

Cinquante nuances d'EAD : réconcilier la théorie et la pratique de la description des manuscrits à la Bibliothèque nationale de France

Maryline Devidal

Département des métadonnées, Bibliothèque nationale de France, Paris, France.
maryline.devidal@bnf.fr

Mélanie Roche

Département des métadonnées, Bibliothèque nationale de France, Paris, France.
melanie.roche@bnf.fr



Copyright © 2018 by Maryline Devidal, Mélanie Roche. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

Résumé :

La description des manuscrits requiert tout autant de soins que leur manipulation : fruit d'arbitrages multiples naviguant sans cesse entre choix intellectuels et techniques, elle nécessite en effet une vision à long terme et une coordination à la fois locale et nationale, au même titre que la conservation, la gestion des collections ou la valorisation des documents. En France, c'est avec l'informatisation des catalogues de manuscrits que s'est développée cette coopération, à la fois à l'échelon local de la Bibliothèque nationale de France (BnF) et à l'échelon plus large du pays tout entier. Cet article porte un regard sur le passé, le présent et l'avenir de la description des manuscrits en EAD, interrogeant les dimensions intellectuelles et techniques que représente l'adoption d'un format d'encodage spécifique.

Mots-clés : EAD, catalogage, coordination, coopération

Introduction

Le catalogage informatisé démarre à la BnF en 1973 pour les documents édités, avec la création du format de catalogage INTERMARC, format maison cousin du MARC21 et de l'UNIMARC. Les documents non édités (au premier rang desquels les manuscrits) font dès lors figure de parent pauvre dans l'environnement informatique de l'établissement, une situation d'autant plus dommageable pour ces collections prestigieuses qu'elle perdure jusqu'au début des années 2000. Au niveau national, la situation est similaire : forte d'une longue tradition de coopération, la France s'est dotée très tôt de catalogues collectifs des manuscrits (le *Catalogue Général des Manuscrits des Bibliothèques publiques de France*, ou

CGM, apparaissant comme l'aboutissement de ces efforts), mais tarde à les faire entrer dans l'ère numérique. Rien d'étonnant donc à ce que, dans la décennie 2000 qui marque les débuts de l'informatisation des catalogues de manuscrits, les réflexions locale et nationale s'articulent, se répondent, voire convergent.

En plus des enjeux techniques interviennent des questions intellectuelles touchant à l'identité même de la profession, attachée au caractère scientifique de ses descriptions : à document unique, description exceptionnelle. Comment la machine pourrait-elle restituer la richesse et la saveur propres au catalogage de tels documents ? On verra que le choix d'un format de catalogage, loin d'être une simple question technique, touche en fait au cœur du métier de bibliothécaire, au même titre que la conservation, la communication ou la valorisation des documents. L'implémentation de l'EAD à la BnF, et plus largement dans les bibliothèques françaises, en est la preuve flagrante.

1. L'EAD, UN CHOIX DECISIF

1.1 Choix techniques, choix intellectuels

Format élaboré dans le cadre de l'Université de Berkeley à partir de 1993 et promu par la Société des Archivistes Américains, l'EAD est un mode de structuration de l'information adapté au traitement des fonds d'archives. Il bénéficie en effet des propriétés du langage XML, sur lequel il repose : souplesse et adaptabilité, gestion de la hiérarchie grâce au principe de la branche XML, intégration dans les systèmes informatiques et sur le web, interopérabilité avec d'autres formats, principe de balisage des données (des balises ouvrante et fermante encadrent les éléments d'information signifiants).

L'EAD offre la possibilité de gérer jusqu'à 12 niveaux hiérarchiques : un niveau de description général pour le fonds, puis des niveaux de description emboîtés les uns dans les autres qui structurent la description. En France, la Direction des archives de France défend depuis plusieurs années son utilisation pour les instruments de recherche des archives nationales et locales. Or le principe convient également à la description des collections de manuscrits, dont il permet de restituer le plan de classement, et aux manuscrits composites, dont il permet de décrire les différentes unités codicologiques. Cette arborescence est restituée dans les interfaces publiques de consultation, offrant à l'utilisateur une navigation dans le fonds, la collection, ou le document.

Au début des années 2000, la question de la description en EAD des manuscrits de la BnF se pose en lien avec l'informatisation des catalogues. Tirant ses origines de la bibliothèque des rois de France, la collection de manuscrits de la BnF mêle manuscrits médiévaux, modernes et contemporains, aux formes et usages variés : copies anciennes, *unica*, textes scientifiques, manuscrits décorés et à peintures, manuscrits littéraires, correspondance, papiers personnels, iconographie, épreuves imprimées annotées... Le catalogue du département des manuscrits reflète la diversité de ces documents, dont le futur format informatique devra rendre compte, tout en en restituant la hiérarchie et en gérant l'héritage de différents niveaux de description. La notion de catalogue informatisé n'existe pas ou peu à cette époque : la décision d'informatiser le catalogue du département des Manuscrits représente donc un véritable changement de paradigme, et l'une des principales craintes concerne alors moins le format de production que la notion d'informatisation elle-même : comme souvent, les considérations techniques passent au second plan par rapport aux réticences intellectuelles.

