

Livre Blanc



inno³

Les licences "Open Data"

Quel(s) usage(s) pour quelle(s) fin(s) ?

Version 1.0

Paris, le 16/04/2015

Table des matières

Propos introductifs.....	3
1 L'Open Data à quelle(s) fin(s) ?.....	4
1.1 L'Open data « subi ».....	4
1.2 L'Open data « évident ».....	4
1.3 L'Open Data comme stratégie de valorisation.....	5
2 Le rôle de la licence dans toute démarche Open Data.....	6
2.1 Qu'est ce qu'une licence en matière d'Open Data ?.....	6
2.2 Quelle utilité ?.....	7
2.3 Les limites.....	7
3 Le choix d'une licence Open Data.....	7
3.1 Qu'est ce qu'une licence Open Data ?.....	8
3.2 Quelle utilité ?.....	9
3.3 Les limites.....	9
4 Les différentes licences Open Data.....	9
4.1 Le binôme français.....	9
4.1.1 ODbL v1.0 – Open Data Commons Open Database License v1.0.....	9
4.1.2 Licence Ouverte.....	10
4.2 Les autres licences.....	10
4.2.1 CC By-SA 4.0 ou Creative Commons Attribution Partage dans les mêmes conditions v4.0.....	10
4.2.2 ODC-By v1.0 – Open Data Commons Attribution License v1.0.....	10
4.2.3 PDDL v1.0 ou Open Data Commons – Public Domain Dedication & License v1.0.....	10
4.3 Les licences spécifiques.....	11
5 L'usage des licences.....	11
5.1 Implémentation du formalisme inhérent aux diverses licences Open Data.....	11
5.2 La combinaison de jeux de données.....	11
5.3 Organiser la gestion des licences.....	12
Licence.....	13
À propos d'inno ³	13
À propos de l'Auteur.....	13

Propos introductifs

Les données, vue comme de simples informations, n'ont que rarement de valeur isolée, alors que leur agrégation au sein de bases de données – parfois énorme et on évoque à ce titre un autre buzzword qui est le BIG DATA – constitue un véritable actif immatériel pour celui qui le détient.

Initié à la fin des années 2000, le mouvement de l'« Open Data » étend les logiques de partage, de collaboration et de mutualisation au champ des bases de données. Depuis lors, il s'est étendu sans faiblir à de nombreux domaines tels que l'environnement, le géospatial, la santé, l'agriculture, le droit, etc. La présence, hier, de monopoles forts, semble entraîner aujourd'hui, un bouleversement profond des secteurs concernés.

Si l'on se place du côté de l'utilisateur ou de l'entrepreneur, force est de constater la richesse et le volume de plus en plus important des jeux de données disponibles, qu'il s'agisse de données publiques (administration et collectivités) ou privées. Ce processus entraîne l'exploitation pleine et entière de certains jeux de données et opère un revirement stratégique pour les acteurs privés comme publics. L'ouverture de ces jeux de données – au potentiel connu ou non – repose sur l'idée de générer une valeur qui, irrémédiablement, profitera à l'économie globale.

Si l'on se place du côté du producteur, force est de constater les bénéfices directs – et surtout indirects – qu'assure une telle diffusion de jeux de données : gain en image, standardisation de ses technologies, meilleurs services aux clients/utilisateurs, création d'un écosystème périphérique d'innovation.

Pour le secteur public plus que pour tout autre, l'Open Data est incontournable. Il induit une logique qui permet à la fois de rendre un meilleur service aux administrés, d'assurer un égal accès aux informations publiques, de mutualiser entre personnes publiques – lesquelles sont souvent les premières utilisatrices de ces informations publiques – et enfin de favoriser l'innovation autour des données ainsi ouvertes. **L'ouverture des données publiques n'est pas seulement un impératif démocratique, c'est aussi un puissant vecteur de modernisation de l'administration et un composant important de l'économie numérique.** Ce n'est donc pas étonnant si l'on compte plus d'une cinquantaine de collectivités françaises entrées dans une démarche Open Data (Brest, Paris, Nantes, Rennes, Toulouse, Bordeaux, Montpellier pour n'en citer que quelques-unes) – mouvement qui s'est traduit par la création de l'association dédiée appelée Open Data France¹. La création d'une fonction de *Chief Data Officer* pour accélérer et systématiser l'Open Data au sein de l'administration prolonge encore cette idée.

