

Sommaire

Vers la révision de l'EAD

Journées d'études aux Archives nationales d'outre-mer, 2-4 juin 2010

Réunion du Comité des normes et bonnes pratiques du Conseil international des archives, 25-27 mai 2010

Séminaire du stage technique international sur la normalisation de la description archivistique, 28 mai 2010

Vers la révision de l'EAD

Les travaux de révision de l'EAD devraient débuter lors du prochain congrès de la Société des archivistes américains (SAA), les 10-15 août prochains, à Washington. En effet, les groupes de travail internationaux hébergés par le Comité des normes de la SAA ont été récemment réorganisés et les nouvelles structures devraient être mises officiellement en place en août. C'est ainsi qu'un sous-comité technique pour l'EAD (TS-EAD), co-présidé par Michael Rush (Université de Yale) et William Stocking (British Library) remplace désormais le groupe de travail EAD. Quant au groupe de travail EAC-CPF, un sous-comité technique pour l'EAC (TS-EAC) devrait prendre prochainement sa suite.

Une équipe restreinte, composée de Daniel Pitti (Université de Virginie), de Karin Bredenberg (Archives nationales de Suède), de Terry Catapano (Université Rutgers du New Jersey), de Michele Combs (Université de Syracuse, New York), de Mark Matienzo (Université de Yale) et des présidents des comités techniques EAD et EAC sera à l'avenir chargée du développement et de la maintenance technique des formats d'échange de la SAA.

Le descriptif du mandat et la liste des membres de ces différentes instances sont consultables sur le site de la SAA : <http://www2.archivists.org/> puis Groups / Subcommittees. Le mandat du nouveau sous-comité technique de l'EAD s'achèvera en août 2015.

Journées d'études aux Archives nationales d'outre-mer, 2-4 juin 2010

En juin 2004, les Archives nationales d'outre-mer rendaient accessible sur Internet une application, IREL (Instruments de recherche en ligne), permettant l'interrogation et la consultation de l'état des fonds d'archives conservés à Aix-en-Provence, ainsi que des instruments de recherche propres à certaines séries (archives notariales, archives du ministère des affaires algériennes). Six ans après, c'est une nouvelle version d'IREL qui est mise à disposition des internautes, avec de nombreux instruments de recherche détaillés encodés en XML et des outils communs d'orientation (listes et notices d'autorité) permettant de guider plus efficacement l'utilisateur (<http://anom.archivesnationales.culture.gouv.fr/>).

A l'occasion du lancement de cette nouvelle version, le Service interministériel des archives de France a souhaité organiser des journées d'études sur des retours d'expériences de mise en œuvre des formats EAD et EAC-CPF (Contexte archivistique encodé – Collectivités, personnes, familles).

Les participants ont présenté les résultats de leurs recherches et leurs idées autour de trois grands axes :

- la normalisation des descriptions, un préalable indispensable à leur structuration en XML,
- les stratégies de production et de publication,
- l'EAC-CPF, un format pour la description contextuelle.

Claire Sibille – de Grimoüard (Service interministériel des Archives de France) a commencé par retracer l'histoire des normes du Conseil international des archives (ICA), qui reflète une évolution dans les pratiques descriptives. ISAD(G) combine toutes les composantes de la description archivistique (les archives et leur contexte). Avec ISAAR(CPF), ISDF, et ISDIAH, on a évolué vers une différenciation des composantes de la description (archives, producteurs/détenteurs des archives et fonctions des producteurs/détenteurs), permettant la création et la gestion de descriptions distinctes (dans un souci d'efficacité et de flexibilité) mais reliées les unes aux autres dynamiquement pour offrir aux utilisateurs une meilleure compréhension des archives et de leur contexte. Un modèle implicite et informel sous-tend cette différenciation.

Conçus pour servir comme des normes de communication pour la description, les formats d'encodage apparentés ont connu la même évolution. L'EAD combine en effet toutes les composantes de la description archivistique (les archives et leur contexte), mais l'EAC-CPF et l'EAG (Guide archivistique encodé) permettent de créer et de gérer des descriptions séparées mais liées des producteurs/détenteurs d'archives, en attendant le développement d'un futur schéma pour structurer en XML les descriptions de fonctions. Les travaux de révision de l'EAD devraient commencer en août (*cf. supra*) pour l'aligner sur le schéma EAC-CPF et la rendre plus interoperable.

Au-delà, il s'agit de mieux répondre aux attentes des utilisateurs en leur offrant un accès homogène aux différentes ressources patrimoniales, qu'elles soient archivistiques, bibliographiques ou muséales. Dans cette optique, l'ICA envisage l'élaboration d'un modèle conceptuel de données réconciliant et harmonisant les quatre normes internationales développées depuis le début des années 1990. Il ne s'agit pas de développer un nouveau format d'encodage, mais d'envisager les relations conceptuelles entre les archives, leurs producteurs/détenteurs et les fonctions des producteurs, et de clarifier les concepts sous-tendant les normes et formats de description. Il s'agira ensuite de comparer le modèle conceptuel élaboré pour les archives avec ceux élaborés pour les musées et les bibliothèques, afin d'harmoniser les processus métiers à travers la compréhension des concepts communs qui se cachent derrière la documentation scientifique des différents métiers patrimoniaux.

