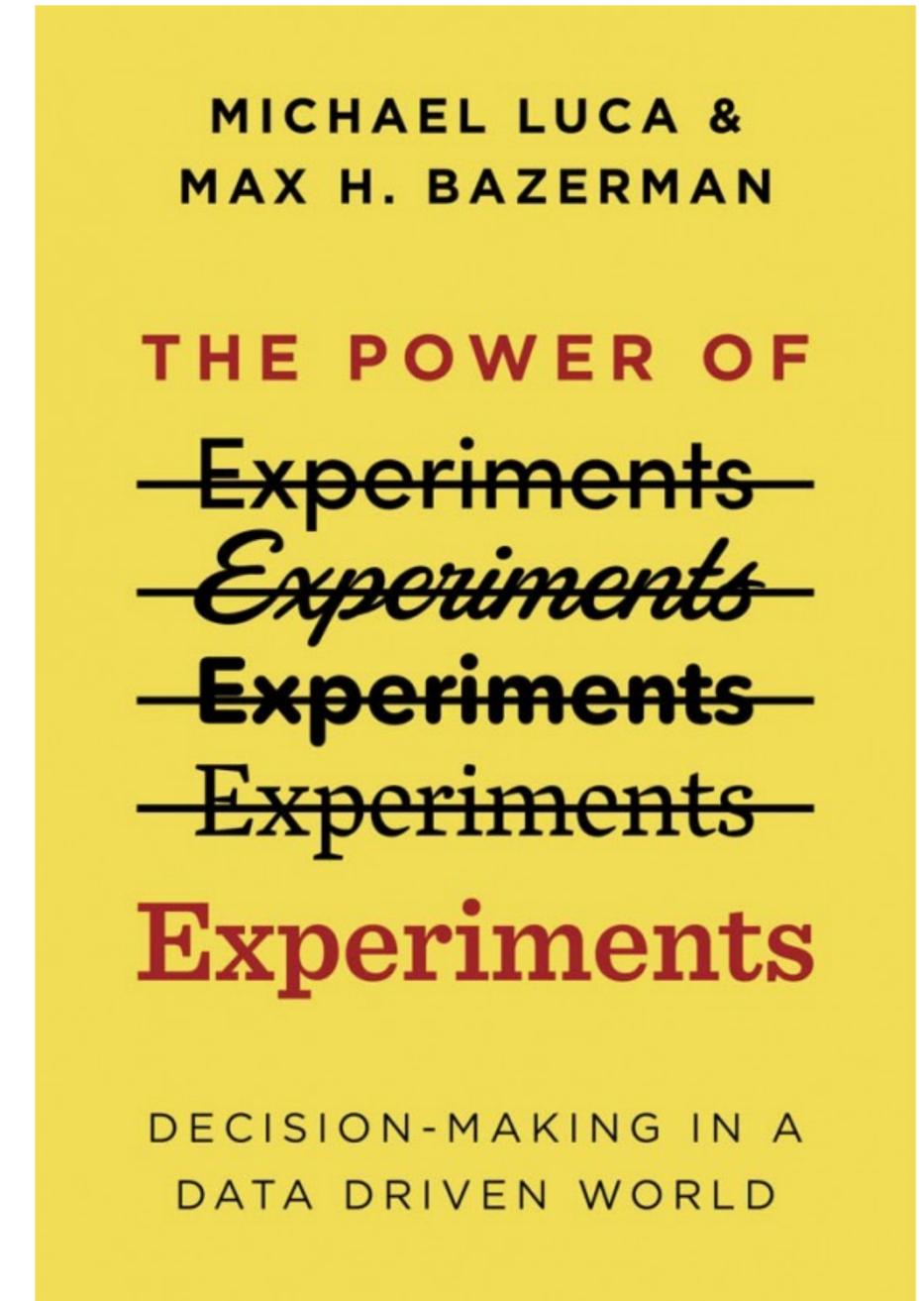
Author(s)

Praise

## Summary

How organizations – including Google, StubHub, Airbnb, and Facebook – learn from experiments in a data-driven world.

Have you logged into Facebook recently? Searched for something on Google? Chosen a movie on Netflix? If so, you've probably been an unwitting participant in a variety of experiments – also known as randomized controlled trials – designed to test the impact of different online experiences. Once an esoteric tool for academic research, the randomized controlled trial has gone mainstream. No tech company worth its salt (or its share price) would dare make major changes to its platform without first running experiments to understand how they would influence user behavior. In this book, Michael Luca and Max Bazerman explain the importance of experiments for decision making in a data-driven world.



# Optimisation ergonomique

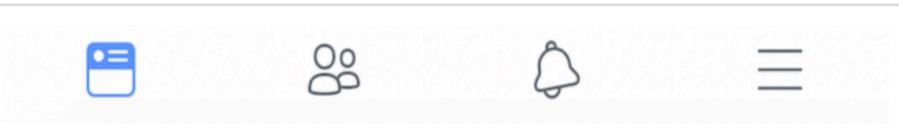
## MOTHERBOARD

DESIGN | By Louise Matsakis | Oct 17 2017, 8:23pm

### There Are Over 60 Different Versions of the Facebook App

Does a tiny tweak in design change user behavior?

These differences exist because it appears Facebook is aggressively testing its mobile navigation bar, or the series of tabs that run along the bottom of the app. There are over 60 distinct versions of the feature worldwide, according to a crowdsourced [spreadsheet](#) created last week by digital designer Luke Wroblewski.

	Oct 13	1	iOS 11.0.3 iPad 10.5, Ukraine
	Oct 13	13	iOS 11.0.1, iPhone 7 plus, Malaysia

[https://motherboard.vice.com/en\\_us/article/8x8wp4/why-navigation-bar-looks-different-facebook](https://motherboard.vice.com/en_us/article/8x8wp4/why-navigation-bar-looks-different-facebook)

# Dérives dans l'expérimentation de terrain

“Facebook published the results of a 2012 study in the Proceedings of the National Academy of Sciences. Unbeknown to users, Facebook had tampered with the news feeds of nearly 700,000 people, showing them an abnormally low number of either positive or negative posts. The experiment aimed to determine whether the company could alter the emotional state of its users”

<https://www.theguardian.com/technology/2014/oct/02/facebook-sorry-secret-psychological-experiment-users>

# ***Toward an understanding of the economics of apologies: evidence from a large-scale natural field experiment***

Basil Halperin, Benjamin Ho, John A. List,  
Ian Muir

**Abstract:** We use a theory of apologies to analyze a nationwide field experiment involving 1.5 million Uber ridesharing consumers who experienced late rides. Several insights emerge. First, apologies are not a panacea: the efficacy of an apology and whether it may backfire depend on how the apology is made. Second, across treatments, money speaks louder than words – the best form of apology is to include a coupon for a future trip. Third, in some cases sending an apology is worse than sending nothing at all, particularly for repeated apologies. For firms, caveat venditor should be the rule when considering apologies.