L'Open Data est donc présent et bien en place, et tout nous laisse présager une adhésion progressive de l'ensemble des acteurs économiques souhaitant tirer profit de ce nouveau courant ou simplement garder une place dans la société de demain.

1 <http://www.opendatafrance.net/>

1 L'Open Data à quelle(s) fin(s) ?

Les chemins qui mènent à l'Open Data ne sont pas tous les mêmes. Force est de constater que l'Open Data est encore bien souvent subi, entraînant une adhésion « à reculons » qui explique parfois certains échecs, mais l'Open Data est aussi et avant tout un choix stratégique.

1.1 L'Open data « subi »

À plusieurs reprises, la loi est venue imposer des contraintes de plus en plus fortes allant jusqu'à forcer un cadre de diffusion Open Data. L'exemple le plus représentatif est sans aucun doute la Loi relative aux documents administratifs et informations publiques (dite loi « CADA ») du 17 juillet 1978. Elle pose les fondements de l'ouverture des données publiques et a évolué jusqu'à imposer la libre réutilisation des informations contenues puis libre réutilisation gratuite de certains jeux de données.

Dans le même esprit, la Charte de l'environnement a été rédigée dans la continuité de la Convention d'Aarhus afin de préciser que « *toute personne a le droit, dans les conditions et les limites définies par la loi, d'accéder aux informations relatives à l'environnement détenues par les autorités publiques et de participer à l'élaboration des décisions publiques ayant une incidence sur l'environnement* ». La directive européenne INSPIRE 2007/2/CE a contraint aussi à l'ouverture de certaines données en faveur de l'établissement et de l'exploitation d'une infrastructure d'information spatiale commune en Europe.

Nul doute que ces Lois ont entraîné la diffusion de la majorité des bases de données Open Data. Néanmoins l'accompagnement très faible (voire inexistant) des acteurs concernés – lorsqu'ils n'ont pas cherché à contourner ces dispositifs ou à minimiser leurs effets – les a souvent rendus impuissants face aux nouvelles questions auxquelles ils devaient répondre, rendant la transformation plus douloureuse.

1.2 L'Open data « évident »

L'Open Data permet de donner une plus grande lisibilité à l'action publique, cette ouverture étant « *une condition de la démocratie participative et une source d'amélioration de l'efficacité des administrations* »² qui répond à l'exigence démocratique de transparence de l'État (le Conseil d'État a d'ailleurs fait du droit d'accès aux documents administratifs une garantie fondamentale accordée au citoyen).

« *Bien informés, les hommes sont des citoyens ; mal informés ils deviennent des sujets* »³. Avec la mise à disposition des données publiques, l'État et les collectivités font montre d'une volonté de transparence afin de conférer une plus grande lisibilité à l'action publique. La compréhension des citoyens passe par l'accès aux dites données publiques. Les objectifs de cette transparence sont

2 Voir l'Avis n°12 du Conseil national du numérique relatif à l'ouverture des données publiques (« Open Data »), juin 2012, http://www.cnumerique.fr/wp-content/uploads/2012/06/2012-06-05_AvisCNum_12_OpenData.pdf.

3 Alfred Sauvy, économiste, démographe et sociologue français du XXème siècle.

multiples, à commencer par l'information des citoyens, la promotion des initiatives et de l'innovation, l'amélioration des services existants, ou encore la lutte contre la fraude et les abus ou encore les erreurs commises au sein de l'administration⁴.

D'un point de vue technique et pragmatique, notons enfin que l'Open Data répond à un besoin pas toujours perçu par les organisations : l'entrée dans une logique d'Open Data entraîne ainsi nécessairement une rationalisation du système d'information, la mise en place de processus automatisés de gestion des données et finalement un meilleur service pour les membres de l'organisation elle-même.

