



Les acteurs publics face à la donnée :
Entre opportunités, nouveaux risques
et nouvelles régulations

Sciences Po Lille - 14 novembre 2019

DATACTIVIST

We open data, we make them useful

Dataactivist est un **pure player de l'open data** créé en 2016, par Samuel Goëta et Joël Gombin. Notre mission : ouvrir les données, les rendre utiles et utilisées.

Se positionnant sur **toutes les étapes du travail d'ouverture des données**, Dataactivist travaille tant avec les producteurs de données qu'avec les réutilisateurs et participe à l'appropriation des données par chacun.

Nous appliquons nos propres **valeurs** : nos supports de formation et nos contenus sont librement réutilisables, publiés en licence Creative Commons.

Nous animons la communauté [**#TeamOpenData**](#)

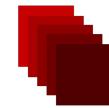


Nos clients - Acteurs publics



Nous travaillons avec de nombreux acteurs publics au niveau central et territorial, et les accompagnons sur des projets de médiation, de stratégie ou encore de formation. Notre travail autour de l'open data se situe souvent à l'intersection de l'innovation économique et de la création de valeur grâce à de nouveaux procédés, services et produits mais aussi aux questions de transparence démocratique et de redevabilité de l'action publique.

Nos clients - Acteurs privés et ONGs



Stat4decision
Data science company



Compagnie des Alpes



GREENPEACE



Nous travaillons également avec des acteurs privés et des ONG, qui sont de plus en plus nombreux à rejoindre le mouvement de l'open data. Nous les aidons à définir leur stratégie, et à augmenter l'impact du partage de leurs données sur leur modèle économique. Dataactivist est aussi missionné par ces acteurs pour rechercher de nouvelles données à utiliser, et mettre en oeuvre des modèles innovants de gouvernance de données.

Introduction à la donnée

Qu'est-ce qu'une donnée ?

- Pouvez-vous lister 10 données que l'on peut trouver dans cette pièce ?

Partage des définitions

- Quelle information est commune à ces deux images ?



Partage des définitions

- Quelle information est commune à ces deux images ?



“Il fait beau”

Partage des définitions

- Quelle information est commune à ces deux images ?



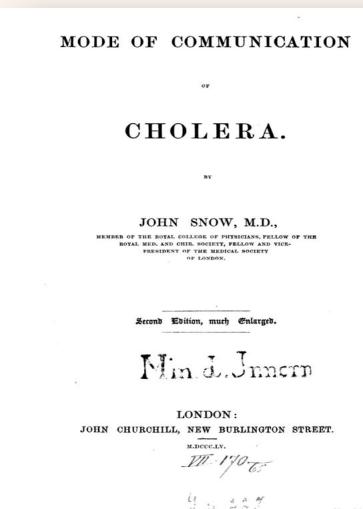
Il fait – 21°C
La nébulosité est de 1%
L'humidité est de 25 %
Coordonnées : -70.302347, 49.089568



Il fait 28°C
La nébulosité est de 6%
L'humidité est de 67%
Coordonnées : -5.144064, 123.215463

Partage des définitions

- Exemple d'un croisement de données



84 THE EPIDEMIC OF 1854,

TABLE VII.

The mortality from Cholera in the four weeks ending 5th August.

Sub-District.	Popula- tion in 1851.	Water Supply.									
		Private Pans & Sinks.	Public Pans & Sinks.	Public Wells.	Public Thems, &c.	Private Wells.	Public Wells.	Public Thems, &c.	Private Wells.	Public Wells.	Uninter- rupted.
St. Saviour, Southwark.	19,709	56	94	—	—	—	—	—	—	—	—
St. Olave, Southwark.	8,015	19	15	—	—	—	—	—	—	—	—
St. John, Horsleydown.	10,000	—	—	—	—	—	—	—	—	—	—
St. Paul, Ludgate Hill.	15,259	29	23	—	—	—	—	—	—	—	—
St. Mary Magdalene.	13,634	30	19	—	—	—	—	—	—	—	—
London Wall.	—	—	—	—	—	—	—	—	—	—	—
Rotherhithe.	17,505	26	17	—	—	—	—	—	—	—	—
Battersea.	10,560	13	10	—	—	—	—	—	—	—	—
Wapping.	—	—	—	—	—	—	—	—	—	—	—
Punney.	—	5,280	1	—	—	—	—	—	—	—	—
Camberwell.	—	17,742	19	15	—	—	—	—	—	—	—
Peckham.	—	19,444	4	4	—	—	—	—	—	—	—
Christchurch, Southw.	16,022	3	2	1	—	—	—	—	—	—	—
Kent Road.	18,126	8	7	1	—	—	—	—	—	—	—
Bowes Road.	—	91	90	—	—	—	—	—	—	—	—
London Road.	17,836	9	6	4	—	—	—	—	—	—	—
Trinity, Newington.	20,500	1	1	—	—	—	—	—	—	—	—
St. Paul, Newington.	14,033	5	5	—	—	—	—	—	—	—	—
Waterloo Road (1st).	—	5	5	—	—	—	—	—	—	—	—
Waterloo Road (2nd).	18,548	5	5	—	—	—	—	—	—	—	—
Lambeth Church (1st).	18,469	5	2	1	—	—	—	—	—	—	—
Lambeth Church (2nd).	—	10	10	—	—	—	—	—	—	—	—
Kennington (1st).	84,281	11	9	1	1	—	—	—	—	—	—
Kennington (2nd).	18,448	3	3	—	—	—	—	—	—	—	—
Brixton.	—	10	10	—	—	—	—	—	—	—	—
Clapham.	—	16,290	5	4	—	—	—	—	—	—	—
St. George, Camberwell.	15,459	9	7	2	—	—	—	—	—	—	—
Norwood.	—	3,077	—	—	—	—	—	—	—	—	—
Streatham.	—	9,023	—	—	—	—	—	—	—	—	—
Dulwich.	—	1,632	—	—	—	—	—	—	—	—	—
Sydenham.	—	4,401	—	—	—	—	—	—	—	—	—
	486,006	334	280	14	4	26	4				

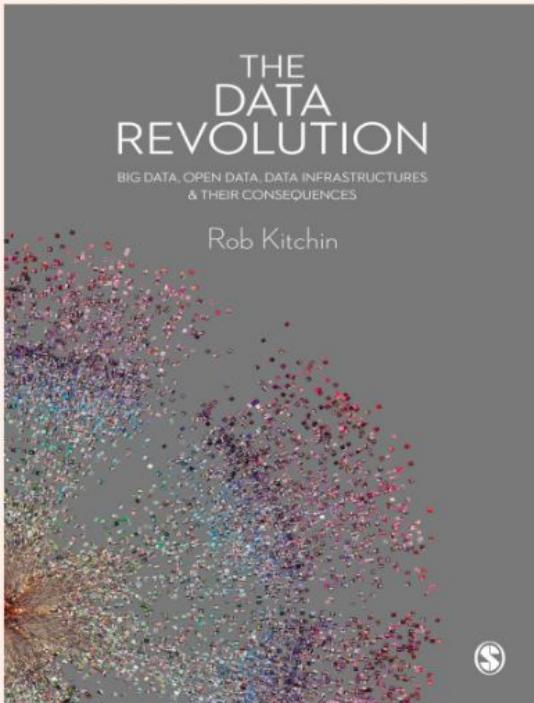
La pyramide Data-Information-Knowledge-Wisdom



Attribuée à Russell Ackoff en 1989, elle signifie que :

- Les **données** sont la matière "brute" de l'information conçues plutôt pour des machines.
- **L'information** pourrait être définie comme des données qui ont été interprétées pour dégager du sens pour des humains.
- En donnant du sens à de l'information, on obtient de la **connaissance**
- En donnant du sens à la connaissance on obtient de la **sagesse**.

Une définition générale des données



Les données sont couramment comprises comme les matériaux bruts produits dans l'abstraction du monde en catégories, mesures et toute autre forme de représentation-nombres, caractères, symboles, images, sons, ondes électromagnétiques, bits qui constituent les fondations sur lesquelles l'information et le savoir sont créés.

Cette définition est une des rares à s'extraire des discours entourant les données ou leurs infrastructures. L'auteur de cet ouvrage de référence en sociologie s'attache à définir précisément l'ontologie et l'épistémologie de la donnée.

<https://dataactivist.coop/ministere-culture/jour1.html#6>

Big data, données personnelles, open data



Big data, données personnelles, open data

L'obsession du volume de données

Quelques chiffres omniprésents :

- le volume de données produit double tous les 3 ans (Gantz & Reisel 2011)
- 90% des données créées dans les deux dernières années (IBM 2012)
- 40% : croissance annuelle de la production de données (Maniyka et al. 2011)

Problèmes de cette approche :

- Estimations guidées par des intérêts commerciaux
- Ne définit pas ce que sont ces données
- Résume le big data au Volume - Explique mal la mise en données du monde



<https://dataactivist.coop/dataliteracy/#65>

Big data, données personnelles, open data

Les promesses du big data

Kitchin (2014) résume les promesses du big data :

- “**Governing people**” : dans la continuité de la statistique, améliorer la connaissance de l’administration et prédire les crimes
- “**Managing organisations**” : améliorer le fonctionnement de toutes les composantes de l’organisation par l’exploitation des données
- “**Leveraging value and producing capital**” : micro-ciblage marketing, optimisation des magasins et des opérations, efficience de la chaîne
- “**Creating Better places**” : gouverner les villes avec des données (smart city)
- **Un nouveau paradigme scientifique** : une nouvelle ère guidée par les corrélations

Big data, données personnelles, open data

Les réalités du Big Data

- Les administrations et les entreprises ne disposent pas toutes de données massives, suffisamment volumineuses pour “faire du Big Data”
- Les données disponibles sont souvent insuffisantes, et leur qualité nécessite un travail de nettoyage important avant de pouvoir être traitées
- Parfois, les besoins réels des différents métiers ne sont pas assez écoutés, et la solution est souvent autre que technologique

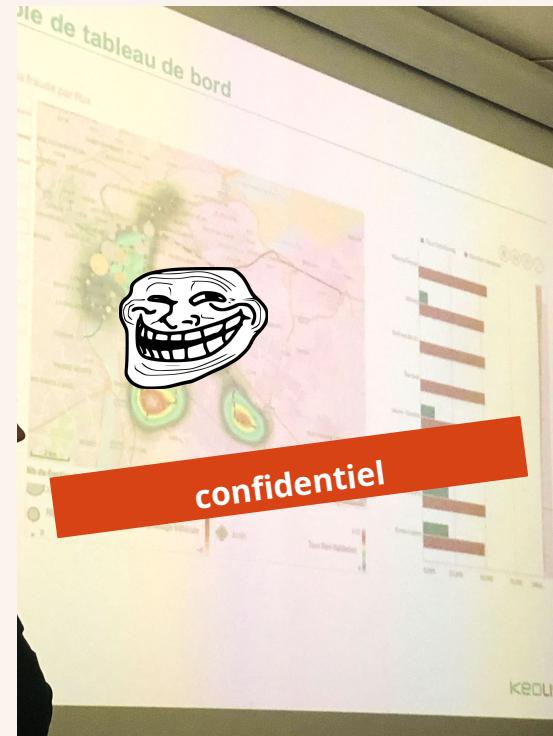
"Le big data, c'est comme le sexe chez les adolescents : tout le monde en parle, personne ne sait vraiment comment le faire, tout le monde pense que tout le monde le fait, donc tout le monde prétend le faire." Dan Ariely

Big data, données personnelles, open data

Carte de la fraude à Lyon =
où envoyer des contrôleurs ?