**Keywords:** apologies, ridesharing, field experiment

**JEL Classification Numbers:** D80, D91, Z13

---

\*This manuscript was not subject to prior review by any party, as per the research contract signed at the outset of this project. The views expressed here are solely those of the authors. Thanks to seminar participants at AFE 2017, AEA 2018, and Williams College, and to Courtney Rosen for helpful comments and assistance. AEA Registry number: AEARCTR-0002342. All errors are our own.

# L'évaluation des politiques publiques

## Une forte réticence juridico-culturelle

La démarche d'expérimentation aléatoire est nécessaire afin d'évaluer rigoureusement l'efficacité d'une action

Elle suppose un traitement différencié dans un temps limité

Obstacle juridique : Dans le cadre des administrations, c'est par définition une **rupture d'égalité** de traitement des usagers

Culturellement, les administrations sont très sensibles à cet argument

# L'évaluation des politiques publiques

Rapport du Conseil d'État en 2019 : « Les expérimentations : comment innover dans la conduite des politiques publiques ? »

Le Conseil d'État pointe notamment le manque de culture de l'expérimentation et de rigueur méthodologique

- **insuffisamment construit suivant les méthodes expérimentales rigoureuses.** La grande majorité des expérimentations n'ont pas recours à l'expérimentation aléatoire (dite parfois « randomisée »), qui consiste à comparer deux groupes formés aléatoirement à partir d'un échantillon d'individus (cf. 2.1). La méthode plus souple, consistant à comparer la situation d'un échantillon avant et après l'expérimentation, ne s'accompagne pas toujours de la simple collecte des données statistiques pertinentes, en amont de l'expérimentation, nécessaire pour dresser un diagnostic d'entrée et permettre l'évaluation des effets de la mesure testée.

<https://www.conseil-etat.fr/actualites/actualites/ameliorer-et-developper-les-experimentations-pour-des-politiques-publiques-plus-efficaces-et-innovantes>

The background of the slide is a photograph of the interior of a grand, ornate dome, likely the Crystal Palace in London. The dome is covered in a complex network of stained glass panels, featuring intricate floral and geometric patterns in shades of blue, gold, and red. The structure is supported by a series of dark, arched ribs that create a sense of depth and scale. The lighting is soft and diffused, highlighting the textures and colors of the glass. The overall atmosphere is one of historical grandeur and architectural complexity.

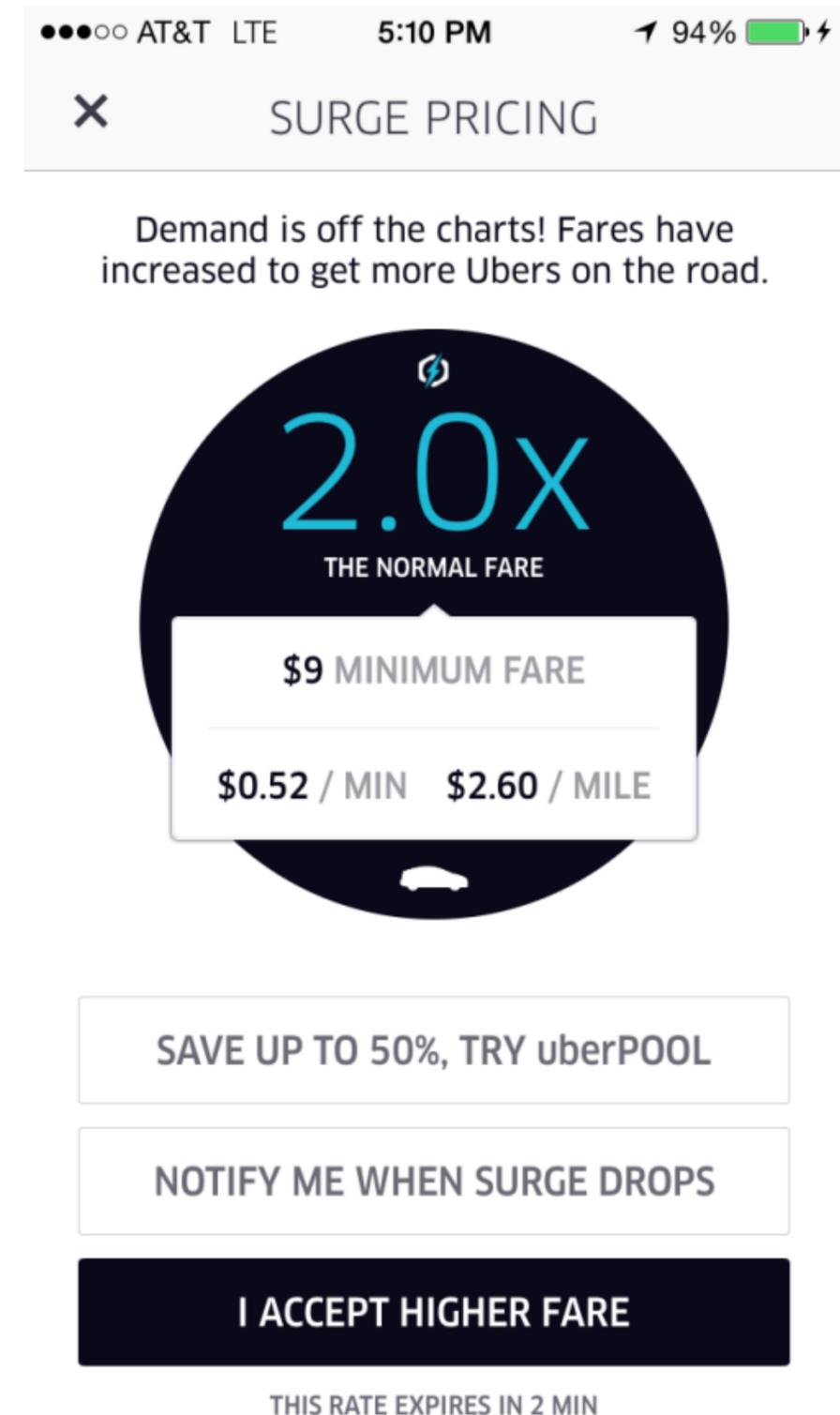
Tarification personnalisée :  
la fin de l'affichage des prix ?

