

LES DATA EN FORME

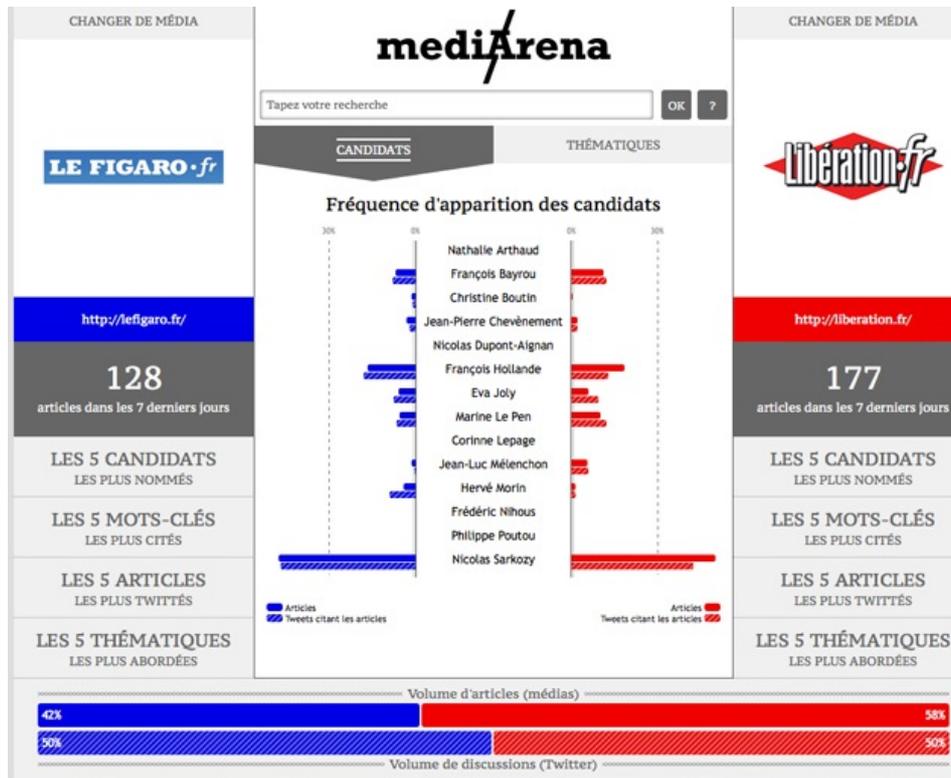
LE 12 DÉCEMBRE 2011 PAULE D'ATHA

Des candidats volants, des data pas encore sur les rails, des rumeurs qui s'analysent, un Monopoly, Tino Rossi et Mozart. C'est le nouvel épisode des data en forme, veille transversale et décalée sur le journalisme de données.

Revenons dans le vif du sujet avec un concours qui a bien buzzé lors de son lancement, il y a moins d'un mois : le **#Googleviz**, clos depuis maintenant une semaine. On avait évoqué l'une des premières web-apps sorties, **Retwhit2012**, dans le dernier épisode des data en forme et le niveau était déjà sérieux. Ça se confirme.

