

Position commune FNPS-GFII-SNE

Cadre de réutilisation des décisions de justice

Les acteurs du monde du droit et de l'ouverture / réutilisation des données publiques que nos organisations représentent expriment depuis plus de dix ans leur attachement à l'ouverture et à la réutilisation des données publiques, dont celles des décisions de justice, et se sont à plusieurs reprises positionnés pour en accompagner la mise en œuvre¹.

L'ouverture des décisions de justice consacrée par la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice et héritée des réflexions entamées dans le cadre de l'adoption de la loi pour une République numérique place en nécessité absolue le respect des données personnelles, de la vie privée et de la sécurité des personnes.

La loi vise à mettre en place un open data au bénéfice **de tous types de publics** sans distinction, qu'il s'agisse de professionnels, en particulier du droit et de la compliance, d'entrepreneurs de la data, de citoyens, etc.

Nos organisations proposent, depuis de nombreuses années, en complément à ce nouvel accès public au flux anonymisé, l'accès à **un flux de décisions « intègres »**, sans occultation préalable, destinées principalement à enrichir et développer des services dédiés aux clients finaux professionnels du droit, en parallèle de la diffusion des décisions anonymisées.

Dès lors que nous aspirons à la création de deux flux différents, il conviendra de développer deux cadres distincts de réutilisation :

1. Un encadrement, assorti d'une pédagogie et d'une attention particulières pour la mise en œuvre de bonnes pratiques s'agissant de l'accès, de la réutilisation et de la diffusion du flux anonymisé destiné au public.
2. Un encadrement renforcé, dans le prolongement des licences déjà existantes, pour l'accès et la réutilisation du flux intègre des décisions, adapté au développement de services destinés à des clients finaux professionnels du droit.

¹ Cf. compte rendu d'audition du SNE annexé (p.116) au rapport sur "l'open data" des décisions de Justice rendu par Loïc Cadet, professeur à l'École de droit de la Sorbonne (université Paris-I), le 9 janvier 2017 à la garde des Sceaux.

1. Encadrement de l'accès au flux anonymisé de données de justice pour une réutilisation par le grand public : open data en application de la loi ° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice

Pour ce qui est de l'open data issu de la loi de 2019, destiné au grand public, il faut noter sa singularité car il se distingue de tout open data basé sur des principes d'ouverture et de libre réutilisation tels que décrits dans la licence ouverte 2.0 par certaines précautions essentielles notamment en matière de protection des données personnelles, qu'il s'agisse de la production des données (occultation de 1^{er} niveau), de la modification des données (occultation de 2^{ème} niveau) et de la prise en compte de la volonté exprimée par les personnes protégées.

Il nous semble dès lors évident que ce type de licence n'est absolument pas conçu pour traiter des données devant faire l'objet de restrictions d'usage ou de modifications imposées. Ce constat semble largement partagé puisqu'il a semblé nécessaire d'y adjoindre des conditions générales d'utilisation. Cependant la construction qui nous est proposée repose sur une base fragile, inadaptée et l'ensemble contractuel est d'un accès peu aisé pour des utilisateurs profanes.

Eu égard à cette situation, il nous apparaîtrait plus clair pour les réutilisateurs comme pour les acteurs publics de rédiger **une licence ad hoc « grand public », comme le permet l'article D323-2-2 du CRPA.**

Il nous semble par ailleurs utile de compléter une mécanique contractuelle améliorée par la **signature d'une charte**, rédigée sous l'égide du ministère de la justice et des responsables de l'ouverture des données de justice, avec l'ensemble des acteurs concernés, dont les différentes catégories d'utilisateurs potentiels.

Cette charte doit :

- expliciter le cadre juridique de l'ouverture des décisions de justice, afin de s'assurer que tous les réutilisateurs des décisions « grand public » auront une approche la plus professionnelle et rigoureuse possible de l'utilisation des données ;
- donner des interprétations aux flous laissés par la loi sur certaines notions (atteinte à la sécurité ou au respect de la vie privée de ces personnes ou de leur entourage)
- mieux définir les activités de profilage et de ré-identification proscrites
- tirer profit des usages et expériences acquis par le passé sur la mise en œuvre de l'open data pour préciser des bonnes pratiques.

Pour compléter ce dispositif, il est souhaité la **création d'une instance de suivi et de régulation**, destinée à réunir les fournisseurs de données (ministère de la justice/Cour de cassation/Conseil d'Etat) et tous les réutilisateurs. Elle permettrait d'échanger sur les aspects techniques (qualité des données, modalité de livraison, post-anonymisation...) et suivre les conditions de réutilisation des données fournies liées à l'exécution de l'open data des décisions.

Cette instance se réunirait régulièrement, idéalement selon une fréquence trimestrielle, et ce pendant toute la durée du projet (donc jusqu'à fin 2025). Son secrétariat pourrait être assuré par la Cour de cassation.

2. Encadrement renforcé pour l'accès et la réutilisation du flux intègre des décisions de justice - Clients finaux professionnels

L'ouverture des décisions de justice ne peut pas se résumer à un Open data visant un public large et réduit à la livraison d'informations tronquées voire dégradées.

