

# LES DATA EN FORME

LE 24 OCTOBRE 2011 PAULE D'ATHA

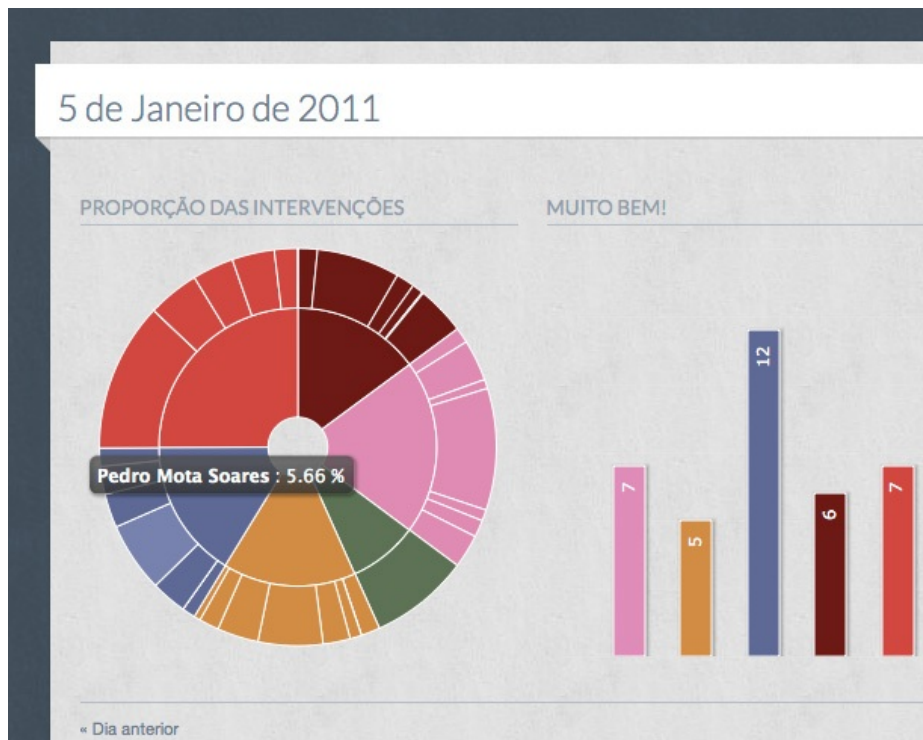
Tour d'horizon de la galaxie data. L'équipe des journalistes de données d'OWNI livre l'essence de sa veille hebdomadaire. Au programme cette semaine : data parlementaires, billets verts et hautes altitudes. Entre autres.

Cette semaine, attaquons par du sérieux : parlons politique. Les 19 et 20 octobre se tenait à Varsovie un événement essentiel dans le monde de l'Open Data : l'**Open Government Data Camp**. Parmi les 400 participants, il y avait notamment les deux créateurs de **Manufactura Independente**, un studio de design basé à Porto. Ils venaient y présenter le dernier projet auquel ils ont participé : **Demo.cratice** (en portugais), un superbe outil permettant d'explorer le parlement portugais à travers les data.

Intéressé par le contenu des sessions parlementaires ? Demo.cratice a ce qu'il vous faut : un calendrier des sessions depuis 2009 permettant de naviguer parmi les transcriptions des débats. Au survol d'une date sur le calendrier s'affiche le mot le plus souvent cité dans les prises de paroles du jour, pratique pour parcourir les thématiques abordées. Lorsque vous vous penchez sur la retranscription d'une session parlementaire, en plus des textes des interventions, Demo.cratice propose une visualisation "statistique" des échanges. Vous pourrez ainsi facilement voir quel groupe parlementaire a le plus pris la parole, via quels députés au sein de ce groupe ou quels sont les grands thèmes abordés.

