

# LA FRANCE ENTR'OUVERTE

LE 10 DÉCEMBRE 2011 NICOLAS PATTE

L'État a lancé son site [data.gouv.fr](http://data.gouv.fr). La France, enthousiaste, ouvre donc ses données publiques comme les États-Unis. Transparence à vérifier. Car pour l'instant, le citoyen est comme absent du grand projet.



Officiellement, le 5 décembre, la République a donc ouvert ses données publiques lors d'une sauterie organisée sous les lambris de la salle de la Chapelle, en l'Hôtel de Cassini à Paris. Au terme de 817 jours de développement, de 691 624 lignes de code déployées par des partenaires technologiques français<sup>1</sup> le site [data.gouv.fr](http://data.gouv.fr) a vu le jour en grandes pompes, mais en version *beta*.

## Viennoiseries

Ce jour-là, semillant comme à son habitude, le jeune (35 ans) directeur du projet gouvernemental Séverin Naudet fait la claque avec **Prezi**, le logiciel de présentation en ligne qui relègue PowerPoint au Paléolithique moyen. Rompu à l'exercice, il parvient miraculeusement à tenir en haleine le parterre de journalistes en déroulant les écrans de tableur. Aimable, cet ancien producteur des *Nada Surf* et pratiquant le Kung-fu sait flatter ses collaborateurs, son audience et ses soutiens anonymes, qui participèrent assidûment à "quatre *workshops* avec l'écosystème Open Data durant les neuf derniers mois". Et ses sponsors aussi - nombreux - dont on entend les noms de Google et de... Microsoft sans en voir la présence sur les documents distribués avec les viennoiseries.

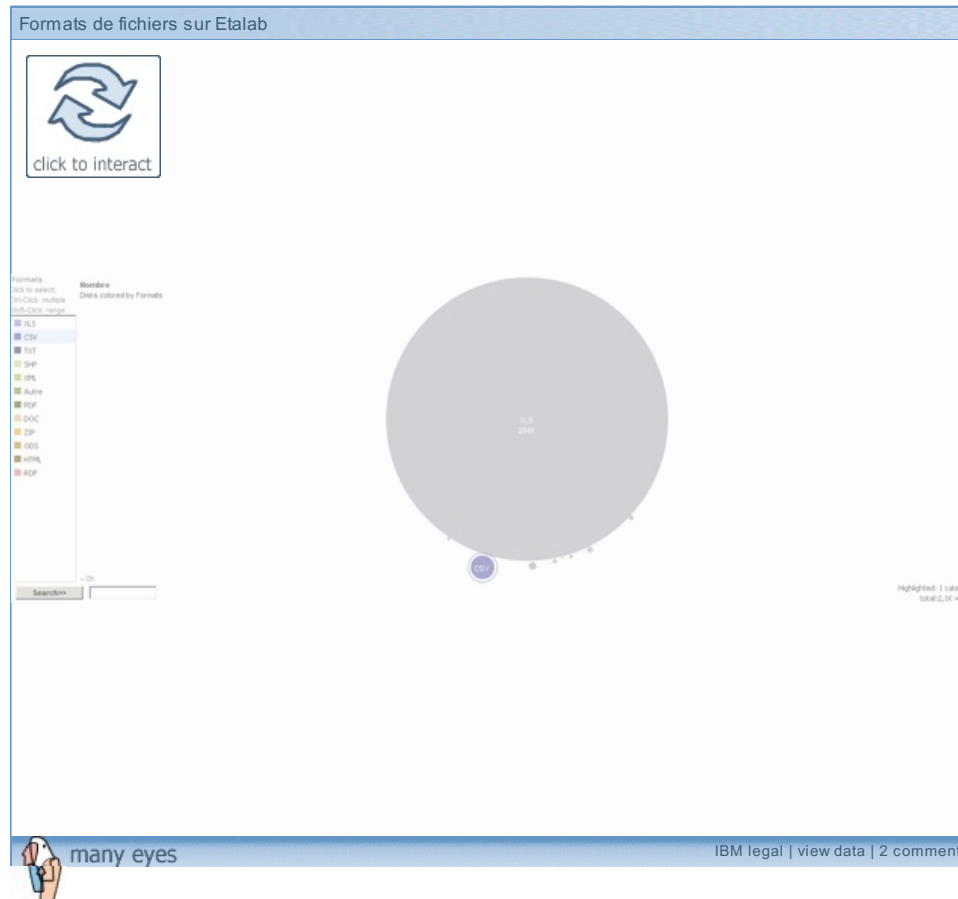
Chargé de l'innovation dans l'équipe, Romain Lacombe est au cœur du projet. Co-auteur du rapport ***Pour une politique ambitieuse des données publiques*** l'été dernier, lui non plus n'est pas avare de bonnes ficelles pour faire passer cette pilule sucrée nommée "**Open Data**". La démonstration (dont il avouera humblement, en privé, que l'idée est de sa pimpante collègue Valérie Schlosser) est aussi brillante que convaincante. En simulant le parcours d'un utilisateur moyen au sein de cette "place de marché globale", l'ingénieur dévoile en quelques clics combien il est simple d'accéder aux plus basiques éléments de la statistique nationale. Dans les faits, il n'est pas certain que la manipulation sera aussi aisée que celle, balisée, préparée, qui fut offerte à la presse. Néanmoins, la salle frémit ; l'enthousiasme est là, et il est transmis sans quasiment aucun accroc.

