

Culture de la donnée... ...données de la culture

SciencesPo Aix — 2025-11
@CharlesNepote

Qui suis-je ?

- Début dans le monde de la culture à Paysalp Écomusée de Savoie, en tant que responsable scientifique et du partage des connaissances (1997-1999)
- 2000-2007 : R&D chez BNP-Paribas : pionnier des wikis, déploiement d'intranets, des services web e-commerce
- Début des années 2000 : co-fondé 1er moteur de wiki francophone, première communauté du web sémantique
- 2007-2018 : Directeur de programmes à la FING (think tank dédié aux usages du numérique)
 - programmes d'action sur l'open data et la littératie de données
 - 3 éditions de la Data Literacy Conference à Aix
- 2019-... : Délégué général d'Open Food Facts

5 minutes, 20 données

Collectivement, identifiez 20 données présentes dans cette salle (sans utiliser d'outil numérique).

(Tout ce que vous voyez ou qui vous passe par la tête.)

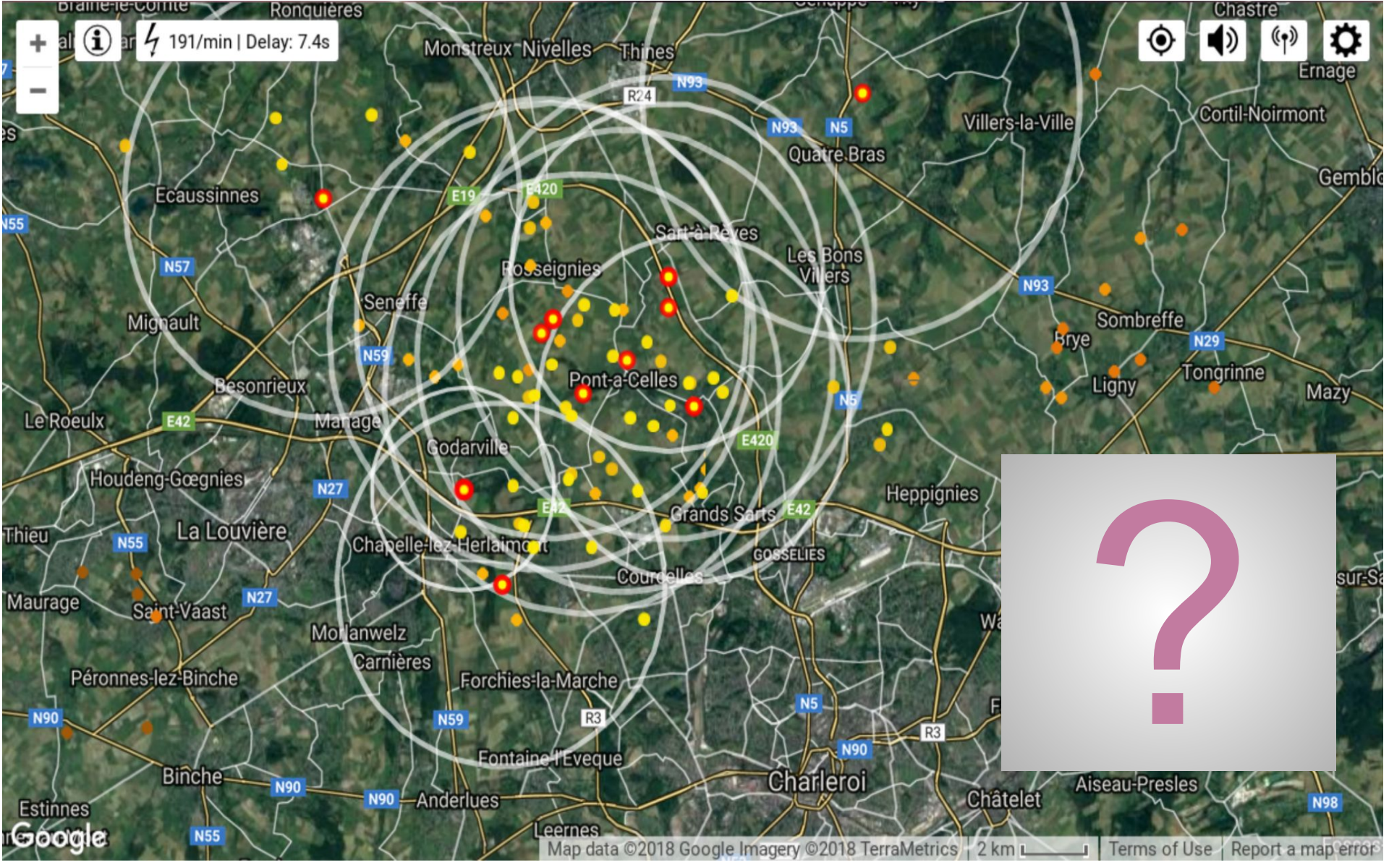
5 minutes, 20 données

Qu'a-t-on appris ?

- nous vivons entourés de données (même en mettant le numérique de côté)
- la polysémie du mot “donné/donnée” complique
- confusion entre un fait (ce qui est donné) et une donnée
 - “quelles sont les données du problème ?” sous-entend “quels sont les faits ?” mais une donnée fausse reste une donnée ; un fait n’est plus un fait s’il est avéré faux
- une donnée n’est qu’une représentation d’une réalité changeante, forgée dans un contexte de production variable (producteur, outils, ...)
- les données peuvent néanmoins avoir une valeur suffisante pour de très nombreux usages

5 minutes, 20 données

- Une méthodologie de temps court
- Un tour de chauffe, un brise glace qui permet :
 - de se placer dans une posture orientée "données"
 - de visualiser concrètement le sujet
 - de se poser des questions sur les données
- Une forme d'introduction à ce qui va nous préoccuper aujourd'hui
 - qu'est-ce qu'une culture ou littératie des données ?
 - quels savoirs et outils pour nous permettre de comprendre et exploiter les données





Maps and statistics

Real Time

- Europe
- Oceania
- America

Information

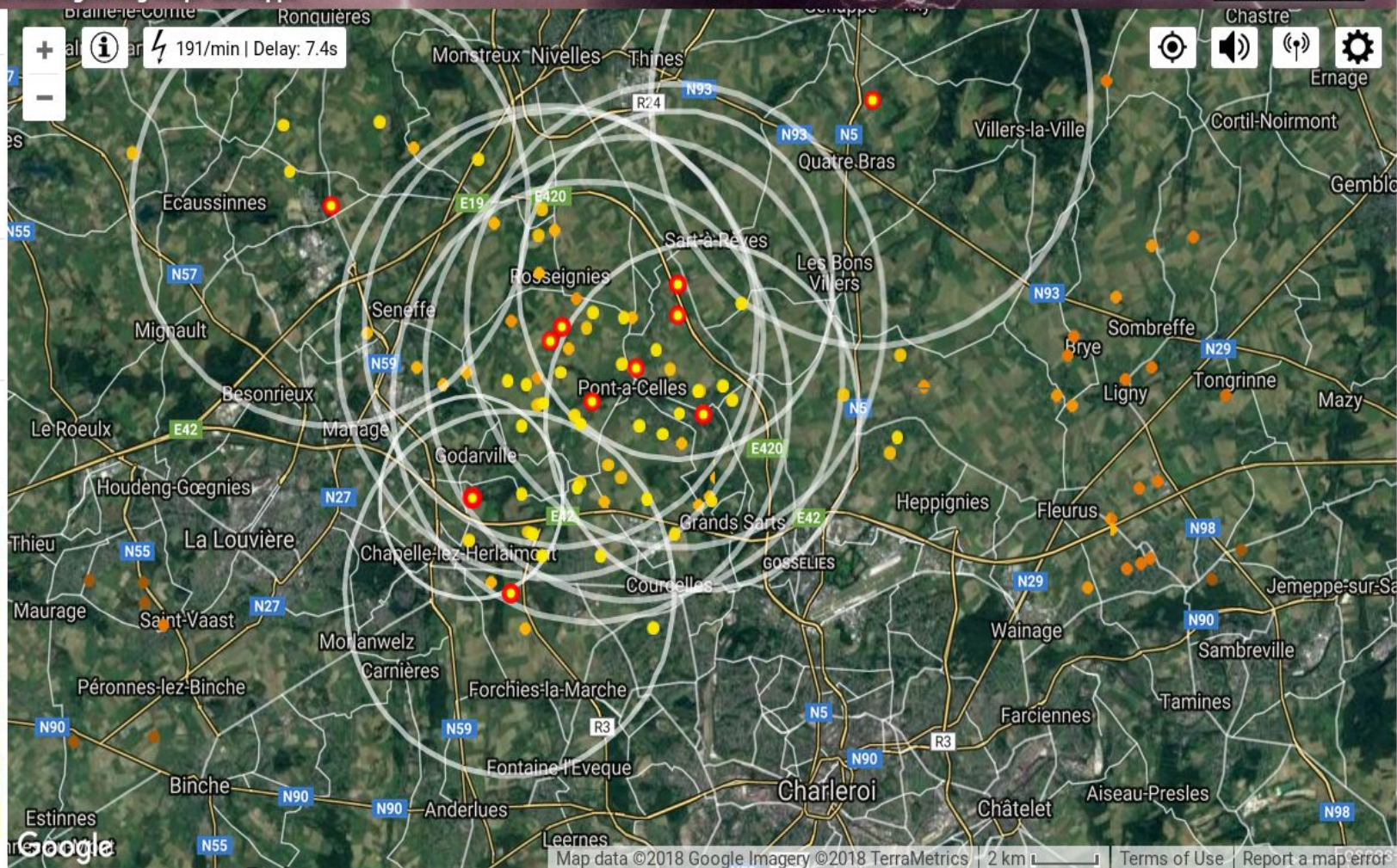
- Extras
- Apps
- About

For Participants

- Login
- MyBlitzortung PHP



Animation



Carte des navires

?
 ?
 ?

