

Ouverture des données publiques : le site « data.gouv.fr » a des progrès à faire

par François Ecalte, économiste, Université Paris I

La mise à disposition, dans de bonnes conditions d'accès, d'un grand nombre de données produites par les administrations publiques est évidemment souhaitable, mais encore faut-il que ces données soient compréhensibles et présentent un minimum de garanties de fiabilité. Presaje.Com a demandé à François Ecalte, chargé d'un cours de politique économique à l'Université Paris 1, de tester le site data.gouv.fr. Supposé devenir la vitrine de l'open data en France, le site a des progrès à faire. Il ne convainc pas que les données rassemblées et publiées ont les propriétés de lisibilité et de fiabilité espérées.

Le site www.data.gouv.fr, placé sous la responsabilité de la « mission Etalab » du secrétariat général pour la modernisation de l'action publique (service du Premier ministre), apparaît d'abord comme un grand bazar de données où se trouvent aussi bien les statistiques démographiques nationales de l'INSEE que la localisation des stations de vélos à Toulouse.

Les données sont regroupées dans neuf grands thèmes (économie et emploi, culture, société, territoires et transports etc.) à l'intérieur desquels il n'y a pas de classement par sous-thèmes. Le thème « économie et emploi » par exemple rassemble ainsi des données en vrac sur le prix des carburants en France, les participations cotées de l'Etat, le chômage de longue durée, le budget de la ville de Montpellier, le bilan social du département de l'Oise, etc.

Beaucoup de producteurs de ces données publiques sont « certifiés » par la mission Etalab, ce qui semble seulement signifier, à la lecture des succinctes indications données sur les conditions de cette « certification », qu'il s'agit d'organismes chargés d'une mission de service public. Tous les services de l'Etat, ses établissements publics et les collectivités territoriales étant chargés d'une mission de service public, le processus de certification ne paraît donc pas très exigeant. Or une information n'est pas nécessairement fiable parce qu'elle provient d'un organisme public.

Label et certification

Rappelons qu'il existe en France un service statistique public, comprenant l'INSEE et les services statistiques ministériels, qui respecte des normes professionnelles, inscrites dans un code de bonnes pratiques prévu par un règlement européen, sous le contrôle de l'Autorité de la statistique publique (ASP). D'autres organismes, privés pour certains, produisent des statistiques publiques qui peuvent faire l'objet d'une labellisation, garantissant leur conformité au code de bonnes pratiques par l'ASP. La « certification » des données publiées sur le site data.gouv.fr est encore loin du contrôle du respect de ces « bonnes pratiques » par l'ASP et risque d'entraîner une confusion entre les « statistiques publiques » labellisées par l'Autorité de la statistique publique et les « données publiques » certifiées par Etalab.

Sur ce site de référence de l'Open Data public, j'ai regardé d'un peu plus près les données publiées dans la catégorie « économie et emploi » que je connais mieux.

J'y ai d'abord trouvé des liens hypertextes vers les sites de l'INSEE, des services statistiques ministériels ou d'Eurostat, ce qui n'appelle aucune observation si ce n'est qu'ils sont nettement moins bien présentés sur data.gouv.fr que sur les sites auxquels celui-ci renvoie.

J'ai également trouvé des données budgétaires en vrac dans des tableaux Excel, sans la moindre explication méthodologique. Personnellement, je sais à quoi correspondent les « AE autres titres du programme contribution au désendettement de l'Etat de la mission gestion du patrimoine immobilier » mais je ne suis pas sûr que beaucoup comprennent ce que cela signifie. Au mieux, ces informations sont inutiles ; au pire, elles seront mal utilisées, car non compréhensibles, par ceux qui ne sont pas initiés au droit budgétaire.

En outre, certaines de ces données sont anciennes (par exemple, des séries de dépenses budgétaires sur la période 2000-2009), alors même que des informations plus récentes existent.

Plus surprenant, j'ai découvert des données sur le nombre de demandeurs d'emplois en fin de mois en Corse (par sexe, catégorie, tranche d'âge etc.) produites par la collectivité territoriale de Corse (producteur « certifié »). Il s'agit d'un fichier Excel dont le contenu est difficilement lisible. La maigre notice explicative précise que la source est « Pôle emploi, Dares, Direccte - retraitement Pôle prospective et observation des mutations économiques du CARIF-OREF ». Il s'agit donc apparemment des statistiques de Pôle emploi et de la DARES, mais la mention du « retraitement Pôle prospective... » est particulièrement inquiétante car elle semble signifier que les statistiques de Pôle emploi et de la DARES, labellisées par l'Autorité de la statistique publique, ont été corrigées dans des conditions obscures par un non moins obscur CARIF-OREF.

Ma navigation sur le site data.gouv.fr a certes été très rapide et les pages que j'ai regardées ne sont sûrement pas représentatives, mais je conseille la plus grande prudence à ceux qui iront collecter des informations dans l'open data public. Il est évidemment souhaitable de rendre disponibles beaucoup plus de données publiques, mais pour être utiles, elles doivent être fiables et compréhensibles, ce qui suppose un minimum de traitements et de contrôles pour garantir cette fiabilité et faciliter cette compréhension. Or les données disponibles sur data.gouv.fr ne me semblent pas posséder ces propriétés de fiabilité et de lisibilité.

Le réseau internet est un extraordinaire instrument de diffusion des connaissances, mais aussi de propagation des erreurs. Le service public devrait donner l'exemple et faire plus d'efforts pour aider les internautes à faire le tri entre ces informations.

 [Télécharger le PDF de l'article](#)

<< [Retour au sommaire](#)

PRES@JE.COM

Une publication de l'Institut PRESAJE
(Prospective, Recherche et Etudes Sociétales Appliquées à la Justice et à l'Economie)
30 rue Claude Lorrain 75016 Paris
Tél. 01 46 51 12 21 - E-mail : contact@presaje.com - www.presaje.com
Directeur de la publication : Michel Rouger

Pour ne plus recevoir d'e-mails de la part de Presaje, [cliquez ici](#) >> [CONSULTER LES PRECEDENTS NUMEROS](#)