

Avant-projet LSI : un nouvel accès aux données publiques Note concernant le chapitre II, articles 3 à 6 de l'avant-projet de loi

La publication officielle sur le site du ministère de l'Économie, des Finances et de l'Industrie, en avril dernier, de l'avantprojet de loi sur la société de l'information (LSI), lève enfin le voile sur les tenants et aboutissants d'une politique gouvernementale d'accès aux données publiques très attendue.

LE TEXTE PUBLIÉ est la dernière mouture d'un grand projet qui a dû faire face à de nombreux arbitrages, tenir compte des fortes demandes de groupes de pressions aux intérêts opposés et naviguer sous l'œil inquisiteur de l'Union européenne et dans le cadre de ses directives ou de ses projets. Bref, après ce parcours difficile, on pourrait s'attendre à trouver un texte en manteau d'Arlequin, or on est agréablement surpris de découvrir un ensemble de mesures, somme toute, cohérentes. Reste que l'on regrettera l'absence de dispositions attendues concernant, par exemple, les droits d'auteur. Mais le périmètre de ce texte risque encore de bouger. Le projet de loi devrait être présenté prochainement en Conseil des ministres et son vote, au dire des observateurs, n'interviendra qu'après la prochaine législature, soit dans au moins un an. Pour chaque domaine traité par le projet de loi, l'analyse des travaux préparatoires et des nombreuses versions du texte qui se sont succédé, amène un éclairage intéressant sur la volonté politique des rédacteurs du projet. C'est notamment le cas pour son chapitre II ?Accès aux données publiques?. L'aménagement de l'accès aux données publiques a donné lieu à de nombreuses analyses. Un sujet longuement débattu est celui du risque de concurrence déloyale que pourrait entraîner la commercialisation de ces données par l'administration en charge de les produire 1. Plusieurs circulaires avaient appelé, sans grand succès, le secteur public à plus de retenue dans ce domaine. La plus célèbre d'entre elles, signée par le Premier ministre Édouard Balladur le 14 février 1994 2, soulignait qu'en matière de diffusion de données publiques, le ?faire faire? sera généralement préférable au ?faire?. Chargé par le Commissariat général du plan d'une large réflexion sur la ?diffusion des données publiques et la révolution numérique?, le Président Mandelkern a rendu son rapport le 16 novembre 1999 ; celui-ci constitue l'une des bases du chapitre II de l'avant-projet de LSI 3. Au vu de tout ce qui précède, l'analyse du projet appelle une réflexion préliminaire. Il n'est plus fait référence, en effet, aux diffuseurs institutionnels (Journal officiel, Imprimerie nationale, Documentation française, Réunion des musées nationaux, IFEN?), contrairement aux textes et rapports préparatoires. Ceux-ci sont donc traités dans l'avant-projet de loi sur le même plan que les autres institutions (publiques ou privées) chargées d'une mission de service public et détenant/produisant des données publiques. Ce qui ne veut pas dire qu'aucune disposition particulière ne s'applique à eux. Seulement, celles-ci ne relèvent pas du domaine législatif, mais plutôt des circulaires. Certaines existent déjà, elles encadrent leur mission de diffuseur public 4. Il existe également dans l'avant-projet de LSI une double ambiguïté apparente, d'une part, de définition et, d'autre part, d'accessibilité. L'idée qui est

développée est que les données essentielles sont mises en ligne gratuitement sur l'internet pour le public (art. 4 II et III) et sous certaines conditions déontologiques et potentiellement financières pour un secteur marchand intéressé par leur valorisation et leur exploitation. Les autres données (considérées comme non essentielles) sont mises à disposition directement par les administrations ou par l'intermédiaire de tiers privés, à des conditions pouvant inclure une valorisation financière. Une institution de médiation, veillant à l'équilibre des règles économiques, devrait être créée. I ?

