

Référentiels, données d'autorité, thésaurus, ontologies... Pour en savoir plus !

Journée d'études BnF / AFNOR CG46 – 28 mars 2008

Éléments pour la table ronde « Sémantique et interopérabilité »

Par Claire Sibille, Direction des Archives de France et membre du Comité des normes et bonnes pratiques du Conseil international des archives

1*) Les référentiels existants dans les archives ont-ils un rôle à jouer dans le web sémantique ?

Chez les professionnels des archives, la prise de conscience que des métadonnées standardisées sont un maillon essentiel de l'interopérabilité et sa gestion est relativement récente. *L'Énoncé des principes relatifs à la description des documents d'archives*, adopté au Congrès international des archives de Montréal en 1992, identifiait ainsi les objectifs des normes de description archivistique :

- assurer la création de descriptions uniformes, pertinentes et explicites
- faciliter le repérage et la communication d'informations sur les documents d'archives
- rendre possible l'échange de fichiers d'autorité
- rendre possible l'intégration de descriptions provenant de plusieurs dépôts dans un système d'information unifié

Différents référentiels ont ainsi été élaborés :

- comme aide à la saisie et à l'homogénéisation des méthodes de description ;
- pour structurer et mettre en commun les informations provenant de plusieurs services d'archives ;
- comme aide à la recherche pour les utilisateurs, qui disposent ainsi de listes de valeurs utilisables pour formuler leurs requêtes.

a) Référentiels structurant les descriptions

A côté du jeu d'éléments Dublin Core, central pour l'interopérabilité, des modèles de documents répondant aux besoins particuliers des archivistes ont été développés. Portés sur le XML, ces modèles décrivent la structure, les éléments et attributs, voire les valeurs autorisées de ces attributs. La DTD a été le modèle utilisé au départ. Ainsi, la DTD EAD (Description archivistique encodée) a été élaborée il y a une dizaine d'années par la Société des archivistes américains. L'EAD formalise en XML la Norme générale et internationale de description archivistique du Conseil international des archives ou norme ISAD(G), parue en 1994 et révisée en 1999. Les principes de la description archivistique ont été appliqués dans l'EAD : respect des fonds et du contexte de production, grand nombre de niveaux hiérarchiques, description allant du général au particulier avec héritage des niveaux supérieurs sans redondance entre eux.

Les données d'autorité peuvent quant à elles être implémentées dans le format EAC (Contexte archivistique encodé), formalisé un peu plus tard, également en XML, et qui s'appuie sur une autre norme du Conseil international des archives, la Norme internationale utilisée par les archives pour les notices d'autorité Collectivités, personnes et familles ou ISAAR(CPF), pour la description des producteurs, entités (collectivités, personnes ou familles) ayant produit, reçu ou conservé des archives dans l'exercice de leurs fonctions. Toutefois, tandis que l'EAD a déjà connu deux versions « officielles » en 1998 et 2002, l'EAC n'en est encore qu'à sa version beta. Un groupe international rattaché au comité des normes de la Société des archivistes américains travaille à son évolution vers une version stable 1.0.

Si l'EAD et l'EAC apparaissent aujourd'hui comme les meilleurs outils pour appliquer concrètement les principes d'ISAD(G) et d'ISAAR(CPF), elles ne remettent absolument pas en cause les deux normes. Les normes du Conseil international des archives ont offert l'instrument nécessaire à la conception des nouveaux outils informatiques. Elles sont un préalable indispensable à la mise en œuvre de l'EAD et de l'EAC. L'informatique n'est qu'un outil, cet outil ne pouvant être mis en œuvre que sur des instruments de recherche ou des notices d'autorité rédigés de manière homogène et structurée.

