

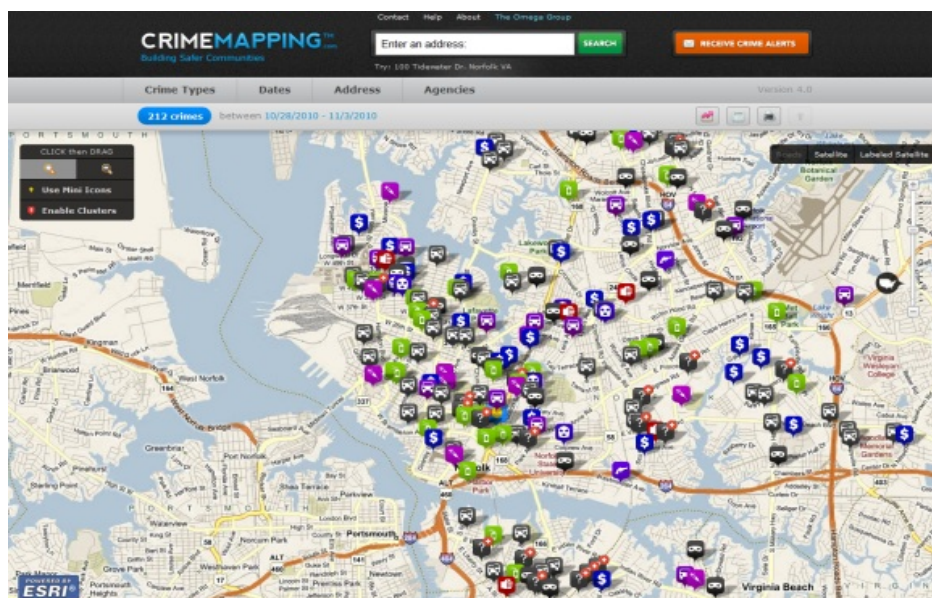
L'OUVERTURE DES DONNÉES PUBLIQUES, ET APRÈS ?

LE 18 JANVIER 2011 DANIEL KAPLAN

Daniel Kaplan, directeur de la Fing, explore les conséquences de l'ouverture des données publiques et leur réutilisation. Une réflexion nécessaire au mouvement de l'open data.

Tous ceux qui, comme nous dans le cadre du programme **Réutilisation des données publiques de la Fondation internet nouvelle génération**, s'engagent en faveur de l'ouverture et de la réutilisation des données publiques, en espèrent des résultats féconds en termes de qualité de vie, de cohésion sociale, d'innovation et de croissance. Mais les choses pourraient se passer tout autrement. Nous devons commencer à penser aux conséquences de l'ouverture des données, pour nous assurer qu'elles soient majoritairement positives.

Ceux qui militent en faveur de l'ouverture des données publiques (ou non, d'ailleurs) et de leur réutilisation par les citoyens, les chercheurs et les entrepreneurs, espèrent qu'il en sortira quelque chose de bon. Ils ont sans doute une idée différente de ce qu'est ce "quelque chose", et de ce que "bon" veut dire : je peux considérer qu'il est bon de créer de la valeur marchande à partir de données publiques gratuites, alors que d'autres ne le penseront pas. Je peux détester les "crime maps" qui se multiplient dans les pays anglo-saxons, parce qu'elles stigmatisent sans rien résoudre, mais d'autres espèrent qu'elles sauveront des vies. Jusqu'ici, tout va bien. C'est même une bonne raison de soutenir l'ouverture des données : parce qu'elle crée le terreau commun sur lequel différents acteurs, avec des motivations différentes, créeront des choses différentes – et Dieu (ou Darwin) reconnaîtra les siens au bout du compte.



Explorer les conséquences

Mais les choses ne s'avéreront pas forcément aussi simples. **L'ouverture des données publiques est un processus vaste, complexe, sans fin, qui implique des milliers d'acteurs. Elle redistribue l'information, le pouvoir, les responsabilités, selon des lignes difficiles à anticiper.** Ses conséquences se feront probablement sentir dans de très nombreux domaines : les métiers des opérateurs de services, les logiciels de gestion, la démocratie participative, les fiches de postes des fonctionnaires, le budget des organismes publics spécialisés dans la production de données (comme l'IGN ou l'Insee), l'économie et le contenu des médias, etc.