# Tarification dynamique selon la demande

En cas d'orage ou d'événements type concerts, les prix augmentent... scandale ?

Côté demande : capter la disposition à payer y compris pure extraction (= faiblesse du client)

Côté offre : attirer les chauffeurs



# Tarification personnalisée

Forbes

Billionaires

Innovation

Leadership

Money

Consumer

Industry

41,921 views | May 25, 2016, 06:30am

## Uber: Users Are More Likely To Pay Surge Pricing If Their Phone Battery Is Low



Amit Chowdhry Contributor ⓘ

Tech enthusiast, born in Ann Arbor and educated at Michigan State



UBER SURGE PRICING SURGE PRICING

PICKUP LOCATION  
29-49 3rd Street

CONFIRMATION  
Type (2.0) to confirm your fare multiple.

MY FARE WILL BE  
2.0  
TIMES THE NORMAL FARE

2.0x  
THE NORMAL FARE  
\$9 MINIMUM FARE  
\$0.52 / MIN \$2.60 / MILE

SAVE UP TO 50%, TRY uberPOOL  
NOTIFY ME WHEN SURGE DROPS  
I ACCEPT HIGHER FARE  
THIS RATE EXPIRES IN 2 MIN

1	2 ABC	3 DEF
4 GHI	5 JKL	6 MNO
7 PQRS	8 TUV	9 WXYZ
	0	✕

# Allocation des ressources

## Comment attribuer 1 Uber pour 6 personnes sous la pluie ?

Par le marché : exclusion par le prix

Enchères

Tirage au sort

File d'attente : premier arrivé, premier servi

Allocation administrative

**Le marché est la pire méthode d'allocation des ressources... à l'exclusion de toutes les autres**



# Algorithmes, open data et justice prédictive

Décision par algorithme

# Décision par algorithme

*Code is law* (Lessig, 2000) et *Law is code*

Les algorithmes permettent avant tout d'automatiser la mise en œuvre de la réglementation

Radars routiers, grilles de critères dans les appels d'offres, accès au crédit, attribution de logements sociaux ou de places dans le supérieur...

**Fin de la corruption, de l'arbitraire, de la tolérance**



*connotations différentes pour une même idée*

# Les algorithmes 1.0 : l'automatisation

Simple automatisation : ce n'est plus un policier qui dresse le PV mais le radar

Il s'agit alors d'automatiser la mise en œuvre, pas la conception

Le positionnement reste déterminé par l'intérêt général et pas par la maximisation des recettes

# La décision par scoring

## Avantages du scoring

Objectivisation des décisions, nécessité de justifier un écart à la suggestion mathématique

Moins de favoritisme

## Limites du scoring

Sentiment d'objectivité qui rend difficile de s'en écarter, framing effect

Transfert de la discrimination de la décision vers l'accès à l'information sur les critères

# Décision par algorithme

À réglementation inchangée, la mise en œuvre est automatisée

- Quel droit à quelle information sur l'automatisation ? Équilibre délicat :
  - Aucune transparence : décision par boîte noire...
  - Transparence totale (code source) : décision incompréhensible, sauf pour quelques initiés
  - Transfert de la discrimination vers l'accès à l'information sur les critères
- Une part importante tient aux modalités de mise en œuvre !

# **Code is law : importance (et limite) de la transparence**

## **connaître la réglementation c'est connaître l'algorithme**

Les associations demandent la diffusion du code source de l'algorithme de ParcoursSup et non le seul cahier des charges

- Cahier des charges : objectifs, intentions et principes, exprimés en français
- Code source : l'implémentation technique

# L'arbitrage efficacité — explicabilité

## L'algorithme le plus efficace est-il le meilleur ?



Admission Post-Bac demandait aux lycéens 3 vœux classés

- Cette information améliore le matching entre lycéens et établissements
- Mais elle n'était pas comprise, ou pas connue, des candidats



ParcourSup retourne une réponse oui/oui si/non pour chaque candidature

Le candidat effectue son choix après avoir vu les résultats (c'est objectivement inutile, mais psychologiquement apprécié)

# Automatisation et faute

La notion de faute disparaît, remplacée par une mesure de la situation ou du comportement

Vous résidez \*

<input checked="" type="checkbox"/> En France métropolitaine	<input type="checkbox"/> En Guyane	<input type="checkbox"/> A Mayotte
<input type="checkbox"/> Dans un autre DOM ?	<input type="checkbox"/> Hors de France	

Votre composition familiale \*

<input checked="" type="checkbox"/> Seul(e) ?	<input type="checkbox"/> Seul(e), avec un enfant ?	<input type="checkbox"/> Seul(e), avec deux enfants ?
<input type="checkbox"/> Seul(e), avec plus de deux enfants ?	<input type="checkbox"/> En couple ?	<input type="checkbox"/> En couple, avec un enfant ?
<input type="checkbox"/> En couple, avec deux enfants ?	<input type="checkbox"/> En couple, avec plus de deux enfants ?	

Votre situation professionnelle

Vous êtes travailleur indépendant ? \*

<input type="checkbox"/> Oui	<input checked="" type="checkbox"/> Non
------------------------------	---

## In Amazon Go, no one thinks I'm stealing

Commentary: I grew up learning that I would be targeted when I shop because of implicit bias. But what if you took away store employees?

BY ASHLEE CLARK THOMPSON | OCTOBER 26, 2018 9:18 AM PDT

# En finir avec la relation de guichet

Dans le cadre d'une démarche administrative :

- L'études pièces en amont
- Leur validation ou rejet automatisé
  - ➔ Fait disparaître la relation administrative traditionnelle (Demande / vérification / fraude)
  - ➔ Fait disparaître le non-recours

Service-Public.fr PAPIERS

Carte d'identité / Passeport

## Quelle photo ?

**QUALITE**

- Photo nette
- Sans pliure
- Ni trace

**DIMENSIONS**

- Taille de la photo : 3,5 cm (largeur) 4,5 cm (hauteur)
- Taille du visage : entre 3,2 et 3,6 cm (soit 70 à 80 % de la photo), du bas du menton au sommet du crâne (hors cheveux)

**CONTRASTES**

- Pas de surexposition ni sous-exposition
- Pas d'ombre
- Contrastes corrects

**COULEURS**

- Fond uni
- Couleur claire mais pas de fond blanc

**EXPRESSION**

- Tête droite et regard face à l'objectif
- Expression neutre
- Bouche fermée

# La loi donne le droit à un regard humain

## ...à quand un droit à l'algorithme ?

Article 10 de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés :

*« Aucune décision de justice impliquant une appréciation sur le comportement d'une personne ne peut avoir pour fondement un traitement automatisé de données à caractère personnel destiné à évaluer certains aspects de sa personnalité.*

*Aucune autre décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé de données destiné à définir le profil de l'intéressé ou à évaluer certains aspects de sa personnalité. (...) »*

L'open data — données ouvertes

« Il y aura presque toujours plus d'intelligence,  
de données, plus d'imagination et de créativité  
à l'**extérieur** qu'à l'intérieur d'une organisation »

Nicolas Colin, Henri Verdier, **L'âge de la multitude. Entreprendre et gouverner après la révolution numérique** (2015)

# L'ouverture des données

## Un changement radical

**Principe :** Publication des données issues de l'activité de l'établissement, en format réutilisable et avec droits de réutilisation

Donc pas de PDF-image, pas de diffusion sélective à des interlocuteurs choisis, pas de revente des données comme source de revenu

# L'ouverture des données

## Un changement radical

Remise en cause de beaucoup de pré-supposés : Les producteurs de données (administrations) ne sont pas les mieux placés pour les exploiter

Exploiter des données =

les analyser, les mettre en forme, les diffuser, en tirer les conséquences

= entrepreneuriat

# L'ouverture des données

## Un changement radical

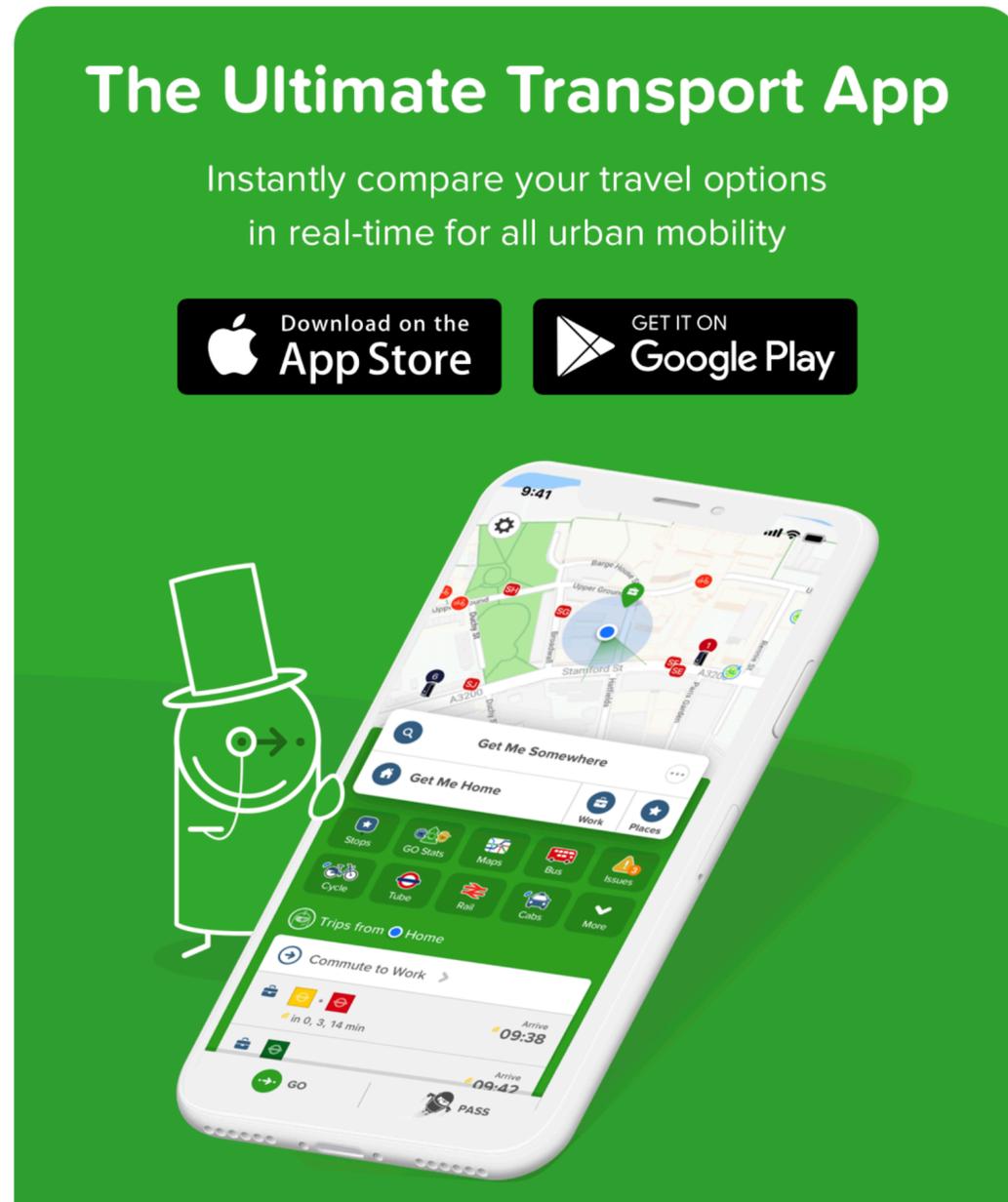
L'analyse des données devient centrées sur **l'utilisateur** et non pas le producteur

Et donc **explicitation** d'éléments sociaux (performance des services publics, délinquance...)