1.3 L'Open Data comme stratégie de valorisation

Se tourner vers l'Open Data peut aussi être une stratégie réfléchie.

Certaines bases de données sont **gérées comme des biens communs**, par réaction aux monopoles contre lesquels elles luttent. Le meilleur exemple en la matière est OpenStreetMap, base de données géographiques mondiale créée en 2004 à la suite du refus de l'*Ordonance Survey* de partager ses données gratuitement. Sans les contributions des citoyens qui éditent eux-mêmes cette « carte », OpenStreetMap n'existerait pas. Or désormais, cette base de données couvre la plupart des pays du monde et concurrence Google Maps. On peut également citer TeleBotanica, réseau de botanistes professionnels et amateurs qui créent collaborativement une base de données partagée sur la faune et la flore partout dans le monde, OpenFoodFacts, base de données qui permet à chacun de scanner les données nutritionnelles présentes sur les emballages des aliments, ou encore OpenMeteoData, base de données en faveur du partage libre de données météorologiques.

L'ouverture des données peut aussi représenter un véritable enjeu économique direct. L'économie numérique est l'un des rares secteurs à maintenir sa croissance et, l'Open Data est appelé à y contribuer formidablement quand ce n'est pas déjà le cas. Si « *l'Open Data à tout prix* » est néfaste, en revanche l'Open Data est bénéfique très largement.

Diffuser ses jeux de données en Open Data peut favoriser le développement de **services innovants et complémentaires autour des données libérées** – songeons par exemple aux développeurs d'applications pour smartphones – ce qui rend d'autant plus attractif les services finaux pour les utilisateurs (par exemple dans le domaine du transport). Alternativement, si la licence Open Data utilisée est suffisamment contraignante, il est même possible de songer à une valorisation économique directe accessoire dans un format de *doubles licences* (la licence commerciale alternative permettant par exemple de ne pas partager ses propres améliorations ou données combinées). La liste n'est pas complète et certaines initiatives ont ainsi vocation à mettre à disposition du public des données pour **déployer un standard, faciliter l'usage d'un autre produit**

⁴ En 2009 au Royaume-Uni, le Guardian a demandé à ses lecteurs d'analyser les dépenses des parlementaires britanniques mises en ligne. En quelques jours, de nombreux abus ont été détectés (notes de frais, subvention pour des résidences secondaires inexistantes, et sept parlementaires ont été incarcérés. De même, aux Etats-Unis, l'administration Obama a créé recovery.org, site permettant de rechercher tous les projets financés par le Recovery Act afin de permettre à chacun de voir les projets en cours dans sa région, pour tel ou tel secteur, en connaître la finalité, le montant ou le statut et ainsi minimiser d'éventuels abus, chacun pouvant rapporter les dépenses injustifiées.

(par exemple une API) ou **renforcer la valeur globale de leur produit en s'appuyant sur une large communauté grâce à la démarche Open Data**. Certains acteurs peuvent choisir de partager leurs jeux de données afin de contribuer au déploiement de standards. Face au succès d'OpenStreetMap, tout le monde à intérêt à contribuer en rajoutant ses données au sein de la base communautaire afin de bénéficier de son effet de réseau (sa communauté) qui en fait une référence incontournable. La création de la Base d'Adresses Nationales Ouverte (BANO – récemment devenue BAN, pour Base Adresse Nationale collaborative) se présente aussi comme un référentiel national.

Lorsque l'on parle d'Open Data, il ne faut enfin pas oublié que la promesse du monde du « Libre » passe nécessairement par à la fois la constitution, la gestion et l'animation d'une communauté et, en parallèle, la mise à disposition d'outils (en premier lieu la base de données structurée et documentée) qui leur permettra de travailler. Un certain nombre d'associations constitue la face active et militante de l'Open Data en France (citons Regards Citoyens, LiberTIC, Open Street Map France, VVL, OKF ou encore Wikimedia France) et jouent un véritable rôle dans les mutations privées et publiques entraînées par l'extension de l'Open Data.

