



HAL
open science

Les limites de l'open data

Lucie Cluzel-Metayer

► **To cite this version:**

Lucie Cluzel-Metayer. Les limites de l'open data. Actualité juridique Droit administratif, 2016, 2. <hal-03455250>

HAL Id: hal-03455250

<https://hal.science/hal-03455250v1>

Submitted on 29 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

LES LIMITES DE L'OPEN DATA

PAR LUCIE CLUZEL-METAYER

PROFESSEUR DE DROIT PUBLIC, UNIVERSITÉ DE LORRAINE, IRENEE - CHERCHEUR ASSOCIÉ AU CERSA

L'ESSENTIEL

L'évolution du cadre juridique de l'*open data* est, enfin, à l'ordre du jour. Des débats qui l'entourent ressortent certaines questions relatives au champ de l'ouverture - quelles données sont concernées? -, mais aussi aux modalités de réutilisation des données. Dans un contexte où la valeur économique et sociale de la donnée est déterminante, des tensions entre droit à l'information et respect de la vie privée, ou encore entre droit à l'information et respect des droits de propriété, traversent un droit de l'*open data* en pleine construction.

Disons-le d'emblée pour ne pas clore ce dossier sur une note exagérément négative: les limites de l'*open data* ne sont rien au regard de son potentiel stratégique, démocratique et économique (C. Bouchoux, *Refonder le droit à l'information publique à l'heure du numérique: un enjeu citoyen, une opportunité stratégique*, Rapport fait au nom de la MCI, Sénat, n° 589, 2014). Les freins qu'il s'agit d'explorer sont juste l'expression d'un droit qui, intervenant dans un contexte incertain, reflète les hésitations du législateur, tiraillé entre des intérêts divergents.

L'*open data* donne indiscutablement de nouveaux moyens de contrôle de l'action publique aux citoyens et participe en cela au renouvellement de la légitimité démocratique; levier économique, il offre en outre de formidables opportunités d'exploitation de gisements de données dans le but de créer des services innovants. Il est donc question d'en favoriser l'essor, voire de consacrer un « droit à l'information d'intérêt public » (C. Paul et C. Feral-Schuhl, *Le droit et les libertés à l'âge du numérique*, rapport d'information, 2015, p. 20), allant bien au-delà du droit d'accès ponctuel « à la demande », et même du droit à réutilisation des données, tels que consacrés par la loi du 17 juillet 1978 modifiée. Mais passer à une culture de la donnée, c'est-à-dire « basculer d'une culture de la détention de l'information à une culture de l'ouverture et de la diffusion » (Conseil national du numérique [CNum], *Ambition numérique, Pour une politique française et européenne de la transition numérique*, 2015, p. 145), correspond à un véritable bouleversement auquel les administrations ne sont pas nécessairement préparées. Tandis que le législateur s'apprête à généraliser le principe, déjà consacré pour les collectivités de plus de 3 500 habitants (CGCT, art. L. 1112-23, issu de la loi NOTRe n° 2015-991 du 7 août 2015, art. 106), de l'ouverture des données « par défaut », des difficultés de plusieurs ordres risquent d'en contrarier l'effectivité. D'abord, bien que peu mentionné, le frein le plus évident est certainement d'ordre politique et psychologique: pour les administrations, en particulier pour les collectivités territoriales, livrer ses données signifie s'exposer à la critique. Abandonner la maîtrise et l'expertise des données, en faire des « communs »¹, peut en effet être perçu comme une perte de pouvoir risquée, surtout en période

électorale (*L'open data en collectivité*, CEREMA, 2015, p. 20). Ensuite, des difficultés techniques, liées notamment au choix délicat des formats de données ou encore à leur actualisation, ne sont pas réglées. Enfin, des limites juridiques, qui correspondent pour la plupart à l'expression d'enjeux économiques, entravent le plein essor de l'*open data*. Ce sont précisément ces dernières limites qui vont retenir l'attention ici.