## Data martyrisée, mais data libérée

Reste que dans les documents de travail que nous avons pu consulter, et dont **Regards Citoyens s'est également fait l'écho**, l'ouverture des données concerne à ce jour, avant tout, la *quantité* des informations. S'agissant des formats de fichiers utilisés sur le site,

l'ouverture est beaucoup moins visible.

Car sur un peu moins de 300 000 fichiers présents dans la base, seuls 2 600 environ sont au **format ouvert (CSV, TXT, SHP, XML, PDF, ODS, HTML et RDF)**, soit moins de 1%. Le reste, soit quasiment l'intégralité des jeux de données présents sur le site, sont téléchargeables aux formats propriétaires **DOC** (une trentaine) et **XLS**. Interrogé par nos soins dans le cadre du *Personal Democracy Forum*, Romain Lacombe se montre à la fois prudent et catégorique. Non, il n'existe aucune volonté de privilégier le format propriétaire de Microsoft. À ce jour, la priorité de la mission Etalab (nom de l'équipe chargée de ce chantier) est de mettre à disposition l'ensemble des données publiques disponibles, et il se trouve à ce titre que le format Excel est celui qui est le plus utilisé au sein des différentes administrations françaises, "notamment par l'Insee".



Rebondissant sur l'anecdote, Lacombe ajoute :

***La mission d'Etalab, c'est coordonner l'ouverture des données publiques des administrations de l'État. Promouvoir l'harmonisation et l'utilisation de formats ouverts et réutilisables en fait partie intégrante.***

En outre, le discours de la Mission est bien goupillé : primo c'est gratuit, ce qui signifie que le fonctionnement du projet ne peut inclure la reconstitution coûteuse de jeux de données inédits. On prend ce qu'on a. Deuxio, le choix des formats de fichiers par les différentes administrations est antérieur à la Mission, on ne peut pas demander à celle-ci de prendre en charge la migration d'une quantité phénoménale de data d'un format propriétaire vers l'univers du libre. Arguant qu'il faut laisser le temps au temps, l'équipe chargée de la libération des données publiques en France a donc choisi de suivre **les recommandations indirectes** [en] de "l'inventeur du Web", Tim Berners-Lee : *balancer les données en vrac*, on verra plus tard pour le reste.