2 Le rôle de la licence dans toute démarche Open Data

La licence intervient afin de sécuriser le déploiement de l'Open Data et assurer le respect du cadre de diffusion choisi par l'entreprise ou la collectivité ayant fait le choix de publier ses données. C'est aussi elle qui scelle le cadre de collaboration choisie. Elle contribue ainsi grandement à la confiance que le contributeur aura vis-à-vis du projet.

2.1 Qu'est ce qu'une licence en matière d'Open Data ?

Les licences sont des contrats de cession non exclusive de droits de propriété intellectuelle, consentis pour le monde entier et toute la durée des droits, par lesquels un titulaire de droit autorise gracieusement un licencié à copier, modifier et distribuer sa création. Reposant sur une assiette plus large que le seul droit d'auteur, ces licences incluent également le droit des contrats et le droit *sui generis* des bases de données :

- **le droit *sui generis*** : existant uniquement en Europe, ce droit est accordé au producteur de bases de données pouvant justifier d'un investissement dans la constitution de sa base de données (de nature financière, matérielle ou humaine pour la constitution, la vérification ou la présentation de la base de données) ;
- **le droit d'auteur** : harmonisé à l'échelle internationale, le droit d'auteur peut trouver à s'appliquer lorsque la structure de la base de données est originale, c'est-à-dire la manière dont sont organisées les données, et les choix liés à cette classification relèvent d'une originalité.

- **le contrat** : lorsque l'accès à certains jeux de données peut être conditionné à l'acceptation de certains termes contractuels.

Avant toute exploitation, il est donc nécessaire de veiller à l'obtention systématique d'une autorisation non ambiguë d'utiliser tout jeu de données, expresse ou sous forme de licence générique, afin de connaître l'étendue des droits et obligations greffés sur ces données ainsi que la pérennité de leur accès.

2.2 Quelle utilité ?

Les licences s'appuient sur un mécanisme clé en matière de propriété intellectuelle : tout ce qui n'est pas cédé expressément est retenu. La licence vient ainsi définir l'étendue de l'autorisation bénéficiant à l'utilisateur.

En cas d'irrespect des obligations imposées par les licences, l'utilisateur de la base de données cesse de bénéficier des droits octroyés par la licence et devient contrefacteur, s'exposant à une sanction pouvant aller jusqu'à 3 ans d'emprisonnement et 300.000 euros d'amende⁵.

2.3 Les limites

Les licences connaissent toutefois certaines limites quant à leur volonté de tout couvrir (l'Open Data repose à l'inverse sur la survenance d'imprévus) et l'austérité que peut caractériser le droit. En effet, ce cadre juridique peut apparaître comme rigide et parfois effrayant vis-à-vis des communautés d'utilisateurs et de contributeurs que l'on cherche à séduire.

Par ailleurs, l'adhésion à la licence doit être la plus légère et plus rapide possible : à cet égard, les premières licences proposées par l'APIE en France furent difficilement acceptées compte tenu de leur cadre classique ici inadapté (jugées trop lourdes, elles nécessitaient une signature individuelle, ne comprenaient pas la dimension communautaire, etc.).

Pour ces raisons, il est généralement conseillé de choisir des licences préexistantes et de s'appuyer sur les efforts de communication dont elles ont fait l'objet.

3 Le choix d'une licence Open Data

Pendant juridique du concept d'Open Data, les licences Open Data sont des contrats prévus pour accompagner et sécuriser ce mouvement d'ouverture et de collaboration, tant pour les ayants droit que les utilisateurs. Le rôle des licences Open Data est donc double, il permet à la fois de favoriser la diffusion des données, et d'autre part d'imposer le respect du cadre d'exploitation défini.

Le choix de la licence initiale est d'autant plus fondamental qu'il a de l'importance en termes de communication à l'égard des communautés déjà existantes, de stratégie, car le modèle économique souhaité peut en dépendre, et enfin, sur le plan juridique⁶.