Par ailleurs, ISAD(G) et ISAAR(CPF) sont des ensembles de données indépendants des formats de sortie, alors que l'EAD et l'EAC ont été conçues comme des supports pour la publication en ligne des instruments de recherche et des notices d'autorité des producteurs d'archives. Les fonctions de support pour une publication électronique (points d'accès et composants permettant de relier les différents niveaux de description les uns aux autres pour l'EAD, éléments et attributs permettant de relier des documents EAD et EAC, etc.) ne sont pas définies dans ISAD(G) et dans ISAAR(CPF). L'EAD et l'EAC comprennent beaucoup plus d'éléments que les normes ISAD(G) et ISAAR(CPF), afin de permettre une manipulation informatique des données. Par exemple, à l'élément « Référence » d'ISAD(G) correspondent trois éléments EAD, afin que la recherche puisse s'effectuer séparément sur le code du pays, le code de l'institution de conservation et les cotes des documents.

Parallèlement, après la norme ISO 15489 sur le Records management, le TC11/SC46 de l'ISO a poursuivi ses travaux avec la publication, en janvier 2006, de la norme ISO 23081, première partie, relative aux processus du records management et plus spécifiquement aux métadonnées pour les « records », c'est-à-dire aux « données décrivant le contexte, le contenu et la structure des documents ainsi que leur gestion dans le temps ».

Les normes ISAD(G) et ISAAR(CPF) ont un recouvrement important avec les métadonnées du records management, même si les métadonnées pour le records management sont généralement plus larges et peuvent inclure d'autres éléments tels que des métadonnées relatives aux processus de records management. En effet, sur les 6 éléments définis comme essentiels par ISAD(G), 5 se retrouvent dans l'ISO 15489 (référence ; intitulé/analyse ; dates ; importance matérielle et support ; nom du producteur). Seul le niveau de description échappe à ce parallèle. Inversement, les quatre métadonnées minimales de la norme ISO (identifiant ; horodatage ; titre ou description ; auteur, expéditeur ou destinataire) trouvent leur corollaire dans les éléments essentiels de la norme ISAD(G). Concernant les entités identifiées par l'ISO 15489 et l'ISO 23081, on constate là aussi un recouvrement partiel par rapport aux normes de ICA : les « records » (« documents d'archives » dans ISAD(G)), les « acteurs » (« entités », collectivités, personnes ou familles, impliquées dans la production des archives dans ISAAR(CPF)), les « missions » et les « activités » (« fonctions » des organismes et individus agissant comme des personnes morales dans la future norme ISDF).

b) Référentiels pour l'homogénéisation des contenus

Les normes ISAD(G) et ISAAR(CPF), l'EAD et l'EAC n'indiquent pas de règles d'écriture. C'est à dessein que ces outils, n'entrent pas dans les détails de présentation. De portée internationale, ils ne peuvent traiter qu'une partie des conditions nécessaires à l'échange des descriptions. Ce sont des outils de structuration de l'information, qui renvoient à des normes de contenu ou à des référentiels, élaborés à un niveau national.

Aussi la direction des Archives de France a-t-elle diffusé dans le réseau des services d'archives publics des règles de ponctuation de l'analyse archivistique venant préciser l'utilisation de l'élément Intitulé/analyse (3.1.2) d'ISAD(G). Pour « dénommer l'unité de description », il s'agit soit de reprendre le titre de l'unité de description quand il existe, soit de rédiger une « analyse », selon le modèle élaboré par Christiane et Gérard Naud pour les archives administratives contemporaines dans leur article fondateur publié en 1981. Ce modèle d'analyse (objet des documents, action exercée sur l'objet, typologie documentaire, éventuellement même agent de l'action s'il est différent du producteur) peut s'appliquer à tous les niveaux de description, dans l'esprit de la norme ISAD(G).

On assiste également depuis peu au développement de normes « hybrides » qui portent à la fois sur la structuration et le contenu des informations. Ainsi, la norme américaine DACS (*Describing Archives: a Content Standard*) a un objectif beaucoup plus large que la description archivistique, qui n'apparaît que dans la première partie de la norme. Pour permettre une meilleure compréhension des documents d'archives par les utilisateurs, il est nécessaire d'en décrire le contexte de production. La deuxième partie de DACS est donc très inspirée d'ISAAR(CPF) et traite de l'identification des producteurs et des notices d'autorité pour les archives. La troisième partie donne des règles de création des formes d'autorité de noms de personnes physiques, de familles, de collectivités et aussi de noms de lieux. Cette dernière partie est très inspirée des Règles anglo-américaines de catalogage (AACR2), avec quelques adaptations au monde des archives. DACS mêle donc structure et contenu et elle s'applique aussi bien à la description archivistique et à la description contextuelle.