L'EAD apparaît pourtant comme tout désigné, car ce format permet aussi de saisir du texte libre, notamment dans les balises de description du contenu, du contexte de production, ou encore de la description physique, ce qui permet de respecter la tradition des notices rédigées des catalogues imprimés. L'EAD offre également la possibilité d'indexer, au fil du texte, des noms de personne, de collectivité, de famille, des noms géographiques, des titres d'œuvres ou des sujets. Par ailleurs, la souplesse des éléments d'indexation permet de saisir à la fois une forme normalisée du point d'accès et une forme littérale pour les usagers. Toutes ces possibilités permettent d'envisager l'EAD non seulement pour le chargement des notices issues de la conversion rétrospective, mais aussi dans une optique de production courante, dans la tradition de notices scientifiques et littéraires du département des manuscrits.

En 2001, une expérimentation menée sur le dernier volume du catalogue des Nouvelles Acquisitions Françaises à la BnF confirme que l'EAD, format conçu par des archivistes, est parfaitement adaptable aux collections de manuscrits présentes non seulement à la BnF mais dans toutes les bibliothèques françaises. En effet, parallèlement à la décision du Département des manuscrits de convertir ses catalogues au format EAD, à l'échelle nationale le Ministère de la culture étudie alors les modalités de la conversion du CGM, et envisage la création, à terme, d'un catalogue informatique collectif des manuscrits, ce qui implique une nécessaire interopérabilité des catalogues.

1.2 Les chantiers de rétroconversion

Le même prestataire (ArchProteus) réalise le balisage dans les deux conversions, locale et nationale. Dès la rédaction des spécifications, il convient d'effectuer des choix de traitement aussi cohérents que possible. Les mêmes personnes se trouvant impliquées dans les deux projets, une certaine homogénéité se fait jour, sans toutefois que cette logique soit systématiquement mise en œuvre. On peut citer comme exemple l'emploi de `<unitid type="numéro de notice">` dans la conversion du CGM : les « numéros de notices » des volumes papier (qui correspondaient généralement aux cotes des manuscrits) ont été encodés comme tels, alors que dans la conversion des catalogues du département des Manuscrits, les cotes ont été saisies en `<unitid type="cote">`.

Au département des Manuscrits de la BnF, le chantier de conversion concerne la quasi-totalité des fonds pour lesquels on dispose de catalogues¹, c'est-à-dire des documents de nature et de types très variés auxquels on applique des règles de catalogage particulières : un certain nombre de différences existent ainsi entre la description d'un estampage chinois et la description d'une Bible latine. Il apparaît donc d'emblée impossible de déterminer un cahier des charges unique pour l'ensemble de la prestation. Au contraire, des spécifications particulières sont rédigées pour chaque fonds, catalogue, et même parfois volume au sein d'un même catalogue. Les spécifications pour l'encodage en EAD doivent ainsi prendre en compte la langue, les tranches de cotes, la présence d'index, etc. et résultent aussi bien de choix communs, de compromis sur l'usage de telle ou telle balise pour exprimer telle ou telle information, que de choix spécifiques à certains fonds.

Le chantier de conversion s'étale sur six ans et se divise en trois phases : numérisation des volumes imprimés, océrisation, et enfin balisage du texte reposant en grande partie sur la mise en forme des notices imprimées. Les catalogues pour lesquels on possède la matrice

¹ Certains petits fonds n'ont pas été inclus dans le chantier de conversion rétrospective et ont fait l'objet d'un encodage manuel en EAD par les chargés de collections.

électronique au format Word sont encodés après qu'une feuille de style leur a été appliquée, permettant de restituer les différents niveaux de l'arborescence ainsi que la nature des informations à baliser (cote, date, titre, etc.). Au fil des relectures, les avantages de l'EAD sont rapidement perçus et sous l'impulsion de Thierry Delcourt, alors directeur du département des Manuscrits, tous les services basculent vers le catalogage courant en EAD. C'est alors que l'on prend conscience que les réalités techniques du catalogage courant contredisent les choix intellectuels pris au moment de la rétroconversion. En effet, la conversion rétrospective des catalogues du département des Manuscrits prévoyait la production d'un fichier EAD par volume. Les fichiers livrés par le prestataire sont donc si volumineux que l'outil de production n'est pas en mesure de les intégrer. Le département des Manuscrits prend alors la décision de faire redécouper les fichiers EAD afin de produire un fichier par manuscrit, décision qui est intégrée aux spécifications de rétroconversion dès 2007.

