

**Directive 2013/37/UE sur la réutilisation des données publiques :**

**Les 9 propositions du GFII pour la transposition en droit français**

La Directive 2003/98/CE concernant la réutilisation des informations du secteur public a fait l'objet d'une révision récente (Directive 2013/37/UE, du 26 juin 2013).

Cette révision traduit un effort d'ouverture des informations publiques et une avancée notamment en matière de droit à la réutilisation, de tarification, d'ouverture des documents des bibliothèques, des archives et des musées, de qualité des informations et d'amélioration des formats. Elle prévoit en outre l'instauration d'une instance de recours indépendante, dotée de pouvoir de sanction. Toutefois, elle reste moins ambitieuse et paradoxalement en retrait par rapport aux précédents travaux de la Commission et aux politiques d'ouverture menées par certains États-membres.

Ainsi, pour les pays ayant opté pour une transposition *a minima* de la Directive 2003/98/CE, ce nouveau texte est une avancée majeure. Pour la France, qui a déjà instauré un droit à la réutilisation, avec la « loi CADA<sup>1</sup> » et posé par la suite un principe de gratuité, la transposition constitue une opportunité pour renforcer la politique de l'Etat, en tenant compte de l'évolution des politiques d'ouverture des données mises en œuvre depuis 2003 et confirmer les politiques volontaristes menées par les gouvernements successifs.

Une transposition ambitieuse de la Directive 2013, qui doit intervenir avant le 18 juillet 2015, aurait en outre pour mérite de s'inscrire dans les plans d'actions concernant la réutilisation des données publiques, auxquels la France s'est engagée à participer dans le cadre de l'exécution de la *Charte du G8 pour l'ouverture des données publiques* qu'elle a signée le 18 juin 2013.

La transposition permettra également de réduire certaines contradictions ou tensions entre les différentes lois qui régissent la mise à disposition des informations publiques : accès et réutilisation, publicité légale versus données personnelles, loi CADA versus loi CNIL,...

Le GFII formule 9 propositions pour réunir les conditions optimales permettant le développement économique à partir de la réutilisation des informations publiques.

---

<sup>1</sup> La loi n° 78-753 du 17 juillet 1978 modifiée, dite « loi CADA », reconnaît à toute personne le droit d'obtenir communication des documents produits ou reçus par une administration dans le cadre de sa mission de service public, quels que soient leur forme ou leur support. Ce droit s'exerce à l'égard de tous les services publics : l'Etat, les collectivités territoriales, et toute autre personne de droit public ou de droit privé chargée d'une mission de service public (art. 1<sup>er</sup>). Depuis 2005 et la transposition de la directive 2003/98/CE concernant la réutilisation des informations du secteur public, la « loi CADA » prévoit également la possibilité d'utiliser les informations publiques, quels que soient leur forme ou leur support, « à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus », avec certaines limites.

## 1. Elargir le périmètre des données réutilisables

A l'exception de l'intégration limitée des données détenues par les établissements culturels, le périmètre de la réutilisation visé par la Directive n'a pas subi de modification sensible. A l'occasion de la transposition de la Directive révisée, un débat pourrait s'ouvrir en France sur les types d'informations qui, dans le cadre des missions de service public, sont utiles à la vie économique et citoyenne et doivent par conséquent entrer dans le champ de la réutilisation.

En effet, suivant les régimes juridiques adoptés (services publics industriels et commerciaux ou services publics administratifs), une même information est considérée ou non comme une donnée publique réutilisable. Cette situation conduit, par exemple, à un manque de cohérence territoriale où, selon les choix des collectivités territoriales, les informations publiques sont ouvertes ou fermées. Le GFII propose que la transposition intègre dans le champ d'application de la réutilisation **certaines catégories d'informations indépendamment des conditions d'organisation du service public**. De telles réglementations catégorielles existent déjà en matière environnementale ou géographique pour lesquelles le régime juridique applicable est dicté par la nature de l'information et non par les conditions d'organisation fluctuantes du service. Une réflexion sectorielle sur les informations « utiles » pour l'exercice de la citoyenneté ou pour les activités économiques est à mener.

Le GFII propose ci-dessous une première liste de ces données à ouvrir en priorité :

- Pour les données sur les entreprises : KBIS, registre des métiers...
- Pour les données juridiques : jurisprudence des conseils de prud'hommes, jugements des tribunaux de Commerce, jugements des tribunaux de grande instance, jugements des tribunaux des Affaires de sécurité sociale ...
- Pour les données de santé : statistique annuelle des établissements de santé, causes de décès, morbidité hospitalière, carte de l'offre de soins, base de données des médicaments...