## Vers l'infini, et au-delà

Etalab n'est donc qu'un moteur de recherche, ni plus ni moins. D'où le soin apporté à bien

communiquer sur la mention BETA du site, qui sonne comme une promesse des jours meilleurs – en espérant évidemment que l'État sera plus prompt que Google, **spécialiste du genre**, à éliminer ce statut qui fleure bon l'inachevé. Un moteur qui recherche dans un catalogue de documents déjà présents, le plus souvent, sur les plates-formes des différents ministères. Etalab peut toujours communiquer (un peu systématiquement) sur **la géolocalisation inédite** des gares SNCF, la réalité est aride : soit les jeux de données disponibles existent ailleurs, soit **de nombreuses données** à ce jour réclamées par les chercheurs, les journalistes et les simples citoyens ne sont pas à disposition du contribuable – alors qu'elles lui appartiennent, *de facto*.



**Un journaliste de données : Data.gouv ? À moi, ça me sert à rien.**



Ceci étant posé, à ce stade, savoir s'il est bon ou pas de pouvoir accéder à ces nombreuses ressources dans un format ou dans un autre semble donc un peu secondaire. Pour plusieurs journalistes de données ("*data-journalists*") interrogés dès le lendemain de l'ouverture d'Etalab, pour eux qui manipulent ces fichiers bruts à longueur de journées – sautant d'un tableur à un autre pour tirer la substantifique moelle de jeux de données – un fichier en format ouvert de type PDF peut s'avérer être parfois plus compliqué à exploiter s'il est mal élaboré qu'un fichier dans un format propriétaire construit avec méthode. "*Les fichiers PDF, c'est sympa, mais ce sont souvent des fichiers inexploitable, conçus à partir de données 'propres'* [qu'elles proviennent de formats ouverts ou pas, ndlr], *et sur lesquels il faut passer un temps fou pour les remettre en format exploitable. C'est kafkaïen !*", argue même François Bancilhon, le directeur général de **Data Publica**, dont le métier, entre autres, est justement de (re)mettre à disposition des données exploitables.

La vraie question qui point au-delà de l'enthousiasme convenu par cette volonté du gouvernement de promouvoir une certaine transparence démocratique, c'est surtout : *et après ?*

## Top-down et bottom-up sont dans un bateau

*Et après*, la question centrale du déploiement et de la promotion de l'Open Data, que nous posions déjà **en février dernier**, c'est la place du citoyen-contribuable au sein de ce dispositif complexe. Derrière les concepts de stratégie de gouvernance un peu pompeux de "top-down" et de "bottom-up" se cache l'enjeu majeur de la réalisation (ou pas) d'un projet de transparence démocratique à la française. Soit les données s'ouvrent *de haut en bas*, considérant le citoyen comme un consommateur passif – on ose l'euphémisme – et par

conséquent bien incapable de contrôler la qualité et la pertinence de la donnée qui lui est offerte par l'État ; soit les données s'ouvrent *de bas en haut*, portées par la demande du citoyen envers son administration, en s'assurant que celle-ci a les moyens – et la volonté – d'y répondre. Ce qui, globalement, aujourd'hui, est loin d'être le cas.



@nicolaskb  
Nicolas Kayser-Bril

#opendata au quotidien: "Le document sollicité n'existe pas en l'état". Offert aujourd'hui par le ministère de l'Intérieur.  
#foutagedeGueule

Pour le ministère de l'intérieur, supprimer une colonne dans un tableur dépasse le "simple traitement automatisé d'usage courant" #opendata

Face au doute, la mission Etalab sort un atout de sa manche : *les DataConnexions*. Le "moteur de recherche" data.gouv.fr doit évoluer, et il le fera "dès janvier" sous les auspices de l'innovation et l'autorité des têtes bien faites de partenaires déjà programmés, tels Orange, Inria, l'Afnic, l'Epita ou encore Oséo. Véritable "programme de soutien à l'innovation", il va permettre "d'enrichir Etalab et sortir du simple moteur de recherche qu'il est actuellement". En clôture du *Personal Democracy Forum*, Séverin Naudet **en remettra d'ailleurs une couche**, évoquant même un programme "enrichi d'espaces collaboratifs et d'échanges entre ses utilisateurs et ses producteurs", un "espace dédié à la mise en valeur des réutilisations des données les plus innovantes".