Un exemple proche d'un usage Big Data : Keolis Lyon

- Pertes de 10 millions d'euros par an
- 200 agents dédiés au contrôle
- Outil Dataviz : un dashboard avec :
 - entrées et sorties (validations)
 - différentiel comparé à la fréquentation
 - cartographie des fraudes
 - remplissage et fréquentation temps réel
- Données à croiser avec :
 - la météo (fraude plus élevée lors de fortes averses : on prend le bus pour une station)
 - période de l'année
 - événements



Big data, données personnelles, open data

Toute information relative à un particulier identifié ou identifiable, directement ou indirectement, grâce à un identifiant ou à un ou plusieurs éléments propres à son identité

Par exemple :



Source : Règlement général sur la protection des données (RGPD) du 27 avril 2016

Big data, données personnelles, open data

Le RGPD ?

Régllement Général de Protection des Données : adopté à la mi-avril 2016 après 4 années de débat, il va considérablement renforcer le droit européen de protection des données personnelles

3 objectifs :

- Renforcer les droits des personnes
- Responsabiliser les acteurs traitant des données personnelles
- Crédibiliser la régulation

Le RGPD propose un **cadre uniifié pour l'ensemble de l'UE** qui s'applique à chaque fois qu'un résident européen est visé par un traitement de données personnelles.

Big data, données personnelles, open data

Exemple de demande de consentement explicite



Big data, données personnelles, open data

Responsabiliser les acteurs traitant des données personnelles

Principes :

- **Privacy by design** : protection des données personnelles dès la conception du produit et par défaut
- **Minimisation** : limiter la quantité de données personnelles dès le départ
- **Accountability** : mettre en place des mesures de protection des données et démontrer cette conformité à tout moment
- Fin des obligations déclaratives sauf si risque accru pour la vie privée

<https://dataactivist.coop/dataliteracy/#77>

Big data, données personnelles, open data

Open data : un terme récent

Le terme est apparu pour la première fois dans les années 1970 dans les accords qu'a signés la NASA avec des pays partenaires en vue du partage de données satellitaires.

Dans le langage courant, cela fait **près de 10 ans** que l'ouverture des données est devenue une composante majeure des politiques publiques numériques.

Big data, données personnelles, open data

07 Décembre 2007 : la rencontre de Sebastopol

 **Quoi ?** Une rencontre de l'Open Government Group à Sebastopol (Californie), siège des éditions O'Reilly

 **Pourquoi ?** : Influencer le futur président des Etats Unis pour faire avancer l'open data

 **Comment ?** En adoptant une déclaration définissant les grands principes de l'Open Government Data



Big data, données personnelles, open data

Les 8 principes de l'Open Government Data

1. Complete

All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.

2. Primary

Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.

3. Timely

Data is made available as quickly as necessary to preserve the value of the data.

4. Accessible

Data is available to the widest range of users for the widest range of purposes.

5. Machine processable

Data is reasonably structured to allow automated processing.

6. Non-discriminatory

Data is available to anyone, with no requirement of registration.

7. Non-proprietary

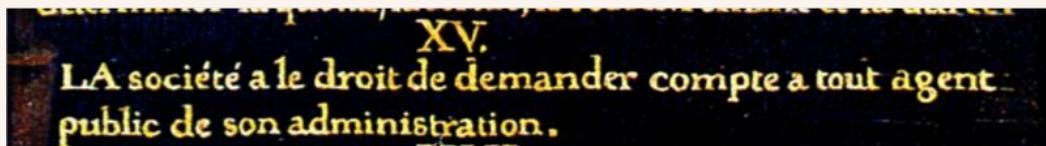
Data is available in a format over which no entity has exclusive control.

8. License-free

Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Big data, données personnelles, open data

Des fondements historiques en France



DATACTIVIST



Big data, données personnelles, open data

Open data : quelques jalons historiques

<https://dataactivist.coop/dataliteracy/#41>



DDHC, Art. 15.
La Société a le
droit de
demander
compte à tout
Agent public de
son
administration.

1789

Loi CADA : droit
d'accès à
l'information
publique

1978

Loi Lemaire :
ouverture des
données
devient
obligatoire

2016



1966

Freedom of
Information Act,
USA

2003

Directive
européenne

2007

Directive
européenne
INSPIRE

2013

Charte du G8
sur l'open data

Big data, données personnelles, open data

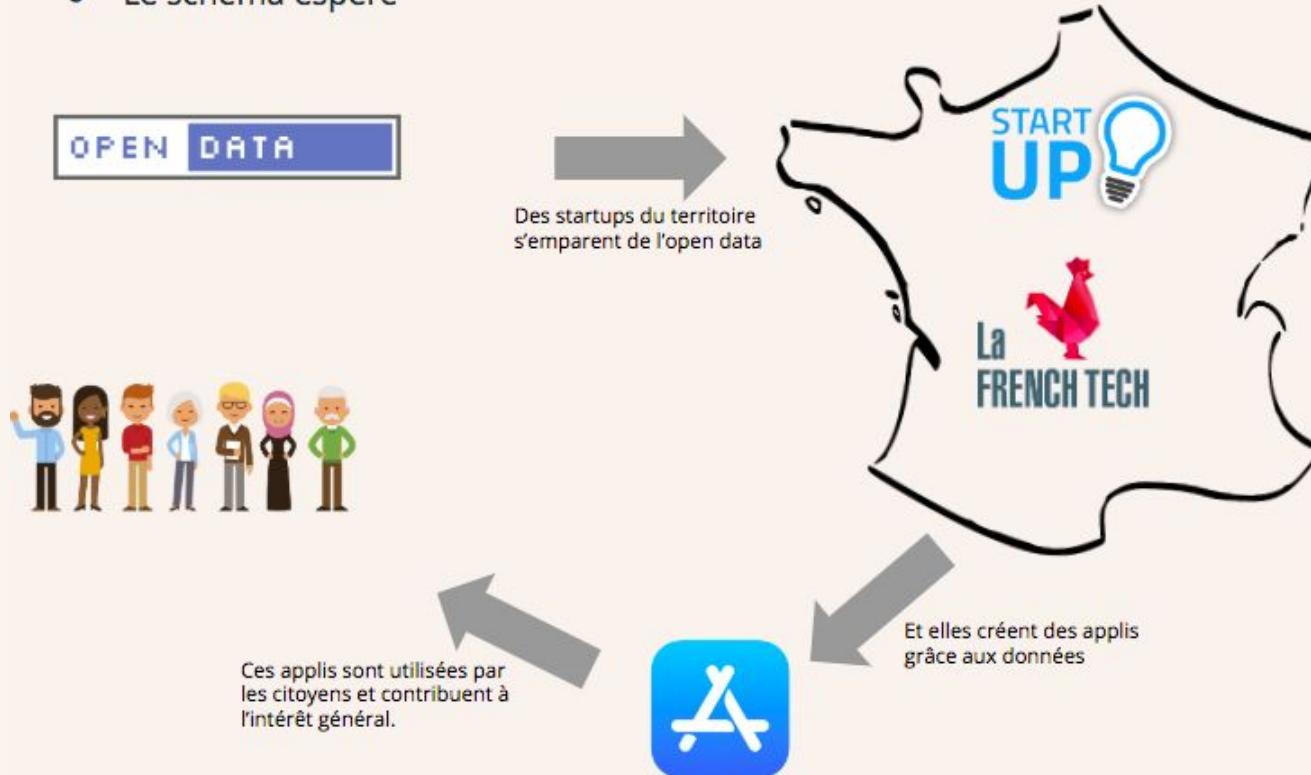
L'open data : pourquoi faire ?

-> réponses dans la dernière partie

La donnée : pétrole du XXIe siècle ?

Des espoirs démesurés de croissance

- Le schéma espéré



Des espoirs démesurés de croissance



La promesse French Tech, symbole d'un espoir de nouvelles sources de croissance

“ la France a lâché dans l'Eureka Park plus de start-up que les Etats-Unis, pays hôte du salon, cinq fois plus peuplé et premier marché technologique occidental. En tout, 381 jeunes pousses tricolores ont fait le déplacement.”

Jamal El Hassani, Journal du Net

- Mais un grand nombre de startups qui utilisent des données publiques ne font pas encore de profit
- Et certains élus et décideurs misent beaucoup sur ces mêmes entreprises

Les administrations et entreprises publiques peuvent-elle vendre leurs données ?

le cas du Grand Lyon

- Licence associée : objectifs et résultats

	LICENCE OUVERTE	LICENCE ENGAGEE	LICENCE ASSOCIEE
Utilisation commerciale de données	✓	✓	✓
Mention de la source et de la date de mise à jour	✓	✓	✓
Déclaration sur l'usage de la donnée	✗	✓	✓
Authentification nécessaire avant téléchargement	✗	✓	✓
Possibilité d'une redevance	✗	✗	✓

Licence associée : une licence unique en France, soumise à redevance.

Jugée non conforme à la Loi pour une République Numérique, ne rapporte pas de véritables recettes, et basée sur un calcul de redevance selon l'utilisateur, qui n'est pas perçu comme suffisamment clair.

Les administrations et entreprises publiques peuvent-elle vendre leurs données ?

- Vendre des données... mais pas directement pour l'argent



→ La redevance sur l'API SNCF est davantage pensée pour empêcher les géants du numérique de rediffuser des données que pour générer des recettes directes (puisque la valeur réside dans l'utilisation des données et l'audience donnée à l'offre SNCF).

Les modèles possibles de vente de données (hors données personnelles)

- **La donnée brute seule a peu de valeur** : les nombreuses promesses du big data vs des données non structurées, pas interopérables, de faible qualité...
- **Une donnée est invendable** s'il est possible de la trouver ailleurs gratuitement ou de la reconstituer soi-même.
- **Besoin de services autour de la donnée :**
 - compétences et métiers (data analysts, data scientists, CDO...)
 - Recherche de données et qualification
 - Mise en qualité de données
 - Garanties (sécurité, qualité, mise à jour, continuité)
 - = service de régie, conciergerie, ingénierie...

Les modèles possibles de vente de données (hors données personnelles)

- La valeur d'une donnée est fixée en fonction de la valeur de son utilisation :

Exemple : des développeurs indépendants développent des apps gratuites pour communiquer (apps vitrines)



Une application conçue par **alex et manon**

Dernière mise à jour : 8 janvier 2018

alex et manon

les artisans de l'application mobile
à Lille

alex et manon vous accompagnent de l'idée à la
réalisation de votre application mobile.
Donnez vie à votre idée !

[NOUS CONTACTER →](#)

à Lille

Les modèles possibles de vente de données (hors données personnelles)

- Pourquoi Waze propose-t-il un échange de données gratuit ?

- Gagner la confiance des collectivités ?
- Proposer des données payantes plus tard ?
- Vendre un logiciel d'analyse de données ?