Alors que la politique d'*open data* est, en France, en plein essor², la fragilité du cadre juridique sur lequel elle se fonde peut être considérée, en soi, comme un frein pour les administrations qui s'interrogent légitimement sur les risques potentiellement induits par l'ouverture de leurs données. L'*open data* repose encore aujourd'hui sur la loi du 17 juillet 1978, modifiée par la loi n° 2000-321 du 12 avril 2000 et surtout par l'ordonnance n° 2005-650 du 6 juin 2005 qui y introduit un nouveau chapitre portant sur la « réutilisation des informations publiques ». Le dispositif se caractérise par une certaine prudence, puisque l'information publique est encore liée au document administratif qui en est le support et que sa diffusion reste à la décision de l'administration. Mais la transparence est devenue une question stratégique, porteuse d'enjeux économiques tels que le cadre ainsi élaboré ne tient plus. L'année 2015 marque ainsi l'adoption de quatre lois: loi n° 2015-990 du 6 août 2015 dite « loi Macron », art. 4 concernant l'ouverture des données de transport; loi NOTRe n° 2015-991 sur les données des collectivités territoriales, art. 106; loi n° XXX du XXX relative à la gratuité et aux modalités de la réutilisation des informations du secteur public, transposant la directive 2013/36/UE du 26 juin 2013; loi n° XXX du XXX de modernisation de notre système de santé, qui promet de bouleverser l'*open data* dans le secteur de la santé. Sans oublier le dépôt du projet de loi sur la République numérique, porté par

1 - « Les communs désignent l'activité des communautés qui s'organisent et se régulent pour protéger et faire fructifier des ressources matérielles ou immatérielles, en marge des régimes de propriété publique ou privée » (CNum, *Ambition numérique*, p. 276).

2 - Signe de son investissement en la matière, la France a été choisie pour prendre la présidence de l'*Open Government Partnership* en 2016.

●●●
Axelle Lemaire, soumis au débat public avant son examen en conseil des ministres. Ces cinq textes sont indiscutablement porteurs d'évolutions significatives.

Tout en saluant le mouvement général d'ouverture dont ils sont les promoteurs, on ne peut s'empêcher de regretter un tel morcellement législatif, qui ne contribue pas à la clarté du débat sur un sujet en soi complexe. Pour schématiser, les projets de loi transversaux portent, d'une part, sur le principe même de mise à disposition des données publiques (projet de loi République numérique) et, d'autre part, sur les modalités de réutilisation de ces données (loi Réutilisation des informations du secteur public), ce qui correspond aux deux étapes au cours desquelles le cadre juridique actuel de l'*open data* s'est forgé. Celui-ci repose en effet sur l'articulation de deux régimes distincts, celui de la mise à disposition des informations par les administrations (en vertu du régime d'accès aux documents administratifs) et celui de leur éventuelle réutilisation; ces deux régimes correspondent aux deux premiers chapitres de la loi CADA (commission d'accès aux documents administratifs), laquelle doit être lue à la lumière de la loi du 6 janvier 1978 (dite «informatique et libertés»), mais aussi des dispositions du code du patrimoine organisant le régime d'accès aux archives publiques.

Le dispositif originel, tel que modifié, est ainsi complexe et l'enchevêtrement textuel qui se profile ne laisse présager aucune clarification. Aussi, des débats³ et multiples rapports qui alimentent cette période de transition juridique, il ressort des inquiétudes, des questions tenant tantôt au principe même de la mise en *open data* – c'est alors le champ de l'ouverture qui est interrogé –, tantôt aux modalités de réutilisation des données, une fois le principe de leur mise à disposition acquis, dans un contexte où la valeur économique et sociale de la donnée ne saurait être négligée (M. Lévy et J.-P. Jouyet, *L'économie de l'immatériel, La croissance de demain*, rapport au ministre de l'économie, des finances et de l'industrie, 2006).

Des tensions entre droit à l'information et respect de la vie privée, ou encore entre droit à l'information et respect des droits de propriété, marquent ainsi le cadre juridique de l'*open data*. Le législateur reste encore prudent sur le principe même de l'ouverture des données et hésitant quant aux conditions de leur réutilisation.