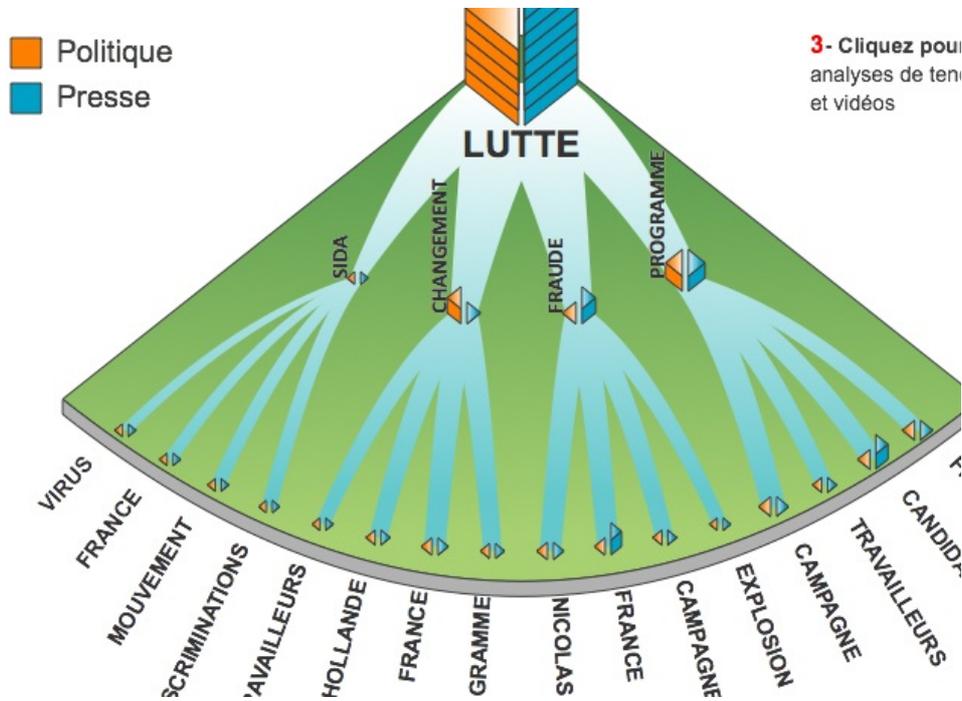
La semaine dernière, alors que le jury délibérait, certains participants n'ont pas résisté à l'envie de partager leur création. Et vu la qualité des projets, le choix du vainqueur risque de se finir à coups de pelles : ça ne sera pas simple.

Mediarena, conçue et développée par **Nils Grünwald**, **Stéphane Raux**, **Alexis Jacomy** et **Ronan Quidu** est une bonne "datagifle". Tout est là au premier coup d'œil : l'angle est clairement identifiable – comment les principaux médias en ligne traitent la présidentielle – et la prise en main plus qu'intuitive. En quelques clics, on joue avec les données et on fait défiler sur ce ring les noms des grands titres. Le cadeau bonux, c'est que derrière cette simplicité gage de lisibilité, Mediarena nous donne accès à bon nombre de données qui donnent du relief et de la profondeur à l'angle choisi.



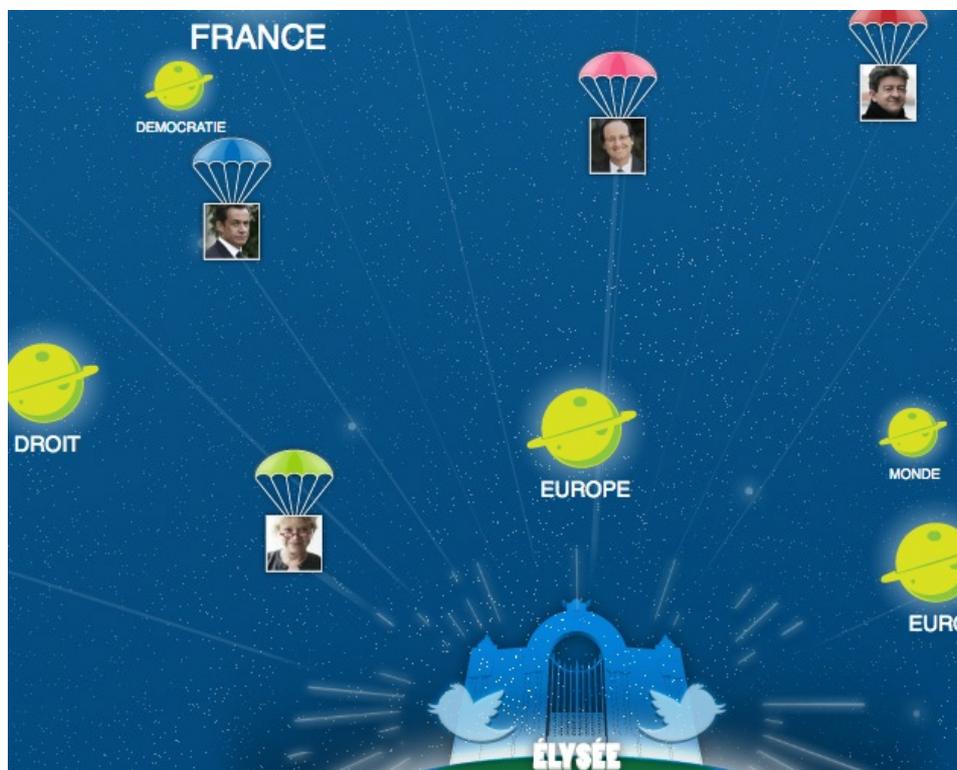
Autre concurrent en lice, les concepteurs de "partie de campagne" (joli hommage à Raymond Depardon et son film "1974, une partie de campagne"). Si le design de leur web-app est un peu moins léché, leur mise en scène des données est vraiment intéressante et innovante. On entre par un nuage de tags qui donne un aperçu des principaux thèmes abordés par les médias et par les politiques (avec un petit graphique les comparant au survol de chaque mot).