Cette décision est lourde de conséquences. En effet, parallèlement à la conversion des catalogues du département des Manuscrits se déroule la conversion du CGM, pour lequel la production d'un fichier EAD par volume est respectée jusqu'au terme du chantier. Les fichiers EAD sont ensuite distribués dans deux catalogues distincts : le Catalogue général des manuscrits d'une part, pour les établissements relevant du Ministère de la culture, et Calames d'autre part pour les établissements relevant du Ministère de l'enseignement supérieur et de la recherche. Or lors de la mise en ligne du portail national « Manuscrits et archives » sur le site du Catalogue collectif de France (CCFr), on s'aperçoit que les choix opérés ont une portée critique pour l'interopérabilité. Ce portail permet de localiser des collections de manuscrits ou d'archives dans tous types de bibliothèques (nationale, territoriales, universitaires...), par l'interrogation simultanée ou indépendante des quatre grands catalogues français informatisés au format EAD (BnF Archives et manuscrits, le CGM, Calames et Palme, le Répertoire des manuscrits littéraires français du XXème siècle). Or les choix divergents quant à la structuration des fichiers EAD représentent un défi technique de taille pour agréger les données dans une même interface, de même que les choix divergents sur la saisie des cotes.

1.3 Le défi de l'ouverture

Parallèlement, d'autres départements de la BnF entrent dans l'aventure. Le département des Arts du spectacle, qui a adopté depuis plusieurs années la norme internationale de description archivistique ISAD(G)² pour produire, sous Word, des instruments de recherche décrivant ses fonds d'archives, est le premier à s'intéresser à l'EAD. Prenant acte de la capacité d'EAD à préserver la cohérence organique du fonds et à en donner une vue d'ensemble hiérarchisée, le département des Arts du spectacle décide de l'adopter pour la saisie rétrospective, puis courante, de ses inventaires de fonds d'archives. Les fonds du département, très riches et très variés, concernent tous les domaines du spectacle vivant : théâtre, marionnettes, danse, mime, musique, music-hall, cirque, arts de la rue, et même pyrotechnie ; mais également les domaines du cinéma, de la radio et de la télévision ; ainsi que le festival d'Avignon, puisque la Maison Jean Vilar est rattachée au département.

C'est ensuite la bibliothèque de l'Arsenal qui rejoint la communauté des producteurs EAD de la BnF. L'Arsenal possède des fonds très proches de ceux conservés au département des

² ISAD (G) : norme générale et internationale de description archivistique, deuxième édition, 1999, https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_FR.pdf

Manuscrits (manuscrits médiévaux, modernes et contemporains), ainsi que des fonds d'archives administratives et littéraires. C'est en fait dans le cadre de la conversion du CGM, réalisée par le service du Catalogue collectif de France, que l'Arsenal produit, ou plutôt reçoit ses premiers fichiers EAD. Il a ensuite été tout naturellement décidé d'encoder en EAD les descriptions des archives administratives de l'Arsenal, ainsi que des fonds qu'elle conserve, soit par conversion rétrospective complémentaire, soit nativement.

Le travail mené par ces départements précurseurs permet d'établir que l'EAD convient aussi bien à la description de collections de manuscrits que de fonds d'archives, dans la granularité souhaitée : description à la pièce pour les manuscrits unitaires, ou hiérarchique pour les fonds. Le projet initial d'informatisation des catalogues des manuscrits de la BnF s'est donc tout naturellement transformé en un projet de catalogue informatisé des fonds d'archives et des manuscrits. Le catalogue est baptisé BAM (pour « BnF Archives et Manuscrits ») afin de refléter cette double identité, et s'ouvre progressivement aux autres départements : le département de la Musique, le département de l'Audiovisuel, et le département des Monnaies, médailles et antiques³.

2. LA PRODUCTION COURANTE : LE DEFI DE L'INTEROPERABILITE

Parallèlement aux chantiers de rétroconversion, la BnF se dotait d'une interface de production des données en EAD (baptisée PiXML) et d'une interface de consultation publique (BAM). Si ces outils sont communs à tout l'établissement, la variété des niveaux de granularité (description unitaire des manuscrits *versus* instruments de recherche complexes), des typologies documentaires et des formes matérielles à décrire (manuscrits en codex, en liasse, en feuilles, rouleaux, estampages, documents iconographiques, papiers personnels, correspondance, matériaux divers liés à des spectacles, costumes, objets, enregistrements sonores, documents audiovisuels ou multimédias, archives numériques...) induit des besoins différents chez les producteurs. Par ailleurs, la question de l'intégration de ces outils dans l'écosystème de la production bibliographique de l'établissement est différée, ce qui n'a pas été sans poser problème par la suite.