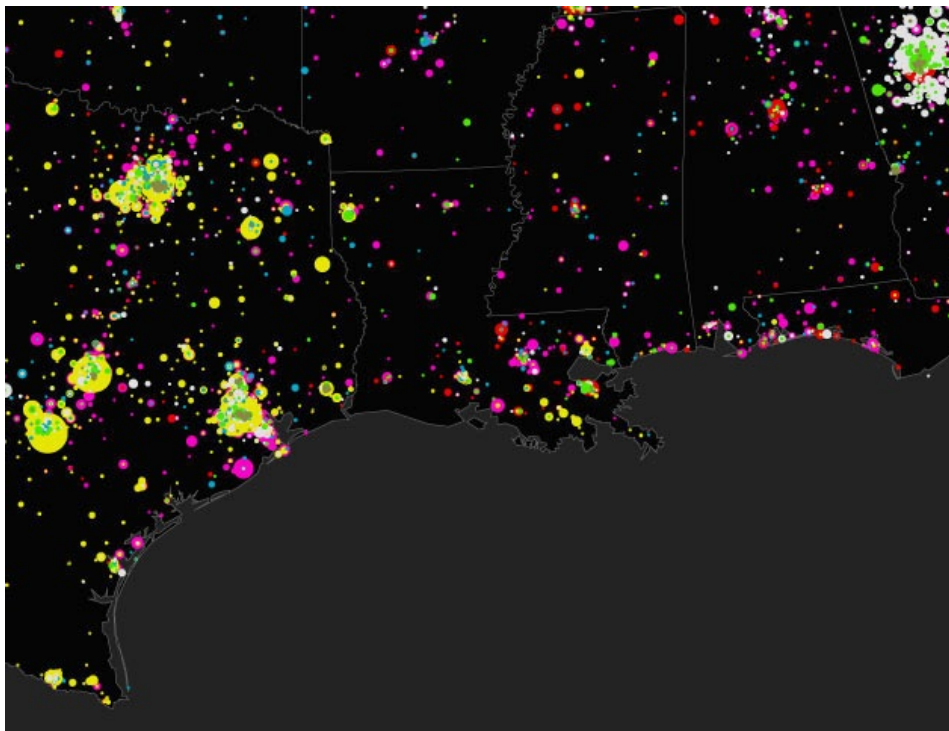
Pour ne rien gâcher, la mise en forme de ces données est plus que léchée et **les sources de Democra.tica** sont regroupées sous la forme d'un logiciel libre et donc disponibles sous licence **AGPL**. Quoi de plus naturel lorsque l'on sait que ce projet est né, il y a 9 mois, au cours d'une session **Hackday à Porto** ?

Depuis le temps qu'on vous le dit : hacker c'est bien, c'est bon.



Continuons avec la politique côté finances. On a pu voir en France que l'exercice des primaires présentait **un solde plus que positif**. Outre-Atlantique **Matt Stiles**, journaliste pour NPR et auteur du blog **The Daily Viz**, s'est amusé avec les data de la Commission nationale électorale. Il s'est posé une question simple : d'où viennent – géographiquement – les 90 millions de dollars engrangés par les candidats républicains depuis le début de l'année ?

Sa réponse tient en **10 cartes** : la première, globale, affiche les données de tous les candidats, les neuf autres présentent la provenance des fonds candidat par candidat. On regrette juste qu'il n'ait pas poussé jusqu'aux quelques lignes de HTML5 qui auraient permis de présenter l'ensemble avec plus de simplicité et d'ergonomie qu'un long scroll.



## Dataviz battles

Ces deux projets pourraient donner quelques pistes de réflexions pour le concours de dataviz qui vient d'être lancé par Google : "**Les élections 2012 autrement**". Le principe est simple :