## 2. Organiser l'ouverture et la gratuité des « données pivot »

Le GFII souhaite que la transposition de la Directive révisée intègre une notion présente dans les débats européens, mais qui n'est pas reprise dans le texte de la Directive : la notion de « données pivot » ou « données de référence ».

Les « données pivot » ou « données de référence » sont des données considérées comme identifiantes, par l'administration ou par l'usage, pour nommer ou identifier des produits, des entités économiques ou des territoires. Ces référentiels sont indispensables pour lier des bases de données de nature hétérogène et construire tout nouveau service intégré au web des données (Linked Open Data).

Un format ouvert spécifique à chaque type de donnée devrait obligatoirement être utilisé, tout en laissant la possibilité de publier également selon certains formats propriétaires très usités. Ces formats ouverts devraient être documentés et stables, toute modification de format devant être précédée de procédures de concertation éliminant tout imprévu pour les réutilisateurs. Les formats de diffusion devraient s'appuyer, autant que possible, sur les normes internationales en vigueur dans les différents métiers.

### **Ces données, de par leur importance, devraient être ouvertes et mises à disposition gratuitement.**

Les lots de données devraient être exhaustifs par rapport au domaine décrit. La pérennité et la périodicité de la fourniture des données et de leurs éventuelles mises à jour devraient être garanties, ainsi que la traçabilité des traitements du producteur initial jusqu'à la diffusion.

Cette initiative mettrait une nouvelle fois la France à la pointe de l'innovation dans le domaine de la diffusion des informations publiques.

Le GFII propose en annexe une première liste de ces données pivot, par secteur et propose ci-dessous quelques exemples :

- Données géographiques : adresse, noms de lieux, limites de communes et limites corrélées, zone de compétences administratives, îlots statistiques...
- Données entreprises : Siren, Siret, numéro RNA, raison sociale, adresse juridique de l'établissement normalisée, code NAF, forme juridique.
- Données juridiques : numéros NOR précédant les textes réglementaires, numéros ECLI (textes officiels)...
- Données de transport : limitation de vitesse routière, horaires et zones tarifaires de transport...
- Données de santé : répertoire des professionnels (RPPS), carte de l'activité hospitalière, répertoire des spécialités pharmaceutiques, base de données des médicaments, liste des produits et prestations (LPP)...

### **3. Impulser la politique nécessaire pour ouvrir les données des collectivités et des établissements publics territoriaux**

Une partie significative des données publiques est aujourd'hui produite ou détenue par les collectivités locales (éducation, tourisme, environnement, données fiscales, données budgétaires, données de santé, données culturelles). Les attentes du grand public et des entreprises pour disposer d'informations locales sont importantes et des services innovants pourraient être créés à partir de ces données.

Dans le cadre de la loi CADA, les collectivités et établissements publics territoriaux sont déjà soumis au régime général en matière de réutilisation des informations publiques. Toutefois, trop peu de collectivités, d'intercommunalités et d'établissements publics territoriaux se sont engagés de façon systématique dans l'ouverture de leurs informations. A titre d'illustration, le réseau des personnes responsables de l'accès aux documents administratifs et des questions relatives à la réutilisation des informations publiques (PRADA), créé par la loi en 2005, compte actuellement seulement 1 500 personnes (annuaire 2013 sur le site de la CADA). Il est capital que le rôle moteur du PRADA soit réaffirmé lors de la transposition de la Directive en confirmant sa compétence quant à la question de la réutilisation des données publiques (en complément au simple droit d'accès) et en faisant l'obligation à toutes les collectivités ou établissements publics territoriaux d'en nommer un. A défaut de nomination, les **missions du PRADA devraient être exercées par le directeur général des services ou, le cas échéant, le directeur juridique.**

## 4. Intégrer les données culturelles au régime général de la réutilisation

La nouvelle Directive a intégré, dans le champ du droit à réutilisation, les données détenues par les « bibliothèques, musées et archives », qui constituent le cœur de l'activité publique dans ce domaine. On peut dès lors s'interroger quant aux raisons justifiant le maintien d'une exception de principe concernant tous les établissements culturels en général.

Les données détenues par les établissements culturels font, aujourd'hui en France, figure de parent pauvre des données publiques en matière de réutilisation, alors même qu'elles portent en elles un potentiel énorme de développement et d'innovation. Le régime dérogatoire (improprement appelé « exception culturelle ») dont bénéficient ces données dans la loi CADA n'a pas permis d'apporter une réponse claire aux confusions relatives aux limites de ce régime dérogatoire ainsi qu'aux questions de propriété intellectuelle et aux principes de tarification des licences de réutilisation.