Notation & Options d'affichage:

☒ Afficher les noms des navires
☐ Ma flotte
☐ Wind
☒ Plus...

☒ Navire à passagers
☒ Cargo
☒ Navires-citernes
☒ Navire à grande vitesse
☒ Remorqueur, Pilote, etc
☒ Navires de plaisance &

☒ Autres
☒ Navires de pêche
☒ Aides à la navigation
☒ Navires Indéfinis
☒ Navires en route
☒ navires ancrés/amarrés

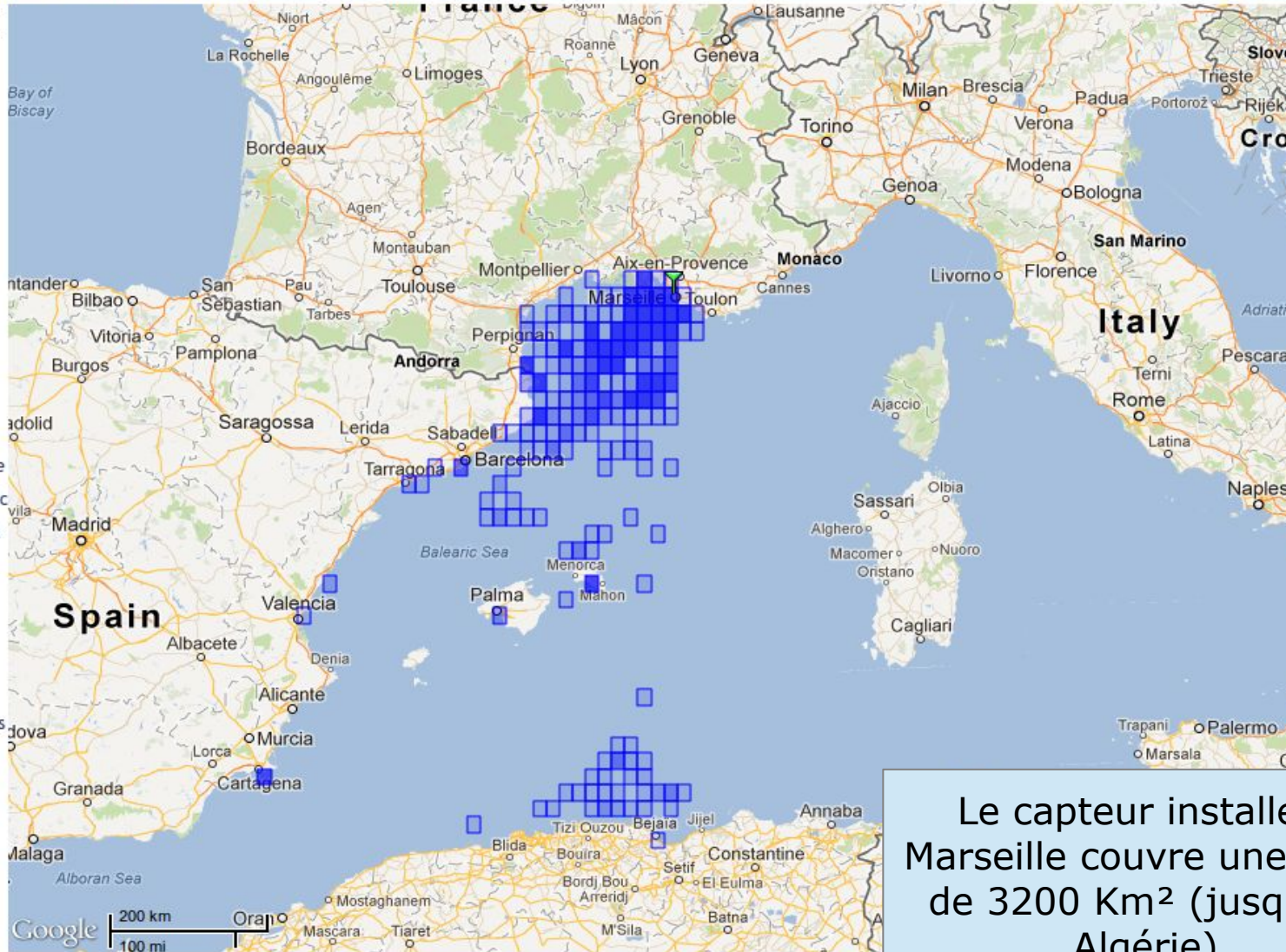
Liens rapides:

[Fourniture gratuite d'un récepteur AIS](#)
[Indiquez votre propre position](#)
[Les stations de réception](#)

Available on the iPhone

App Store

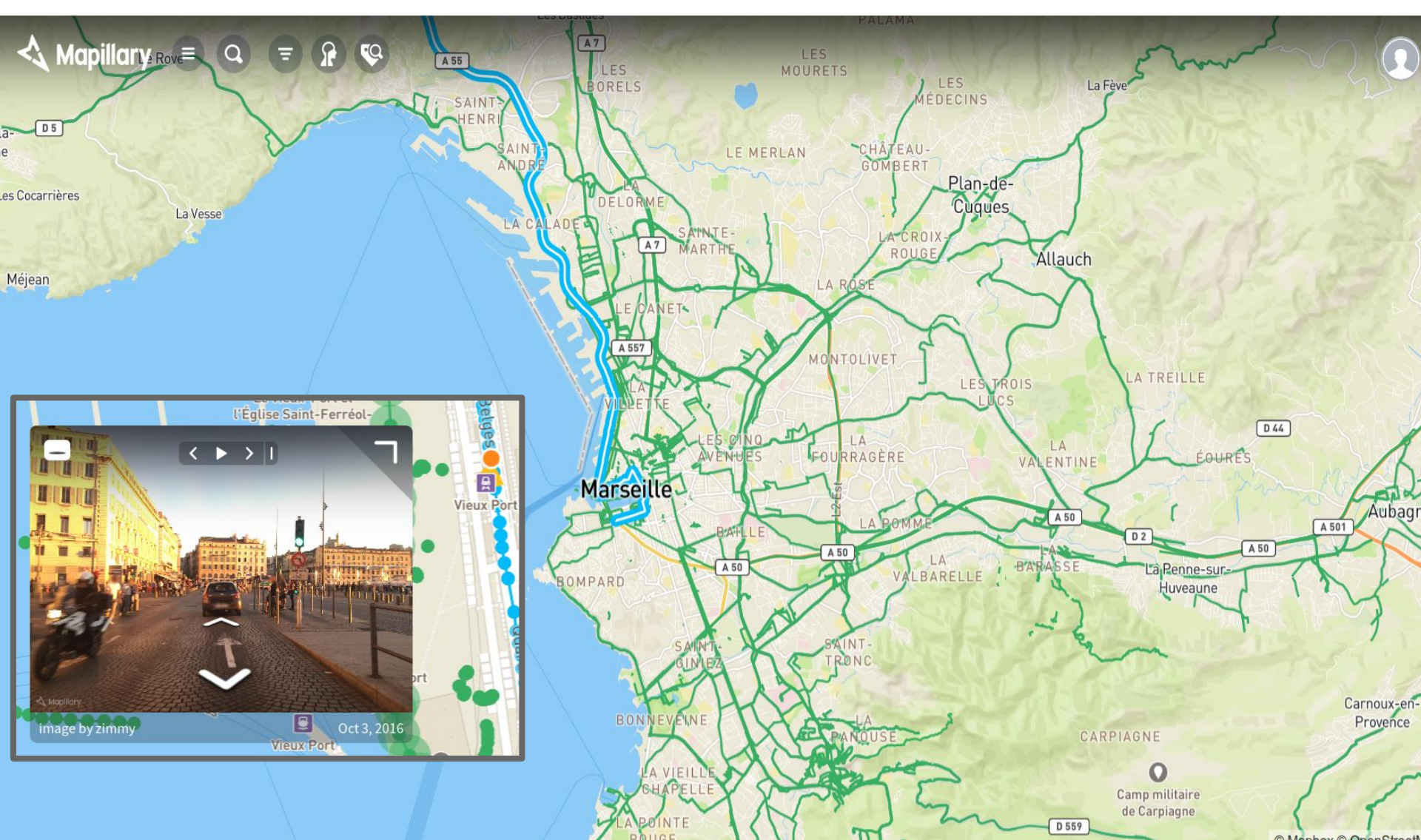
Aidé par:

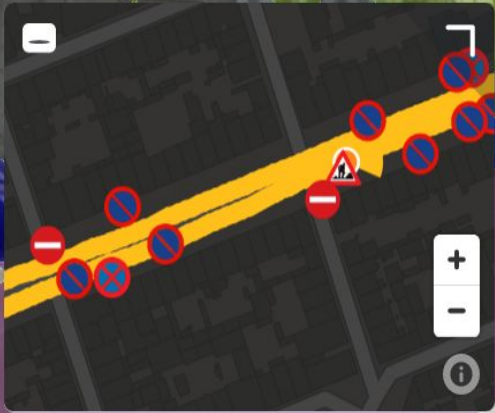
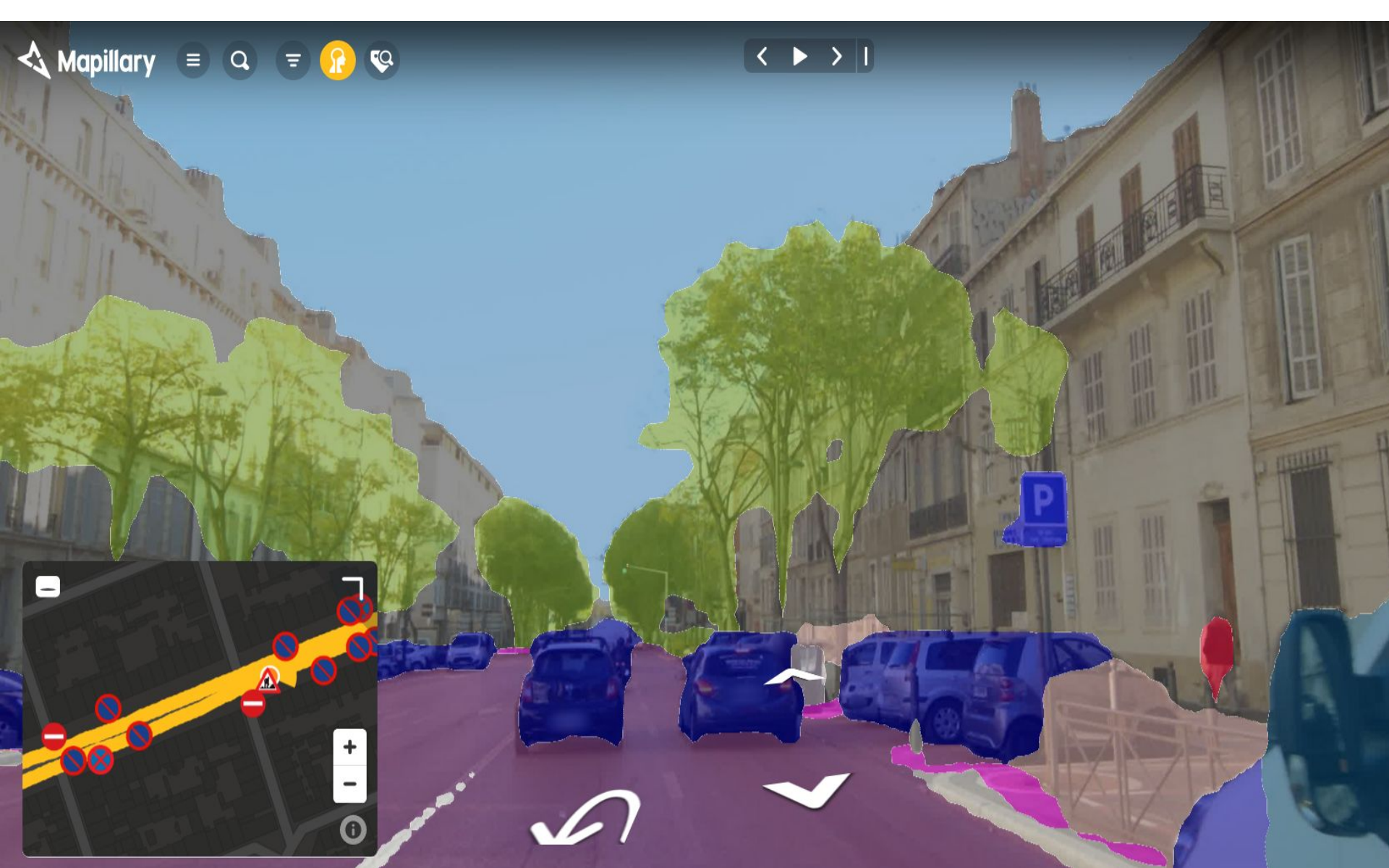


Le capteur installé à Marseille couvre une zone de 3200 Km² (jusqu'en Algérie)

Actualiser dans: " [Actualiser maintenant!](#) " [Navires dans le réseau:](#) [Navires affichés:](#) .

© MarineTraffic.com, [Lisez Les Conditions d'utilisation](#) **AVIS:** Les positions des navires peuvent dater d'une heure ou être incomplètes. Les données sont fournies à titre indicatif. Tous droits réservés. Aucune partie de MarineTraffic peut être copiée, reproduite ou retransmise sous tous supports ou formes, ou par tout autre moyen, sans l'autorisation écrite de MarineTraffic.com.





☐ Only show ideas closely related to my search terms [?](#)

[+ Advanced Options and Filters](#)

Locations: Australia [✕](#)

Lan

Type the characters that appear in the picture below.
Or [sign in](#) to get more keyword ideas tailored to your account. [?](#)

versofts



Letters are not case-sensitive

Search

Select all images with **street signs**.



Verify

Want an easier challenge?

[Why?](#)

La révolution des données en 7 points

1. La **production** : digital labor, capteurs, crowdsourcing, etc.
 2. Les **producteurs** et usagers : les M. Jourdain, les communautés, les objets
 3. La **quantité** : Big Data mais aussi les small data
 4. La **pluralité** des sources (les *proxies*)
 5. La **démultiplication** des usages : l'autonomisation des données
 6. Le **partage** des données (notamment l'open data)
 7. **L'infrastructure** de partage : d'un web documentaire à un web des données
- > L'exclusivité et la vente de données se fait dure à défendre.
 - > Si vous ne partagez pas vos données d'autre le feront à votre place.
 - > Le bouleversement complet du paysage de la data pose un problème culturel



Les données, c'est pas trivial

Les difficultés que l'on rencontre,
les défis que cela nous pose

L'objet “donnée”

- On ne croise pas des données tous les matins en sortant dans la rue
- On ne peut pas “toucher” des données (intangible)
- Un objet ingrat, aride
- Une polysémie problématique qui ne favorise pas le recul critique
 - Une représentation de la réalité et non des faits
 - Bruno Latour avait proposé, sans succès, de parler **d'obtenues**

La réalité des données (souvent)

