

Objet : licences de données et logiciels, et messages aux producteurs

Date : 25/05/2018

Version : 1

Émetteur : L. Coudercy

NOTE GCIB

Contexte

Le CPUSIE du 05/2011 a l'importance pour le SIE de l'open data, et en particulier que les données du SIE doivent être librement accessibles, gratuitement, sauf restrictions prévues par la loi.

Depuis la loi Lemaire de 2016 a clairement modifié le paradigme de la mise à disposition des données publiques. Désormais l'open data est la règle générale pour les données publiques, la limitation d'accès est l'exception. Cette loi conforte la position prise par le CPUSIE en 2011.

Dans le cadre du SIE, les principales limitations d'accès prévus par la loi et les réponses principales à appliquer sont :

- La sécurité civile (pour les captages AEP) 🚰 floutés, mais accès après identification
- Les données sensibles pour l'environnement (espèces spécifiques) 🚰 floutées sauf pour services publics
- Le secret industriel et commercial hors émissions (volumes prélevés) 🚰 agrégées
- Le secret de la vie privée 🚰 anonymisées, information sur les cookies

Par ailleurs, la loi Lemaire impose une liste de licences limitative, sauf justification devant la COEPIA (Décret n° 2017-638 du 27 avril 2017).

Proposition d'application de ces règles dans le cas des banques du SIE

Le SIE ayant une politique d'open data, il n'y a pas de raison valable de demander une exception aux licences proposées par le décret cité précédemment.

Il est proposé d'appliquer les licences suivantes :

- licence ouverte Etalab 2.0 : pour les données (on pourrait envisager la licence ODBL, qui impose la rediffusion des données déduites sous cette même licence (licence dite virale), mais cela limiterait sans raison valable l'usage de nos données).
- licence ouverte Etalab 2.0 : pour les documents (même remarque)
- une des licences non virale pour logiciel (Cecill B, Apache, ...)
- ou une licence virale (Cecill ou GNU), s'il est souhaitable de garantir que les modifications apportées par un tiers au logiciel libre de l'administration (par exemple les codes e calcul d'indicateurs DCE) seront accessibles sous les mêmes conditions.

Si des restrictions sont prévues par la loi, il faut alors prévoir la licence libre pour la forme autorisée, et sinon une licence rappelant les limites d'emploi pour les versions réservées, et en particulier indiquant qu'il y a un secret que le nouvel utilisateur doit appliquer.

Sur les sites d'accès grand public (front office), il faut rappeler la licence employée sans rajouter de conditions spécifiques, sauf pour les documents ou photos grevées par une propriété intellectuelle d'un tiers privé (dans ce cas, mettre la propriété intellectuelle visée sur le document, en exergue).

Il faut par ailleurs appliquer les règles du RGPD, avec les rappels qui vont bien concernant la vie privée ; un point particulier concerne les cookies : il faut permettre de les refuser, et donc les signaler dans une pop up en entrée.

On peut s'inspirer pour l'ensemble de ces textes de ce qui est fait par exemple sur bdlisa.eaufrance.fr.

Sur les sites d'alimentation producteur (back office), il est important de rappeler la licence employée en sortie pour els données, afin que les producteurs aient pleinement conscience que leurs données seront diffusées en open data.

Enfin là aussi, il faut appliquer le RGPD.