

## Applications concrètes de l'EAD : le chantier de dématérialisation des instruments de recherche et le chantier « modèles de production » des Archives nationales

Isabelle Aristide-Hastir

---

### Citer ce document / Cite this document :

Aristide-Hastir Isabelle. Applications concrètes de l'EAD : le chantier de dématérialisation des instruments de recherche et le chantier « modèles de production » des Archives nationales. In: La Gazette des archives, n°220, 2010. Les instruments de recherche : évolutions, publics et stratégies. pp. 195-218;

[http://www.persee.fr/doc/gazar\\_0016-5522\\_2010\\_num\\_220\\_4\\_4758](http://www.persee.fr/doc/gazar_0016-5522_2010_num_220_4_4758)

---

Document généré le 15/03/2017

# Applications concrètes de l'EAD : le chantier de dématérialisation des instruments de recherche et le chantier « modèles de production » des Archives nationales

---

Isabelle ARISTIDE-HASTIR

## Introduction

Dans la perspective de l'ouverture au public de leur nouveau bâtiment à Pierrefitte-sur-Seine, les Archives nationales ont engagé depuis 2006 plusieurs grands chantiers connexes au chantier de construction proprement dit : chantier des fonds (reconditionnement, numérisation, déménagement), chantier du système d'information archivistique (SIA) et chantier de dématérialisation des instruments de recherche.

Au cœur de ces chantiers connexes ont été développées des réflexions scientifiques autour de la normalisation des instruments de recherche.

Les instruments de recherche anciens (antérieurs à 2010) sont traités par le chantier de dématérialisation, piloté par Isabelle Aristide-Hastir, conservateur en chef.

Pour les nouveaux instruments de recherche, le futur système d'information archivistique (SIA) piloté par Béatrice Hérold, conservateur en chef<sup>1</sup>, met en place des stratégies de normalisation et d'approche simplifiée d'un EAD accessible par le plus grand nombre du personnel scientifique.

---

1 Aujourd'hui conservateur général.

## Le chantier de dématérialisation des instruments de recherche

Plusieurs expériences ont été précédemment menées au sein des Archives nationales pour intégrer l'XML-EAD lors de la rétroconversion en XML-EAD des instruments de recherche existants. On citera notamment le projet NOE (normalisation océrisation encodage) engagé en 2003 qui a fait encoder dix-sept instruments de recherche imprimés – soit 11 000 pages – prenant en compte la diversité des fonds conservés aux Archives nationales, site de Paris, du XIV<sup>e</sup> à la fin du XX<sup>e</sup> siècle.

### *Définir des objectifs*

Le but du chantier de dématérialisation des instruments de recherche créé en 2006 n'était pas *a priori* de tout faire encoder en XML-EAD mais de mettre à la disposition des lecteurs dans la future salle des inventaires virtuelle un maximum d'instruments de recherche soit sous forme de documents numériques, soit sous forme de documents PDF texte<sup>1</sup>, soit sous une forme plus élaborée et mieux adaptée à la diffusion sur Internet, à savoir un document encodé en XML ou mieux en XML-EAD.

Deux contraintes fortes avaient été posées dès le départ :

- la dématérialisation concernera les instruments de recherche de tous les fonds, ceux qui déménagent comme ceux ne déménageant pas, et concernera surtout aussi bien les instruments de recherche des fonds actuellement conservés à Paris que les instruments de recherche des fonds actuellement conservés à Fontainebleau<sup>2</sup> ;

---

<sup>1</sup> C'est la solution actuellement adoptée sur le site Internet des Archives nationales.

<sup>2</sup> Rappelons que les Archives nationales conservent actuellement sur le site de Paris les archives de l'origine à 1958, les archives privées, les cartes et plans, la bibliothèque historique et le minutier central des notaires de Paris ; le site de Fontainebleau accueille les versements contemporains depuis 1976.

Dans le cadre du projet Pierrefitte, les fonds des Archives nationales seront ainsi répartis : à Paris resteront les archives antérieures à 1790, la bibliothèque historique et le minutier central des notaires de Paris ; Pierrefitte et Fontainebleau se partageront les fonds postérieurs à 1790 : le domaine « Équipement, Territoire, Agriculture, Écologie, Industrie, Commerce, Tourisme, Économie, Finances » sera conservé et communiqué sur le site de Fontainebleau sous la dénomination « Hommes et territoires » ; les domaines « Pouvoir législatif, Pouvoir exécutif (Présidence, Premier ministre et secrétariat général du gouvernement), Assemblées, Grands Corps de l'État, Autorités administratives indépendantes, Temps de guerre et anciens combattants »,

▪ la dématérialisation est étroitement liée à un projet ambitieux du SIA : permettre au lecteur de commander ses documents *via* l'instrument de recherche en cliquant sur la cote du document.

Or, dans la majorité des cas, la cote est incomplète (elle ne comporte que le numéro d'article) et même si elle est complète, elle est rédigée sous sa forme traditionnelle (avec des espaces et des exposants) et non dans une grammaire des cotes stricte, compréhensible par le système informatique.

Exemple : extrait de l'instrument de recherche des Archives  
des Forces françaises libres conservées sous la cote 3AG3

*Instruments de recherche d'origine :*

1. Fragments (dactylographiés) d'un historique des Forces françaises libres (période du 18 juin 1940 au 31 décembre 1941)  
Inventaire dactylographié des archives concernant l'Afrique française libre (période du 28 août 1940 au 1<sup>er</sup> août 1943) détenues par le général commandant supérieur des troupes de l'AEF Cameroun  
Répertoire dactylographié des noms des militaires et civils pour lesquels une correspondance existe aux Archives du cabinet militaire du Haut-Commissaire (AFL) ; s.d.

*Instrument de recherche avec la cote reconstituée et libellée selon la grammaire des cotes informatique :*

**AG/3(3)/1**

Fragments (dactylographiés) d'un historique des Forces françaises libres (période du 18 juin 1940 au 31 décembre 1941)  
Inventaire dactylographié des archives concernant l'Afrique française libre (période du 28 août 1940 au 1<sup>er</sup> août 1943) détenues par le général commandant supérieur des troupes de l'AEF Cameroun  
Répertoire dactylographié des noms des militaires et civils pour lesquels une correspondance existe aux Archives du cabinet militaire du Haut-Commissaire (AFL) ; s.d.