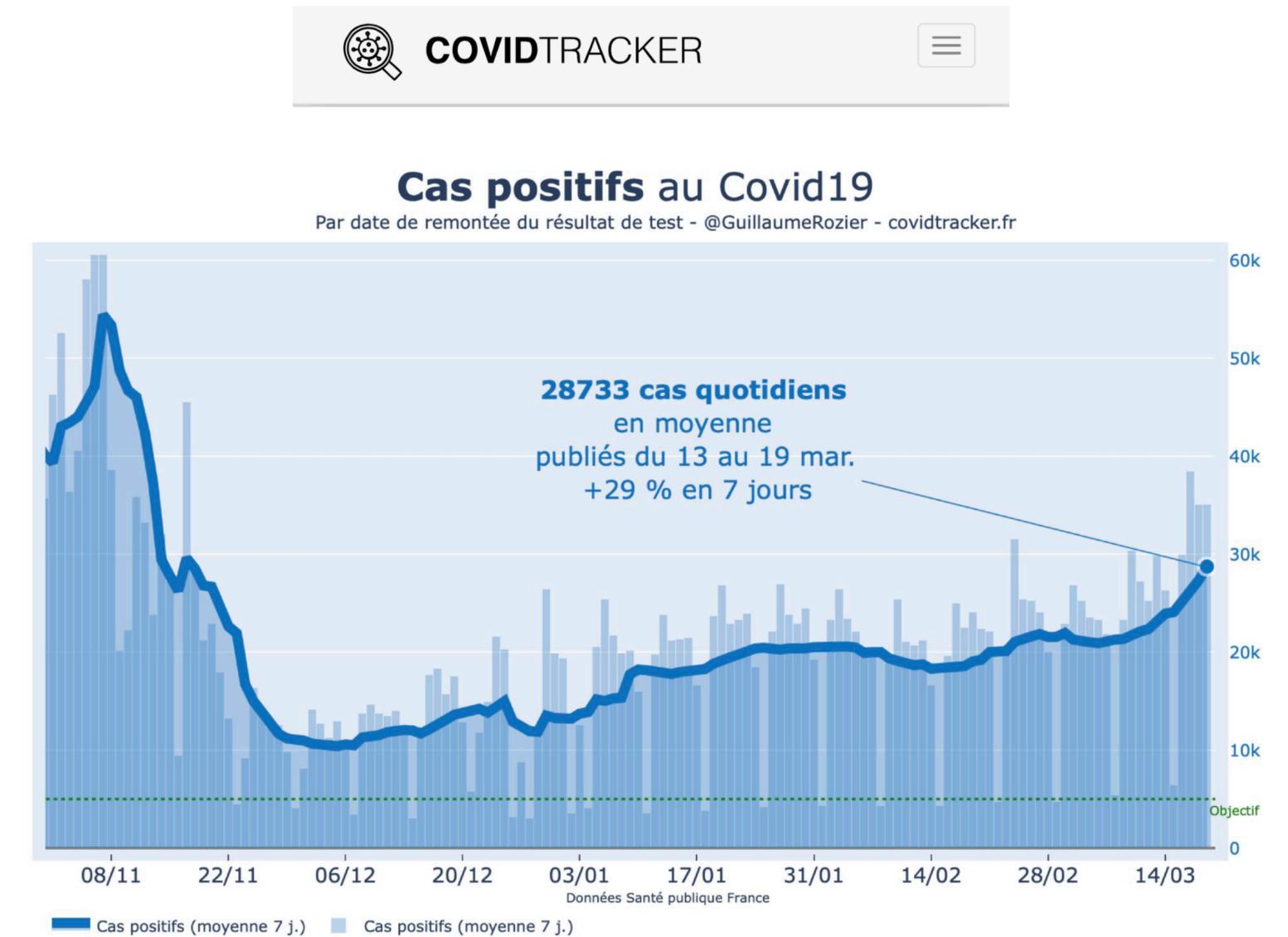
- L'information n'est plus réservée aux initiés
- Risque d'effets auto-réalisateurs

# Open data et entrepreneuriat

Citymapper



CovidTracker & VaccineTracker



# Open data et entrepreneuriat

Éducation nationale et parcours des élèves,  
classement des établissements

data.education.gouv.fr

Bienvenue sur data.education.gouv.fr — Éducation Nationale - Accueil

Merci à toutes et à tous pour votre participation, vous étiez 206 à contribuer à l'amélioration de votre plateforme data.education ! Les évolutions apparaîtront au fur et à mesure dans les prochaines semaines.

RÉPUBLIQUE FRANÇAISE  
Liberté  
Égalité  
Fraternité

data.education.gouv.fr

Connexion

Données Data-visualisations Démarche Créer une carte Créer un graphique Nous contacter

## Plateforme ouverte des données Éducation, Sports et Jeunesse

Porté par le Ministère de l'Éducation nationale, associé au Ministère des Sports et des jeux olympiques et paralympiques, ce portail se positionne pour une **gouvernance toujours plus transparente, et au service des usagers.**

### Éducation

Données du Ministère de l'Éducation nationale

### Sports

### Comprendre les données

Pour un gouvernement toujours plus transparent, des outils de visualisations sont à votre disposition pour vous permettre de vous approprier les données !

Le Monde

ÉDUCATION • SOCIÉTÉ

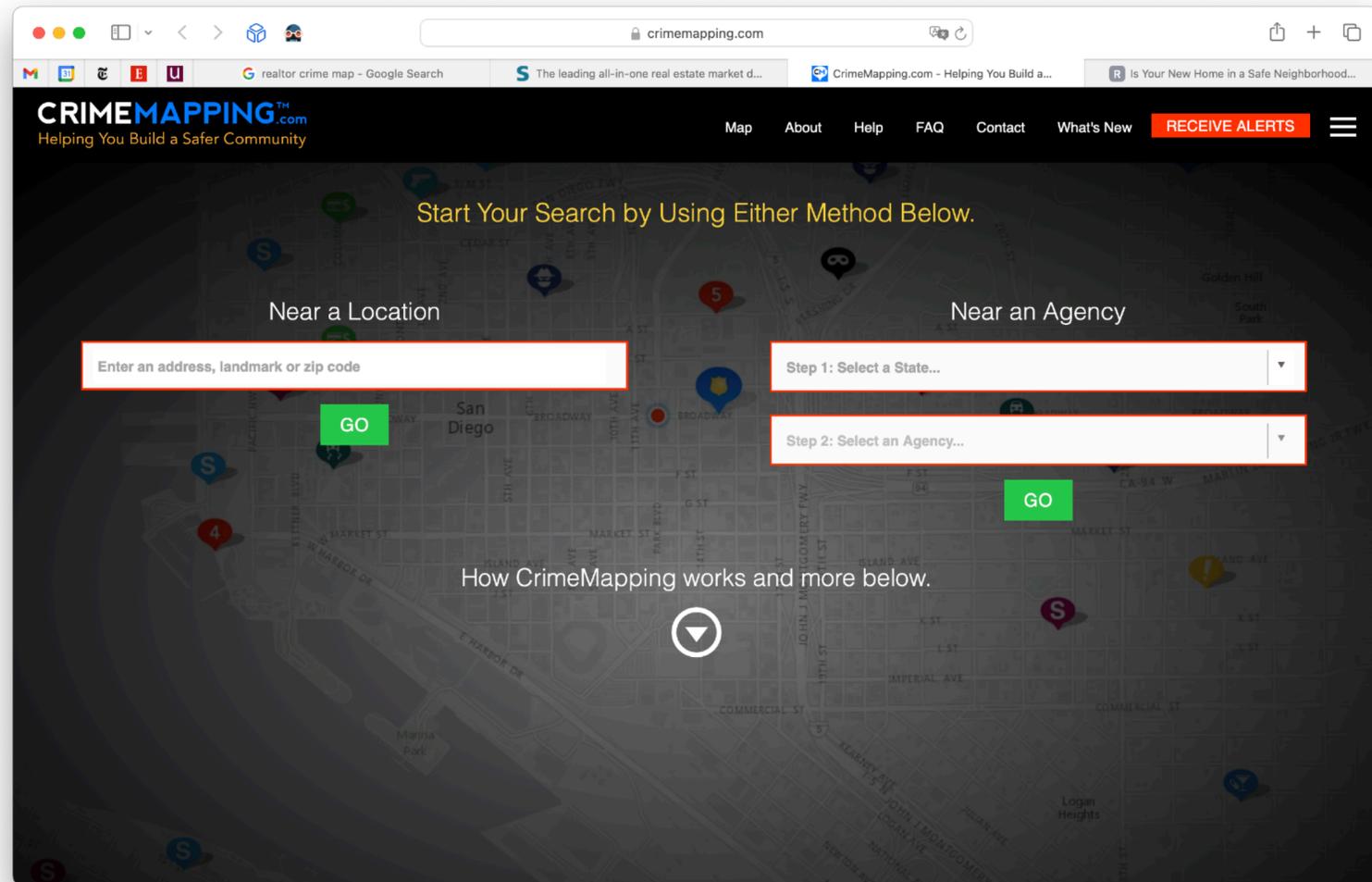
## Les données personnelles détenues par l'éducation nationale, un trésor convoité