I - LA FRAGILITÉ DU PRINCIPE D'OUVREMENT PAR DÉFAUT

La notion de « donnée publique », produite et utilisée par les administrations, fait exploser le cadre juridique posé en 1978. Du droit à obtenir un « document administratif », sous certaines conditions, auquel s'est ajoutée la possibilité d'avoir accès et de réutiliser l'« information publique » en 2005, on s'achemine aujourd'hui vers la reconnaissance d'un véritable droit d'accès aux « données publiques ». La consécration du

principe d'ouverture par défaut des données publiques qui s'imposerait aux administrations, suppose l'adoption d'un texte de valeur législative, puisqu'il renverserait la règle actuelle, simplement incitative, de diffusion (Conseil d'Etat, *Le numérique et les droits fondamentaux*, rapport annuel 2014, p. 309). Dans ce contexte d'ouverture par défaut, la question doit être posée de savoir si les exceptions prévues par la loi CADA perdurent. Alors que le concept de « données privées » (art. 10 de la loi) semble devoir s'entendre de manière plus stricte afin d'élargir le champ de l'*open data*, le risque de « ré-identification » après anonymisation des données personnelles continue de soulever une difficulté au regard de la protection de la vie privée (art. 6 et 13).

A. Des données « privées » aux « données d'intérêt général »

Bien que la charte pour l'ouverture des données publiques, signée le 18 juin 2013 par les Etats membres du G8 dont la France fait partie, ait érigé le principe de l'ouverture par défaut en principe premier, un certain nombre d'exceptions demeurent. En conformité avec la directive 2013/37/UE, certaines données ne sont pas considérées comme entrant dans le champ des informations publiques communicables, au sens de la loi de 1978. Le projet de loi pour une République numérique entend cependant restreindre le champ de ces exceptions.

Aux termes de l'article 10 de la loi CADA, ne sont pas considérées comme des informations publiques les informations contenues dans des documents dont la communication ne constitue pas un droit: les documents préparatoires à une décision administrative en cours d'élaboration, les avis du Conseil d'Etat et des juridictions administratives, les documents qui porteraient atteinte à certains secrets (défense nationale, politique extérieure, sûreté de l'Etat, monnaie...) en font partie (art. 6). De même, les informations produites ou reçues par les établissements gérant un service public industriel et commercial sont exclues du champ de l'ouverture, tout comme celles sur lesquelles les tiers détiennent des droits de propriété intellectuelle.

Le projet de loi pour une République numérique pose le principe de l'ouverture obligatoire des données en modifiant l'article 7 de la loi CADA (la formule « les administrations [...] peuvent rendre publics les documents » devient « les administrations [...] ~~diffusent~~ **rendent public** publiquement en ligne [...] les documents »), mais sous réserve des dispositions de l'article 6⁴, ce qui signifie que les exceptions tenant aux secrets, précédemment évoquées, sont toujours d'actualité. De même, les informations contenues dans des documents sur lesquels des tiers détiennent des droits de propriété intellectuelle demeurent dans le champ de l'exception. En revanche, le projet entend lever le voile sur les données détenues par des personnes gérant un service public industriel et commercial ainsi que sur celles détenues par les délégataires de service public. Cela signifie que toute donnée reçue ou produite dans le cadre d'un service public devra être mise en *open data*. Veolia, Velib', l'IGN, l'INSEE devront ainsi rallier la RATP et la SNCF, déjà contraintes de mettre en *open data* leurs données par la loi « Macron » du 6 août 2015 (art. 4, sur les transports).

³ - V. la concertation publique mise en place dans le cadre de la discussion du projet de loi pour une République numérique, grâce à laquelle tout internaute pouvait produire des observations pendant deux mois, jusqu'au 18 octobre 2015.

Parce que ces données entrent dans la catégorie des « données d'intérêt général », le projet de loi entend imposer au délégataire de fournir à l'autorité administrative « dans un standard ouvert et aisément réutilisable, les données et bases de données collectées ou produites à l'occasion de l'exploitation du service public » (art. 8 du projet de loi). Le même article prévoit par ailleurs que, sauf stipulation contraire inscrite dans la convention de délégation, le délégataire autorise l'administration à « extraire et exploiter librement tout ou partie de ces données et bases de données, notamment en vue de leur mise à disposition à titre gratuit à des fins de réutilisation à titre gratuit ou onéreux ». En l'état actuel du droit, l'absence d'inscription de la mise à disposition des données dans les conventions est indiscutablement un frein à l'obtention, pour la personne publique, des données détenues par ses cocontractants, lesquels préfèrent en conserver la maîtrise, allant même jusqu'à créer leur propre plate-forme d'*open data*, sans prévoir de lien avec celle des collectivités et sans utiliser des licences identiques, ce qui entrave l'interopérabilité et en complique nécessairement l'usage (v. la plate-forme créée par l'entreprise JC. Decaux pour les données « Vélos en libre service » : <https://developper.jcdecaux.com/#/home>).