---

« Intérieur, Outre-mer, Justice », « Instruction-Éducation, Recherche, Culture, Jeunesse et Sports », « Santé, Action sociale, Affaires sociales, Travail, Population », ainsi que les archives privées seront conservées et communiquées à Pierrefitte.

L'objectif de la cote « cliquable » s'avérait donc primordial dès le démarrage du chantier ; il entraînait la reprise de la quasi-totalité des instruments de recherche des Archives nationales avec comme corollaire de devoir trouver une méthode simple et efficace.

### *Évaluer le travail*

La dématérialisation des instruments de recherche a commencé par une évaluation très précise des ressources tant sur le plan quantitatif que qualitatif.

Certes, les Archives nationales ont publié à partir de 1985, à la suite de l'état général des fonds, un état général des inventaires ventilé par section (section ancienne, section du XIX<sup>e</sup> siècle, section du XX<sup>e</sup> siècle, section des cartes et plans, section du minutier central) ainsi qu'un état sommaire des archives privées. Mais cet état des inventaires imprimés ne prenait en compte ni les instruments de recherche des archives conservées à Fontainebleau ni les nouveaux instruments de recherche, leurs modifications, refonte, etc.

Un recensement effectué par une équipe de dix personnes pendant neuf mois a permis de couvrir l'ensemble des instruments de recherche sur les deux sites, sous toutes leurs formes (papier, fiches cartonnées, microfiches, microfilms, électronique) et de donner pour chacun d'eux des données très précises d'ordre matériel (format, nombre de pages, qualité de l'écriture, qualité du support) et d'ordre intellectuel (type d'instrument de recherche, cotes vacantes, instruments de recherche associés, etc.).

Voici les résultats du recensement des instruments de recherche des Archives nationales effectué en 2007 :

- 23 500 instruments de recherche (5 400 à Paris et 18 100 à Fontainebleau) ;
- 1,2 million de pages papier ;
- 607 microfilms ;
- 7 600 microfiches ;
- 238 fichiers ;
- 3,5 millions de fiches papier.

Les résultats du recensement ont été chargés dans une base de données servant à la fois d'état des inventaires à jour et d'outil de suivi du travail de dématérialisation.

*Mettre au point une chaîne de traitement*

En 2008 a été développé un marché « test » portant sur environ 1 000 instruments de recherche (100 000 pages papier et un « petit » fichier de 56 000 fiches) qui a permis de forger une méthode et de définir précisément le résultat de la dématérialisation.

Les résultats ont été très encourageants : tous les instruments de recherche, numérisés dans un premier temps, pouvaient, quels que soient leur support et leur écriture, être restitués en texte soit par saisie manuelle soit par océrisation ; le texte pouvait être encodé nativement en XML dès la saisie. Il était dès lors possible d'envisager de ne mettre en ligne que des documents en mode texte, à l'exclusion des images numériques ; l'enjeu serait dans un deuxième temps de les normaliser en XML-EAD.

Même si les opérations de numérisation, océrisation, saisie, encodage sont externalisées, les Archives nationales restent les concepteurs de la dématérialisation de leurs instruments de recherche qu'elles contrôlent à plusieurs étapes. Chaque instrument de recherche est considéré comme un projet à part entière, quels que soient sa taille et son contenu ; l'équipe projet rédige donc pour chaque instrument de recherche une fiche individuelle de consignes destinée au prestataire<sup>1</sup>. Cette méthode permet de s'adapter à toutes les variantes de rédaction et de présentation.

L'instrument de recherche est contrôlé à plusieurs niveaux : au niveau de l'image numérique, de la qualité de la restitution du texte, de la qualité de la restitution de la cote, du respect des consignes de restitution et notamment de l'encodage.

La chaîne de traitement de la dématérialisation est donc une chaîne pilotée à toutes les étapes par les Archives nationales :

- préparation matérielle : Archives nationales ;
- préparation intellectuelle (rédaction des fiches individuelles de consignes de restitution) : Archives nationales ;
- numérisation : prestataire/contrôle Archives nationales ;
- saisie du texte : prestataire/contrôle Archives nationales ;
- encodage : prestataire/contrôle Archives nationales ;
- saisie de la cote : prestataire/contrôle Archives nationales.

---

<sup>1</sup> Ill. 1 et 2, p. XIV à XVI.

*Adapter l'XML-EAD à la rétroconversion de masse*

Les expériences antérieures de rétroconversion ont montré la nécessaire adaptation de l'instrument de recherche à l'EAD. Dans le cadre de la dématérialisation de masse, il n'était pas envisageable de reprendre, même sommairement, chacun des instruments de recherche :

- les Archives nationales produisent depuis plus de deux cents ans des instruments de recherche (quelques-uns même datent du XVIII<sup>e</sup> siècle et sont toujours utilisés) ; ceux-ci sont donc nécessairement très divers dans leurs supports et leur présentation, et la grande majorité n'est pas conforme à l'ISAD(G) ;

- reprendre un instrument de recherche requiert un long travail d'investissement intellectuel, parfois même de reclassement et de reprise des cotes qui aurait demandé un temps considérable incompatible avec l'objectif de l'ouverture de la salle des inventaires virtuelle en 2012 ;

- la dématérialisation de masse fait intervenir une équipe projet entièrement dédiée à cet objectif mais qui ne se substitue pas au personnel scientifique, détenant la meilleure connaissance des fonds ; l'équipe projet s'est fixé comme contrainte déontologique de ne pas intervenir dans la conception, la rédaction ou la correction des instruments de recherche existants et de ne procéder qu'à des opérations de « toilettage » portant sur la ponctuation, la mise en forme des titres ou les dates extrêmes si elles font défaut.

