

Bulletin des Archives de France sur la conservation à long terme des documents électroniques N° 3 octobre 2000

L'archivage des données de la nouvelle chaîne pénale de Paris et de la région parisienne (Françoise Banat-Berger, responsable des archives du ministère de la justice)

Une réunion a été tenue le 11 juillet 2000 au bureau des services informatiques (sous-direction de l'informatique, direction de l'administration générale et de l'équipement) avec le responsable des archives et les services informatiques du ministère de la justice, des représentants de la Direction des Archives de France, du centre des archives contemporaines de Fontainebleau et le directeur des services d'archives de Paris.

La réunion portait sur les modalités de versement dans les services d'archives départementales de la base archives de la nouvelle chaîne pénale (N.C.P.).

Présentation sommaire de la NCP

La N.C.P. implantée au TGI de Paris et dans les 7 TGI de la Région parisienne comprend 265 tables dont 5 tables principales : affaires / personnes- alias / infractions –nature des affaires- nature des infractions / événements- décisions / mesures- peines. L'élément d'identification est le numéro de l'affaire (numéro du parquet).

Si la N.C.P. est un véritable outil d'information, certains bouts de la chaîne pénale lui échappent :

- les affaires portées et réglées au Cabinet du juge des enfants (pénal et civil) (application Wineurs) : l'enregistrement de l'affaire figure dans la NCP et n'y revient que si l'affaire passe, après instruction du juge des enfants ou du juge d'instruction, devant le tribunal pour enfants
- le suivi des affaires portées devant le juge d'instruction (application Instru) : l'enregistrement de l'affaire figure dans la NCP mais ne revient qu'en cas de renvoi devant le tribunal correctionnel. De même, sont enregistrées dans la NCP les décisions mettant fin à l'instruction (dessaisissement, renvoi, non-lieu...) ainsi que les mesures de sûreté et les mesures de mise en liberté.
- les affaires portées devant le tribunal de police (application Minos) ou encore les affaires renvoyées après instruction devant le parquet général (cour d'assises).

En revanche, sont également enregistrées dans la NCP des décisions dans le domaine des affaires (au civil) économiques et financières prises en amont par le tribunal de commerce ou le mandataire de justice.

Les données enregistrées sur la NCP dans la base vivante sont destinées à être transférées dans une base archives intermédiaires au terme de délais qui ont été définis avec les juridictions (entre 18 mois et 3 ans). Ce transfert est transparent pour les utilisateurs : il n'y a aucune modification des données entre les deux bases, le transfert permet simplement d' "alléger " la base vivante. C'est après un délai supplémentaire que les données sont transférées dans une base dite " archives " avec, cette fois, une suppression de données (de gestion), le principe étant de ne conserver que les décisions significatives (juridiques) pour chaque affaire. C'est ainsi qu'une affaire après archivage pèse 90% moins qu'une affaire dans les bases vivante ou de préarchivage.

400 000 données sont enregistrées actuellement dans la base vivante de la NCP, 1 200 000 dans la base préarchivée, 6 000 000 dans la base archives. Le transfert sur la base archives se fait une

Bulletin des Archives de France sur la conservation à long terme des documents électroniques N° 3 octobre 2000

semaine sur 3 pour la région parisienne, chaque semaine pour Paris.

La base archives de la NCP

Principes d'archivage

Les modules d'archivage n'ayant pas encore été définis pour les modules jugement / recours / exécution des peines, les décisions prises à l'étape 1 (ministère public ou doyen des juges d'instruction) pour des affaires qui vont vivre aux étapes suivantes (instruction, audiencement, jugement) ne sont pas concernées ici.

Seront donc présentes dans la base archives, les affaires qui se sont terminées à l'étape 1 (bureau d'ordre ou doyen des juges d'instruction) ou à l'étape 2 (instruction), le principe étant de ne conserver pour chaque affaire que les décisions significatives (juridiques) et non les événements de gestion.