2.1 À la BnF, la balkanisation des pratiques

En 2004 est créé un groupe de travail ayant pour mission d'identifier les balises à utiliser dans le cadre du catalogage courant en EAD au département des Manuscrits. Le groupe recense les balises minimales requises grâce au travail en cours pour la conversion rétrospective, et élargit l'éventail des balises autorisées en étudiant l'encodage nécessaire à l'enrichissement de ces notices. Cette seconde étape repose sur des projections établies à partir des processus et règles de catalogage en vigueur au département, et des fonctionnalités souhaitées pour l'exploitation des données dans BAM (gestion de l'arborescence, exploitation de l'héritage, constitution des index, etc.). Le groupe ne retient en fin de compte qu'un peu plus du tiers des 146 éléments de la DTD.

Au-delà de l'identification des balises à utiliser, le groupe a pour mission de définir le contenu des balises et leur mode d'emploi. Il est aussi chargé de définir les attributs à utiliser

³ La mission pour la Gestion de la production documentaire et des archives de l'établissement constitue un cas à part, dans la mesure où les instruments de recherche en EAD décrivent, non pas les collections de la BnF, mais ses archives administratives, c'est-à-dire l'ensemble des documents professionnels, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par l'établissement et ses personnels dans l'exercice de leur activité.

et les référentiels à leur associer. Enfin, une réflexion sur les points d'accès et sur la cohérence avec les autres catalogues de la BnF est envisagée. On s'interroge notamment sur l'emploi des éléments d'indexation <persname>, <corpname>, <famname>, <title>, <geogname>, pour lesquels on décide de créer des liens vers le fichier d'autorité de la BnF dans les points d'accès. La fonction exercée sur le document se fera grâce à l'attribut ROLE, fondé sur le référentiel INTERMARC des codes de fonction pour le Catalogue général, ceci afin de réunir les conditions d'une recherche fédérée sur les différents catalogues de la BnF qu'il reste à inventer.

Dans l'intervalle, le département des Arts du spectacle a commencé à décrire ses fonds en EAD dans PiXML et, les besoins spécifiques aux documents qu'il conserve n'étant pas encore abordés, il instruit et produit ses propres consignes d'encodage. Lors de leur entrée successive dans BAM, les départements, confrontés à des besoins propres, adoptent des pratiques *ad hoc*. La production dans BAM se balkanise, et les écarts se creusent dans le traitement des données. Face à cette balkanisation des pratiques, le besoin d'une instance de coordination transverse réunissant les départements producteurs, le service informatique et le département des métadonnées se fait de plus en plus pressant. Faute de support de poste dédié, une telle coordination peine toutefois à se mettre en place. De plus, la coordination EAD ne s'intègre pas dans la coordination bibliographique de la BnF, centrée autour du Catalogue général et du MARC, contribuant ainsi à faire du catalogage en EAD un objet à part dans l'organisation de la production bibliographique. Avec le temps, PiXML et BAM connaissent en outre des dysfonctionnements croissants au fur et à mesure que la production augmente. Très rapidement, les questions techniques prennent donc le pas sur la coordination intellectuelle.

2.2 Une tentative de normalisation nationale

C'est finalement d'une tentative de normalisation nationale que viendra une première réponse. Au début des années 2000, un groupe de travail AFNOR⁴ « Description bibliographique des manuscrits modernes et contemporains » est créé afin d'élaborer une recommandation pour la description des manuscrits et fonds d'archives modernes et contemporains en bibliothèque. Le groupe est constitué de représentants de la Bibliothèque nationale de France (département des Manuscrits, département des Arts du spectacle, bibliothèque de l'Arsenal), de bibliothèques universitaires et de grands établissements, ainsi que de bibliothèques municipales. Le Service interministériel des archives de France est également représenté. La recommandation *DeMarch*⁵ est publiée en 2010.

En 2008, un autre groupe de travail national se met en place autour du projet de rédaction d'un guide national d'application de l'EAD dans les bibliothèques. La conversion rétrospective du CGM vient alors de s'achever et, dans la perspective de la réalisation d'un catalogue national des fonds de manuscrits et d'archives interrogeant simultanément le CGM, Calames et BAM, on souhaite accompagner les bibliothèques dans l'enrichissement et la mise à jour des notices, afin de garantir l'homogénéité du futur catalogue. Le projet associe des représentants des bibliothèques de lecture publique, des bibliothèques universitaires, ainsi que des représentants des départements des Manuscrits et des Arts du spectacle de la BnF. Des personnalités du monde des archives sont également invitées à participer aux travaux, de manière à conserver une cohérence avec l'EAD pratiqué par le monde des archives.