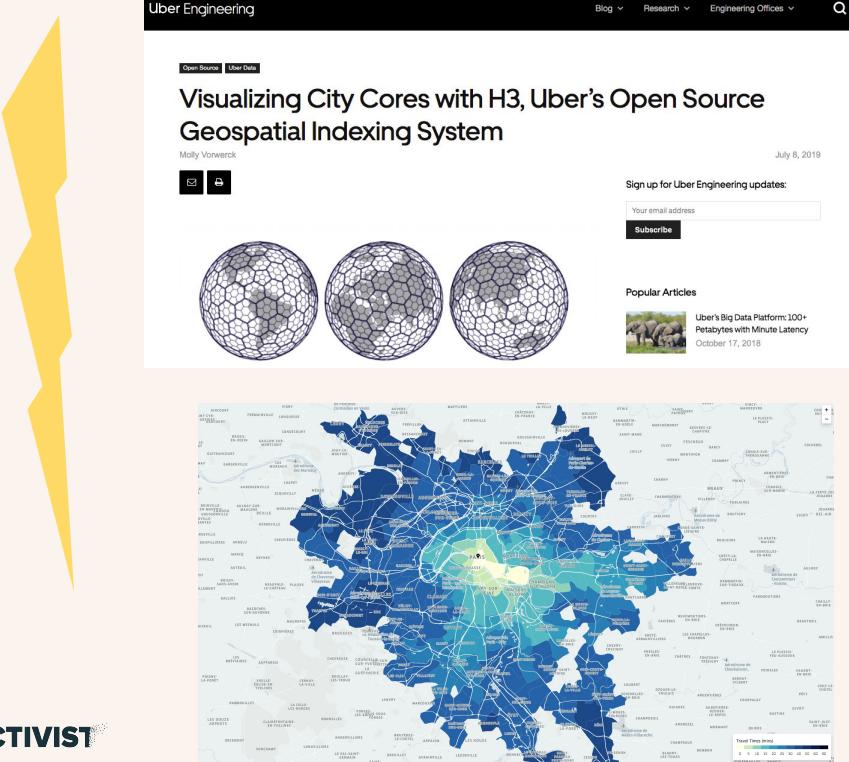
Les modèles possibles de vente de données (hors données personnelles)

- Uber, entreprise de VTC ou géant de la donnée ? inventer de nouveaux standards en révolutionnant la modélisation spatiale



SOURCE, février 2019

DATACTIVIST

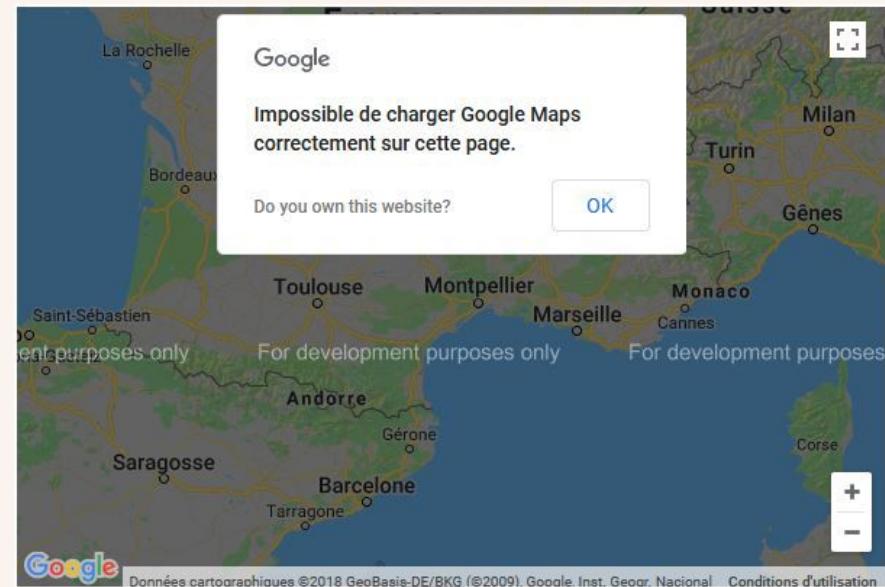


Les modèles possibles de vente de données (hors données personnelles)

- Rendre une donnée unique et nécessaire pour pouvoir la vendre ensuite, en offrant un service d'abord gratuit (dumping).
- Plus d'un million de sites utilisent Google Maps en fond de carte

→ Nécessité des contre-pouvoirs

Lorsque Google Maps augmente ses prix, les cartes Google affichées sur les sites (boutiques par exemple) deviennent indisponibles. Des contre-pouvoirs, comme OpenStreetMap (collaboratifs, libres et gratuits), sont plus que nécessaires pour proposer des services de manière pérenne.



Dérives des géants d'Internet et tentatives de régulations publiques

Tromper les algorithmes ? Le cas de Waze

Les "No Waze Streets"

On commence à apercevoir des annonces immobilières pour des biens qui vantent le fait de ne pas subir les redirections de trafic opérées par Waze via des quartiers résidentiels.

[Source : Antoine Courmont](#)



TRANSPORTATION URBAN DEVELOPMENT

Waze Hijacked L.A. in the Name of Convenience. Can Anyone Put the Genie Back in the Bottle?

Traffic apps turned the city's neighborhoods into "shortcuts." Now furious residents are attempting to take them back, street by street

By Jonathan Littman • August 20, 2019



...is where you start. The higher echelons of control come only with

3915 Longview Valley Rd
Sherman Oaks, CA 91423
Status: Closed Sale

\$1,336,974
Redfin Estimate | **\$1,250,000**
Last Sold Price | **3** Beds | **2** Baths | **1,816 Sq**
Built

Overview Redfin Estimate Listing Details Tour Insights Property History Public Facts

SOLD JUL 18, 2018

Street View

3809 Longview Valley Rd
Sherman Oaks, CA 91423
Status: Closed Sale

\$1,365,789
Redfin Estimate | **\$1,200,000**
Last Sold Price | **3** Beds | **3** Baths | **1,827 Sq**
Built

Overview Redfin Estimate Listing Details Tour Insights Property History Public Facts

SOLD OCT 31, 2018

Street View

Mid Century Traditional with contemporary updates on large grassy lot with a pool!

Tucked away on a cul-de-sac (no waze street); this Sherman Oaks South of the Blvd hills home is both Private and has easy access to both the West Side and the Valley. This updated home features a remodeled kitchen with Professional Range and oven, custom cabinetry and stainless steel appliances. Recessed lights have been added throughout. Upon Entry the formal living area features a grand fireplace finished with marble and traditional mantle. Hardwood floors are featured. This home has a flexible layout with a din

en
the quiet cul-de-sac Street (no waze drivers)
fre

dual closets – one a full walk in – and a travertine finished Master bath with Jacuzzi soaking spa tub. The 2nd bedrooms are both well sized and the mid bath has been updated as well. Roof was replaced in 2017 and Electrical panel updated. A Beautiful home in the Hills with space to run and play and a lovely pool to frolic in!

Stunning Mid-century Traditional; remodeled throughout! Located in the Hills of Sherman Oaks; South of the boulevard; with easy access to both the West Side & the Valley. This warm home is on a quiet cul-de-sac Street (no waze drivers) and is complete with a chef's garden and a solar heated swimming pool. The home has an entertainer's open floor plan; and a true chef's kitchen. Viking stove, soft close cabinets, stainless steel appliances and Marble counters – there is space to cook and a comfortable sitting nook in the light and airy room. Gorgeous hardwood floors extend throughout the home and throughout the private side of the house that can be closed off; containing all 3 bedrooms. The 3 bathrooms have been updated keeping with the mid century style but adding

c
c
c Tucked away on a cul-de-sac (no waze street);
f

plumbing included. The Direct access 2 car garage has a finished floor and plentiful cabinets – park the cars or create a gym. The gardens surrounding this home are both practical and lovely – please see the complete list of all edible items in the supplement section.

Tromper les algorithmes ? Le cas de Waze

Solution 1 : tenter de négocier avec Waze pour modifier les données ou l'algorithme ?

- Des collectivités ont déjà demandé à Waze ou Google de rayer une voirie de la carte
- D'autres collectivités tentent de demander à Waze une adaptation de leur algorithme, et de **prendre en compte la hiérarchie des voies** (Ex : Grand Lyon)
- **Jusqu'à aujourd'hui, Waze n'a jamais cédé**

Tromper les algorithmes ? Le cas de Waze

Solution 2 : prendre l'algorithme (et le conducteur) pour un idiot -> ça marche

L'exemple de Lieusaint
en Seine-et-Marne



[SOURCE](#)

Tromper les algorithmes ? Le cas de Waze

Solution 3 : se lier d'amitié avec l'algorithme

Waze teste la régulation des flux touristiques dans le Sud

L'application de guidage automobile Waze et le Comité régional de tourisme Provence-Alpes-Côte d'Azur testent une solution pour fluidifier la circulation dans les sites touristiques les plus fréquentés. Elle va adresser des messages à ses utilisateurs en amont des déplacements, quand leur destination est saturée.

[SOURCE](#)

Tromper les algorithmes ? Le cas de Waze

Solution 4 : mettre en oeuvre une vraie politique publique pour agir sur la source du problème

A fascinating new scheme to create walkable public spaces in Barcelona

Check out "superblocks" y'all.

By David Roberts | @drvox | david@vox.com | Updated Apr 22, 2017, 8:06am EDT

f t SHARE

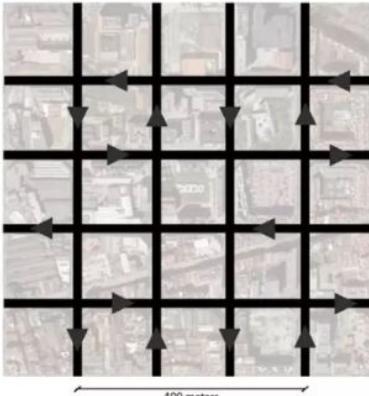


Barcelona's Eixample neighborhood, just waiting to be superblocked. | Shutterstock

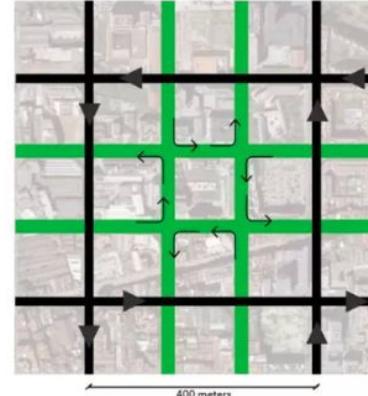
[SOURCE](#)

Road hierarchy in a Superblock model

CURRENT SITUATION



SUPERBLOCK



EXERCISE OF ALL THE RIGHTS THAT THE CITY OFFERS. HIGHEST AIM: CITIZEN.

DATACTIVIST

Tromper les algorithmes ? Le cas de Waze

Solution 4 : mettre en oeuvre une vraie politique publique pour agir sur la source du problème



Mathieu Chassignet
@M_Chassignet

En seulement un an, à Louvain (Belgique) :

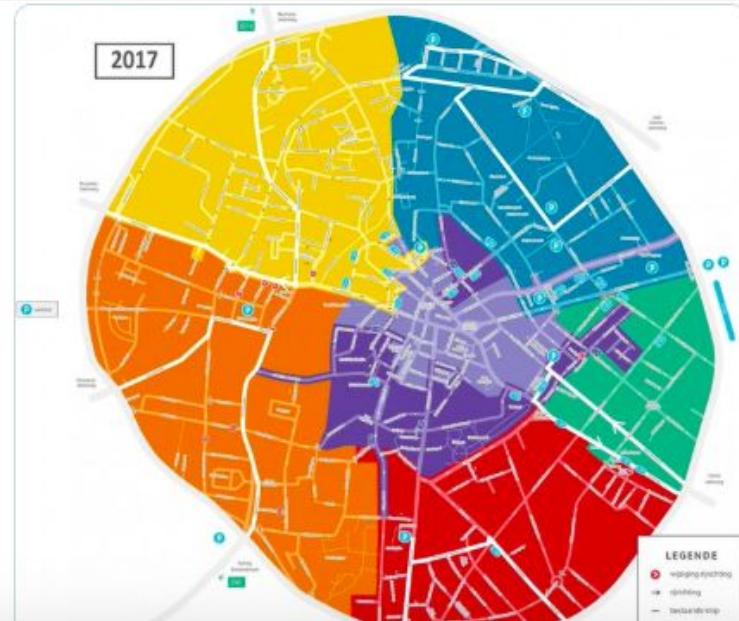
- ➔ Vélo : +32%
- ➔ TC : +12%
- ➔ Voiture : -8% en centre-ville
- ➔ Pollution de l'air : jusqu'à -60%

Comment? Grâce à un nouveau plan de circulation qui coupe le trafic de transit
ecf.com/news-and-event...