Béatrice Hérold et Isabelle Aristide (Archives nationales – Pierrefitte) ont ensuite montré comment la normalisation était au cœur des préoccupations des travaux de préparation du chantier du futur Système d'information archivistique (SIA) des Archives nationales, avec la mise en place d'outils transitoires permettant de répondre à un axe fort de la politique scientifique de l'établissement : produire à terme tous les instruments de recherche en XML/EAD. Un programme (appelé Sosie) de conversion d'instruments de recherche produits sous OpenOffice en fichiers au format EAD est ainsi en cours de déploiement aux Archives nationales (Paris et Fontainebleau).

Parallèlement, les spécifications détaillées du futur outil de production des instruments de recherche sont en cours d'écriture avec le prestataire chargé de la réalisation du SIA. Les éléments et attributs retenus pour Sosie vont devenir la structure EAD de référence du SIA, mais des déclinaisons de ce modèle ont été mises au point dans le cadre d'un chantier scientifique spécifique réunissant des agents des Archives nationales et du Service interministériel des Archives de France. Un instrument de recherche pourra combiner plusieurs modèles de description, génériques ou spécialisés (pour les documents sériels ou nécessitant une description physique plus poussée). Il a été décidé d'en limiter le nombre et de privilégier l'établissement de règles de rédaction détaillées permettant de renseigner tel ou tel élément.

Dans les bibliothèques, l'EAD est utilisée aussi bien pour les ensembles structurés (fonds d'archives ou collections de documents) que pour des manuscrits isolés. Toutefois, **Florent Palluault** (Service du livre et de la lecture) a montré à travers des exemples récents de mise en œuvre du format que la souplesse d'encodage pouvait prêter à des interprétations qui peuvent aller à l'encontre d'un affichage homogène dans les catalogues collectifs et d'un échange pertinent de données, d'où l'importance de la formation. Un Guide des bonnes pratiques de l'EAD en bibliothèque devrait bientôt proposer aux bibliothécaires des recommandations pour l'encodage courant et pour la mise à jour des données issues des rétroconversions de catalogues. Des formations ont également été mises en place, par la Bibliothèque nationale de France (en interne et pour les pôles associés), par l'Enssib, l'Agence bibliographique de l'enseignement supérieur et d'autres organismes comme Médiacitain, Médiadix, mais les résultats sont inégaux. Dans les bibliothèques municipales, on constate encore une faible mise en œuvre, peu de visibilité en ligne des projets et une absence de réseau de partage des pratiques.

Pour **Bertrand Rey** (Archives départementales de Haute-Savoie), l'apprentissage de l'EAD passe par de la formation en externe et en interne et par la mise en place de nouveaux outils. Au-delà d'une présentation de l'EAD et de l'encodage en général, les Archives départementales de Haute-Savoie ont donc rédigé un mémento pour une description normalisée à usage interne. Ce document devrait être complété prochainement par un rappel des règles d'indexation. L'objectif est de procurer à tout rédacteur d'instrument de recherche des outils lui permettant de s'approprier la chaîne de production et de publication de descriptions archivistiques.

La présentation du logiciel ICA-AtoM par **Jean-François Moufflet** (Service interministériel des Archives de France) est venue clore la première demi-journée. ICA-AtoM est un projet international coopératif qui tend à développer un outil libre permettant aux archivistes de produire des descriptions archivistiques et de publier sur le Web des instruments de recherche archivistiques (voir : <http://ica-atom.org/>). Le logiciel est conçu pour être multilingue, basé sur le Web et conforme aux normes de description de l'ICA. Il est développé pour l'ICA par une compagnie canadienne, Artefactual Systems, et le projet est conduit par un Comité directeur, composé de représentants de Bibliothèque et Archives Canada, des Archives nationales d'Écosse, des Archives de France, de l'Unesco et de l'École d'archivistique des Pays-Bas. Avec ICA-AtoM, il est possible de créer, de modifier et de publier quatre types de notices descriptives (description archivistique, notices d'autorité, services d'archives, fonctions). Des exports et des imports EAD sont également possibles. Dans le cas d'imports EAD, les données saisies dans <origination> et dans <repository> sont intégrées dans la base sous forme de notices distinctes de producteurs et d'institutions de conservation. Des fonctionnalités pour l'import/export de thésaurus au format SKOS ainsi que de notices d'autorité utilisant le schéma XML EAC-CPF devraient être intégrées prochainement au logiciel. La version actuelle (beta 1.0.9) devrait migrer vers une version stable 1.1 à l'automne, mais ICA-AtoM ne sera publié officiellement qu'après la réalisation d'un audit technique sur la montée en charge du logiciel, l'écriture des codes, etc.