⁴ Association Française de Normalisation, miroir de l'ISO au niveau national.

⁵ *Description des Manuscrits et Fonds d'archives modernes et contemporains en bibliothèque.*

Le guide est conçu comme un profil d'application du format EAD. Les éléments sont déclinés en autant d'éléments obligatoires, recommandés, autorisés ou déconseillés (voir figure 1) pour réaliser des instruments de recherche à des niveaux de granularité variable : du signalement contenant uniquement les niveaux supérieurs <eadheader> et <archdesc> et les éléments obligatoires ou recommandés peuplant le <did>, au répertoire détaillé contenant tous les niveaux de description et tous les éléments nécessaires à une identification fine. On indique également pour chaque élément ses attributs obligatoires, recommandés, autorisés, déconseillés, ses éléments fils et pères, et un lien à la bonne pratique (voir figure 2). L'utilisation des attributs est préconisée dans un certain nombre d'éléments afin de préciser la nature de l'information encodée ou d'en donner une forme normalisée. Des listes de valeurs possibles ou des règles de normalisation sont généralement proposées. Des consignes d'indexation sont également fournies.

Afficher l'intégralité | Cacher les éléments et attributs déconseillés

<p>Éléments obligatoires <archdesc> <did> <ead> <eadheader> <eadid> <filedesc> <titleproper> <titlestmt></p> <p>Éléments recommandés <author> <c> <change> <creation> <descrules> <dsc> <language> <language> <origination> <profiledesc> <publicationstmt> <repository> <unitid> <unittitle></p> <p>Éléments autorisés <abbr> <accessrestrict> <accruals> <acqinfo> <address> <addressline> <altformavail> <appraisal> <archref> <arrangement> <bibliography> <bibref> <bioghist> <blockquote> <controlaccess> <corpname> <custodhist> <dao> <daodesc> <daogrp> <daoloc> <date> <dimensions> <div> <edition> <editionstmt> <emph> <expand> <extent> <extptr> <extref> <extrefloc> <famname> <fileplan> <function> <genreform> <geogname> <head> <item> <langmaterial> <lb> <linkgrp> <list> <materialspec> <name> <note> <num> <occupation> <originalsloc> <otherfindaid> <p> <persname> <physdesc> <physfacet> <physloc> <phystech> <prefercite> <processinfo> <ptr> <publisher> <ref> <relatedmaterial> <revisiondesc> <scopecontent> <separatedmaterial> <seriesstmt> <sponsor> <subject> <subtitle> <title> <unitdate> <userrestrict></p> <p>Éléments autorisés pour des projets spécifiques <bibseries> <container> <frontmatter> <imprint> <notestmt> <odd></p> <p>Éléments déconseillés <abstract> <arc> <archdescgrp> <c01> <c02> <c03> <c04> <c05> <c06> <c07> <c08> <c09> <c10> <c11> <c12> <chronitem> <chronlist> <colspec> <defitem> <descgrp> <dscgrp> <eadgrp> <entry> <event> <eventgrp> <extptrloc> <head01> <head02> <index> <indexentry> <label> <legalstatus> <listhead> <namegrp> <ptrgrp> <ptrloc> <refloc> <resource> <row> <runner> <subarea> <table> <tbody> <tgroup> <thead> <titlepage></p>	<p>Attributs obligatoires CALENDAR COUNTRYENCODING DATEENCODING ERA LANGCODE LANGENCODING REPOSITORYENCODING SCRIPTENCODING</p> <p>Attributs recommandés ABBR EXPAN MAINAGENCYCODE</p> <p>Attributs autorisés ACTUATE ARCROLE AUDIENCE AUTHFILENUMBER CERTAINTY CONTINUATION COUNTRYCODE HREF ID IDENTIFIER LABEL LEVEL LINKTYPE NORMAL OTHERLEVEL RELATEDENCODING RENDER ROLE RULES SCRIPTCODE SHOW SOURCE TARGET TITLE TYPE UNIT XMLNS XPOINTER</p> <p>Attributs déconseillés ALIGN ALTHEAD ALTRENDER CHAR CHAROFF COLNAME COLNUM COLS COLSEP COLWIDTH DATECHAR FRAME FROM MARK MOREROWS NAMEEND NAMEST NUMERATION OTHERTYPE PARENT PGWIDE PLACEMENT PUBLICID REPOSITORYCODE ROWSEP TO TPATTERN URL URN VALIGN</p>
---	---

Légende : obligatoire | recommandé | autorisé | autorisé sur projet spécifique | déconseillé


Figure 1 : la liste des éléments EAD dans Le guide des bonnes pratiques de l'EAD en bibliothèque

Affichage des éléments et attributs par **nature ou type** | **ordre alphabétique** | **statut**

