



CONCOURS INTERNE D'INGENIEUR EN CHEF TERRITORIAL

SESSION 2014

Analyse d'un document

Option : Systèmes d'information géographique (SIG), topographie

EPREUVE N° 16

Durée : 4 h

Coefficient : 4

SUJET : Ouverture des données géoréférencées (Open Data Géo).

Vous êtes nouvellement recruté(e) en qualité de responsable de la direction de l'information géographique (DIG) de la Communauté urbaine de « METROPOSUD », établissement public de coopération intercommunale (EPCI) de 18 communes représentant plus d'un million d'habitants.

La Communauté urbaine de « METROPOSUD » souhaite s'engager dans une logique d'ouverture de ses données publiques. Le directeur général des services (DGS) de la Communauté urbaine souligne le caractère stratégique de cette orientation, non seulement en termes de transparence de l'action publique mais également en termes de développement de services innovants sur les territoires et de dynamisme de la filière des industries numériques. Il s'agit donc d'une démarche essentielle pour l'institution.

En effet, tout d'abord, la loi du 17 juillet 1978 institue diverses mesures d'amélioration des relations entre l'administration et le public notamment en ce qui concerne la liberté d'accès aux documents de caractère non nominatif et la création d'une commission dite « Commission d'accès aux documents administratifs » (CADA) chargée de veiller au respect de la liberté d'accès aux documents administratifs.

Ensuite, la création de la mission Etalab (<http://data.gouv.fr>), celle d'une licence ouverte OL/LO et l'obligation pour les services de l'Etat de faire de l'open data, renforcent cette démarche de transparence.

Enfin, conformément aux orientations récentes prises par le gouvernement, il semble donc que l'acte III de la décentralisation comme le prévoit l'article 111 du projet pose aussi le principe d'obligation d'open data pour les collectivités territoriales.

Sensibilisé à ce contexte innovant, le DGS, informé des actions mises en place dans certaines villes de France ayant déjà privilégié cette démarche, souhaite amorcer ce processus en se limitant, dans un premier temps, à des données géographiques, notamment celles qui sont relatives à l'aménagement et à la connaissance du territoire, à l'environnement, au transport ...

Question 1

Analyse critique du dossier « Open Data »

Le DGS vous demande un argumentaire stratégique, sur la base des documents mis à votre disposition, afin de sensibiliser les élus communautaires à l'ouverture des données géoréférencées. Fort de votre expérience et expertise technique dans les SIG, vous aborderez le principe organisationnel et technique à mettre en œuvre.

Question 2

Mise en œuvre opérationnelle de l'ouverture des données géoréférencées:

Vous rédigerez un ensemble de recommandations et argumentaires thématiques propres à guider votre hiérarchie sur la politique de mise en œuvre de l'ouverture des données publiques (Open Data) à l'échelle de l'intercommunalité. Vous décrierez, à cet effet, les différentes étapes (techniques, organisationnelles, administratives, sociales....) dans sa mise en œuvre afin de promouvoir le rayonnement de l'institution et d'améliorer la démocratie participative.

Barème de notation :

- Compte rendu critique de l'analyse : 8 points
- Orientations prospectives : 10 points
- Lisibilité et cohérence du document : 2 points

DOCUMENTS JOINTS

Fiche n° 1	Eléments de contextes généraux	Page 4
Document n° 1	Feuille de route du gouvernement sur les données publiques – Mars 2013	Page 7
Document n° 2	Fiche 20 : Communication des documents administratifs - L'administration et la gestion.	Page 12

Document n° 3	Extrait du guide pratique de l'ouverture des données publiques territoriales – FING Fondation Internet Nouvelle Génération – Janvier 2011.	Page 16
Document n° 4	Les défis de l'ouverture des données publiques : entretien avec Henri Verdier - Le portail de la modernisation de l'action publique – Février 2013	Page 37
Document n° 5	Exemple de licence Etalab : Licence ouverte – Octobre 2011.	Page 39
Document n° 6	Décisions du CIMAP du 12 décembre 2013	Page 42

NOTA :

- 2 points seront retirés au total de la note sur 20 si la copie contient plus de 10 fautes d'orthographe ou de syntaxe.
- **Les candidats ne doivent porter aucun signe distinctif sur les copies** : pas de signature (signature à apposer uniquement dans le coin gommé de la copie à rabattre) ou nom, grade, même fictifs. Seuls la date du concours et le destinataire, (celui-ci est clairement identifié dans l'énoncé du sujet) sont à porter sur la copie.
- Les épreuves sont d'une durée limitée. Aucun brouillon ne sera accepté, la gestion du temps faisant partie intégrante des épreuves.
- Lorsque les renvois et annotations en bas d'une page ou à la fin d'un document ne sont pas joints au sujet, c'est qu'ils ne sont pas indispensables.

FICHE n° 1

Éléments de contextes généraux.

1° Le principe organisationnel au sein de l'institution :

La représentation géographique des informations traitées dans le système d'information de « METROPOSUD » est une nécessité pour la gestion quotidienne, l'aide à la décision ou la communication. A l'exception des logiciels fonctionnels, toutes les solutions logicielles mises en place ces dernières années intègrent une composante SIG. Aussi, il est nécessaire que « METROPOSUD » soit davantage réactif sur les dossiers SIG et en capacité d'apporter des solutions novatrices et ouvertes aux citoyens.

La gestion et la diffusion de l'information géographique sont des domaines totalement transversaux. L'offre SIG doit proposer des outils faciles d'accès et d'utilisation de manière à démocratiser son usage.

L'étude très récente d'un schéma directeur de l'information géographique (SDIG) confiée à un prestataire, spécialiste dans ce domaine, a démontré notamment la nécessaire évolution vers un SIG ouvert, facilitateur d'une accessibilité accrue pour les citoyens et dont l'open data serait le catalyseur.

La direction de l'information géographique de « METROPOSUD », sous votre autorité, est organisée comme suit :

Une assistance de direction

Missions : Chargée de la gestion des dossiers administratifs, du personnel et de la communication interne et externe.

Un Service « Administration des données géographiques ».

Missions : Le service est chargé de la modélisation des bases, de l'affectation des droits, de la gestion des échanges et de la régie de recettes, créée en 2007 pour la vente de données géographiques numériques (conforme à l'avis de la Commission d'Accès aux Documents Administratifs formulé à cette même date). Cette dernière activité sera très probablement revue compte tenu du projet « Open Data ».

Un Service « Géoservices et promotion ».

Missions : Le service est chargé des relations avec les partenaires internes et externes, de l'animation du réseau de référents SIG (soit environ 50 correspondants) et de la mise en œuvre de la directive européenne INSPIRE notamment en respectant les échéances sur la diffusion des données. Elle a été amenée à conduire la mise en œuvre du schéma directeur de l'information géographique communautaire afin d'orienter la politique d'acquisition et de diffusion des données, à court et moyen termes.

Un service « Topographie et 3D ».

Missions : Le service est chargé de l'acquisition des données topographiques à moyennes (1/1000) et grandes (1/200) échelles. Celle-ci s'effectue par le biais de marchés de topographie, de photogrammétrie et par la brigade topographique interne. Le prochain nouvel objectif est l'acquisition d'un scanner laser topographique afin de répondre au besoin émergent de modélisation 3D.

L'organisation de la Direction de l'Information Géographique de « METROPOSUD » repose également sur un réseau de référents SIG. En effet, chaque direction et communes membres de la Communauté urbaine ont désigné un référent SIG, sur les bases d'un profil élaboré par la direction. Il est chargé de recenser et promouvoir le SIG au sein de sa direction d'appartenance ou de sa commune. Il est, par ailleurs, le référent privilégié auprès de la DIG et régulièrement invité à des réunions de groupes de travail thématiques.

2^e La gouvernance du SIG :

La dynamique du SIG communautaire repose sur une instance de travail dénommée « échelon permanent SIG » essentiellement composée de l'ensemble des référents SIG, eux-mêmes faisant partie de groupes de travail thématiques.

Un comité de suivi SIG est également constitué, essentiellement d'experts techniques, administratifs, ressources humaines, juridiques... Il est animé par le Directeur qui impulse la stratégie de développement à travers l'échelon permanent SIG.

Enfin, un comité de pilotage SIG constitué de l'élu(e) délégué(e) aux Technologies de l'Information et de la Communication (TIC), du DGS, des DGA, de représentants des communes, d'experts et du Directeur de la DIG, arbitre les choix et orientations proposés par le comité de suivi.

3^e La politique d'acquisition de données géographiques :

Les données de références existantes :

Le Plan Cadastral Informatisé (PCI).

Un plan de ville.

Une orthophotographie aérienne.

Spécifications de l'orthophotographie :

Orthophotographie vraie avec recouvrements importants.

Résolution = 15 cm.

Précision planimétrique = 30 cm.

Précision altimétrique = 50 cm.

Réalisation d'un Modèle Numérique de Terrain (MNT) de 5 m de résolution.

6

**D'un Modèle Numérique de Surface (MNS) de 50 cm de résolution.
Et d'un Modèle Numérique d'Élévation (MNE) de 50 cm de résolution.**

Un référentiel des voies (voies et adresses).

Une base de données photogrammétrique à moyenne échelle (1/1000) couverte sur la totalité du territoire, (classe de précision de 10 cm) avec restitution des éléments principaux du sursol.

Une base de données topographique à grande échelle (1/200) existante sur les territoires à enjeux, de classe de précision 2cm (plan régulier au 1/200).

Les données d'intérêt commun :

- Le Plan Local d'Urbanisme (PLU).**
- Le Schéma de Cohérence Territorial (SCOT).**
- Les Zones d'Aménagement Concerté (ZAC).**
- Données environnementales (dont l'occupation du sol).**
- Transport en commun.**
- Réseau...**

Les données de l'agence d'urbanisme.

Toutes les études produites par l'agence d'urbanisme sont intégrées au sein du SIG communautaire.

Les données métiers :

- Données de gestion de voirie.**
- Données de gestion patrimoniale**
- Données de circulation, d'assainissement...**

En général, les données métiers sont sous la responsabilité des directions opérationnelles au sein desquelles elles sont produites. Néanmoins, elles peuvent constituer un intérêt de partage entre directions. Elles sont classifiées alors en données d'intérêt commun.

La feuille de route du Gouvernement en matière d'ouverture et de partage des données publiques

L'ouverture des données publiques, liberté publique et levier d'innovation

L'ouverture des données publiques est fondée sur la loi n° 78-753 du 17 juillet 1978 relative au droit d'accès aux documents administratifs, ainsi définis : "(...) quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, 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. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions et décisions. (...)" (article 1^{er}).

Le droit d'accès aux documents administratifs a été reconnu comme une "liberté publique" par le Conseil d'État (CE, 29 avril 2002, U., n° 228830). En 1997, le Gouvernement en a élargi le principe en décidant la mise en ligne gratuite des "données publiques essentielles". En 2003, la directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public, transposée par l'ordonnance du 6 juin 2005 et le décret du 30 décembre 2005, a facilité la réutilisation des documents existants détenus par des organismes du secteur public. Le décret n° 2011-577 du 26 mai 2011 relatif à la réutilisation des informations publiques détenues par l'État et ses établissements publics administratifs a posé le principe de la gratuité du droit à réutilisation.

Le 21 février 2011, le décret n° 2011-194 portant création d'une mission "Etalab" chargée de la création d'un portail unique interministériel des données publiques créait la structure en charge d'animer et d'héberger la diffusion de ces données publiques.

Avec la révolution numérique – qui facilite la production, la diffusion et le traitement de données de masse – l'ouverture et le partage des données publiques sont devenus un puissant levier pour :

- renforcer la confiance des citoyens grâce à une plus grande transparence de l'action publique (par exemple, avec la mise en ligne d'informations sur les comptes publics, la pollution ou les délits constatés) ;
- permettre de nouvelles formes de co-production avec la société et soutenir l'innovation sociale (comme le projet Handimap qui a permis, grâce aux données des villes de Rennes et de Montpellier, de développer une application de calcul d'itinéraire pour handicapés moteurs) ;
- améliorer le fonctionnement administratif (comme en témoigne la très forte utilisation des données publiques par l'administration elle-même) ;
- améliorer l'efficacité de l'action publique en développant de nouveaux modes d'organisation et de nouveaux processus de travail (comme le suivi des accidents de la route par la sécurité routière permettant d'améliorer l'aménagement de la voirie) ;
- soutenir le dynamisme économique, en créant de nouvelles ressources pour l'innovation et la croissance (comme en témoignent les jeunes entreprises innovantes primées par Dataconnexions) ;
- contribuer à l'influence et au rayonnement international de la France (initiative *Open Government Partnership* – OGP –, initiative sur la transparence de l'aide internationale – IATI – ou sur la transparence des industries extractives – EITI).

L'engagement du Gouvernement en faveur de la transparence de l'action publique et de l'ouverture des données

Dès le premier Conseil des ministres, le 17 mai 2012, le Président de la République a fait signer, à tous les membres du Gouvernement, une charte de déontologie rappelant notamment leur "devoir de transparence", leur "scrupuleux respect des dispositions garantissant l'accès des citoyens aux documents administratifs" et leur engagement à "mener une action déterminée pour la mise à disposition gratuite et commode sur Internet d'un grand nombre de données publiques".

En lançant le chantier de la modernisation de l'action publique, le Premier ministre a réaffirmé le 31 octobre 2012 les principes d'ouverture et de partage des données publiques prévus par la circulaire du Premier ministre du 26 mai 2011, notamment le droit pour tout citoyen de pouvoir réutiliser librement et gratuitement les données publiques, ou l'encadrement strict des redevances de réutilisation de données publiques.

En intégrant la mission Etalab au SGMAP, il a constitué une force globale d'innovation, associant la conception des systèmes d'information de l'État, l'ouverture des données publiques, l'évaluation des politiques publiques et l'innovation en matière de politiques publiques.

Dès le premier comité interministériel pour la modernisation de l'action publique (CIMAP), le 18 décembre 2012, le Gouvernement a pris trois décisions importantes pour l'ouverture des données publiques :

- » "Réaffirmer le principe de gratuité de la réutilisation des données publiques et l'étendre, en concertation, aux collectivités, aux délégataires de service public, à la sphère sociale et aux autorités administratives indépendantes" ;
- » "Lancer, dès 2013, des travaux sur la transparence de l'action publique et des institutions, notamment en ouvrant le débat sur la mise à disposition de données en matière d'éducation, de risques environnementaux, de transparence du système de santé, de l'offre de transport, de logement, de lisibilité des prestations sociales et des dépenses publiques" ;
- » "Confier au SGMAP, en lien avec le ministère du budget, la mission d'évaluer les modèles économiques de chaque redevance existante, notamment en auditant les coûts et les recettes associés. Le SGMAP doit remettre ses propositions au Premier ministre au printemps 2013".

L'ouverture des données publiques est engagée. Outre les informations partagées (mais pas toujours réutilisables) par le service public de l'accès au droit (Légifrance), le cadastre ou encore Géoportail, de nombreuses bases de données sont mises en libre accès par les administrations, les établissements publics et les collectivités locales.