Malgré ces contraintes, le marché test réalisé en 2007 avec un partenaire extérieur a mis en évidence qu'au-delà de la diversité apparente des instruments de recherche, voire, selon certains, de leur incohérence, le cœur de l'instrument de recherche repose sur ce que l'équipe projet a appelé « notice », à savoir :

- la cote ;
- la description ;
- les dates extrêmes.

Ces trois éléments existent dans tous les instruments de recherche, avec des variantes : numéros de dossiers ou d'analyses, dates extrêmes portant sur les dossiers et non sur la totalité de l'article, description très sommaire ou très détaillée.

Ces trois éléments (cote, description, dates) correspondent parfaitement aux trois « u » de l'encodage en XML-EAD :

- <unitid> ;
- <unittitle> ;
- <unitdate>.

L'équipe projet a donc décidé d'utiliser la DTD-EAD sous une forme très simplifiée<sup>1</sup>, à savoir une trentaine d'éléments seulement, dont des éléments peu utilisés en encodage direct (<note>, <index>, <odd>) afin de « coller » le plus possible à l'instrument de recherche d'origine sans le reprendre et sans le structurer :

- toute l'introduction, parfois très développée dans les instruments de recherche des Archives nationales, est placée dans le *scopecontent* de niveau supérieur sans utiliser les éléments <bioghist> [historique du producteur], [historique du fonds] ;
- si le cas se présente nettement, on peut utiliser <bibliography> [bibliographie] ou <otherfindaid> [sources complémentaires] ;
- on distingue deux <unitid> : <unitid type="cote-de-consultation"> est utilisé pour la cote « cliquable », <unitid> simple est utilisé pour tous les compléments de la cote (numéro de dossier, cote d'origine des services versants, numéro d'analyse, etc.) ;
- l'élément <unittitle> est répétable autant de fois que l'instrument de recherche d'origine comporte de paragraphes ; la distinction <unittitle> suivi de <scopecontent> dans le même <did> est possible si elle est évidente ;
- l'élément <note> est utilisé pour reprendre les notes des instruments de recherche mais utilisé seulement dans le <scopecontent> ; les notes sont numérotées à 1 à chaque <c><did> ;
- l'élément <unitdate> avec normalisation de la date est utilisé seulement si les dates extrêmes sont celles de l'article ; il est dans tous les cas et *a minima* utilisé dans le titre et dans les niveaux supérieurs de la structure ;
- la structuration est hiérarchisée à trois niveaux maximum au moment de la saisie<sup>2</sup> ;
- l'élément <odd> qui est autant à proscrire que « divers » dans une description est néanmoins utilisé pour toutes les parties se trouvant matériellement à la fin des instruments de recherche : table des matières, *errata*, annexes ;

---

<sup>1</sup> Annexe 1 et 2, p. 213 à 215.

<sup>2</sup> La structure peut ensuite être développée lors de la reprise de l'instrument de recherche. Les fonds contemporains versés à Fontainebleau depuis 1976 par les missions des Archives nationales auprès des ministères disposent tous d'un sommaire (accessible dans PRIAM 3) ; les autres fonds, notamment ceux de Paris, disposent généralement d'un instrument de recherche structuré (avec titres, sous-titres, etc.) ; à défaut, l'état général des fonds peut être utilisé pour monter une structure de base. On peut donc considérer que la quasi-totalité des instruments de recherche des Archives nationales est structurée.



▪ l'élément <index> est utilisé pour encoder les index se trouvant à la fin des instruments de recherche ; il n'est pas prévu en dématérialisation d'indexer au sein de l'instrument de recherche avec <controlaccess> ou avec les éléments <persname> ou autres ; il n'est pas prévu non plus de faire un lien entre les mots de l'index et les occurrences dans le texte de l'instrument de recherche.

### *Fixer des priorités*

Le programme de dématérialisation est déterminé par tranches annuelles. En ce qui concerne Fontainebleau (versements depuis 1976), le programme est chronologique : 1976-1985, 1986-1993, 1994-2000, 2001-2010 (environ 100 000 pages par an). Pour Paris, la sélection des instruments de recherche à traiter chaque année se fait en collaboration avec les sections qui répartissent leurs instruments de recherche selon des priorités 1, 2, 3. Les consignes étaient que chaque fonds ait au moins un instrument de recherche permettant de décrire les cotes, puis les instruments de recherche plus détaillés étaient désignés par les sections<sup>1</sup>.

### *Mettre en place un post-traitement*

L'équipe de dématérialisation a anticipé très en amont les opérations de post-traitement indispensables à la publication dans la salle des inventaires virtuelle :

- vérification de la table des cotes présentes dans l'instrument de recherche avec la table des cotes du récolement ;
- reprise et normalisation des titres ;
- mise en place de la structure si elle n'a pu être posée lors de la saisie ;
- mise en place de dates extrêmes au moins dans les niveaux supérieurs ;
- « toiletage ».

Ces opérations se font en commun avec le chantier chargé du système d'information et avec le personnel scientifique des deux sites de Paris et de Fontainebleau.

---

<sup>1</sup> Il est fréquent sur le site de Paris, notamment pour la section ancienne et la section du XIX<sup>e</sup> siècle, d'avoir des instruments de recherche « à tiroir » : un instrument de recherche général, puis, selon les groupes de cotes un ou plusieurs instruments de recherche détaillés. Pour certains fonds, il est utile de pouvoir accéder directement à l'instrument de recherche détaillé.

## **Le chantier « modèles de production » des instruments de recherche**

Parallèlement à la construction du nouveau système informatique (SIA) qui se déploiera sur les trois sites de Paris, Pierrefitte et Fontainebleau, ont été lancés des chantiers scientifiques ayant pour objectif d'affiner et de préciser les pratiques, pour que le SIA soit au plus près des besoins des utilisateurs. Parmi ces chantiers, le chantier « modèles de description » s'attache à la production des instruments de recherche avec comme enjeu l'harmonisation des pratiques d'élaboration des instruments de recherche, tout en offrant aux utilisateurs des outils facilitant une saisie normalisée. Il s'agit donc de définir les modèles de description, non des modèles de rédaction ou de mise en forme et affichage.