[Afficher l'intégralité](#) | [Cacher les éléments et attributs déconseillés](#)

<unittitle> Intitulé de l'unité documentaire

Consulter le dictionnaire des balises [<ead>](#)

Consulter la bonne pratique 

Attributs :

ID
ALTRENDER
AUDIENCE
ENCODINGANALOG
LABEL
TYPE

Contenu textuel : Autorisé

Statut des éléments fils de <unittitle> :

Autorisé : <ptr> <extptr> <emph> <lb> <abbr> <expan> <ref> <extref> <bibref> <title> <archref> <corpname>
<famname> <geogname> <name> <occupation> <persname> <subject> <genreform> <function> <unitdate> <num> <date>
Autorisé sur projet spécifique : <bibseries> <edition> <imprint>
Déconseillé : <linkgrp>

Statut de l'élément <unittitle> :

Recommandé au sein de : <did>
Déconseillé au sein de : <p> <ref> <refloc> <extref> <extrefloc> <archref> <event> <item> <label> <entry>

Légende : **obligatoire** | **recommandé** | **autorisé** | **autorisé sur projet spécifique** | **déconseillé**

Figure 2 La liste des attributs possibles dans un élément EAD

Le *Guide des bonnes pratiques de l'EAD en bibliothèque*⁶, mis en ligne en 2012, est compatible avec ISAD(G) et *DeMarch*. Il définit un profil d'application commun suffisamment souple pour ne pas être inutilement contraignant, et suffisamment précis pour garantir la production de données interopérables. On peut noter cependant que le *Guide* ne couvre pas tout à fait l'ensemble des cas d'usage que l'on rencontre en bibliothèque. En effet, à l'époque où le groupe se réunit, les départements de l'Audiovisuel et de la Musique de la BnF ne sont pas encore entrés dans BAM, ce qui explique qu'un certain nombre de besoins ne soient pas pris en compte : par exemple, la description des particularités physiques et techniques des documents sonores, audiovisuels, multimédias ou numériques natifs ; ou encore la description de la présentation et/ou de la distribution musicale pour les manuscrits musicaux.

2.3 La mise en place d'une coordination

Dès leur parution, la recommandation AFNOR et le *Guide des bonnes pratiques de l'EAD en bibliothèque* deviennent les documents de référence à la BnF, et les bonnes pratiques sont peu à peu diffusées dans les consignes de production et dans l'offre de formation commune à tous les départements. Chacun est donc censé produire le même EAD, quelle que soit la composante à laquelle il appartient. Dans les faits cependant on est encore loin du compte. Le projet de refonte de BAM, initié en 2015, met au jour le manque flagrant de cohérence des

⁶ <http://bonnespratiques-ead.net/>

données, et le groupe de travail composé à cette occasion se heurte à un nombre considérable de problèmes.

Le poids des conversions et le recours inégal en production courante à l'indexation des noms, lieux, titres d'œuvres et sujets, font qu'il existe des centaines de milliers d'entités balisées sans distinction en <name> et plus encore d'entités non balisées. L'absence de liens vers un fichier d'autorités, de points d'accès normalisés dans l'attribut NORMAL et de saisie du code de fonction dans l'attribut ROLE, produisent des index à la fois incomplets et difficilement exploitables. Les différences de pratiques dans la saisie des valeurs de LEVEL interdisent de se fonder sur cet attribut pour gérer la hiérarchie, laquelle repose uniquement sur l'emboîtement des composants. Les différences de pratiques dans la saisie des cotes, ajoutées à la mauvaise utilisation du LEVEL « item », empêchent l'identification *a priori* du niveau correspondant à l'unité communicable. Le recours à l'élément <genreform> est également inégal, de sorte qu'il est en fait impossible de connaître la proportion de documents décrits dans BAM qui relèvent du texte, de l'image fixe ou animée, du son, de la musique notée, des objets, etc.

À la lumière de ces constats, on décide en 2016 de consolider la coordination transverse autour de l'utilisation de l'EAD à la BnF. Un support de poste « Administration des données EAD » est créé au sein du département des Métadonnées, mêlant étroitement les problématiques liées à l'application du format et les problématiques liées à la cohérence des données. La coordinatrice s'entoure d'un groupe d'experts issus de l'ensemble des départements producteurs de données, de manière à représenter tous les points de vue. Une experte en modélisation ainsi qu'une informaticienne complètent le groupe. Le périmètre de cette coordination permet d'envisager l'application de l'EAD dans tous ses aspects : définition du profil d'application dans le respect des bonnes pratiques, suivi des chantiers nécessaires à la cohérence des données, formalisation et diffusion des consignes de catalogage, suivi des applications PiXML et BAM. Au plan national et international la coordinatrice participe également aux instructions ayant un impact sur l'EAD. Cette fiche de poste permet d'identifier et d'intégrer la coordination de l'EAD dans l'écosystème plus large de la coordination bibliographique à la BnF.