SUR LA DÉFINITION DES DONNÉES PUBLIQUES

Qu'apporte l'avant-projet de loi sur la société de l'information par rapport à la loi du 17 juillet 1978 sur l'accès aux données publiques ? La réponse est, clairement, la réutilisation commerciale des gisements (bases) de données publiques. La loi de 1978 excluait cette réutilisation commerciale et parlait de documents individuels auxquels les citoyens pouvaient avoir accès. La LSI va transformer ce concept en l'ouvrant à la diffusion et l'exploitation. Les deux lois ont en commun la définition des données concernées qui reste celle retenue par la loi de 1978. Toutefois, les données sont regardées dans la LSI comme des gisements pouvant constituer des bases, et non plus comme des documents. Néanmoins, la rédaction de l'article 3 de l'avant-projet de loi (alinéas 1 : « tiennent à la disposition du public » et 3 « données dont la communication est demandée ») montre bien que l'on se situe encore majoritairement dans le cadre des dispositions de la loi de 1978, ce qui peut décevoir ceux qui espéraient une obligation plus forte de mise en ligne, idée pourtant maîtresse du PAGSI (Programme d'action gouvernementale pour la société de l'information). La définition des données essentielles (article 4 I), même si celle-ci reste dans la ligne du rapport Mandelkern (importantes pour la vie du citoyen, dont la production est permanente, systématique?), est plus restrictive que ce qu'on pouvait attendre. Il s'agit avant tout des données ayant une obligation de publicité (données juridiques principalement), des données concernant le fonctionnement des administrations, des études et rapports concernant l'amélioration des relations entre l'administration et le public (données d'état civil dont la mise à disposition était déjà prévue par la loi de 1978, par exemple). Cette lecture restrictive exclut, par exemple, la jurisprudence des tribunaux comme faisant partie des données essentielles, bien que le Comité interministériel pour la réforme de l'État (CIRE) du 12 octobre dernier l'avait retenue dans sa définition. Il semble en fait que les rédacteurs du projet de loi aient voulu engager le moins possible l'État par une définition législative. Mais l'annonce de la dénonciation anticipée de la concession Jurifrance (courant 2002 avant son échéance contractuelle de 2004) prouve, en fait, que l'administration travaille à la mise en ligne des décisions de justice dans un portail Légifrance agrandi 5. Autre preuve de cette prochaine mise en ligne des arrêts des cours et tribunaux, les fonds de concours ? sorte de droits d'expédition que les greffes avaient institués avec le temps et dont les revenus sont devenus aujourd'hui nécessaires à leur travail ? devraient être remplacés par une augmentation du budget de fonctionnement des juridictions. On devrait trouver ainsi, de façon partielle pour les cours d'appel et de façon exhaustive pour les décisions des hautes cours, les grands arrêts de la République. La mise en ligne des arrêts des tribunaux de première instance se heurtant à des problèmes de numérisation, ceux-ci ne seraient

pas proposés sur Légifrance et leur disponibilité, pour les opérateurs du marché intéressés, serait plus aléatoire. L'article 4 I, al. 5 de l'avant-projet de loi prévoit que de nouvelles catégories de données essentielles pourront être définies par décret en Conseil d'État. Le périmètre de la mise à disposition du public de ces données continue donc de dépendre de la volonté gouvernementale. A contrario, toutes les données publiques qui ne sont pas essentielles et dont la communication n'est pas interdite ou réglementée par la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, sont à la disposition du public. Encore faut-il les connaître ! On peut regretter que l'obligation faite aux détenteurs de données publiques de diffuser les annuaires des données communicables accompagnés d'une offre véritable (d'un tarif applicable), mesure imaginée lors des travaux préparatoires, n'apparaît pas dans ce projet de loi.

II ? SUR L'ACCESSIBILITÉ AUX DONNÉES PUBLIQUES

Les producteurs de données essentielles auront l'obligation de les mettre en ligne gratuitement (modèle Légifrance). Et toutes les données publiques, y compris non essentielles mais constituant une base (un gisement) seront mises à la disposition des opérateurs du marché qui en feront la demande, sous certaines conditions.