* Soit, pour les affaires terminées à l'étape 1 : enregistrement des

- Décisions d'entrée :

Saisine du ministère public (23 codes possibles)

Saisine du doyen des juges d'instruction (2 codes)

- Décisions de sortie (archivage 3 ans après décision finale sauf les classements sans suite : 18 mois)

- Décisions prises par le ministère public

Dessaisissement

Saisine du juge des enfants (en revanche, si le juge pour enfants a saisi le juge d'instruction ou le tribunal pour enfants, l'affaire reviendra dans la NCP)

Saisine du tribunal de police (s'il y eu opposition, l'affaire reviendra dans la NCP)

Classement /médiation/injonction thérapeutique

Jonction

- Décision prise par le doyen

Désistement de partie civile

Incompétence

Irrecevabilité de constitution de partie civile

- Il existe des affaires essentiellement dans le domaine économique et financier (décisions prises en amont dans les affaires par les tribunal de commerce ou le mandataire de justice), pour lesquelles la

Bulletin des Archives de France sur la conservation à long terme des documents électroniques
N° 3 octobre 2000

décision personne est différente de la décision affaire : sera par conséquent conservée la dernière décision- personne juridique (bilan, liquidation en cours, règlement judiciaire, concordat, liquidation judiciaire, redressement judiciaire, liquidation de biens...)

* Pour les affaires terminées à l'étape 2 (instruction)

- Décisions d'entrée :

décisions de saisine du ministère public et saisine du doyen des juges d'instruction (énuméré à l'étape 1)

décisions de saisine du juge d'instruction

- Décisions de sortie :

décisions prises par le juge d'instruction qui mettent réellement fin à l'affaire (incompétence, non-lieu, récusation, dessaisissement...)

décisions prises par le juge d'instruction qui font que l'affaire sort de NCP (renvoi devant le juge des enfants, renvoi devant le tribunal de police, transmission au parquet général)

- Sont également archivées les décisions personnes qui sont différentes des décisions de fin d'affaire (non-lieu partiel, contrainte ou condamnation de témoin, désistement ou irrecevabilité de partie civile) ou qui concernant la mise éventuelle en détention du mis en examen : on archivera les décisions de début (mesures de sûreté) et fin de détention (mise en liberté).

La base archives se compose :

- d'une table des affaires archivées : les données ont été déstructurées et sont conservées en mode texte séquentiel, le seul accès possible se faisant par le numéro d'affaire (chaque affaire a une longueur de 80 caractères). Les libellés sont en clair (à la différence des tables de recherches qui fonctionnent sur la base de nombreux codes et tables de références). Les données ayant été déstructurées, elles ne se retrouvent pas, selon les affaires, à la même place

- de 4 tables de recherches permettant des recherches croisées (outil RAFFA) : les données ont été sélectionnées et ont été intégrées dans des tables indexées : affaires / personnes / véhicule / chèques.

Type des informations stockées dans les 2 tables de recherche principales (affaires et personnes) :

Table affaires

Numéro de l'affaire (n° du parquet)

Numéro du procès-verbal

Date du début des faits

Nature de l'affaire (nomenclature NATINFF)

Code du dernier événement (ayant provoqué l'archivage : décisions significatives de sortie de l'étape 1 ou 2, voir supra)

Bulletin des Archives de France sur la conservation à long terme des documents électroniques
N° 3 octobre 2000

Date de ce dernier événement
Code du service dans lequel l'affaire était traitée avant son archivage
Nom de la première personne (qui apparaît dans le procès-verbal) [champ créé pour faciliter la recherche]
Code de l'élément de structure à l'origine de l'affaire (commissariat, gendarmerie)
Code du premier événement à l'origine de l'affaire (décisions de saisine)
Date de ce premier événement
Mot clé de l'affaire [champ créé pour faciliter la recherche] (type affaire un tel, affaire réservée, affaire mineurs...) : à partir d'une table de références + saisie libre
Indication (selon une table de références) du motif du dernier événement (notamment pour les classements sans suite)
Type de l'infraction (table de codification) : indicateur statistique (infraction, crime, contravention...)
Indicateur statistique si affaire relative à l'enfance délinquante
Nature de l'affaire (table NATAFF)

Table personnes

Indication du rang de la personne dans l'affaire
Indication de l'année de naissance de la personne physique
Indication du rang du déclarant (pour personnes morales)
Date du début des faits
Nom de la personne physique ou morale (y compris les alias)
Nom phonétisé de la personne physique/morale
Indication du prénom (personne physique) ou du sigle (personne morale)
Nom phonétisé du prénom/sigle
Indication du nom marital (personne physique) ou de l'enseigne (personne morale)
Phonème du nom marital/enseigne
Qualité de la personne (code) : monsieur, madame, mademoiselle
Rôle de la personne (table de références) : auteur, victime, prévenu, mis en examen...
Code de l'événement qui a provoqué l'archivage
Code du service dans lequel l'affaire était traitée avant son archivage
Date du dernier événement
Code indiquant si personne physique ou morale
Indication des liens entre noms et rangs, déclarants et personnes morales