FRÉQUENTATION DE 2006 à 2010 (Sources : Musées et Sagemor Morbihan en 2009)																
N°REF	ETABLISSEMENTS	LOCALITE	TOTAL 06	GRAT 06	TOTAL 07	GRAT 07	Evolution 07/06 en %	TOTAL 08	GRAT 08	Evolution 08/07 en %	TOTAL 09	GRAT 09	Evolution 09/08 en %	TOTAL 10	GRAT 10	Evolution 10/09 en %
2205001	MUSEE DU CHATEAU DE DINAN	DINAN	22 598		23 000	7 144	2	21 395	8 237	-7	23 706	8 368	11	23 536	18 677	168
2209301	MUSEE MATHURIN MEHEU	LAMBALLE						8 657	1 305		5 081	1 008	-41	6 340	1 208	25
2227901	MUSEE D'ART ET D'HISTOIRE	SANT-BREUC	41 811	41 811	38 992	38 992	-8	35 859	35 859	-7	43 668	43 668	22	39 422	39 422	-10
2301001	MUSEE DES BEAUX ARTS	BREST	17 119	14 102	15 116	12 983	-12	23 175	18 389	53				22 382	18 767	
2303801	ECOMUSEE DES MONTS D'ARREE	COMMANA	23 519	5 867	20 302	4 650	-14	21 522	6 183	6	19 961	5 477	-7	18 594	4 990	-7
2303901	MUSEE DE LA PECHE	CONCARNEAU	35 955	2 343	34 252	3 497	-5	33 844	3 540	-1	32 554	4 475	-4	32 576	4 427	0
2404801	LE PORT-MUSEE	DOUARNENEZ	46 488	10 325	51 648	13 289	11	44 496	13 364	-14	44 657	9 873	0	68 869	39 364	54
2615101	MUSEE DES JACOBINS	MORLAIX	17 836	10 426	19 203	9 605	8	34 410	17 374	79	18 857	8 819	-45	19 792	11 701	5
2915501	ECOMUSEE DE L'ILE D'Ouessant	ILE D'OUESSANT	13 318		11 181	4 799	-16	10 714	3 931	-4	9 864	3 975	-8	10 151	4 336	3
2915502	MUSEE DES PHARES ET BALISES	ILE D'OUESSANT	21 969		20 800	4 711	-5	16 262	3 358	-22	16 174	3 460	-1	16 556	4 041	2
	2	ILE D'OUESSANT	35 287		31 981	9 510		26 976	7 289		26 038	7 435		26 707	8 377	-19
2921701	MUSEE DE PONT-AVEN	PONT-AVEN	36 609		40 902	9 173	12	43 575	9 459	7	53 290	11 480	22	43 273	11 858	-19
2922001	MUSEE BIGODEUEN	PONT-LABBE	11 107	2 031												
2923201	MUSEE DES BEAUX-ARTS	QUIMPER	46 822	20 438	79 123	31 176	70	55 644	26 895	-30	65 048	32 184	17	57 174	28 456	-12
2923202	MUSEE DEPARTEMENTAL BRIETON	QUIMPER	41 153	18 177	49 397	19 649	20	45 111	20 208	-9	42 445	17 711	-6	40 175	11 844	-5
	2	QUIMPER	87 775	38 615	128 520	50 825		100 755	46 903		107 493	49 895		97 348	40 400	
2924901	MUSEE DE L'ECOLE	TREGARIVAN	11 439	1 151	11 756	1 390	3	9 297	1 242	-21	12 012	1 857	29	12 763	2 239	6
3518801	ECOMUSEE DU PAYS DE MONTFORT	MONTFORT-SUR-MAEU									4 516	1 547		5 002	1 451	11
3523801	MUSEE DE BRETAGNE	RENNES	48 277		54 281	14 837	12	99 553	50 774	83	68 943	13 887	-31	47 236	12 806	-31
3523802	MUSEE DES BEAUX-ARTS	RENNES	55 897	42 092	59 316	43 084	6	56 339	33 025	-5	35 146	23 416	-38	57 510	45 479	92
3523803	ECOMUSEE DU PAYS DE RENNES	RENNES	35 042	24 710	41 637	29 401	19	39 026	26 649	-6	46 002	33 635	18	45 885	33 138	-0
	3	RENNES	139 216	66 802	155 234	87 322		194 918	110 448		150 091	70 838		160 631	91 223	
3529801	MUSEE INTERNATIONAL DU LONG-COURS	SAINT-MALO	10 150	3 283	8 937	2 632	-12	10 712	3 462	20						
3529802	MUSEE D'HISTOIRE ET D'ETHNOGRAPHIE	SAINT-MALO	26 764	7 503	44 821	14 142	67	27 355	7 698	-39						
	2	SAINT-MALO	36 914	10 786	53 758	16 774		38 067	11 160							
3519401	MUSEE DE LA PAULICOMNAIE	VITRE	2 303	697	1 641	504	-29	1 292	490	-21	1 374	410	6	1 996	1 611	-86
3536001	MUSEE SAINT-NICOLAS	VITRE	2 886	914	2 209	799	-23	1 960	565	-11	1 961	621	1	1 904	1 621	-9
3536002	MUSEE DU CHATEAU	VITRE	25 744	7 701	17 669	5 970	-31	16 817	6 697	-5	21 363	7 628	27	20 992	6 978	-2
3536003	CHATEAU-MUSEE DES ROCHERS-SEVIGNE	VITRE	8 996	3 075	7 661	2 173	-15	7 200	2 257	-6	8 024	2 384	11	9 223	5 422	15
	4	VITRE	39 929	12 387	29 180	9 446		27 269	10 009		32 742	11 043		32 215	14 202	
5602301	ECOMUSEE DE SAINT-DEGAN	AURAY-BRECH	12 969	1 046	14 529	2 487	12	14 727	2 937	1	14 824	2 867	1	13 031	2 306	-12
5603401	MUSEE DE LA PREHISTOIRE	CARNAC	32 273	5 328	32 317	6 136	0	31 314	6 157	-3	30 036	5 510	-4	31 156	5 824	4
5606901	ECOMUSEE DE L'ILE DE GROIX	GROIX	7 870	1 355	9 520	1 496	21	8 196	1 028	-14	7 171	940	-13	5 171	980	-28
5609001	ECOMUSEE INDUSTRIELLES FORDS	INZINZAC LOCHIST	5 472	1 328										5 000	2 270	
5618102	MUSEE DE LA COMPAGNIE DES INDES	PORT-LOUIS	61 443	22 162	62 050	20 625	1	58 312	19 194	-6	68 683		18	61 299	25 472	-11
5622801	MUSEE DE LA RESISTANCE BRETONNE	ST MARCEL MALESTROFF														
5624201	MUSEE DU CHATEAU DE SUSCINO	SARZEAU	90 721		95 482	15 204	5	91 603	16 303	-4	103 472		13	Retrait appellation		
5626003	MUSEE DE LA COHUE	VANNES	30 447	21 974	44 747	28 723	47	36 662	26 191		15 677	9 083		18 667	9 513	
5626002	MUSEE D'HISTOIRE ET D'ARCHEOLOGIE	VANNES	15 155	12 594	4 468	1 441	-71				5 271	2 903		5 827	3 154	
	2	VANNES	45 602	34 568	49 215	30 164		36 662	26 191	-26	20 948	11 986	-43	24 494	12 667	17
33	MUSEES		898 595	287 618	973 270	360 503		952 395	380 210		921 947	293 688		860 732	372 350	
		rs réponses par année	34		34			34			34			33		

Données confidentielles - Contacter le chef d'établissement

Augmentation de la fréquentation en 2010 grâce à l'exposition consacrée à Doisneau du 10 juillet au 30 septembre qui a accueilli 37 357 personnes

Fréquentation 2007 en baisse à cause de la dislocation d'un marché du ferroc sur le site de St-Rivoal, perte d'environ 3 000 visiteurs durant l'été, des conditions climatiques qui ont amoindri certaines visites et actions en plein air, et de la fermeture pour travaux durant 4 mois

Travaux sur bateaux à flot, ouverture début juillet au lieu de début avril 2008, ce qui explique la baisse de la fréquentation. Progression de la fréquentation en 2010 grâce aux fêtes maritimes, entrée libre au musée du 22 au 25 juillet

Augmentation de la fréquentation du musée en 2008 grâce à l'exposition "Mathurin Méheut" organisée du 13 juin au 10 novembre qui a accueilli 20 624 visiteurs

L'exposition Yves Tanguy du 29 juin au 30 septembre a accueilli 50 251 visiteurs en 2007

Les gratuits de 2009 correspondent aux jeunes de moins de 14 ans, à la nuit des musées et aux journées du patrimoine

Réouverture mi-mars 2006. Fréquentation importante de l'exposition "Le Roi Arthur, une légende en devenir" du 17/07/08 au 04/01/09 : 67 970 visiteurs en 2008

Des données de qualité variable

Remarque : outre celles que nous voyons ici hors de France, nombreuses sont situées dans le Golfe Guinée, dans une île appelé **Null Island** ([Wikipedia](#))



Les Mairies de France ? (par Christian Quest)

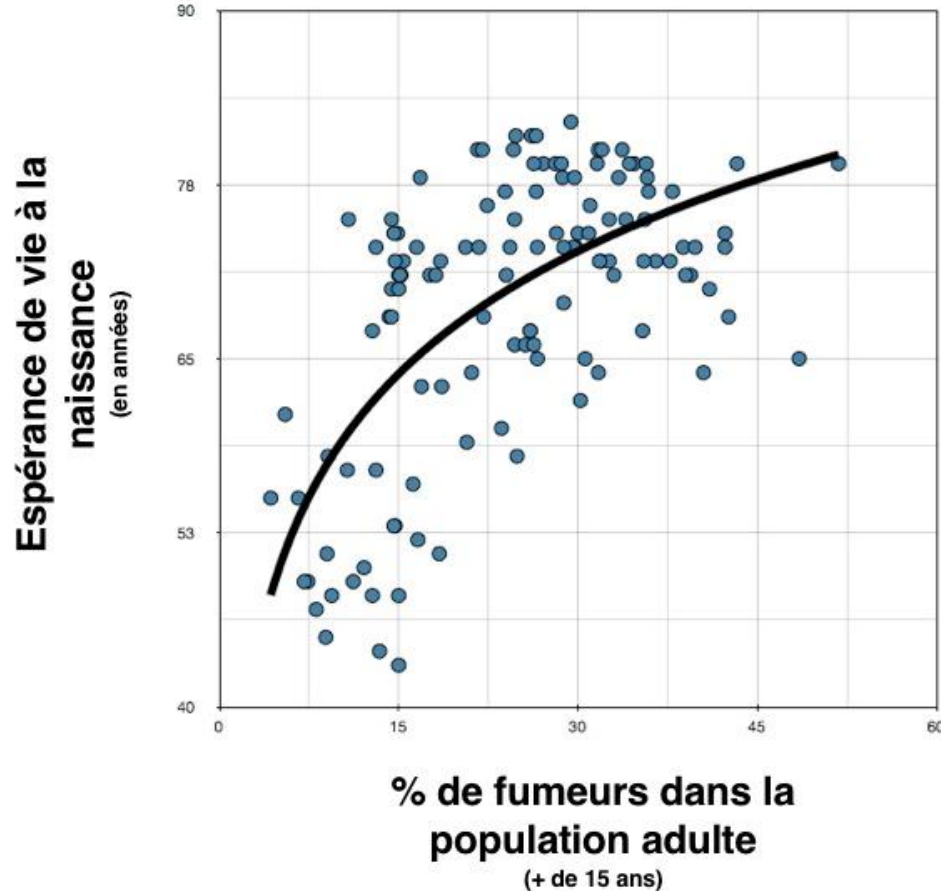
Des données peu standardisées

- Touche la plupart des données hors quelques domaines comme :
 - les transports (GTFS)
 - les données géographiques (GPX, KML, etc.)
 - les oeuvres (Dublin core, MARC, etc.)
 - ...
- Un des gros problèmes de “l’offre” actuelle
- Des conséquences lourdes
 - pas de capitalisation des savoirs techniques
 - des données difficiles à croiser
 - un marché et des usages qui peinent à décoller

Des données pas faciles à trouver

- Les liens profonds sont souvent mal référencés sur les moteurs de recherche, il faut donc *d'abord* identifier la source qui peut héberger les données...
- Exemple : je cherche une liste un peu exhaustive des lieux culturels du Calvados
 - [Google](#) ne va pas m'apprendre grand chose...
 - ... ah si j'ai les [lieux culturels en Normandie](#) sur le site data.culture.gouv.fr mais il va me falloir trier tout ça
 - [Enfin](#) (en y passant un peu de temps)

Des données avec lesquelles on raconte n'importe quoi ?