### *La production des instruments de recherche dans le cadre du nouveau système d'information archivistique (SLA)*

Le SIA est un système centralisé composé de trois groupes de fonctionnalités : gestion logistique et archivistique (récolement, gestion des espaces, des communications, des lecteurs, etc.) ; production d'instruments de recherche ; interrogation des instruments de recherche dans le cadre de la salle des inventaires virtuelle (SIV).

La production des instruments de recherche dans le SIA est décomposée en quatre phases :

- initialisation de l'instrument de recherche par le responsable des fonds, qui charge un ou plusieurs de ses collaborateurs de la rédaction de l'instrument de recherche ; ce travail se fait dans un espace collaboratif, où les rédacteurs ont à leur disposition tous les outils nécessaires à la rédaction (modèles de description, plans de classement éventuels, référentiels d'indexation, etc.) ;
- rédaction de l'instrument de recherche par les agents désignés, qui ont seuls la responsabilité de leur façon de rédiger ; ils ont la possibilité de visualiser leur travail, de l'envoyer au responsable de fonds, avec lequel des échanges ont lieu jusqu'à ce qu'une forme finale satisfaisante soit mise au point, et peuvent également proposer des descripteurs s'ils n'en trouvent pas qui conviennent dans les référentiels ;
- validation de l'instrument de recherche : technique (conformité à l'EAD) et archivistique ;
- demande de publication dans la SIV : soumission de l'instrument de recherche au comité de publication.

Dès son initialisation, l'instrument de recherche est inscrit dans le référentiel des instruments de recherche, contenant l'ensemble des instruments de recherche existant dans la base interne au SIA, qu'ils soient validés, publiés, en cours de rédaction, etc. On pourra donner accès à ces instruments de recherche soit totalement à tous les agents, soit partiellement, soit pas du tout. Ce référentiel permettra notamment aux présidents de salle d'avoir accès au moins au titre des instruments de recherche en cours de rédaction, pour pouvoir contacter le responsable de fonds si nécessaire. En revanche, le public n'y aura pas accès.

La publication des instruments de recherche dans la SIV peut être partielle si certaines informations doivent rester confidentielles (exemple : données nominatives) – l'instrument de recherche intégral restant accessible en interne – soit *via* une SIV locale, à partir de laquelle aucun téléchargement ne sera possible (alors que ce sera le cas de tous les instruments de recherche publiés dans la SIV Internet), soit à partir du référentiel interne des instruments de recherche : les deux possibilités sont à l'étude.

La rédaction des instruments de recherche dans le SIA est soumise aux deux principes de conformité à ISAD(G) et à l'EAD ; l'usage de l'EAD avec ses 146 éléments disponibles étant trop lourd, il a été convenu d'élaborer un schéma de référence excluant les éléments et attributs n'ayant pas vocation à être utilisés aux Archives nationales<sup>1</sup>. À partir de ce schéma simplifié seront déclinés plusieurs modèles de description.

Le SIA a défini des modalités pratiques de production afin que les utilisateurs fassent de l'EAD sans s'en rendre compte : soit en utilisant un logiciel de traitement de texte avec attribution de styles (correspondant aux balises EAD) et visualisation du plan de classement sur une partie de l'écran au fur et à mesure de la rédaction, soit en utilisant un formulaire avec des masques de saisie adaptés au modèle de description sélectionné.

#### *Le groupe de travail du chantier « modèle de description »*

Les objectifs fixés au groupe de travail « modèle de description » étaient les suivants : déterminer le nombre de modèles de description nécessaires pour la création des instruments de recherche dans le SIA, identifier ceux à implémenter pour la mise en production et fixer la liste de ceux à élaborer après la mise en production ; construire ces modèles, en précisant les éléments

---

<sup>1</sup> Annexe 3, p. 216.

EAD utilisés, leur ordre, leur caractère obligatoire ou facultatif et leur répétabilité ; décliner ces modèles en masques de saisie (l'utilisateur du SIA pourra saisir ses instruments de recherche en mode formulaire, quel que soit le type d'instrument de recherche) ; établir des règles de rédaction à associer à chaque modèle sous la forme d'un manuel d'utilisateur.

### *Méthodologie du groupe de travail*

- Enquête par service et par type d'instrument de recherche

Afin d'établir clairement les besoins en terme de modèles d'instrument de recherche, le groupe a procédé, dans un premier temps, à un état des lieux. En effet, les Archives nationales et les missions des Archives de France produisent depuis des décennies des instruments de recherche, et un certain nombre de modèles, grilles de description, consignes de rédaction, feuilles de style, etc. existaient déjà.

Force a été de constater une grande homogénéité des pratiques sur le fond, avec une plus grande disparité sur la forme. Tous les modèles ou exemples présentés, quels que soient la typologie de l'instrument de recherche, les documents décrits ou l'outil utilisé pour guider le travail du rédacteur, étaient conformes à ISAD(G). Les seules différences trouvées relèvent de la présentation des informations ou du niveau de détail de certains aspects de la description. De la même manière, la longueur ou le caractère dense ou non d'une description n'est pas apparu comme un critère toujours pertinent pour distinguer des modèles d'instrument de recherche.

- Prendre en compte les instruments de recherche des missions

Les Archives nationales reçoivent les versements d'archives contemporaines provenant des missions implantées dans les ministères. Dans le cadre de la production de nouveaux instruments de recherche, les répertoires de versements envoyés par les missions représentent la première source d'accroissement des nouveaux instruments de recherche dans la salle des inventaires virtuelle (environ 900 par an).

Or, si, au sein des Archives nationales, les rédactions d'instruments de recherche sont confiées à du personnel scientifique, cela n'est pas toujours le cas dans les missions, composées pour bonne partie de personnel administratif peu familiarisé au vocabulaire archivistique, en dépit d'une pratique quotidienne conforme aux principes archivistiques. Les modèles devront donc être simples, non seulement dans leur conception mais aussi dans leur utilisation.

▪ Tirer parti de l'expérience Sosie<sup>1</sup>

SOSIE (saisie en Open office pour la structuration d'instruments de recherche en EAD) est un groupe de travail lancé en octobre 2007 dans le but de réfléchir à un outil intermédiaire de production normalisée d'instruments de recherche simples sous Open Office, avant la mise en place du nouveau système d'information, afin de familiariser le personnel avec la normalisation des futurs instruments de recherche.