Une fois le choix défini, le second écran nous immerge dans le "delta" (version HTML5 + canvas), les pieds dans la thématique et ses sujets affluents. En cliquant sur chaque terme on accède aux données : analyse de la tendance, affichage des sources (politiques et médiatiques) et même un listing des vidéos Youtube référentes. De quoi fouiller le sujet.



La team haploid a choisi elle un angle légèrement plus ludique, du moins dans la visualisation des données, avec son **“Qui sera parachuté à l’Élysée ?”**. Les candidats sont physiquement parachutés vers l’Élysée, point de gravité qui les attire tous, inexorablement. La base data pour chaque candidat étant son compte Twitter officiel associé à celui de sa formation politique, une corrélation entre le nombre de followers et le nombre de retweets permet de déterminer en live lequel est le plus proche du Saint-Graal.

Autour d’eux flottent les planètes des thématiques quotidiennes identifiées sur Twitter. L’interaction est, là aussi, assez intuitive avec, au survol, un affichage des liens entre candidats-volants et planètes-thématiques et, au clic sur chaque élément, l’affichage du fil Twitter correspondant.



La bonne data derrière ces premiers exemples, c’est que ce concours **#Googleviz** a déjà permis de remarquer de sacrées équipes. De quoi **émoustiller** le timide milieu de la dataviz hexagonal.

Open Data sans data

Avant de filer vers les datas non-hexagonales, sortons le saint-tag **#opendata** pour regarder du côté de ceux qui essaient de nous faire préférer le train. **data.sncf.com**, dix lettres et

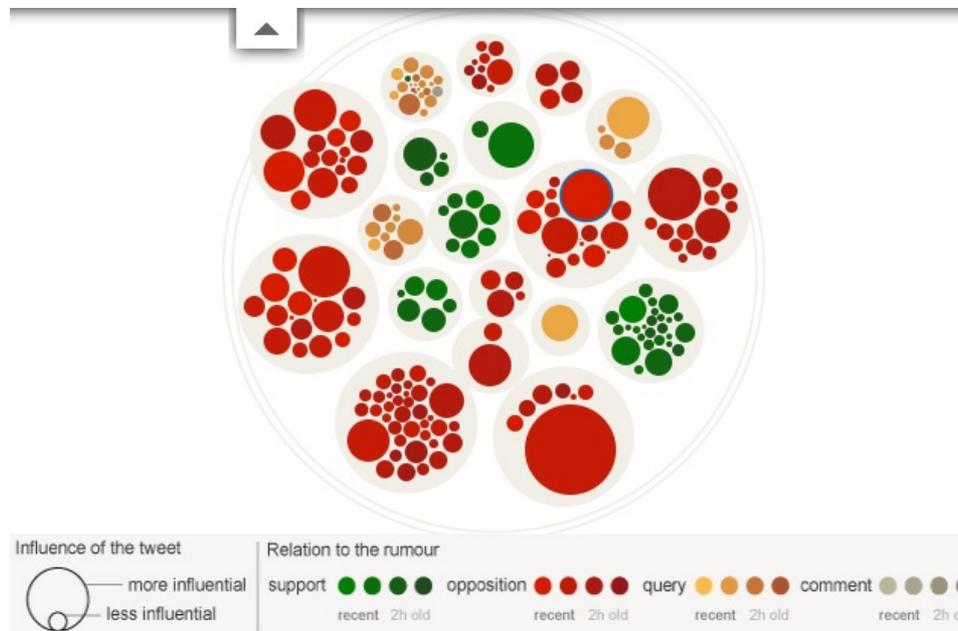
deux points qui ont de quoi faire saliver quelques csv-dépendants. Seulement voilà, une fois la page d'accueil chargée, on reste pantois : pas un seul jeu de données à se mettre sous l'tableur. Le site est un appel à débat comme le souligne cette bondissante baseline : "Open data, open débat". Pour le modèle "gagnant-gagnant" prôné dans le court texte de présentation, on repassera.



