

Laure Lucchesi, Etalab : « plus de 1000 réutilisations de données pour data.gouv.fr 2.0 »

Laure Lucchesi est directrice adjointe d'**Etalab**, la mission chargée de piloter l'ouverture et le partage des données publiques (Open Data), sous l'autorité du Premier ministre. Dans cet entretien, Laure Lucchesi s'exprime sur les applications concrètes de l'Open Data et la démocratisation des usages.

Silicon.fr – **Des projets concrets de réutilisation des données publiques ont été récompensés lors du [concours Dataconnexions 5](#). Mais il semble que les citoyens soient encore peu impliqués dans le mouvement Open Data. Pour quelles raisons, selon vous ?**

Laure Lucchesi – Comme sur le web, il faut distinguer les visiteurs des contributeurs et réutilisateurs : sur Wikipédia comme sur Tripadvisor ou sur les forums de Silicon.fr, les contributeurs représentent en général environ 1/1000e des visiteurs. Très souvent, les citoyens sont au contact de données publiques sans toujours le savoir : prévisions météorologiques, données géographiques et cadastre, statistiques de l'INSEE, résultats des élections, jurisprudence sur Légifrance, etc.

Ce qui est plus nouveau, c'est qu'au-delà de ces données construites, connues de tous, nous travaillons aussi à la mise à disposition des données brutes existant dans les systèmes d'information des administrations, dont le potentiel peut être exploité plus largement. Cette matière brute, qui doit être diffusée dans les formats les plus ouverts afin qu'elle puisse être lisible par des machines, prend sens lorsqu'elle est réutilisée. Savoir ouvrir un fichier CSV, produire un graphe ou calculer une distribution peuvent alors s'avérer nécessaires. Ces démarches ne sont pas tout à fait « grand public », mais nous rencontrons et travaillons avec de très nombreux réutilisateurs, qui peuvent être journalistes, chercheurs, étudiants, membres de très petites entreprises et aussi des collectifs comme OpenStreetMap, OpenFoodfacts, etc.

La mécanique sociale adoptée pour [data.gouv.fr](#) vise à attirer sur la plateforme cette communauté de réutilisateurs qui y publient le fruit de leurs travaux – applications, datavisualisations, analyses – sur la même page que les données brutes, et participent ainsi à la médiation entre les données et les citoyens. Plus d'un an après l'ouverture de la nouvelle version de data.gouv.fr, plus de 1000 réutilisations citoyennes donnent vie à ces données publiques. C'est encourageant !

Plus l'ouverture des données se poursuit, plus on met à disposition des ressources accessibles à tous, plus on avance vers une appropriation de ces ressources par les citoyens. Les réutilisateurs [candidats](#) et [finalistes](#) de Dataconnexions 5 en sont la preuve : cette année plus encore que les précédentes, il n'y avait pas seulement des startups. De nouvelles communautés se sont emparées des données publiques grâce à des ouvertures récentes dans les domaines juridiques ou de la santé par exemple. Par ailleurs, beaucoup de porteurs de projets se présentaient à titre individuel. Le cas du projet « [SuperLachaise](#) », coup de coeur du jury, est un exemple : un frère et une soeur se

sont lancés bénévolement dans la cartographie du cimetière.

C'est l'une de nos ambitions : la production de données et leur analyse peuvent être le fait de tous. L'un des objectifs de l'open data est bien de stimuler la démocratie, en donnant de nouveaux pouvoirs d'agir aux citoyens. Il s'agit non seulement d'être transparent et de rendre compte, mais aussi d'accepter le dialogue avec ceux qui utilisent les données, parfois pour proposer des points de vue contradictoires. C'est en cela que l'open data engage une transformation plus large, celle de l'« open government » ou de la « gouvernance ouverte », qui promeut l'ouverture de l'action publique à de nouveaux modes de concertation et de collaboration avec la société civile et les citoyens.