Surmonter cet écueil par la création d'un nouveau statut appliqué aux « données d'intérêt général », détenues par les SPA mais aussi par les SPIC, laisse cependant perplexe. D'abord pour des questions de définition : le point de savoir ce qu'est une « donnée d'intérêt général » n'est pas abordé dans le projet de loi. Tout au plus peut-on déduire du texte qu'il s'agit de données reçues ou produites par le délégataire de service public et par les structures ayant reçu des subventions publiques. Il semblerait donc que les données d'intérêt général renvoient essentiellement à un critère organique et ne soient dès lors pas synonymes de « données de référence » ; celles-ci semblent correspondre à un sous-ensemble des « données d'intérêt général », visant plus précisément les données qui sont susceptibles de faire l'objet d'« une utilisation fréquente par un grand nombre d'acteurs tant publics que privés et dont la qualité, en termes notamment de précision, de fréquence de mise à jour ou d'accessibilité, est essentielle pour ces utilisations » (art. 7). Ces données de référence, que l'on appelle aussi « données pivot », correspondent à des « briques de base » nécessaires à la construction de services autour des données qu'il s'agit de relier entre elles pour les enrichir (« *Linked open data* »). Le projet *Open Law* porté par la Direction de l'information légale et administrative s'inscrit dans cette démarche, en créant un « réseau de données liées » qui vise à produire des services innovants autour des données juridiques ouvertes. ~~Le régime des données d'intérêt général reste, ensuite, assez obscur. Selon qu'elles sont émises dans le cadre d'un service public ou dans celui d'une activité subventionnée, ce régime est différent : dans le premier cas, elles sont obligatoirement mises en *open data*, alors que dans le second, ce n'est qu'une faculté (art. 9). Aussi, parmi ces données, les « données de référence » seraient l'objet d'un traitement particulier, puisqu'il est question de les confier à un « service public de la donnée », probablement créé à partir de la fusion de la CNIL et de la CADA.]~~

↳ Parmi les « données d'intérêt général », un sort particulier est d'ailleurs réservé à ces « données de référence » dans la mesure où leur mise à disposition et leur publication, en vue d'en faciliter la réutilisation, constituent « une mission de service public » relevant de l'Etat.

A l'heure où le principe de la soumission des SPIC et des délégataires de service public à l'*open data* semble acquis, une clarification sémantique et juridique des données qu'ils devront effectivement livrer est sans aucun doute nécessaire. Le champ des « données privées » et des « données d'intérêt général » mérite d'être précisé pour sécuriser la pratique des acteurs. La question de la protection des données personnelles suscite peut-être encore davantage d'inquiétudes.

B. Les données personnelles face au risque de ré-identification

En principe, l'*open data* exclut toute diffusion de données à caractère personnel : la protection de la vie privée prime en effet sur l'exigence d'information publique. De la combinaison des articles 6, 7 et 13 de la loi CADA avec les dispositions de la loi Informatique et libertés du 6 janvier 1978 modifiée, résulte donc une interdiction de principe, assortie de trois exceptions. L'interdiction de mise en ligne de données personnelles par l'administration et de réutilisation par des tiers peut en effet être levée si l'intéressé consent à leur diffusion, s'il existe une obligation légale de publication ou s'il est procédé à l'anonymisation des données publiées. Compte tenu des potentialités de l'anonymisation, qui permet aux administrations de diffuser des bases entières de données sans avoir à recueillir au préalable le consentement des intéressés, le procédé est très prisé. Il n'est cependant pas exempt de risques dans la mesure où, à l'heure actuelle, aucun dispositif d'anonymisation n'est infaillible.

“ En principe, l'*open data* exclut toute diffusion de données à caractère personnel : la protection de la vie privée prime ”

Testées à l'aune des critères posés par le G29 (avis du Groupe européen des autorités de protection du 10 avril 2014 sur les techniques d'anonymisation), à savoir l'individualisation (est-il toujours possible d'isoler l'individu ?), la corrélation (est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?) et l'inférence (peut-on déduire de l'information sur un individu ?), les techniques classiques d'anonymisation présentent, en effet, des risques de ré-identification. Qu'elles aient pour objet de transformer les données pour qu'elles ne se réfèrent plus à une personne réelle (« pseudonymisation » et « masquage ») ou de généraliser les données pour qu'elles ne soient plus spécifiques à une personne, mais communes à un ensemble de personnes (« agrégation »), aucune de ces techniques n'est infaillible. En définitive, tout dépend du niveau de sécurité des outils utilisés. Ainsi, plus le niveau d'agrégation est élevé et plus les outils de codage sont poussés, moins il y a de risque de ré-identification.