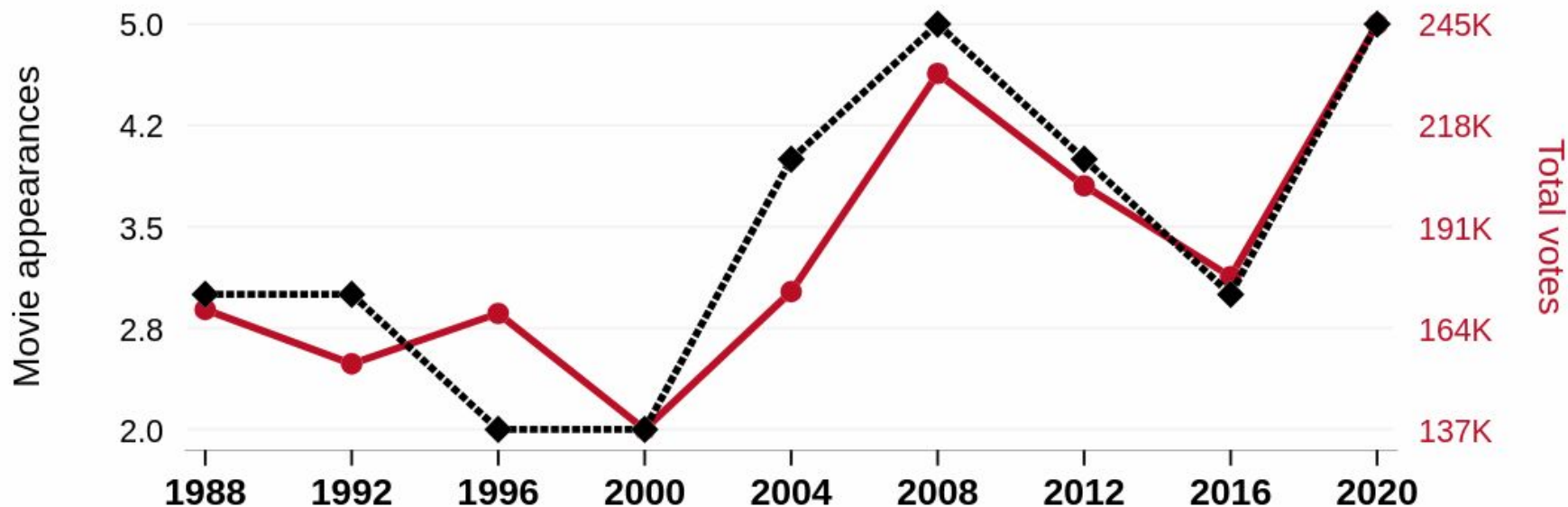
Corrélation ne fait pas causalité !

Source des données : OMS via Gapminder

The number of movies Tom Cruise appeared in

correlates with

Votes for the Democratic Presidential candidate in Montana



◆ The number of movies Tom Cruise appeared in · Source: The Movie DB

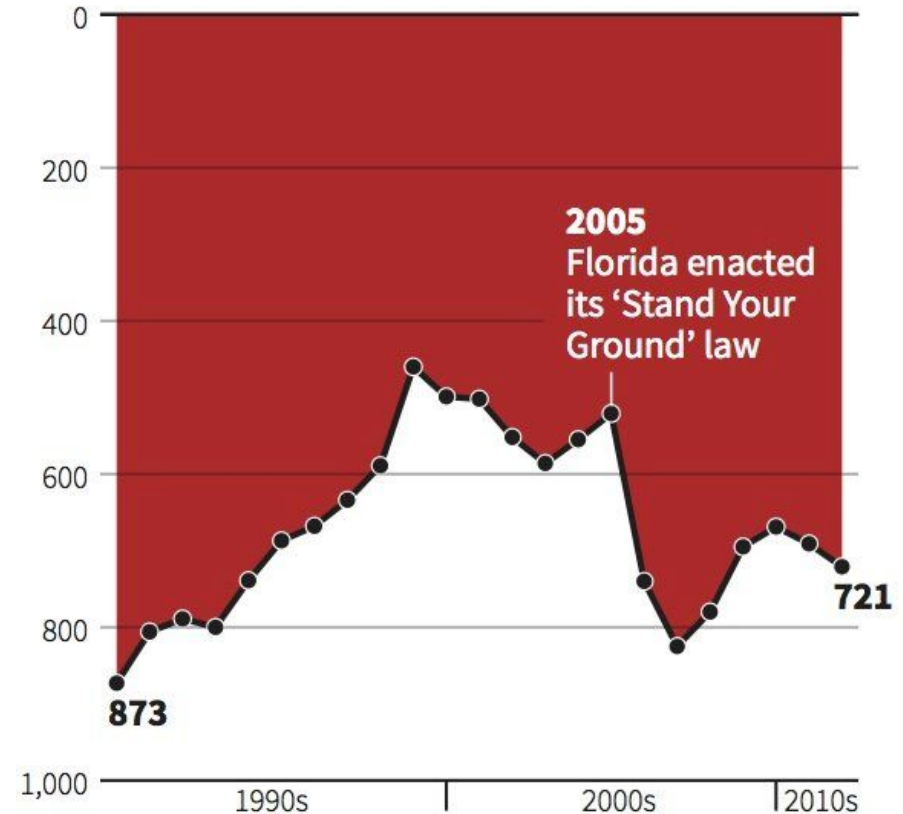
● Total votes cast for the Democrat Presidential candidate in Montana · Source: MIT Election Data and Science Lab, Harvard Dataverse

1981-2020, $r=0.897$, $r^2=0.804$, $p<0.01$ · tylervigen.com/spurious/correlation/5862

Les morts par arme à feu
après le vote de la loi
“Stand Your Ground” en
Floride ?

Gun deaths in Florida

Number of murders committed using firearms



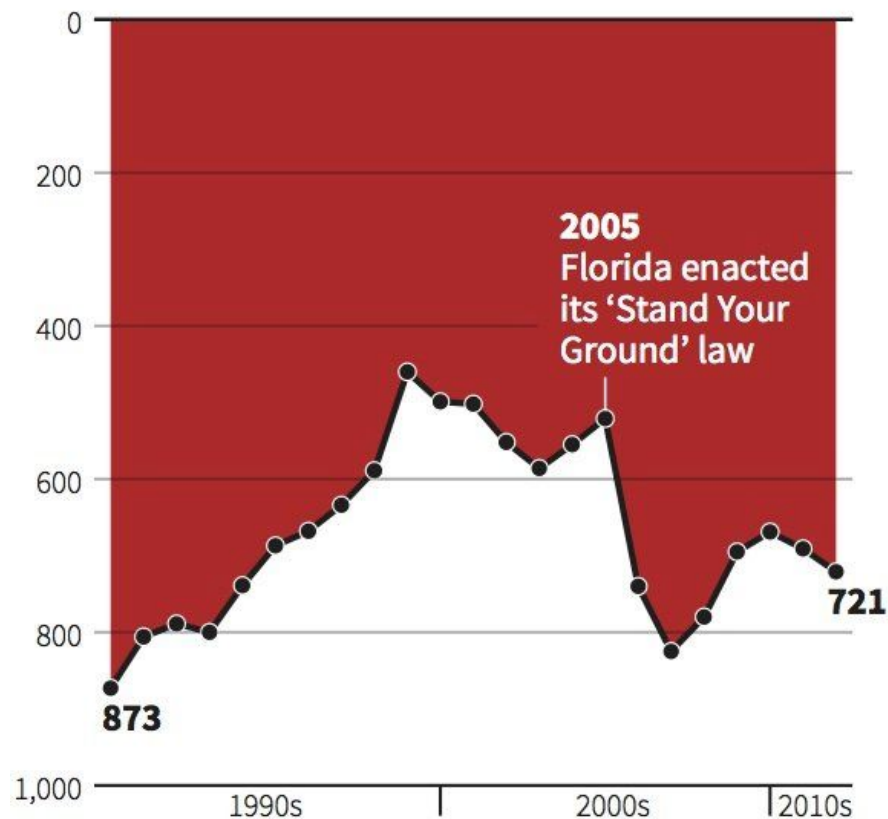
Source: Florida Department of Law Enforcement

**Des
représentations
qui manipulent**

**L'échelle est
inversée par
rapport à
l'usage
classique**

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

FausseS informations en ligne : les adolescents « facilement dupés », selon une étude

Des chercheurs de l'université de Stanford qualifient la capacité de raisonnement des jeunes sur l'information en ligne de « désolante », étude à l'appui.