Comme l'a évoqué le président Sommer lors du dernier groupe de travail, l'Open data n'est pas le seul moyen de diffusion des données publiques. C'est d'ailleurs en dehors de ce cadre que les éditeurs juridiques accèdent depuis de nombreuses années à des flux de décisions de justices.

Les besoins des professionnels du traitement de l'information, en particulier de l'information juridique, tant pour assurer leur mission naturelle d'analyse et de commentaires doctrinaux de la jurisprudence, que pour le développement de services innovants, sont distincts de ceux du grand public et supposent l'accès à ce flux intègre.

A ce sujet, en premier lieu, la fourniture de données intégrées aux professionnels est indispensable pour identifier et éviter tout biais algorithmique dans le traitement par l'intelligence artificielle (IA) de bases de données judiciaires. D'autre part, **il est important que toute diffusion, de quelque manière que ce soit, d'une décision non anonymisée ou de quelque donnée personnelle que ce soit, à qui que ce soit ne puisse être opéré que dans un cadre professionnel strictement délimité.**

Cet accès aux décisions de justice, qui va au-delà de l'open data, si elle est correctement encadrée, procure un avantage de compétitivité déterminant pour les entreprises qui en sont bénéficiaires car il est une condition de leurs capacités d'innovation.

En outre, les enjeux relatifs à cette ouverture ne sont pas seulement économiques. La nature des données concernées pose des questions fondamentales relatives au fonctionnement de notre démocratie, qu'il s'agisse, du droit à l'information, ou plutôt à une information de qualité, vérifiée et non tronquée ou de l'égalité entre les acteurs du droit, qui élaborent la jurisprudence par leurs décisions ou qui expriment des arguments en demande ou en défense. Ces enjeux justifient selon nous le maintien de l'existence d'un flux intègre, strictement encadré et sécurisé, parallèle au flux anonymisé.

La donnée juridique intègre n'est donc, à l'évidence, pas une donnée comme les autres et doit donc obéir à un cadre spécial apportant des garanties et répondant à des préoccupations éthiques, opérationnelles et techniques spécifiques.

C'est pourquoi la FNPS, le GFII et le SNE soutiennent la mise en œuvre, dans le prolongement des relations contractuelles existantes, d'une solution hybride rendant possible l'accès à un flux intègre, destiné aux réutilisateurs professionnels, avec des garanties renforcées dans l'utilisation de ces flux telles que : l'exercice d'un contrôle en amont et aval par une instance dédiée ainsi que la signature d'un contrat de licence ad hoc.

L'exercice d'un contrôle amont/aval par une instance de contrôle dédiée

- En amont, au moment de la signature de la licence, cette instance vérifierait que certains prérequis techniques et juridiques sont réunis sur la base d'engagements formels du réutilisateur
- En aval, cette instance procéderait à un contrôle effectif qui se doublerait d'une analyse de l'adéquation et de la qualité des outils ou applications mis en œuvre par les réutilisateurs.

La signature d'un contrat de licence

Le cadre conventionnel est absolument essentiel dans la mesure où il fixe avec précision les modalités d'accès, de réutilisation et de diffusion des décisions, de même que les modalités de contrôle.

Les réutilisateurs professionnels sont parfaitement conscients des nécessités juridiques et opérationnelles attachées au traitement d'un flux intègre de données.

Les licences en cours, qui ne concernent pas le périmètre de l'open data introduit par la loi de 2019, peuvent *dès lors* subsister en parallèle d'un open data des données anonymisées qui ne répond pas aux mêmes standards d'exigence.

Qu'il s'agisse d'adapter les licences en cours ou de formaliser des licences pour l'avenir, les réutilisateurs mentionnés ci-dessus sont favorables à un renforcement des exigences en matière de stockage, collecte, traitement ou transfert de données, en prévoyant notamment :

- des limitations à certaines utilisations afin de prévenir des utilisations indésirables
- les modalités de contrôle du bon usage et de la sécurité des données :
 - o contrôle à tout moment de l'usage qui est fait des données en conformité avec les conditions fixées par la loi et la licence,
 - o contrôle sur la sécurité de l'hébergement et des traitements automatiques ou manuels effectués,
 - o contrôle des « produits » et services développés par le réutilisateur à partir de ces données,
- des sanctions spécifiques en cas de manquement aux obligations de bon usage et sécurité des données :
 - o interruption du flux pour non résolution du manquement,
 - o interruption définitive du flux en cas de manquement grave,
- L'obligation pour le réutilisateur de décrire ses activités :
 - o canaux de diffusion des prestations et décisions anonymisées,
 - o description des prestations et services développés avec les données concernées,
 - o description des algorithmes utilisés et des bonnes pratiques mises en œuvre pour prévenir, détecter, et alerter sur de possibles biais,
 - o description des processus permettant d'assurer la sécurité et la mise à jour des données.

Ce cadre renforcé d'utilisation du flux intègre des décisions de justice à destination des professionnels brièvement esquissé, nous semble de nature à offrir des garanties répondant aux préoccupations soulevées par la nature même de la matière.