Et, selon la manière dont nous ouvrirons l'accès aux données, et selon ce que nous ferons après avoir ouvert cet accès, certaines de ces conséquences pourraient être moins positives que nous ne le souhaiterions.



Commençons donc à penser aux conséquences. Imaginons que nous avons gagné : une part très significative des “données de service public” sont désormais accessibles et réutilisables, brutes, en un format lisible par des machines, à un coût faible, voire (le plus souvent) nul. Avançons de quelques années : en quoi cela a-t-il changé la vie quotidienne des citoyens et des entreprises ? Et notre potentiel d’innovation et de croissance ? Et la capacité des acteurs publics à atteindre des objectifs collectifs (environnementaux ou sociaux, par exemple) ? Et la démocratie, l’esprit public ?

Un avenir radieux

Le récit dominant de l’ouverture des données publiques décrit un avenir radieux. Des groupes de citoyens mettent à jour les abus de pouvoir, et participent de manière active à des débats publics où l’échange d’avis informés permet d’aboutir à de meilleures décisions. Ils contribuent même à enrichir le stock des données publiques, comme ils le font déjà aujourd’hui sur **Open Street Map**. Des grandes entreprises comme des jeunes pousses créent toutes sortes de nouveaux services et contribuent à la croissance et à la qualité de la vie. Chercheurs et “data-journalistes” passent en revue des masses de données pour enrichir la compréhension, par exemple, des changements climatiques ou des dynamiques urbaines. Les institutions publiques collaborent entre elles pour éliminer des procédures redondantes et avec les entreprises et les associations pour produire de meilleurs services publics. Nourrie par un aller et retour constant d’informations, d’évaluations et d’échanges, la démocratie trouve une nouvelle jeunesse.

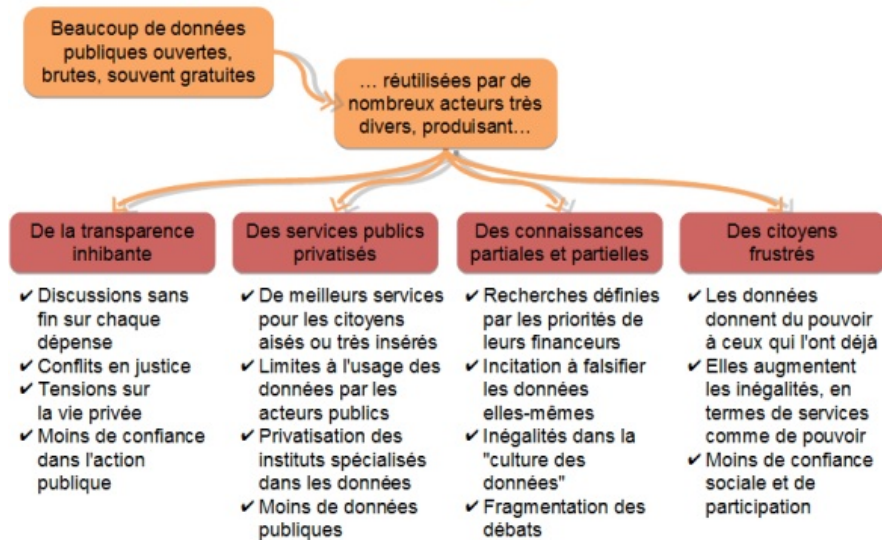


Nous pouvons bien sûr produire plusieurs exemples et quelques études à l’appui de ce récit. Mais si nous décidons au contraire de nous concentrer sur le côté sombre, il devient assez facile, à partir du même point de départ, de raconter une histoire tout à fait différente.