Surtout, la possibilité d'établir des liens entre les données anonymisées constitue la principale faiblesse de l'ensemble de ces dispositifs, puisqu'il suffit de croiser les informations pour qu'elles redeviennent identifiantes : à titre d'exemple, il a été prouvé que 89 % des patients

...

•••

d'un hôpital peuvent être identifiés à partir de leur code postal, du mois et de leur année de naissance, de leur sexe et du mois de sortie de l'hôpital en question (P.-L. Bras, A. Loth, *Rapport sur la gouvernance et l'utilisation des données de santé*, 2013, p. 27). De même, la technique du « carroyage » utilisée par l'INSEE, qui consistait à diviser le territoire en carrés de 200 mètres de côté pour calculer l'imposition moyenne des contribuables, a été abandonnée dès lors que les résultats mis en ligne permettaient d'identifier les contribuables résidant dans un espace faiblement peuplé, ne comptant qu'un seul foyer fiscal (G. Gorce, F. Pillet, *L'open data et la protection de la vie privée*, rapport d'information, Sénat, n° 469, avr. 2014, p. 49).

Le risque de ré-identification par recouplement de données, dans un contexte d'explosion quantitative – « Big Data » – et de diffusion sans limite, est ainsi avéré. Les fuites peuvent alors être lourdes de conséquences, pour les citoyens, bien sûr, mais aussi pour les administrations. Leur responsabilité pourrait en effet être engagée pour négligence dans l'anonymisation, par un citoyen victime d'une ré-identification rendue possible par une mise en *open data* mal maîtrisée.

Ce risque serait néanmoins considérablement limité par l'adoption de techniques d'anonymisation *ab initio* – ou « privacy by design » – des bases de données des administrations, à l'instar de la base SNIIRAM (système national d'informations inter-régimes de l'Assurance maladie), conçue dès l'origine dans le souci de protéger l'anonymat des bénéficiaires grâce au codage initial du numéro de sécurité sociale. Mais cela implique un changement profond de l'approche des administrations et une mobilisation technique, financière et humaine importante. Le principe de l'ouverture par défaut promet d'être ainsi complexe à mettre en œuvre pour les administrations. Il en va de même du principe de gratuité dans la réutilisation des données.

II - LES HÉSITATIONS TENANT AUX CONDITIONS DE RÉUTILISATION

L'ordonnance du 6 juin 2005, qui transpose la directive européenne du 17 novembre 2003 et modifie la loi du 17 juillet 1978, crée un droit de réutilisation des données publiques. Au fondement d'un tel principe se trouve l'idée que ces données correspondent à des « communs », qui doivent être régulées en marge des régimes de propriété publique ou privée (CNum, p. 276, préc.), pour permettre aux citoyens de s'en emparer et de produire des services innovants tels que la création d'une carte interactive des lieux accessibles aux personnes handicapées (Rennes) ou encore le recensement de la biodiversité (Montpellier). Afin d'exploiter tout le potentiel économique et social de l'*open data*, le gouvernement semble aujourd'hui favorable à la consécration du principe de gratuité dans la réutilisation des données (loi réutilisation des informations du secteur public, art. 4). Mais, tout comme le principe de gratuité se heurte à la résistance des administrations qui l'assimilent à la disparition d'importantes ressources, le principe de libre réutilisation achoppe sur les droits d'auteur, ce que le juge administratif n'a pas manqué de relever.