Après les préalables de mise en œuvre de l'EAD, le colloque s'est poursuivi avec la présentation de différentes stratégies de production et de publication. **Nathalie Vidal** et **Eric Froissart** ont ainsi présenté l'expérience des Archives départementales du Doubs, qui ont connu une informatisation tardive, avec au départ de très nombreux inventaires produits avec des outils bureautiques mais très bien structurés, des agents peu nombreux mais bien formés à l'archivistique et à l'EAD. Le cahier des charges pour l'informatisation du service a été axé sur l'EAD, dans une perspective d'import de documents réalisés avec un logiciel non spécialisé. Le logiciel Mnesys, acquis en 2008 par le service, admet trois modes de production EAD : création native, permettant une lisibilité immédiate de la structure de l'inventaire, conversion rétrospective de documents bureautiques, très adaptée pour les agents ayant peu de formation archivistique ou informatique, et intégration de fichiers XML. La mise en œuvre de l'EAD a modifié les pratiques, avec notamment une plus grande attention portée à la question des niveaux et à la séparation des informations entre intitulé et présentation du contenu ainsi qu'un retour systématique à la cotation à l'article.

Tout autre est le contexte des Archives départementales du Var, qui sont informatisées depuis 1984, comme l'ont rappelé **Pascale Bugat** et **Benoît Jegouzo**. De 1984 à 2010, le service a fait le choix d'évoluer avec les nouvelles technologies, le logiciel utilisé passant du MS-DOS monoposte au réseau Windows puis au Web. A la première base de données consacrée aux archives contemporaines ont succédé d'autres bases documentaires ou de gestion (documents iconographiques, notaires, état civil, bibliothèque, presse, archives communales ou hospitalières), en cours de migration. Un premier projet concerne la conversion de 1400 unités de description correspondant à des sous-fonds d'archives contemporaines. Un deuxième exemple concerne la reprise de la base de données du fonds iconographique, qui doit répondre à la fois aux besoins de consultation sur Internet et à la contextualisation des fonds constitués.

L'EAD peut être mise en relation avec d'autres formats. **Florence Clavaud** (École nationale des chartes) a ainsi décrit un projet d'humanités numériques, celui des chartes médiévales de l'abbaye de Saint-Denis, conduit par l'École nationale des chartes et permettant de numériser, d'encoder, de mettre en relation et de publier sur le Web inventaires archivistiques et éditions critiques des

documents décrits. Cette présentation a insisté sur les choix qui ont été faits en matière de modélisation : utilisation du format EAD pour l'inventaire, définition d'un modèle spécifique conforme à TEI P5 pour l'édition critique du cartulaire et utilisation de fichiers XML/METS pour consigner les liens entre composants du corpus. Ce corpus devrait être publié au sein d'une application Web unique dotée de moyens de recherche avancée et offrant des moyens de navigation de l'un à l'autre des composants. Pour en savoir plus : <http://saint-denis.enc.sorbonne.fr/>.

Pour **Pascale Heurtel** (Bibliothèque centrale du Muséum national d'histoire naturelle) et **Enrico Cima** (Agence bibliographique de l'enseignement supérieur), présenter Calames (Catalogue accessible en Ligne des archives et manuscrits de l'enseignement supérieur), ce n'est pas seulement présenter une réalisation technique qui a été ouverte au public dans le premier semestre de 2008, c'est présenter à la fois un catalogue commun qui offre au public plusieurs façons d'explorer les richesses patrimoniales des bibliothèques participantes et un outil de catalogage partagé à distance. La multiplication des intervenants dans un système très ouvert et une DTD très permissive ont rendu nécessaire la création d'un véritable réseau d'entraide, piloté par l'Agence bibliographique de l'enseignement supérieur (ABES) et soutenu par un groupe de travail, mais aussi la mise au point de manuels de catalogage, de listes de discussion et de listes fermées pour faciliter la saisie des attributs et améliorer ainsi la recherche. Plus récemment, un programme XSL a été développé, permettant à chaque contributeur de tester la conformité de son travail avec les règles communes.