Table des affaires archivées

Indication du numéro de l'affaire
Indication du rang de la ligne

Texte (libellé en clair sur 80 caractères) : le texte comprend les informations portées dans les différentes tables, à l'exception d'informations portées dans des champs créés uniquement en vue

Bulletin des Archives de France sur la conservation à long terme des documents électroniques N° 3 octobre 2000

de faciliter la recherche ou encore des indicateurs statistiques.

Modalités du transfert des données contenues dans la base archives de la NCP aux services d'archives départementales : propositions (premières ébauches)

Le transfert des données répond à deux impératifs :

* Pouvoir utiliser un outil informatique d'accès aux données papier

* Pouvoir fournir aux chercheurs un outil de recherche sur le traitement de la délinquance pour le dernier quart du 20ème siècle . Cet outil nominatif n'est pas consultable immédiatement (il est soumis, comme les dossiers de procédure, au délai de 100 ans et n'est consultable, avant ce délai, que sur dérogation accordée par le directeur des Archives de France, sur avis du procureur de la République de la juridiction concernée).

Pour répondre à ce second objectif, le transfert de la table des affaires archivées est suffisante. Cette table présente de nombreux avantages pour les services d'archives : déstructuration des données (mise à plat) qui rend l'information indépendante des matériels et logiciels qui évoluent très vite dans le temps, libellés en clair ce qui évite aux services d'archives d'avoir à archiver également les tables de références et nomenclatures accompagnant la base. Ce transfert pourra se faire sur le support préconisé par le service des archives départementales d'accueil (CD-R, bandes DLT...). Celui-ci devra assurer la pérennité de ces données par des migrations régulières (migrations sur d'autres supports, voire dans d'autres formats), en s'appuyant d'une part sur l'expérience acquise en la matière par le service Constance des Archives nationales et, d'autre part, sur ses propres services informatiques.

Ce transfert devra obligatoirement s'accompagner de toute la documentation technique relative à la N.C.P. qui sera fournie tant par les services informatiques de la sous-direction de l'informatique que par les maîtres d'ouvrage (bureau de l'informatisation des juridictions de la direction des services judiciaires). Cette documentation devra notamment comprendre les spécifications détaillées de l'application, les dictionnaires de données, les modèles conceptuels des données, les données disponibles sur les modules d'archivage...

Pour répondre au premier objectif, le service d'archives départementales d'accueil devra disposer, outre cet archivage de la table déstructurée, d'un certain nombre de données en ligne lui permettant d'effectuer un certain nombre de recherches. Ces données seraient extraites des 2 tables de recherche existantes (affaires/personnes) à savoir, dans l'ordre suivant, les données correspondants aux champs :

Table affaire

Numéro de l'affaire (n° du parquet)

Nature de l'affaire (table NATAFF)

Date du début des faits

Code de l'élément de structure à l'origine de l'affaire (commissariat, gendarmerie)

Code du premier événement à l'origine de l'affaire (décisions de saisine)

Date de ce premier événement

Nature de l'infraction (nomenclature NATINFF)

Code du dernier événement (ayant provoqué l'archivage : décisions significatives de sortie de

Bulletin des Archives de France sur la conservation à long terme des documents électroniques **N° 3 octobre 2000**

l'étape 1 ou 2, voir supra)

Date de ce dernier événement

Code du service dans lequel l'affaire était traitée avant son archivage

Indication (selon une table de références) du motif du dernier événement (notamment pour les classements sans suite)

Table personne

Numéro de l'affaire (n° du parquet)

Indication de l'année de naissance de la personne physique

Nom de la personne physique ou morale (y compris les alias)

Indication du complément du nom (par exemple, pour les noms de sociétés, dont le nombre de caractères dépasse 30)

Nom phonétisé de la personne physique/morale

Indication du prénom (personne physique) ou du sigle (personne morale)

Nom phonétisé du prénom/sigle

Indication du nom marital (personne physique) ou de l'enseigne (personne morale)

Phonème du nom marital/enseigne

Qualité de la personne (code) : monsieur, madame, mademoiselle

Rôle de la personne (table de références) : auteur, victime, prévenu, mis en examen...

