

L'Open data, pour quoi faire ?

Description

Interview d' Henri Verdier, propos recueillis par Francis Balle

Henri Verdier est administrateur général des données, directeur de la mission Etalab au sein du Secrétariat général pour la modernisation de l'action publique (SGMAP). Francis Balle est professeur à l'Université Paris 2 et directeur de l'IREC (Institut de recherche et d'études sur la communication).

La mission Etalab, installée au sein du SGMAP (Secrétariat général pour la modernisation de l'action publique), service du Premier ministre, est chargée de la politique d'ouverture des données publiques (*open data*), et de conduire la participation française au « Partenariat pour un gouvernement ouvert » (association de 65 pays coopérant pour le développement de pratiques de gouvernement plus transparentes, plus concertées et plus collaboratives, dont la France prendra la présidence en 2016). Il nous a semblé intéressant d'interroger Henri Verdier, nommé administrateur général des données en septembre 2014, afin de dresser un premier bilan de son action et d'envisager les enjeux de l'*open data* dans les années à venir.

Administrateur général des données depuis septembre 2014, « afin, dites-vous, que les bonnes personnes disposent de bonnes données au bon moment pour prendre de bonnes décisions ». Pouvez-vous préciser ?

Je voudrais tout d'abord préciser que cette mission est différente de l'ouverture des données publiques, même si l'intérêt est toujours de conjuguer les deux missions : la force qu'il y a à ce que ceux qui ouvrent les données les connaissent et les manipulent eux-mêmes, et la force qu'il y a à ce que ceux qui définissent une gouvernance de la donnée n'oublent pas que l'ouverture est une forme essentielle de circulation de la donnée, même si ce n'est pas la seule. La création du poste d'administrateur général des données (AGD) répond en effet à l'une des conséquences imprévues de la révolution numérique en cours. L'abondance de données disponibles, leur finesse et leur collecte immédiate permettent d'utiliser les données de plus en plus profondément dans les organisations, et de coller de plus en plus aux décisions individuelles. On est passé en quelques années des statistiques (analyse descriptive pouvant éclairer la décision stratégique du dirigeant), aux analyses prédictives (qui permettent d'éclairer la décision opérationnelle), puis à des dispositifs extrêmement fins qui peuvent intervenir au fil de la décision et éclairer des choix très précis (par exemple pour mieux cibler des contrôles de sécurité). Quand Amazon vous recommande un livre qui correspond à vos achats passés, il va bien au-delà d'une sorte de sociologie des pratiques culturelles. Il crée un dispositif de données très inséré dans l'action concrète. **L'administrateur général des données est chargé de travailler à la gouvernance des données.**

[/pullquote]De nombreuses organisations essaient de tirer les conséquences de cette nouvelle donne, en créant des fonctions nouvelles (le *chief data officer*), en apprenant à utiliser différemment les données, avec des statisticiens post-numériques qu'elles appellent *data scientists*, et surtout avec des réflexions sur de nouvelles stratégies d'action (les *data-driven strategies*). Mais immédiatement, la non-disponibilité des données, leurs règles d'interopérabilité, leurs modes de recueil et de conservation apparaissent comme l'un des freins au développement de ces pratiques. Et c'est pourquoi l'administrateur général des données est également chargé de travailler à la gouvernance des données, c'est-à-dire, pour simplifier, aux règles d'échange et de circulation de ces données. La France est le premier Etat à avoir institué cette fonction. L'équipe de l'administrateur général des données travaille avec les différents ministères afin de les aider à aborder de façons diverses certains enjeux. Au passage, elle soutient également la diffusion de ces pratiques dans les administrations et essaie de fluidifier la circulation des données entre administrations, dans le respect des secrets légaux.

Pourriez-vous illustrer l'état d'avancement du chantier ouvert en septembre 2014 ?

L'équipe s'est mise au travail avec la conviction que la révolution des données, comme la révolution numérique, est une révolution pratique et que l'important est d'amorcer le mouvement. Fondamentalement, sa mission comprend deux missions complémentaires : – développer la culture de l'analyse prédictive et des *data sciences* au sein de l'administration. Pour cela, elle travaille pour le compte des administrations qui la sollicitent, ou en appui de ces administrations. Elle a par exemple mené un important travail avec le service des achats de l'Etat (SAE) pour mieux suivre, et donc contrôler plus efficacement, les factures d'électricité de l'Etat (<http://agd.data.gouv.fr/2015/05/17/analyser-les-consommations-energetiques-des-batiments-publics/>). Elle travaille aujourd'hui avec Pôle emploi sur une méthode permettant de mieux caractériser les entreprises qui vont sans doute recruter prochainement. Plusieurs autres projets sont en cours de démarrage. [pullquote]**Développer la culture de l'analyse prédictive et des *data sciences* au sein de l'administration.**[/pullquote] – **soutenir l'émergence d'une culture de la donnée propice au recours aux *data sciences*.** Pour cela, l'équipe de l'AGD travaille avec la DISIC (Direction interministérielle des systèmes d'information et de communication), qui est chargée notamment de la numérisation du système d'information de l'Etat. Elle a mis en place un processus de saisine à partir de son blog (<http://agd.data.gouv.fr/>) et a aidé quelques dizaines de postulants. Officiellement saisie par le ministère des affaires sociales et de la santé, elle a remis un premier avis qui a nourri la préparation du projet de loi sur la santé et a commencé une étude sur la gouvernance de l'AGD. Enfin, à la suite des premières analyses de l'AGD, le Premier ministre vient de lancer une mission de réflexion sur le coût induit, au sein de l'Etat, par la facturation des données entre les administrations.