Valérie Brousselle (Archives départementales du Val-de-Marne) a souligné les enjeux de la mise en ligne, par les services territoriaux, de leurs inventaires et des images numériques des documents décrits, malgré la baisse de fréquentation des salles de lecture (on est passé de 5000 séances en 2004 à 3230 en 2008 aux Archives départementales du Val-de-Marne) : visibilité par rapport aux autorités de tutelle, valorisation des fonds conservés et enjeux de management des équipes. Le système d'information des Archives départementales du Val-de-Marne était éclaté entre des outils bureautiques et quatre bases de données. A l'hétérogénéité des formats de description s'ajoutait celle du niveau de détail des descriptions (au niveau des versements pour la série W, au niveau du fonds pour les fonds audiovisuels, jusqu'au niveau des articles voire des pièces pour d'autres ressources), ainsi que l'absence d'expertise avancée ou partagée en EAD, de sorte que le service est très lié aux prestataires. La première version du portail documentaire des Archives départementales du Val-de-Marne date de 2008, une nouvelle mise à jour est prévue cet été. Les modes de production ont eu des répercussions fortes sur la mise en ligne. Il a fallu notamment relier des fiches descriptives de groupes d'articles aux notices descriptives de fonds. Des séances pratiques sont organisées pour former les lecteurs à la consultation du portail documentaire, d'autant que l'on s'est aperçu à l'usage que les internautes passaient davantage par la liste des fonds que par le formulaire de recherche avancée et la recherche guidée via le thésaurus pour l'indexation des archives locales et le thésaurus Motbis.

Mettre en oeuvre l'EAD, c'est aussi permettre le moissonnage de métadonnées normalisées par des portails. **Frédéric Rolland** (MCC, sous-direction des systèmes d'information) et **Lucile Grand** (Service interministériel des Archives de France) ont commencé par rappeler les différents niveaux d'interopérabilité définis par la direction générale pour la modernisation de l'Etat : organisationnel (définir les rôles et responsabilités, mettre en place les ressources), sémantique (s'entendre sur la signification des données qu'on échange), syntaxique (s'entendre sur la forme) et technique (véhiculer des informations). La mise en place d'un protocole commun basé sur les standards du Web, l'OAI-PMH (Open Archive Initiative - Protocol for Metadata Harvesting), facilitant l'accès au contenu par un ensemble de services extérieurs, permet l'ouverture de l'architecture technique des entrepôts d'informations contenant des documents numériques.

Les apports de la mise en oeuvre de l'OAI-PMH dans le cadre de projets de fédération de bases du ministère de la Culture et de la Communication (portails Généalogie, Collections) sont multiples : facilité de mise en oeuvre, standardisation de la relation avec les bases sources, réduction des coûts d'exploitation et de maintenance. Ainsi, le portail Collections sera agrégateur de contenus pour Europeana, portail européen multi-domaine sur le patrimoine culturel, donnant accès à des objets numérisés et à leur description (archives, musées, bibliothèques et institutions audiovisuelles). Le portail européen des archives (APEnet) sera lui-même l'intégrateur dans le domaine archives pour Europeana. Tous les partenaires d'APEnet devront fournir leurs données sous la forme de fichiers EAD XML, quel que soit le format original de leurs données. Un schéma "européen" APEnet EAD est

en cours d'élaboration, afin de permettre un affichage correct et homogénéisé des données, de pouvoir chercher correctement dans les données, de pouvoir aligner les métadonnées et développer des fonctionnalités précises, et d'alimenter Europeana par OAI-PMH.

Le 5 mars dernier, le Groupe de travail international pour l'élaboration de l'EAC annonçait la publication du schéma EAC-CPF (Contexte archivistique encodé – Collectivités, personnes, familles) sur son site officiel. L'EAC-CPF permet de créer et de gérer la description de producteurs d'archives de manière autonome, en la séparant de la description des documents d'archives en EAD. **Anila Angjeli** (Bibliothèque nationale de France) a retracé l'historique et les objectifs de ce projet international collaboratif entre archives et bibliothèques. L'accent a ensuite été mis sur les nouveautés par rapport à la version beta de 2004 et aux atouts du nouveau schéma : clarté, simplicité d'implémentation et d'utilisation, capacité de répondre aux besoins d'encodage des notices d'autorité des archives comme des bibliothèques, intégration de technologies plus récentes (espaces de noms, identifiants, mécanismes de liens) pour la mise en relation des informations et le dialogue avec d'autres applications.

Annabel Montécot (Archives départementales du Morbihan) a ensuite présenté une expérience d'intégration de notices d'autorité de collectivités dans un moteur de recherche, à travers l'exemple du guide des sources de la Seconde Guerre mondiale et de la reconstruction dans le département du Morbihan. En parallèle de l'exposition "Le Morbihan en guerre" inaugurée en octobre 2009, un important travail de recensement des sources produites pendant cette période a été réalisé. Ce recensement a été l'occasion d'élaborer un guide des sources en XML-EAD et de rédiger des notices d'autorité de collectivités au format XML-EAC (version beta de 2004), offrant un aperçu assez précis de l'histoire, notamment administrative, du département pendant cette période troublée. La rédaction de ce guide en EAD et de notices d'autorité en EAC a permis d'approfondir les connaissances du service en XML, notamment par l'utilisation d'un logiciel d'encodage, ATES, basé sur l'éditeur XMLMind. Elle a été également l'occasion de réfléchir à une bonne répartition des informations et à la génération de liens entre les différents fichiers XML, les notices d'autorité n'étant pas initialement prévues dans le moteur de recherche. C'est ainsi qu'une distinction a été opérée entre les organismes du temps de guerre (notices au format PDF liées aux fichiers EAD avec <corpname> et l'attribut @authfilenumber) et ceux du temps de paix (pas de notices spécifiques, informations insérées dans l'élément <bioghist> de l'EAD). La rédaction du guide a permis d'élargir la réflexion à la répartition des informations et à la génération de liens entre l'état des versements de la série W, les instruments de recherche et les notices d'autorité.