LE MONDE | 23.11.2016 à 13h12 • Mis à jour le 24.11.2016 à 06h40 |

Par **Morgane Tual**

... et les adultes aussi



"death by gps"



Web

Images

Shopping

Vidéos

Actualités

Plus ▾

Outils de recherche

Environ 19 800 résultats (0,23 secondes)

'Death by GPS' in desert - Travel - The Sacramento Bee

www.sacbee.com/2011/.../death-by-gps-in-desert.htm... ▾ Traduire cette page

30 janv. 2011 - "It's what I'm beginning to call **death by GPS**," said Death Valley wilderness coordinator Charlie Callagan. "People are renting vehicles with ...

'Death By GPS' Increasing In America's Wilderness - Slashdot

tech-beta.slashdot.org/.../death-by-gps-increasing-in-a... ▾ Traduire cette page

The Sacramento Bee quotes Death Valley wilderness coordinator Charlie Callagan: "It's what I'm beginning to call **death by GPS** ... People are renting vehicles ...

Experts Warn of 'Death by GPS' as More People Visit ...

www.foxnews.com/tech/2011/02/04/death-gps-rise/ ▾ Traduire cette page

4 févr. 2011 - "It's what I'm beginning to call **death by GPS**," Callagan told the Sacramento Bee. "People are renting vehicles with GPS and they have no idea ...

"Death by GPS" again. Almost. - RoadTrip America

www.roadtripamerica.com ▸ ... ▸ Topics ▸ Gear-Up! ▾ Traduire cette page

10 nov. 2013 - 6 messages - 5 auteurs

I just read a story from the western Montana newspaper "Missoulian". It tells of a young

Des données pas faciles à manipuler

- Gros fichiers : exemple [fichier des licenciés sportifs par commune de France](#) : fichier .csv de 236 Mo pour 2022
 - Ne s'ouvre pas dans votre Excel/LibreOffice
- TRÈS GROS fichiers : exemple, le [DAMIR](#) (Dépenses d'assurance maladie hors prestations hospitalières)
 - un fichier de 30 Go pour une année
 - 1 milliard 500 millions de lignes pour 6 ans
- formats très techniques (OpenStreetMap, Inspire, etc.)

Littératie ...

Acculturation ...

Médiation ...

Éducation ...

... aux données



**Surpasser américains et européens en partant d'un pays
exsangue ?**

Japon 1960' : l'invention d'une culture, la qualité.

La littératie de donnée, un avantage concurrentiel ?



DAVID SELINGER
CONTRIBUTOR

Founder and CEO of
RichRelevance

Data Driven: What Amazon's Jeff Bezos Taught Me About Running a Company

While conventional wisdom has held that customer service is Amazon's secret sauce, Bezos's core innovation was to place data at the center of his corporate culture. Despite big data's hype, Bezos is unique in his data focus. The rest of us are using buzzwords to try to keep up.

David Selinger

<https://www.entrepreneur.com/article/237326>

1. Metrics define your corporate culture.
2. Listening to data starts at the top.
3. Democratize data.



Dan Ariely ✓

6 janvier 2013 · 🌐

📡 S'abonner



Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...

👍 J'aime

💬 Commenter

➦ Partager

👍 😂 ❤️ 2,5 K

1 231 partages

124 commentaires

Les données massives c'est comme le sexe à l'adolescence : tout le monde en parle, personne ne sait vraiment comment ça marche, tout le monde pense que les autres le font, alors tout le monde clame qu'il le fait ...

Dan Ariely, Université de Duke

<https://www.facebook.com/dan.ariely/posts/904383595868>

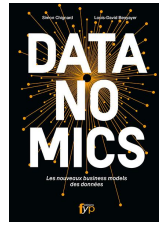
L'approche “stratégique”

À la recherche du “data mindset”

Le “data mindset”

- Une composante critique des organisations qui utilisent le mieux les données
- Kenneth Cukier : “*The biggest challenge is not technology, but mindset*” (Big Data, 2013)
- Le “penser données” c’est la capacité d’une organisation à tirer parti des données, des siennes mais aussi celles des autres (détenir des données ne suffit pas)
- Le “data mindset” est-il inné ou acquis ?

Développer un “data mindset”



- **Synchroniser** des compétences parfois éloignées (stratégique, sociale, statistique, informatique) - recruter des datascientists ne suffit pas
- Penser la donnée **comme un actif**, sous l'angle de **la liquidité** (et non comme un patrimoine) : buy, build, invest, sell
- Développer une stratégie de **partage et d'ouverture** (selon 3 univers de diffusion), animer un écosystème
- Changement de posture : expérimenter et pas seulement planifier (**on cherche à apprendre**, pas seulement à réussir) - expérimenter c'est poser puis tester 10 hypothèses contradictoires
- Explorer et exploiter simultanément : des boucles courtes (ex. tests a/b)



***Toutes les équipes
devront désormais
exposer leurs données
et leurs
fonctionnalités par
des interfaces de
services à travers le
réseau (API). [...]***

***Quiconque ne
respectera pas ces
principes sera viré.
(2002)***

Amazon : l'entreprise "API"



Les tests AB, un des secrets de Google



50% of visitors
see variation A



Variation A



26%
Conversion Rate



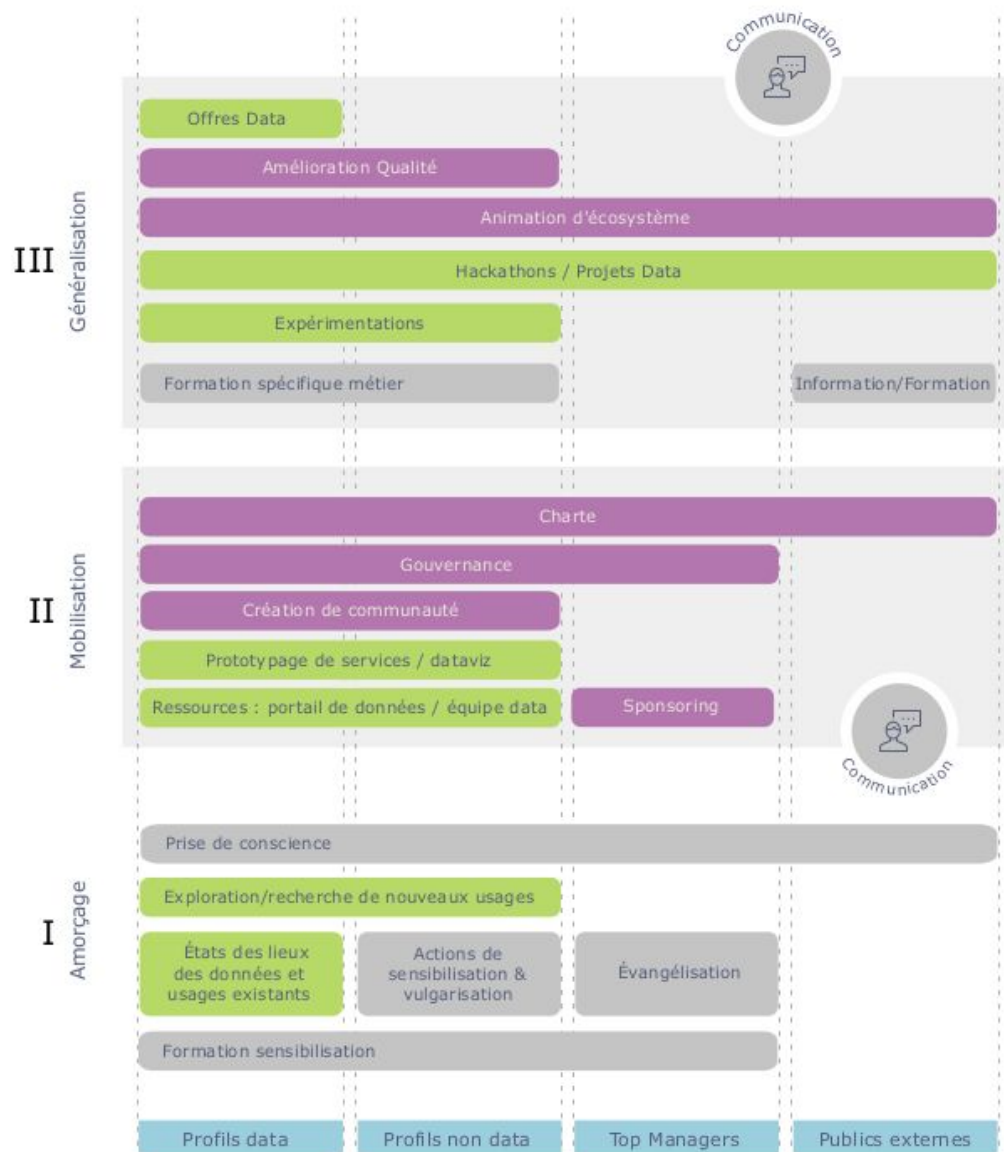
50% of visitors
see variation B



Variation B



18%
Conversion Rate



Déployer une culture de la donnée dans l'organisation, publics et phases

<http://infolabs.io/cd16>

Les données ouvertes (open data)

À la croisée des idéaux français et du pragmatisme anglo-saxon

- 1789 : DDHC, art. 15
 - La société a le droit de demander compte à tout Agent public de son administration.
- 1966 : Freedom of Information Act, É.-U.
- 1978 : loi CNIL mais aussi loi CADA
- 2003-2007 : directive PSI et INSPIRE
- 2007 : rencontre de Sébastopol (grand principes de l'Open Government Data)
- 2016 : Loi Lemaire : ouverture obligatoire

data.gouv.fr



Le catalogue des données françaises

Découvrez des données sur tous les sujets, produites par l'administration et la société civile.