⁵ Article L335-3 du Code de la propriété intellectuelle.

⁶ Voir le Petit guide à destination des administrations souhaitant basculer du payant à l'Open Data, juillet 2014, <http://www.regardscitoyens.org/documents/notes/20140713-RegardsCitoyens->

3.1 Qu'est ce qu'une licence Open Data ?

L'Open Data est une application des concepts formalisés par l'*Open Knowledge Foundation* au Royaume-Uni et la *Sunlight Foundation* (laquelle a pour objet de rendre les gouvernements "transparents et responsables") aux États-Unis.

En vertu de l'Open Definition rédigée sur le modèle de l'*Open Source Definition*, la licence utilisée sur le jeu de données sera considérée comme Open Data si elle garantit :

- **Le libre accès des bases de données** : les bases de données doivent être disponibles entièrement, pour un coût raisonnable en cas de distribution matérielle (gratuitement lorsque le coût est négligeable, comme pour un transfert par Internet). Elles doivent être disponibles dans une forme adaptée et modifiable.
- **La libre redistribution des bases de données** : la licence ne peut, par exemple, exiger le paiement d'une redevance supplémentaire au titre de la redistribution.
- **La libre réutilisation des bases de données** : la licence doit permettre la création et la redistribution de bases de données dérivées (éventuellement sous la même licence lorsqu'elle est de type share alike ou copyleft).
- **L'absence de restrictions techniques** : il ne doit pas y avoir de restriction technique empêchant la jouissance des droits conférés par la licence.
- **L'attribution des auteurs et contributeurs** : la licence peut soumettre au bénéfice de la licence l'attribution de la paternité des créateurs et contributeurs.
- **L'intégrité de la base doit être préservée** : un tiers ne peut pas s'approprier le travail d'un autre et les contributions de chacun sont clairement attribuées.
- **L'absence de discrimination entre les personnes ou les groupes** : toute personne détentrice d'une copie de la base de données bénéficie des termes de la licence tant qu'il s'y conforme lui-même.
- **L'absence de discrimination entre les domaines d'application** : la licence se limite à la propriété intellectuelle : elle ne peut en aucun cas réguler d'autres domaines « politiques » ou commerciaux.
- **La licence s'applique sans dépendre d'autres contrats** : par exemple on ne peut pas ajouter un NDA (Non-Disclosure Agreement ou accord de confidentialité) lors de l'accès à la base.

Contrairement aux licences Open Source, il n'existe pas encore d'organisations telles que l'Open Source Initiative qui viendrait labelliser les licences conformes.

3.2 Quelle utilité ?

Bénéficiant de l'expérience tirée des logiciels libres et Open Source, ces licences ont pour force de reposer sur une base juridique solide et une « communauté d'utilisateurs » croissante (tout l'intérêt, et le succès, des licences étant la confiance et harmonisation qu'elles engendrent), leur permettant d'être utilisées en toute confiance et en mutualisant sur les efforts de promotion et de vulgarisation réalisés par chaque diffuseur.

Choisir la licence u

3.3 Les limites

Jusqu'à présent, les licences Open Data n'ont pas fait l'objet de contestations judiciaires qui auraient permis d'affirmer ou infirmer leur « conformité » au regard des droits nationaux. Pour autant, l'absence de contestations témoigne en un sens de l'acceptation générale à l'égard des licences Open Data et les décisions judiciaires afférentes aux licences libres et Open Source laissent supposer que les juges français se positionneraient en faveur du respect de celles-ci. Par ailleurs, puisque tout ce qui n'est pas permis est expressément interdit, personne n'aurait intérêt à contester une licence sans laquelle il n'aurait plus de droits.

Dernière limite, la résistance au changement que subissent les projets Open Data (en interne, mais aussi en externe) ne doit pas faire perdre de vue l'intérêt général et la souplesse qu'offrent les licences Open Data : toutes mises à disposition différenciées (licences négociées, format spécifique, interprétations différentes, etc.) doivent ainsi être évitées.