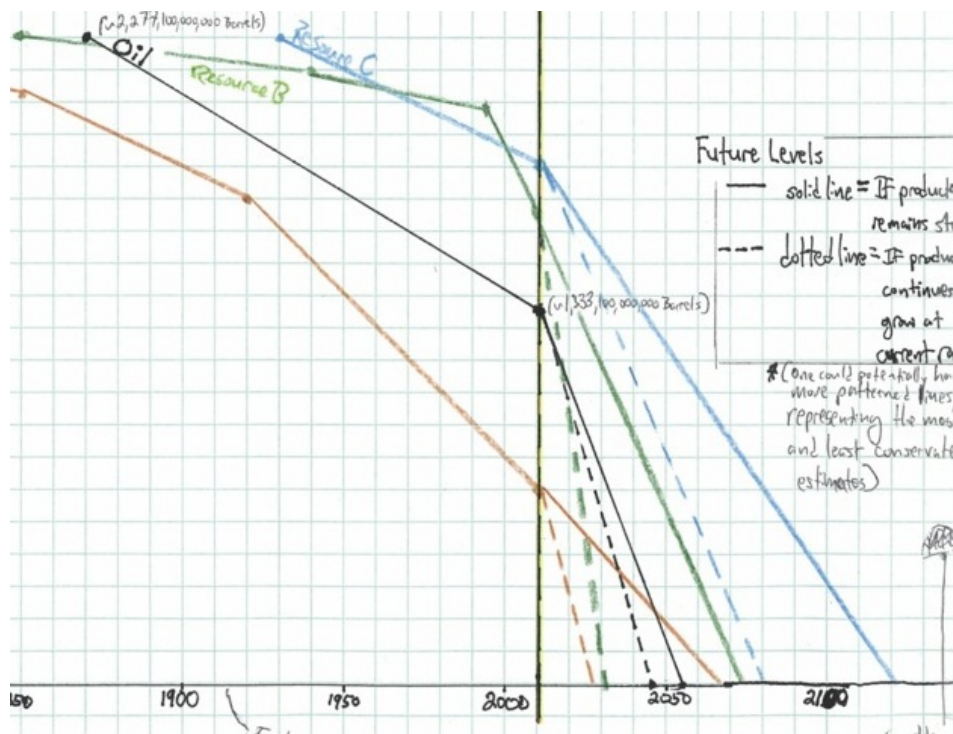


***Proposer une application web interactive qui utilise des données de Google ou de Twitter pour proposer un nouveau regard sur l'élection présidentielle française de 2012.***



Seule contrainte, donc, utiliser au moins un jeu de données provenant de Google ou Twitter, libre à chacun ensuite de laisser s'exprimer sa créativité. API fortement conseillées. Le tout (les sources de la webapp) est à envoyer sur [dataviz2012@gmail.com](mailto:dataviz2012@gmail.com) avant le 7 décembre 2011.

Les concours de dataviz ont d'ailleurs tendance à se multiplier ces derniers temps. On notera celui organisé par l'un des papes de la dataviz, David Mc Candless, tout simplement nommé : **The Information is beautiful Awards**. Le premier concours portait sur l'évolution des stocks d'énergies non-renouvelables dans les années à venir et les vainqueurs sont sur le point d'être annoncés. En attendant on peut aller se balader sur les projets des finalistes des deux catégories : **interactive challenge**, visant à produire une web-app restituant les data, et le **napkin challenge**, qui rassemble les projets à l'état d'intentions graphiques crayonnées sur des nappes en papier ou tout autre support un tant soit peu lisible, du jus de cerveau en open source en somme.



## OWS DVZ

La dataviz, c'est à chacun de s'en servir et, bien utilisée, c'est un excellent outil de communication. Le monde de l'entreprise l'a compris en assommant des armées de commerciaux à coup de powerpoints graphomaniques durant de nombreux séminaires "Chiffre d'Affaires", "Chiffres de Vente" ou "Parts De Marché" – car chacun sait qu'une image passe toujours bien mieux. Toutefois avec deux sous d'inventivité, un soupçon de **DIY** et un message plus politique à faire passer, on trouve tout de suite des idées bien plus intéressantes.

Au cœur mouvement **Occupy Wall Street**, la datavisualisation est apparue en toute logique car **quoi de plus efficace** (photo, à voir) pour représenter le rapport de force entre les 1% les plus riches et les 99% restants ?

Une autre belle idée, celle d'**Occupy George**, est d'avoir collé des visualisations – grâce à des tampons-encreurs à la papa dûment bricolés – sur le support même qui circule le plus dans nos (99%) mains de consommateurs assoiffés pour finir dans les poches des 1% : le "George-Washington", emblématique billet de un dollar.