71K

JEUX DE DONNÉES ET API

375K

FICHIERS

6K

ORGANISATIONS

ILS PUBLIENT DES DONNÉES SUR DATA.GOUV.FR



Voir les données

Comment publier des données ?

JEUX DE DONNÉES



Base Sirene des entreprises

[Insee](#)



Répertoire national des élus

[Ministère de l'Intérieur et des Outre-Mer](#)



Données hospitalières COVID-19

[Santé Publique France](#)



Prix des carburants

[Ministère de l'Économie, des Finances et de la S...](#)



Parcelles en agricultures bio

[Agence Bio](#)

L'approche “opérationnelle”

Profession Data
Scientist

Vont-ils sauver votre
organisation ? le
territoire ? le monde ?

	Sale	Buy	Grow
Gold	\$647.00	\$984.51	39.80%
Platinum	\$1881.00	\$2559.78	36.10%
Silver	\$774.00	\$1061.93	37.20%
Copper	\$616.00	\$837.76	36.00%
Steel	\$449.00	\$537.90	19.80%
Beryllium	\$243.00	\$795.34	33.00%
Manganese	\$598.00	\$795.34	33.00%
Aluminum	\$299.00	\$354.61	18.60%
Chrome	\$666.00	\$727.23	9.20%
Nickel	\$421.00	\$453.84	7.80%
Bauxite	\$240.00	\$293.33	22.20%

Cotton	\$162.00	\$196.34	21.20%
Flax	\$372.00	\$182.32	5.00%
Troities	\$243.00	\$330.48	36.00%
Wool	\$269.00	\$359.69	33.80%
Fur	\$116.00	\$118.55	2.20%
Sateen	\$201.00	\$246.43	22.60%
Silk	\$177.99	\$184.79	4.40%

Oil	\$609.00	\$811.19	33.20%
Gas	\$516.00	\$708.98	37.40%
Electric power	\$578.00	\$808.04	39.80%





*“The Data
Literacy
Company”*

Enabling Companies to Build a **Data Literate** Workforce

Learn About Tuva's Data Literacy Solutions

K-12 Schools &
Districts >

Business >

Sustainable
Development >

Data Literacy Will Be the Most Critical Skill for the Workforce in 2030

That's why we started a global community dedicated to creating a data-literate world.

READY TO START YOUR JOURNEY? →



WE OPEN DATA, WE MAKE THEM USEFUL

**QUI SOMMES
NOUS ?**

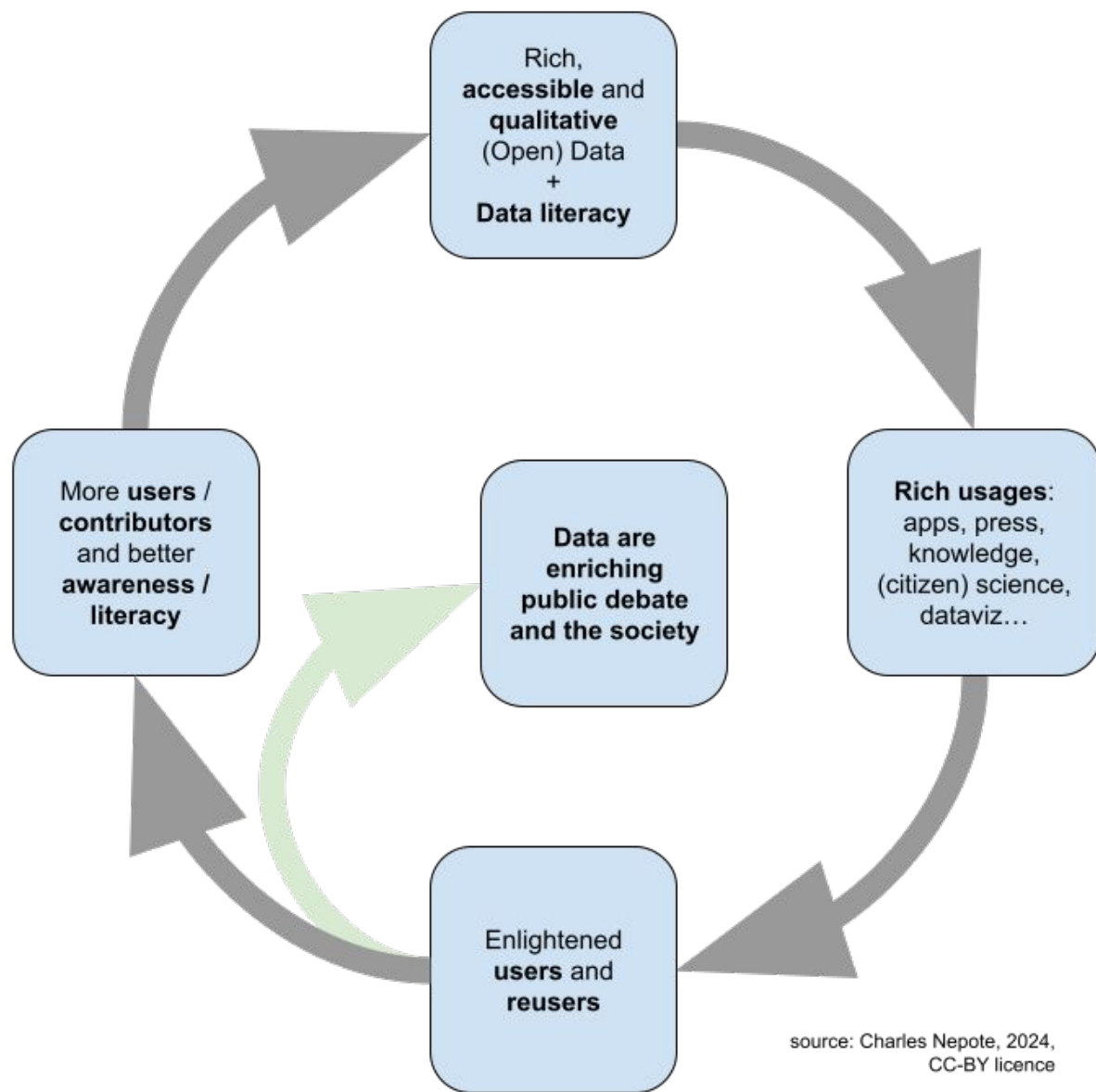


La médiation aux données : quelques principes

1. Rendre lisible, tangible, sensible, à l'échelle des individus
2. Manipuler, **apprendre par le “faire”**
3. **Produire des données** : un bon producteur/ouvreur de données est un bon usager/réutilisateur
4. Une démarche **ouverte et collective**, parce qu'avec des choses si neuves l'apprentissage collectif est tellement plus efficace

Le nouveau monde de données requiert de nouvelles médiations qui sont en cours d'invention, de test.

Le cercle vertueux de l'acculturation aux données



source: Charles Nepote, 2024,
CC-BY licence

Expérimentation orientée "projet"

Zooniverse

Challenges de
data science
communautaires

[Kaggle.com/comp
etitions](https://www.kaggle.com/competitions)

Tests A/B

Le mode beta,
délivrer mais sans
s'engager

Wanted Data
Workshop, parcours
d'évaluation

*Méthodo : liste des
données idéales*

Projet MAIF & GO

Outils de dataviz

Datalab/infolab
interne

Hackathon interne

L'infolab externe
le Tubà, la Coop
Infolab, ...

Hackathon

APIs & Système
d'information
*Service Oriented
Architecture*

Kits de dataviz
physique

Cartopartie : carto
collaborative

Scan partie :
collecte de
produits

Expédition de
données

Datablitz : collecte
collaborative
intensive

Plateforme data
ouverte : Open
Data, Datanova...

*Où sont les data ?
Recensement
express*

Data Storytelling

Datacamps,
barcamps, camps...

Concours de
réutilisation :
Dataconnexions...