Une première version agrégée de la base de données de santé DAMIR de l'Assurance maladie (Cnamts) est disponible sur data.gouv.fr depuis mi-décembre. Quels acteurs vont bénéficier de l'ouverture d'autres jeux de données de santé ?

L'ouverture des données de santé est un chantier majeur de l'open data en France. À la suite du débat décidé par le gouvernement en décembre 2012 et de l'ouverture la Commission Open Data en Santé lancée en novembre 2013, la mise à disposition des données progresse. On y retrouve un triple bénéfice pour les citoyens : il y a des données qui relèvent d'un savoir directement utile à tous, telles que celles sur l'offre de soins, l'épidémiologie, les caractéristiques des médicaments... Il y a ensuite des données auxquelles des réutilisateurs avertis, startups, chercheurs donnent du sens. Enfin, il y a l'amélioration du service public qu'apporte le simple fait d'entrer dans une démarche d'ouverture et de dialogue avec ces innovateurs.

L'ouverture des données de dépenses d'assurance maladie par la Caisse nationale d'assurance maladie le mois dernier a notamment constitué une étape importante : avec 1,5 milliards de lignes CSV, c'est [l'un des plus gros fichiers mis à disposition à ce jour](#) (ndlr : la base complète est disponible depuis le 13 février 2015). Comme l'ont déclaré la [ministre Marisol Touraine](#) et la CNAM, ce mouvement d'ouverture est incontournable et va se poursuivre. Au-delà des données de la CNAM, d'autres jeux de données utiles sont mis à disposition sur data.gouv.fr : le répertoire FINESS par exemple, qui renseigne sur les établissements sanitaires et sociaux, ou encore la Base publique des médicaments de l'ANSM.

Comme le [Hackathon données de santé \(#HackDDS\) du 26 janvier](#) l'a prouvé, toute une diversité d'acteurs profitera de ces données : des associations, des startups, comme par exemple la toute jeune medicat.io lauréate de Dataconnexions 5, ou des acteurs plus installés comme Withings, des bureaux d'étude, des chercheurs... Mais aussi des acteurs publics et des administrations elles-mêmes : c'est ce qu'a montré l'ARS Paca à Dataconnexions, avec sa [datavisualisation](#) présentant de façon dynamique l'évolution des dépenses de santé dans la région.

La France fait partie des leaders mondiaux de l'Open Data. Les communautés techniques et administratives sont les premières à en bénéficier. Comment élargir l'audience et les usages des données publiques ouvertes ?

La diffusion de la culture de la donnée prendra du temps, mais nous multiplions les rencontres où nous sensibilisons à la donnée : « Open Data Camps » que nous organisons presque tous les mois désormais, événements « Bonjour Data » où l'équipe d'Etalab se tient à la disposition de tous une

fois par mois, ateliers... Nous intervenons également dans des formations dans plusieurs écoles : Sciences Po, l'Essec, Epitech... mais aussi dans des formations continues dans l'administration, avec l'Ecole de la Modernisation de l'Etat. Des initiatives s'organisent également du côté de la société civile, avec par exemple le très beau projet « [Ecole des Données](#) » de l'Open Knowledge Foundation France.

Une grande partie de la stratégie numérique du gouvernement, comme l'enseignement du code, le soutien à la FrenchTech, les investissements d'avenir, les plans filières, contribuera à augmenter l'intensité numérique de l'économie et de la société, et donnera une importance croissante à l'ouverture des données publiques. L'open data constitue une ressource pour des citoyens mieux éclairés, capables de construire leurs propres points de vue et d'organiser de nouvelles solidarités. C'est aussi à la société – c'est-à-dire à nous tous – de s'en emparer et de la faire fructifier.

Lire aussi :

[Open Data : la France fait bien partie des leaders mondiaux](#)

[Open Data : la France se hisse au 3e rang mondial](#)

[Henri Verdier, Etalab : « L'Open Data est un droit de l'Homme »](#)