Chers transporteurs d'humanités, sachez que l'Open Data est un débat depuis quelques temps déjà et que le meilleur moyen de le faire avancer eût été de nous lâcher vos données. Pour l'innovation, côté des détenteurs de données, on attendra.

Traquer les rumeurs

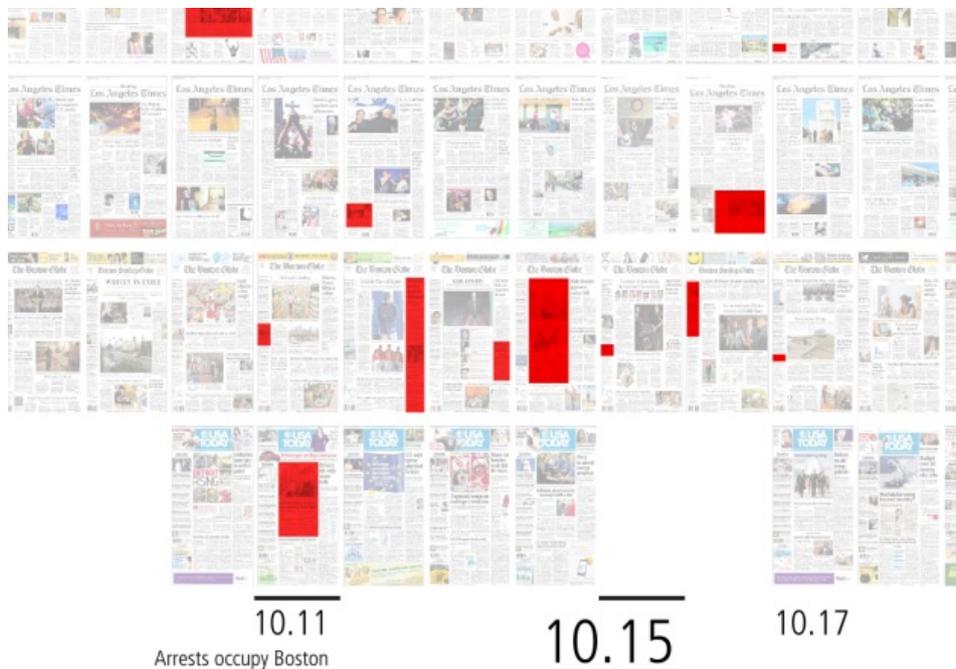
Prenons l'avion pour aller faire un tour outre-manche. Nos amis du *Guardian*, qu'on ne présente plus en matière de journalisme de données, ont encore fait joujou avec quelques chiffres pour notre plus grand plaisir. Le concept de longue traîne ne leur étant pas non plus inconnu, ils sortent **une très belle visualisation interactive** sur un événement qui a près de quatre mois : les émeutes londoniennes.



Alastair Dant et ses collègues ont décidé d'analyser l'évolution des rumeurs sur Twitter pendant ces événements. "*Les émeutiers ont libérés les animaux du zoo de Londres*", "*les émeutiers se font leurs propres sandwiches dans les Mc Do*"... À vous de choisir parmi sept rumeurs – cinq fausses, une infondée et une avérée – pour visualiser leur évolution. Le replay est intelligemment construit, avec notamment l'identification visuelle des tweets favorables, opposés, interrogatifs ou simplement commentant les faits et la mise en avant de moments clés où la diffusion de la rumeur se modifie. Et, comme au *Guardian* ils aiment partager, ils nous livrent un **making of** de cette datavisualisation qui permet notamment de comprendre l'importance d'un travail d'équipe intégrant : journalistes, développeurs, designers et universitaires.

