



**CONSEIL DE
L'UNION EUROPÉENNE**

Bruxelles, le 19 juin 2006

10673/06

JURINFO 21

NOTE

de	la délégation française
au:	Groupe "Informatique juridique"
Objet:	L'archivage à long terme des données juridiques numériques

Introduction

I. L'archivage par les pouvoirs publics:

- A) Une contrainte juridique: l'authentification
- B) Une contrainte technique: des méthodes standardisées

II. L'archivage pour les utilisateurs:

- A) La recherche
- B) Les conditions de la réutilisation

Conclusion

Introduction:

Il y a une dizaine d'années, à l'ouverture des « autoroutes de l'information », la question de la conservation à long terme des informations numériques ne se posait pas ; il n'était question que de leur accessibilité et de leur coût. Aujourd'hui, devant la masse considérable de données numériques désormais disponible, la fixation de règles communes pour leur archivage et leur conservation dans la durée est devenue un objectif prioritaire.

Dans le domaine juridique, la question est encore plus sensible : l'exhaustivité des données, aussi anciennes soient-elles, est en effet une condition indispensable à la pertinence du résultat d'une recherche et chacun sait que les lacunes en la matière peuvent être lourdes de conséquence.

C'est une évidence pour tous : la conservation à long terme des données juridiques électroniques de manière sûre et fiable est la garantie d'une juste application du droit. Aujourd'hui, nous savons bien que ce qui est en jeu, c'est la survie des collections imprimées et reliées de lois et des jugements qui garnissent nos bibliothèques de droit depuis des décennies : cette révolution à haut risque n'est acceptable qu'à la condition que la conservation et l'accessibilité permanente et durable de collections électroniques exhaustives soient assurées de manière aussi incontestable que les collections imprimées.

Sécurité juridique et sécurité technique sont en l'espèce intimement liées.

Les questions majeures que nous nous posons sont les suivantes :

Quelles sont les contraintes juridiques et techniques propres au processus de conservation des données juridiques électroniques ?

Il est couramment admis que les émetteurs institutionnels des textes officiels (gouvernements, parlements, cours et tribunaux) sont naturellement les mieux placés pour conserver, à long terme et de manière certaine, les données juridiques qu'ils produisent : cette conservation est-elle assurée à ce jour ?

Quel rôle peuvent jouer les usagers et de leurs intermédiaires professionnels (éditeurs juridiques, bibliothécaires de droit etc.) dans ce domaine ?

Ont-ils un accès permanent aux données juridiques conservées par les émetteurs institutionnels ?

Ont-ils le droit de réutiliser ces données à des fins personnelles, professionnelles ou commerciales et donc la possibilité de les archiver eux-mêmes avec leur propre valeur ajoutée ?

I. L'archivage par les pouvoirs publics:

A) Une contrainte juridique: l'authentification

En France, une loi du 13 mars 2001 a admis que l'écrit sur support électronique avait la même force probante que l'écrit sur papier, à la double condition qu'il soit authentifié, c'est-à-dire signé électroniquement, et qu'il soit archivé de manière à garantir son intégrité, afin que sa valeur probante ne soit pas être altérée par le temps.

Dans le cas des données juridiques, ce sont les pouvoirs publics (les institutions émettrices des textes) qui garantissent que cette double condition est remplie : une ordonnance du 20 février 2004 a redéfini les règles en la matière qui dataient pour la plupart du 19^e siècle, et c'est le 2 juin 2004 qu'a débuté la publication quotidienne simultanée de deux versions du Journal Officiel, l'une sur papier et l'autre sur support électronique authentifié.

Sur le plan juridique, cette publication concomitante des deux versions (papier et électronique authentifiée) est à la fois reconnue comme publication officielle et comme condition nécessaire à l'entrée en vigueur de textes normatifs.