La mission Etalab a développé une plateforme nationale d'hébergement et de partage de données publiques ouvertes, gratuites et réutilisables (www.data.gouv.fr). Elle coordonne un réseau de correspondants Open data et de fournisseurs de données qui ont permis la mise en ligne de plus de 350 000 jeux de données. La mission Etalab a également travaillé à structurer l'écosystème des réutilisateurs de données (startups, associations,...), notamment par le biais de l'organisation du concours Dataconnexions, qui lui a permis, en deux éditions, de récompenser 11 projets particulièrement prometteurs.

L'ambition du Gouvernement en matière d'ouverture des données publiques

Le Gouvernement entend donner sa pleine mesure à cette politique, et en faire un levier de modernisation de l'action publique et de soutien au dynamisme économique.

1. Travailler à l'ouverture de jeux de données stratégiques

Le Gouvernement souhaite favoriser l'ouverture de données stratégiques et de qualité, qu'elles permettent de créer des services utiles au quotidien (comme pour le transport, l'environnement, la santé, la consommation énergétique, le tourisme...), qu'elles soient réclamées par les citoyens en raison de leur caractère démocratique ou pratique (budgets et dépenses détaillées par administration...), ou qu'elles répondent à une exigence légale ou réglementaire (publication des attributaires de marchés publics conformément à l'article 138 du code des marchés publics).

Pour ce faire, la mission Etalab du SGMAP organisera, dès 2013, six débats thématiques et ouverts, afin d'identifier et de publier de nouveaux jeux de données. Ces débats seront confiés à une commission composée par le SGMAP et le ministère concerné, pouvant bénéficier d'un rapporteur ou d'un secrétariat assuré par Etalab ; ils donneront lieu à un rapport, remis au Premier ministre et au ministre intéressé, qui formuleront des recommandations en termes d'ouverture cohérente des données et préciseront les bénéfices attendus.

Dans ce cadre, le Gouvernement étudiera les possibilités :

- » d'élargissement de la stratégie d'ouverture des données publiques à tous les établissements publics administratifs, aux délégataires de service public, aux collectivités territoriales et aux autorités administratives indépendantes ;
- » d'extension de la politique d'ouverture des données publiques aux missions de service public à caractère industriel et commercial, aux entreprises publiques et à la sphère associative.

Le Premier ministre a par ailleurs demandé à tous les ministres de préciser leur stratégie d'ouverture des données publiques dans les programmes ministériels de modernisation et de simplification (PMMS), à transmettre d'ici la mi-mars.

2. Faciliter et améliorer le processus d'ouverture des données publiques dans les ministères, les établissements publics, et les collectivités territoriales

Etalab ouvrira, dès 2013, une nouvelle version de la plateforme www.data.gouv.fr. Tout en conservant sa vocation de diffuser largement les données brutes, celle-ci deviendra à terme une vitrine grand public de l'ouverture et du partage des données publiques. Elle proposera en particulier des outils techniques utiles pour s'approprier ces données (outils de visualisation de données, de traitement statistique, etc.).

Une attention toute particulière sera apportée à l'indexation des données, à leur traitement sémantique et aux liens entre ces données. Etalab, en collaboration avec les collectivités locales et les autres producteurs de données publiques, travaillera à répertorier et à présenter l'ensemble des données publiques librement accessibles et gratuitement réutilisables, quelle qu'en soit la source. De même, Etalab poursuivra avec les fournisseurs de données la réflexion sur l'importance des formats ouverts et la possibilité de les faire converger, et s'attachera à construire une plus grande interopérabilité entre la plateforme www.data.gouv.fr et les systèmes d'information de l'État, des collectivités territoriales ou des institutions européennes.

Par ailleurs, Etalab lancera également dès 2013 le chantier de la diffusion de données par des interfaces programmables ("API").

3. Accompagner les réutilisations innovantes et la communauté des réutilisateurs

L'État encourage la réutilisation la plus large des données publiques afin que cette politique d'ouverture soutienne une stratégie d'innovation, y compris sociale, et de croissance économique.

» Le service public est le premier utilisateur de données publiques de qualité

Les données publiques sont produites dans le cadre d'une mission de service public et le service public est fréquemment le premier bénéficiaire de leur ouverture. La production, la collecte et la publication de données par l'administration doivent répondre aux besoins d'un service public de qualité. Etalab doit donc être à la disposition des administrations pour leur permettre de saisir ces opportunités. En particulier, elle organisera des sessions de formation, pourra monter des ateliers *ad-hoc*, pour analyser le potentiel de certaines applications de l'ouverture des données publiques (lutte contre le chômage, par exemple), et publiera des notes d'analyses de cas exemplaires.

» L'État favorise la réutilisation des données publiques par les entreprises pour créer des services innovants

Etalab continuera à soutenir les entreprises innovantes, l'économie sociale et solidaire ainsi que les grands groupes industriels, dans leurs projets de réutilisation de données publiques. Elle organisera en 2013 trois concours Dataconnexions, permettant de repérer et de faire connaître des réutilisations exemplaires, et soutiendra l'organisation d'Assises nationales de l'open data, en lien avec le ministère chargé du redressement productif, le ministère chargé des PME et le ministère chargé de l'économie sociale et solidaire.

» L'État soutient l'engagement des citoyens et des associations autour des données publiques au profit d'une action publique moderne et renouvelée

Le soutien à l'engagement des citoyens, pour coproduire certaines informations, concevoir de nouveaux services ou aider les acteurs publics à relever certains défis, fait partie intégrante des missions d'Etalab, qui organisera en particulier des concours mobilisant les citoyens autour des données et des objectifs partagés par l'État.

» Suivre et encourager la recherche en sciences de la donnée

Les "Big Data", la transformation des chaînes de valeur, les stratégies et les organisations fondées sur la "culture de la donnée" caractérisent les formes actuelles de la révolution numérique. Pour suivre cette transformation rapide, Etalab accentuera sa coopération avec le monde de l'enseignement supérieur et de la recherche et se dotera, avant la fin de l'année 2013, d'un collège d'experts rassemblant des experts technologiques, des spécialistes de datasciences, des juristes, des économistes ou des réutilisateurs (entrepreneurs ou responsables associatifs), à moyens constants.

4. La transparence et la culture de la donnée au service de la modernisation de l'action publique

L'ouverture des données publiques est un outil majeur au service de la transformation du service public, pour en améliorer la qualité et en permettre l'évaluation par tous ses usagers. Le SGMAP veillera à construire, chaque fois que cela sera pertinent, des référentiels quantitatifs, si possible en temps réel, pour les évaluations des politiques publiques, afin qu'elles permettent de construire des outils utiles au pilotage de ces politiques publiques. Comme prévu par le CIMAP du 18 décembre 2012, le SGMAP travaillera à la conception d'innovations dans les services publics en ayant recours aux données publiques.

5. L'évaluation des redevances existantes

À l'occasion du CIMAP du 18 décembre 2012, le Premier ministre a confié au SGMAP, en lien avec le ministère du Budget, une mission d'évaluation des modèles économiques des redevances existantes, notamment en auditant les coûts et les recettes associés. Ces recommandations seront remises au premier semestre 2013.

Etalab et le SGMAP accompagneront les administrations qui commercialisent de l'information publique dans l'évolution de leurs modèles économiques. En effet, le respect du principe d'accès libre et gratuit du citoyen à des données publiques réutilisables n'empêche pas de concevoir des stratégies de monétisation de services à haute valeur ajoutée ou de création de plateformes.

6. Évolutions du cadre juridique et administratif

Un *vade-mecum* rappelant le cadre juridique de la mise à disposition des données publiques sera diffusé aux ministères, afin de faciliter leur engagement dans la démarche d'ouverture de leurs données.

À l'occasion, notamment, de la révision de la directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public, le Gouvernement lancera des travaux de concertation, visant à actualiser si nécessaire les définitions légales encadrant le droit d'accès et de réutilisation des données publiques.

7. L'action internationale de la France

Enfin, le Gouvernement attache une importance toute particulière au rayonnement de la France sur le sujet de la transparence, de l'ouverture des données publiques ("Open Data") et de la gouvernance ouverte ("Open Government"). Etalab prendra part au pilotage des projets européens de plateformes de données ouvertes communautaires (notamment le site ouvert au début de l'année 2013 par la Commission européenne (<http://open-data.europa.eu/open-data/>).

La France étudie par ailleurs l'opportunité de rejoindre l'*Open Government Partnership* ("OGP"), pour engager une coopération fructueuse avec les autres gouvernements engagés dans cette voie. Elle se rapprochera du partenariat pour examiner les conditions de son adhésion. La France s'engagera par ailleurs largement dans la coopération multilatérale autour des approches de transparence et de gouvernement ouvert.

COMMUNICATION DES DOCUMENTS ADMINISTRATIFS

I. DOCUMENTS ADMINISTRATIFS COMMUNICABLES

- 1 - DOCUMENTS NON NOMINATIFS
- 2 - DOCUMENTS NOMINATIFS
- 3 - DISPOSITIONS COMMUNES

II. EXERCICE DU DROIT D'ACCES

- 1 - DISPOSITIONS COMMUNES
- 2 - DEMANDE DE COMMUNICATION

III. REFUS DE COMMUNICATION

13

La loi n° 78-753 du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal, garantit le droit de toute personne à l'information en ce qui concerne la liberté d'accès aux documents de caractère non nominatif. Cette loi a ouvert à toute personne le droit d'obtenir communication des documents administratifs qui émanent "des administrations de l'État, des collectivités territoriales, des établissements publics et des organismes, fussent-ils de droit privé, chargés de la gestion d'un service public" (art. 1er et 2). Les EPLE entrent naturellement dans le champ d'application de ce texte.

Toutefois les règles que pose ce texte et qui sont développées ci-après, ne sont applicables qu'en l'absence de textes spécifiques. Ainsi, la communication du dossier d'un agent ou d'un élève, dans le cadre d'une procédure disciplinaire, est régie par les dispositions réglementaires et les principes jurisprudentiels propres à cette procédure.

La loi du 17 juillet 1978 a créé une commission dite "Commission d'accès aux documents administratifs" (C.A.D.A.). Celle-ci est chargée de veiller au respect de la liberté d'accès aux documents administratifs.

I. LES DOCUMENTS COMMUNICABLES

1 - LES DOCUMENTS NON NOMINATIFS

1. La loi du 17 juillet 1978 considère comme documents administratifs "tous dossiers, rapports, études, compte rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles qui comportent une interprétation du droit positif ou une description des procédures administratives, avis, à l'exception des avis du Conseil d'État et des tribunaux administratifs, prévisions et décisions revêtant la forme d'écrits, enregistrements sonores ou visuels, de traitements automatisés d'informations non nominatives" (art. 1er). Ces documents sont donc communicables de plein droit aux personnes qui en font la demande (art. 2).

- La notion de documents administratifs est plus large que celle d'actes administratifs, le champ d'application de la loi du 17 juillet 1978 s'étendant à des documents qui n'ont pas le caractère de décision. Elle exclut cependant les actes de droit privé (actes notariés, contrats commerciaux etc.).

2 - LES DOCUMENTS NOMINATIFS

2. Le législateur n'a pas défini la notion d'informations nominatives. Cependant la CADA considère comme nominatif tout document portant une appréciation ou un jugement de valeur sur une personne physique nommément désignée ou facilement identifiable ou incluant la description du comportement d'une personne dès lors qu'il s'avère que, d'une manière ou d'une autre, la divulgation de ce comportement pourrait lui porter préjudice. D'une manière plus générale, doit être considéré comme nominatif tout document dont les mentions ne peuvent être communiquées sans qu'une atteinte soit portée au droit qu'a chaque individu de ne pas divulguer à autrui des renseignements sur sa personne, ses activités ou ses biens.

A titre d'exemples sont considérés comme nominatifs, les notes attribuées aux élèves ou aux candidats d'un concours, le dossier scolaire d'un élève, une copie remise à l'occasion des épreuves du baccalauréat, la notation d'un agent public.

En revanche, les listes de noms (listes de membres du conseil d'administration, listes d'agents ou d'élèves) ne sont pas considérées comme des documents nominatifs à condition de ne comporter aucune mention personnelle relative aux intéressés (date de naissance, n° INSEE, adresse, téléphone, situation familiale... etc.). Ces éléments de la vie privée sont protégés par le Code civil (art. 9) et la Convention européenne des droits de l'homme et des libertés fondamentales (art. 8). En revanche, rien ne s'oppose à ce que les coordonnées communiquées soient celles de l'établissement scolaire.

3 - DISPOSITIONS COMMUNES

3. L'article 6 de la loi du 17 juillet 1978 permet, néanmoins, à l'autorité administrative de refuser la communication d'un document administratif lorsque celle-ci porterait atteinte aux secrets protégés par la loi et notamment aux secrets des délibérations du Gouvernement et des autorités responsables relevant du pouvoir exécutif, de la défense nationale et de la politique extérieure, de la monnaie et du crédit public, de la sûreté de l'État et de la sécurité publique, de la vie privée et en matière commerciale et industrielle, ainsi qu'au déroulement d'une procédure engagée devant une juridiction ou à la recherche des infractions fiscales et douanières.

II. EXERCICE DU DROIT D'ACCES

1 - DISPOSITIONS COMMUNES

4. Le droit à communication ne s'applique qu'à des documents achevés et non aux états préparatoires, partiels ou provisoires d'un document tant qu'il est en

FICHE N° 20

cours d'élaboration ⁽¹⁾. De plus, le droit d'accès ouvert par la loi du 17 juillet 1978 nécessite que le demandeur soit en mesure d'identifier précisément le document dont il souhaite obtenir communication, son droit d'accès ne pouvant concerner qu'un document existant dans l'état dans lequel il se trouve. En aucune façon, le législateur n'a entendu, à cette occasion, imposer à l'administration de procéder à des recherches en vue de fournir au demandeur une documentation, ni même d'effectuer un travail de reconstitution ou de compilation d'informations.

2 - DOCUMENTS NON NOMINATIFS

5. Toute personne, physique ou morale, peut demander la communication d'un document administratif non nominatif. Elle n'est tenue ni de motiver sa demande, ni de justifier l'intérêt qu'elle porte à ce document ou l'usage qu'elle entend en faire.

En outre, l'article 3 de la loi du 17 juillet 1978 consacre le droit de toute personne "de connaître les informations contenues dans un document administratif dont les conclusions lui sont opposées".

3 - DOCUMENTS NOMINATIFS

6. Le droit d'accès aux documents à caractère nominatif est réservé à la personne concernée par ces documents.

- S'agissant des documents nominatifs intéressant les enfants mineurs, seules la ou les personnes qui sont les représentants légaux de l'enfant et qui justifient être titulaire de l'autorité parentale peuvent recevoir communication de ces documents. ⁽²⁾

- Les documents qui comportent des mentions à caractère nominatif peuvent être communiqués à des tiers à condition que ces mentions soient divisibles et qu'elles puissent être occultées sans que l'administration ne soit contrainte d'établir un document distinct de l'original.