Le groupe de travail a étudié longuement les pratiques de rédaction d'un instrument de recherche et l'utilisation des éléments en EAD et a établi les correspondances entre ISAD(G), l'EAD et les styles Open office. Il a sélectionné 24 éléments de la DTD-EAD parmi les éléments structurants de l'EAD<sup>2</sup> (les éléments de mise en forme, par exemple, ont été écartés), a structuré un modèle et rédigé un manuel d'utilisateur (75 pages) assorti d'un mémo plus concis pour l'utilisation courante.

SOSIE permet le passage simplifié du traitement de texte en XML-EAD, l'intégration des éléments forts de normalisation dans la rédaction (par exemple, rédaction des motifs d'incommunicabilité) et l'utilisation de l'EAD.

Une fois le modèle déterminé, le groupe technique a rédigé le cahier des charges du programme de conversion permettant de passer d'un document Open office stylé à un document EAD. L'accent a été mis sur les processus de

---

<sup>1</sup> La solution retenue a été d'utiliser Open office pour les raisons suivantes : c'est l'outil choisi par le ministère de la Culture ; l'utilisation de traitement de texte est bien intégrée chez les utilisateurs ; Open office contient un filtre XML permettant de faire des conversions directes en XML, contrairement à Word 97.

<sup>2</sup> Quelques caractéristiques de SOSIE pour le passage du modèle à l'EAD :

- *Distinction cote/identifiant* : la cote est l'élément que le lecteur pourra commander grâce à la cote cliquable ; l'identifiant est utilisé pour tous les autres niveaux (supérieur : groupe d'articles, ou inférieur : numéro de dossier). Toutes les descriptions doivent commencer par le style cote ou identifiant.

- *Définition de champs obligatoires et de listes de valeurs autorisées* pour certains champs (exemple : langue, conditions de communication, etc.).

Concernant les champs obligatoires, une réflexion est en cours pour les signaler au rédacteur (en soulignant le libellé ou par un jeu de couleur) afin de faciliter la saisie. Une orientation du rédacteur vers les annexes spécifiques à certains champs (par une insertion dans le libellé) est également proposée.

- *La structuration* : le modèle permet au total sept niveaux pour le corps de l'instrument de recherche.

- *L'index* est construit par saisie manuelle à la fin de l'instrument de recherche, avec les cotes (ou groupe de cotes) dans leur intégralité.

contrôle et validation des instruments de recherche produits par SOSIE, avec des listes d'erreurs pour aider l'utilisateur<sup>1</sup>. Les Archives nationales ont tous les droits sur ce programme (utilisation, diffusion, évolution, etc.).

Les travaux du groupe « modèles de description » ont pu s'appuyer sur les conclusions du groupe SOSIE, qu'il s'agisse du modèle SOSIE lui-même ou de toutes les questions soulevées lors de son élaboration, qu'elles soient archivistiques, pratiques ou organisationnelles<sup>2</sup>. Les enseignements tirés de SOSIE sont précieux et c'est dans le même esprit qu'ont été conçus les modèles proposés par le groupe « modèles ».

Le travail avec SOSIE a commencé à préparer les esprits et la construction d'un nouvel outil pour la production des instruments de recherche est une occasion idéale de poursuivre sur cette voie. SOSIE a montré que nous faisons tous, de la section ancienne aux missions en passant par les cartes et plans, le même travail archivistique sur le fonds, ce qui a été confirmé par l'enquête menée par le chantier « modèles ». En plus d'homogénéiser la production d'instruments de recherche, les modèles peuvent être un moyen de rapprocher des personnes qui, en faisant la même chose avec des outils différents, ont fini parfois par croire qu'elles faisaient un travail différent.

Enfin, SOSIE a montré que des typologies très variées peuvent être prises en charge par un modèle généraliste. Les principales difficultés tiennent à la précision de la description physique ou encore à la description des dossiers sériels, type dossiers nominatifs ; d'autres difficultés seront levées par un changement du mode de rédaction (exemple des notes de bas de page) ou par une évolution de la présentation des informations.

SOSIE a donc été considéré par le groupe de travail « modèles de description » comme le modèle de base à partir duquel pouvaient se décliner des modèles plus simples ou plus perfectionnés.

---

<sup>1</sup> À noter que les erreurs sont signalées mais n'empêchent pas la création du fichier XML.

<sup>2</sup> Le groupe « modèle de description » a également tenu compte des limites de SOSIE ; par exemple certains choix (tel le choix de la cote comme élément pivot) peuvent le rendre inutilisable pour d'autres formes d'instruments de recherche (type répertoire méthodique).

*Principales réflexions abordées dans le groupe « modélisation des instruments de recherche »<sup>1</sup>*

▪ Réfléchir au lien entre saisie et recherche

Si la finalité des modèles est de faciliter la saisie pour le rédacteur et d'uniformiser la production, les modèles servent avant tout à structurer une description avec, comme aboutissement logique, la publication de l'instrument de recherche dans la salle des inventaires virtuelle et donc la recherche et la consultation par les lecteurs. La tentation est parfois grande au moment de la saisie d'exploiter au maximum les possibilités de structuration apportées par l'EAD, ce qui alourdit la saisie et exige une grande vigilance pour appliquer la même structuration aux instruments de recherche de même catégorie<sup>2</sup> sans que le résultat soit réellement pertinent pour la recherche. D'autres moyens peuvent être mis en œuvre pour guider le rédacteur (règles de rédaction, listes d'autorité, etc.). En revanche, des éléments de description ou d'indexation ont pu être ajoutés car ils correspondent à des recherches fréquemment effectuées par les lecteurs et à des formulaires de recherche spécialisés en salle des inventaires virtuelle<sup>3</sup>.