Cependant, sur le plan du contenu, il y a déjà des divergences qui démontrent que le Journal officiel publié en tant que folio unique a déjà disparu, laissant place à diverses publications :

En effet, certains textes ne figurent déjà plus dans la version imprimée et ne sont plus publiés officiellement que par voie électronique : il s'agit notamment des textes relatifs à l'organisation administrative de l'Etat, aux fonctionnaires et agents publics, aux magistrats et aux militaires ainsi qu'au budget de l'Etat ; on considère qu'ils n'intéressent pas les citoyens dans leur vie quotidienne et que les fonctionnaires auxquels ils s'appliquent ont tous un accès à Internet dans leurs administrations respectives pour les consulter.

A l'inverse, est exclue la publication par voie électronique de certains actes individuels relatifs notamment à l'état et à la nationalité des personnes, dans un souci de protection de la vie privée ; la version imprimée sera donc maintenue au moins pour ces textes mais disparaîtra lorsque que l'état des technologies permettra d'empêcher tout usage malveillant des données personnelles en ligne.

L'authenticité de la version électronique est assurée par une chaîne de confiance permettant d'en authentifier la source émettrice : au cœur du dispositif, la Direction des Journaux Officiels qui relève du Secrétariat général du Gouvernement et des services du Premier Ministre. L'acte d'authentification est matérialisé par une clé de certification faisant figurer la date et l'heure de publication ainsi que l'origine de la certification.

Les nouvelles règles de 2004 imposent que la version électronique soit accessible de manière permanente et gratuite comme la version imprimée : je ne parlerai pas ici de la gratuité qui a fait l'objet, en France comme partout ailleurs, de débats passionnés qu'il serait trop long de détailler ici. S'agissant de l'obligation d'accès permanent, elle suppose que les données juridiques électroniques ne soient donc pas seulement archivées mais aussi maintenues en ligne.

L'accès en ligne est assuré sur deux sites officiels :

- une version électronique authentique sur le site du Journal Officiel ;
- une version électronique non certifiée sur le site Legifrance.

L'archivage est assuré de la manière suivante :

- sur le site du Journal Officiel ne sont pas conservés de textes authentifiés antérieurs à la version électronique ; les lois anciennes qui ne sont pas numériques d'origine sont scannées et diffusées sur CD-ROM/DVD depuis 1940, sans certificat. La numérisation est réalisée progressivement et se poursuivra jusqu'aux premiers textes publiés en 1848 ; il faudra ensuite prendre en compte les lois antérieures à la création du Journal Officiel jusqu'au texte le plus ancien, une ordonnance de François 1^{er} du 15 août 1539. L'archivage des textes juridiques certifiés authentiques a lieu par l'envoi mensuel d'un DVD aux Archives Nationales (avec signature électronique). Il est différent du dépôt légal qui est effectué auprès de la Bibliothèque Nationale de France (BNF) : le dépôt légal est une obligation pour toute publication éditée ; elle est effectuée par un robot, mais sans vérification de signature électronique. L'objectif n'est pas de conserver le document authentique mais de constituer une mémoire patrimoniale.
- sur Legifrance, il ne s'agit pas d'archivage à proprement parler : le site met à disposition, à des fins de recherche documentaire, l'intégralité des lois et règlements applicables quelle que soit leur ancienneté, dans leur version d'origine mais aussi dans leur version consolidée intégrant les modifications successives des textes.

Il est important de préciser ici qu'une récente enquête menée auprès des utilisateurs a montré que l'authentification des textes officiels diffusés sur support électronique n'était pas du tout une priorité pour eux et que les problèmes techniques qu'ils pouvaient rencontrer pour afficher un certificat en ligne ne constituait pas un frein à leurs recherches. En effet, il ne semble pas que la diffusion de textes non certifiés sur Legifrance soit dissuasive puisque le site a reçu trente et un millions de visites en 2005. On constate que l'authenticité est le plus souvent présumée par les usagers dès lors que la diffusion est assurée par un site « .gouv ».

La complémentarité des deux sites est donc confirmée à l'usage : d'un côté le site du Journal Officiel qui constitue en quelque sorte le « sanctuaire » où l'on va consulter la version authentique des textes en cas de doute ou de litige et d'autre part le site de documentation juridique Legifrance, doté d'un moteur de recherche puissant.