Ainsi, des lettres adressées par des parents d'élèves à un chef d'établissement comportant des appréciations sur les conditions dans lesquelles un enseignant exerçait ses fonctions et sur son comportement à l'égard des élèves peuvent être communiquées à celui-ci, après suppression des noms des parents et des éléments pouvant permettre l'identification des élèves ⁽³⁾

- Le caractère personnel de la communication des documents nominatifs ne s'oppose pas à ce que la

personne concernée donne expressément mandat à un tiers pour recevoir communication, pour son compte, du document concerné ou qu'elle soit accompagnée d'un tiers lors de cette consultation.

- En ce qui concerne les informations à caractère médical elles ne peuvent être communiquées à l'intéressé que par l'intermédiaire d'un médecin qu'il désigne à cet effet. Ainsi, l'administré qui demande la communication d'un document à caractère médical le concernant doit être invité à indiquer les coordonnées du médecin de son choix. Le document sera communiqué à ce médecin à qui il appartiendra, dans le respect des règles de la déontologie médicale, d'en divulguer la teneur à l'intéressé. En l'absence d'indication par l'intéressé des coordonnées d'un médecin, l'administration est tenue de refuser la communication du document.

4 - MODALITES DU DROIT D'ACCES

7. La demande de communication doit être adressée à l'autorité administrative qui détient le document soit qu'elle en est l'auteur, soit que le document a été établi pour son compte et lui a été transmis. Si un EPLE est saisi d'une demande tendant à la communication d'un document qui émane d'une autre administration (service académique, commune, collectivité territoriale de rattachement ...), il lui appartient de transmettre cette demande à l'administration compétente et d'en aviser l'intéressé.

"L'accès aux documents administratifs s'exerce :

a) par consultation gratuite sur place, sauf si la préservation du document ne le permet pas ou n'en permet pas la reproduction ;

b) sous réserve que la reproduction ne nuise pas à la conservation du document, par délivrance de copies en un seul exemplaire, aux frais de la personne qui les sollicite, et sans que ces frais puissent excéder le coût réel des charges de fonctionnement créées par l'application du présent titre" ⁽⁴⁾.

- Un arrêté du ministre du budget du 29 mai 1980 a fixé à un franc par page les frais de photocopie d'un document dans les services et les établissements publics de l'État. Les chefs d'établissement sont habilités à instituer des régies de recettes notamment pour l'encaissement des remboursements de services rendus, parmi lesquels figurent les frais de photocopie (arrêté interministériel du 11 octobre 1993).

(1) CE, 11 février 1983 *Ministre de l'urbanisme et du logement c/ Association Atelier libre d'urbanisme de la Région Lyonnaise*. C'est le cas des tableaux de répartition des moyens par discipline (TRMD) dont le caractère préparatoire implique qu'ils doivent être regardés comme inséparables et indissolublement liés à une décision à intervenir et ne sont donc communicables que lorsque la décision finale est prise.

(2) CE - 31.07.1996 - M. ZANONE.

(3) CE - 14.10.1992 - *Ministre de l'éducation nationale c/ PONTIUS*.

(4) arrêté interministériel du 11 octobre 1993.

FICHE N°20

- Le choix du mode d'accès est laissé au demandeur. Mais la communication du document ne doit pas perturber le fonctionnement de l'administration (1). Lorsque les documents dont la communication est demandée représentent un volume important, la communication des pièces peut-être étalée dans le temps. Dans cette hypothèse, la communication peut être limitée à une consultation sur place suivie de la délivrance de copies pour les seuls éléments sélectionnés par le demandeur (2).

de saisir, sans formalisme, la Commission d'accès aux documents administratifs (3) dans un délai de deux mois à compter du refus, exprès ou implicite, qui lui a été opposé. Cette commission émet dans le délai d'un mois un avis qu'elle adresse à l'autorité compétente sur le caractère communicable ou non du document demandé. L'autorité administrative doit alors faire connaître dans le délai d'un mois à la Commission la suite qu'elle entend donner à la demande. Si elle ne donne pas satisfaction à l'intéressé dans un délai de deux mois à compter de la saisine de la CADA, elle est réputée maintenir son refus. L'intéressé pourra alors, le cas échéant, saisir le tribunal administratif du litige relatif à la communication de ce document.

III. REFUS DE COMMUNICATION

8. Faute d'avoir satisfait dans un délai d'un mois à la demande de communication d'un document, l'autorité administrative est réputée avoir rejeté cette demande. Il appartient alors à l'administré intéressé

- Le mauvais vouloir de l'administration à communiquer un document administratif est susceptible d'engager sa responsabilité. Dans le cas d'un EPLE, c'est la responsabilité de l'établissement lui-même qui serait engagée.

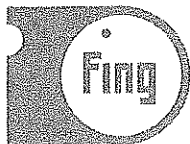
TEXTES DE REFERENCE

- Loi n° 78-753 du 17 juillet 1978 modifiée portant diverses dispositions d'amélioration des relations entre l'administration et le public (RLR n° 104-9) ;
- Décret n° 88-465 du 28 avril 1988 relatif à la procédure d'accès aux documents administratifs (RLR n° 104-9) ;
- Arrêté du ministre du budget du 29 mai 1980 (RLR n° 160-3) ;
- Arrêté interministériel du 11 octobre 1993 relatif à l'habilitation les chefs d'EPLÉ à instituer des régies de recettes et des régies d'avances (RLR n° 364-6).

INDEX ALPHABETIQUE

<p>Actes de droit privé : 1. CADA : 2, 8. Consultation : 7. Délais : 8. Document : - nominatif : 2, 6. - non nominatif : 1, 5. - inachevé : 4 ; - inexistant : 4 ;</p>	<p>- médical : 6. Dossier d'un élève : 2. Enfants mineurs : 6. Mandat : 6. Photocopie : 7. Recours : 8. Reproduction : 7. Responsabilité : 8. Tiers : 6.</p>
--	--

(1) La Commission apprécie en fonction de la taille et des moyens de l'administration concernée. Mais elle recommande en tout cas aux administrations d'offrir une alternative négociée au demandeur auquel la plage horaire ne convient pas. La consultation des documents doit avoir lieu dans les locaux du service qui en assure la conservation, sans que l'administration soit tenue de transférer des documents dans les locaux d'un autre service pour les besoins de cette consultation (CE, 26 octobre 1988, Lalande)
(2) En effet, la CADA considère qu'il appartient à l'administration d'établir par tous moyens la preuve matérielle de la communication.
(3) Commission d'accès aux documents administratifs (C.A.D.A.), 64 rue de Varenne - 75700 PARIS
Tél. : 01.42.75.79.99 - Fax : 01.42.75.80.70



GUIDE PRATIQUE DE L'OUVERTURE DES DONNEES PUBLIQUES TERRITORIALES

Guide à l'usage des territoires
et de leurs partenaires

VERSION DE TRAVAIL N° 1
Janvier 2011

Note : ce guide pratique se fixe pour objectif de faciliter l'engagement des collectivités territoriales dans une démarche d'ouverture de leurs données publiques. Il se fonde sur une observation des expériences françaises et européennes, et notamment sur les démarches des Communautés urbaines de Bordeaux et Rennes. Cette version doit être considérée comme une version de travail, encore imparfaite. Elle sera enrichie et améliorée tout au long de l'année 2011.

Nous accueillons avec plaisir vos commentaires, remarques et contributions par email ou bien en ligne à l'adresse : <http://fing.org/vademecum2011>

Amandine BRUGIÈRE, abrugiere@fing.org
Charles NÉPOTE, charles.nepote@fing.org

Table des matières

Introduction – Contexte 3

 Qu'est-ce que l'ouverture des données publiques? 3

 Pourquoi ce guide pratique? 4

 Pourquoi y a-t-il un processus d'ouverture des données publiques aujourd'hui? 5

1. Ouvrir et utiliser des données publiques territoriales 10

 1.1. Une donnée publique, c'est quoi au juste? 10

 1.2. Qu'attend-on de l'ouverture et la réutilisation des données publiques? 12

 1.3. Un nouvel écosystème d'acteurs interdépendants 27

 1.4. Un cadre juridique très favorable 34

2. Comment s'y prendre? 36

 2.1. Engager une politique de réutilisation des données publiques 36

 2.2. Définir sa politique d'ouverture et de diffusion des informations publiques : les grands choix à faire 39

 2.3. Comment rendre les données accessibles? 47

Annexes 55

 A1. "L'ouverture des données et après..." 55

 A2. Petit lexique de l'OpenData 61

 A3. Répertoire des acteurs, organismes et institutions 63

 A4. Bibliographie – Webographie 66

2. COMMENT S'Y PRENDRE ?

2.1. ENGAGER UNE POLITIQUE DE REUTILISATION DES DONNEES PUBLIQUES

2.1.1. Une démarche à la fois volontariste, pragmatique et ouverte

La réutilisation des données est un droit qui s'impose aux acteurs publics, mais ceux-ci ont tout intérêt à engager une démarche volontariste, sans attendre. Il s'agit en effet de maîtriser le cheminement plutôt que le subir : identifier les priorités plutôt que de se les faire dicter ; expérimenter pour prendre les bonnes décisions au lieu de livrer sous la contrainte ; industrialiser la production des données plutôt que risquer de se voir dépassé par la demande.

La démarche gagne à être pragmatique. Tout d'abord, il s'agit d'un sujet nouveau et prospectif. On aura donc tout intérêt à avancer progressivement. De petits réajustements valent mieux que de grands changements de cap. L'acteur public a tout intérêt à démarrer dans un mode expérimental, avec des cycles courts, qui permettent notamment de prendre le temps de pérenniser ce qui marche et de réfléchir à ce qui ne marche pas. Une démarche pragmatique permet ensuite de réduire un champ qui est potentiellement immense, en ne traitant au départ que ce qui offre les meilleures chances de résultats. Elle démontre la volonté de l'acteur public et permet d'anticiper des logiques de confrontations : tout le monde peut admettre qu'une institution qui s'engage sans réticence ne soit pas non plus en mesure de tout mettre à disposition tout de suite. Le pragmatisme est la base de tous les exemples réussis actuellement, notamment dans les pays anglo-saxons.

Dernier point, mais non des moindres, l'ouverture des données publiques relève d'un processus d'innovation ouverte : il s'agit de donner libre cours à l'imagination des acteurs économiques et sociaux, pour qu'ils créent de nouveaux services, de nouvelles connaissances, de nouvelles interprétations. Dans ce contexte, la diffusion des données publiques ne fait que la moitié du chemin et c'est seulement la réutilisation de ces données qui crée de la valeur. L'acteur public ne peut pas, légalement, choisir lui-même qui pourra réutiliser quelles données, à quelles fins. En revanche, il est important, d'un côté, de stimuler la demande de données pour que l'ouverture produise effectivement des résultats ; et de l'autre, d'orienter autant que faire se peut les réutilisateurs vers des usages socialement et économiquement féconds. Le dialogue et la collaboration entre les acteurs publics et les réutilisateurs sont donc des conditions premières de réussite.

Ces processus d'innovation ouverte, s'ils sont bien connus dans l'industrie³⁶ – notamment dans celle du numérique – sont moins bien connus dans les administrations, quand ils ne sont pas même tout simplement vécus comme dérangeants. Pourtant, dans pareil processus il n'y a pas d'autre option que de savoir lâcher prise, d'admettre que les résultats ne soient pas donnés à l'avance, d'accepter d'être surpris ! Cela ne veut pas dire que l'acteur public doive se résigner à ne rien maîtriser du tout : il a un vrai rôle dans l'organisation de "l'écosystème" qui s'organise autour des données publiques³⁷.

³⁶ Voir par exemple <http://www.internetactu.net/2010/09/23/eric-von-hippel-il-y-a-2-a-3-fois-plus-dinnovations-de-la-part-des-consommateurs-qui-l-ny-en-a-dans-lindustrie/> et <http://www.internetactu.net/2009/06/26/les-innovations-ouvertes-sont-elles-compatibles-avec-les-systemes-dinformation/>

³⁷ Voir à ce titre <http://www.internetactu.net/2010/08/30/concevoir-pour-perdre-le-controle/> ainsi que l'ouvrage : <http://www.internetactu.net/2010/06/01/pour-un-design-de-politiques-publiques/>

2.1.2. Les grandes étapes de la démarche

Selon la nature de l'acteur public (Etat, collectivité locale, établissement public...), son activité, ses moyens, son système d'information, l'ouverture des données publiques prendra des formes très différentes. On peut cependant, pour tous, distinguer trois grandes étapes dans cette démarche de création de valeur à travers la réutilisation des données publiques.

► Etape 1 : Esquisser les enjeux, repérer les potentiels de création de valeur

La première étape s'attache à un travail de compréhension du sujet, d'écoute et de dialogue avec les réutilisateurs potentiels, notamment pour évaluer avec eux les potentiels de création de valeur. Elle consiste en un repérage fin des acteurs clé et, en particulier, des premiers réutilisateurs potentiels : qui sont ces gens qui pourraient vouloir de vos données ?

Une démarche simple peut consister à faire le tour des "têtes de réseaux" de votre territoire ou de votre domaine d'action afin de comprendre les acteurs et les problématiques qui pourraient être impactés par la diffusion des données publiques. Ne vous en tenez pas aux acteurs les plus institutionnels, qui ne sont pas toujours les plus innovants : pensez aux réseaux d'entrepreneurs, aux chercheurs, aux associations dynamiques, aux communautés de développeurs informatiques "citoyens", aux artistes...

Ainsi repérés, un bon moyen pour démarrer le dialogue avec les réutilisateurs potentiels consiste à effectuer avec eux le recensement des données dignes d'intérêt. Ce recensement est guidé par le pragmatisme, on regardera plus spécifiquement, au moins pour un début, aux données qui déclenchent des réflexes créatifs : "Ah si j'avais ces données là, je pourrais...". Il permet également, toujours avec les réutilisateurs potentiels, de prioriser ce qui pourra être ouvert, en tenant compte :

- des contraintes de l'acteur public : d'ordre technique, financier, juridique, etc.,
- des motivations et des contraintes des réutilisateurs potentiels,
- et enfin, des autres sources possibles de données ouvertes (autres sources publiques, données privées, données coproduites par les utilisateurs...).

Cette étape conduit aussi à s'interroger sur les partenariats utiles à envisager pour la suite des opérations. Dans le cas de données de transports, par exemple, une administration unique est rarement la seule détentrice de toutes les données utiles. Des partenariats croisés avec d'autres administrations, avec des opérateurs privés, peuvent donc être précieux, si ce n'est indispensables !

► Etape 2 : Ouvrir des premiers jeux de données

La seconde étape consiste à ouvrir des données, à livrer progressivement mais régulièrement et aussi à le faire savoir. Nous détaillons cette étape, relativement "technique", dans la partie 2.3, p. 47.

Le premier "lâcher de données"³⁸ peut être limité et simplifié, ce qui ne signifie pas que les conditions de base que nous décrivons plus bas (documentation des données, formats ouverts ou au moins communs, licences ou conditions d'usage claires et simples) ne s'applique pas : si le but est d'amorcer une dynamique, il vaut mieux abaisser les barrières...