[Translate Tweet](#)

[SOURCE](#)

DATACTIVIST



La Smart City et les partenariats publics privés

Le cas du Sidewalk Labs à Toronto

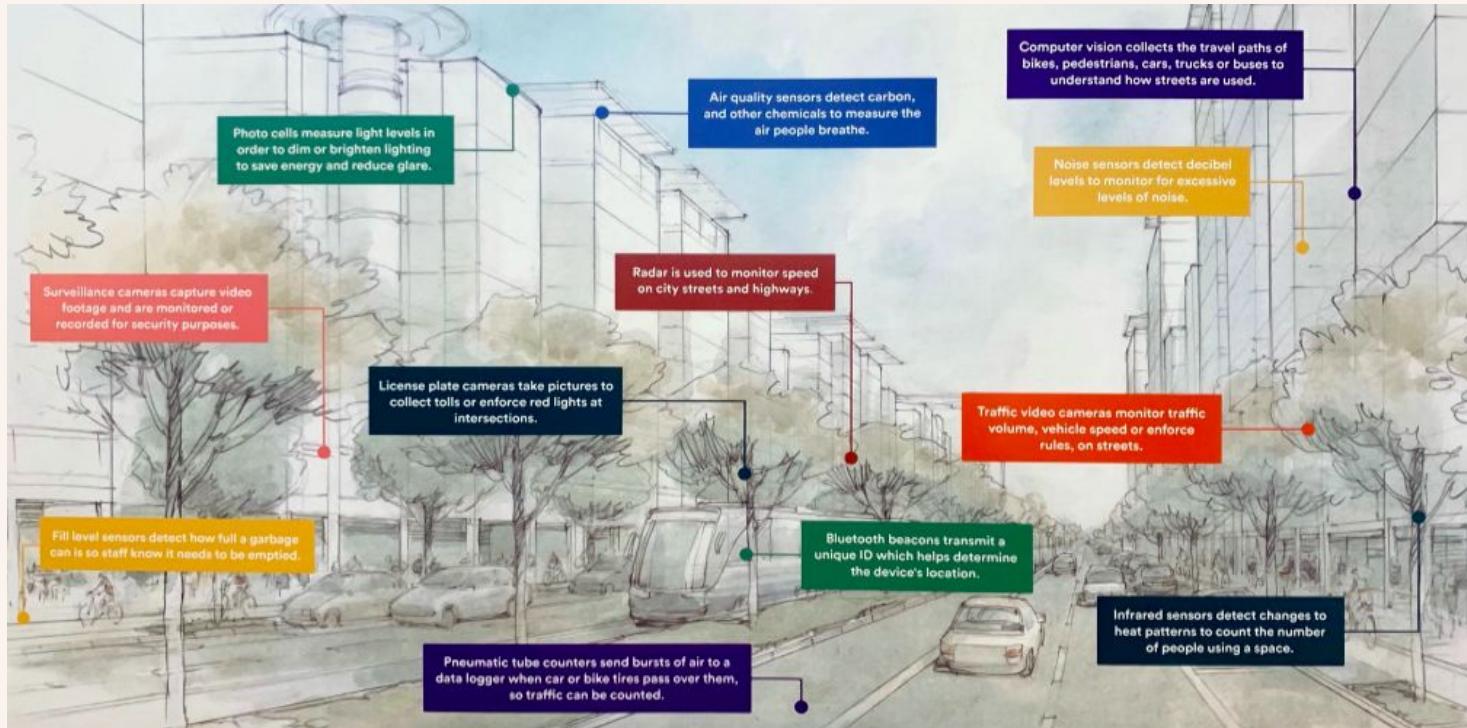


La Smart City et les partenariats publics privés

Le cas du Sidewalk Labs à Toronto

Problème n°1 : une captation de données personnelles importante :

- capteurs vidéo, lecteurs de plaques, capteurs de sons, collecte de géolocalisation



La Smart City et les partenariats publics privés

Le cas du Sidewalk Labs à Toronto

Problème n°2 : un projet porté par Alphabet

Aujourd'hui, Sidewalk Labs garantit que les données collectées resteront dans la maison Sidewalk Labs - Mais Bianca Wylie (opposante au projet), préfère supposer qu'à tout moment, Google peut se rapprocher du projet (enjeux de monétisation des données personnelles)

Un projet et une équipe très influente :
la directrice de la stratégie => c'est l'ex Chief Data Officer de la ville de New York pendant 10 ans.



La Smart City et les partenariats publics privés

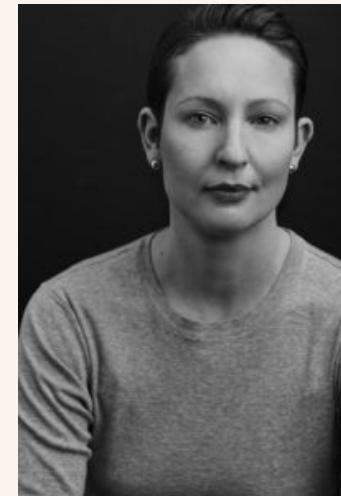
Le cas du Sidewalk Labs à Toronto

Problème n°3 : des citoyens pas assez impliqués dans le projet

*“Alors qu'il attirait peu l'attention des citoyens au départ, le projet a réellement commencé à soulever l'inquiétude générale en août 2018, lorsque la surface réelle du développement annoncé s'est avérée bien supérieure à celle prévue par la proposition initiale. **De nombreux habitants se sont alors rendu compte de l'ampleur du projet, de son opacité et de leur faible pouvoir de décision.***

Les citoyens demeurent largement désinformés et dépourvus de tout pouvoir de décision sur la collecte, les traitements et les réutilisations des informations dont ils sont eux-mêmes producteurs. “

Intervention de Bianca Wylie et Nasma Ahmed, Toronto, avril 2019
Voyage d'Étude organisé par CIVITEO Conseil



Bianca Wylie

Son article sur le master plan du Sidewalk Labs

La Smart City et les partenariats publics privés

Le cas du Sidewalk Labs à Toronto

Des avancées intéressantes cependant :

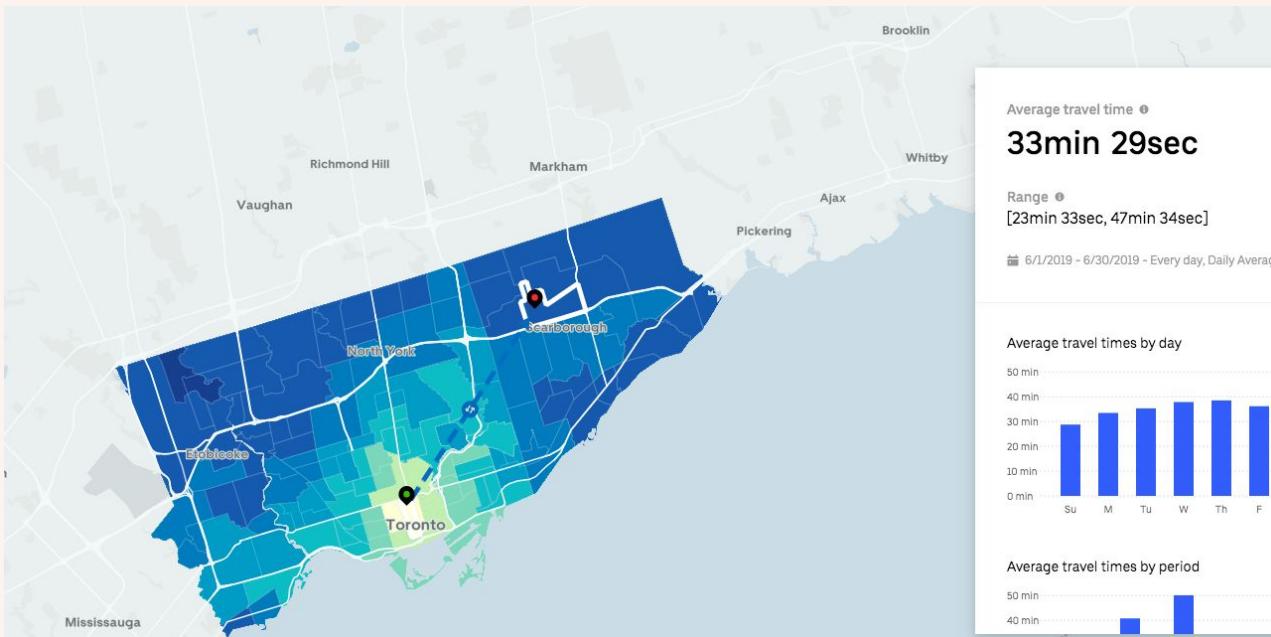
- Un effort de transparence sur la collecte des données (des panneaux positionnés dans la rue informant de façon pédagogique la captation de données aux habitants). **Une pédagogie inégalée jusque là dans le monde sur le sujet des données personnelles dans le cadre de projets urbains.**
- Une mise en **open data** des données pouvant servir à la communauté
- Une réflexion menée sur la question d'un **Civic Data Trust** (un cadre de confiance public, citoyen et privé - autour des données et de leur utilisation).
- Un master plan extrêmement détaillé, qui **clarifie mieux le cadre du projet**



Pour aller plus loin : voir la conférence passionnante de Jacques Priol à Nantes

Réguler avec la donnée

Le cas de la ville de Toronto avec Uber



- Chaque trajet Uber est taxé \$ 0.30
- Pour collecter la taxe, la ville récupère toute les données d'Uber
- Une base de données de 80,000,000 de trajets
- La ville utilise les données pour planifier les transports urbains

Monopoles de la donnée cartographique

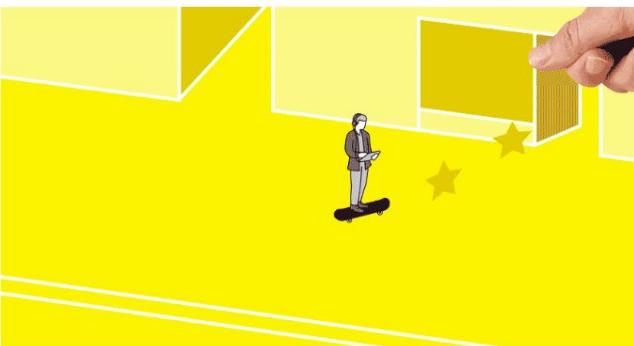
Le cas de Google Maps

La détention d'un monopole sur la collecte et le partage de données peut avoir un impact bien plus élevé que toutes les politiques publiques d'une collectivité locale - Exemple :

Espace public : Google a les moyens de tout gâcher — et pas qu'à Toronto

vraiment Vraiment Follow
Oct 18 · 19 min read

Par Vraiment Vraiment



SOURCE



Monopoles de la donnée cartographique

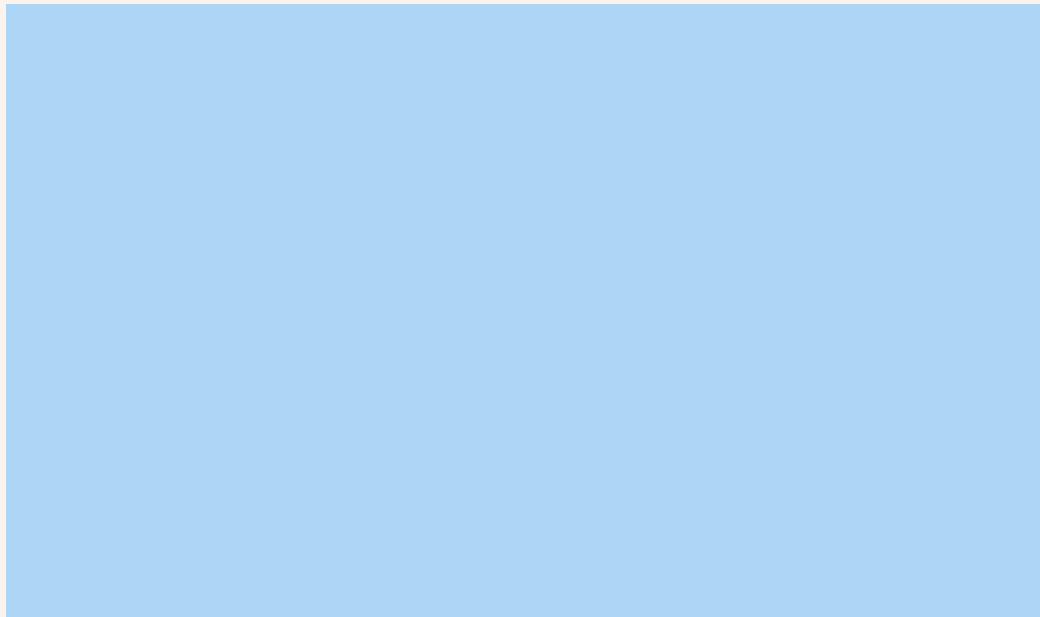
Le cas de Google Maps

Plus d'**un milliard**
d'utilisateurs actifs

Monopoles de la donnée cartographique

Le cas de Google Maps

Des zones d'intérêt, principalement déterminées par la présence de commerces



[SOURCE](#)

Monopoles de la donnée cartographique

Le cas de Google Maps

Des zones d'intérêt, principalement déterminées par la présence de commerces

“Lorsqu'une rue commerçante jouxte un centre commercial, la représentation graphique de Maps attire l'attention sur l'intérieur du centre commercial – alors même que la Rue Neuve, ici, est une des rues les plus commerciales de Belgique.”