A. La résistance à la gratuité

Le principe de gratuité de la réutilisation des données est encore aujourd'hui davantage un idéal qu'un principe prescriptif. L'article 15 de la loi de 1978, en prévoyant que la réutilisation « peut donner lieu au versement de redevances », signifie certainement, en creux, que la gratuité doit être le principe ; le décret du 26 mai 2011, qui dispose que la liste des informations soumises à redevance doit être prévue par décret, fait également du paiement l'exception. Mais ce décret ne concerne que les informations détenues par l'Etat et ses établissements publics administratifs. Le principe de gratuité n'a ainsi jamais été positivement consacré de manière générale par un texte. Au titre de l'article 15, la perception de redevances est possible pour couvrir les coûts d'anonymisation, les coûts de collecte, de production et de mise à disposition des informations, mais aussi pour permettre une rémunération raisonnable des investissements de l'administration comprenant, « le cas échéant, une part au titre des droits de propriété intellectuelle » ; en outre, l'instauration de redevances est prévue lorsque la réutilisation porte sur des informations issues de la numérisation des fonds de bibliothèques, musées et archives. Ainsi que l'a souligné le Conseil d'Etat dans un avis du 23 juillet 2015, ce dispositif est conforme à la directive PSI du 27 juin 2013, qui autorise les redevances excédant les coûts marginaux.

La loi Réutilisation des informations – dont l'intitulé est en lui-même emblématique puisqu'il est « relatif à la gratuité et aux modalités de la réutilisation des informations du secteur public » – modifie l'article 15 de la loi de 1978, dont le premier phrase affirme que « La réutilisation d'informations publiques est gratuite ». La suite de l'article révèle néanmoins que la gratuité ne s'impose pas avec tant d'évidence : tout en étendant l'encadrement prévu par le décret du 26 mai 2011 à l'ensemble des redevances perçues par les collectivités publiques, il en maintient le principe, lorsque les administrations « sont tenues de couvrir par des recettes propres une part substantielle des coûts liés à l'accomplissement de leurs missions de service public ». Il faut dire que les motivations avancées par les administrations pour assujettir la réutilisation à redevance sont nombreuses. Des motifs budgétaires, d'abord : dans un contexte financier contraint, l'objectif d'accroissement des ressources propres est évidemment primordial et, à court terme, le système des redevances est nécessairement plus avantageux que la gratuité. Des préoccupations qualitatives, telles que la recherche de la valorisation du patrimoine immatériel public et l'amélioration de la qualité de l'offre de données ainsi que des systèmes de mise à disposition de ces données, sont également avancées. Les administrations considèrent en outre que la mise en place d'une redevance régule efficacement les demandes de réutilisation, leur évitant d'être submergées par les requêtes. Enfin, des considérations liées à la préservation d'un « écosystème » stable freinent la gratuité : d'une part, sur un marché déjà concurrentiel, injecter des données gratuites risque de déstabiliser les produits payants des entreprises en place sans nécessairement stimuler la concurrence ; d'autre part, des voix s'élèvent contre la mise à disposition gratuite des données à des entreprises, bien souvent des multinationales

« La
en affirmant
que

en position dominante (Google, par exemple), qui en retirent de considérables profits, abandonnant au contribuable le soin de les financer (M.-A. Trojette, *Rapport au Premier ministre sur l'ouverture des données publiques*, 2013, p. 36).

Les partisans du « gratuit » contrent ces arguments en expliquant que les données des administrations sont produites à l'occasion de leur mission de service public, financées par l'impôt, et qu'elles sont par conséquent déjà payées. En outre, la mise en place de redevances ne permet pas la réutilisation non discriminatoire des données puisqu'elle en interdit concrètement l'accès à certains utilisateurs. Enfin, et peut-être surtout, le retour économique à moyen et long terme, fondé sur le potentiel d'innovation de l'ouverture gratuite des données, ne saurait être négligé.

Des solutions de compromis, proposant de moduler les redevances suivant le type de réutilisateur et la valeur ajoutée de l'offre, émergent aujourd'hui. Le rapport Trojette propose ainsi de ménager un espace de gratuité pour les citoyens et les *start-up*, et un espace payant pour les entreprises « consolidées », le tarif étant fonction de la valeur ajoutée de la plateforme (volume de transactions et des accès concurrents, « fraîcheur » des données, type de licence...). Il est préconisé d'articuler ce dispositif à un système de double licence pour les réutilisations soumises à redevances : d'une part, le recours à une licence ouverte (par exemple, celle d'etalab sur le site data.gouv.fr, qui autorise la reproduction, la redistribution, l'adaptation et l'exploitation commerciale des données, sous réserve de mentionner expressément leur paternité) et, d'autre part, l'utilisation d'une licence de « partage à l'identique », ou « share-alike » (licence ODbL), qui impose aux réutilisateurs, notamment ceux en position dominante, de ne pas « ré-enfermer » les données libérées (phénomène d'« enclosure »), en les obligeant à utiliser la même licence pour maintenir ouverte la base de données. Ce schéma contributif, associant administrations et partenaires économiques, a pour but de protéger les utilisateurs finaux que sont les citoyens, en faisant des données publiques de véritables « biens communs ». Mais des questions de propriété intellectuelle demeurent.