Troisième après le Royaume-Uni et le Danemark selon Open Knowledge Foundation, Quatrième selon l'e-government Survey de l'ONU, après la Corée du Sud, l'Australie et Singapour. Ces pays sont-ils des exemples à suivre ? Afin de les dépasser ?

Et même deuxième d'après l'OCDE... [pullquote]**Le principal intérêt de l'ouverture des données publiques est de favoriser et de permettre la modernisation du fonctionnement de l'Etat.**[/pullquote] La

question n'est pas tellement le classement, et il faut savoir le prendre avec les réserves nécessaires. Quand on regarde de près, les pays n'ont pas tous les mêmes données (le Royaume-Uni, par exemple, a privatisé de nombreux services publics dont les données, de ce fait, échappent à l'*open data*, d'autres pays ont perdu la capacité de produire une information statistique, géographique ou météorologique, par exemple). Ils n'ont pas tous les mêmes organisations administratives. Ils ont parfois même de réelles divergences dans les objectifs qu'ils assignent à leurs politiques d'*open data*, qui peuvent être, selon les sensibilités, plus axées sur la transparence du gouvernement, sur la fourniture de ressources à l'économie, sur la création de lieux d'échange avec la société civile. Nous avons la chance de connaître nombre de nos collègues, notamment grâce au Partenariat pour un gouvernement ouvert. Nous avons tout à apprendre des échanges que nous avons avec eux. Mais pas tellement pour les imiter. Plutôt parce que le principal intérêt de l'ouverture des données publiques est de favoriser et de permettre la modernisation du fonctionnement de l'Etat, les relations entre l'Etat et la société, et la vie démocratique, économique et sociale. Les expériences qui travaillent sur les référentiels clés, les études sur l'impact économique de l'ouverture des données, les expériences pour utiliser les données publiques au service de la régulation, du développement économique, de l'innovation ouverte ou de l'engagement des citoyens, par exemple, sont riches d'enseignements pour nos propres projets.

Quelles sont les grandes lignes de votre plan pour le « gouvernement ouvert » ?

En adhérant au Partenariat pour un gouvernement ouvert, la France s'est engagée à définir avec la société civile un plan d'action pour les deux prochaines années. Nous travaillons le nôtre depuis plus de huit mois et devrions le présenter fin juin. Il résulte d'une concertation conséquente, avec plus de 80 ateliers, forums ou auditions, de la consultation numérique portée par le Conseil national du numérique, sans oublier l'intense travail interministériel. C'est un plan ambitieux qui comporte de nombreuses mesures. Globalement, elles sont regroupées autour de trois grandes thématiques :- rendre des comptes à la société, pour renforcer, grâce à la transparence, l'intégrité de la vie publique, et donc la confiance ; – ouvrir la décision publique, en la rendant plus lisible, mieux documentée et surtout plus concertée ; – et ouvrir l'action publique, en partageant plus de ressources, en encourageant l'innovation ouverte et en stimulant la contribution des citoyens, notamment grâce à l'*open data*, mais aussi en proposant des modèles (comme nous l'avons fait avec OpenFisca.fr, qui présente le modèle du droit fiscal et social en open source), ou en construisant des ressources collaboratives (comme nous l'avons fait avec la base nationale des adresses). L'ouverture à l'innovation et l'ouverture des carrières ne sont pas absentes de ces projets. Outre ces mesures générales, le plan s'attache à appliquer cette vision à un enjeu crucial pour les années à venir : le réchauffement global, en cette année où la France va accueillir la conférence sur le climat. Il propose un projet d'ouverture de données, d'innovation collaborative et de transparence accrue sur les débats. C'est aussi un plan qui ouvre une dynamique puisque, dès l'année prochaine, il faudra en évaluer la progression et commencer à préparer le plan suivant prévu pour 2017.

Entre les données accessibles à tous (*open data*) et les secrets légaux, y a-t-il une place et un plan pour les données dont l'accès est subordonné à une autorisation ? Une porte ne doit-elle pas être ouverte ou fermée ?

Si les données personnelles ne sont disponibles et accessibles qu'avec l'autorisation des personnes concernées et contre rétribution, les données « ouvertes » pourront-elles encore donner lieu à des analyses pertinentes ? Incomplètes, ces données ne pourraient alors donner lieu qu'à des analyses « biaisées ».