Les premiers travaux ont consisté à mettre les référentiels internes en adéquation avec les listes de valeurs préconisées par le Guide des bonnes pratiques. Des chantiers d'alignement des données accompagnent chaque redéploiement des référentiels consolidés. De nouvelles procédures d'encodage sont mises en place pour le signalement des documents numérisés de façon à mieux gérer la conversion des données vers Gallica, la bibliothèque numérique de la BnF et de ses partenaires⁷. La définition de ces nouvelles procédures s'accompagne du suivi de leur implémentation technique dans l'application de production, d'importants travaux de reprise rétrospective, et enfin d'un volet de formalisation des nouvelles consignes et d'accompagnement des utilisateurs. La création du poste d'administration des données EAD a enfin été l'occasion de refondre l'offre de formation EAD aux professionnels, déclinée désormais en un tronc commun et des modules spécifiques pour les manuscrits médiévaux et les fonds d'archives.

⁷ <https://gallica.bnf.fr/accueil/?mode=desktop>

3. LES TRAVAUX A VENIR

3.1 Au niveau local : la qualité des données

Force est de constater que la coordination de l'EAD à la BnF se heurte rapidement à des limites, notamment en raison de choix historiques ou plus proches de nous. En premier lieu, l'EAD et BAM n'ont été que marginalement intégrés à la réflexion globale de l'établissement autour de la qualité des données. Alors que la politique de catalogage dans le Catalogue général a fait l'objet d'une récente mise à jour, l'idée même d'une politique de catalogage ou de charte sur la qualité des données reste étrangère à l'environnement EAD de la BnF. L'articulation entre les deux catalogues est donc problématique : en théorie, on considère que les documents édités seront décrits dans le Catalogue général, tandis que les documents non édités seront décrits dans BAM. En réalité, la frontière entre les deux catalogues est bien plus poreuse, certains types de documents non édités étant décrits dans le Catalogue général, qui s'était entre-temps doté d'un format de production sur mesure avec l'INTERMARC intégré.

Une réflexion sur la qualité des données est actuellement en cours à la BnF, à laquelle l'EAD sera nécessairement incluse : c'est l'avantage de disposer d'une coordination clairement identifiée. Au cœur de cette réflexion reviendra la question de la césure entre BAM et le fichier d'autorité du Catalogue général. Si le fichier est bien commun aux deux bases (une décision prise dès les commencements de BAM et de laquelle on ne cesse de se féliciter), les mises à jour des points d'accès ne sont pas synchronisées, entraînant des disparités entre le fichier d'autorités et l'affichage dans BAM. Partant, une recherche fédérée sur les deux catalogues reste impossible à ce jour. Certes, depuis 2013 data.bnf.fr joue le rôle de pivot documentaire entre les deux bases ; mais l'algorithme de l'interface repose sur des liens à des notices d'autorité validées du fichier d'autorités. Or le poids des rétroconversions limite le nombre de ces liens, la source n'étant pas toujours donnée dans les notices. En 2018, sur les 2,4 millions de points d'accès recensés dans les Instruments de recherche, seuls 8% sont liés au fichier d'autorité de la BnF, ce qui explique que data.bnf.fr ne contienne qu'un peu plus de 70 000 liens à BAM.

Définir une politique de qualité des données permet également de mettre au jour l'intrication des choix intellectuels et des obstacles techniques. Actuellement, la mise en place d'un système de communication automatisée des documents de BAM se heurte à des obstacles dits « techniques » mais qui relèvent en réalité plus de questions d'utilisation du format : car dans un catalogue où les valeurs de LEVEL sont employées de manière hétérogène, et où des <unitid> de TYPE « cote » existent à tous les niveaux de la description, il devient difficile de déterminer où se trouve l'unité communicable. À court terme, cela aboutit à la rédaction de spécifications « à la carte ». À plus long terme, il faut dès à présent envisager de nouvelles consignes d'encodage des composants qui auront été ciblés par ces spécifications comme étant les niveaux à communiquer, ce qui ne va pas de soi dans un format comme l'EAD, conçu comme un format de description et non de gestion technique de la communication des documents.

3.2 Au niveau national : la co-production des données

Aujourd'hui comme hier, la solution aux problèmes de cohérence locale trouve à la fois un écho et une solution dans son pendant national. Suite à la dissolution du groupe des bonnes pratiques à la parution du *Guide*, la coordination nationale était pour ainsi dire non existante pendant plusieurs années. En 2017, un projet d'outil national de production, rapidement

abandonné, oblige néanmoins à se reposer la question de la coordination au niveau national, tant les questions de cohérence des données et d'interopérabilité des données restent d'actualité. Un groupe de travail national sur l'interopérabilité des données en EAD réaffirme alors la position centrale du *Guide des bonnes pratiques* ; or le site, figé depuis sa mise en ligne, ne joue pas le rôle de fédérateur de communauté escompté à ses débuts, et l'écart entre la théorie et la pratique se creuse de nouveau.