Projet MesInfos,
phase d'expé

creadata.fr

WTF CSV

*L'interview de la
donnée*

Expérimentation
orientée "données"

Datalake

Projet Data Shaker
SNCF

en italique les projets/outils "live" ou prêts à l'emploi

Projet
"léger"

Projet
"lourd"

À titre d'exemple : *l'interview de la donnée*

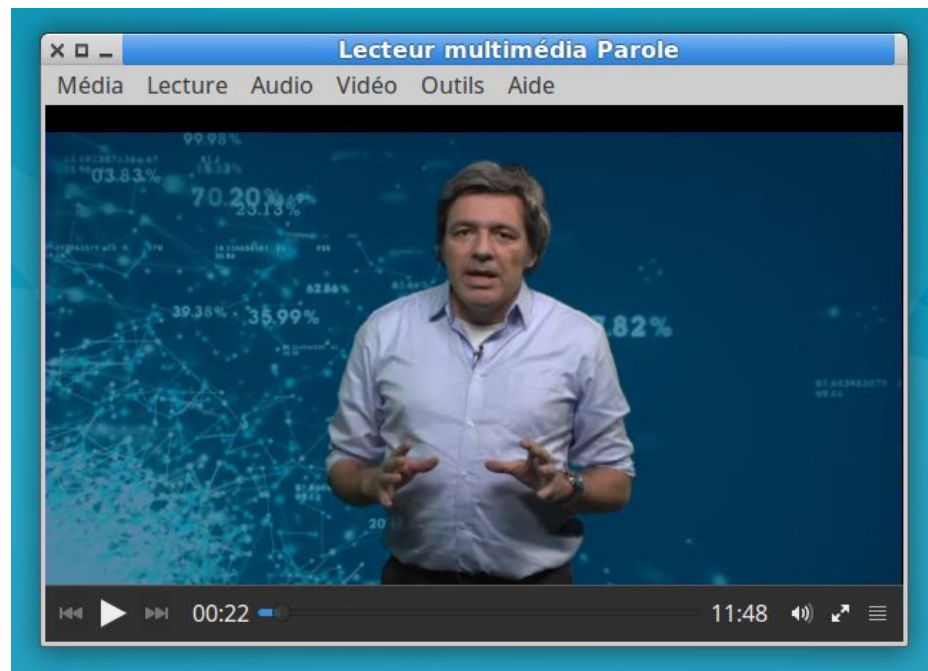
- Une méthodologie de découverte et de compréhension d'un jeu de données
- Une personne ou une petite équipe joue le rôle d'un jeu de données comme s'il s'agissait d'une personne interviewée par un journaliste/un enquêteur
- Ce dernier cherche à connaître le plus de choses d'elle en un minimum de temps
- Intérêts :
 - fait appel à des compétences sociales, partagées par le plus grand nombre et non-techniques
 - sa forme "bottom-up", ludique et interactive, qui renforce l'attention et la motivation des participants
 - une très bonne base pour réaliser une fiche descriptive d'un jeu de données

Exemple : **la base de données Open Food Facts**, à vos questions !

Évangélisation

Exemple : le COOC Data de la MAIF

Plus de 2 heures
d'épisodes vidéo qui
introduisent les
basiques de la
donnée dans
l'entreprise





Le développement des données transforme la société. La Poste s'inscrit dans cet

Pour cela, Le Groupe La Poste s'engage à :

1 / APPORTER TOUJOURS PLUS DE SERVICES À SES CLIENTS

/ Services numériques :

avec les données clients, simplifier les relations et opérer les échanges tant dans le monde physique que dans le monde du numérique.

/ Logistique :

avec les données de flux et les préférences de livraison des clients, adapter nos processus industriels aux contraintes et besoins de la vie moderne.

/ Activités bancaires :

2 / DONNER AUX PERSONNES LE CONTRÔLE DE LEURS DONNÉES

/ Information :

donner une information claire et complète pour permettre un accord averti.

/ Liberté :

accéder à ses données pour pouvoir les rectifier et exprimer librement ses choix (opt in / out).

/ Continuité :

assurer la pérennité et la continuité des services, garantir

3 / PROTÉGER LA DONNÉE ET RENDRE COMPTE DE SON TRAITEMENT

/ Exemplarité :

appliquer la réglementation et les contrats en vigueur, respecter la charte éthique du Groupe et contribuer à l'émergence de nouveaux labels ou certifications.

/ Expérience client :

par la qualité et l'usage pertinent des données, améliorer l'expérience client, personnaliser la relation et développer

LE #DATA LAB

au cœur de GrDF

INNOVER PAR LES DONNÉES

Au sein de la DSI de GrDF, ingénieurs en informatique, en mathématiques appliquées, en statistiques et analystes sémantiques décortiquent les **données de l'entreprise**. Ils utilisent les Data sciences et les outils Big Data pour élaborer des modèles prédictifs au service des métiers de GrDF.



Le sprint qualité : MAIF, Etalab, etc.



ID	Type de problème	Observateurs	Problème rencontré ?	Commentaire + exemple réel dans le jeu de données	A quel niveau ou fréquence le problème est-il observé ?	Impact du problème sur les usages	Effort de correction estimé
39	Manque de métadonnées : la fraîcheur des données n'est pas explicité : * le délai entre le réel et la mise en base de la donnée * le délai entre le réel et la publication de la donnée	Binôme 4	oui	on a effectivement une date de mise à jour cette date n'est pas necessairement identique à la date réel de modification de l'équipent	1. concerne l'ensemble du jeu	important	important
40	Manque de métadonnées : la langue des textes n'est pas spécifiée	Binôme 4	oui	suffisamment parlant	1. concerne l'ensemble du jeu	nul ou négligeable	très faible
41	Métadonnées imprécises : le format de date n'est pas spécifié	Binôme 4	oui	pas de mention du format	1. concerne l'ensemble du jeu	faible	faible
42	Métadonnées imprécises : unités non spécifiées	Binôme 4	non	unité précisée en M2			
43	Métadonnées imprécises : système de coordonnées non spécifié	Binôme 5	non	pas de coordonnée GPS			
44	Métadonnées imprécises : nom de colonnes ambigus	Binôme 5	non	nom des colonnes décrit dans méta données			
45	Métadonnées fausses	Binôme 5	oui	incomplètes : exemple manque Insnovoie	4. problème fréquent	faible	très faible
46	La taille maximale d'un champ dépasse celle qui est spécifiée dans la documentation	Binôme 5	oui	taille non indiquée dans les méta données : ne permet pas de contrôler	4. problème fréquent	faible	
	L'ordre des colonnes ne correspond pas à l'ordre donné dans la			on a été obligé de regrouper et trier les méta	4. problème		

Bienvenue sur dataNOVA

Le portail Open Data du Groupe La Poste

Vous souhaitez



Avoir un aperçu du patrimoine de données de La Poste et d'autres acteurs



Explorer les données disponibles, et faire des analyses simples



Télécharger les données ou utiliser les API pour intégrer les données ouvertes
(nombre limité de requêtes)



Partager des données postales et non postales



Proposer des applications ou des services utilisant les données

dataNOVA est votre plateforme !

Vous souhaitez aussi

Aller plus loin

CAHIER #1 : CULTURE DES DONNÉES

Vers une organisation *data driven* : mettre la culture des données au cœur de l'organisation

Ce cahier porte sur la place que la "culture des données" peut acquérir dans une organisation. Il montre non seulement la transversalité de la culture des données, mais encore la diversité des actions qui peuvent être menées pour favoriser son partage et sa circulation.

Ce cahier est le résultat d'un travail et d'échanges collectifs réalisés tout au long de l'année 2016 au sein de la Campagne Infolab de la FING. Le groupe a été animé par Sarah Labelle, avec l'étroite collaboration d'Armelle Gilliard et Charles Népote et la complicité de Simon Chignard. Le groupe était composé de : Marine Guillon et François Herlent pour la MAIF, Axelle Reynier pour le Groupe Up, Stéphane Gouffier pour l'ADEME, Isabelle Micheu pour La Poste. Le cahier a été rédigé sous la direction de Sarah Labelle.

Retrouver ce document en ligne :

<http://infolabs.io/cd16>

=> <https://culture-d.rn7.net>



Aller plus loin

CAHIER #2

1, 2, 3 data, expérimentez !

Ce cahier propose un panorama du monde de l'exploration et de l'expérimentation avec les données. Il montre comment l'expérimentation peut être à la portée de tous, avec des moyens adaptés à toutes les organisations.

Ce cahier est le résultat d'un travail et d'échanges collectifs réalisés tout au long de l'année 2016 au sein de la Campagne Infolab de la FING. Le groupe a été animé par Charles Népote, avec l'étroite collaboration d'Armelle Gilliard et la complicité de Sarah Labelle et Simon Chignard. Le groupe était composé d'Emmanuelle Katz et Axelle Reynier (Groupe UP), Isabelle Micheu et Zouhroua Hadji (Groupe La Poste), Érik Arnaud, Véronique Morault et Sophie Vitrat (MAIF). Le cahier a été rédigé sous la direction de Charles Népote.

Retrouver ce document en ligne :

<http://infolabs.io/xp16>

=> <https://culture-d.rn7.net>



Ressources

- [Liste thématique](#) d'une douzaine de ressources sur la culture de la donnée
 - Des ressources générales
 - Des ressources thématisées par objectif
- Des professionnels
 - La coopérative [Dataactivist](#)
 - Le [collectif Data Publica](#)
- Des lieux d'échange
 - Forum [Team Open Data](#)
 - [data.gouv.fr](#) est très social : commentaires, valorisations des usages





Data Literacy Conference

2016, 2017, 2018

- 2 jours
- 170+ participants
- 15 intervenants internationaux de haut niveau
- 8 ateliers, 2 masterclass
- Une première mondiale



Les données de la culture

- Vendredi prochain 9h
- 45 minutes de préparation
- Interview de la donnée : identifiez un jeu de données relatif à la culture. Astuce : portails de données vus aujourd'hui.
 - Parcourez-le, regardez/explorez : les données, leur contexte de production, leur producteur (qui, pourquoi il fait ça ?), des éventuels problèmes qualité, **les usages**, les usagers, ...
- Alternative : “les données et moi” : une de vos expériences d’usage des données : qu’avez-vous appris ?
- Alternative : trouvez un exemple de bouleversement du monde de la culture par les données

Sources

Cette présentation reprend pour partie les travaux effectués pendant la campagne Infolabs de la Fing (2015-2018), menés par Armelle Gilliard, Sarah Labelle, Simon Chignard et Charles Nepote.