Les informations personnelles numérisées sur les élèves et les personnels de l'éducation représentent une manne pour de nombreux éditeurs de logiciels. Le ministère promet de rédiger un « code de conduite ».

Par Soazig Le Nevé  
Publié le 11 mars 2019 à 14h42, modifié le 12 mars 2019 à 11h30 • Lecture 6 min.

# Open data et entrepreneuriat

Immobilier et data délinquance aux US



# L'open data des décisions de justice et leur exploitation

La plupart des décisions (première instance) ne sont pour le moment pas facilement accessible aux tiers : **seules les grandes décisions sont largement diffusées**

L'open data consiste à mettre en libre accès dans un format exploitable la quasi-totalité des décisions

Ainsi, le montant accordé en prestation compensatoire après un divorce, les dommages-intérêts, la décision de détention provisoire etc. pourront être exploitées statistiquement

- Quel montant dans quelle ville ? Selon quelles caractéristiques du demandeur ?
- Ce juge précis décide-t-il plus que la moyenne la détention provisoire ?
- Les juges femmes accordent-elles plus d'indemnité ?
- etc

# L'open data des décisions de justice et leur exploitation

<http://www.conseil-etat.fr/Actualites/Discours-Interventions/La-justice-predictive>

## **I. La justice prédictive est porteuse de transformations majeures, mais ambivalentes.**

### **A. Elle promet certes des évolutions bénéfiques pour la qualité et l'efficacité de la justice.**

1. Les algorithmes prédictifs, fondés sur l'ouverture progressive, mais massive et gratuite des bases de jurisprudence à tous – l'*open data* –, visent à accélérer le règlement des litiges et à accroître la sécurité juridique, en améliorant la prévisibilité des décisions de justice. En effet, par leur utilisation les juges connaîtront mieux les pratiques juridictionnelles de leurs collègues et les parties pourront déterminer plus précisément les chances de succès d'une procédure juridictionnelle, ainsi que les moyens les plus pertinents à soulever. En retour, le couple *open data*/algorithmes devrait favoriser l'accès au droit et l'égalité devant la justice ainsi que la stabilisation, l'harmonisation et la convergence de la jurisprudence. Il est certes des méthodes plus classiques et tout aussi efficaces pour parvenir à cette fin. Mais l'unité et la cohérence de la jurisprudence peuvent, c'est vrai, gagner aux développements technologiques prévisibles à court terme.

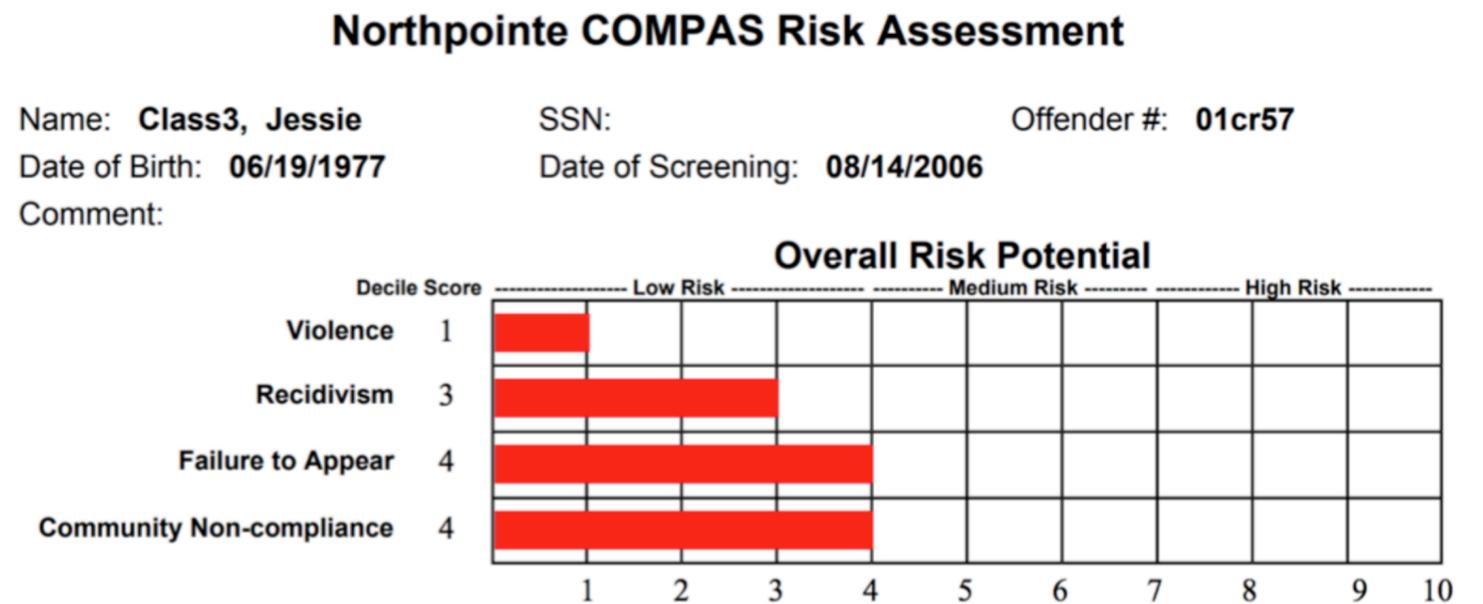
2. Le recours à des algorithmes pour le traitement des dossiers les plus répétitifs et les plus simples, ceux par exemple qui ne nécessitent que l'évaluation d'un dommage, l'application d'un barème ou d'une trame prédéterminée, encouragerait aussi le règlement de nombreux litiges en amont même du recours au juge, par le développements des modes alternatifs de règlement, comme la médiation ou la conciliation. En réduisant le temps passé aux recherches fines fondées sur des