[SOURCE](#)



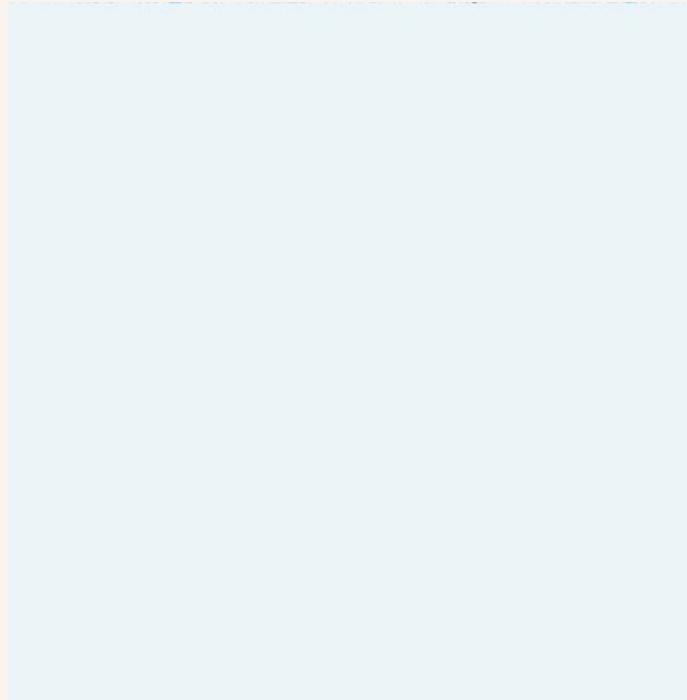
Monopoles de la donnée cartographique

Le cas de Google Maps

Une représentation des données qui ne reflète pas les efforts des politiques urbaines

*“Le boulevard Anspach,
nouveau grand piétonnier
bruxellois de plus de 30m de
large, disparaît de la carte en
étant relayé au maigre statut
de sente piétonne alors qu'il
s'agit de l'axe structurant de la
ville en termes de renouveau de
l'espace public, de commerce et
de représentation mentale de
la ville.”*

[SOURCE](#)

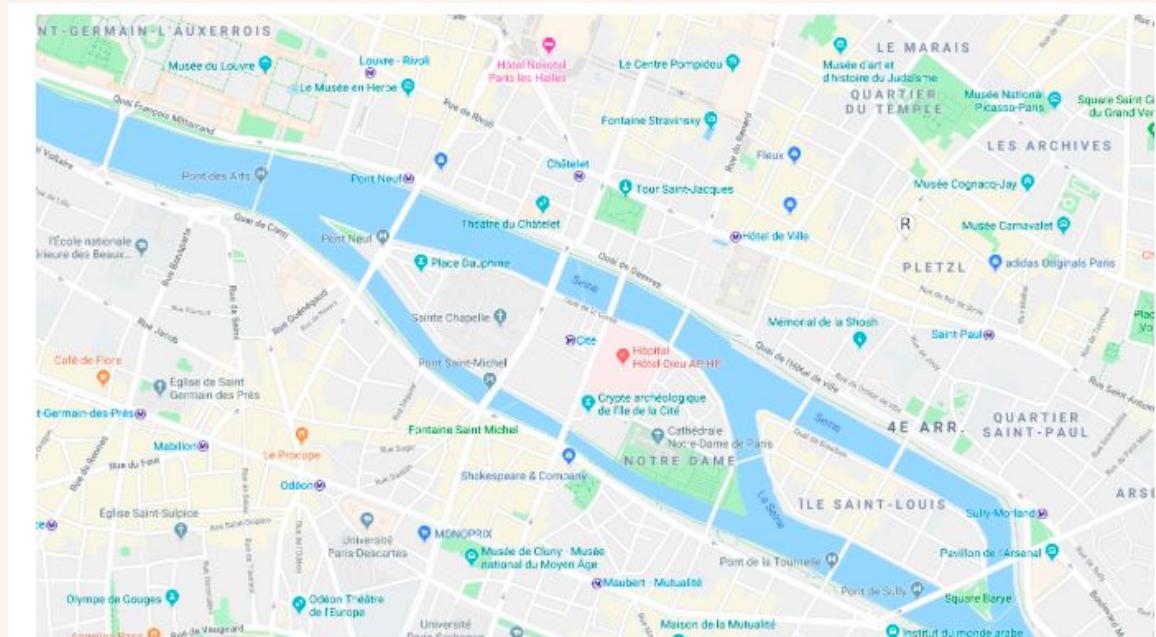


[SOURCE](#)

Monopoles de la donnée cartographique

Le cas de Google Maps

Une représentation des données qui ne reflète pas les efforts des politiques urbaines



SOURCE

Les quais de Seine piétonnisés de Paris, pourtant très fréquenté, ne seront jamais mis en avant comme "Zone of interest" parce qu'ils ne sont pas bordés par du foncier privé (seulement par des péniches et des structures temporaires qui n'apparaissent pas dans le cadastre)

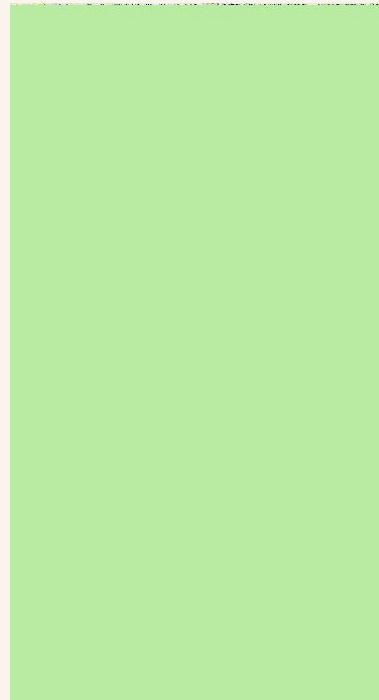
Monopoles de la donnée cartographique

Le cas de Google Maps

Des affichages de données qui correspondent à un référencement payant



Vous pouvez déjà ouvrir un commerce sans vitrine et une enseigne ultra-minimale...à condition d'être bien référencé sur Maps et d'avoir un bon "ranking" Deliveroo.



[SOURCE](#)

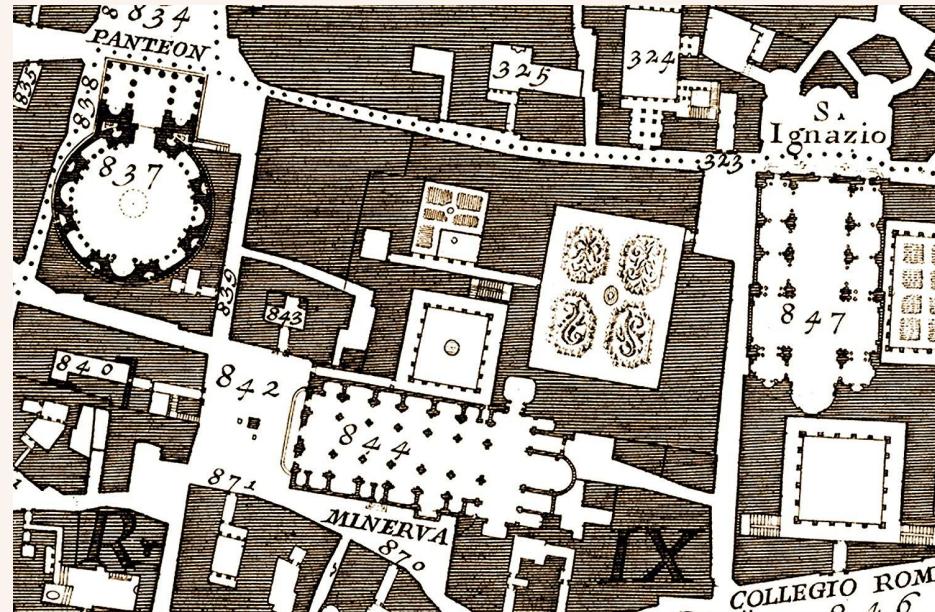
Monopoles de la donnée cartographique

Le cas de Google Maps

Reprendre le contrôle sur la production et la diffusion des données :
un contre-pouvoir nécessaire

L'article termine avec cette illustration issue du Plan de Nolli, (Giambattista Nolli), un plan de Rome représentant les espaces intérieurs accessibles au public (églises, bâtiments publics) en blanc, comme les rues.

[SOURCE](#)