## Paroles, paroles

On a forcément envie d'y croire, à cette ouverture de l'ouverture. D'autant que dans le rapport rédigé par Romain Lacombe, qui tient nécessairement une place importante dans la mise en place du portail gouvernemental – rapport déjà mentionné plus haut – figurent quelques passages qui pourraient inciter à l'optimisme. Comme *OWNI* **le relevait** en juillet dernier :



***L'État devra donc réfléchir à la possibilité de passer d'un modèle "à sens unique" (diffusion des données du secteur public vers la société civile) à un modèle d'écosystème où les données de l'État et des collectivités, ouvertes à la société civile, pourraient être enrichies en retour de façon collaborative ("crowdsourcing").***



Collaboration certaine de centres de recherche et de l'enseignement supérieur, participation éventuelle du citoyen à la constitution de bases de données et d'applications réutilisant vertueusement ces bases de données. On n'est pas loin du but. À condition, là encore, d'aller jusqu'au bout de la démarche de transparence et de coupler ces efforts et ces ambitions d'Open Data avec celles de l'Open Gov – ou "gouvernance ouverte" – comme le rappelle régulièrement l'association LiberTIC à travers son porte-voix **Claire Gallon**. Une libération des données qui prépare la gouvernance ouverte : là, on en demande peut-être *un peu trop* à la France.

Invitée par Barack Obama et Dilma Rousseff, la présidente du Brésil, à prendre siège autour de la table du projet *Open Government Partnership* (OGP), la France ne fait aujourd'hui pas partie de **la cinquantaine de pays** [en] s'étant engagés fermement à suivre **les intentions vertueuses** de cette initiative promue par l'Onu – dont le programme est pourtant alléchant : “engagement à la disponibilité accrue d'informations relatives aux activités gouvernementales”, “engagement à promouvoir la participation civique”, “engagement à faire appliquer par les administrations les normes les plus strictes d'intégrité professionnelle”, ou encore “engagement à intensifier l'accès aux nouvelles technologies à des fins de transparence et de responsabilisation”. Le reste est à l'avenant. Des pays européens comme la Grande-Bretagne, la Grèce, l'Espagne ou l'Italie, la Suède, la Norvège ou encore le Danemark ont franchi le pas vers l'avant que le couple franco-allemand aura décidé de ne pas faire.

## Transparence à la française

Ce manque de volontarisme apparent de la France serait officiellement une question d'agenda, la demande de l'OGP ayant vraisemblablement été envoyée au cabinet de Christine Lagarde – alors à Bercy – qui n'y aurait pas répondu dans les temps. Selon une source proche du dossier, cette absence remarquée pourrait toutefois s'expliquer par d'autres raisons plus politiques. Articulé autour d'une vision anglo-saxonne de la transparence (*Transparency & Accountability*), le projet *Open Government Partnership* pourrait imposer cette vision dans laquelle la frontière entre transparence de l'État et transparence de la vie privée est si ténue qu'elle s'opposerait illico au modèle de protection des données personnelles qui fait loi dans l'Hexagone. Méfiance donc. Méfiance justifiée ? En vérité, les grands principes de l'OGP sont évidemment compatibles avec une vision cocorico de la transparence de l'État ; il semblerait que les inventeurs du clavier AZERTY aient aujourd'hui surtout besoin d'une “transparence à la française”.

De plus, un engagement de la France dans le programme Obama nécessiterait l'existence d'un agenda politique dédié à la gouvernance ouverte. En Grande-Bretagne, il existe une équipe “Transparency” au sein du *Cabinet Office*, ainsi qu'un “**Public Sector Transparency Board**” [en] qui oriente la démarche Open Data depuis 2010, et qui compte parmi ses membres des éminences telles que Tim Berners-Lee et Nigel Shadbolt (les pères spirituels du portail [data.gov.uk](http://data.gov.uk), en ligne depuis 2009) ou Rufus Pollock de l'**Open Knowledge Foundation**. En France, la nécessité de relier *Open Data* à *Open Gov* semble encore avoir un bout de chemin à faire, malgré l'énergie déployée par Séverin Naudet et son équipe à passer pour de radicaux rénovateurs de la société. Ou même mieux, si l'on en croit **Franck Riester** :