La solution est intellectuelle : on envisage en effet la refonte du *Guide des bonnes pratiques*, avec comme objectif de créer une plateforme centrale unique recensant l'intégralité de la documentation sur l'EAD pour les bibliothécaires, des informations sur les réseaux et l'offre de formation, une boîte à outils du catalogueur, etc. Cette offre visant à l'exhaustivité signifierait la disparition des documentations locales. Un important travail de mise à jour des bonnes pratiques sera en outre nécessaire, car beaucoup ne répondent plus aux besoins actuels de catalogage courant, du fait de leur ancrage historique sur les spécifications de conversion, mais aussi parce que la pratique de l'EAD en catalogage courant n'était pas encore mature à l'époque de la création du *Guide*, et qu'un certain nombre de types de documents n'avait pas été pris en compte (archives numériques natives, documents audiovisuels et multimédias, etc.).

Les évolutions techniques permettent d'appuyer ces décisions intellectuelles. Ainsi, le futur outil de production TAPIR développé par le service informatique de la BnF pour le CCFr permettra d'envisager davantage de mutualisation, notamment peut être dans le cadre d'opérations de traitement en masse des données. À cet effet, il a été décidé d'intégrer dans TAPIR, et partant dans PiXML, une fonctionnalité de validation selon les bonnes pratiques, en plus de la seule validation du format proposée par l'éditeur. Ce travail de validation est mené par l'équipe projet TAPIR et la coordination EAD à la BnF. La plupart du temps les contrôles seront non bloquants et signaleront les écarts par rapport aux bonnes pratiques, en invitant le catalogueur à les corriger.

CONCLUSION

L'expérience française montre bien l'importance de mettre en place une coordination, et d'en définir le périmètre et les responsabilités à chaque niveau : si les instructions propres à des problématiques internes relèvent de la coordination locale, la coordination nationale a vocation à instruire tout besoin commun à plusieurs établissements, et plus particulièrement tout besoin lié à l'actualité des travaux de modélisation internationale. Au-delà des interrogations techniques conjoncturelles, l'enjeu de la coordination est bien plus pérenne, bien plus ancré dans notre cœur de métier, car il en va de produire des données de qualité, dont on prend aussi soin que les documents que l'on décrit. De même que les conservateurs anticipent l'évolution des matériaux et leur potentielle fragilité face à l'épreuve du temps, les spécialistes des données envisagent leur préservation sur le long terme.

Remerciements

Les auteures de cet article remercient chaleureusement le groupe EAD pour leur précieuse expertise sans laquelle il n'y aurait de coordination que le nom. Un grand merci également à Vincent Boulet, Laurence Le Bras et Françoise Leresche pour leur mémoire sans faille sur l'historique de l'implémentation de l'EAD à la BnF, ainsi que pour leur disponibilité.

Références

Description des manuscrits et fonds d'archives modernes (DeMArch). Disponible en ligne : http://www.bnf.fr/documents/ead_demarch.pdf (consulté le 10 juin 2018)

Description archivistique encodée : Dictionnaire des balises, 2004. Disponible en ligne : https://francearchives.fr/file/0def64f5a10f3f1ae03fdea59399a3e0755ef157/static_1066.pdf (consulté le 10 juin 2018)

Guide des bonnes pratiques de l'EAD en bibliothèque. Disponible en ligne : <http://bonnespratiques-ead.net/> (consulté le 10 juin 2018)

ISAD (G) : norme générale et internationale de description archivistique, deuxième édition, 1999. Disponible en ligne : https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_FR.pdf (consulté le 10 juin 2018), ISBN 0-9696035-6-8

PALLUAULT, Florent. « Le Catalogue général des manuscrits des bibliothèques publiques de France : informatisation et avenir ». *Bulletin des bibliothèques de France (BBF)*, 2009, n°1, p. 68-72. Disponible en ligne : <http://bbf.enssib.fr/consulter/bbf-2009-01-0068-010> (consulté le 10 juin 2018). ISSN 1292-8399

PLAZANNET, Fabien. « Le Catalogue général des manuscrits des bibliothèques publiques de France ». *Bulletin des bibliothèques de France (BBF)*, 2003, n°5, p. 74-78. Disponible en ligne : <http://bbf.enssib.fr/consulter/bbf-2003-05-0074-002> (consulté le 10 juin 2018). ISSN 1292-8399