Unes VS Twitter

Direction le sud de l'Europe et l'Espagne : autre mouvement, OccupyWallStreet, et **autre visualisation** signée Numeroteca. Si le rendu manque cruellement de possibilités d'interaction pour mieux saisir les données, le principe mérite que l'on s'y arrête. Le but est de matérialiser un comparatif entre le traitement du sujet à la une des grands journaux américains et le nombre de tweet/jour sur ce même sujet. Le rendu visuel permet non seulement de comparer les deux présences médiatiques sur un graphique mais on visualise également la place physiquement accordée sur la page de une. **Le même type de travail** a été mené pour comparer le traitement à la une des principaux journaux espagnols et sur Twitter du printemps arabe.



Des recherches à suivre, notamment lorsque l'on sait qu'un Hackathon sur le sujet #OccupyData s'est tenu le 9 décembre dernier et que les pistes qui s'y sont dessinées sont accessibles directement dans un **GoogleDoc** en libre accès.

Au-delà du chaos

Passons outre-atlantique pour aller crowdsourcer le futur de l'informatique. C'est ce que propose le *New York Times* à travers **une efficace timeline verticale** qui répartit le sujet sur quatre grands thèmes : calcul, intelligence artificielle, transports et mode de vie et communication. Rien ne sortant de rien, ce sont les bases qui apparaissent en premier avec tout l'historique de l'informatique au sens large depuis **les bâtons de Napier** en 1617 jusqu'à 2011 année historique où Watson, super-ordinateur conçu par IBM, a battu les deux champions du jeu américain "**Jeopardy!**".

2012

THE FUTURE

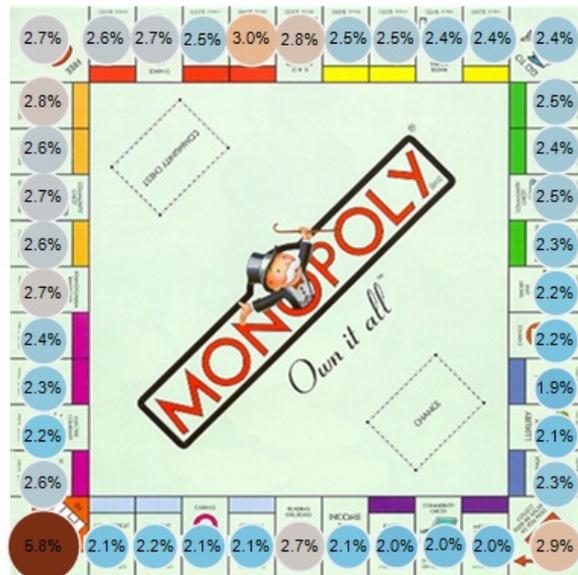


Puis c'est le grand saut. De quoi demain sera fait ? À partir de 2012, c'est un grand tableau noir où s'affichent les prédictions proposées par les audiences. L'espace pour laisser sa propre prospective est aujourd'hui clos mais vous pouvez toujours interagir de deux façons différentes. Soit en déplaçant les événements affichés sur la timeline post-2011, soit en votant pour les propositions qui vous semblent les plus intéressantes, réalistes #oupas. Les mieux notées seront progressivement insérées dans la partie prédictive de la timeline.

DataTriche et WTF

L'hiver approche et avec lui le cliché des longues soirées au coin feu qu'il va bien falloir occuper. Si certain(e)s d'entre vous aiment le jeu, possible que vous ressortiez dans les semaines qui viennent ce bon vieux **Monopoly** qui prend la poussière depuis un an. Alors laissez-moi vous donner un tuyau. Il y a un développeur californien nommé **Ben Jones** qui s'est amusé à modéliser les statistiques issues de 60.000 parties aléatoires. Son "**Dominate Family Game Night**" présente un tableau de bord des différentes stratégies de jeu en fonction de grandes tendances. Chut, je ne vous ai rien dit.

Which Board Spaces are Landed on Most Frequently?



Hover for Instructions

Property Scenario
With a Hotel

Railroads You Own
 1
 2
 3
 4

Utilities You Own
 1
 2

Created by: Ben Jones
@DataRemixed

Avant de finir, on ne va pas, nous aussi, lancer un concours mais juste un appel : saisissez-vous des données WTF ! L'Internet mondial en est rempli et elles n'attendent que de folles petites équipes pour être visualisées. Tiens par exemple : **les meilleures ventes de 45T/Singles de tous les temps** (merci @Pirhoo) ou comment refaire un TOP50 2.0 avec en tête Tino Rossi et J.J Lionel.