1. Le texte prévoit une tarification

Concernant les données essentielles, le texte précise qu'elles seront mises en ligne gratuitement et seront réutilisables et diffusables gratuitement sous certaines conditions de qualité (art 4 II et III). Mais il réside une ambiguïté : les dispositions de l'article 3, al. 5 qui prévoient une rémunération, valent pour l'ensemble des données publiques. Deux systèmes de rémunération/contribution financière par des opérateurs privés sont prévus à cet alinéa : ? La rémunération proportionnelle aux ressources tirées d'une exploitation commerciale des données transmises semble fondée sur l'idée de favoriser sur un marché des nouveaux entrants qui seraient, a contrario, pénalisés par une rémunération forfaitaire/ fixe. Ce qui est a priori choquant est la mise en place, par ce biais, d'un intéressement des administrations à la production de données commercialisables. Or, celles-ci risquent d'être amenées à privilégier la production des données les plus ?rentables? Bien que moins ?spectaculaire?, la participation forfaitaire prévue par le texte risque néanmoins d'être la plus dangereuse. Celle-ci devant inclure : « le cas échéant une participation forfaitaire aux dépenses de création, de maintenance et de mise à jour des systèmes d'information nécessaires à leur collecte ou à leur traitement. » Cela veut dire que l'idée initiale, consignée d'ailleurs par le rapport Mandelkern, qui prévoyait une rémunération possible de la fourniture de données cantonnée au coût de leur mise à disposition ? du ?packaging? des données pour les besoins de l'utilisateur qui en a fait la demande ? est abandonnée, puisque la participation au financement des données publiques pourra servir à financer le prix de leur production. Cette rédaction est récente. L'avant-dernière mouture a vu son texte augmenté des ?dépenses de création?, soulignant là la volonté des rédacteurs d'amener le secteur privé à participer au financement de toute la chaîne de production des données publiques. On peut penser que des opérateurs institutionnels comme l'INSEE ou l'INPI sont derrière cette volonté de tarification. L'INPI, par exemple, a une politique commerciale vis-à-vis des autres opérateurs du secteur et, en tant que fabricant de données et opérateur lui-même, fabrique des outils spécifiques à la diffusion de ces données. La logique

développée par l'administration voudrait que si les autres opérateurs du marché ne participent pas à l'élaboration de ces données intelligibles, il n'y aura plus de données disponibles.² Le texte insiste également sur les garanties demandées au diffuseur de données (intégrité, fiabilité) Une fois les données cédées, que peut-on en faire ? Les articles 3 al.4 et 4 II al.2 demandent aux opérateurs de garantir l'intégrité des données. Des licences de réutilisation devront donc être signées par les professionnels qui voudront réexploiter ces données. Elles devraient interdire la cession des données acquises à d'autres professionnels de la diffusion (il ne pourrait y avoir d'opérateurs ?grossistes?), comme elles exigeraient des garanties sur la qualité de mise en ligne des données (fiabilité, intégrité, mise à jour). Ceci est le résultat de l'expérience du minitel où nombre de données économiques servaient de pur produit d'appel ? de faire valoir ? imparfaitement mis à jour, pour attirer le client sur toutes sortes de services. En plus de cette garantie de qualité, il serait également demandé que l'environnement soit acceptable (ex : pas de minitel rose diffusant des données publiques). Les licences d'exploitation écarteraient également l'utilisation de données à caractère personnel. Il s'agit du problème de l' ?anonymisation? des bases de jurisprudence ou de l'exploitation qui peut en être faite et qui ne présente pas un intérêt citoyen (ex. : créer des bases répertoriant les sociétés les plus procédurières ou les mauvais payeurs), Il est intéressant de noter enfin à l'article 3, alinéa 4 l'allusion faite au respect des éventuels droits de propriété intellectuelle des créateurs de ces données publiques. Si les fonctionnaires ont un droit d'auteur, celui-ci est-il automatiquement cédé dans le cadre de leur statut ? L'administration a-t-elle la capacité de signer librement des licences d'exploitation de données publiques ? Détient-elle l'ensemble des droits sur ces données ? C'est l'une des questions que devra débattre le tout nouveau Conseil supérieur de la propriété littéraire et artistique (CSPLA) réuni par la ministre Catherine Tasca le 11 mai dernier 6. Il n'y a pas, a priori, derrière cette volonté de faire signer un contrat d'exploitation, d'idée d'exercer une régulation économique d'un secteur. Seule la qualité des données est recherchée. Mais il conviendrait que l'établissement de ces ?contrats de confiance? entre producteurs publics et diffuseurs privés soient d'accès facile comme les licences de type ?click-and-use? mises en place le premier avril dernier au Royaume-Uni dans le cadre du lancement de sa politique de diffusion de données publiques. III ? SUR LA CRÉATION D'UNE INSTANCE DE MÉDIATION Concernant la création de l'instance de médiation (art 3, dernier alinéa et article 5), celle-ci sera compétente sur la définition des données essentielles communicables, les garanties demandées concernant leur utilisation, le prix de leur mise à disposition ainsi que les normes d'accessibilité de ces données. Cette instance ne sera en revanche pas compétente pour connaître des questions de concurrence déloyale entre le public et le privé dans ce domaine. Ce point d'ailleurs est totalement abandonné par le texte. Certes, le Conseil de la concurrence existe et pourrait, dans le cadre d'actions spécifiques, s'appuyer sur des rapports émis par cette instance, relatifs à des obstructions dans la fourniture de données, par exemple, mais les éditeurs privés espéraient une instance de médiation, par nature plus légère et rapide que le Conseil de la concurrence, compétente également sur ces thèmes. On ne résiste pas, en conclusion, à rappeler que notre perfide Albion vient de nous