- *Une exception de principe aux fondements fragiles*

La notion de « données culturelles », improprement utilisée, entretient une confusion avec les œuvres de l'esprit, qui peuvent être protégées par le droit d'auteur. Or les spécificités des œuvres de l'esprit qui sont détenues par des établissements culturels sont déjà prises en compte dans la loi CADA, qui retire du champ de la réutilisation toutes les données sur lesquelles des tiers détiendraient des droits de propriété intellectuelle.

Par ailleurs, le principe suivant lequel des données relèveraient de statuts juridiques différents en fonction des institutions qui les détiennent conduit à un risque de **régression sur l'ouverture des données publiques, en fonction de leur cycle de vie** : ainsi, une donnée publique librement réutilisable conformément au régime général, qui fait l'objet d'un versement à un service d'archives, n'est plus réutilisable aux mêmes conditions, étant détenue par un « établissement culturel ».

Enfin, si l'on peut comprendre que des données soient exclues du droit à réutilisation en raison de leurs caractéristiques propres, il conviendrait pour plus de clarté de **renoncer à une exception générale devenue sans objet et de procéder par voie d'exceptions**, de portée strictement définie, pour un petit nombre de catégories très particulières et pour des raisons précisément justifiées.

- *Des principes de tarification peu effectifs*

Pour les données détenues par les bibliothèques, les musées et les archives, la Directive prévoit un régime dérogatoire de tarification basé sur le principe d'un plafond intégrant le coût de conservation et d'acquisition des données ainsi qu'un retour raisonnable sur investissement.

Le **principe d'un régime dérogatoire en matière de tarification basé sur la nature d'établissements (archives, musées, bibliothèques) apparaît peu justifié**. En France, ce régime dérogatoire, déjà prévu par la loi CADA, a souvent constitué un moyen de faire obstacle à l'obligation d'ouverture des données par la mise en place de tarifications prohibitives. Si un tel principe devait toutefois être maintenu, il conviendrait de mettre en place des mécanismes efficaces permettant de contrôler que les principes de tarification particuliers dont bénéficient bibliothèques, musées et archives ne peuvent en aucun cas constituer un obstacle à la réutilisation des données.

En conclusion, l'argument parfois utilisé consistant à évoquer à cet égard l'« exception culturelle » ne semble pas opérant sauf à renverser la logique d'une politique faite, justement, pour assurer une plus grande diffusion des œuvres. Il serait paradoxal que des institutions culturelles publiques se voient conférer, de droit, par un régime dérogatoire, ou de fait, par une trop grande liberté en

matière de tarification, un droit à refuser l'accès et la réutilisation de leurs données alors que leur mission est au contraire d'en assurer la plus large circulation.

Il conviendrait au contraire de **favoriser l'accessibilité des données publiques constitutives du patrimoine culturel français par une politique volontariste**. La levée des dernières barrières tarifaires et juridiques favorisera l'innovation et la création en laissant libre cours à l'imagination des réutilisateurs.

## 5. Généraliser et organiser la pratique des licences pour la réutilisation des données publiques

Le GFII souhaite que soit **systematisé l'usage de licences pour toute réutilisation, y compris lorsque les données sont mises à disposition gratuitement** afin de sécuriser producteurs de données et réutilisateurs en rappelant le cadre juridique et notamment l'obligation de mentionner la source et les dates de mise à jour des données.

**Le GFII est favorable à l'utilisation de licences-types** telles que la licence ouverte (dite licence Etalab), la licence APIE ou encore certaines licences Creative Commons. A cet effet, il est nécessaire de rappeler que certaines conditions des Creative Commons sont contraires à l'esprit de la Directive sur la réutilisation des données publiques :

- « Pas d'utilisation commerciale » exclut les opérateurs économiques du champ de la réutilisation en leur interdisant de commercialiser les données concernées ;
- « Partage dans les mêmes conditions » crée une contrainte à la réutilisation qui va à l'encontre des logiques d'investissements des acteurs économiques. En effet, le réutilisateur serait contraint de mettre à disposition gratuitement ses données dès lors qu'il aurait utilisé dans son service ou son application certaines données diffusées avec ce type de licence ;
- « Pas de modification » limite les possibilités de mash-up, d'actualisation et d'enrichissement des données.