D'autres référentiels ont été développés pour la construction de points d'accès normalisés et la constitution de fichiers d'autorité.

Par exemple, au Royaume-Uni, le Conseil national des archives a élaboré des *Règles nationales pour la construction de noms de personnes, d'institution et de lieux*. Un portail national, le National Register of Archives, donnant accès aux fonds conservés dans les services d'archives britanniques, met à disposition de l'utilisateur des notices d'autorité décrivant des institutions, des personnes ou des familles reliées aux descriptions de fonds ou de parties de fonds. Ces notices d'autorité, très réduites (elles se limitent en général à l'identifiant de la notice, à la forme autorisée et aux autres formes de noms éventuelles et aux dates d'existence) sont reliées pour certaines d'entre elles à des notices plus complètes du dictionnaire biographique d'Oxford.

Dans les archives françaises, il existe également quelques référentiels d'indexation. Toutefois, thésaurus et listes d'autorités de descripteurs-matières font défaut pour les archives anciennes et modernes. D'où par exemple le travail en cours aux Archives nationales d'outre-mer et aux Archives nationales (Paris) pour l'élaboration de listes et de notices d'autorité communes à partir du vocabulaire utilisé dans les inventaires détaillés de ces deux services. Les normes AFNOR NF Z 44-060 et 44-061, utilisées dans les bibliothèques pour la constitution de fichiers d'autorité, pourraient également jouer le rôle de normes de contenu pour aider à renseigner l'élément Forme(s) autorisée(s) du nom (5.1.2) de la zone d'identification de la norme ISAAR(CPF).

Quelques exemples de référentiels :

- Fichiers d'autorité

Le site des Archives nationales offre aux chercheurs un instrument de recherche leur permettant de s'orienter à travers les centaines de minutiers produits par les notaires parisiens depuis le XV^e siècle et conservés au Minutier central. ETANOT (ETAt des NOTaires parisiens) est une base d'autorités, décrivant non pas des documents d'archives, mais des producteurs d'archives (notaires parisiens du XV^e siècle à nos jours). Dès l'origine, elle a été conçue en conformité avec la norme ISAAR(CPF). ETANOT s'inscrit dans un projet plus général, baptisé NOEMI (NOTaires et MINutes). NOEMI doit être, à terme, un ensemble de bases de données coordonnées et accessibles sur internet, sous la forme d'un portail. ETANOT est l'une de ces bases, les autres étant :

- ETAREP (ETAt des REPERTOires), qui donne accès à l'image numérisée des répertoires de notaires conservés au Minutier. Le travail de numérisation est achevé et la saisie des notices est en cours pour les dernières études.
- la base des états numériques (actuellement dénommée ETAMIN) qui donne, par étude et par notaire, la cote des liasses et des registres conservés au Minutier. Elle a été réalisée sous CINDOC et une version en XML-EAD est actuellement en cours de réalisation.
- MINUTES qui regroupe les descriptions d'actes réalisées dans le cadre de travaux de dépouillement partiels portant sur des thématiques particulières. Elle comprend actuellement 60 000 références et n'est encore consultable qu'au Minutier sous CINDOC.

Les Archives nationales ont également élaboré pour l'administration ministérielle contemporaine, un « codage » désigné par le terme « N1 » transcrivant la hiérarchie des organigrammes. Pour chaque service, on aura un code alphanumérique, un intitulé, des informations concernant la période d'activité du service et les relations chronologiques entre services. Dans le cadre du système d'information du nouveau centre de Pierrefitte-sur-Seine, il est prévu de reprendre la structure de ce fichier pour une mise en conformité avec ISAAR et une simplification de l'identification des producteurs dans le référentiel.