B. Les droits de propriété intellectuelle

Les droits de propriété intellectuelle, pouvant emporter droits exclusifs sur la réutilisation des données, constituent des freins naturels à l'*open data*. De la confrontation entre logique d'ouverture et protection de ces droits, émergent deux questions sensiblement différentes.

D'abord, la question doit être posée de savoir si les agents publics peuvent être considérés comme des tiers détenteurs de droits de propriété intellectuelle sur certains documents, au sens de l'article 10 de la loi de 1978, empêchant leur diffusion et leur réutilisation. S'ils sont bel et bien titulaires de droits d'auteur (CPI, art. L. 111-1), le droit d'exploitation d'une œuvre créée par un agent dans l'exercice de ses fonctions est cédé de plein droit à son administration (CPI, art. L. 131-3-1). Mais pour l'exploitation commerciale, l'administration ne dispose que d'un droit de préférence pour en disposer,

à charge pour elle de prévoir ensuite un mécanisme d'intéressement de l'agent aux produits tirés de l'exploitation de l'œuvre (art. L. 131-3-3). En l'absence de décret en Conseil d'Etat permettant la mise en œuvre de cette disposition, la question demeure de savoir quel est exactement le statut de ces bases de données. C'est ainsi que, dans le doute, la base Mérimée du ministère de la Culture, regroupant des données sur les monuments français, a été expurgée de nombreux éléments rédigés par les agents du ministère.

Une autre question concerne la possibilité pour l'administration d'opposer elle-même des droits de propriété intellectuelle. En effet, une base de données peut être assimilée à une œuvre originale ; sa structure relèvera alors du droit d'auteur de l'administration. D'autre part, si sa constitution a donné lieu à un investissement financier, matériel et humain substantiel, son contenu relèvera du droit du producteur (CPI, art. L. 341-1). Or, dès lors que l'administration entre dans ces catégories d'auteur ou de producteur de bases de données, elle peut légalement en restreindre la réutilisation (CPI, art. L. 342-1). Concilier ces dispositions avec la logique d'*open data* n'est, dès lors, pas aisé (P. De Filippi et D. Bourcier, La double face de l'*open data*, LPA 10 oct. 2013, n° 2013, p. 6).

« Les droits de propriété intellectuelle constituent des freins naturels à l'*open data* »

Saisi de la question par la société de généalogie Notre-famille.com, qui s'était vu refuser l'accès à la base d'archives du département du Cantal, le juge administratif a d'abord jugé que l'« exception culturelle » ne conférait pas le droit de déroger au principe général de libre réutilisation (TA Clermont-Ferrand, 13 juill. 2011, n° 1001584, AJDA 2012, 375, note D. Connil; CAA Lyon, 4 juill. 2012, n° 11LY02325, *Département du Cantal*, AJDA 2013, 301, note D. Connil). Mais sa position s'est depuis lors infléchi : dans une affaire identique, opposant la même société au département de la Vienne, il a considéré que ce dernier était, compte tenu de l'investissement réalisé, « producteur de la base de données » et pouvait par conséquent interdire d'en « scraper » le contenu (c'est-à-dire, de l'aspirer de manière automatique), d'autant que d'autres modalités de consultation – sur place ou sur internet – étaient prévues (CAA Bordeaux, 26 févr. 2015, n° 13BX00856, *NotreFamille.com c/ Conseil général de la Vienne*, AJDA 2015, 1078).

Le Conseil d'Etat, saisi d'une demande de QPC portant sur la conformité de l'article L. 342-1 CPI aux droits et libertés garantis par la Constitution, a jugé que la disposition en cause se borne à tirer les conséquences nécessaires des dispositions précises et inconditionnelles d'une directive européenne, sans mettre en cause une règle ou un principe inhérent à l'identité constitutionnelle de la France, et qu'il n'y avait dès lors pas lieu de transmettre la question au Conseil constitutionnel (CE 14 sept. 2015, n° 389806, *Société NotreFamille.com*, Lebon T.; AJDA 2015, 2441, note F. Rassy; D. 2015, 2214, obs. J. Larrieu, C. Le Stanc et P. Tréfigny).