Dans les cas où des contraintes spécifiques seraient liées à la nature des données (données contenant des informations à caractère personnel...) ou lorsque la réutilisation est soumise au versement d'une redevance, le GFII préconise la **mise en place de licences plus détaillées** qui, en plus d'un rappel général du cadre légal, définiront les engagements respectifs des producteurs et réutilisateurs en matière de fourniture et de réutilisation des données, par exemple :

- données fournies en l'état conformément aux missions du producteur,
- format et périodicité de mise à disposition,
- respect des données personnelles,
- respect des droits de propriété intellectuelle, *etc.*

Le producteur d'information sera ainsi en mesure de contacter les réutilisateurs en cas de manquement manifeste au respect des obligations prévues par la licence ou si des modifications doivent être répercutées (dans le cas de problème lié à des défauts d'anonymisation par exemple).

De son côté, l'utilisateur sera sécurisé quant à l'étendue exacte de ses obligations contractuelles.

## 6. Réaffirmer le principe de gratuité et expliciter les principes de tarification dans les cas particuliers

Devançant le texte européen, la France a, depuis 2011, instauré un principe de gratuité des données publiques des administrations de l'Etat, qui ne cède que si l'administration concernée bénéficie d'un régime dérogatoire ou obtient une autorisation pour percevoir des redevances. Le GFII se félicite de cette décision qui constitue un grand pas en avant dans l'ouverture des données.

Le texte européen, en revanche, pose un principe de tarification aux « coûts marginaux de reproduction, de mise à disposition et de diffusion », assorti d'exceptions :

- pour les « organismes du secteur public tenus de générer des recettes destinées à couvrir une part substantielle des coûts liés à l'accomplissement de leurs missions de service public »,
- pour les bibliothèques (y compris les bibliothèques universitaires), les musées et les archives,
- pour les documents pour lesquels l'organisme du secteur public concerné est tenu de générer des recettes.

Dans la quasi-totalité des cas, les ressources tirées d'une tarification au coût marginal de diffusion proposé par la Directive sont bien en-deçà du coût de perception de ces redevances. C'est pourquoi **le GFII ne souhaite pas la reprise dans la transposition d'une tarification au coût marginal, mais que le principe de gratuité des données publiques soit réaffirmé.**

Mais, le GFII insiste aussi sur la nécessité de veiller à ce **que la mise en œuvre de ce principe ne conduise pas de fait à la disparition de certaines informations ou à l'altération de la qualité des données produites** : fréquence de mise à jour, complétude, exhaustivité, formats... La qualité des données est en effet une condition indispensable au développement d'un écosystème pérenne autour de ces données.

Ainsi le GFII propose que la loi de transposition de la nouvelle Directive reprenne le principe de gratuité, tout en ménageant la possibilité de tarification dans un certain nombre de cas particuliers, sous réserve que :

- Soient très clairement repris et détaillés dans la loi de transposition les critères de tarification « objectifs, transparents et vérifiables » prévus par la Directive ;
- Que les conditions d'accès soient neutres et ne comportent pas de discrimination entre utilisateurs publics ou commerciaux, tant en ce qui concerne la tarification que la qualité des données, la fréquence de mise à jour, la structuration, le périmètre offert *etc.* Un utilisateur externe doit bénéficier, par exemple, des mêmes conditions de mise à jour ou de structuration que l'administration pour ses usages propres et le calcul des coûts doit être identique dans les deux cas.

Une tarification pourrait donc être autorisée dans les cas suivants :

- **les organismes tenus de générer des recettes**

Dans les cas où des organismes publics, tenus de générer des recettes destinées à couvrir une part substantielle des coûts liés à l'accomplissement de leurs missions de service public (Directive art 6-2a), sont autorisés à percevoir des redevances, ceux-ci doivent pratiquer une politique tarifaire transparente s'appuyant sur une analyse des coûts et respecter la libre concurrence tout en

permettant l'émergence de nouveaux acteurs (il ne doit pas y avoir de barrière à l'entrée pour les nouveaux entrants). Dans ce cas, l'adoption d'une comptabilité analytique par les administrations qui tarifient les informations publiques est une condition *sine qua non*. La loi doit également prévoir que l'organisme soit susceptible d'être audité par l'instance de régulation, pour justifier de la conformité des coûts pris en compte dans les modes de calcul de la redevance. En conclusion, il est important que les coûts soient justifiés, équitables (égalité de traitement entre tous les réutilisateurs) et permettent le développement de projets économiquement viables, les coûts devant pas constituer une barrière à la réutilisation.