coiffer sur le fil en rendant opérationnel son « guichet unique » d'accès aux données publiques au début du mois d'avril. Celui-ci est assez proche de ce que les équipes ministérielles en charge de ces questions en France souhaitent mettre en place. À quelques exceptions près, et celles-ci sont de taille : le coût facturé aux licenciés de la fameuse licence « click-and-use » doit être strictement limité aux coûts de duplication ou d'extraction des données et ne peut incorporer aucune contribution aux coûts de production de ces données ; les administrations centrales peuvent, si elles le souhaitent, développer elles-mêmes des produits d'information à valeur ajoutée, commercialisés sur le marché, mais à la condition de fournir également les données de base à tout opérateur privé concurrent au strict coût marginal ; enfin le dispositif britannique s'appuie sur un guichet central diffusant un registre central des données publiques dans lequel toute administration a l'obligation d'identifier ses collections de données. Même si le projet de loi sur la société de l'information risque d'être voté plus tardivement que prévu, ses possibles retombées sont à analyser dès aujourd'hui. Car il s'agit enfin d'une loi et non plus de simples circulaires dont la portée juridique est restreinte. Ceci veut dire que même les collectivités territoriales seront tenues de prendre les dispositions nécessaires à la diffusion des données publiques qu'elles produisent et détiennent ! Le champ est donc vaste. L'impact d'une telle loi est susceptible d'influer fortement sur le fonctionnement et les finances des administrations locales. De nouveaux opérateurs, pour qui le coût de numérisation des données publiques était un frein à l'entrée sur ce secteur, pourront développer des activités ; de nouvelles demandes d'information pourront être satisfaites, favorisant chez les éditeurs un regain de créativité ; à l'inverse, certains services, en place depuis des décennies, risquent de disparaître. La partie « données publiques » de la LSI va influencer le développement de la presse. Mais pour que ceci puisse se faire dans un environnement économique « sain », il faut que ces nouvelles règles soient stables et acceptées par tous. Les débats parlementaires lèveront certainement les zones d'ombre qui subsistent encore.

Auteur(s) :

Charles-Henri DUBAIL - Directeur de Légipresse

Notes de bas de page :

2. Voir « La commercialisation des données publiques », par C.-H. Dubail, Légipressen° 148-II, p. 11, janvier-février 1998.
3. « Modalités de diffusion des données publiques » Légipressen° 110-I, p. 40, Légipressen° 110-IV, p. 31,
4. Voir extraits des recommandations du rapport du Commissariat général du Plan « Diffusion des données publiques et révolution numérique » et commentaire C.-H. Dubail., Légipressen° 167-IV, p. 137, décembre 1999.
5. Circulaire du 20 mars 1998 relative à l'activité éditoriale des administrations et des établissements publics de l'État, JO22/03/98, p. 4301 Légipressen° 151-IV, p. 41 5. Légifrance est un portail de données juridiques officielles en libre accès ; Jurifrance propose des données juridiques en accès payant dans le cadre d'une convention OR Télématique.
6. Voir le discours de Catherine Tasca d'inauguration du Conseil supérieur de la propriété littéraire et artistique, cahier IV de ce numéro, p. 41