Ou encore, pour rester dans le domaine des mélodies inoubliables, toute la correspondance de Wolfgang Amadeus Mozart et sa famille. Près de 1 400 lettres triées en fonction des dates, lieux, expéditeurs, destinataires, œuvres mentionnées. De quoi scraper...



Addio, porte-toi bien. Mon seul plaisir consiste actuellement à faire des pas anglais, des cabrioles et le grand écart. L'Italie est un pays de dormeurs, on y somnole toujours !

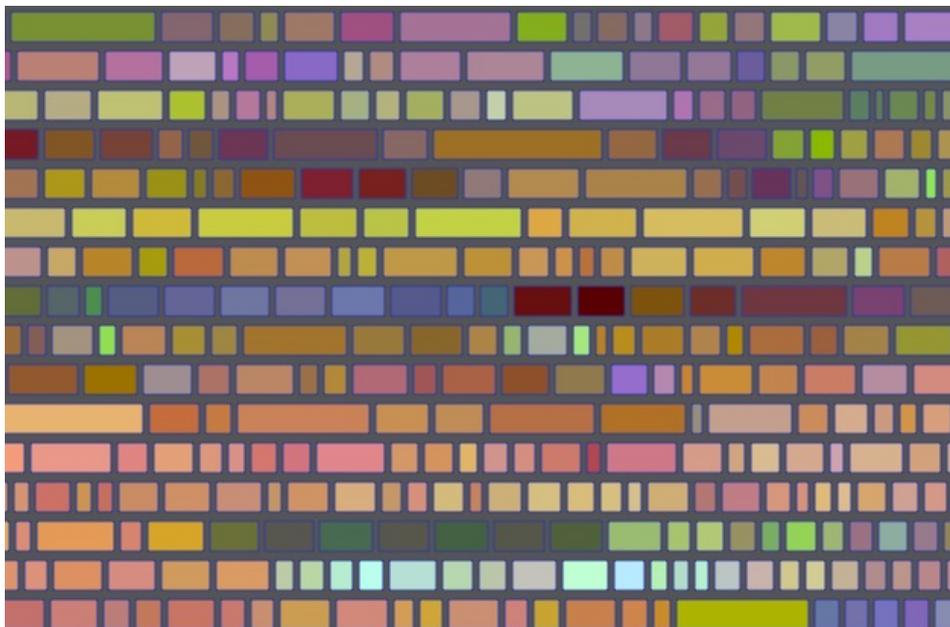
Addio, adieu !

Le 4 août 1770.

Wolfgang Mozart

BRoeimn hhaapdsy

Reprenant la chronologie inversée du précédent "**Les data en forme**", terminons cette semaine en musique avec une dataviz pour les oreilles. C'est du Queen et c'est la mythique Bohemian Rhapsody qui est joliment déstructurée. Dans **Bohemian Rhapsicord**, créée lors du **Music Hack Day** de Boston, Jennie and Paul Lamere ont concrètement morcelé le morceau en une multitude de séquences pour nous laisser le rejouer à notre manière. Soit vous appliquez un des filtres qu'ils proposent (durée, volume, inversion, similarité), soit vous définissez une touche de votre clavier pour chaque segment et à vous de reconstruire le puzzle musical. Seul bémol : la web-app ne fonctionne que sous Chrome.



Allez, que l'#opentata et le #dadajournalisme vous inondent et à la semaine prochaine.

—

Retrouvez les précédents épisodes des **Data en forme** !

PLACEBO

le 12 décembre 2011 - 19:52 • SIGNALER UN ABUS - PERMALINK



Nos* amis du Guardian.

Sympa les applis sur la présidentielle.

C'est fou comme ça tourne autour de twitter n'empêche, beaucoup de mouvement, de statistiques exploitables, de tweets reproductibles... faut en profiter.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

BEN JONES

le 13 décembre 2011 - 7:06 • SIGNALER UN ABUS - PERMALINK



Merci beaucoup pour remarquer mon projet de monopole. S'il vous plaît ne dites pas mes fils mon secret! ;)

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

WILNOCK

le 13 décembre 2011 - 8:57 • SIGNALER UN ABUS - PERMALINK



L'application du Guardian est tres impressionnante. Un peu lourde au chargement peut-etre, ou est-ce due a la vetuste de mon equipement...