- **les données demandant des traitements spécifiques préalables à leur ouverture**

De nombreux jeux de données restent aujourd'hui inaccessibles car les établissements publics ne bénéficient pas des budgets suffisants pour les mettre à disposition dans des formats lisibles par machine. Dans ces cas spécifiques, il serait souhaitable qu'un régime dérogatoire soit mis en place afin de permettre l'investissement qui rendra ces données accessibles (cf. Directive art. 6-2b). Le périmètre de l'investissement pris en compte pour l'établissement d'un tarif devrait se limiter strictement aux traitements spécifiques et au développement de plateformes allant au-delà d'une simple mise à disposition des données produites dans le cadre de la mission de service public. Le GFII suggère que les coûts pouvant être pris en compte pour définir l'assiette soient clairement définis : licence des logiciels utilisés spécifiquement pour l'anonymisation, l'exportation ou la conversion des données, adaptation de la base de données notamment pour l'exportation des données, améliorations/accroissement des capacités d'hébergement (hardware), infrastructures réseaux

En raison de la nature des investissements pris en compte, les régimes dérogatoires ne pourraient excéder une période de 5 ans.

- **la mise en place de services accompagnant la mise à disposition de données, dans le cadre d'une stratégie de plateforme**

Suivant les recommandations du rapport Trojette relatif à l'ouverture des données publiques (novembre 2013), le GFII suggère que, dans le cadre de la mise en place d'une plateforme offrant des services à valeur ajoutée, l'administration puisse tarifier le service rendu par la plateforme, au titre d'un retour raisonnable sur l'investissement, en tenant compte des critères suivants :

- du volume des transactions sur la plateforme (nombre de connexions sur une durée limitée),
- des accès privilégiés aux données dans le temps (durée d'embargo, avec un accès payant avant le terme de l'embargo),
- de la fraîcheur des données (fréquence et/ou rapidité de mise à jour),
- du débit de la connexion,
- du recours au service d'information en ligne sur la structuration des données (*hotline*)...

## 7. Mettre en place une instance de recours disposant de moyens juridictionnels et budgétaires

Les réutilisateurs et les organismes publics ont besoin d'une institution unique de référence garante des règles applicables à la réutilisation des données publiques afin de faciliter les démarches, d'accélérer les contractualisations entre acteurs privés et organismes publics et de limiter les recours juridictionnels classiques.

Cette instance pourrait prendre la forme d'une **autorité administrative indépendante spécialisée à compétence nationale étendue** (administrations centrales, collectivités territoriales, établissements publics...). **Ses missions seraient l'information, la régulation, la protection, le contrôle, la sanction et l'anticipation.** Elles s'exerceraient aussi bien auprès des réutilisateurs que des organismes publics et concerneraient l'intégralité des problématiques liées à la réutilisation : accès, dérogations, formats, licences, tarifications...

Dans le paysage administratif français actuel, le GFII estime que **la CADA** bénéficie d'une histoire et d'une compétence qui la désignent comme la plus à même de devenir cette autorité indépendante sous réserve qu'elle dispose de pouvoirs renforcés, notamment d'auto-saisine, de moyens, notamment financiers, accrus et de véritables pouvoirs d'investigations, d'injonction et de sanction vis-à-vis des administrations et des ré-utilisateurs. Un élargissement de sa composition serait sans doute nécessaire afin d'intégrer d'autres expertises en matière technique mais également en matière de tarification et de concurrence.

Un parallèle utile peut être fait avec la CNIL pour illustrer la compétence et le champ d'actions de cette future autorité, notamment en termes d'auto-saisine.

## 8. Préciser les règles relatives à la réutilisation des informations publiques pouvant contenir des données à caractère personnel

Le dispositif actuellement en vigueur en France est beaucoup plus restrictif que celui posé par la Directive et la Directive révisée sur la réutilisation des informations publiques. Dans ces textes, seule est rappelée l'exigence de réutiliser les données dans le respect des principes relatifs à la protection des données à caractère personnel, conformément à la Directive de 1995 applicable en la matière. Le projet de règlement européen sur la protection des données personnelles n'apporte pas de modification sur le sujet.

- **Dissonances des lois CADA/CNIL**

L'article 13 de la loi CADA pose deux séries de conditions à la réutilisation des données publiques à caractère personnel :

- le 1er alinéa prévoit que cette réutilisation est possible lorsque la personne intéressée y a consenti, ou si l'autorité détentrice est en mesure de les rendre anonymes, ou à défaut d'anonymisation, si une disposition législative ou réglementaire le permet ;
- le 2ème alinéa prévoit que la réutilisation est subordonnée au respect des dispositions de la loi CNIL. L'article 7 indique notamment que tout traitement de données à caractère personnel doit avoir reçu le consentement de la personne concernée, sauf s'il satisfait à certaines conditions. Parmi celles-ci figure « la réalisation de l'intérêt légitime poursuivi par



le responsable du traitement ou par le destinataire, sous réserve de ne pas méconnaître l'intérêt ou les droits et libertés fondamentaux de la personne concernée ».