***L'Open Data, c'est une révolution silencieuse.***



Entre contrepoint, par précaution, on rappellera **une récente intervention** de David Eaves, célèbre conseiller canadien sur les questions d'ouverture des données publiques :



***Un risque majeur pour l'Open Data, c'est que tout notre travail soit réduit à n'être qu'une initiative pour la transparence et aurait par conséquent pour unique objet de mettre en conformité des structures gouvernementales. Si c'est ainsi que se joue notre destin, je suspecte que dans 5 à 10 ans les gouvernements, désireux de pratiquer des coupes budgétaires, n'inscrivent les portails Open Data dans la liste des économies à réaliser.***



Construit vingt ans avant la grande Révolution française, l'Hôtel de Cassini porte le nom

d'une célèbre famille d'astronomes royalistes, savoyards puis français par naturalisation, dont le patriarche Giovanni Domenico fut notamment le premier à diriger l'observatoire de Paris. C'est lui qui découvrit la grande tache rouge de Jupiter, qui mesura la distance de la Terre au Soleil et qui mourut aveugle, après avoir passé sa vie à observer l'infini et tenté d'organiser les étoiles.

Photos au mobile par Nicolas Patte /-)



Retrouvez **notre Une Open Data**

## La France entr'ouverte Vous n'habitez pas chez vous Un état des lieux de l'Open Data

1. Logica, Mondeca et Exalead [↔]

### MICKAËL JORDAN

le 11 décembre 2011 - 13:07 &bullet; SIGNALER UN ABUS - PERMALINK



1. En tant que développeur d'une plateforme Open Data internationale, citizensup, je ne pense pas que le format XLS soit un si grand problème que ça.

En effet, on peut très facilement convertir un fichier XLS en CSV, puis exploiter ce CSV pour faire de la DataViz par exemple (en le couplant à un fichier XML).

Donc cette phobie du fichier Excel n'est pas justifiée je pense.

2. Les jeux de données de data.gouv.fr sont intéressants, mais le format "moteur de recherche de fichiers de données à télécharger" va encore échouer à susciter l'intérêt du citoyen normalement constitué. Personnellement je vais lancer un tableau de bord interactif où chaque citoyen peut suivre les données gouvernementales du pays du monde entier sous forme de graphes, cartes, bubble chart...

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### AN391

le 12 décembre 2011 - 15:53 &bullet; SIGNALER UN ABUS - PERMALINK



Pour que l'open data se mette vraiment en place il faudrait déjà que l'informatique sorte de son paroxysme ou syndrome du cordonnier toujours le plus mal chaussé, et reconnaisse que comme n'importe quel autre domaine technique de publications ou de nommages, elle a besoin d'un stock d'identifiants genre codes barre GS1 (et principes de distribution qui vont avec), pour s'écrire et fonctionner.

Un peu plus développé là :

<http://iiscn.wordpress.com/about/>

ou là en anglais :

[http://groups.google.com/group/comp.lang.lisp/browse\\_thread/thread/1b182a31305bf0b9#](http://groups.google.com/group/comp.lang.lisp/browse_thread/thread/1b182a31305bf0b9#)

La plupart des idiots informatiques n'ont toujours pas compris qu'XML ou quoi que ce soit ça ne sert à rien si il n'y a pas de partage de codes (au sens identifiants) entre les deux côtés.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

## 10 pings

Data.gouv.fr by guish - Pearltrees le 10 décembre 2011 - 11:40

[...] La France entr'ouverte » OWNI, News, Augmented Officiellement, le 5 décembre, la République a donc ouvert ses données publiques lors d'une sauterie organisée sous les lambris de la salle de la Chapelle, en l'Hôtel de Cassini à Paris. Au terme de 817 jours

de développement, de 691 624 lignes de code déployées par des partenaires technologiques français le site data.gouv.fr a vu le jour en grandes pompes, mais en version beta . Viennoiseries [...]