La structuration en XML-EAC de notices d'autorités de producteurs est également un sujet d'actualité pour les Archives nationales, dans le contexte de l'élaboration du futur Système d'information archivistique. **Marie-Françoise Limon-Bonnet** et **Michel Thibault** (Archives nationales) ont montré comment la construction d'un référentiel "Producteurs" normalisé répond à deux besoins fondamentaux : identifier des fonds cotés selon des systèmes de cotation différents à Paris et à Fontainebleau, et relier entre eux les fonds de producteurs entretenant des relations hiérarchiques, chronologiques, familiales ou d'association. La construction de ce référentiel partira d'outils déjà disponibles, à savoir la nomenclature N1, créée à la fin des années 1970 pour décrire les producteurs effectuant des versements sur le site de Fontainebleau, et la base ETANOT, en ligne depuis 2005, identifiant chaque notaire parisien du XV^e siècle à nos jours. La migration de ces deux outils vers le référentiel Producteurs soulève des questions spécifiques : non conformité du N1 avec la norme ISAAR(CPF), définition de la forme autorisée du nom pour les producteurs publics et celle de l'élément Lieu, essentiel pour l'identification des notaires parisiens, choix de ne pas adopter de code signifiant pour retrouver un producteur (contrairement au code N1, permettant une identification rapide du ministère d'appartenance, mais se révélant trop complexe à gérer au fur et à mesure des évolutions d'organigrammes).

Les journées d'études se sont achevées avec la présentation, par **Martine Cornède**, **Jean-Luc Vannier**, **Anne-Isabelle Vidal**, **Eléonore Bozzi** (Archives nationales d'outre-mer) et **Nader Boutros** (Pass-Technologies) des principes qui ont présidé aux évolutions récentes de de l'application IREL, à savoir la production de notices d'autorité et de référentiels pour guider plus efficacement les utilisateur dans leurs recherches. Les référentiels ont été construits a posteriori, à partir des termes utilisés dans les instruments de recherche et les index existants. Les mots-matières proviennent donc du vocabulaire utilisé dans les analyses et ont été organisés par domaines, dans différentes listes d'autorité hiérarchisées. En ce qui concerne les personnes et les collectivités, seuls les acteurs

importants de l'histoire coloniale ont fait l'objet d'une notice d'autorité, qu'ils soient producteurs de fonds ou qu'il en soit simplement question dans les documents d'archives. Dans les fichiers EAD, les mots-matières et les noms de personnes, de collectivités et de lieux géographiques sont systématiquement encapsulés dans des balises d'indexation <subject>, <persname>, <corpname> et <geogname>. Un outil de saisie de notices EAC (version beta, mais les équivalences avec l'EAC-CPF sont en cours d'étude) a été développé en interne, ainsi que des outils de contrôle et des guides de bonnes pratiques pour la construction des formes de nom personnes, collectivités ou lieux géographiques. L'accès aux notices d'autorité se fait à partir des instruments de recherche ou de la base de données nominative.

Pour en savoir plus: se reporter aux supports de présentation accessibles sur le site des Archives de France (<http://www.archivesdefrance.culture.gouv.fr> Gérer les archives / Classement et description / Colloques et journées d'études).

Réunion du Comité des normes et bonnes pratiques du Conseil international des archives, 25-27 mai 2010

Les 25-27 mai derniers, les Archives de France accueillait la réunion annuelle du groupe de travail du Comité des normes et bonnes pratiques (CBPS) sur la description archivistique.

Lors de sa dernière réunion, à Marbourg (Allemagne), les 25-27 mai 2009, le groupe de travail avait décidé de développer un recueil de normes de description se concentrant sur les relations entre les différents objets archivistiques à décrire (archives, producteurs/détenteurs d'archives et fonctions des producteurs). L'élaboration de ce document reste la priorité du sous-comité pour le mandat actuel 2008-2012, et le recueil devrait être publié pour le prochain Congrès international des archives à Brisbane (Australie), en 2012. Son objectif est de favoriser un usage cohérent des normes et d'expliquer à la communauté comment les quatre normes de l'ICA s'articulent entre elles comme un unique ensemble de normes permettant de produire des descriptions archivistiques séparées mais liées. Des changements substantiels dans les normes interviendront pendant le mandat 2012-2014, ainsi que l'élaboration d'un modèle conceptuel pour la description archivistique. Par ailleurs, le recueil de normes pourrait servir de point de départ aux travaux du groupe de la SAA chargé de réviser l'EAD d'ici 2015.