Par contre, je ne suis pas sur d'avoir réussi a demeler le vrai du faux dans toutes ces rumeurs. Pour certaines, c'est facile, car il est possible de suivre les dementies, mais pour d'autres, ce n'est pas si evident que ca. Un travail de journaliste sur le terrain s'avere plus que necessaire pour apporter une conclusion parfois.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

5 pings

Visualisation by kimio - Pearltrees le 13 décembre 2011 - 9:07

[...] Les data en forme » OWNI, News, Augmented La team haploid a choisi elle un

angle légèrement plus ludique, du moins dans la visualisation des données, avec son "Qui sera parachuté à l'Élysée ?". Les candidats sont physiquement parachutés vers l'Élysée, point de gravité qui les attire tous, inexorablement. La base data pour chaque candidat étant son compte Twitter officiel associé à celui de sa formation politique, une corrélation entre le nombre de followers et le nombre de retweets permet de déterminer en live lequel est le plus proche du Saint-Graal. Autour d'eux flottent les planètes des thématiques quotidiennes identifiées sur Twitter. L'interaction est, là aussi, assez intuitive avec, au survol, un affichage des liens entre candidats-volants et planètes-thématiques et, au clic sur chaque élément, l'affichage du fil Twitter correspondant. [...]

data en forme by vcroissa - Pearltrees le 13 décembre 2011 - 12:50

*[...] Les data en forme » OWNI, News, Augmented On avait évoqué l'une des premières web-apps sortie, Retwhit2012, dans le dernier épisode des data en forme et le niveau était déjà sérieux. Ça se confirme. Visualisation Réutilisations – Applications The art of data visualization Mind Maps/Thinking Maps/Graphic Organizers Data compiler [+]
data visualization japan project visualizations wonder viegas Data visualization 2
visualizing Data Visualization software data journalism backstage sense a clasificar
mapping nathan tools Data visualisation ideas information kiesler Data Visualization
majormode script monitor information info design times Visual Thinking The future of
User Interfaces bumttop holographic Data Visualization scraping online Outils
opencorporates understand wordle Gapminder gapcast Quelques visualisations
visualisons visualisation Public Data Sources forum religion Outils, ressources think
webtools screenr InfoGraphics (Visual Maps) social vizualize Data Visualization
infographics interesting ultimate Information Visualisation ideas DATOS ocean BEST
INFOGRAPHICS infographie libya visualization manovich INFOGRAFICAS occupy Info
design enterprises blogs i webs sobre mapas aesthetics mapping Data & Visualisation
directgov hartig dataviz mapping Data Art information DataViz dataveyes generation
Data Visualization simile resource viegas Data Visualization visual quantitative ~temp~
create κατοικίες Presentations mining General usage triptico worditout Datavision david
create Les 3 blogs indispensables & le Framapack periodic visualization evaluation
information visualization aesthetics 2.0 services google Infographics media Visualisation
de données vidéo home • contact • blog • fb • twitter to experience pearltrees activate
javascript. [...]*

Vise-moi cette Googleviz » datablog.owni.fr le 14 décembre 2011 - 16:19

[...] concours Google dataviz (organisé par Google et YouTube Elections 2012) a récompensé – comme nous le pressentions – l'équipe de Mediarena dans une ambiance de bonne humeur à l'intérieur des locaux [...]

Actus Généralistes 2011 S50 | La Mare du Gof le 18 décembre 2011 - 17:48

[...] en forme, veille transversale et décalée sur le journalisme de données (...).» Source : owni.fr/2011/12/12/data-googleviz-2012-monopoly-riots-london-queen/ Billets en relation : 12/12/2011. Open Data : nouvelle stratégie pour Bruxelles : [...]

Régions.news #15 – Edition du vendredi 23 décembre 2011 | Régions.news le 25 décembre 2011 - 18:24

[...] Les data en forme publié le 13/12/2011 par OWNI [...]