4 Les différentes licences Open Data

Les premières licences Open Data sont apparues pour répondre aux besoins des projets communautaires ainsi qu'à l'ouverture des informations publiques. Peu nombreuses à l'origine, il en existe désormais un nombre raisonnable (comparativement à l'Open Source qui en compte plusieurs centaines).

En présence de licences Open Data, un certain nombre de questions doivent être posées : la licence est-elle copyleft ? Cette licence impose-t-elle un formalisme spécifique ? Ces licences sont-elles compatibles entre elles ? Quels usages sont susceptibles d'en découler ? Ses caractéristiques, ou leurs combinaisons, sont traitées différemment par chaque licence Open Data existante – à la charge de celui qui la choisit ou qui souhaite s'y soumettre de bien comprendre les libertés et obligations qu'elle contient.

4.1 Le binôme français

Complémentaires et compatibles les licences ODbL et la Licence Ouverte sont les deux licences les plus utilisées en France.

4.1.1 ODbL v1.0 – Open Data Commons Open Database License v1.0

Il s'agit de la première licence rédigée spécifiquement pour un Open Data en matière de bases de données s'appuyant sur les législations européennes. Elle prévoit un encadrement complet comprenant les droits d'auteurs, mais aussi le droit des contrats et le droit *sui generis* des bases de données.

La licence ODbL a pour particularité d'être une licence de type « Share Alike » ou *copyleft*. À ce titre, elle impose la redistribution de toute base de données – modifiées (on parle de *bases de données dérivées*) ou non – sous les mêmes conditions, permettant ainsi aux titulaires de droits de conserver un certain contrôle sur leurs bases de données (tout en autorisant néanmoins la combinaison de plusieurs bases au sein d'une base plus importante dite base de données collaborative⁷).

Enfin, la licence ODbL inclut un mécanisme de compatibilité expresse qui ouvre la voie à une certaine interopérabilité. À noter qu'elle contient un dispositif de « proxy » qui permet de laisser ouverte certaine potentialité de compatibilité en se réservant ou en déléguant à une autre entité la possibilité d'ajouter de rendre la licence compatible avec toute autre licence (Open Data ou plus traditionnelle).

4.1.2 Licence Ouverte

Initialement rédigée pour l'administration centrale dans le cadre du projet Etalab, la Licence Ouverte est une licence très permissive permettant tout type de réutilisation des données.

Elle impose seulement de maintenir un formalisme simple se traduisant par l'affichage de certaines mentions lors de la distribution de la base, modifiée ou non (mention de la base d'origine et de sa fraîcheur, mention de la licence). En plus des licences expressément désignées, sont compatibles toutes les licences exigeant mention de la paternité.

4.2 Les autres licences

4.2.1 CC By-SA 4.0 ou Creative Commons Attribution Partage dans les mêmes conditions v4.0

La dernière version des licences Creative Commons est désormais adaptée aux bases de données.

À l'instar de la licence ODbL, la licence CC By-SA est *copyleft* ou *share alike*, toute base de données dérivée doit être conservée sous la même licence. Pour cela, lors de toute communication de la base, il convient d'indiquer la source de la base de données, ainsi que la mention de la licence.

Sans être compatible avec les autres licences, elle contient un dispositif de compatibilité lui permettant à terme de devenir compatible avec d'autres licences similaires, notamment l'ODbL. Elle reste néanmoins encore peu utilisée sur des projets Open Data, l'ODbL ayant rapidement occupé la place qui aurait pu être celle de la CC By SA.

⁷ Voir <http://www.vvlibri.org/fr/Analyse/open-database-license-10-analyse>.

4.2.2 ODC-By v1.0 – Open Data Commons Attribution License v1.0

Cette licence fait partie de la famille de licences (*Open Data Commons*) rédigée par l'Open Knowledge Foundation.

À l'instar de l'ODbL, l'ODC-By tient compte des législations européennes et prévoit un encadrement complet comprenant le droit d'auteur, le droit des contrats et le droit *sui generis* des bases de données portant sur la base de données.