▪ Privilégier la souplesse d'évolution et la simplicité d'utilisation

La possibilité d'évolution du système est un critère qui a été jugé déterminant. L'intégration d'une dizaine de modèles est prévue à l'ouverture du système d'information archivistique (SIA). Même si ce chiffre paraît aujourd'hui suffisant, il ne faut pas exclure une évolution des besoins voire, tout simplement, une nécessité de modifier les modèles existants (par exemple, passer à une nouvelle version de l'EAD ou rajouter un élément). Que cela soit faisable est une chose, que cela soit simple et avec le moins de conséquences possibles, techniques ou en terme d'ergonomie, en est une autre. Le groupe a donc été attentif à ces points qui lui ont fait prendre les orientations suivantes :

- sous-structuration minimale des éléments : les grilles de description utilisées aujourd'hui pour guider le rédacteur<sup>4</sup> sont en général très précises (notamment pour les éléments relatifs à la description physique des documents).

---

<sup>1</sup> Pour un complément d'information, voir le rapport du groupe de travail rédigé par Vanessa Szollosi et remis à Isabelle Neuschwander, directrice du service à compétence nationale Archives nationales.

<sup>2</sup> C'est le cas par exemple de l'élément <physdesc> : le groupe a réfléchi à la façon de l'utiliser sans alourdir la saisie, selon le type de fonds décrit et selon le type d'instrument de recherche.

<sup>3</sup> Exemple : élément <title> dans <scopecontent> pour le formulaire sceaux, afin de permettre une recherche sur la légende du sceau.

<sup>4</sup> Cas des documents figurés, des cartes et plans, etc.

Le rédacteur pourrait avoir l'impression de disposer d'un outil moins complet si le formulaire de saisie du SIA était d'apparence plus pauvre que les outils actuels. Mais une structuration trop rigide de certains éléments EAD pourrait s'avérer plus contraignante qu'utile, notamment si les éléments sont notés comme obligatoires. Cette structuration est souvent liée à l'utilisation de listes de vocabulaire contrôlé, qui elles aussi ont fait l'objet de discussions ;

- constitution de listes de vocabulaire contrôlé sous la forme de référentiels : le choix était à faire entre une inscription en dur, dans chaque modèle, d'une liste de valeurs autorisées, qui aurait ainsi pu être affectée précisément à un élément, voire à un attribut, et l'appel à des listes présentées comme des référentiels. La première solution, qui permet au rédacteur d'avoir une aide à la saisie très pointue, est également celle qui rend l'évolution du modèle délicate puisque, pour rajouter un terme à la liste, il faudrait modifier le modèle. La seconde, en revanche, en rendant indépendants modèle et liste de valeurs autorisées, permet une évolution plus facile de l'un ou de l'autre. La liste pourra être structurée et donc facile d'utilisation et le rédacteur y fera appel de la même manière que pour les référentiels proprement dits.

La combinaison de ces deux options aboutit certes à un encadrement moindre du rédacteur mais est un gage de souplesse d'évolution et d'utilisation du système. Au terme de ses réflexions, le groupe est ainsi arrivé à une simplification des modèles, avec un abandon de la sous-structuration de certains éléments et la constitution de listes de vocabulaire contrôlé dans lesquelles le rédacteur ira chercher les descripteurs correspondant aux différents éléments de description dont il a besoin : cela a semblé un compromis acceptable aux différents membres.

- Ne pas tout reporter sur l'outil

Au cours de ses travaux, le groupe a été amené à plusieurs reprises à trancher en faveur d'une gestion humaine d'opérations qui auraient pu être traitées par l'outil mais de façon non satisfaisante (lourdeur et rigidité des processus, difficultés d'évolution, déresponsabilisation des chefs de projet, etc.) ou qui ne trouvaient pas de solution technique, tout simplement parce que la question était organisationnelle. Le rôle de chef de projet a donc été réaffirmé en tant que responsable du suivi et du contrôle de la production des instruments de recherche, sans s'en remettre totalement à l'outil de production documentaire. Ce qui a permis une simplification des modèles, ne servant qu'à encadrer et guider le travail du rédacteur mais ne permettant pas un contrôle strict en dehors du respect des règles associées aux éléments EAD.



Ceci donne également une plus grande liberté à l'utilisateur, avec deux écueils : un utilisateur qui a l'impression d'avoir perdu ses repères et d'être confronté à un outil trop lâche, et un utilisateur qui serait tenté d'utiliser les ressources techniques de façon trop personnelle. Ces dangers seront à nouveau endigués par le contrôle du chef de projet.

### *Peu de modèles mais des règles de rédaction détaillées*

L'écueil auquel s'est vite trouvé confronté le groupe de travail, en partant de l'analyse des instruments de recherche actuellement produits, a été de vouloir construire un modèle à partir de chacun des exemples présentés : les utilisateurs veulent souvent que les modèles reflètent leur façon de travailler mais la réflexion sur les modèles doit également s'accompagner d'une adaptation des pratiques de chacun.

Par ailleurs, même si la majorité des instruments de recherche paraît pouvoir rentrer dans le modèle dit moyen (SOSIE), multiplier les modèles, dont certains auraient un usage très limité, pourrait laisser perplexes et hésitants les futurs utilisateurs. Or, l'idée est de simplifier le travail des rédacteurs et de trouver un juste milieu entre aider à la saisie et contraindre la rédaction.

Le groupe s'est néanmoins attaché à créer des modèles pour trois catégories d'instruments de recherche : ceux fréquemment produits (les instruments de recherche constitués de listes de dossiers par ordre alphabétique de noms de personnes ou de noms de lieux), ceux au contraire très peu nombreux mais restant une caractéristique forte des Archives nationales (les instruments de recherche des sceaux) et ceux en voie de multiplication (les instruments de recherche électroniques). Ont été exclus le modèle état des fonds<sup>1</sup>, le modèle objet<sup>2</sup> et le modèle méthodique<sup>3</sup>.

On notera qu'à une ou deux exceptions près, le groupe de travail s'est éloigné de la possibilité de créer un modèle à partir de la typologie archivistique (répertoire

---

<sup>1</sup> Ce modèle peut être obtenu par extraction et n'est pas un modèle de production d'instruments de recherche.