L'ouverture des données publiques n'a pas été décidée à seule fin de construire des systèmes de représentation du monde, mais pour organiser la « redevabilité de l'Etat » : son devoir de rendre des comptes. Vos questions portent donc sur des problèmes nouveaux, très récents, sur lesquels il va falloir construire une position collective. Les secrets légaux ont été instaurés pour protéger des personnes ou des entreprises. Ils ont des raisons d'exister et la question de leur transposition dans un monde de données numériques doit être traitée avec précaution et discernement. En France, la règle qui prévaut pour l'instant est que nous essayons d'ouvrir massivement les données qui ne sont pas couvertes par ces secrets.[pullquote] **Organiser la « redevabilité de l'Etat » : son devoir de rendre des comptes.**[/pullquote] Cela dit, les données couvertes par les secrets légaux sont également importantes, et peuvent recéler une grande valeur pour la recherche, l'innovation, voire l'efficacité de l'action publique. Paradoxalement, c'est avec la protection de la vie privée que les choses sont le plus facile, puisque la CNIL sait organiser l'accès aux données pour les usages importants et notamment la recherche. De même, sur des sujets précis, le législateur peut organiser l'accès plus large à des données qui étaient jusqu'ici protégées par ces secrets. Pour les autres secrets, il va peut-être falloir apprendre, progressivement, à organiser des ouvertures sélectives, comme on sait déjà le faire pour la recherche en santé. Bien sûr, il y aura des asymétries d'information entre les personnes ou les entreprises qui n'auront accès qu'aux données publiques ouvertes et les personnes qui détiendront des données personnelles. Mais ce n'est pas une raison pour briser toutes les barrières. Il faudra apprendre à discerner le moment où ces asymétries deviendront des monopoles de fait susceptibles d'entraver la concurrence normale. Les autorités de régulation de la concurrence et le législateur commencent à apprendre à ordonner le partage des données. Cela a commencé avec les obligations en cas de dégroupage, mais on a vu la même chose pour la géolocalisation des taxis, par exemple.

Pour une personne, partager des données sur sa santé, c'est un service rendu à tous, une marque d'altruisme. Jusqu'où peut-on aller sans enfreindre le droit de chacun à l'intimité de sa vie privée – son *right to privacy* ?

Je pense que chacun est libre de partager ses propres données s'il le souhaite. Simplement, celui qui traiterait ces données à caractère personnel devrait le signaler à la CNIL et avoir une autorisation. On entend souvent les protestations de certains « libertariens » qui contestent cette sécurité. Mais il y a bien des secteurs où l'Etat met des garde-fous pour empêcher certaines dérives. Si l'utilisation des données librement partagées était libérée, êtes-vous certain qu'il n'y aurait pas rapidement de subtiles formes de pression pour

pousser à cette ouverture ?

Etes-vous certain qu'il y aurait les sécurités permettant de faire machine arrière ?

Nous sommes entrés dans une nouvelle ère, où les données disponibles s'accroissent rapidement, où les méthodes de traitement changent aussi très vite, en même temps que les stratégies économiques et les valeurs et les aspirations sociales. Il est normal que la société prenne un peu de temps pour définir de nouveaux équilibres. D'un certain point de vue, nous avons connu la même situation, il y a une trentaine d'années, avec l'éthique biomédicale, et il a fallu près de quinze ans pour définir de nouveaux consensus, qui sont d'ailleurs régulièrement ajustés.

En mars dernier, l'Association des journalistes de l'information sociale s'opposait à la soumission à deux instances, et notamment à un comité d'experts, avant publication, de leur méthodologie et des résultats de leur enquête : comment s'est faite la conciliation entre l'ouverture des données publiques de santé et le respect de la vie privée ?

Ici, c'est plutôt un débat entre la protection de la vie privée et la liberté de la presse. Le projet de loi sur la santé est toujours devant le Parlement. Il semble qu'un amendement satisfaisant pour les journalistes ait été adopté. Nous verrons comment le législateur tranchera en dernière instance...

Y aura-t-il un jour, en France, un Tripadvisor des médecins ?

Peut-être, mais là encore, il y aura plusieurs libertés à concilier. Les médecins ont droit, eux aussi, à la protection de leurs droits fondamentaux, comme nous tous. Que diriez-vous d'un Tripadvisor des enseignants ? Ou des agriculteurs ?

Comment, dans le même ordre d'idées, éviter une nouvelle affaire Médiateur ? Serait-ce possible ?

L'analyse, en temps réel, des prescriptions et des surprescriptions, et l'utilisation de ces résultats pour améliorer la santé publique est un enjeu essentiel. La transparence peut y contribuer, mais c'est probablement l'organisation du système de soins dans son ensemble qui devra apprendre à adopter l'usage, en temps réel, de données massives et la construction de processus guidés par ces données. C'est évidemment un objectif qu'il est nécessaire d'atteindre.

Categorie

1. Articles & chroniques

date créée

22 septembre 2015

Auteur

henriverdier