Il en résulte que, si la réutilisation de données publiques à caractère personnel ne respecte pas les conditions de l'article 13 de la loi CADA, alors que le traitement des données aurait été autorisé au titre de la loi CNIL, cette réutilisation est impossible.

Cette situation a donc une double conséquence. D'une part, elle est source d'insécurité juridique, certains pouvant avoir une interprétation libérale des textes qui risquerait d'être contredite par une interprétation restrictive. D'autre part, elle peut constituer aujourd'hui un obstacle au développement de la réutilisation des données publiques en France, en étant un prétexte à la non-diffusion de certains jeux de données.

La position du GFII obéit un double objectif : développer de nouveaux services tout en garantissant la protection de la vie privée. Le GFII estime nécessaire d'encadrer la réutilisation des informations contenant des données à caractère personnel, notamment par le moyen de licences pour garantir la protection de la vie privée et le respect des règles encadrant l'utilisation des données à caractère personnel.

- **Trancher entre les règles CADA et CNIL, en faveur des seules règles CNIL, sans cumul entre-elles**

Tenant compte des blocages qui ont pu être constatés et en s'appuyant sur les avis rendus par la CNIL, le GFII estime que le caractère réutilisable d'une information contenant des données à caractère personnel ne doit pas être jugé en fonction :

- de la nature de l'établissement qui la détient ou l'a collectée (établissements publics),
- de l'aspect commercial ou non de la réutilisation,

Pour dépasser les contradictions entre les lois CADA et CNIL, il semblerait pertinent que seuls les éléments suivants soient pris en compte dans le cas où les personnes n'auraient pas consenti préalablement à la réutilisation des informations :

- la destination première de l'information d'origine : des données personnelles destinées à être rendues publiques par l'administration, notamment pour satisfaire à l'information nécessaire et minimale des citoyens ne peuvent faire l'objet d'une anonymisation ou d'une interdiction de réutilisation, par exemple les noms des représentants légaux ;
- la sensibilité des informations concernées au sens des « données sensibles » définies par la CNIL (origines raciales ou ethniques, opinions philosophiques, politiques, syndicales, religieuses, vie sexuelle ou santé des personnes) ;
- la nature du traitement effectué sur les données à caractère personnel ;
- la finalité du traitement réalisé par le réutilisateur ;
- la conformité des dispositifs mis en place par le responsable de traitement avec les recommandations de la CNIL en matière de déclaration ou d'autorisation préalable.

- **Faire prévaloir la compétence de la CNIL sur le sujet des informations publiques contenant des données à caractère personnel**

Pour éviter toute dissonance, le GFII préconise que la réutilisation des informations publiques contenant des données à caractère personnel relève de la seule compétence de la CNIL.

- **Préciser le cadre de réutilisation des informations contenant des données à caractère personnel avec des licences**

Dans les cas où des contraintes spécifiques sont liées à la nature des données (données contenant des informations à caractère personnel...), le GFII préconise la mise en place de licences- type exemples de questions résolues par les licences qui, en plus d'un rappel général du cadre légal, définiront les engagements respectifs des producteurs et réutilisateurs en matière de fourniture et de réutilisation des données : par exemple, données fournies en l'état conformément aux missions du producteur, format et périodicité de mise à disposition, respect des données personnelles, respect des droits de propriété intellectuelle etc.

Le producteur d'information sera ainsi en mesure de contacter les réutilisateurs en cas de manquement manifeste au respect des obligations prévues par la licence ou si des modifications doivent être répercutées.

Les membres du GFII, qui mettent à disposition ou rediffusent ces données dans le cadre de licences soulignent la nécessité de pouvoir informer et être informés en cas de problème lié à des défauts d'anonymisation par exemple ou de demande de suppression d'informations par les personnes concernées.

- **Sécuriser le cadre des données ayant été anonymisées**

L'anonymisation ne consiste pas simplement à supprimer le nom d'une personne dans un document ; il s'agit d'ôter du document tout caractère qui permettrait d'identifier une personne, directement ou indirectement. Dans certains cas, cette opération peut s'avérer complexe (exemples).

A l'heure des big data et du croisement de gros volumes d'informations de sources multiples, il est envisageable que des jeux de données qui ont été préalablement anonymisés puissent être associés à des faisceaux de données non directement personnelles mais cependant reliées aux personnes (habitudes de vie et de consommation, utilisations d'Internet, contenus des courriers électroniques, numéros de téléphone, etc.). Il en résulte des moyens accrus d'enrichissement des données qui peuvent aboutir à la ré-identification de personnes.