Les exigences des utilisateurs vont en priorité vers la qualité rédactionnelle des textes et leur accessibilité : l'accès à des textes compréhensibles et à jour est donc un objectif prioritaire pour les pouvoirs publics en France, mais la question de leur authentification n'est pas négligée pour autant ; à noter cependant que la signature électronique des lois anciennes, qui ne sont pas numériques d'origine mais par scannage ou par migration des supports microfiches, n'est pas actuellement envisagée ; seul le stockage sécurisé est assuré.

L'archivage des données juridiques n'est bien entendu pas l'exclusivité de l'Etat : les professionnels de l'information du secteur privé, commercial ou non, ont aussi un rôle à jouer. Pour éviter des redondances coûteuses ou l'obsolescence technologique des systèmes d'archivage, un dialogue, une collaboration, un partenariat apparaît de plus en plus indispensable entre ces différents acteurs (institutions émettrices des textes officiels, bibliothèques universitaires et sociétés commerciales notamment) afin de développer des standards commun de conservation des documents juridiques et de migration de ces documents pour s'adapter aux nouvelles versions de logiciels et de matériels.

B) Une contrainte technique: des méthodes standardisées

La difficulté majeure de l'archivage électronique en général tient à la nécessité d'élaborer une stratégie de conservation des données à long terme dans un contexte technologique très instable. Jusqu'à présent le support microfiche représentait le seul support d'archivage fiable et légal des éditions imprimées du Journal Officiel ; d'une durée de vie estimée à 100 ans et d'une manipulation devenue compliquée, il est désormais obsolète .

La capacité des nouveaux supports augmente mais leur durée de vie est variable avec un risque de rupture de lisibilité. Pour éviter la perte de données, une duplication des supports en des lieux distincts est à envisager et une surveillance régulière est nécessaire ainsi que la reproduction préventive de leur contenu à un rythme périodique, par exemple tous les cinq ans.

La méthode retenue pour les JO en Europe est d'intégrer les documents dans une structure commune XML avant de les exporter, soit dans une base de données soit dans différentes structures XML avec un portail permettant de les retrouver grâce à un système de repérage des méta-données.

Le format PDF 1.4 (PDF/A-1) préconisé par la norme ISO 19005 pour la conservation des données doit permettre à l'avenir un archivage de documents juridiques authentiques, fiables, complets, intacts et exploitables ; ce format doit permettre en effet l'utilisation et l'accessibilité des fichiers au fil de l'évolution des multiples générations successives de supports technologiques.

Toute initiative des institutions françaises en matière d'archivage électronique doit par ailleurs se conformer au programme « ADELE » (ADministration ELEctronique) dont le chapitre « Archivage et cycle de vie du document électronique » vise à réaliser une plate-forme pilote d'archivage électronique pour les Archives Nationales et à produire un référentiel normatif décliné selon cinq points :

- le cadre juridique :
 - la valeur légale des documents électroniques
 - la législation et la réglementation propre aux archives (Code du Patrimoine, Code général des collectivités territoriales)
 - les jurisprudences en matière d'obligations d'archivage
- le processus d'archivage :
 - les normes d'archivage
 - les règles de bonnes pratiques
- le dispositif organisationnel supportant le processus :
 - le rôle et les responsabilités des acteurs
 - les contrats ou conventions les reliant
 - leurs statuts juridiques
- les métadonnées accompagnant le document pour donner à l'archivage le rôle qu'il est censé assurer (juridique, administratif, historique)
 - les métadonnées utiles à la consultation ultérieure
 - les métadonnées utiles à la gestion technique du document
 - les métadonnées utiles à la traçabilité du document

- les moyens informatiques de mise en œuvre jugés sur leurs qualités techniques au regard des exigences juridiques (il s'agit de l'imputabilité, l'intégrité, la traçabilité, la fiabilité) et fonctionnelles (il s'agit de la durée de la conservation et la facilité de la consultation) de l'archivage :
 - les formats des données
 - les formats des métadonnées
 - les technologies de sécurisation
 - les supports et matériels
 - l'architecture technique

II. L'archivage pour les utilisateurs:

A) La recherche

Comme on l'a vu, les nouvelles règles de 2004 imposent en France que la version électronique soit accessible de manière permanente et gratuite comme la version imprimée, ce qui suppose que les données juridiques électroniques ne soient pas seulement archivées mais aussi maintenues en ligne. Le citoyen doit en effet disposer de moyens à sa mesure pour accéder aux règles de droit qui lui sont applicables, sans limite à caractère technique, juridique ou financier ni de restrictions quant aux contenus. Depuis 2002, Legifrance constitue ce service public de diffusion du droit par l'Internet permettant l'accès gratuit du public aux textes en vigueur.