<sup>2</sup> Ce modèle, qui peut être utile pour les collections du Musée ou les pièces à conviction dans les fonds judiciaires, sera peut-être intégré ultérieurement ; le groupe a estimé qu'en raison du petit nombre d'instruments de recherche actuellement produits, il dispose de trop peu d'éléments de réflexion.

<sup>3</sup> Ce modèle peut, sur le fond, être englobé dans les modèles existants ; reste la question des cotes multiples pour un article et de l'éclatement d'une même cote entre plusieurs articles ; ces points doivent être abordés dans la construction du système d'information archivistique.

numérique, répertoire numérique détaillé, inventaire sommaire, inventaire analytique, etc.).

Au final, il a semblé au groupe que les modèles de description, étant donné le rôle qui leur est assigné et les options proposées ci-dessus, pouvaient être en nombre limité, tout en couvrant le champ fixé :

- modèle dématérialisation ;
- modèle généraliste (SOSIE) ;
- modèle entrées/simplifié ;
- modèles documents audiovisuels/sonores ;
- modèle dossiers nominatifs ;
- modèle dossiers noms géographiques ;
- modèle documents figurés ;
- modèle sceaux.

À ces modèles sont associés de nombreux éléments qui les enrichissent nécessairement : listes de vocabulaire contrôlé, référentiels, contrôle scientifique et règles de rédaction. Ces dernières, faisant également partie des objectifs du chantier, seront déterminantes. Le manuel d'utilisateur, qui pourra être développé à partir de celui créé pour SOSIE, définira plus précisément pour chaque élément de chaque modèle les règles (élément obligatoire ou non, répétable ou non), le contenu attendu et des exemples de rédaction. C'est dans ces règles de rédaction que l'on pourra donner toutes les consignes indispensables pour produire un instrument de recherche satisfaisant.

Pour chaque modèle est définie une DTD (adoption de la DTD-EAD), listant les éléments utilisés et leurs règles d'utilisation (obligatoire, répétable ou non, etc.).

Un certain nombre de règles seront valables quel que soit le modèle. Il s'agit notamment du caractère obligatoire ou facultatif et de la répétabilité des éléments, des imbrications d'éléments permises (notamment les « p », « list », etc.) et des champs dont le nombre de caractères sera limité, définis par la DTD-EAD SIA.

*Conclusion : des instruments de recherche souples d'utilisation et évolutifs*

Même si la normalisation était attendue dans le cadre du futur système d'information archivistique, les travaux des groupes d'utilisateurs, dont celui du chantier « modèles de description », ont révélé des attentes fortes d'encadrement humain face à une technique exigeante, conjuguées à une mise en œuvre souple et évolutive.

Certains modèles d'instruments de recherche seront proposés en version formulaire de saisie. Les modèles peuvent être combinés entre eux ; ils ne sont plus figés mais peuvent évoluer : ils peuvent être fusionnés, entrer en relation. Ils pourront faire l'objet de reprises, de détails, de suppléments, en fonction de la recherche, des demandes du public, etc. La production et la mise en ligne rapide des instruments de recherche pourront générer la production de nouveaux instruments de recherche mieux adaptés à la demande, et, peut-être, une nouvelle organisation du travail de description.

Isabelle ARISTIDE-HASTIR

Conservateur en chef, chargé de la dématérialisation des instruments de recherche dans le cadre du projet de nouveau bâtiment des Archives nationales à Pierrefitte-sur-Seine  
isabelle.aristide@culture.gouv.fr

## ANNEXE 1

### **Exemple de masque de saisie Inventaires de dossiers nominatifs papier/électroniques**

#### **Identification**

Caractère confidentiel du nom en intitulé <c> @audience

Référence <did><unitid>

Intitulé <did><unittitle><persname> (saisie libre)

Dates <did><unitdate> [règles génériques]

Niveau de description <c> @level [règles génériques]

Importance matérielle <did><physdesc><extent>

#### **Contexte**

Histoire administrative/biographie <bioghist>

Historique de la conservation <custodhist>

Dates d'existence, de nomination, de cessation de fonctions <bioghist><date type="type de date"> [règles génériques] = une liste de type de dates devra être définie

Modalités d'entrée <acqinfo> [règles génériques]

#### **Contenu et structure**

Présentation du contenu <scopecontent>

Évaluation, tris et éliminations <appraisal>

Mode de classement <arrangement>

#### **Conditions d'accès et d'utilisation**

Conditions d'accès <accessrestrict> [règles génériques]

Conditions de reproduction <userrestrict> [règles génériques]

Instrument de recherche associé <otherfindaid>

#### **Sources complémentaires**

Mentions relatives aux originaux <originalsloc>

Mentions relatives aux copies <altformavail>

Sources complémentaires de même provenance <separatedmaterial>

Sources complémentaires sans lien de provenance <relatedmaterial>

Bibliographie <bibliography> [règles génériques]

## Indexation

Collectivité <controlaccess><corpname role="rôle" source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [AFNOR NF-Z-44-060 et référentiel]

Personne physique <controlaccess><persname role="rôle" source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [AFNOR NF-Z-44-061 et référentiel]

Famille <controlaccess><famname role="rôle" source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [AFNOR NF-Z-44-061 et référentiel]

Nom de lieu <controlaccess><geogname source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [AFNOR NF-Z-44-081 et référentiel EHESS]

Mot-matière <controlaccess><subject source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [référentiel]

Fonction administrative <controlaccess><function source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [référentiel]

Activité (d'une personne) <controlaccess><occupation source="référentiel" authfilenumber="identifiant\_référentiel" rules="règles"> [référentiel]

Titre <controlaccess><title rules="règles"> [règles génériques]