C'est pourquoi, afin de prévenir tout risque, le GFII suggère qu'une licence soit également proposée pour les jeux de données anonymisées, qui, du fait de leur nature, ne doivent pas être considérés comme des jeux de données open data. Cette licence pourrait mentionner :

- le fait que le jeu de données a fait l'objet d'une anonymisation ;
- l'interdiction de recoupements d'informations anonymisées ou toute autre pratique permettant de rendre ces données identifiantes (loi du 6 janvier 1978 modifiée) ;
- la responsabilité du réutilisateur qui volontairement croiserait les données en vue d'identifier les personnes concernées.

Le producteur d'information doit être aussi en mesure de contacter les réutilisateurs si des modifications doivent être répercutées (dans le cas de problème lié à des défauts d'anonymisation par exemple) ou en cas de manquement manifeste au respect des obligations prévues par la licence.

De son côté, l'utilisateur sera sécurisé quant à l'étendue exacte de ses obligations contractuelles.

- **Assouplir les règles concernant la charge de l'anonymisation**

Dans le cas où les données doivent être anonymisées, la loi CADA prévoit que l'opération incombe à l'administration qui ne peut s'y soustraire que si l'opération entraîne des efforts disproportionnés. En revanche, la CNIL indique que l'anonymisation peut être réalisée par les réutilisateurs eux-mêmes, sous certaines conditions.

Le GFII souhaite que la loi soit modifiée : lorsque l'administration n'est pas en mesure de procéder aux opérations d'anonymisation, notamment si ces opérations engendrent des efforts disproportionnés, le réutilisateur pourrait procéder lui-même aux opérations d'anonymisation, dans un cadre défini et contrôlé par l'administration concernée, via une licence de réutilisation. Une telle licence, en plus des clauses traditionnelles, devrait prévoir *a minima* :

- L'obligation expresse d'anonymisation ;
- La possibilité pour l'administration de procéder à des contrôles réguliers ;
- L'obligation pour le réutilisateur de fournir un échantillon d'informations anonymisées ;
- L'obligation pour le réutilisateur d'expliquer lors de la conclusion de la licence les modalités d'anonymisation ;
- L'obligation pour le réutilisateur d'informer à échéance régulière l'administration sur la mise en œuvre concrète du processus d'anonymisation ;
- L'obligation pour le réutilisateur de transmettre à l'administration toute « plainte » de personne intéressée, de procéder sans délai à l'anonymisation des documents similaires qui ne l'auraient pas été à tort et d'en informer l'administration ;
- Des sanctions en cas de manquement aux obligations prévues dans la licence ...

Le GFII souhaite en outre rappeler que le fait que des jeux de données contiennent des données à caractère personnel ne pouvant pas être réutilisées ne doit pas conduire à l'interdiction de réutilisation de l'ensemble du jeu de données ou à une anonymisation systématique mais à la mise en place de moyens d'extraction permettant d'exclure ces données.

## **9. Définir une véritable gouvernance des données et accompagner les administrations dans la publication**

L'effort entrepris par le Secrétariat général à la modernisation de l'action publique pour développer une véritable gouvernance des données, avec un cadre d'architecture des référentiels de données, doit être poursuivi.

Cela signifie un recensement voire une rationalisation des référentiels de données existants et l'obligation pour les administrations de les convertir, au moins pour l'avenir, dans des formats de type XML ou RDF favorisant une réutilisation maximale (enrichissement, croisement, mélange, interconnexion avec le web des données).

Cela passe également par la mise en place de référentiels d'autorité communs à toutes les administrations.

Ce travail de normalisation exige des moyens, mais c'est un facteur de modernisation de l'action publique. Par ailleurs, il est indispensable qu'il soit mené au sein de l'administration et diffusé largement pour éviter l'élaboration et la privatisation de référentiels par de grands opérateurs du web.

Publier des données et des métadonnées (données sur le contenu, la source, le format) est désormais un point de passage incontournable de toute stratégie de diffusion de données. Aussi, l'obligation de publier des métadonnées est-elle maintenant souvent inscrite dans les textes, sans en fournir la méthode, laissant les agents publics responsables de ces publications malheureusement désarmés. Des outils et des méthodes standardisés, appuyés par des actions de formations doivent être mis en œuvre, permettant ainsi à l'administration de mener ses processus de manière intégrée.