Il est plus utile de s'intéresser à un petit nombre de données en profondeur plutôt qu'à un nombre très large, mais en surface. Traiter à fond un petit nombre de données permet de se confronter à toutes les problématiques juridiques et techniques de la publication avant d'appliquer l'expérience acquise à d'autres jeux de données. Il ne s'agit pas pour non plus de passer plusieurs mois à vouloir documenter chaque information, à produire la meilleure API possible, etc. Même si elles ne sont pas parfaites, des livraisons progressives mais régulières assurent un *feedback* régulier des réutilisateurs et permettent de corriger rapidement des problèmes qui coûteront autrement plus cher s'ils sont détectés en fin de process.

³⁸ Pour reprendre la jolie expression inventée par l'association Regards Citoyens

► **Etape 3 : Accompagner : stimuler, valoriser, aider, dialoguer**

La dernière étape est au moins aussi importante que les précédentes. Ouvrir des jeux de données ne suffit pas : sans un minimum d'accompagnement, cela ne produit pas grand chose³⁹ et l'acteur public minimise ses chances de voir émerger des usages réellement innovants de ses données. Il doit également tâcher d'orienter, de stimuler et de valoriser les usages de ses données.

Les moyens sont nombreux :

- Certains sont classiques et bien connus des acteurs publics comme les appels à projets, l'animation de réseaux d'innovateurs, les rencontres publiques...
- D'autres formes, plus volontaristes ou plus innovantes, peuvent être plus pertinentes. Les **concours** ont été expérimentés aux États-Unis, en Grande-Bretagne, au Canada et en Australie – ainsi qu'à Rennes : les résultats sont souvent très intéressants, comme à Ottawa où, lancé en septembre 2010, Apps4Ottawa⁴⁰, a vu 101 applications⁴¹ concourir pour 50 000\$ canadiens de prix !
- Les **ateliers créatifs** sous toutes leurs formes sont également un moyen très efficace : "Barcamps", forums créatifs, "coding parties", compétitions amicales, etc. Si l'acteur public n'est pas toujours très au fait de ces méthodes, les réseaux d'innovateurs ou des médiateurs spécialisés les connaissent bien. Ces ateliers créatifs sont particulièrement utiles pour sortir des usages trop classiques et pour faire émerger des projets particulièrement innovants.

Plus en aval, la valorisation des projets existants peut prendre diverses formes. Au delà de la labellisation d'applications de qualité, démarche bien connue mais un peu lourde à mettre en œuvre, elle peut prendre la forme d'une **galerie**⁴² **des usages et/ou des applications**. Idéalement publiée à proximité de la description de la démarche et des jeux de données ouverts, elle permet de comprendre concrètement tout le potentiel créatif des données et montre synthétiquement les résultats de l'opération. Cette galerie peut aussi mettre en lumière des réalisations particulièrement emblématiques, choisies par l'acteur public, par un jury ou par un vote du public.

L'acteur public portera une attention toute particulière à la mise en relation entre certains réutilisateurs imaginatifs mais dépourvus de moyens de compétences techniques, et les acteurs capables de réaliser (ingénieurs, start-ups, "geeks"...), permettant ainsi à des problématiques de rencontrer des solutions insoupçonnées.

³⁹ "Ils ont cru que la communauté allait se bâtir toute seule" se plaint un critique des expériences (médiocres) d'Edmonton, Toronto et Vancouver : <http://weait.com/content/tragedy-edmonton-couver-open-data>

⁴⁰ www.apps4ottawa.ca/

⁴¹ www.ottawacitizen.com/technology/City+Ottawa+contest+attracts+double+expected+entries/4065650/story.html

⁴² Ou "showcase" comme disent les anglo-saxons. On verra par exemple celui de San Francisco : <http://datasf.gov/showcase/>

2.2. DEFINIR SA POLITIQUE D'OUVERTURE ET DE DIFFUSION DES INFORMATIONS PUBLIQUES : LES GRANDS CHOIX A FAIRE

Quelles données publier ? Sous quelles conditions et dans quelles limites ? Gratuites ou payantes ? Certaines de ces questions sont bien balisées comme celle, en amont, de la possibilité même de réutiliser des données publiques. D'autres questions, comme par exemple la problématique de la redevance, font l'objet de discussions parfois vives. Cette partie vise à éclairer les choix des collectivités territoriales, sans pour autant les faire à leur place, dans trois domaines : la nature des données à ouvrir, les conditions juridiques de leur mise en accès, et l'éventuelle tarification de ces données.

2.2.1. Quelles données ouvrir ?

► Ce que dit la loi

Du point de vue de la loi, on le rappelle, toutes les données produites ou reçues par un acteur public dans l'exercice de ses missions de service public ont vocation à devenir réutilisables, à l'exception :

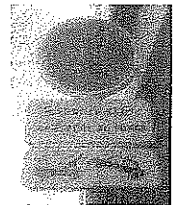
- des données personnelles,
- et des données susceptibles de porter atteinte à la sécurité publique.

Sont également exclues de l'application "automatique" de la loi, en matière de réutilisation :

- Les données auxquelles sont associées des droits d'auteur, sauf bien sûr si cette réutilisation est explicitement consentie,
- Les services publics industriels et commerciaux, comme par exemple les transports. Cela n'interdit pas aux acteurs publics de *choisir* de rendre ces données réutilisables, ou de l'imposer dans leurs cahiers des charges aux entreprises délégataires de ces services.

Les panneaux de signalisation, une donnée sensible ???

Qui penserait trouver par exemple des données sensibles dans les panneaux de signalisation urbains ? L'un d'entre eux, pourtant, signale les zones d'arrêt des convois de fonds. Cette information, même si elle est publique, peut être sensible.



► Quelles données méritent-elles d'être ouvertes ?

Il est sans doute bien rare qu'une donnée n'ait pas quelque intérêt à être ouverte. L'expérience des pays anglo-saxons et les témoignages des réutilisateurs nous permettent d'affirmer que quels que soient sa forme, toute donnée ou presque peut faire l'objet d'une réutilisation pertinente. C'est pourquoi il est particulièrement utile d'associer très tôt les réutilisateurs potentiels au recensement des données. Les données les plus pertinentes ne sont pas toujours celle que l'on imagine au départ !

Dans ce cas, comment choisir quelles données ouvrir en priorité ? Plusieurs critères de choix peuvent conduire la réflexion :

- **Les priorités publiques** : la mise à disposition de données peut contribuer à mobiliser les énergies pour atteindre un but collectif. Par exemple, tel acteur public pourrait vouloir sélectionner les données qui peuvent avoir un impact concret sur l'accessibilité de la ville et de services aux personnes à mobilité réduite : cartographie des services disposant de dispositif d'accessibilité, etc.
- **La volonté de soutenir une filière économique** : en ouvrant ces données, la collectivité peut fournir un terrain d'expérimentation à des entreprises d'un secteur qu'elle souhaite soutenir, ou pour lequel elle souhaite devenir attractive.
- **La demande** : si la demande spontanée demeure aujourd'hui assez faible, la situation pourrait rapidement changer. Un nombre croissant d'acteurs associatifs et économiques est informé de ce qu'il se passe en France et dans le monde autour des données publiques. Des acteurs

militants s'en saisissent, des entreprises les réclament pour développer leurs services – par exemple dans des domaines tels que la mobilité, ou les services géolocalisés. Et localement, en interrogeant les acteurs, on fera aisément émerger des attentes.

- **Le pragmatisme**, enfin : certaines données sont plus faciles à publier que d'autres (elles sont déjà informatisées d'une manière qui en facilite la mise à disposition) – on aura intérêt à travailler avec les services informatiques du territoire, pour identifier non ce qui difficile (les informaticiens savent très bien faire ça !), mais au contraire ce qui est facile. D'autres données pourraient rapidement susciter l'émergence d'applications, voire bénéficier d'un retour des utilisateurs qui les enrichiront. Dans d'autres cas, enfin, des opérateurs extérieurs, liés ou non à la collectivité, peuvent se montrer volontaires : à Rennes, c'est au départ l'opérateur des vélos en libre-service qui a ouvert ses données, avec le soutien de la métropole. L'acteur public a ensuite étendu la démarche à un grand nombre d'autres données.

► Des données intuitivement plus importantes que d'autres ?

Le champ des données publiques réutilisables est donc immense. Il est encore difficile d'en dresser une liste approchant de l'exhaustivité.

Que peut-on partager dans une ville ? (liste non exhaustive)

Des informations et des données	
<ul style="list-style-type: none"> • La description du territoire (cartes, cadastre...) • Des fonds documentaires (études, réglementation, statistiques, archives...) • Les données de la décision publique (projets, enquêtes, délibérations, subventions...) • Le fonctionnement des réseaux urbains (eau, énergie, transports, logistique, télécoms...) • La localisation et les horaires d'ouverture des services et des commerces 	<ul style="list-style-type: none"> • L'occupation des ressources et des capacités (voirie, bâtiments, espaces, parkings...) • Des mesures (environnement, trafic...) • Des événements (culture, sports, citoyenneté...) • Des informations touristiques, culturelles, • Les flux urbains (circulation...) • Des données de surveillance • ...
Des applications et des services	
<ul style="list-style-type: none"> • Des systèmes d'information géographiques (SIG) • Des modèles (représentation, prévision...) • Des applications permettant de calculer des droits (sociaux, etc.) 	<ul style="list-style-type: none"> • Des applications propres à chaque "métier" public • Des applications transversales : identification, localisation, paiement, sécurisation • Des services de paiement, de billetterie...

On peut cependant penser que certaines données peuvent revêtir une importance particulière. A partir de l'expérience d'autres collectivités françaises et européennes, on en proposera ici une première liste :

- Les codes et les règles produits par l'administration : ces informations constituent le socle de l'information administrative et la base de nombreux d'usages. Nous pensons par exemple : aux codes des communes, aux codes identifiant donnés par l'administration pour normer, ou identifier des produits des entités économiques etc. ; aux codes et ou normes propres à certains produits de consommation, de santé publique (médicaments, produits chimiques, etc.) ; aux codes Siren, Siret, etc.
- Les données devenues essentielles pour des domaines transverses : par exemple, les points de l'IGN, utilisés chaque jours par des milliers de géomètres pour établir des mesures correctes ; ou encore, la localisation des services publics.
- Les données de base de la citoyenneté et de la vie politique : limites géographiques des circonscriptions administratives, des circonscriptions politiques ; noms des représentants politiques et leur territoire ou leur domaine d'action ; bureaux de vote...
- Les données nécessaires au débat démocratique: budgets, marchés publics, délibérations, études et statistiques...
- Les données utiles à une meilleure gestion des énergies et des ressources : pics de consommation, localisation des équipement de recyclage des services de retraitement des déchets et leurs horaires d'ouverture ;

- Les données utiles à la santé et la sécurité publique : données épidémiologiques, pollution, adresses des services de santé, etc.
- Les données culturelles et touristiques, etc.

La réutilisation de ces données devrait être aussi aisée et libre que possible.

► Les mauvaises raisons de ne pas mettre à disposition des données

En dehors des cas prévus par la loi, les raisons apparentes de ne pas partager sont nombreuses. Là encore, l'expérience sur le terrain conduit à penser qu'elles sont souvent mal fondées, voire qu'elles servent de prétexte à une réticence de fond à partager l'information

Mes données ne sont pas de bonne qualité, elles comportent des erreurs

Qui prétend ne jamais en faire ? Et même avec de nombreuses erreurs, les données peuvent tout de même intéresser les réutilisateurs : certaines sont très faciles à corriger, elles peuvent intéresser à titre de comparaison, etc. En outre, on constate que la publication de données peut, dans certains cas, conduire les utilisateurs eux-mêmes à les corriger, au plus grand bénéfice de l'acteur public. Le producteur aurait en revanche intérêt, pour tout jeu de données, à indiquer un taux d'erreur, même grossier.

Mes données ne sont pas de bonne qualité car elles sont incomplètes.

Il est très probable qu'elles soient en réalité complètes sous un certain angle : par exemple restreintes à une zone géographique bien définie. Mêmes "incomplètes", ces données peuvent être utiles aux réutilisateurs pour initier un travail. Là encore, le producteur doit documenter clairement la couverture de ses données.

Mes données ne sont pas de bonne qualité car elles ne sont pas assez précises.

La précision recherchée ne dépend que de l'usage. Aux réutilisateurs de dire si la précision leur convient.

Ces données sont de bonne qualité mais trop anciennes.

Des données périmées pour leur usage initial peuvent néanmoins intéresser de nombreux acteurs - chercheurs, étudiants, consultants... - désireux d'effectuer par exemple des travaux de synthèse, d'analyse historique, etc.

Ces données sont de bonne qualité mais leur mise à jour n'est pas assez régulière.

Là encore, elles peuvent présenter un intérêt pour des usages d'analyse comparative, de travaux de synthèses ou autres.

Le format de ces données n'est pas standard.

Si le format est documenté, ce n'est pas un problème. Si les réutilisateurs y voient un intérêt, convertir des données bien documentées d'un format à un autre n'est pas si difficile.

Mais le format de ces données est mal documenté !

Un réutilisateur vraiment intéressé trouvera toujours moyen de le faire a posteriori. Peut-être une bonne occasion pour vous de dialoguer avec ces réutilisateurs pour produire une documentation de qualité.

Ces données n'intéressent personne !

Chiche ! Dans tous les cas, ne pas préjuger à l'avance qu'elles n'intéressent personne.

2.2.2. Les conditions juridiques de réutilisation des données

La loi encadre la réutilisation des données publiques en précisant les droits et obligations des acteurs publics comme des réutilisateurs (voir chapitre 1.4, p. 34). Pour autant, les producteurs de données peuvent vouloir, à l'intérieur de ce cadre légal, effectuer des choix plus précis. Tel voudra rendre aussi libre que possible la réutilisation ; tel autre voudra imposer des conditions strictes en matière d'intégrité des données, d'identification de la source, etc. Tel autre choisira de différencier les conditions qui s'appliquent aux usages commerciaux et non-commerciaux.

Préciser les conditions d'usage de ces données requiert des compétences juridiques. Dont toutes les collectivités ne disposent pas aisément. C'est pourquoi, afin de faciliter leur choix, différentes organisations ont élaboré des **cadres juridiques "clés en mains"** qui ont pris la forme, soit de licences, soit de conditions générales d'utilisation⁴³.

Plusieurs points peuvent en particulier faire l'objet de choix spécifiques, qui conditionneront le choix de tel ou tel cadre juridique :

- Les restrictions à la réutilisation, "toute restriction devant impérativement être motivée par l'intérêt général, respecter le droit de la concurrence et être proportionnée à la réutilisation envisagée"⁴⁴
- Le choix de percevoir une redevance pour la réutilisation des données
- Le choix d'ouvrir plus largement certaines possibilités, qui doivent être explicitement indiquées : autoriser l'altération des données, lever l'obligation d'indiquer la source des données leur date de mise à jour...
- La volonté de s'inscrire dans un cadre juridique commun avec d'autres acteurs, par exemple dans un domaine professionnel (on pense aux données de mobilité), ou en choisissant une licence de portée internationale.