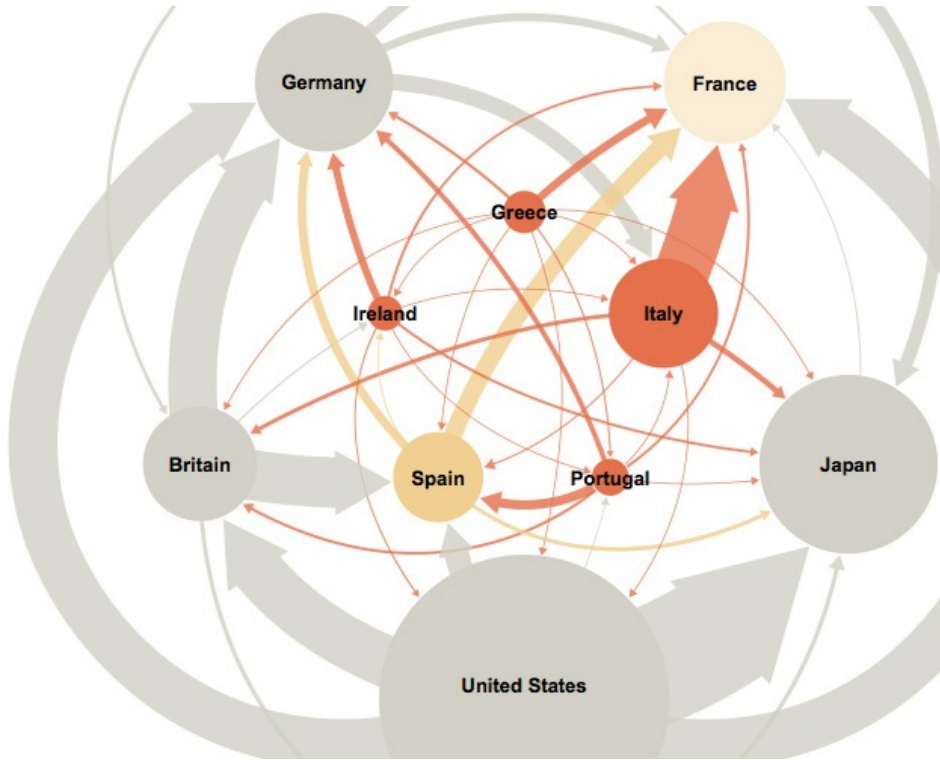


Tiens d'ailleurs, pendant que l'on parle politique, économie et crise dans notre datarticle hebdomadaire, le *New York Times* nous gratifie d'un très beau travail sur la crise de l'Euro.



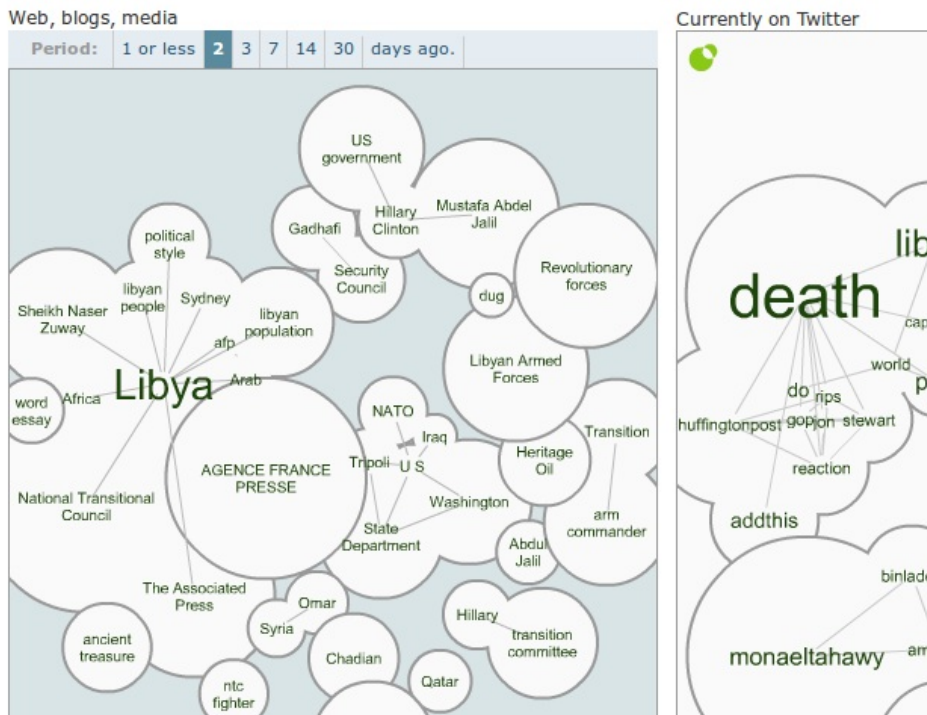
Leur visualisation interactive met en évidence l'interdépendance des différents acteurs (actifs ou passifs) : **It's all connected.**

En plus des chiffres propres à chaque pays (qui détient quoi ?), plusieurs onglets permettent d'explorer les différents aspects : des problèmes actuels aux risques de contagion en passant par les scénarios possibles.



Avant de vous remettre au boulot, allez jeter un œil sur l'outil développé par icosystem : **Infomous**. Pointé par l'excellent Simon Rogers, sur **le datablog du Guardian**, ce système propose un principe de navigation au sein de l'information assez novateur. Il permet de visualiser les sujets les plus importants, les plus partagés en temps réel en les organisant par mots-clés.

Des "galaxies" de "faits" se dessinent ainsi en donnant accès aux contenus en profondeur (les articles liés). Infomous est plutôt bien pensé car il intègre également quelques outils pour personnaliser cette visualisation : paramètres de la visualisation (zoom, nombre de sujets...), exclusion de mots, de types de mots, screenshot, embed ou encore choix des sources (malheureusement parmi une liste plutôt courte de 13 médias internationaux anglophones).