Au niveau du Conseil de l'Union Européenne, deux initiatives sont destinées à faciliter l'accès des citoyens de l'Union aux données juridiques des 25 Etats-membres :

La première a consisté à développer une interface de recherche standard pour les services nationaux. Ce projet, baptisé N-Lex, va permettre à tout citoyen d'un Etat membre d'accéder, dans sa propre langue et d'une façon aussi conviviale que possible, à la législation d'un autre Etat membre.

La deuxième est la création de "LINE"(Legal Information Network in Europe) qui vise à coordonner l'accessibilité des bases de données juridiques européennes et des bases de données nationales et à établir un code de bonnes pratiques communes à tous les Etats-membres dans l'élaboration, la consolidation et la codification de leurs règles de droit.

A noter aussi pour la jurisprudence le réseau des présidents des Cours suprêmes judiciaires de l'Union européenne et de l'Association des Conseils d'État et des juridictions administratives suprêmes de l'Union européenne, ainsi que le Réseau judiciaire européen en matière civile et commerciale.

B) Les conditions de la réutilisation

On l'a vu, les Etats et leurs institutions ne sont pas les seuls acteurs en matière d'archivage des données juridiques et les professionnels de l'information du secteur privé, commercial ou non, ont aussi un rôle à jouer. La valeur ajoutée qu'ils apportent aux données qu'ils traitent et l'aide qu'ils apportent de ce fait aux citoyens dans la compréhension du droit nécessite que les Etats et leurs institutions leur facilitent l'accès aux données juridiques qu'ils produisent. En Europe, la liberté de réutilisation des données juridiques, à des fins essentiellement commerciales, est reconnue par une directive européenne dite « Service Public de l'Information » ; elle ne peut cependant s'exercer que moyennant quelques contraintes, principalement juridiques qui sont posées par les dispositions du code de la propriété intellectuelle relatives d'une part au droit d'auteur et d'autre part aux droits des producteurs de bases de données. Voici les règles en résumé :

- Au-dessous du seuil « substantiel » de quelques milliers de documents par base de données, les téléchargements sont libres ;
- Au-dessus de ce seuil, une licence gratuite est obligatoire ;
- Pour disposer de la totalité de la base, une licence payante doit être demandée, au coût de la mise à disposition des données.

Une convention conclue entre l'Etat et le demandeur précise les conditions d'utilisation des données et, notamment, les engagements pris par le bénéficiaire afin de garantir que l'usage qui en sera fait répond à l'exigence de fiabilité.

En France, la décision d'accorder la licence est prise sur avis du comité éditorial de Legifrance, dénommé comité du service public de la diffusion du droit par l'internet et créé par un décret d'août 2002.

Qu'elles donnent lieu ou non au paiement d'un coût de mise à disposition, toutes ces licences poursuivent le même but, celui de garantir la fiabilité des données utilisées ou réutilisées : ainsi, la licence ne peut être accordée que si le candidat s'engage :

- 1) à respecter l'intégrité des données mises à sa dispositions, réutilisées, rediffusées, en veillant à ce que la teneur et la portée des textes et décisions de justice ne soient pas altérées, notamment par des retraitements (modifications des textes, insertion de commentaires sans que ceux-ci puissent être clairement distingués du texte authentique, coupes altérant le sens de l'acte) de nature à induire le lecteur en erreur. En particulier, le licencié doit respecter, le cas échéant, les droits d'auteur qui s'attachent aux données considérées : ainsi, la reproduction des titrages, abstrats ou résumés de jurisprudence doivent en indiquer la source et ne comporter aucune modification.
- 2) à indiquer
 - la source des données rediffusées (Legifrance) ;
 - la date de leur dernière mise à jour ;
 - les indications utiles pour apprécier la nature et l'origine de l'acte (date, intitulé, auteur...).
 - pour la jurisprudence, mention de la juridiction et de la date de l'arrêt ;
 - pour les textes, mention de leur nature (code, loi, décret etc.) et de leur date
- 3) le licencié s'abstient de tout usage contraire aux lois et règlements ou portant atteinte à l'ordre public et aux bonnes mœurs.