► Pourquoi c'est important

Un cadre juridique n'est pas seulement là pour *protéger* mais aussi pour *faciliter la rencontre entre le détenteur des données et ses réutilisateurs*. Ainsi, un cadre trop restrictif peut conduire à une valorisation médiocre voire, nulle. Un cadre trop imprécis peut créer de l'incertitude, ou favoriser des usages néfastes qui nuisent à tous les acteurs – pensons à l'impact d'informations sur le trafic, si elles sont rediffusées trop tardivement. Un cadre très précis mais incompréhensible par des non-spécialistes chassera les petits acteurs, souvent les plus innovants...

Il est généralement utile de choisir un cadre juridique connu (voir ci-dessous) plutôt que de réinventer la roue. Produire soi-même ses conditions d'utilisation ou sa propre licence est contre-productif. Un cadre juridique particulier est un objet juridique non identifié, obligeant tous les utilisateurs à l'analyser et le discuter, faisant ainsi perdre un temps précieux à tout le monde. Vous avez en outre toutes les chances de commettre des erreurs alors qu'il existe de nombreux cadres déjà éprouvés

On le voit, le choix d'un cadre juridique a donc un impact réel sur le succès des usages de données publiques.

Mais comment choisir ? Nous nous proposons de faire le tour des différents critères et fonctionnalités qui distinguent les cadres juridiques "prêts à l'emploi".

⁴³ Pourquoi parlons-nous de "cadre juridique" et pas de "licence de réutilisation" ? Tout simplement parce qu'une licence n'est pas le seul cadre juridique possible. La réutilisation n'a pas nécessairement besoin de licence, la loi de 1978 fournissant un cadre par défaut. Plus récemment, l'APIE a produit des "conditions générales d'utilisation" qui jouent un rôle proche de la notion de licence.

⁴⁴ Les données publiques. Guide juridique et pratique, AEC.

► **Les principaux cadres juridiques en France**

Nous avons retenu 8 cadres juridiques, ou plutôt 7 cadres et une "situation par défaut".

La situation par défaut est celle qui voit l'acteur public ne mentionner aucun cadre : c'est alors la loi du 17 juillet 1978, modifiée en 2005, qui s'applique⁴⁵. Si l'acteur public fait le choix de s'en tenir à la loi, il n'est pas indispensable de le préciser, mais cela facilitera néanmoins la vie des réutilisateurs. A noter que la décision d'appliquer une redevance ou des conditions restrictives de réutilisation impose la production d'une licence.

D'autres cadres juridiques précisent les conditions de réutilisation :

- **Les Conditions générales de l'APIE⁴⁶** (octobre 2010). Ces Conditions générales sont une reformulation pédagogique du cadre standard de la loi.
- **Les licences-types de l'APIE⁴⁷**, qui sont des cadres à compléter pour des données la réutilisation est soumise à des conditions particulières (par exemple de mise à jour) et/ou au paiement d'une redevance.
- **La licence OdbL⁴⁸** (dite aussi ODC-ODbL, 2010). Licence internationale établie par l'organisation Opendatacommons. A l'heure où nous écrivons ces lignes, la ville de Paris travaille à une traduction française dont un premier jet peut être consulté et commenté.
- **La licence ODC-by⁴⁹** (2010). Idem.
- **La licence PDDL 1.0⁵⁰** (2010). Idem.
- **Les licences Creative Commons⁵¹** (depuis 2004). Il s'agit en fait d'un groupe de licences et qui définissent différentes manières pour un auteur de faciliter la circulation et la réutilisation de son œuvre, dans le cadre de la législation sur le droit d'auteur. Son intérêt dans le domaine des contenus littéraires et artistiques n'est plus à démontrer mais ces licences semblent moins bien adaptées au domaine des bases de données. Nous les évoquons à titre d'information et de comparaison.
- **La licence IP du ministère de la Justice⁵²** (mai 2010). Une des premières licences type en France, inspirée des Creative commons.

Ces différentes licences partagent certains points communs et se différencient sur d'autres :

Des fonctionnalités communes

- **L'obligation de la mention d'auteur (BY)** semble être une des fonctionnalités les mieux partagées, puisqu'un seul cadre juridique dispense de la mention d'auteur, la licence PDDL 1.0 – qui par essence dispense de toute obligation.
- La plupart des cadres juridiques **permettent un usage commercial tel quel** : autrement dit, ils rendent possible à un réutilisateur de revendre des données publiques sans les transformer, ni les enrichir. La licence IP du ministère de la Justice fait seule exception et interdit une réutilisation telles quelles de ses données à des fins commerciales : pour tirer un bénéfice commercial de la réutilisation, il faut les enrichir d'une manière ou d'une autre.

Des fonctionnalités clivantes

Certaines fonctionnalités sont adoptées par certains cadres juridiques et non par les autres. Ces fonctionnalités établissent une sorte de clivage sur quelques grands sujet d'usage des données.

La loi française de 1978 exige, sauf consentement explicite, que les réutilisateurs indiquent la source des données et leur date de mise à jour, et respectent l'intégrité des données. Certaines licences facilitent la levée de telle ou telle de ces obligations :

- **Permettre l'altération des données** est une fonctionnalité explicite des licences de l'Opendatacommons (ODC). Cette fonctionnalité est bien adaptée aux données qui proviennent de sources multiples et dont la mise en valeur exige des retraitements. Elle simplifie la réutilisation des données en ce sens qu'un auteur modifiant une donnée n'a pas à en

⁴⁵ <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068643&dateTexte=20101103>

⁴⁶ <https://www.apiefrance.fr/sections/actualites/des-conditions-generales-pour-la-reutilisation-des-informations-publiques/view>

⁴⁷ https://www.apiefrance.fr/sections/acces_thematique/reutilisation-des-informations-publiques/licence-type/view

⁴⁸ <http://www.opendatacommons.org/licenses/odbl/>

⁴⁹ <http://www.opendatacommons.org/licenses/by/>

⁵⁰ <http://www.opendatacommons.org/licenses/pddl/>

⁵¹ <http://creativecommons.org/about/licenses>

⁵² http://www.rlp.justice.fr/information_publicque_librement_reutilisable

demander l'autorisation au cas par cas. C'est le cas par exemple des données d'OpenStreetMap : la carte étant enrichie et modifiée en permanence par des dizaines de milliers de contributeurs, il fallait un mécanisme simple et rapide pour laisser chaque contributeur enrichir la carte à loisir.

La loi française de 1978 exige, sauf consentement explicite, que les réutilisateurs indiquent la source des données et leur date de mise à jour, et respectent l'intégrité des données. Certaines licences facilitent la levée de telle ou telle de ces obligations.

- **L'obligation de mention de la date de dernière mise à jour des données** n'est pas prévue par les licences de l'Open Knowledge Foundation. Cela dit, ces licences ne sauraient remplacer la loi et l'on peut penser qu'un réutilisateur de données publiques sous licence ODbL, par exemple, devra tout de même se conformer à cette obligation. Les licences de l'OKFN gagneraient peut-être en efficacité, en France, à expliciter clairement ce point.

D'autres fonctions différencient également certaines licences :

- **L'interdiction de l'usage commercial** des données est une fonctionnalité d'une des licences Creative Commons (CC-NC), la seule à l'adopter. Les données concernées par cette obligation ne peuvent pas être réutilisées dans un cadre commercial. Cependant, la réutilisation des données publiques remplit très largement un but économique : la loi ne permet donc pas, sauf exception qui doit être motivée par l'intérêt général, d'interdire de manière générale les réutilisations commerciales.
- La licence IP propose une fonctionnalité originale en **permettant un usage commercial conditionné par l'obligation d'enrichissement des données**.
- **L'obligation de fournir les données sans entrave technique**, comme par exemple des systèmes techniques de gestion des droits de propriété intellectuelle (DRM), est une fonctionnalité originale de la licence ODbL.

Enfin, les licences-types de l'APIE sont, dans ce paysage, les seules à recouvrir le **paiement d'une redevance** sur les données publiques. La question de la tarification des données est abordée au paragraphe suivant.

2.2.3. Faut-il tarifier la réutilisation des données publiques ?

Les données publiques doivent-elles être mises à disposition des réutilisateurs (qui peuvent être des entreprises installées, mais aussi des "jeunes pousses", des associations, des citoyens, des chercheurs) de manière gratuite ou payante ? Et lorsqu'elles sont payantes, sur quelles bases ?

► Dans les pays anglo-saxons, le choix de la gratuité

Au Royaume-Uni et aux Etats-Unis, la question est pour l'essentiel tranchée en faveur de la gratuité, qui ne fait pratiquement plus débat. Plusieurs arguments ont conduit les acteurs publics à ce choix :

- Un argument de principe : la production de ces données est financée par l'impôt. Cet argument comporte deux faiblesses : d'une part, la mise à disposition en dehors de l'administration peut entraîner des coûts ; et d'autre part, les transférer aux entreprises pourrait s'assimiler à une forme de "privatisation des profits", tandis que les coûts demeurent publics.
- Deux arguments économiques : d'une part, le coût de distribution en ligne de la plupart des données publiques est nul, ou presque – et par conséquent son prix "optimal" est très faible, inférieur au coût qu'entraînerait le fait de facturer les réutilisateurs, relancer les mauvais payeurs, etc. D'autre part, les données publiques forment la base d'activités économiques qui contribuent à la croissance et génèrent des rentrées fiscales supérieures aux revenus directs escomptés de la vente des données⁵³.

En France, et ailleurs en Europe, la question est plus ouverte. Si les voix en faveur de la gratuité se font entendre, le déséquilibre des comptes publics favorise également la tentation inverse. Comme l'indique ce compte-rendu d'un conseil des Ministres de juin 2010⁵⁴ portant sur la revue générale des politiques publiques, " *La mise en œuvre d'une politique volontariste de licences, notamment dans le cadre de leur réutilisation commerciale, peut rapidement augmenter les produits pour l'État et ses établissements publics. En effet, les licences déjà octroyées qui concernent, à l'heure actuelle, essentiellement les secteurs traditionnels de l'information géographique, météorologique, économique et financière, rapportent chaque année environ 50 millions d'euros à l'État.*"

► Ce que dit la loi

La loi prévoit que les détenteurs des données publiques peuvent exiger une redevance dans le but de couvrir les coûts de collecte, de production et de mise à disposition, ainsi que les investissements consentis. Ce principe établit un cadre qui peut sembler contraignant, mais laisse pour l'instant d'assez nombreuses zones d'ombre :

- Le but ne peut pas être de dégager des bénéfices sur la mise à disposition des données publiques. Mais que revêt la couverture des coûts ? Si des données doivent de toute manière être produites dans le cadre normal de l'activité de service public (financée par l'impôt), il paraît difficilement justifiable d'en faire payer autre chose que les coûts spécifiquement liés à leur mise à disposition à des fins de réutilisation – coûts qui seront souvent assez faibles, du moins si les systèmes informatiques ont été conçus comme il faut.
- La rémunération doit respecter les règles de la concurrence et être "équilibrée, proportionnelle et non discriminatoire". L'exclusivité est interdite, ainsi que la discrimination entre deux acteurs ayant des besoins équivalents. Un réutilisateur public ne peut par exemple pas être favorisé par rapport à une entreprise, du moins pour le même type de réutilisation. Il est cependant possible d'établir des différences, par exemple entre usage commercial et non-commercial, ou encore entre des "gros" réutilisateurs et des "petits".

⁵³ En 2008, une étude préalable avait été commandée par le gouvernement britannique à 3 chercheurs de l'université de Cambridge (Pollock, Newbery, Bentley 2008). L'Etat devait-il, du point de vue de l'efficacité économique, donner ses données ou les faire payer ? Selon leurs conclusions, "la chose la plus intelligente à faire pour l'Etat est de faire payer le coût de distribution, ce qui tend vers zéro pour une distribution en ligne et donc revient à les mettre à disposition gratuitement". Cette analyse des coûts et bénéfices du modèle commercialisé a donc conclu en faveur d'une mise à disposition gracieuse.

⁵⁴ <https://www.apiefrance.com/sections/actualites/conseil-des-ministres-du-30-juin-2010-revue-generale-des-politiques-publiques/view>

En revanche, il est clair que ce sont les coûts qui doivent guider la tarification, et non un objectif de maximisation du revenu direct susceptible d'être tiré de la mise à disposition des données.

► Comment définir sa stratégie de tarification ou de gratuité ?

Plusieurs facteurs doivent être pris en compte⁵⁵ :

- **Combien coûte réellement la mise à disposition des données – et combien coûte le fait de les facturer ?** Pour un système informatique correctement conçu, qui sépare efficacement "données" et "traitements" (un principe appliqué depuis plusieurs décennies), la réponse pour des fichiers relativement simples et pérennes est clairement : rien, ou presque – en tout cas bien moins que le coût de facturation et de recouvrement. En revanche, s'agissant de bases de données complexes ou encore de données en temps réel, il ne s'agit plus d'exporter des fichiers mis en téléchargement, mais d'ouvrir des "portes d'entrée" sur les infrastructures informatiques de l'acteur public, qu'il aura souvent fallu adapter à cet effet. L'investissement peut s'avérer significatif, et si la demande est forte, le coût d'exploitation ne sera pas non plus négligeable⁵⁶.
- **Quel est le consentement à payer, de la part de qui ?** Les grands acteurs du "marché de l'information" seront prêts à payer plus que les petites entreprises, et *a fortiori* que les associations – y compris, d'ailleurs, pour évincer les nouveaux entrants. On payera plus volontiers des données enrichies que des données brutes. Bref, tout dépend de ce que les acteurs publics veulent privilégier, ainsi que de *qui* ils veulent privilégier.
- **Quels sont les objectifs de long terme de l'acteur public ?** Celui-ci peut en effet poursuivre plusieurs objectifs différents, et pas toujours compatibles :
 - Développer des revenus extrafiscaux, en facturant le plus possible. Cet objectif invite à favoriser au plus près du "consentement à payer", en flirtant avec les limites de la loi.
 - Favoriser le développement de l'activité économique, de l'innovation sociale et des connaissances que peuvent induire les données publiques. C'est ce que visent les Anglo-saxons, c'est également la raison pour laquelle un puissant mouvement (dans lequel la France se montre active) se développe en faveur de la gratuité des données scientifiques. Cet objectif invite plutôt à la gratuité, sauf dans le cas où le coût de mise à disposition est très élevé.
 - Atteindre des objectifs politiques plus larges, en matière d'éducation, de recherche, de culture, de développement durable, de participation démocratique, etc. Cette orientation tend également vers la gratuité, du moins dans les domaines considérés comme prioritaires.
- **Enfin, comment éviter que la grille tarifaire elle-même ne dissuade la demande ?** C'est le cas en particulier lorsqu'un dispositif trop "intelligent" crée une complexité incompréhensible pour la plupart des acteurs. Une administration nationale a par exemple récemment diffusé une grille tarifaire dans laquelle elle distingue une demi-douzaine de cas (et en leur sein, une multitude de sous-cas), dont certains ne peuvent que créer une incertitude totale pour les réutilisateurs. S'agissant des entreprises, elle distingue par exemple entre un usage "interne" et un usage "externe" (commercial) ; quant à l'usage externe, le tarif est proportionnel au chiffre d'affaires lié à la réutilisation des données, alors que cette réutilisation se fait très souvent en combinant plusieurs données de plusieurs sources, rendant impossible la mesure de la contribution d'une donnée précise...