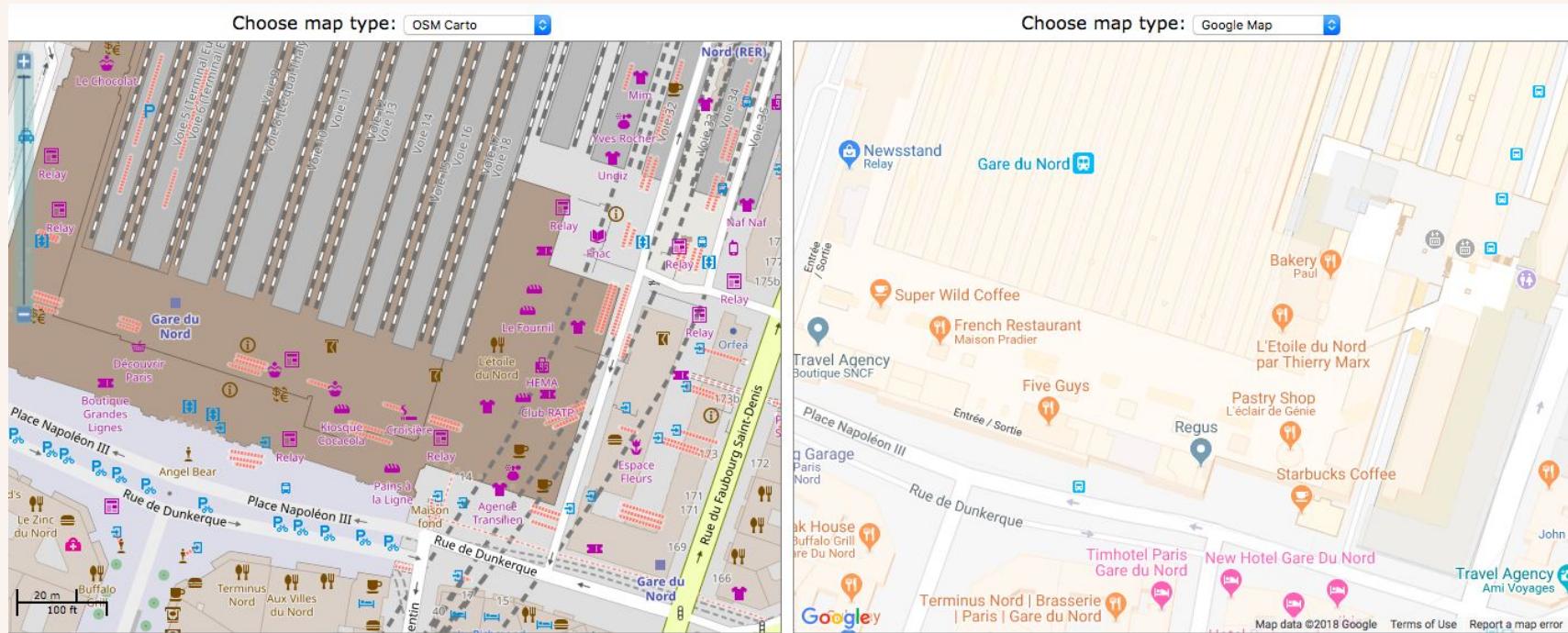


L'open data : un contre-pouvoir, et levier d'action publique

L'open data au service de l'économie et de l'efficacité de l'action publique

Être plus efficace grâce au crowdsourcing

- OpenStreetMap vs Google Maps



L'open data au service de l'économie et de l'efficacité de l'action publique

Créer de nouveaux services



A screenshot of a mobile application interface. On the left, there's a cartoon illustration of a woman holding a smartphone. On the right, there's promotional text and icons. The text reads: "DÉCUPLEZ LES POSSIBILITÉS DE VOTRE CANNE BLANCHE ET FACILITEZ VOS DÉPLACEMENTS". Below this are four circular icons with text: "OÙ SUIS-JE ?" (location), "NAVIGATION INTELLIGENTE" (smart navigation), "BALISES SONORES" (sound markers), and "TRANSPORTS EN COMMUN" (public transport). At the bottom, it says "PROCHAINEMENT DISPONIBLE SUR WWW.HANDISCO.COM".

A dark banner featuring the text "Just'In Biker" in large white letters at the top. Below it is the logo for "CITIZEN CLAN" with a stylized infinity symbol. At the bottom, the website address "http://citizenclan.org" is displayed. The background of the banner shows a blurred image of a bicycle wheel.

L'open data au service de l'économie et de l'efficacité de l'action publique

Gagner du temps

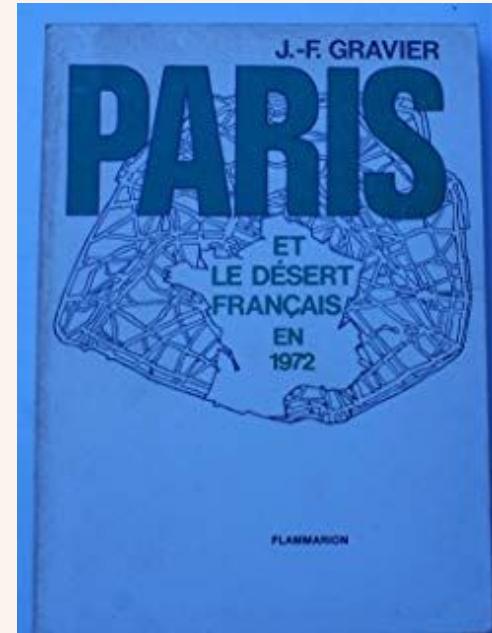
- Répondre moins souvent aux demandes individuelles du fait de la publicité des données



L'open data au service de l'économie et de l'efficacité de l'action publique

Attractivité des territoires

- Les services s'installent dans les villes où les données sont disponibles, complètes et de qualité
- Augmentation des services disponibles dans les villes moyennes, au même niveau que les métropoles mondiales au niveau numérique



L'open data au service de l'économie et de l'efficacité de l'action publique

Efficacité des transports (Transport for London)

- **Pas de frais d'investissements** dans une application d'itinéraires : un point central
- **Pas de frais de campagne** sur ces applications
- Données enregistrées depuis 1971, **excellente connaissance des mobilités des londoniens**
- La qualité des données temps réel a permis d'envoyer **moins d'alertes SMS aux usagers : £ 3M économisés par an**
- L'amélioration de la lisibilité de l'offre du réseau TfL a permis de **générer plus de trajets = £ 20M de bénéfices par an**



L'open data au service de l'économie et de l'efficacité de l'action publique

Efficacité des transports

£ 130 000 000

de valeur créée chaque année grâce aux données de TfL
(étude Deloitte)

L'open data au service de l'économie et de l'efficacité de l'action publique

Efficacité des transports

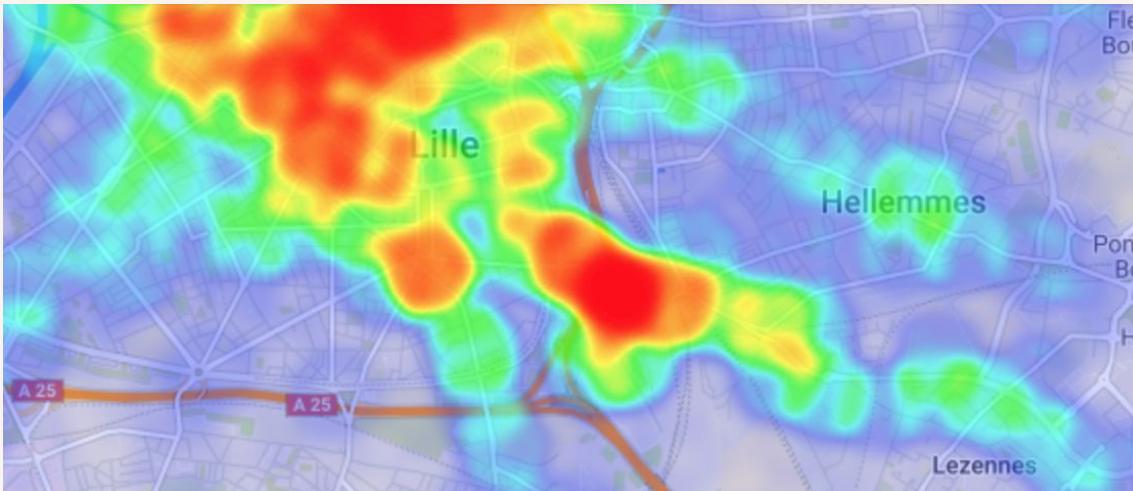
- Exemple lillois n°1 : gestion des vélos en libre service



L'open data au service de l'économie et de l'efficacité de l'action publique

Efficacité des transports

- Exemple lillois n°2 : connaissance des pratiques cyclables



L'open data au service de l'économie et de l'efficacité de l'action publique

Connaissance du territoire par les investisseurs

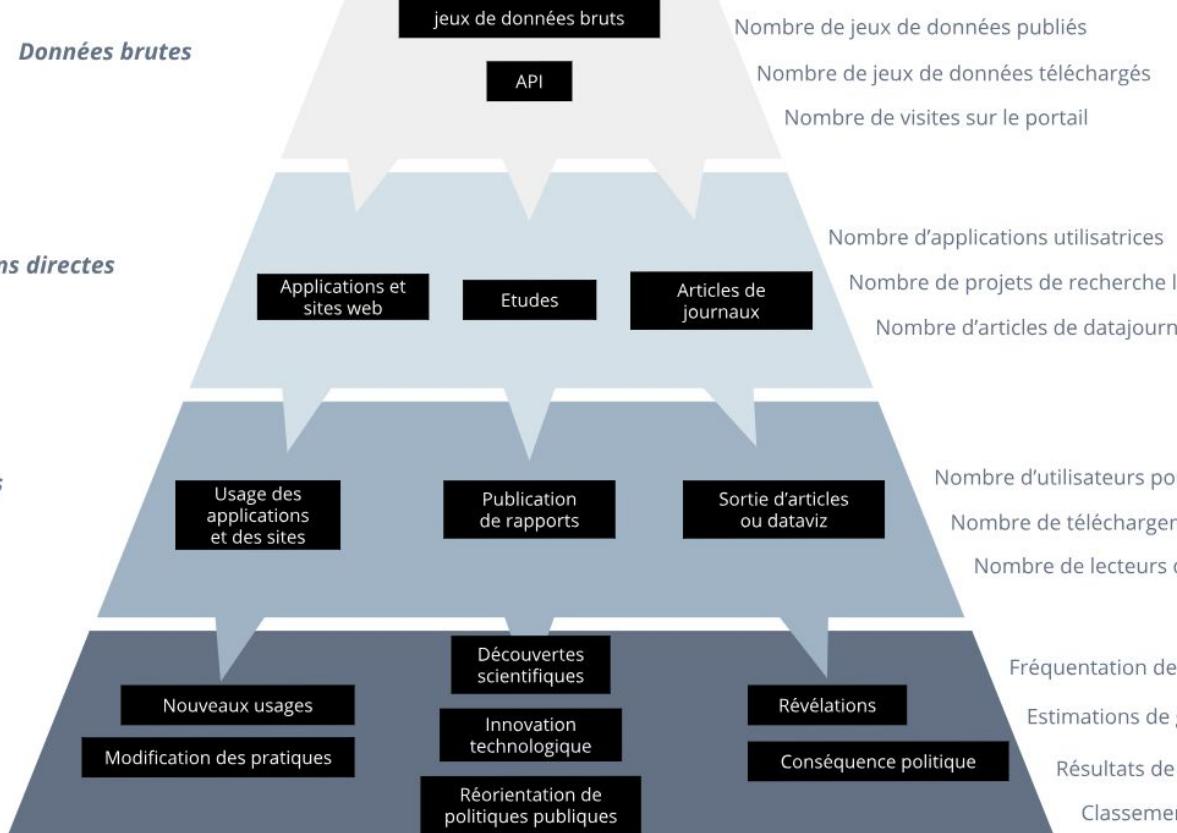
- Exemple lillois n°3 : Gobeebikes



Stade de l'utilisation des données

Portail open data

Avec quels indicateurs évaluer l'impact de l'open data ?



Parcours de la donnée : de sa publication à son impact

- Difficulté de l'évaluation +

L'open data pour créer de la valeur sociale

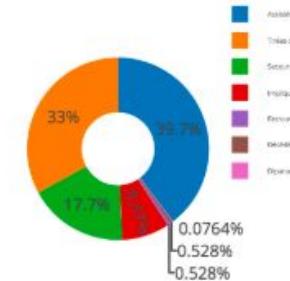
- Exemple : sauver des vies en mer en prévenant les accidents maritimes



5299 interventions géolocalisées a sur la période sélectionnée et sur l'affichée

Choisissez votre visualisation

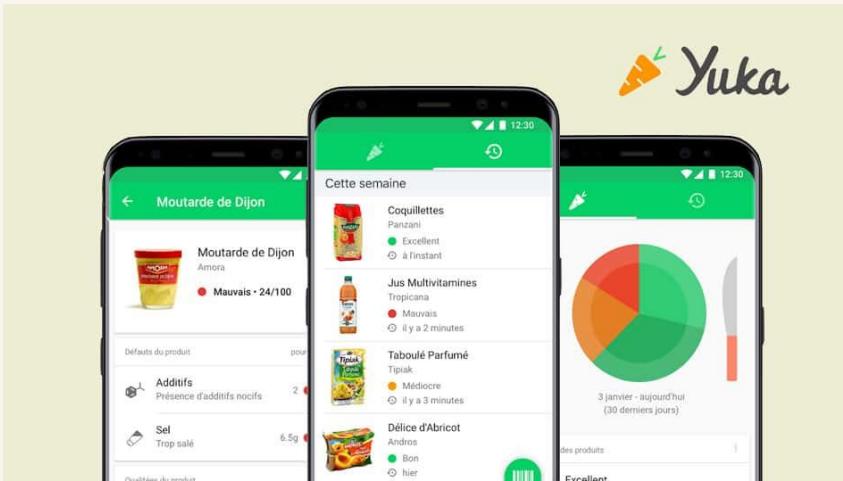
Répartition du bilan humain



Choisissez votre visualisation

Force du vent

L'open data pour créer de la **valeur sociale**



Yuka

8,5 millions d'utilisateurs

2 millions de produits scannés chaque jour

Top 10 des applis gratuites sur l'Appstore et le Playstore

<https://dataactivist.coop/dataweek/#7>

L'open data pour créer de la valeur sociale



Open Food Facts

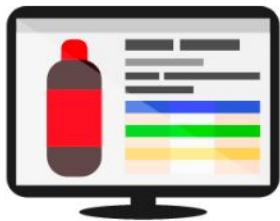
La base de données libre sur les produits alimentaires.