ANNEXE 2

## Le modèle « archdesc »

|  |  |
|--|--|
| <b>Identification</b>                      | c/@level   |
|  | did/unitid @type="identifiant" ou "cote-de-consultation"                                 |
|  | did/unittitle  |
|  | did/unitdate ou did/unittitle/unitdate   |
|  | did/physdesc   |
|  | did/repository   |
|  | did/physloc  |
| <b>Contexte</b>                            | did/origination corpname/persname/famname/origine inconnue                               |
|  | c/bioghist   |
|  | c/custodhist   |
|  | c/acqinfo  |
| <b>Contenu et structure</b>                | c/scopecontent   |
|  | c/appraisal  |
|  | c/arrangement  |
|  | c/accruals   |
| <b>Conditions d'accès et d'utilisation</b> | c/accessrestrict   |
|  | c/userrestrict   |
|  | did/langmaterial/language  |
|  | c/phystech   |
|  | c/otherfindaid   |
| <b>Sources complémentaires</b>             | c/originalsloc   |
|  | c/altformavail   |
|  | c/separatedmaterial  |
|  | c/relatedmaterial  |
|  | c/bibliography   |
| <b>Indexation</b>                          | c/controlaccess/ corpname persname famname geogname<br>subject function occupation title |

ANNEXE 3

**Tableau récapitulatif d'utilisation des éléments EAD  
pour les modèles d'instruments de recherche**

|                       | modèle général (SOSIE)   | modèle simplifié | documents figurés                 | audio-vidéo          | sceaux                            | dossiers nominatifs     | dossiers par nom de lieu |
|-----------------------|--|------------------|-----------------------------------|----------------------|-----------------------------------|-------------------------|--------------------------|
| <b>Identification</b> | c/@level   | x                | x                                 | x                    | x                                 | x                       | x                        |
|                       |  |                  |                                   |                      |                                   | c/@audienc<br>e         |                          |
|                       | did/unitid<br>@type="identifiant" ou "cote-de-consultation"<br>+ "dossier-piece" (ce dernier est à ajouter au modèle SOSIE existant) | x                | x                                 | x                    | x                                 | x                       | x                        |
|                       | did/unittitle  | x                | x                                 | x                    | x                                 | did/unittitle /persname | did/unittitle /geogname  |
|                       | did/unitdate ou did/unittitle /unitdate  | x                | x                                 | x                    | x                                 | uniquement did/unitdate | uniquement did/unitdate  |
|                       | did/physdesc   |                  | + extent / physfacet / dimensions | + extent / physfacet | + extent / physfacet / dimensions | + extent                | + extent                 |
|                       |  |                  | did/materialspec                  | did/materialspec     |                                   |                         |                          |
|                       | did/repository   |                  |                                   |                      |                                   |                         |                          |
|                       | did/physloc  |                  |                                   |                      |                                   |                         |                          |

|  | modèle général (SOSIE)  | modèle simplifié | documents figurés | audio-vidéo    | sceaux    | dossiers nominatifs                               | dossiers par nom de lieu                          |
|--|---|------------------|-------------------|----------------|-----------|---|---|
| <b>Contexte</b>                            | did/origination<br>corp-name/persname/<br>fam-name/origine inconnue |                  | x                 |                | x         |   |   |
|  | c/bioghist  |                  | x                 | x              | x         | + bioghist/<br><b>date</b><br>type="type de date" | + bioghist/<br><b>date</b><br>type="type de date" |
|  | c/custodhist  |                  | x                 | x              | x         | x   | x   |
|  | c/acqinfo   |                  | x                 | x              | x         | x   | x   |
| <b>Contenu et structure</b>                | c/scopecontent  |                  | + p/title         | + p/title      | + p/title | x   | x   |
|  | c/appraisal   |                  | x                 | x              | x         | x   | x   |
|  | c/arrangement   |                  | x                 | x              | x         | x   | x   |
|  | c/accruals  |                  | ?                 | ?              | x         | ?   | ?   |
| <b>Conditions d'accès et d'utilisation</b> | c/accessrestrict  |                  | x                 | x              | x         | x   | x   |
|  | c/userrestrict  |                  | x                 | x              | x         | x   | x   |
|  | did/langmaterial/language   |                  | x                 | x              | x         | x   | x   |
|  |   |                  |                   | c/<br>phystech |           |   |   |
|  | c/otherfindaid  |                  | x                 | x              | x         | x   | x   |



|                                | modèle général (SOSIE)   | modèle simplifié | documents figurés  | audio-vidéo   | sceaux   | dossiers nominatifs  | dossiers par nom de lieu   |
|--------------------------------|--|------------------|--|---|--|--|--|
| <b>Sources complémentaires</b> | c/originalsloc   |                  | x  | x   | x  | x  | x  |
|                                | c/altformavail   |                  | x  | x   | x  | x  | x  |
|                                | c/separated material   |                  | x  | x   | x  | x  | x  |
|                                | c/related material   |                  | x  | x   | x  | x  | x  |
|                                | c/bibliography   |                  | x  | x   | x  | x  | x  |
| <b>Indexation</b>              | <p>c/<br/>controlaccess/corpname<br/>persname<br/>famname<br/>geogname<br/>subject<br/>title<br/>function<br/>occupation</p> <p>(ces deux derniers sont à rajouter au modèle SOSIE existant)</p> |                  | <p>+ corp-name@role<br/>pers-name@role<br/>fam-name@role<br/>@source (référentiel)<br/>@authfile<br/>number (lien vers notice d'autorité)<br/>@rules</p> <p>(règles suivies pour forme normalisée)<br/>genreform</p> | <p>+ corp-name@role<br/>pers-name@role<br/>fam-name@role<br/>+ corp name @role<br/>pers name @role<br/>fam name @role</p> | <p>+ corp-name@role<br/>pers-name@role<br/>fam-name@role<br/>@source (référentiel)<br/>@authfile<br/>number (lien vers notice d'autorité)<br/>@rules</p> <p>(règles suivies pour forme normalisée)<br/>genreform</p> | <p>+ corp-name@role<br/>pers-name@role<br/>fam-name@role<br/>@source (référentiel)<br/>@authfile<br/>number (lien vers notice d'autorité)<br/>@rules</p> <p>(règles suivies pour forme normalisée)</p> | <p>+ corp-name@role<br/>pers-name@role<br/>fam-name@role<br/>@source (référentiel)<br/>@authfile<br/>number (lien vers notice d'autorité)<br/>@rules</p> <p>(règles suivies pour forme normalisée)</p> |