Les objectifs de la description et des systèmes d'information archivistiques seront rappelés au début du recueil de normes :

- nature des archives (elles ont des caractéristiques uniques, elles sont créées, rassemblées, reçues dans le cadre de l'exercice d'activités, etc., elles sont conservées à titre de preuve, etc.),
- objectifs de la description archivistique (elle documente l'historique de la conservation des documents pour établir la preuve de leur authenticité, elle aide les utilisateurs à localiser et à comprendre la signification des archives comme ressources pour la recherche, etc.),
- nature de la description archivistique (elle organise les documents en ensembles parce qu'ils ont été créés et gérés de cette manière et parce que on ne peut les comprendre qu'en relation avec d'autres documents, elle reflète le classement à plusieurs niveaux, les archivistes décrivent à la fois le contenu et le contexte des archives, les mêmes principes s'appliquent aux archives produites par des personnes, des collectivités et des familles, etc.),
- bénéfices de la description archivistique (elle permet aux services d'archives de présenter leurs collections d'une manière cohérente, elle permet l'échange d'informations descriptives entre systèmes d'information, elle encourage les bonnes pratiques par les archives, etc.),
- nature des systèmes de description archivistique (qu'ils soient manuels ou automatisés, ils permettent aux utilisateurs de localiser, d'identifier, d'accéder et d'utiliser des documents pertinents, les quatre normes de l'ICA décrivent un système d'information dans lequel les différents aspects de la description sont créés et gérés dans des systèmes séparés qui sont reliés entre eux d'une manière ou d'une autre).

Le groupe a également commencé à modéliser les relations entre objets archivistiques à travers un diagramme inter-reliant les entités (archives, acteurs et fonctions) par une action. Par exemple, un acteur exécute l'action d'assigner une fonction à un autre acteur (collectivité, personne, famille) ; un détenteur d'archives, catégorie spécifique d'acteur, a un mandat pour détenir des archives / exécute

l'action de détenir des archives ; les archives sont la preuve d'actions exécutées par des acteurs ; les archives peuvent concerner des acteurs ; les archives peuvent concerner d'autres archives, etc.

Après la publication officielle du recueil de normes, le développement d'un modèle conceptuel pour les archives pourrait être envisagé pour le prochain mandat 2012-2016.

Un modèle conceptuel n'est pas une « cartographie de concepts », c'est un outil, un langage contrôlé pour décrire les états possibles d'objets dans un domaine donné, il est formulé comme une « ontologie formelle ».

Le terme d' « Ontologie formelle » est un terme utilisé en informatique. Le système informatique a la capacité de réagir de façon intelligente. C'est un moyen d'exprimer des conceptualisations du monde réel de manière non équivoque et stable dans le temps, et de discuter objectivement des implications de ces conceptualisations.

Les modèles conceptuels ne sont pas intuitifs, mais analytiques. Ils sont constitués d'entités ou de « classes ». Les relations entre les classes sont appelées « propriétés ». Les propriétés sont définies (ou « déclarées ») comme s'appliquant entre un domaine et une série, elles lient une instance de la classe qui fait office de domaine (ou source du lien) avec une instance de la classe qui sert de cible (vers laquelle pointe le lien). Les attributs sont également modélisés en utilisant les propriétés. Les classes et les propriétés s'inscrivent dans une hiérarchie, quand une classe est déclarée comme sous-classe d'une autre classe, alors chaque instance de la sous-classe est également une instance valide de la super-classe, et hérite (au sens strict) les propriétés de cette super-classe.

Un modèle n'est pas un format, il explique les formats existants. Un modèle explique les données existantes, il peut les transformer sous n'importe quelle forme, il peut faire un lien de cause à effet entre toutes ces données, il peut rendre les formats plus compréhensibles. Ainsi, avant la généralisation et l'intégration, il est nécessaire de comprendre comment les concepts actuels reflètent des fonctionnalités et des processus. La terminologie peut être un obstacle à la compréhension (par exemple, est-ce que les mandats sont des fonctions de haut niveau ?). Il faut différencier la nature et la désignation des choses (par exemple, la zone d'identification de la norme ISAAR (CPF) comprend des éléments précisant le type de l'entité décrite et d'autres contenant les différentes formes de noms de l'entité).

Pour résumer, le but est de créer et de partager une compréhension commune de la structure, du contenu et des utilisations prévues des normes et des formats. Il vise à permettre de mieux comprendre certains domaines de la connaissance, il permet à d'autres personnes de comprendre certains domaines de la connaissance, il permet de rendre explicites les concepts les sous-tendant. A noter : le monde des données utilisées par les archivistes dans leurs descriptions est plus vaste que le monde représenté dans les normes : les quatre normes de l'ICA permettent de décrire des documents d'archives, des collectivités, des personnes physiques, des familles ou des fonctions, mais l'élément « Présentation du contenu » d'ISAD(G) mentionne aussi les lieux et les sujets dont il est question dans les documents d'archives.