L'expérience tirée des réglementations de la sécurité des systèmes d'informations (RGS) constitue un bon exemple d'une pratique réglementaire qui doit s'appliquer aux métadonnées : un processus de définition ouvert des formats à respecter puis l'adoption de textes contraignants précis et la mise en œuvre d'une politique de sensibilisation auprès de l'administration.

La mise en place de lieux de rencontres et échanges public/privé sur les référentiels et les formats semble indispensable.

\*\*\*\*\*

Le GFII considère que les pouvoirs publics doivent saisir l'opportunité de la transposition pour créer un écosystème économique innovant, associant producteurs, diffuseurs et agrégateurs d'informations, laboratoires de recherche, professionnels de la dématérialisation, start-up et nouveaux acteurs de l'économie numérique. Ces 9 propositions définissent, selon le GFII, les conditions optimales permettant le développement économique à partir de la réutilisation des informations publiques.

### A propos du GFII, le think tank des acteurs du marché de l'information et de la connaissance

Le GFII rassemble l'ensemble des acteurs de la chaîne de l'information spécialisée, des secteurs public et privé : producteurs d'information, éditeurs scientifiques, éditeurs juridiques, éditeurs de contenus professionnels diffuseurs, agrégateurs, agences d'abonnements, éditeurs et prestataires de solutions (veille, text mining, sémantique...), cabinets de conseil, cabinets d'avocats, cellules de veille, départements R&D, département d'études et prospective centres de documentation de grandes entreprises, bibliothèques universitaires... Ce positionnement transversal fait du GFII un lieu unique pour les échanges entre professionnels de l'information numérique.

Le GFII permet aux acteurs de l'industrie de l'information de se rencontrer, de confronter et d'échanger leurs points de vue sur les aspects juridiques, techniques et économiques du secteur. Le GFII anime des groupes de travail sur les sujets stratégiques : open data, open access, web sémantique, big data, social data, datamining et text mining, veille et traitement avancé des données, économie de la connaissance, enjeux juridiques...

### A propos du groupe de travail Données publiques

Le GFII a créé en 1997 un groupe de travail consacré à la diffusion des informations publiques. Le GFII a ensuite élargi son action en accueillant d'autres associations au sein de ce groupe :

- AFIGEO : l'association Française pour l'Information Géographique
- FIGEC : Fédération Nationale de l'Information d'Entreprise et de la Gestion de Créances
- FING : Fondation Internet Nouvelle Génération
- FNPS : Fédération française de la Presse Spécialisée
- GESTE : Groupement des Editeurs de Services en Ligne
- SPDG : Syndicat Professionnel De la Géomatique
- FEDAE : Fédération des Auto-Entrepreneurs

Le groupe inter-associations rassemble ainsi les principaux acteurs économiques du secteur, publics ou privés, participant à la chaîne de valeur de l'information : producteurs, réutilisateurs publics et privés, éditeurs de logiciels sémantiques ou cartographiques permettant d'améliorer la qualité et le potentiel de cette information, utilisateurs en entreprises, chercheurs... Les participants du groupe de travail sont actifs dans tous les secteurs de l'information : juridique, géographique, propriété industrielle, informations sur les entreprises, territoriale, environnementale, scientifique, santé, transport, tourisme ...

Parmi les organismes participants figurent : Altarès, Bilans Service, Caisse des dépôts, Canope, CCI Ile de France, Coface Services, DILA, IGN, INPI, Editions Francis Lefebvre, Editions Législatives, Groupe Moniteur, Guillaume Prigent Avocat, Guy Lambot Avocat, HBS Research, Infolegale, Lawcracy, LexisNexis, Notrefamille.com, OIEAU, Orange Business Services, SVP, Temis, Thibault du Manoir de Juaye Avocats, Transpacity, Vecteur Plus, Wolters Kluwer France, Zebaz...

Le groupe est animé par Bernard Marx (GFII) et Denis Berthault (LexisNexis).

Ce groupe est un lieu de rencontre, de débats et d'échanges qui favorise les synergies entre les acteurs afin de favoriser une meilleure réutilisation des nombreux gisements d'informations publiques, et ainsi contribuer de manière dynamique au développement de l'économie de la connaissance. Les membres du groupe de travail bénéficient par ailleurs d'une veille quotidienne sur les données publiques via le réseau social du GFII, AMICO (les Acteurs du Marché de l'Information et de la Connaissance).

Le GFII participe notamment aux travaux du Comité d'Orientation de l'Édition Publique et de l'Information Administrative (COEPIA).

Contact : Ruth Martinez, Déléguée générale du GFII, 06 08 83 25 01, ruth.martinez@gfii.fr