Dis-moi ce que tu manges et je te dirai ce que tu es.

— Jean Anthelme Brillat-Savarin - 1825

Ouvrez votre nourriture pour savoir ce que vous mangez

Participez à notre base de données collaborative, libre et ouverte des produits alimentaires du monde entier !



Une base de produits alimentaires

Open Food Facts est une base de données de produits alimentaires qui répertorie les ingrédients, les allergènes, la composition nutritionnelle et toutes les informations présentes sur les étiquettes des aliments.



Faite par tout le monde

Open Food Facts est une association à but non lucratif composée de volontaires.

Plus de 9000 contributeurs comme vous ont ajouté 600 000 produits de 200 pays en utilisant notre app [Android](#), [iPhone](#) ou [Windows Phone](#) ou leur appareil photo pour scanner les codes barres et envoyer



Pour tout le monde

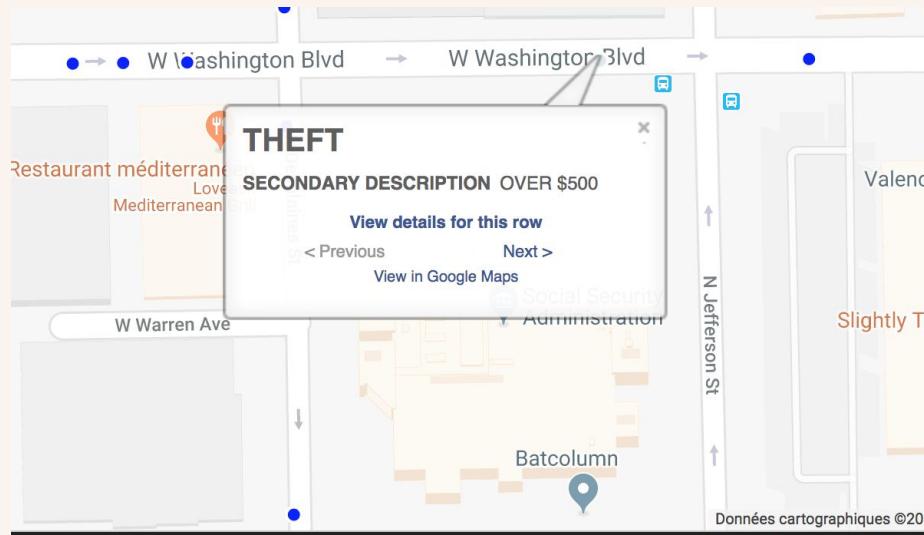
Les données sur la nourriture sont d'intérêt public et doivent être libres et ouvertes.

Toute la base de données est publiée sous forme de données ouvertes (open data) qui peuvent être utilisées par tous et pour tous usages. Allez voir les [réutilisations](#) ou créez la vôtre !

L'open data au service de la transparence

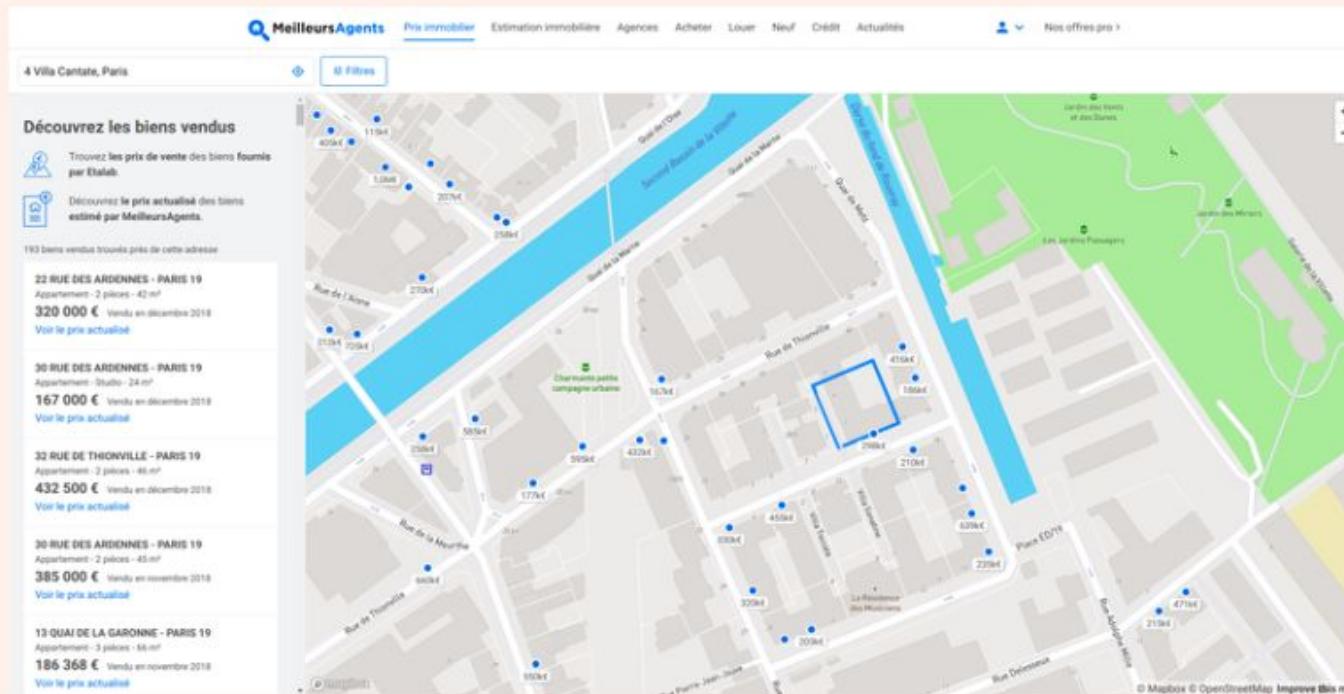
Crime map de Chicago

Une cartographie en libre accès de tous les crimes et délits ayant eu lieu dans la ville



L'open data au service de la transparence

- Les données des **valeurs foncières** reprises par le site meilleurs agents, avril 2019



L'open data au service de la transparence

Data journalisme

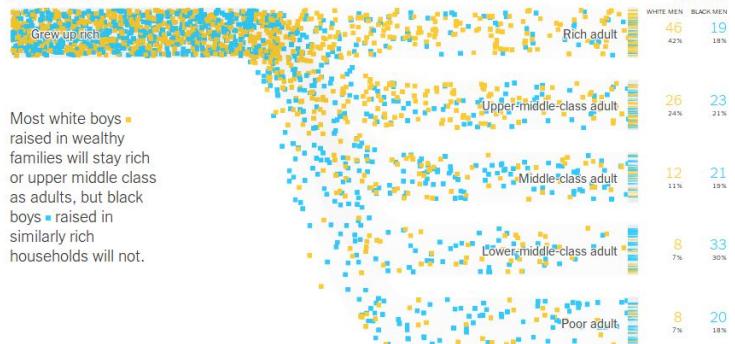
Extensive Data Shows Punishing Reach of Racism for Black Boys

By EMILY BADGER, CLAIRE CAIN MILLER, ADAM PEARCE and KEVIN QUEALY MARCH 19, 2018

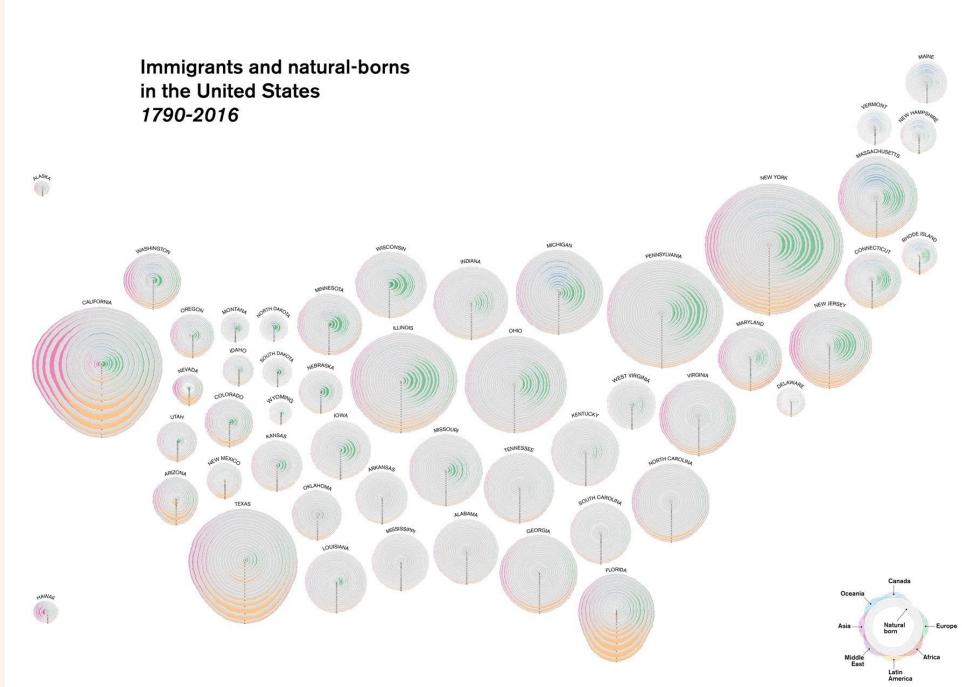
Black boys raised in America, even in the wealthiest families and living in some of the most well-to-do neighborhoods, still earn less in adulthood than white boys with similar backgrounds, according to a sweeping new study that traced the lives of millions of children.

White boys who grow up rich are likely to remain that way. Black boys raised at the top, however, are more likely to become poor than to stay wealthy in their own adult households.

Follow the lives of 2,423 boys who grew up in rich families ...



Immigrants and natural-borns in the United States 1790-2016



[source](#)

DATACTIVIST

[source](#)

L'open data : un enjeu d'intérêt général ?

- **Continuité** : des données accessibles en permanence et mises à jour
- **Adaptabilité** : des données adaptées à la demande (formats, nature, services)
- **Egalité** : des données accessibles à tous, sans restriction ni discrimination

Ouverture : 3 politiques publiques françaises en projet autour de la donnée

Question :

Vrai ou faux ?

Le Ministère de l'Economie et des Finances souhaite lutter contre la fraude en utilisant l'intelligence artificielle pour collecter des données sur les réseaux sociaux

Collecte de données sur les réseaux sociaux pour lutter contre la fraude



Gerald Darmanin • 1st

Ministre chez Ministères de l'Économie, des Finances, de l'Action et des C...

3w

...

NOUS ALLONS ACCÉLÉRER LA LUTTE CONTRE LA FRAUDE

Afin de renforcer la lutte contre les fraudes les plus lourdes telles que les fausses domiciliations fiscales à l'étranger ou le commerce de marchandises prohibées, nous voulons expérimenter la collecte ciblée et informatisée des données en accès libre des réseaux sociaux.

Ce système existe aux Etats-Unis et au Royaume-Uni où il a fait ses preuves : il a permis de recouvrer plusieurs milliards de livres.

La lutte contre la fraude est une priorité pour le Gouvernement. Elle doit s'adapter aux évolutions de la société, ce qui nécessite d'en moderniser les outils. Par cette expérimentation, le Gouvernement entend apporter une réponse aux limites de nos systèmes de détection actuels, dans un cadre rigoureux et contrôlé par le Parlement et la CNIL.