Définitions/concepts/réflexions by josselinhenno - Pearltrees le 10 décembre 2011 - 18:22

[...] Et après, la question centrale du déploiement et de la promotion de l'Open Data, que nous posions déjà en février dernier, c'est la place du citoyen-contribuable au sein de ce dispositif complexe. Derrière les concepts de stratégie de gouvernance un peu pompeux de "top-down" et de "bottom-up" se cache l'enjeu majeur de la réalisation (ou pas) d'un projet de transparence démocratique à la française. Soit les données s'ouvrent de haut en bas, considérant le citoyen comme un consommateur passif – on ose l'euphémisme – et par conséquent bien incapable de contrôler la qualité et la pertinence de la donnée qui lui est offerte par l'État ; soit les données s'ouvrent de bas en haut, portées par la demande du citoyen envers son administration, en s'assurant que celle-ci a les moyens – et la volonté – d'y répondre. La France entr'ouverte » OWNI, News, Augmented [...]

data.gouv « le café clic le 11 décembre 2011 - 13:54

[...] lire aussi sur Owni un article intéressant sur [...]

Libertés by olivier93fr - Pearltrees le 11 décembre 2011 - 14:08

[...] L'État devra donc réfléchir à la possibilité de passer d'un modèle "à sens unique" (diffusion des données du secteur public vers la société civile) à un modèle d'écosystème où les données de l'État et des collectivités, ouvertes à la société civile, pourraient être enrichies en retour de façon collaborative ("crowdsourcing"). Collaboration certaine de centres de recherche et de l'enseignement supérieur, participation éventuelle du citoyen à la constitution de bases de données et d'applications réutilisant vertueusement ces bases de données. On n'est pas loin du but. La France entr'ouverte » OWNI, News, Augmented [...]

Open Data by showloulou - Pearltrees le 11 décembre 2011 - 15:55

[...] La France entr'ouverte » OWNI, News, Augmented Ce jour-là, semillant comme à son habitude, le jeune (35 ans) directeur du projet gouvernemental Séverin Naudet fait la claque avec Prezi, le logiciel de présentation en ligne qui relègue PowerPoint au Paléolithique moyen. Rompu à l'exercice, il parvient miraculeusement à tenir en haleine le parterre de journalistes en déroulant les écrans de tableur. Aimable, cet ancien producteur des Nada Surf et pratiquant le Kung-fu sait flatter ses collaborateurs, son audience et ses soutiens anonymes, qui participèrent assidûment à "quatre workshops avec l'écosystème Open Data durant les neuf derniers mois". Et ses sponsors aussi - nombreux - dont on entend les noms de Google et de... Microsoft sans en voir la présence sur les documents distribués avec les viennoiseries. Chargé de l'innovation dans l'équipe, Romain Lacombe est au cœur du projet. [...]

Etalab by le CNNum ou CNNumérique - Pearltrees le 13 décembre 2011 - 23:20

[...] La France entr'ouverte » OWNI, News, Augmented Non, il n'existe aucune volonté de privilégier le format propriétaire de Microsoft. À ce jour, la priorité de la mission Etalab (nom de l'équipe chargée de ce chantier) est de mettre à disposition l'ensemble des données publiques disponibles, et il se trouve à ce titre que le format Excel est celui qui est le plus utilisé au sein des différentes administrations françaises, "notamment par l'Insee". [...]

Régions.news #16 – Edition du vendredi 30 décembre 2011 | Régions.news le 30 décembre 2011 - 19:53

[...] La France entr'ouverte publié le 13/12/2011 par Owni [...]

Le Parti Pirate voudraient bien hacker l'Open Data » revue du web, Just another weblog le 1 juin 2012 - 16:08

*[...] quelques mois) souhaite prolonger la libération des données entamée par l'Etat français (lire ici ou encore là) en poussant le concept de transparence jusqu'au [...]*

L'Open Data payant s'ouvre à la gratuité des débats - Rennes 1720 le 30 octobre 2012 - 23:34

*[...] la circulaire Fillon ayant officiellement lancé la mission d'ouverture de données publiques Etalab, l'Open Data est devenu un non-sujet au [...]*

Vol noir de corbeaux sur l'Open Data français | loKaliz.me le 4 novembre 2012 - 9:17

*[...][...]*