Un tour du côté sombre

Supposons, par exemple, que la plupart des services innovants fondés sur l’usage de données publiques sont proposés par des entreprises, et financés soit par leurs utilisateurs, soit par la publicité. Pour vivre, ces entreprises se focalisent sur les territoires et les populations les plus solvables. Après avoir montré que leurs services sont bien meilleurs que ceux des institutions, elles obtiennent que celles-ci n’aient plus le droit de leur faire concurrence. Le secteur public se retrouve cantonné à servir les populations délaissées par le marché, et les inégalités devant les services fondamentaux s’accroissent.

Le côté sombre des données publiques ouvertes : plausible ?



Supposons ensuite que les instituts publics dont le métier consiste à produire des données, géographiques ou statistiques par exemple, ne parviennent plus à tirer de revenus significatifs de la vente de ces données. Elles tentent, sans y parvenir, de concurrencer les cabinets d'étude ou les éditeurs professionnels. L'état des finances publiques ne permettant pas d'augmenter leur dotation budgétaire, elles doivent réduire leurs activités. Beaucoup de données cessent d'être produites, ou bien le sont par des entreprises. L'ouverture des données publiques réduirait ainsi le stock de données publiques...

La transparence peut également produire des conséquences problématiques. On peut imaginer qu'au lieu de s'engager dans des discussions ouvertes et constructives sur des sujets d'intérêt général, les groupes d'intérêt s'en servent pour contester la moindre décision, la moindre ligne de dépense publique. **Au lieu d'aboutir à des décisions plus sages et plus largement soutenues, la transparence totale dresse les intérêts particuliers les uns contre les autres, inhibe la prise de décision et réduit la confiance.** La nouvelle visibilité des données invite même à les trafiquer à la source, pour faire en sorte que même des analyses indépendantes, fondées sur des moyens informatiques massifs, produisent les conclusions qu'attendent les producteurs de données. L'ouverture des données publiques réduirait ainsi la fiabilité des données publiques...

Imaginons enfin que la capacité à extraire du sens à partir des masses de données disponibles demeure très inégalement distribuée – ce qui n'apparaît pas invraisemblable. Dans un article intitulé "Du pouvoir à ceux qui l'ont déjà ?" ("**Empowering the empowered ?**"), Mike Gurstein cite l'exemple d'une recherche sur la numérisation des registres fonciers de Bangalore (Inde) :



La mise en accès de l'information sur la propriété foncière à Bangalore a principalement permis aux personnes les plus aisées à quelques entreprises à évincer les populations pauvres et marginalisées de leurs terres.



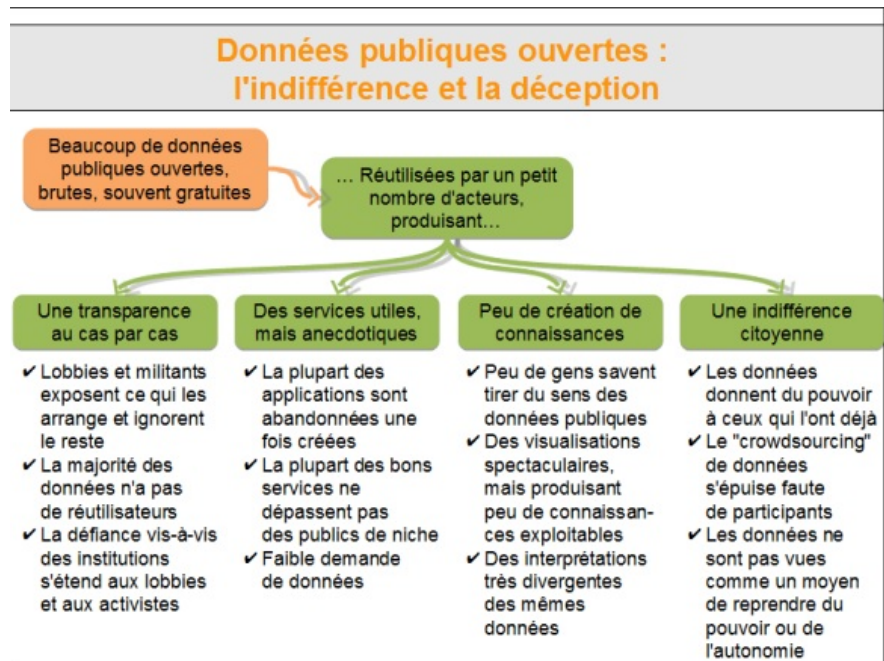
Et ici, en Europe, des lobbies bien financés ne sauront-ils pas, aussi, faire un usage plus efficace des données que de petits groupes locaux ? Les chercheurs qui voudront traiter des données ne devront-ils pas chercher leurs financements chez des acteurs qui auront intérêt à leurs résultats ? Peut-être la "sagesse des foules" corrigera-t-elle ces déséquilibres, mais rien n'est moins sûr.