Sans faire de la gratuité un principe, on ne saurait trop inviter les acteurs locaux à ne tarifier leurs données qu'en cas de réelle nécessité, après avoir exploré toutes les autres pistes. Dans la plupart des cas (au moins dans les premiers temps), les inconvénients de la tarification dépassent ses avantages.

⁵⁵ Pour aller plus loin, on pourra se référer à l'étude économique réalisée en 2010 par le Bureau d'Economie Théorique et Appliquée (BETA) de l'université de Strasbourg pour le compte de l'APIE, mais non sans en exercer une lecture critique. Le BETA part en effet de l'idée que la mise à disposition de données sous une forme "brute" (même documentées) en limite très fortement la réutilisation, alors que la plupart des réutilisateurs, même petits, semblent réclamer des données aussi brutes que possible.

https://www.apiefrance.fr/sections/acces_thematique/reutilisation-des-informations-publiques/etude-economique/view

⁵⁶ Sur son service Google Maps, Google a tranché d'une manière intéressante : les petits utilisateurs ne sont pas facturés, tandis que ceux qui consomment une part mesurable des ressources informatiques le sont.

2.3. COMMENT RENDRE LES DONNEES ACCESSIBLES ?

La diffusion des données est une étape clé du processus de partage des données publiques. Face à ces problématiques qui comportent parfois des dimensions assez techniques, nous présentons ici les grands traits de ce sujet, pour y voir plus clair et pour commencer à avancer sans attendre que tout soit résolu.

Le W3C s'intéresse à la publication de données publiques

Le W3C, l'organisation internationale qui édicte les standard du web – la norme HTML, c'est lui ! – a publié en septembre 2009 un "Working Draft" (document de travail) sur la publication des données publiques : [Publishing Open Government Data](#)⁵⁷. Si le processus de standardisation de ce document en est encore à ses débuts, ce document pourra néanmoins intéresser les lecteurs avancés qui voudraient anticiper à sa rédaction, voire y participer.

2.3.1. Formats et méthodes d'accès aux données

(note : les non-techniciens pourront se contenter de lire les "grands principes" avant de passer à la partie suivante)

► Les grands principes

Pour être facilement réutilisables, vos données doivent partager quelques règles simples.

Des données documentées

Les données doivent d'abord être documentées, en ce qui concerne leur contenu, leur format, leurs moyens et leurs conditions d'accès.

Il s'agit d'une étape importante car elle conditionne la réutilisation pertinente de vos données. Une donnée n'est jamais trop décrite, expliquée.

Prenons par un exemple un champ intitulé "nom" : s'agit-il d'un nom de personne, d'un nom de produit, du nom d'un auteur, d'un fabricant, d'un usager, ou d'autre chose encore ? Prenons un exemple apparemment moins problématique, un champ "date". S'agit-il d'une date de création, de changement, d'une date approximative, d'une date précise, d'une date de modification de la donnée ? Et comment cette date s'écrit-elle : 2010-12-02 ? 02/12/2010 ? 12-dec-2010 ? La documentation des données facilite leur compréhension et donc leur réutilisation.

La documentation concerne tous les niveaux de la diffusion des données :

- La sémantique des données : que représente cette donnée, que signifie chaque champ d'un fichier ?
- La syntaxe des données : comment s'écrit chaque information ?
- Le format des fichiers contenant les données
- La méthode d'accès aux données

Pour chaque donnée on tâchera de préciser les éléments suivants :

- Une description libre en bon français tâchant de décrire au mieux chaque donnée
- L'éventuel format ou norme formelle adoptée (par exemple pour une date)
- Eventuellement la taille maximum ou fixe de la donnée
- Le type de donnée : alphanumérique, numérique, binaire, etc.

On précisera si le schéma global des données est lui-même normalisé ou repris à l'identique d'une autre base.

Enfin, il est important de documenter le jeu de données en lui-même : d'où proviennent ces données ? Sont-elles brutes ou issues du traitement de données "primaires" – et dans ce cas,

⁵⁷ <http://www.w3.org/TR/gov-data/>

lesquelles ? Quelle est leur dernière date de mise à jour ? Leur fréquence de mise à jour ? Qui en est l'auteur ? Y a-t-il des conditions de réutilisation, des licences ? Etc.

Données brutes ou retraitées ?

Les réutilisateurs revendiquent la plupart du temps des "**données brutes**", pour une meilleure transparence. Ils peuvent aussi apprécier des données "retravaillées", "enrichies", à partir du moment où les traitements sont correctement documentés.

Dans la plupart des cas, on aura intérêt à livrer les données brutes par défaut, quitte à proposer également des données plus élaborées.

Privilégier les standards ouverts

Enfin, au-delà de la documentation et de la diffusion de données brutes, le producteur de données regardera les **standards ouverts** relatifs aux données diffusées. Certains domaines ont vu naître de tels standards qui concourent à une meilleure interopérabilité et diffusion des savoirs. Par exemple, le fameux format RSS, servant de format de syndication de contenu web, est devenu un incontournable. Il serait dommage de réinventer la roue...

Un format ouvert est caractérisé par "des spécifications techniques documentées, publiées, non payantes, sans brevet dessus, sans royalties dessus" pour des données dont l'usage est "indépendant d'un logiciel particulier, d'un système d'exploitation ou d'une société⁵⁸". Un format ouvert garantit donc la maîtrise des données, leur pérennité, leur archivage et leur interopérabilité – et facilite donc la réutilisation.

Standard ouvert : une définition juridique pour aider les acteurs publics

Le terme de "standard ouvert" est défini par la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, comme suit : "tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérables et dont les spécifications techniques sont publiques et sans restriction d'accès ni de mise en œuvre" (article 4).

Dans ses relations avec ses fournisseurs technologiques, un acteur public peut avoir intérêt à exiger la livraison d'outils qui produisent des données selon un standard ouvert, au sens de cette loi.

► **La sémantique des données**

La documentation de la sémantique ("que signifie tel champ d'un fichier ?") est parfois négligée car elle peut paraître évidente au producteur des données. Même une simple date peut poser un problème de sémantique : s'agit-il du calendrier grégorien ou d'un calendrier musulman ? L'encyclopédie Wikipedia recense plus d'une trentaine de calendriers⁵⁹...

Normaliser la sémantique des données sur le web : le modèle RDF

Le format RDF⁶⁰ permet de documenter, de décrire de façon standard des jeux de données sur le web en utilisant des propriétés publiques et non ambiguës, regroupées en "vocabulaires". Par exemple, le "vocabulaire RDF" appelé FOAF décrit la propriété "name" comme le nom propre d'une personne. Pour être sûr qu'il n'y ai pas d'ambiguïté, cette propriété "name" est elle-même documentée à une adresse web unique (URI) contrôlée par l'auteur du vocabulaire FOAF : <http://xmlns.com/foaf/0.1/name>.

Le format RDF, en plus d'être un format ouvert, améliore l'interopérabilité des données. Il est à la base de ce qu'on appelle le "web sémantique" ou "web des données".

⁵⁸ <http://formats-ouverts.org/blog/2004/07/01/12-un-article-de-loi-definit-ce-que-sont-les-formats-ouverts>

⁵⁹ <http://fr.wikipedia.org/wiki/Calendrier>

⁶⁰ <http://www.w3.org/RDF/>

► La syntaxe des données

La syntaxe des données doit aussi faire l'objet d'une attention particulière. Beaucoup d'éléments codifiés peuvent l'objet de syntaxes très différentes qui, si on ne les connaît pas, peuvent altérer la compréhension des informations. On l'a vu plus haut pour l'exemple des dates mais cela vaut aussi pour de très nombreuses autres données : ISBN, ISSN, numérotations en tous genres (numéros de département, codes postaux, numéros INSEE...), etc. Il est donc utile de documenter clairement la syntaxe de chaque champ.

Là encore, il n'est pas rare de voir des standards ouverts couvrant la syntaxe de certaines données.

► Les méthodes d'accès aux données

Il existe deux façons de donner accès à des données sur le web : le téléchargement d'un fichier, ou l'accès direct à la source de la donnée.

Les fichiers téléchargeables

L'accès direct à un fichier contenant les données est souvent la méthode la plus simple car les outils informatiques proposent généralement une fonction d'export des données sous forme de fichier. Elle est cependant moins bien adaptée lorsque les données changent souvent et qu'il y en a beaucoup : cela oblige les réutilisateurs à recharger fréquemment l'ensemble du fichier et peut être consommateur en ressources système et réseau.

Les interfaces de programmation (API)

Cette limite explique pourquoi se développe fréquemment un mode d'accès direct aux données uniquement nécessaires à travers une "interface de programmation" (API en anglais). Une API va permettre à une application "demandeuse" (celle du réutilisateur) de demander à l'application "fournisseuse" (celle du détenteur des données "sources") les informations dont elle a besoin. Les grands acteurs du numérique en général et du web particulier, comme Google, Amazon, Facebook, etc., ont établi leur succès grâce à leurs APIs : leurs données sont en quelque sorte "réutilisables", sans que leurs bases de données ne quittent jamais leurs ordinateurs.

Il existe deux grands types d'APIs dans le monde du web : les APIs REST et SOAP.

- SOAP est riche en fonctionnalités mais complexe à mettre en œuvre, en particulier pour les réutilisateurs.
- A l'inverse, les APIs REST sont plus légères et plus simples à mettre en œuvre. Ses dernières sont d'ailleurs plébiscitées par les producteurs de données comme par les réutilisateurs. SOAP connaît quelque succès dans les réseaux internes des grandes entreprises mais il est presque inexistant sur le web : la plupart des grands acteurs du web ou des grands producteurs de données publiques ne l'utilisent pas.

Pour être efficace, l'API doit être clairement documentée. Par ailleurs, selon vos jeux de données, il est possible de choisir une API existante plutôt que de réinventer la roue. Si vous utilisez des données décrites en RDF, il peut être intéressant de mettre en œuvre l'API standard SPARQL⁶¹. SPARQL est à la fois un protocole et un langage de requête pour les données en RDF. Il permet d'interroger toutes les données d'une base de données accessible sur le web à l'aide d'une seule technique : il n'est donc pas nécessaire de développer une API pour chaque type de données. En outre, il existe déjà de nombreux outils de diffusion et d'interrogation qui reposent sur SPARQL, évitant ainsi un travail fastidieux.

Une API standard pour le suivi d'anomalies impliquant les services publics du territoire ?

<http://open311.org>

Accès aux services d'enlèvement des objets encombrants, signalement d'éclairages défectueux ou de graffitis, lieux et horaires des centres de tri : autant d'activités que les territoires traitent aujourd'hui, souvent, à l'aide de centres d'appels dédiés : "Allo mairie", "Allo propreté", etc. Open311, un nom inspiré par le numéro de téléphone standard pour ce genre de services aux États-Unis – le 311 –, est une API ouverte qui permet de standardiser les données et les

⁶¹ <http://fr.wikipedia.org/wiki/SPARQL>

échanges de données dans le suivi d'anomalies impliquant les services publics d'un territoire.

Cette API est d'ores et déjà utilisée par quelques grandes villes américaines comme New York ou San Francisco. Elle permet, non seulement d'accéder à l'information, mais également d'en envoyer : déclarer un problème, par exemple. *"En rendant publique cette information, cela produit de la transparence et une meilleure prise en compte de la part de ceux qui ont la charge de ces problèmes. La transparence garantit aussi que la voix de tous est écoutée et encourage à plus de participation"*⁶², indique le site d'Open311.

Cette API rencontre un certain succès avec plus d'une vingtaine d'applications qui l'exploitent pour des usages parfois très originaux : la cartographie du bruit dans la ville, un tableau de bord hyperlocal (l'activité publique autour de mon pâté de maison), des jeux participatifs, etc.

► **Dans quel format de fichier publier ses données ?**

"N'importe quel format", demandent de nombreux réutilisateurs, "du moment qu'il est ouvert et documenté". Les standards ouverts sont importants, ils sont toujours préférables, mais un fichier bien documenté qui ne respecte aucun standard est toujours préférable à pas de fichier du tout. L'absence de standard ne constitue pas une bonne raison d'attendre...

Les formats de fichiers

Nous l'avons dit, les formats ouverts garantissent une meilleure réutilisabilité des données. Certains formats de fichiers propriétaires peuvent cependant être acceptables lorsqu'ils sont extrêmement répandus et qu'on trouve facilement des outils à bas coût pour les manipuler. Par exemple, le format .xls (pour les documents issus du tableur Excel de Microsoft), s'il n'est pas ouvert, peut éventuellement convenir et sera toujours "mieux que rien". A coût de production égal, on lui préférera cependant le format Open Document⁶³ qui est, lui, réellement ouvert.

Les formats d'API

Dans le monde des API web, les formats ouverts XML et JSON⁶⁴ règnent en maîtres.

- On ne présente plus XML et ses qualités dans la gestion des documents structurés, mais beaucoup d'experts considèrent qu'il est plus adapté au traitement de documents que de données.
- Le format JSON, plus récent, est né de ce constat et vise tout particulièrement le traitement des données. Ses promoteurs assurent qu'il présente l'intérêt d'être plus simple et plus compact que XML et de mieux interagir avec les langages de programmation. Depuis 2010, par exemple, le site de microblogging Twitter.com, succès planétaire avec plus de 100 millions de comptes, ne délivre plus de données en XML au profit de JSON. JSON a aussi ses détracteurs et les défauts d'un nouvel entrant face à un XML bien implanté et qui continue de progresser⁶⁵.

2.3.2. Informer sur les données disponibles : répertoires et portails

Il ne suffit pas de mettre ses données en ligne : encore faut-il le faire savoir, et permettre à tous ceux qui s'y intéresse de savoir quelles données sont disponibles, sous quelle forme et dans quelles conditions. Il existe plusieurs moyens et plusieurs "espaces" pour informer sur les données disponibles, tous n'étant pas nécessairement créés et gérés par les détenteurs des données.

► **Une obligation : informer sur les informations publiques que l'on détient**

En France, la loi oblige tous les acteurs publics à **informer le public des informations publiques qu'ils produisent ou détiennent**, sous la forme d'un *répertoire des informations publiques*. S'ils disposent d'un site internet, ce répertoire doit alors exister sous une forme électronique.

⁶² <http://open311.org/learn/>
⁶³ <http://fr.wikipedia.org/wiki/OpenDocument>
⁶⁴ <http://www.json.org/>
⁶⁵ E4X, Ecmascript pour XML, est une extension d'Ecmascript qui rend plus simple la lecture du XML. Elle d'ores et déjà intégrée dans Firefox et Google Chrome : http://fr.wikipedia.org/wiki/ECMAScript_pour_XML

Ce répertoire doit mentionner pour chaque document ou jeu de données, au moins : le titre, l'objet, la date de création, les conditions de réutilisation, la fréquence et la nature des mises à jour.