Ses droits et obligations ainsi que ses éléments déclencheurs sont sensiblement les mêmes que pour l'ODbL à ceci près que l'ODC-By n'est pas copyleft et les bases de données dérivées ou les créations qui en sont issues n'ont donc pas à être licenciées sous la même licence.

4.2.3 PDDL v1.0 ou Open Data Commons – Public Domain Dedication & License v1.0

La PDDL place au plus près du domaine public tout travail sur les bases de données, ce qui en fait la plus permissive des licences Open Data Commons.

Cette licence prend la forme d'une renonciation au profit du domaine public lorsque c'est possible et lorsque ça ne l'est pas, la licence opère une cession non exclusive mondiale, gratuite et pendant toute la durée de protection des droits d'auteur.

Très permissive, cette licence est compatible avec de nombreuses licences.

4.3 Les licences spécifiques

De nombreuses autres licences spécifiques continuent néanmoins d'exister (à ne pas confondre avec les licences d'APIs généralement associées à des services offerts au travers de protocoles standards sur Internet) et doivent ainsi être comprises au travers du filtre classique : délimitation de l'objet de droit appréhendé par la licence, autorisations données aux utilisateurs, contraintes associées et cadre d'exploitation possible.

De manière générale, ces licences ne sont pas adaptées à une réutilisation à l'extérieur du cadre initial et sont progressivement remplacées au profit de licences harmonisées et compatibles.

5 L'usage des licences

Concrètement, l'usage de jeux de données sous licence Open Data impose de comprendre chaque licence afin de répercuter le formalisme requis par chacune, de combiner sans risque plusieurs jeux de données et, enfin, de gérer cet usage dès la sélection du jeu de données.

5.1 Implémentation du formalisme inhérent aux diverses licences Open Data

En pratique, le respect du formalisme se traduit par l'affichage d'une série de mentions à destination des utilisateurs finaux afin de :

- lui assurer l'accès à certaines informations (clause de paternité ou de publicité) ;
- lui permettre une appréhension matérielle et effective de l'œuvre (clause copyleft/de partage à l'identique ; clause spécifiant les infos à communiquer, leur format, etc.). Ce deuxième point est particulièrement important puisqu'il s'agit généralement de mettre à disposition le ou les jeux de données Open Data utilisés – modifiés ou non – pour générer le résultat ou la création à destination de l'utilisateur final.

Ce formalisme est très souvent simple à mettre en œuvre, il est d'autant plus dommage de risquer par mégarde une violation des licences (entraînant l'annulation immédiate de ces dernières).

5.2 La combinaison de jeux de données

La combinaison de jeux de données est l'un des usages souhaités, peut être le plus intéressant, dans le domaine de l'Open Data. Cela induit néanmoins un rapprochement en termes de licences et pose la question de l'incompatibilité (lorsque les termes de plusieurs licences Open Data seraient contradictoires – rendant la combinaison souhaitée impossible).

En raison de sa clause copyleft, la licence ODbL est la plus contraignante à cet égard (soit qu'elle étende ses termes à des jeux de données qui ne sont pas Open Data soit qu'elle provoque une situation d'incompatibilité). Globalement, il est permis de considérer qu'il est possible de combiner plusieurs bases dès lors que cette combinaison étend ou complète (telle que l'ajout de calques ou encore la séparation franche de plusieurs bases en fonction du territoire), mais n'améliore pas la base initiale. Toutes ces améliorations sont directement concernées par la licence ODbL et doivent être disponibles sous la même licence.

Au-delà, il est nécessaire de s'appuyer sur les usages et pratiques développés et acceptés au sein des communautés des projets Open Data utilisés. Par exemple, la communauté *OpenStreetMap* publie depuis peu une série de *Community Guidelines* qui permettent de préciser certains usages de la base OSM qui seraient soumis ou non à la licence ODbL sous laquelle OSM est distribué. Cela porte sur deux critères précis : le caractère substantiel de l'extraction faite de la base (en deçà, l'usage serait permis) ainsi que la notion de *Derivative Database* qui donne les limites de l'étendue de l'ODbL.