# Les outils d'aide à la décision : justice prédictive

The Impact of Algorithms on Judicial Discretion (Cowgill 2018)

## Tables and Figures

Figure 1: Sample of COMPAS Output: Michigan



**Notes:** The above is a screenshot of a sample report shown to judges in Michigan criminal court. An example of a complete report can be viewed at: [https://www.michigan.gov/documents/corrections/Timothy\\_Brenne\\_Ph.D.\\_Jessie\\_Report\\_297502\\_7.pdf](https://www.michigan.gov/documents/corrections/Timothy_Brenne_Ph.D._Jessie_Report_297502_7.pdf)

# Les outils d'aide à la décision : justice prédictive

## Quels sont les avantages de l'outil précédent ?

La fonction est déterministe : la même situation donnera la même fiche, pas de subjectivité immédiate (fatigue, ambiance dans le tribunal)

## Quels sont les problèmes avec l'outil ? Ils sont nombreux...

La fiche n'est pas un simple outil : biais d'encrage (*framing effect*), illusion de contrôle, et tous les biais cognitifs vus précédemment

- ▶ Pour produire ces fiches : biais dans les données ayant servies à l'apprentissage
- ▶ Effet de renforcement

# L'Intelligence Artificielle : principes et enjeux

# IA : l'apprentissage machine

Fonctionnement (très) schématique de l'IA :

1. Codage d'une base de données
  - Variables explicatives
  - Variable de résultat, outcome

*Le labeling est très intensif en main d'œuvre !*

2. Apprentissage sur cette base pour faire ressortir des corrélations
  - Boîte noire : ces corrélations ne sont pas toujours explicites
3. Mise en œuvre avec d'autres données afin d'obtenir une prédiction

---

**Quels sont les biais possibles et leurs sources ?**

---

# IA : l'apprentissage machine

## 1. Le choix des variables peut impacter le résultat ou l'interprétation

- Un codage fin des variables donnera un résultat détaillé dans ces situations
  - Ex. employeur screenne les CV avec rang des écoles mais pas des universités

## 2. L'algorithme peut révéler nos propres biais

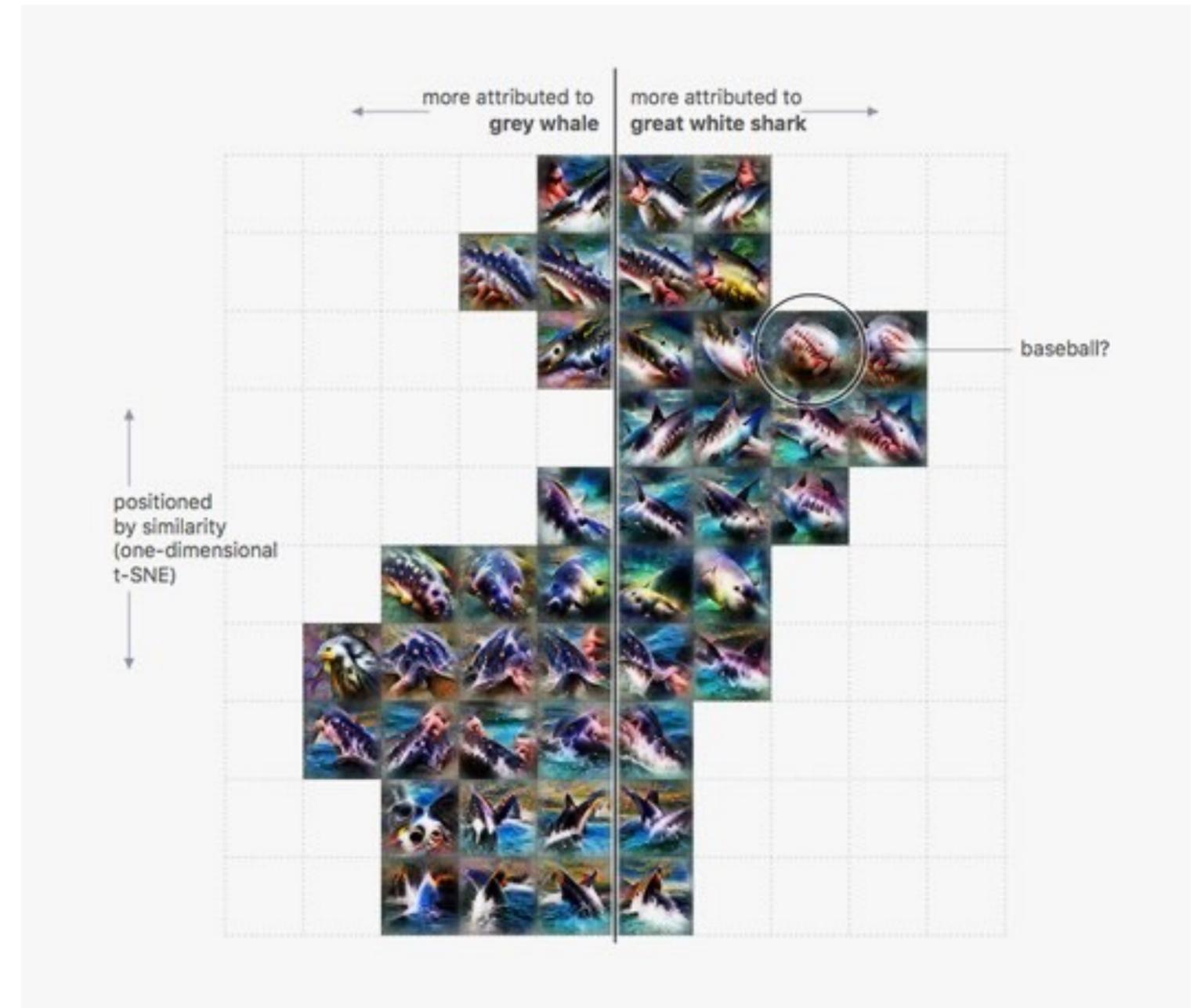
## 3. Corrélations fallacieuses

- Si dans l'échantillon de vérification les tests sont réalisés le mercredi, ce jour ressortira comme une variable explicative
- Les langages de programmation anciens ressortent comme explication du succès professionnel (car sur le CV de la génération d'ingénieurs qui a accédé aux postes à responsabilités)