En signant la convention, le licencié souscrit également aux engagements suivants:

- 1) Le droit d'usage reconnu au licencié n'implique aucun transfert du droit de propriété sur les données et les bases de données concernées.
- 2) Le licencié diffuse les données, objet de la licence, sous sa seule responsabilité.
- 3) Il utilise les données dans le cadre de l'activité décrite dans la convention, laquelle précise, le cas échéant, si le licencié entend réaliser une ou des bases de données éditoriales destinées à l'édition de produits en ligne ou hors ligne, en vue d'une diffusion publique ou restreinte, gratuite ou payante.
- 4) Le licencié ne peut consentir de sous-licences, c'est à dire autoriser un tiers à réutiliser les données considérées.

Cette disposition ne fait cependant pas obstacle à la diffusion de produits éditoriaux réalisée à partir de la licence, dès lors que la valeur ajoutée éditoriale, apportée par le licencié aux données objet de la licence, leur a conféré un caractère d'œuvre originale.

En vertu de l'ordonnance de 2005, les administrations sont astreintes à la transparence quant au mode de calcul des redevances, quant aux principales informations susceptibles de réutilisation, qui figurent dans un répertoire ad hoc et quant à d'éventuels détenteurs de droits de propriété intellectuelle.

La directive interdit que le montant de la redevance perçue à l'occasion d'une réutilisation excède la totalité des coûts supportés par l'administration, majorés d'un retour sur investissement raisonnable. Pour une licence avec coût de mise à disposition des données juridiques de Legifrance, un tableau complet des fonds disponibles et des coûts pour le stock et le flux sont accessibles à la rubrique Licences du site Legifrance.

Conclusion:

La disparition à terme de l'édition papier du Journal Officiel et la généralisation de la publication électronique officielle des textes nous impose de mettre en place sans tarder des règles communes d'archivage.

Il convient d'en distinguer deux sortes :

- l'archivage, par les pouvoirs publics, des textes publiés au Journal Officiel électronique, qui sont authentifiés par un certificat et destinés à servir de référence en cas de doute ou de litige ;
- l'archivage des multiples agrégats, toujours plus riches et variés, que les technologies permettent et permettront de faire à partir de ces textes (textes codifiés, consolidés dans leurs versions successives, reliés entre eux etc) :

Si cet archivage incombera pour une grande part aux pouvoirs publics qui les maintiendront en permanence en accès libre, l'archivage de leur propre production par les agrégateurs de contenu en ligne reste incertain.

Comme on l'a vu, dans la pratique, la recherche de l'exhaustivité du droit applicable à une situation donnée et à un moment précis prime pour les usagers sur l'authentification des documents électroniques fournis.

Il serait donc souhaitable de faire porter la priorité sur :

- le développement de supports de stockage et de recherche,
- l'archivage dynamique et la conservation à long terme des versions successives consolidées des documents juridiques, sans que leur certification ne soit une préoccupation majeure ni ne génère des investissements supplémentaires coûteux dans la mesure ou elle sera déjà appliquée aux textes d'origine.

D'une manière générale, on a actuellement tendance à imposer aux données juridiques électroniques des contraintes de certification trop lourdes, d'autant qu'on n'en demande pas aux éditions imprimées.

Enfin, il faut dire encore une fois l'importance d'un dialogue constant et d'une réelle collaboration entre les différents acteurs dans ce secteur, qu'ils soient producteurs, diffuseurs, utilisateurs, conservateurs, juges, ou législateurs. En matière d'archivage électronique, les coûts sont élevés et la question très sensible de la répartition des charges entre les finances publiques, le secteur concurrentiel et le « consommateur » reste à débattre.