Séminaire du stage technique international sur la normalisation de la description archivistique, 28 mai 2010

La réunion du Comité des normes et bonnes pratiques de l'ICA a été suivie d'un séminaire sur la normalisation de la description archivistique, dans le cadre du stage technique international des archives organisé par le département de la formation scientifique et professionnelle de la direction générale des patrimoines. Plus de 40 archivistes d'Amérique du Nord, d'Europe centrale et de l'Est, du Maghreb et d'Afrique noire y participaient. Cette manifestation a été l'occasion d'échanges entre les concepteurs et les utilisateurs des normes.

Daniel Pitti (Université de Virginie) a commencé par dresser une typologie des normes de description, opérant une distinction entre normes de contenu intellectuel et normes de communication. Les normes de contenu peuvent elles-mêmes soit des normes de structuration de données (schémas ou cadres conceptuels élaborés au niveau international) soit des normes de contenu détaillées, plutôt locales qu'internationales et s'appuyant sur les normes de structuration internationales. Quant aux

normes de communication, elles permettent de représenter le contenu intellectuel sous une forme lisible par machine.

Les modèles conceptuels de description des ressources représentent une nouvelle tendance dans le développement de normes de structuration. Ils s'appuient sur des méthodes formelles d'identification et d'interrelation des composantes de la description (archives, producteurs/détenteurs, fonctions des producteurs) et ils sont souvent appelés « ontologies » (*cf. supra*). Les partisans de la modélisation conceptuelle envisagent l'utilisation des données de description comme support des applications du Web sémantique. Ils considèrent également les modèles conceptuels comme un moyen d'intégrer l'accès aux ressources du patrimoine culturel. Quoi qu'il en soit, cette méthodologie constitue un moyen très efficace de comprendre la description des ressources de manière plus formelle et logique.

Les normes archivistiques de contenu et de communication s'appuient donc sur un modèle conceptuel implicite et informel. Ce modèle implicite sous-tend la séparation et l'interrelation des composantes de la description. L'ICA s'efforce à présent de rendre explicite et formel ce modèle implicite et informel en utilisant les méthodes de modélisation conceptuelle. Un modèle explicite et exprimé de manière formelle pour la description archivistique permettra d'améliorer la qualité des normes de description de l'ICA et des formats apparentés.

En conclusion, les normes « métiers » incarnent les principes et facilitent les pratiques des communautés professionnelles. Elles ne sont pas figées mais en continuelle évolution. Elles sont des « champs de négociation », où des professionnels définissent les principes et les pratiques qui définiront à leur tour les pratiques professionnelles.

Il ne suffit pas de participer aux instances de normalisation internationales, encore faut-il adapter les normes aux pratiques locales pour se les approprier. C'est ce qu'a montré **Vitor da Fonseca** (Archives nationales du Brésil), à travers l'exemple de la norme NOBRADE, adaptation d'ISAD(G) aux pratiques brésiliennes. La norme ISAD(G) peut en effet être utilisée telle quelle dans les pays ne disposant pas de normes, ou servir de base pour l'élaboration de normes nationales comme cela a été le cas au Brésil.

Le Brésil participe depuis 1996 aux travaux de normalisation de la description archivistique au sein de l'ICA. Une commission technique nationale a été créée en 2001 pour concevoir une norme nationale à partir d'ISAD(G). NOBRADE présente plusieurs nouveautés par rapport à la norme internationale : ajout d'une zone de points d'accès et d'indexation des matières et d'un élément de description « Notes sur la conservation », division de l'élément « Dates » en deux éléments « Date(s) de lieu(x) » et « Date(s) chronologique(s) », caractère obligatoire de l'élément « Conditions d'accès », et insertion d'un schéma des niveaux de description en annexe, avec un niveau « zéro » (description des fonds et collections d'une institution de conservation).

De nouvelles normes ou recommandations sont en cours d'élaboration, pour l'attribution de codes d'identification aux services d'archives, la construction de formes de noms de collectivités et l'élaboration d'instruments de recherche.