Data lassitude

Lorsque j'ai présenté ces deux premiers scénarios à l'équipe de la Fing, elle en a

immédiatement produit un troisième, qui décrit une adoption modeste des données publiques, produisant des résultats anecdotiques ou décevants.

Les données publiques sont si nombreuses, complexes, hétérogènes, difficiles à interpréter, que très peu d'acteurs parviennent à réunir le courage ou les ressources pour les travailler sérieusement. Les premières applications que nous voyons aujourd'hui candidater à des concours tels qu'**Apps for Democracy** ou **Show Us a Better Way** en restent à l'état d'aujourd'hui, celui de prototypes, rarement exploitables par des citoyens ordinaires. La plupart ne se convertissent jamais en services opérationnels. De magnifiques visualisations font plus pour la célébrité de leurs concepteurs que pour créer une compréhension partagée de phénomènes complexes. Seuls quelques lobbies plongent dans les données, pour en ramener les résultats qu'ils savaient vouloir y trouver, et sans grand risque de trouver contradictoire. Après quelques contributions sur des sites comme **FixMyStreet**, les citoyens se fatiguent de faire le boulot qu'ils attendent de leurs municipalités, surtout quand celle-ci est plus indifférente que réactive. L'ouverture des données publiques ne fait pas beaucoup de mal, mais pas beaucoup de bien non plus.



Bifurcations



Chacun pourra croire qu'un de ces scénarios (ou d'autres encore) a plus de chances de se réaliser que les autres. Mais sur quelles bases ? Quelles actions, décisions, conditions, ont le plus de chances de nous engager dans une voie plutôt que l'autre ? Peut-on avoir des conséquences positives sur certains plans, négatives ailleurs, nulles ailleurs encore ?



Nous commençons à peine à envisager ces questions. Et pourtant, il faut le faire, maintenant, ou courir le risque de graves déconvenues. Il y a plusieurs manières d'ouvrir les données publiques, et beaucoup de manières d'accompagner cette ouverture. Si, par exemple, on n'ouvre les données que dans l'espoir de réduire les dépenses publiques en transférant la fourniture de nombreux services au privé (voire au secteur associatif), alors la probabilité de voir s'engager les cercles vicieux décrits dans le scénario "sombre" devient très élevée.

Nous devons soulever quelques questions difficiles. Comment le rôle des acteurs publics doit-il évoluer si les entreprises et les citoyens gagnent en capacité d'action grâce aux données qu'ont produites ces mêmes institutions ? Que faudrait-il faire pour élever la capacité de tous de lire et exploiter des données, et qui doit-il le faire ? **Quelles sont les limites de la transparence : la vie privée, la sécurité, la possibilité de peser**

tranquillement des décisions difficiles – ou bien n’y a-t-il pas de limite ? Comment financerons-nous demain la production de données du domaine public ? Comment développer une “culture des données” qui nous rendra même capables de discuter comment les données “brutes” elles-mêmes sont produites – puisque nous savons (au moins depuis Einstein) qu’une donnée est toujours construite par l’observateur.

Acteurs et leviers

La liste qui précède n’est pas close. Bien d’autres défis se feront jour à mesure que les Etats, les collectivités locales, les laboratoires et même des entreprises livreront leurs données, et que ces citoyens, des entreprises, des chercheurs, des activistes, des lobbyistes ou des artistes s’en empareront.