Indication des liens entre noms et rangs, déclarants et personnes morales

Les données correspondantes à ces champs seraient mises à plat de manière à ce qu'elles puissent être remontées sous forme de tables indexées dans l'environnement informatique du service des Archives départementales. Dans ce cas, l'accès serait strictement réservé au service d'archives départementales. Serait ainsi constitué dans ce service une nouvelle forme d'instrument de recherche informatisé (suivant la méthodologie mise en place aux archives nationales de Fontainebleau pour la gestion informatisée des dossiers individuels de personnels ou encore pour l'enregistrement des dossiers de contentieux du Conseil d'Etat).

Parallèlement devront être transférées au service d'archives départementales d'accueil, les nomenclatures, codes et tables de références sur support papier si possible ainsi que sur support numérique (dans ce cas, les données seront également mises à plat) permettant de comprendre les données stockées.

La Direction des Archives de France sollicitera un des départements de Paris et de la couronne parisienne afin de tenter cette opération qui sera une première dans un service d'archives territoriales et est, pour cette raison d'une importance toute particulière.

Création d'un groupe Archivage au sein de l'association Aristote

Claude Huc, directeur au Centre National d'Etudes Spatiales

Bulletin des Archives de France sur la conservation à long terme des documents électroniques **N° 3 octobre 2000**

(CNES) d'un département dédié à la gestion et à la valorisation des données issues des observations spatiales, vient de créer un groupe de réflexion sur la pérennisation des informations numériques.

Ce groupe est mis en place dans le cadre de l'association Aristote (<http://www.aristote.asso.fr/>), association visant au partage de savoirs et de savoir faire entre les organismes membres. Parmi eux, se trouve un certain nombre de grands organismes à caractère technique, scientifique ou industriels (CEA, CNES, CNRS, EDF,).

Le groupe "archivage" rassemble quant à lui, à la fois des représentants des organismes précités, certains d'entre eux ayant déjà une expérience durable (de plus de 25 ans) dans le domaine de la pérennisation des données sous forme numérique, et des représentants d'organismes ou d'institutions publiques dont l'expérience dans le domaine est plus récente mais qui, par contre, sont riches d'une longue pratique de l'archivage (Archives de France, Bibliothèque nationale de France...), ainsi que des sociétés privées spécialisées dans ces questions et des consultants.

Cette interpénétration des compétences et savoir faire sur des sujets d'intérêt communs peut être très fructueuse.

L'idée de constituer ce groupe est née de la rédaction d'une recommandation élaborée dans le domaine spatial par un comité de normalisation qui rassemble les agences spatiales internationales. Cette recommandation intitulée 'Reference Model for an Open Archival Information System' (OAIS) n'a rien de spécifique au domaine spatial et a très vite suscité

l'intérêt de nombres institutions en charge de questions d'archivage dont par exemple celui des responsables du projet européen NEDLIB auquel participe la Bibliothèque nationale de France. Ce modèle de Référence subit actuellement le processus de normalisation ISO. L'archivage est ici défini dans son sens le plus large, en incluant l'ensemble des fonctions d'entrée des documents dans les systèmes et services d'archivage, de gestion, de stockage, d'administration et d'accès à ces documents par les utilisateurs des archives.

La version anglaise du document est accessible sur http://www.ccsds.org/ccsds/red_books.html (la référence du document est CCSDS 650.0-R-1 en version pdf).

Un forum de discussion par mail a été ouvert par Claude Huc.

La première réunion du groupe a eu lieu le 28 août et la prochaine aura lieu le 29 novembre.

Nous vous tiendrons informés dans ce Bulletin des travaux du groupe.

Pour toute information complémentaire : Claude.Huc@cnes.fr

Nouvelles du DLM-Forum

Le DLM-Monitoring Committee (émanation du DLM-Forum, groupe de travail de la Commission européenne sur les Données Lisibles par Machine) a rédigé le document qu'il avait été décidé

Bulletin des Archives de France sur la conservation à long terme des documents électroniques **N° 3 octobre 2000**

d'adresser au monde de l'industrie, lors de la dernière réunion du DLM-Forum à Bruxelles en octobre 1999. La version finale de ce texte a été envoyée au Secrétaire général de la Commission européenne, Mr David O'SULLIVAN, et au Commissaire pour la Société de l'information, Mr Erkki LIIKANEN.