HAL -
01026207

bornait

Doit-on y voir une réticence du Conseil d'Etat à faire prévaloir l'ouverture des données sur le droit de la propriété intellectuelle? On ne saurait sérieusement lui reprocher son refus de faire du droit à l'information publique un principe inhérent à l'identité constitutionnelle de la France. Ces arrêts, qui vont à rebours des jurisprudences antérieures et des avis de la CADA, illustrent une certaine forme de résistance à l'ouverture, les administrations n'ayant pas toutes les moyens de procéder à la mise en *open data*. Le gouvernement entend

lever ces obstacles en leur interdisant de se retrancher derrière leurs droits de propriété intellectuelle pour refuser la réutilisation du contenu des bases de données qu'elles ont obligation de publier (art. 6 de la loi). Sur ces questions sensibles, sous-tendues par des enjeux économiques importants, rien ne dit que le législateur ne fera pas, lui aussi, de la résistance à l'*open data*.

7 du projet
de loi
République
numérique

NOTE À L'ATTENTION DES AUTEURS

Les articles proposés à l'AJDA doivent être adressés par courrier électronique exclusivement à l'adresse aj@dalloz.fr (et non aux adresses personnelles ou professionnelles du directeur scientifique, des membres de la rédaction ou des responsables de rubriques).

Ils doivent être envoyés sous un format permettant de mesurer la longueur précise de l'article (word, wordperfect...) et non au format pdf.

Les études ou actualités législatives doivent avoir une longueur comprise entre 15 000 et 50 000 signes, espaces et notes de bas de page compris. Celles-ci doivent être aussi peu nombreuses que possible, les principales références pouvant être regroupées dans un encadré.

Les notes de jurisprudence et les conclusions doivent être d'une longueur allant de 10 000 à 30 000 signes et suivre les mêmes règles pour les références et notes de bas de page.

En ce qui concerne les notes de jurisprudence, si les commentaires des membres du barreau sont toujours les bienvenus, le conseil scientifique de l'AJDA a posé la règle que ne devaient jamais être acceptées les notes rédigées par les conseils des parties. Cette règle, qui ne connaît aucune exception, s'applique à l'ensemble du cabinet consultant une partie et, *a fortiori*, aux parties elles-mêmes.

Les auteurs sont priés d'indiquer leurs nom, prénom, qualités et coordonnées personnelles (adresses électronique, postale et téléphone).

Les articles seront examinés par la direction scientifique et une réponse donnée dans un délai d'un mois au plus tard, pendant lequel l'auteur s'engage à ne pas soumettre son article à une autre revue.

RESPONSABLES DE RUBRIQUE

Actes unilatéraux. Delphine Costa, professeur de droit public à l'université d'Aix-Marseille. **Action sociale et aide sociale.** Hervé Rihol, professeur de droit public à l'université d'Angers. **Collectivités territoriales.** Bertrand Faure, professeur à l'université de Nantes, directeur du laboratoire « Droit et changement social » (UMR 6297). **Contentieux.** Florent Blanco, professeur de droit public à l'université d'Orléans. **Contrats.** Guylain Clamour, professeur à l'université de Montpellier, directeur du CREAM (EA 2038). **Domaine.** Norbert Foulquier, professeur de droit à l'université Paris I – Panthéon-Sorbonne, codirecteur du SERDEAUT. **Droit des étrangers.** Emmanuel Aubin, professeur de droit public, faculté de droit et des sciences sociales de l'université de Poitiers. **Droit électoral.** Romain Rambaud, professeur de droit public à l'université Grenoble-Alpes. **Droit public de l'économie.** Sophie Nicinski, professeur à l'université Paris I – Panthéon-Sorbonne. **Libertés publiques.** Xavier Bioy, professeur à l'université de Toulouse Capitole. **Responsabilité.** Anne Jacquemet-Gauché, professeur de droit public à l'université d'Auvergne. **Urbanisme et environnement.** Yves Jégouzo, professeur émérite à l'université Panthéon-Sorbonne.