Pour commencer à identifier et à traiter ces questions, nous avons d’ailleurs intérêt à les penser à partir des acteurs : qui est concerné, qui a intérêt à la production, la circulation, la compréhension ou l’exploitation de quelles données ? Ouvertes ou non, les données n’existent pas toutes seules, elles ne circulent pas en apesanteur, mais au contraire dans un espace parcouru de courants, de champs de force, où des intérêts, des habitudes, des capacités, coopèrent ou luttent entre eux. Personne ne tire le même pouvoir des mêmes données, dans les mêmes circonstances.

	Produc- teurs	Réutili- sateurs	Inter- médiaires	Intérêt éco.	Intérêt public
Institutions publiques					
Nationales / Continentales ("Gouvernements")	++	+	+	+	++
Locales	++	+		+	++
Opérateurs de services publics	++	+		++	++
Entreprises					
Fournisseurs spécialisés de services en ligne	+	++	++	++	
Autres entreprises existantes	+	++		++	
Startups		++	+	++	
Fournisseurs de technologies		+	+	+	+
Autres					
Associations, activistes, lobbyistes		++	+		++
Médias, blogueurs		++	+	+	++
Chercheurs	+	++		++	++
Citoyens	+	+		+	++

Nous pouvons également penser en termes de leviers, c’est-à-dire des décisions, des structures et des actions qui peuvent influencer sur les résultats de l’ouverture des données publiques.

Certains de ces leviers sont déjà connus, parce qu’ils émergent naturellement dès que l’on pense à ouvrir les données : les conditions d’accès (licences et tarifs), le financement des producteurs spécialisés de données, les manières de stimuler l’“écosystème” des réutilisateurs de données publiques... **Tout le monde organise des concours afin de stimuler l’inventivité des entrepreneurs et développeurs, mais comment aller chercher d’autres groupes, moins motivés a priori ?** Faut-il organiser des ateliers ou des sessions de formation à destination d’associations, de PME non numériques, de médias locaux ? Faut-il aider certains réutilisateurs plutôt que d’autres ? Faut-il, **comme le propose Tim Davies**, travailler sur des interfaces communes qui facilitent l’accès aux données de la part de non-spécialistes, sans trop influencer sur l’interprétation possible de ces données ?

D’autres leviers ont un caractère plus politique, plus large. Comment diffuser une “culture des données” et éviter que la fracture sociale ne passe par la capacité à comprendre les données ? Cette “culture” doit-elle inclure une réflexion sur le bon niveau de transparence à obtenir des acteurs publics (et privés) ? Et d’ailleurs, certaines données produites par les acteurs privés doivent-elles devenir publiques ? – je pense à Adam Greenfield, s’étonnant que les données que produisent des capteurs privés installés dans l’espace public demeurent entièrement invisibles. Quel est, aussi, le statut des données produites par les citoyens ? En définitive, verrons-nous émerger une fonction de régulation du paysage des données ouvertes ?

A nouveau, la liste n’est pas close. Cet article a précisément pour but de l’ouvrir. Commençons dès aujourd’hui à explorer ensemble ce qu’il se passe après l’ouverture des données publiques. Pour pouvoir, dans quelques années, nous montrer fiers d’avoir contribué à engager ce mouvement.¹

1. La Fing, l'ePSI plateforme (la plateforme d'information du secteur public européen), la Métropole de Rennes, la Région Bretagne, e-megalis et le GFII organisent le 29 novembre à Rennes une journée d'échanges européenne autour des bonnes pratiques d'ouverture et de réutilisation des données publiques. Elle présentera de nombreuses initiatives locales européennes en provenance du Royaume-Uni, de Finlande, d'Espagne, de Pologne et de France... **Si le sujet vous intéresse, ne manquez pas cet événement !** [+]

OURSIN

le 21 janvier 2011 - 11:43 • SIGNALER UN ABUS - PERMALINK



*Mais d'abord qu'est-ce qu'une donnée publique?
Qui décrète son statut de publique ou pas?
Et qui décide de les envoyer au public justement?
Celui qui envoie l'information est toujours celui que ça intéresse de le faire.
Et il n'envoie que ce qu'il veut puisqu'il est le seul à savoir ce qu'il a.*

Plus simplement nous n'aurons comme données publiques que celles que les collectivités, l'état voudront bien nous laisser en pâture, et comme le soulignent justement les contributions antérieures, elles seront à nos frais, ou pillées du travail des employés du secteur public.