Enfin, **Stefano Vitali** (Surintendance archivistique d'Emilie-Romagne) a présenté un exemple concret d'implémentation des normes : le portail SAN (Système archivistique national), qui doit devenir à terme un portail commun d'accès aux ressources Web des archives italiennes. Le portail devrait comprendre un glossaire archivistique, une base de données de ressources bibliographiques concernant les archives italiennes, des guides de recherche, des tutoriels et d'autres ressources pour les différents publics (enseignants et étudiants, généalogistes, historiens, etc.), des expositions et des publications numériques, il devrait également utiliser les technologies du Web 2.0. Le portail comprendra aussi des contenus fournis par les institutions et les organismes participant au projet : des archives numériques donnant accès aux reproductions de documents d'archives, et un catalogue des ressources archivistiques italiennes (CAT), sorte d'état collectif des fonds et sous-fonds conservés par les services d'archives italiens et par d'autres détenteurs d'archives, publics et privés. Un moteur de recherche sera également développé pour permettre des recherches simultanées dans les bases données archivistiques en ligne et l'interrogation de services Web spécifiques.

Ces dix dernières années, de nombreux systèmes d'information ou des bases de données, pour décrire et publier sur le web des descriptions archivistiques, ont été développés par la Direction

générale des archives (SIAS, SIUSA, Guida generale), des archives d'Etat locales, comme les Archives d'Etat de Florence (SIASFI), de Milan (Guida on line), de Bologne, des régions (comme l'Émilie-Romagne, la Lombardie, le Piémont, l'Ombrie) et d'autres collectivités locales, diverses institutions culturelles (Archivi del Novecento, Archos). Ces systèmes partagent une architecture commune et des formats de données très similaires, puisqu'ils s'appuient en général sur les normes de description internationales ISDIAH, ISAD (G) et ISAAR (CPF). Toutefois, pour des raisons "politiques", institutionnelles et techniques, il n'était pas possible d'unifier ces systèmes dans un seul système. Il s'agissait donc d'assurer leur interopérabilité et la base de données CAT a été conçue comme un point commun d'accès à ces systèmes, qui resteront séparés et autonomes.

Les entités décrites dans le portail CAT seront : les détenteurs d'archives, les « ensembles archivistiques » (groupes de fonds, fonds et séries organiques), les instruments de recherche, les producteurs d'archives, et un fichier d'autorités « producteurs ».

Les détenteurs d'archives ne sont pas seulement des services d'archives mais aussi des organismes privés, des personnes et des familles qui possèdent ou détiennent des documents d'archives. Chaque détenteur fera l'objet d'une notice descriptive conforme aux éléments essentiels de la norme ISDIAH. Les données d'identification et les informations descriptives seront fournies par les systèmes qui participeront au projet. Elles seront normalisées afin de produire une liste d'autorités des détenteurs d'archives.

Les descriptions des ensembles archivistiques, des instruments de recherche et des producteurs seront importées depuis les systèmes archivistiques existants. Seuls les éléments de description indiqués comme essentiels ou obligatoires dans les normes de l'ICA seront importés. Plusieurs descriptions de la même entité peuvent être importées si l'entité a été décrite dans deux systèmes ou plus. Chaque notice importée sera liée à une brève description du système à partir duquel la notice a été importée afin d'identifier sa provenance. La notice comprendra aussi un lien direct vers la description dans le système à partir duquel la notice a été importée. En suivant le lien, l'utilisateur accédera à la description complète de l'ensemble archivistique, du producteur ou à l'instrument de recherche complet – s'il existe – dans son contexte original.

Le CAT sera alimenté par des données exportées des systèmes participant au projet ; ces données seront importées dans la base de données grâce à différentes procédures d'export et d'import, en fonction de la capacité technique et des outils des fournisseurs de données. Ces procédures comprendront le moissonnage de données conformément au protocole OAI-PMH, l'envoi par mail ou le téléchargement de fichiers XML et aussi une interface Web pour intégrer les données dans la base de données CAT.

Pour exporter les descriptions depuis les systèmes existants vers la base CAT, un format d'échange XML a été développé. Il s'appuie sur trois schémas, chacun incluant un ensemble d'éléments de l'EAD (pour les ensembles archivistiques et les métadonnées d'instruments de recherche) ou de l'EAC-CPF (pour les producteurs). Un quatrième schéma ad hoc a également été développé pour importer les informations essentielles sur les détenteurs à partir des systèmes qui sont en mesure de fournir ces informations.

L'architecture générale du portail a été réalisée et a été présentée à la 11e Conférence archivistique nationale qui s'est tenue à Bologne les 19-21 novembre 2009. La structure de la base de données CAT a été réalisée et les procédures d'import ont été testées avec l'import de fichiers XML à partir des trois principaux systèmes d'information archivistiques nationaux et à partir de trois systèmes régionaux. La prochaine étape sera la mise en œuvre des procédures de moissonnage de données, de contrôle et d'intégration de données. Le CAT devrait être accessible en ligne en octobre 2010. Les archives numériques devraient ensuite être développées en adoptant la même approche que pour le CAT : seules les métadonnées essentielles seront importées, avec un lien vers l'objet numérique dans le système à partir duquel les métadonnées ont été importées.

Merci de faire parvenir toute information ou tout article pour le prochain numéro à : Claire Sibille –de Grimoüard, Direction générale des patrimoines, Service interministériel des Archives de France