# Transparences des algorithmes

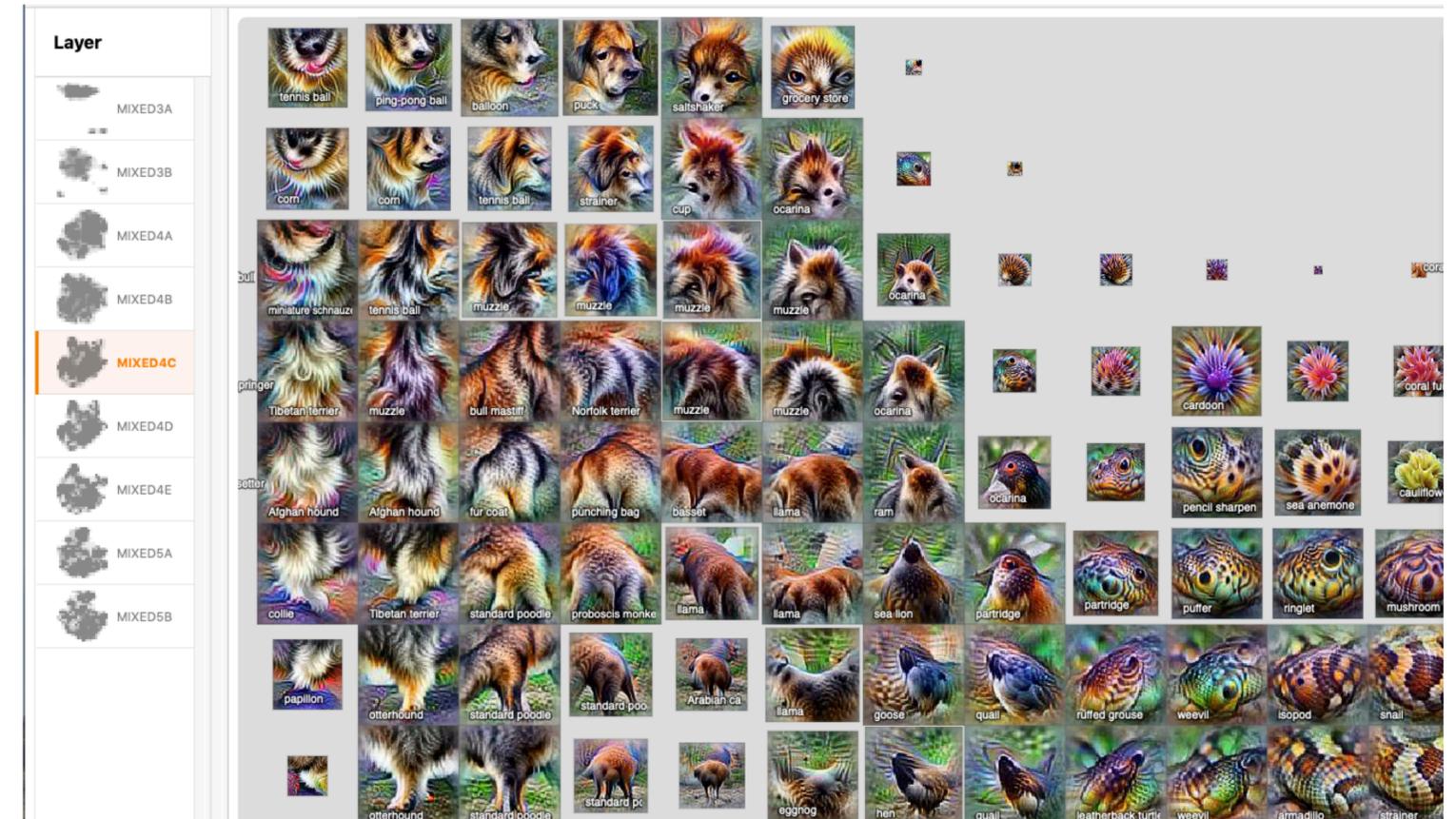
L'apprentissage machine n'est pas aussi explicite qu'un scoring avec pondérations

Wired, 6 mars 2019 : <https://www.wired.com/story/inside-black-box-of-neural-network/>



# Transparences des algorithmes

Bien différencier un chien d'une balle de tennis, d'une plante et d'un serpent...



Wired, 6 mars 2019 : <https://www.wired.com/story/inside-black-box-of-neural-network/>

# Effets économiques et sociaux de l'IA

L'IA remplace-t-elle l'intervention humaine ?

Oui ponctuellement pour des tâches mécaniques

**Mais** l'apprentissage machine a besoin d'apprendre

- Pour identifier des corrélations **l'IA doit connaître la bonne réponse**
- Et si la compétence est perdue par les humains ?

# Voiture autonome

## Moins d'accidents, mais ceux qui restent sont causés par une machine

Les humains sont de très mauvais conducteurs

une voiture autonome apportera une baisse drastique des accidents

...mais suscitera des accidents spécifiques à l'IA

Une décision mortelle a alors été prise par un véhicule autonome : est-ce acceptable ?

---

Généraliser la conduite autonome est-il une horreur techniciste aveugle ou au contraire une obligation éthique ?

---

# Voiture autonome

## Il n'y a pas d'explication des décisions

On imagine souvent que les décisions d'IA de véhicules autonomes sont basées sur des schémas éthiques :

- Aristotéicien : sacrifier le conducteur
- La règle : suivre le code de la route sans considération des conséquences
- Utilitariste : minimiser le nombre d'années de vie perdues

Mais c'est plutôt une « boîte noire » : la décision ne sera ni explicable ni clairement réductible à des principes définis *ex-ante*

# Voiture autonome

## Tesla's Autopilot Could Save the Lives of Millions, But It Will Kill Some People First

The complicated ethics of Elon Musk's grand autonomous vehicle experiment.

<https://www.bloomberg.com/news/features/2019-10-09/tesla-s-autopilot-could-save-the-lives-of-millions-but-it-will-kill-some-people-first>