Ce répertoire concerne *toutes* les informations de service public détenues ou mises à jour par les acteurs publics, qu'elles soient ou non mises à disposition sous la forme de données informatiques accessibles en ligne. Le répertoire facilite à la fois le droit d'accès, et celui de réutilisation.

Mais, bien évidemment, une action délibérée de mise en ligne des données publiques, telle que nous la décrivons dans ce guide, ne sera pas complète sans un répertoire complet, accessible et bien organisé des données.

Dans tous les cas, il est dans l'intérêt de l'acteur public, qu'il s'engage ou non dans une politique volontariste de réutilisation de ses données, de fournir des informations supplémentaires tels que la personne à contacter pour demander l'accès aux données, le mode d'accès aux données, le format technique des données, etc. L'APIE mentionne sur son site plusieurs répertoires à titre d'exemple, et propose un guide technique très complet⁶⁶.

► Premiers pas : un répertoire simple

Au-delà de l'obligation légale, un bon **répertoire des informations publiques** peut offrir une **bonne base de publication et de référencement des données** :

- Il liste en principe toutes les données détenues (qu'elles soient en ligne ou non) ;
- Il donne une vue d'ensemble qui permet aux réutilisateurs de comprendre ce qui existe et ce qui est plus ou moins aisément disponible ;
- Il décrit précisément chaque jeu de données ;
- Il peut, soit constituer la plate-forme sur laquelle certains jeux de données seront physiquement disponibles, soit "pointer" sur les sources des données qu'il référence.

Certains acteurs peuvent faire le choix de distinguer le répertoire exhaustif des informations de service public (exigé par la loi de 1978) et le répertoire plus restreint des "données publiques" qu'ils mettent en ligne. Cette option a du sens du point de vue des réutilisateurs. Encore faut-il ne pas confondre deux objectifs :

- Recenser les informations mises en ligne à des fins de réutilisation ne dispense pas de l'obligation légale de répertorier *toutes* les informations de service public détenues ;
- Un répertoire des données publiques mises en ligne doit être pensé comme un *service* destiné aux réutilisateurs, pas comme une simple liste.

► Du répertoire au "portail"

La plupart des acteurs publics qui ont ouvert leurs données en France ont ainsi choisi de créer une **plateforme ou un portail de publication spécifique**. Ces plateformes sont conçues dans le but de cataloguer et diffuser des données réutilisables. Elles vont à l'essentiel sans perdre de temps à lister ce qui n'est pas disponible. Par ailleurs, ces plateformes intègrent des services qui facilitent la réutilisation. Souvent il s'agit :

- De "métadonnées" précises destinées à informer les réutilisateurs : description des ressources, conditions de réutilisation, contacts...
- De la documentation technique précise des méthodes d'accès (fichiers ou APIs)
- De forums de discussion pour aider les réutilisateurs dans leurs démarches
- D'espaces de promotion des usages et des applications réalisées grâce aux données

Comme souvent pour ce genre d'outil, il existe des solutions de plateformes/portail très différentes, qui répondent aux différents besoins des acteurs publics.

Certains acteurs pourront choisir un **développement spécifique interne (ou sous-traité)**, comme c'est le cas de certains grands portails nationaux aux États-Unis, en Australie, et bientôt en France. Un développement spécifique peut être utile dans plusieurs cas : par exemple quand l'acteur public doit intégrer des fonctionnalités spécifiques correspondant à son environnement ou plus simplement quand il ne trouve pas d'autres solutions adaptées à ses besoins et ses contraintes. Il présente en revanche l'inconvénient de coûter cher, de prendre du temps et

⁶⁶ <http://bit.ly/a5rRk0>

d'aboutir à un dispositif difficile à faire évoluer. Cette voie n'est pas recommandée aux acteurs territoriaux.

Depuis 2010, plusieurs initiatives de **plateformes "open source"** permettent de monter un portail à peu de frais :

- Après l'avoir développée à l'origine pour Rennes Métropole, la société In-Cité a publié récemment sa solution de plateforme libre sous licence "libre" (GPL) : <http://opendata.in-cite.net/>. Ce projet, basé sur un outil très répandu de gestion de contenus web (Typo3), fournit une solution complète pour acteur public, qui inclut le catalogage des données, leur diffusion, ainsi que des outils de dialogue avec les réutilisateurs.
- Le logiciel libre CKAN⁶⁷, créée sous l'égide de l'*Open Knowledge Foundation* britannique, se focalise quant à lui sur la partie catalogage et diffusion des jeux de données⁶⁸. Il connaît déjà un certain succès puisqu'il motorise le portail officiel du gouvernement data.gov.uk ainsi que plusieurs autres sites anglais, allemands, canadiens et français.

Enfin, il est aussi possible de faire appel à un acteur tiers qui créera et hébergera le portail sur sa propre infrastructure informatique. Là encore on distingue plusieurs cas de figures possibles :

- Les **plateformes privées** en ligne sont encore rares. En France nous avons la chance d'avoir un projet riche et abouti : Data Publica (www.data-publica.com)
- Si l'offre n'est pas encore disponible, ces tous prochains mois vont voir apparaître les **plateformes publiques multi-acteurs** : elles sont opérées par des acteurs publics pour le compte d'autres acteurs publics. Le projet le plus significatif est le futur portail "Etatlab" de l'APIE qui doit voir le jour en 2012. Certaines collectivités publiques, comme le département de la Gironde, sont en train de réfléchir à un tel modèle.

Enfin, si cette offre n'est pas non plus disponible actuellement, nous savons qu'il existe chez plusieurs grands intégrateurs informatiques, des réflexions sur des **plateformes en "marque blanche"**. Il s'agit d'outils génériques, conçus pour pouvoir s'adapter facilement aux besoins d'une collectivité publique particulière.

► Les portails indépendants et "portails de portails"

Enfin, on assiste également à l'émergence d'initiatives publiques, privées et associatives qui visent à répertorier des sources de données provenant de plusieurs acteurs :

- Le portail français "Etatlab" recensera des données provenant de l'ensemble des acteurs publics nationaux, mais n'exclut apparemment pas d'intégrer également des données provenant, par exemple, de collectivités locales.
- Le site français Data Publica poursuit le but de publier ou référencer n'importe quel type de donnée ouverte, y compris celles venant d'acteurs privés, à travers de nombreux mécanismes de publication et d'échange.
- Le projet CKAN⁶⁹ - pour *Comprehensive Knowledge Archive Network* - de l'Open Knowledge Foundation s'est donné pour mission de favoriser la diffusion et la réutilisation automatique de jeux de données à travers le monde. Il s'agit d'un registre "ouvert" qui référence et mutualise des données ou toute autre source de connaissances. Une application facilite la recherche, le partage et la réutilisation de données.

De son côté, le collectif Regards Citoyens a ouvert en novembre 2010 un portail de recensement des données publiques librement accessibles en France ; nosdonnees.fr. Le site propose pour l'instant un registre libre de jeux de données ainsi qu'une "plateforme collaborative de chasse aux trésors", qui vise à identifier les données publiques disponibles même lorsque celles-ci sont difficiles à identifier. Ce portail s'appuie sur le logiciel CKAN, en l'adaptant au contexte français (langue et jeux de données).

⁶⁷ <http://okfn.org/projects/ckan/>

⁶⁸ <http://knowledgeforge.net/ckan/trac/#Features>

⁶⁹ www.ckan.net/

Bienvenue sur NosDonnees.fr

NosDonnees.fr vise à apporter plus de visibilité aux données publiques librement accessibles à chacun.

Le site est constitué de deux sites :

- un répertoire de jeux de données : centré sur le répertoire GDD, développé par l'Agence Française pour l'Accès aux Données Publiques, et sur les données locales, le portail offre une sélection de données et permet de télécharger les données directement.
- un système informatique de gestion des données : permet la mise à jour, la consultation et la diffusion des données.

Le contenu des articles est soumis à l'approbation de l'Agence Française pour l'Accès aux Données Publiques, par exemple pour valider les données et les rendre disponibles publiquement, ou pour valider les données soumises à l'Agence Française pour l'Accès aux Données Publiques.

Les données déjà recensées

13 jeux de données enregistrés.

Rechercher

Les données par mots-clés

résultats_nictoriaux, fiscalité, communes, élections, import, export, régions

vos dernières modifications

Le contenu des articles est soumis à l'approbation de l'Agence Française pour l'Accès aux Données Publiques, par exemple pour valider les données et les rendre disponibles publiquement, ou pour valider les données soumises à l'Agence Française pour l'Accès aux Données Publiques.

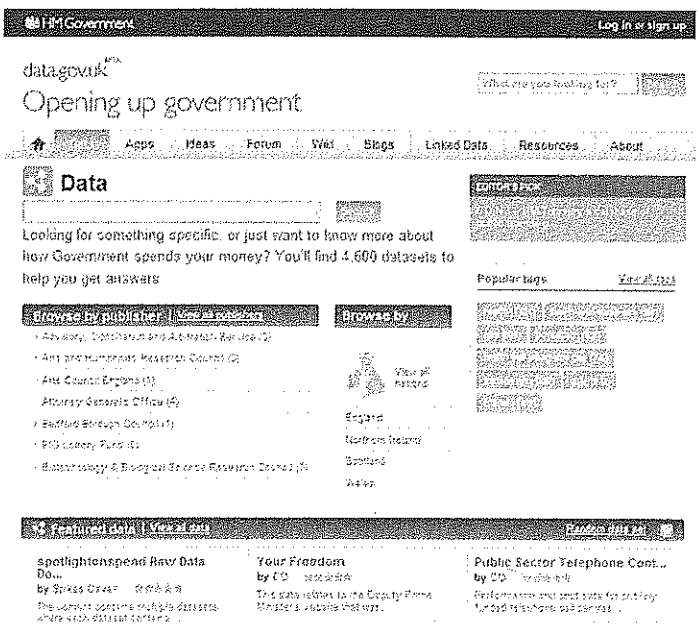
Datalift : un projet de recherche de plateforme d'interconnexion de données

Datalift⁷⁰ est un projet de recherche expérimentale financé par l'agence nationale de la recherche. Son but est de développer une plateforme pour publier et interconnecter des jeux de données sur le web de données. Datalift tout à la fois publie des jeux de données provenant d'un réseau de partenaires et propose un ensemble d'outils facilitant le processus de publication de jeux de données. L'interconnexion des jeux de données va permettre de rapprocher des données qui ne l'étaient pas au départ, pour des usages toujours plus riches.

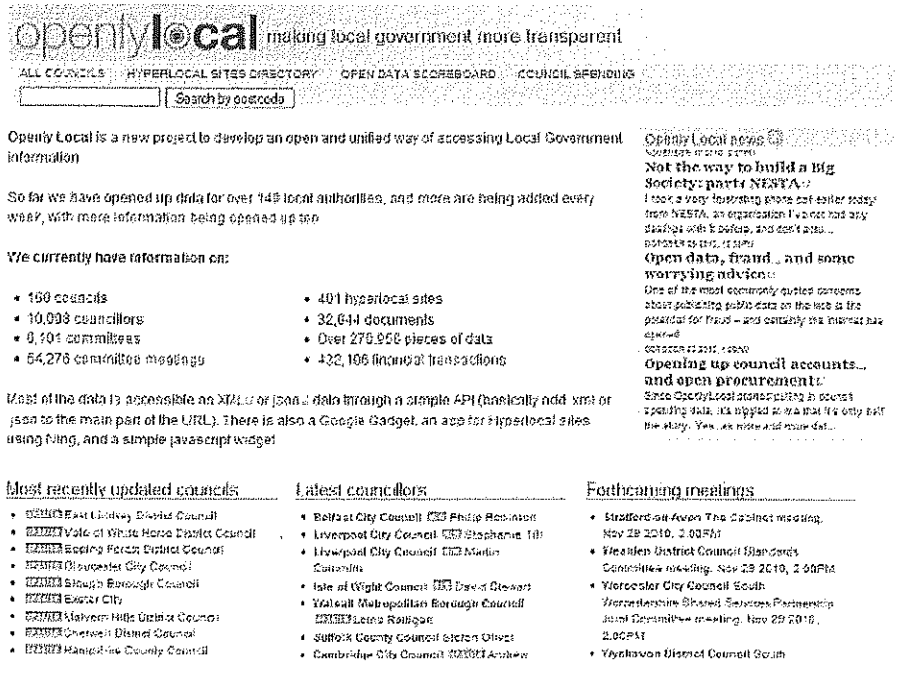
⁷⁰ <http://datalift.org/>

Les portails britanniques de recensement de données publiques

Le portail gouvernemental data.gov.uk dresse la liste des données publiques ouvertes. Actuellement il héberge ou référence environ 3200 bases de données et une cinquantaine d'applications dérivées. Il référence plus d'informations que le portail américain data.gov. Le portail britannique se différencie aussi de son homologue américain en ayant fait le choix de formats standardisés favorisant le développement du web sémantique⁷¹. Les données peuvent être recherchées par thèmes, par mots-clés ("tags"), par sources, par régions géographiques, etc. La gestion du portail est partagée entre le cabinet du Premier ministre et les archives nationales.



Le portail [Openlylocal](http://Openlylocal.com)⁷², porté par l'association Countculture, entend donner une meilleure visibilité et faciliter l'accès aux données publiques locales. Chris Taggart et son équipe mènent un travail de recensement des données existantes au sein des *councils* britanniques (plus de 140 d'entre eux étaient "couverts" au début 2011), auquel s'ajoute un travail de mise en forme des données, pour que celles-ci puissent être facilement réutilisées et croisées.



⁷¹ Cf. Claire Goulard "Les bons ingrédients pour une ouverture des données réussie", Owni - <http://owni.fr/2010/05/31/opendata-12-data-gov-ou-data-gov-uk/>

⁷² <http://openlylocal.com/>

Les défis de l'ouverture des données publiques : entretien avec Henri Verdier

15.02.13

37

Henri Verdier est à la tête de la mission Etalab depuis janvier. Au sein du Secrétariat général pour la modernisation de l'action publique, Etalab coordonne l'action des services de l'Etat et de ses établissements publics pour faciliter la réutilisation la plus large possible de leurs informations publiques. Henri Verdier revient ici sur sa mission et nous présente les projets et défis à venir.

Vous dirigez depuis janvier la mission Etalab au sein du SGMAP. Pouvez-vous nous présenter votre structure ?

Rattaché au Secrétariat général pour la Modernisation de l'Action Publique (SGMAP), placé sous l'autorité du Premier Ministre, Etalab est la mission chargée de l'ouverture des données publiques françaises (politique de données ouvertes / Open data). Etalab coordonne le travail des administrations pour permettre la mise en ligne et la réutilisation libre, facile et gratuite la plus large possible des données publiques produites par l'administration. Elle opère la plateforme nationale www.data.gouv.fr et apporte son appui aux autres entités dotées d'une mission de service public qui souhaitent participer à ce mouvement.