Il est disponible en ligne sur le site du DLM-Forum à l'adresse <http://www.dlmforum.eu.org>.

Il existe pour l'instant en version anglaise et allemande. La version française sera disponible sous peu.

Site européen sur les normes et standards pour l'archivage des documents électroniques

Signalé par Nadine Gilbert (Archives départementales des Pyrénées-Orientales), un site européen très fourni et très soigné sur les normes dans tous les domaines qui nous intéressent (gestion de l'information, formats d'échange, normes à l'usage des archives, bibliothèques et musées, métadonnées, archivage électronique, commerce électronique, etc...) :

Diffuse Project (Finnish IT Development Centre, IC Focus et le SGML Centre, sous l'égide de la Commission européenne, dans le cadre du programme "Information Society Technologies").

à l'adresse :

<http://www.diffuse.org/index.html>

Naissance de l'Institut des sciences du numérique

On nous annonce la naissance en Rhône-Alpes de l'Institut des sciences du document numérique (Isdn) qui regroupe 17 laboratoires de la région travaillant sur cet objet avec des approches disciplinaires différentes (Sciences de l'ingénieur et Sciences humaines). Son objectif est de faciliter les échanges entre les équipes de façon à enrichir les recherches en cours et développer de nouvelles actions.

Vous trouverez sur son site les informations utiles concernant ses orientations de recherche, ses appuis, les

laboratoires participants, son organisation, etc.

Url : <http://isdn.enssib.fr>

Ressources sur le site de la MTIC

La MTIC (Mission interministérielle de soutien technique pour le développement des technologies

Bulletin des Archives de France sur la conservation à long terme des documents électroniques **N° 3 octobre 2000**

de l'information et de la communication dans l'administration) a ouvert un forum sur le thème suivant :

"Réflexion publique sur la conservation des documents et des informations"

<http://www.mtic.pm.gouv.fr/php/forum/index.php3?code=fconsdoc99741>

Pour ceux qui ne seraient pas convaincus que l'essence même du métier d'archiviste et même les missions de nos institutions de conservation publiques sont en train d'évoluer, avec l'intégration de la fonction archives dans des systèmes d'information qui peuvent assurer aussi la conservation des données et avec la définition d'un métier de "tiers archiviste" (qui reste à préciser), ce forum comporte d'intéressants messages de Gérard Weiz et Michel Lesourd sur le thème de l'archivage et des supports de conservation électronique.

On y trouve aussi diffusé le texte suivant :

EDIFOCAS & IALTA, Guide de l'archivage électronique sécurisé, 65 p. Sept. 2000, en format pdf :

http://www.mtic.pm.gouv.fr/php/forum/delivre_fichier.php3?code=fconsdoc99741&id=35

Ce texte a été mis au point par un groupe de travail présidé par Michel Lesourd, du Conseil supérieur de l'ordre des experts-comptables. Il s'agit d'un document de référence, donnant les divers aspects méthodologiques et techniques qui

devront être mis en oeuvre lors de la conception des systèmes d'archivage électronique, dans la ligne tracée par la loi sur la signature électronique et dans le souci d'assurer la conservation à long terme des données.

Ce groupe de travail est patronné par les institutions suivantes :

CS-OEC (experts-comptables), CNCC (Commissaires aux comptes), EDIFICAS (CS-OEC et SSIIs), IALTA, CNHJ (huissiers), CSN (notaires), CNG (Greffiers de tribunaux de commerce), APROGED

AFAI (audit informatique), CIGREF (club informatique des grandes entreprises françaises).

Travaux de la National Library of Medicine (USA) sur les métadonnées

La National Library a mis en ligne une conférence de Diane Boehr sur l'utilisation de métadonnées, "Developing the Use of Metadata at the National Library of Medicine: From Planning to Implementation", le 7 juillet 2000. Elle contient la définition des métadonnées minimales pour assurer la diffusion et la sauvegarde sur le web des ressources très diverses conservées à la Bibliothèque et aider à une prise en compte plus grande par les moteurs de recherche des données du Dublin Core.

Vous trouverez cette conférence à l'adresse :

Bulletin des Archives de France sur la conservation à long terme des documents électroniques
N° 3 octobre 2000

<http://www.nlm.nih.gov/tsd/cataloging/metadata/sld001.htm>

Merci d'envoyer tout texte ou information pour le Bulletin n° 4 à catherine.dherent@culture.gouv.fr