Il est donc nécessaire de s'assurer que l'usage et l'architecture de l'outil permettent de séparer sans ambiguïté les différentes bases pouvant être utilisées concomitamment. Par ailleurs, ces choix doivent être motivés et publiés sur le site de l'entreprise/l'association.

5.3 Organiser la gestion des licences

La bonne gestion de l'Open Data passe par l'adoption d'un certain nombre de bonnes pratiques. L'industrialisation (l'usage systématique) de jeux de données Open Data vient renforcer l'importance des transformations nécessaires au sein de l'organisation.

Sans être exhaustive, la liste des préconisations qui suit donne un aperçu des changements

structurels et organisationnels nécessaires :

- centralisation et mise à jour des informations relatives aux différentes licences Open Data associée aux jeux de données utilisés ;
- documentation des contraintes relatives aux licences des bases de données utilisées ;
- documentation, au sein d'une page dédiée, de la démarche entreprise pour assurer le respect des licences (en termes de respect des licences, architecture adoptée et processus mis en œuvre au sein de l'organisation) ;
- mentions légales spécifiant les jeux de données utilisés et les licences associées ;

Enfin, cela peut reposer sur la mise en place d'une gouvernance dédiée qui repose sur un comité Open Data en charge de la définition et de l'évolution de la politique, et d'un responsable Open Data qui prend le rôle de référent institutionnel. In fine, une telle organisation permettra d'autoriser et approuver le déploiement de jeux de données au sein des solutions ou produits de l'organisation sur la base de critères techniques, juridiques et commerciaux prédéterminés et moyennant le respect d'un formalisme spécifique au contexte du projet.

Licence

Ce livre blanc a été écrit par Benjamin Jean pour la société Inno³ sous licence Creative Common By SA 4.0. Pour obtenir des informations complémentaires sur cette licence, veuillez consulter le site : <http://creativecommons.fr/>

À propos d'inno³



Inno³ est un cabinet français indépendant accompagnant les organisations souhaitant tirer pleinement profit de l'Open Source/ au sein de leur IT et de leur R&D.

Leader du domaine, inno³ dispose de toutes les compétences nécessaires pour adresser les trois piliers sur lesquels repose cette mutation : la dimension outil (par la gestion des progiciels Open Source), la dimension juridique (au travers de la gestion des licences Open Source et des contrats associés) et la dimension humaine (dans l'organisation des rapports avec les communautés dont sont issus les projets). Les offres de services proposées vont de la simple expertise ponctuelle (notamment audit et analyse d'opportunité) à l'accompagnement au changement (définition de stratégie, mise en place de politique et gouvernance dédiées) et au transfert de compétences (au travers de cursus préconstitués ou de modules « à la carte »).

La société anime enfin un réseau d'experts métiers qui permettent de porter les transformations liées à l'Open Source dans tous les secteurs industriels ou publics.

À propos de l'Auteur



Fondateur du cabinet inno³, Benjamin Jean travaille essentiellement sur les nouveaux usages et collaborations qui se développent autour de la création et de l'innovation (Open Source, Open Data ou plus généralement toute la mouvance en faveur d'une innovation ouverte). Actif depuis près de dix ans dans ce domaine, il a notamment œuvré pour leur prise en compte au sein des gouvernances de multiples acteurs privés et publics, de tailles et de domaines variés, et formé de nombreuses équipes internes aux aspects juridiques afférents.

Il enseigne par ailleurs la propriété intellectuelle au sein de plusieurs Masters spécialisés, intervient comme consultant au sein du cabinet Gilles Vercken et maître de conférences à Sciences Po.

Il est administrateur de PLOSS Paris Région, membre du Directoire de Systematic et Président de l'association *Open Law, le droit ouvert*. Il est également très présent dans les communautés du Logiciel Libre, administrateur de Framasoft et cofondateur de Veni Vidi Libri.