Comment vous situez-vous au sein de la modernisation de l'action publique ?

L'ouverture des données publiques est un mouvement mondial qui est bien souvent un très fort levier de modernisation. Il permet de fonder de nouvelles relations avec les citoyens, basées sur la confiance et la coopération. Il permet de simplifier le fonctionnement des administrations, et souvent de stimuler l'économie.

En rattachant Etalab au SGMAP, le gouvernement a regroupé la DISIC - DSI de l'Etat - la DIMAP et Etalab permettant ainsi de réunir ensemble tous les services qui œuvrent pour la modernisation de l'action publique.

Où en est la politique d'ouverture des données publiques ?

L'ouverture des données publiques a une longue histoire : loi CADA de 1978, décisions du gouvernement Jospin de libérer des « données publiques essentielles », directive européenne « public sector information » (PSI) de 2003 transposée en droit français en 2005, engagement des collectivités locales à partir de 2008, création de la mission Etalab en 2011 ou Cimap du 18 décembre 2012.

Première étape nécessaire, Etalab a réussi ce qui a permis de construire une solution technique permettant d'héberger plus de 350 000 jeux de données publiques avec la mise en place de la plateforme www.data.gouv.fr, de bâtir un processus pour mettre en ligne ces données, et de constituer un réseau de correspondants open data dans l'administration notamment. Nous démarrons maintenant la seconde étape de l'histoire d'Etalab. L'objectif du SGMAP est de faire évoluer le contenu même des politiques publiques grâce aux données, et de mettre plus de données publiques à disposition, dans une optique de qualité des données, afin de susciter des réutilisations toujours

plus nombreuses.

Quelle est votre méthode pour convaincre les administrations d'ouvrir leurs données ?

Vous venez en fait d'employer le mot qui résume notre méthode : la conviction. Les correspondants open data du service public se sont mobilisés de façon remarquable en deux ans, au-delà de leurs responsabilités déjà engagées. En plus de leur mission, certains d'entre eux ont volontairement pris en charge le sujet Open data pour le faire avancer dans leurs administrations. Lorsqu'on a la chance d'avoir des agents du service public qui se mobilisent sur un sujet majeur de modernisation de l'action publique, il semble logique de poursuivre l'engagement des administrations avec la conviction qui est la nôtre.

Les administrations ont globalement compris qu'elles avaient beaucoup à gagner en contribuant à l'ouverture de leurs données publiques : l'open data c'est avant tout un outil formidable pour faire travailler les administrations ensemble, c'est aussi un formidable outil de valorisation du travail des agents publics, qui permet en transparence, d'entrevoir la réalité du travail de l'administration. L'open data c'est du gagnant gagnant pour les administrations.

Quels sont vos principaux défis, projets pour l'année 2013 ?

Le Comité interministériel pour la modernisation de l'action publique du 18 décembre 2012 a déjà fixé à Etalab deux importantes missions :

- organiser une évaluation des redevances existantes ;
- organiser un débat ouvert sur les applications de l'ouverture des données publiques à 6 secteurs prioritaires (santé, transport, environnement, logement, éducation, lisibilité des politiques publiques).

Au-delà de ces priorités, je peux d'ores et déjà vous annoncer notre ambition de mettre à disposition de nouveaux jeux de données publiques, de travailler à la transformation des politiques publiques grâce aux données publiques, et de nous rapprocher des collectivités locales et du monde de la recherche



LICENCE OUVERTE
OPEN LICENCE

Vous pouvez réutiliser « l'Information » rendue disponible par le « Producteur » dans les libertés et les conditions prévues par la présente licence.

LA RÉUTILISATION DE L'INFORMATION DIFFUSÉE SOUS CETTE LICENCE

Le « Producteur » garantit au « Réutilisateur » le droit personnel, non exclusif et gratuit, de réutilisation de « l'Information » soumise à la présente licence, dans le monde entier et pour une durée illimitée, dans les libertés et les conditions exprimées ci-dessous.

VOUS ÊTES LIBRE DE RÉUTILISER « L'INFORMATION » :

- Reproduire, copier, publier et transmettre « l'Information » ;
- Diffuser et redistribuer « l'Information » ;
- Adapter, modifier, extraire et transformer à partir de « l'Information », notamment pour créer des « Informations dérivées » ;
- Exploiter « l'Information » à titre commercial, par exemple en la combinant avec d'autres « Informations », ou en l'incluant dans votre propre produit ou application.

SOUS RÉSERVE DE :

- Mentionner la paternité de « l'Information » : sa source (*a minima* le nom du « Producteur ») et la date de sa dernière mise à jour.

Le « Réutilisateur » peut notamment s'acquitter de cette condition en indiquant un ou des liens hypertextes (URL) renvoyant vers « l'Information » et assurant une mention effective de sa paternité.

Cette mention de paternité ne doit ni conférer un caractère officiel à la réutilisation de « l'Information », ni suggérer une quelconque reconnaissance ou caution par le « Producteur », ou par toute autre entité publique, du « Réutilisateur » ou de sa réutilisation.



RESPONSABILITÉ

« L'Information » est mise à disposition telle que produite ou reçue par le « Producteur », sans autre garantie expresse ou tacite qui n'est pas prévue par la présente licence.

Le « Producteur » garantit qu'il met à disposition gratuitement « l'Information » dans les libertés et les conditions définies par la présente licence. Il ne peut garantir l'absence de défauts ou d'irrégularités éventuellement contenues dans « l'Information ». Il ne garantit pas la fourniture continue de « l'Information ». Il ne peut être tenu pour responsable de toute perte, préjudice ou dommage de quelque sorte causé à des tiers du fait de la réutilisation.

Le « Réutilisateur » est le seul responsable de la réutilisation de « l'Information ». La réutilisation ne doit pas induire en erreur des tiers quant au contenu de « l'Information », sa source et sa date de mise à jour.

DROITS DE PROPRIÉTÉ INTELLECTUELLE

Le « Producteur » garantit que « l'Information » ne contient pas de droits de propriété intellectuelle appartenant à des tiers.

Les éventuels « Droits de propriété intellectuelle » détenus par le « Producteur » sur des documents contenant « l'Information » ne font pas obstacle à la libre réutilisation de « l'Information ». Lorsque le « Producteur » détient des « Droits de propriété intellectuelle » sur des documents qui contiennent « l'Information », il les cède de façon non exclusive, à titre gracieux, pour le monde entier et pour toute la durée des « Droits de propriété intellectuelle », au « Réutilisateur » qui peut en faire tout usage conformément aux libertés et aux conditions définies par la présente licence.

COMPATIBILITÉ DE LA PRÉSENTE LICENCE

Pour faciliter la réutilisation des « Informations », cette licence a été conçue pour être compatible avec toute licence libre qui exige *a minima* la mention de paternité. Elle est notamment compatible avec les licences « Open Government Licence » (OGL) du Royaume-Uni, « Creative Commons Attribution 2.0 » (CC-BY 2.0) de Creative Commons et « Open Data Commons Attribution » (ODC-BY) de l'Open Knowledge Foundation.

DROIT APPLICABLE

La présente licence est régie par le droit français.



DÉFINITIONS

DROITS DE PROPRIÉTÉ INTELLECTUELLE*

Il s'agit des droits identifiés comme tels par le Code de la propriété intellectuelle (droit d'auteur, droits voisins au droit d'auteur, droit sui generis des bases de données).

INFORMATION*

Il s'agit des données ou des informations proposées à la réutilisation dans les libertés et les conditions de cette licence.

INFORMATIONS DÉRIVÉES*

Il s'agit des nouvelles données ou informations qui ont été créés soit directement à partir « d'Informations », soit à partir d'une combinaison « d'Informations » et d'autres données ou informations qui ne seraient pas soumises à cette licence.

PRODUCTEUR*

Il s'agit de l'entité qui produit « l'Information » et l'ouvre à la réutilisation dans les libertés et les conditions prévues par cette licence.

RÉUTILISATEUR*

Il s'agit de toute personne physique ou morale qui réutilise « l'Information » conformément aux libertés et aux conditions de cette licence.

À PROPOS DE LA LICENCE OUVERTE

Etalab est la mission chargée sous l'autorité du Premier ministre d'ouvrir le plus grand nombre de données publiques des administrations de l'Etat et de ses établissements publics. Elle a réalisé la Licence Ouverte pour faciliter la réutilisation libre et gratuite de ces informations publiques, telles que définies par l'article 10 de la loi n°78-753 du 17 juillet 1978.

Dans le cadre de leurs missions de service public, les administrations produisent ou reçoivent des informations publiques qui peuvent être réutilisées par toute personne physique ou morale à d'autres fins que celles de la mission de service public.

Ne sont pas des informations publiques au sens de la loi du 17 juillet 1978 les informations contenues dans des documents dont la communication ne constitue pas un droit (en application de la loi du 17 juillet 1978 ou d'autres dispositions législatives, sauf si ces informations font l'objet d'une diffusion publique), celles contenues dans des documents produits ou reçus par les administrations dans l'exercice d'une mission de service public à caractère industriel ou commercial, et celles contenues dans des documents sur lesquels des tiers détiennent des droits de propriété intellectuelle.

Ne sont également pas des informations publiques susceptibles d'être réutilisées celles qui contiennent des données à caractère personnel, sauf lorsque les personnes intéressées y ont consenti, ou lorsqu'elles ont fait l'objet d'une anonymisation par l'administration, ou lorsqu'une disposition légale ou réglementaire le permet (dans ces trois cas, la réutilisation est subordonnée au respect de la loi n°78-17 du 6 janvier 1978).

Cette licence est une version 1.0 de la Licence Ouverte. Etalab se réserve la faculté de proposer de nouvelles versions de la Licence Ouverte. Cependant, les réutilisateurs pourront continuer à réutiliser les informations disponibles sous cette licence s'ils le souhaitent.

Les décisions sur l'open data du CIMAP du 18 décembre 2013

Jeudi 18 décembre 2013, le quatrième Comité interministériel pour la modernisation de l'action publique a permis au gouvernement de prendre plusieurs décisions sur l'open data, notamment à la suite du travail sur les redevances initié à l'occasion du rapport Trojette. Le relevé de décisions du CIMAP, disponible sur le site du gouvernement, précise, concernant l'open data :

Open data et Gouvernement ouvert : faire de l'administration une plateforme pour l'innovation économique et sociale

Le Gouvernement réaffirme son engagement dans le projet d'open data et de Gouvernement ouvert, levier de confiance démocratique, de stimulation de la croissance, d'innovation et de modernisation de l'action publique.

Le Gouvernement ouvre ce jour un nouveau portail d'open data, à l'adresse data.gouv.fr

Ce nouveau portail ouvre la voie d'une coproduction des données d'intérêt public. Il rassemble des données publiques et des données produites par les forces vives de la société, et met en avant les réutilisations et enrichissements de ces données apportés par la société civile (citoyens, associations...), les innovateurs, les chercheurs, les journalistes, etc.

Pour accélérer l'ouverture et le partage des données publiques, le Gouvernement clarifie sa doctrine en matière de redevances sur la réutilisation de données publiques :

Décision n°24 : Le Gouvernement réaffirme le principe de gratuité de la réutilisation des données publiques et décide de ne plus autoriser la création de nouvelle redevance.

Décision n°25 : Sur la base du rapport remis par M. Trojette relatif au modèle économique des redevances de réutilisation des données publiques et après examen des différentes catégories de données publiques soumises à redevance, le Gouvernement décide de supprimer plusieurs redevances. Les redevances supprimées sont celles perçues sur :

- les données relatives aux dotations globales de fonctionnement versées aux collectivités locales par la direction générale des collectivités locales (ministère de l'Intérieur) ;
- les données comptables brutes gérées par la direction générale des finances publiques – balance des comptes définitifs des collectivités et établissements publics locaux recensés à compter de l'arrêté définitif des comptes 2013. Les fiches de situation financière des collectivités locales restent consultables sur le site collectivites-locales.gouv.fr
- le parcellaire agricole de l'Agence de services et de paiement (ministère de l'agriculture, de l'agroalimentaire et de la forêt) ;
- les données cadastrales de l'Institut national de l'origine et de la qualité (ministère de l'agriculture, de l'agroalimentaire et de la forêt) ;
- les données brutes diffusées par l'Institut français du cheval et de l'équitation ;
- les données relatives aux adresses postales des établissements scolaires (ministère de l'éducation nationale) ;
- les données de l'Office national d'information sur les enseignements et les professions – ONISEP (ministère de l'éducation nationale) ;

- les données de la CADA

- les redevances sur la réutilisation professionnelle des informations suivantes diffusées par la DILA : ASSOCIATIONS, DEPOTS DES COMPTES DES ASSOCIATIONS, JORF, LEGI, KALI, JADE, CONSTIT, CASS-INCA, CAPP, CNIL, CIRCULAIRES.

Par ailleurs, le ministère de l'économie et des finances fournira en open data, sur une base hebdomadaire, la base des prix des carburants.

Décision n°26 : Le Gouvernement précise sa doctrine en matière d'exceptions au principe de gratuité :

- - aucune redevance ne saurait être exigée sur des données résultant des missions de service public des administrations générales. Il est demandé aux administrations d'accroître leurs efforts en matière de publication de données leur permettant de mener à bien leurs missions, de motiver leurs décisions, de mesurer leurs activités ou leur bilan. Le site data.gouv.fr, est en mesure d'accueillir ces données ;
- - le respect du secret statistique et donc l'anonymisation des informations pouvant éventuellement permettre d'identifier des personnes est un principe central du droit français. Certaines redevances ont été instaurées pour financer cette tâche, qui est indispensable avant ouverture de la donnée. Le Gouvernement demande au SGMAP d'apporter tout le soutien possible aux services publics concernés, afin d'identifier les moyens de réduire ces coûts sans dégrader la qualité de cette anonymisation ;
- - enfin, le Gouvernement réaffirme que les opérateurs dont la mission même est de produire des données doivent rechercher des modèles économiques leur permettant de faire face à un paysage économique en profonde reconstitution. Conformément aux conclusions du rapport Trojette, il leur demande d'engager, dans les meilleurs délais, avec l'appui du SGMAP et du ministère du Budget, une réflexion sur les évolutions de leurs modèles économiques. Il leur demande de rechercher des modèles stimulant l'innovation autour de leurs données, favorables aux entrepreneurs innovants, et soutenables à l'heure de l'économie numérique, de la production de nombreuses données par les citoyens eux-mêmes, et des stratégies de plateformes.

Afin de mener cette transition - qui vise à, maintenir et améliorer la qualité des informations publiques, à garantir la pérennité du service public et à développer des stratégies numériques pérennes, efficaces et innovantes, **le Gouvernement demande au SGMAP d'accompagner au mieux ces opérateurs, tel l'IGN (Institut géographique national), le SHOM (Service hydrographique et océanographique de la marine) ou Météo France, en lien avec les ministères de tutelle pour définir avec eux des trajectoires d'investissement dans l'innovation compatibles avec le redressement des finances publiques.**

ÉPREUVE N° 16