Que l'info soit livrée de manière brute pose certes le problème essentiel de l'incapacité à traiter l'information, mais cela est déjà vrai pour l'internet en général.

Tant que la chose publique ne sera pas vraiment sous contrôle démocratique du citoyen, on assistera impuissants à des politiques de pseudo-information, partielles, tendancieuses, manipulatrices et commerçantes. mais on pourra quand même s'amuser avec, comme on le fait déjà.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

DANIEL KAPLAN

le 21 janvier 2011 - 18:18 • SIGNALER UN ABUS - PERMALINK



@oursin : c'est la loi qui dit ce qu'est une donnée publique, pas le bon vouloir de chaque institution. Une donnée publique une information produite ou reçue "dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission" [article 1 de la loi de 1978]. Il y a des exceptions assez précises : en gros, les données personnelles, les données auxquelles s'appliquent des droits d'auteur (sauf accord des auteurs), les données dont la diffusion poserait un problème de sécurité publique.

Donc en principe (et sur ce point la France est juridiquement plus volontariste que pas mal d'autres pays), on peut demander toutes les données qui rentrent dans ce cadre. Et en cas de refus, se plaindre à la Commission d'accès aux documents administratifs, puis au Tribunal administratif qui peut forcer l'acteur public à livrer ces données.

Donc pas trop d'arbitraire du point de vue du droit. Ensuite, ça ne signifie pas que les acteurs publics feront tous un effort pour rendre accessibles en ligne toutes leurs données, d'un simple clic, bien documentées, etc.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

OURSIN

le 21 janvier 2011 - 22:55 • SIGNALER UN ABUS - PERMALINK



*Merci pour cette réponse.
Il est évident qu'on préfère que les données publiques soient en libre accès, ce qui ne veut pas dire gratuites d'ailleurs.
Rajoutons en peu du côté sombre, pour reprendre votre expression.
Que se passera-t-il en définitive?
L'info brute ne sera plus traitée que par des entreprises privées, ce qui est une façon de la privatiser, de la rentabiliser.
Les statistiques publiques disparaîtront, il est d'ailleurs paradoxal qu'à l'heure où on prétend livrer au public ce qui en définitive lui appartient de droit, on affaiblisse et dévoie les organismes de statistiques publiques : diminution de 25% du budget en 2010 de l'INSEE, la DEPP réduite à la propagande ministérielle.
On assiste aussi à la construction de bases de données personnelles, qui elles ne sont*

pas consultables par le citoyen qu'on fiche, mais sans doute seront-elles aussi revendues à des entreprises sous prétexte de traitement et d'optimisation des services publics? comme cela se pratique déjà.

On voit bien quand même que sous couvert d'ouverture, il s'agirait plutôt d'une libéralisation des données publiques, qui n'excluera pas le fichage des citoyens.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

OURSIN

le 22 janvier 2011 - 20:31 • SIGNALER UN ABUS - PERMALINK



Pour finir, je dirais que par exemple nous risquons la publication des résultats aux évaluations nationales scolaires aux portes des écoles, ce qui entraîne la mise en concurrence des établissements, et évidemment cela sera étendu à tous les organismes publics. Puisqu'aucune clause restrictive à la publication ne s'y applique : pas de données confidentielles, pas de mise en péril de la sécurité de l'état. Quand on se bat contre le pilotage par l'évaluation systématique et selon des critères discutables, on peut se poser légitimement des questions sur les risques de ce qui apparaît comme une liberté nouvelle.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

1 ping

OpenData (Owni.fr) | Navicorp le 11 mars 2012 - 2:20

[...] par les citoyens, les journalistes, les universitaires et les communautés open source. Reste à déterminer un futur avec des données en libre accès. Une sélection d'articles d'OWNI pour comprendre la [...]