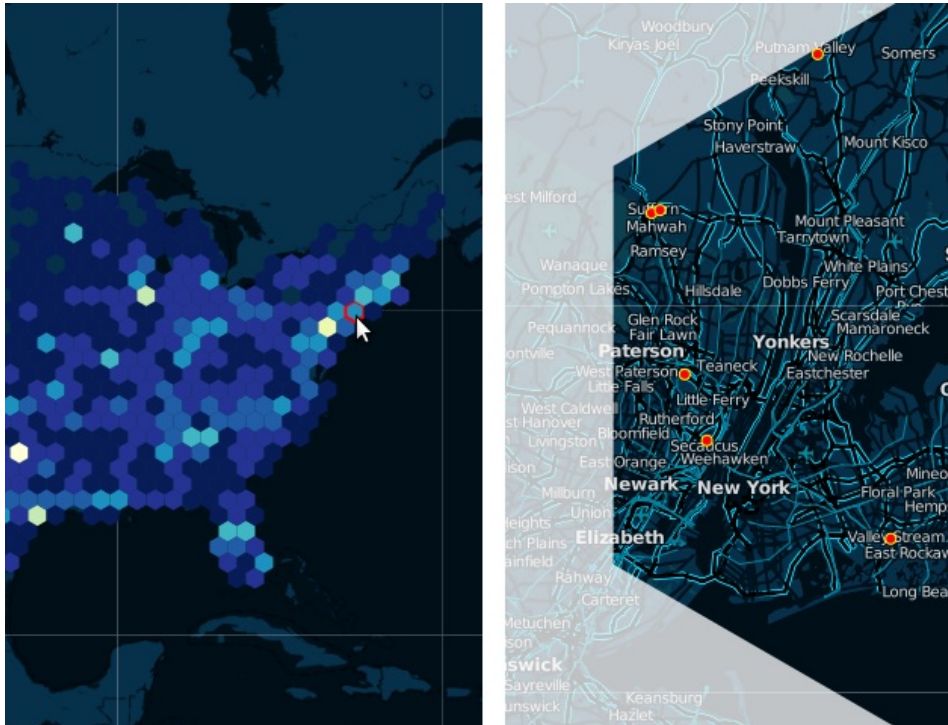


## Inspirez, visualisez

Histoire de garder quelques belles images en tête, finissons sur de l'expérimentation, possible source d'inspiration.

Une première piste en matière de géolocalisation proposée par **Zachary Forest Johnson** (repérée sur **Information Aesthetics**). Pour faciliter la visualisation de nombreux points sur une vaste zone géographique, il a appliqué un principe délicatement nommé : **"hexagonal binning"** [PDF].

Le but est de diviser les zones contenant des data en hexagones différenciés graphiquement, par exemple avec des nuances de couleur, en fonction du nombre de points contenus dans la zone en question. Sur **l'exemple de la répartition des magasins Wall-Mart** sur le sol US, plus une zone contient de magasins, plus l'hexagone correspondant tire vers le clair, moins il y en a plus sa couleur est sombre. Le tout permet de simplifier la visualisation de données nombreuses sur un espace restreint.



Petit bonheur : HexBin, l'outil développé par Zachary F. Johnson est **disponible sur github** et il peut s'utiliser sous forme de fonction d3.js et s'intégrer à **PolyMaps**.

Restons dans l'inspiration, allons prendre l'air. L'hiver approche, c'est l'occasion d'aller côtoyer les sommets enneigés mais même encordés, n'oublions pas les data. La raréfaction de l'air au niveau des 6,962 m d'altitude du **Mont Aconcagua** a sans doute inspiré le photographe **Michael Najjar**.

Il s'est servi de ses clichés comme base pour produire d'étonnantes visualisations : **les chemins de crête** dessinent les cours du Dow Jones, Nikkei, Nasdaq ou autres Lemhan Brothers.



Enfin, expérimentation totale. Si votre carte graphique supporte le WebGL, allez faire un tour sur le module **wire.2x.io** et amusez-vous ! Dessinez des courbes de data sur la grille et baladez-vous en 3D à l'intérieur à l'aide du clavier. Ça laisse rêveur...

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

#### SERGE

le 24 octobre 2011 - 16:03 &bullet; SIGNALER UN ABUS - PERMALINK



L'URL [wire.2x.io](http://wire.2x.io) mériterait un <http://>

VOUS AIMEZ



VOUS N'AIMEZ PAS



LUI RÉPONDRE

#### ADMIN

le 24 octobre 2011 - 22:16 &bullet; SIGNALER UN ABUS - PERMALINK



Merci Serge, réparé.

VOUS AIMEZ



VOUS N'AIMEZ PAS



LUI RÉPONDRE

#### CANYON

le 26 octobre 2011 - 12:55 &bullet; SIGNALER UN ABUS - PERMALINK



Sympa comme tout.

VOUS AIMEZ



VOUS N'AIMEZ PAS



LUI RÉPONDRE