[See translation](#)

Collecte de données sur les réseaux sociaux pour lutter contre la fraude



marc rees 
@reesmarc

Pour se rendre compte de la surface d'attaque de cette collecte/traitement de masse, il suffit de lister certains des sites visés. Oh, des broutilles :

- Facebook,
- Twitter
- LeBonCoin
- eBay
- Instagram
- YouTube
- DailyMotion
- Les Marketplace (Amazon, etc.)
- Airbnb
- etc.

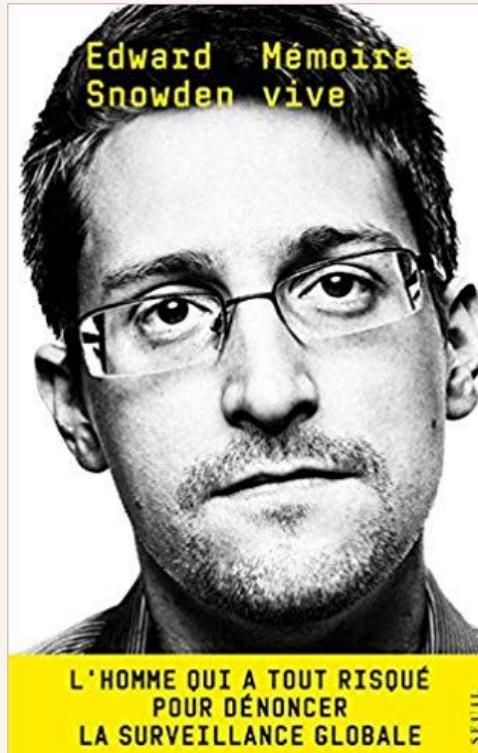
- Pause Edward Snowden -

TOKYO

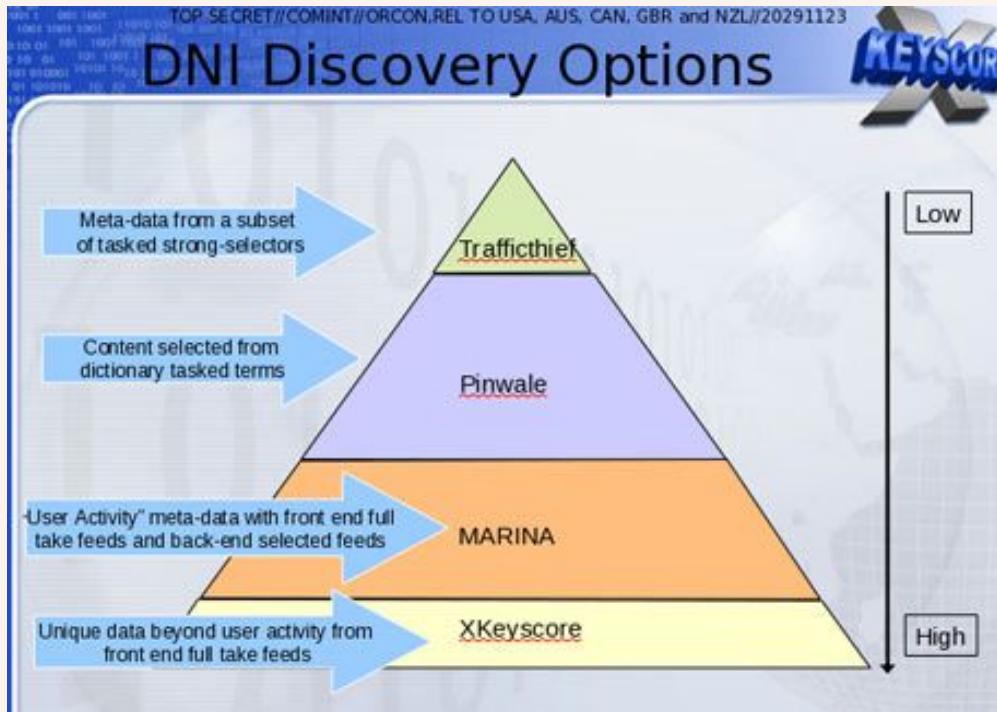
destinataire, etc. Les métadonnées peuvent permettre à celui qui vous surveille de connaître l'endroit où vous avez passé la nuit et à quelle heure vous vous êtes réveillé ce matin-là. Elles permettent de retracer ce que fut votre parcours dans la journée, combien de temps vous avez passé dans chaque endroit visité et avec qui vous avez été en contact.

Tout cela vient contredire l'affirmation gouvernementale selon laquelle les métadonnées ne constituaient pas un accès direct à nos communications. Vu le nombre vertigineux de messages numériques échangés dans le monde, il est tout bonnement impossible d'écouter tous les appels téléphoniques et de lire tous les e-mails. Et même si c'était faisable, ça ne servirait à rien car les métadonnées, en opérant un tri, rendent la tâche inutile. Voilà pourquoi il ne faut pas envisager les métadonnées comme des abstractions inoffensives mais comme l'essence même du contenu : elles sont précisément la première source d'information exigée par celui qui vous surveille.

Par ailleurs, le



- Pause Edward Snowden -



- Pause Edward Snowden -

TOP SECRET//SI//ORCON//NOFORN

Gmail Hotmail™ msn YAHOO! Google™ skype paltalk.com YouTube AOL mail

(TS//SI//NF) PRISM Collection Details

SPECIAL SOURCE OPERATIONS

What Will You Receive in Collection
(Surveillance and Stored Comms)?
It varies by provider. In general:

Current Providers

- Microsoft (Hotmail, etc.)
- Google
- Yahoo!
- Facebook
- PalTalk
- YouTube
- Skype
- AOL
- Apple

Complete list and details on PRISM web page:
Go PRISMFAA

TOP SECRET//SI//ORCON//NOFORN

A large green arrow points from the 'Current Providers' list to the 'What Will You Receive in Collection' section.

Question :

Vrai ou faux ?

Le gouvernement va bientôt proposer une application pour accéder aux services publics en ligne qui oblige les citoyens à utiliser la reconnaissance faciale pour s'identifier

Justifications du projet :



[Article de Libération \(CheckNews\)](#)

- Simplifier l'accès aux services publics en ligne ("*plus besoin de mémoriser plusieurs mots de passe*")
- Permettre une identité numérique authentique et sécurisée
- Selon Christophe Castaner, Ministre de l'Intérieur, cette identité numérique devrait permettre de lutter contre l'anonymat sur Internet et les contenus haineux.
[Source](#)
- "[Dans un entretien paru dans Le Monde du 15 octobre](#), le secrétaire d'Etat au numérique Cédric O, ancien cadre du groupe Safran, a notamment estimé qu'« expérimenter » la reconnaissance faciale était « nécessaire pour que nos industriels progressent »."

Critiques du projet



- La CNIL désapprouve le projet car il oblige tout utilisateur de l'application ALICEM à la reconnaissance faciale. Aucun autre moyen n'est proposé (une visio avec un agent, un face à face...) à la place de la biométrie : incompatible avec l'article 9 du RGPD.
 - *"Il n'y a pas de système à 100% sécurisé. La question n'est pas de savoir quand est-ce qu'on va trouver la première faille, mais combien on va en trouver."*
- Baptiste Robert**, chercheur en cybersécurité
[Source](#)
- Si l'outil et son périmètre d'utilisation sont pour l'instant limités, peut-on garantir que ce sera toujours le cas ?

Question :

Vrai ou faux ?

La ville de Nice est en train d'expérimenter la reconnaissance faciale via la vidéosurveillance

Nice expérimente la reconnaissance faciale via son système de vidéosurveillance



- Un événement qui a accéléré les politiques sécuritaires (déjà en place) : l'attentat de Nice en 2016
- Une expérimentation en cours autorisant le recours à l'intelligence artificielle et la reconnaissance faciale
- La CNIL attend toujours des éclaircissements, “*qu'il s'agisse des taux d'erreur de l'algorithme utilisé, de la qualité des images ou des risques de discriminations (liés aux biais de l'algorithme).*” Sans ces éclaircissements, la CNIL ne pourra pas donner un avis favorable. Sera-t-elle écoutée ?

Source

... Et Saint-Etienne tente d'écouter la rue avec des micros.

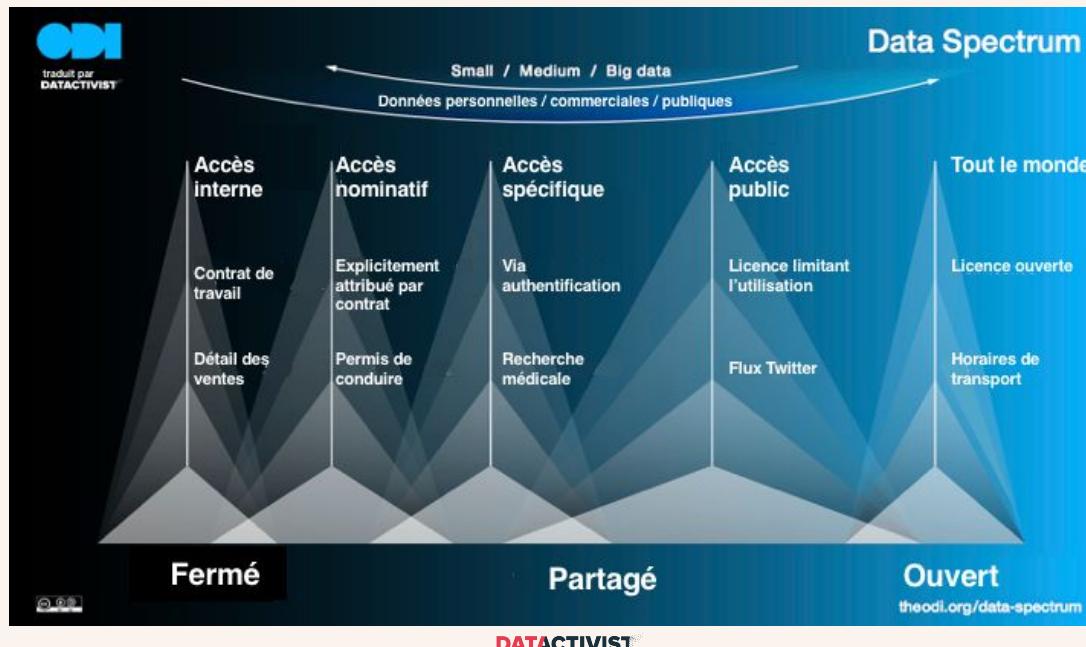


- Des micros, de la taille d'une pièce de deux euros, installés dans un quartier populaire, pour détecter : "klaxons, bris de vitre, coups de feu, cris, bruits de perceuse, perforateurs, chocs, coups de sifflet, bombes aérosols (tags, bombe lacrymogène), crépitements (incendie), explosion, accidents... "jusqu'au chant des oiseaux" cité par un des responsables du fournisseur..." [Source](#)
- La CNIL s'est positionnée contre ce projet, qui non seulement peut rapidement porter atteinte aux libertés individuelles, mais aussi : « conduit à renforcer l'intrusivité du système et le niveau de surveillance dont fait l'objet la population vivant, circulant ou travaillant dans la zone concernée »
- Un projet, s'il est mis en place qui peut aller jusqu'à "altérer les comportements des habitants du quartier" toujours selon la CNIL. [Source](#)
- Une collection de données, permanente, justifiée par le caractère expérimental du projet, sous couvert d'innovation.

En résumé :

Des données qui devraient être publiques mais qui ne le sont pas encore, et **des données qui ne devraient jamais être collectées** malgré les tentations qui existent (autant du secteur privé que public).

À vous de jouer, (avec nous !) en tant que futur.e.s décideuses et décideurs ?



Merci !

Ces slides sont en accès libre :

etienne@dataactivist.st



@Etienne010101

<http://dataactivist.coop>

<http://teamopendata.org>