- Thésauri

Le thésaurus Priam 3 s'applique aux documents gérés par les Archives nationales à Fontainebleau et envoyés par les missions des Archives nationales établies dans les différents ministères. Ce thésaurus est donc utilisé dans la base de données du même nom par le site de Fontainebleau et les missions. Il sert à compléter la partie analyse des bordereaux de versement et s'inscrit sur la fiche des unités documentaires. Il a été établi à partir d'un échantillon représentatif de termes introduits dans le champ *objet* de la base de données *Priam 3* depuis 1984.

Chacune de ces rubriques peut comporter jusqu'à sept niveaux de description : un terme générique, trois termes spécifiques et des termes faisant partie d'une liste ouverte mais dont l'ajout est facultatif. Le plus couramment, les archivistes utilisent cinq niveaux.

Le thésaurus *Priam 3* comprend aussi des noms propres d'organismes ou d'entreprises. Pour ces noms, il n'a pas de contrôle d'autorité, mais beaucoup d'informations sont déjà enregistrées à ce titre sur les bordereaux de versement, comme les changements de noms et de responsabilités.

Le thésaurus W paru en 1987 a fait l'objet de refontes en 1989 et 1997 et d'une mise à jour en 2000 (une refonte est en cours), découlant des réflexions formulées par les services d'archives après utilisation. Son utilisation est obligatoire pour tous les services d'archives territoriaux. Il se compose de 1015 descripteurs et de non-descripteurs répartis selon quatre niveaux de hiérarchisation :

- chapitres (11 chapitres correspondant aux grands domaines d'intervention de l'administration),
- rubriques,
- sous-rubriques,
- descripteurs « simples ».

Le thésaurus W comporte aussi trois listes d'autorités :

- une liste d'actions administratives (anciennement « mots outils »)
- une liste de typologie documentaire
- une liste de contextes historiques.

Ces termes ne sont pas en eux-mêmes des termes d'indexation mais ils peuvent être associés à un descripteur du thésaurus pour préciser le contenu d'un dossier.

2*) Quelles étapes faut-il franchir pour rendre les référentiels des archives plus faciles à exploiter par les moteurs de recherche ?

a) Une meilleure articulation des normes et formats d'échange

Si la nécessité de normaliser les descriptions archivistiques fait aujourd'hui consensus, pour permettre l'échange des informations à l'échelle nationale et internationale, la discussion porte dorénavant sur le rôle et la portée des différents outils normatifs.

Les normes ISAD(G) et ISAAR(CPF) sont complémentaires, elles ont chacune un champ bien défini. ISAD(G) fournit les lignes directrices pour la description des fonds d'archives et de leurs composantes (description archivistique) ; ISAAR(CPF) traite de la description des producteurs des fonds d'archives (description contextuelle).

Il est possible soit d'utiliser exclusivement la norme ISAD(G), soit d'appliquer conjointement ISAD(G) et ISAAR(CPF).

L'application de la norme ISAD(G) constitue le niveau minimal de mise en œuvre de descriptions normalisées. Il est tout à fait possible d'utiliser ISAD(G) sans ISAAR(CPF) et de décrire le producteur d'un fonds en même temps que l'on décrit les documents d'archives, comme on le fait dans les instruments de recherche papier traditionnels. Dans ce cas, tous les éléments de la zone du contexte d'ISAD(G) seront à renseigner, notamment l'élément Histoire administrative/notice biographique (3.2.2).

Toutefois, l'utilisation conjointe d'ISAD(G) et d'ISAAR(CPF) permet plus de souplesse dans l'élaboration et la gestion des informations et constitue un deuxième niveau dans la normalisation des descriptions archivistiques. Des descriptions séparées mais liées des producteurs et de leurs archives permettent de ne pas dupliquer les informations, et les deux types de notices peuvent être mises à jour indépendamment les unes des autres. De plus, il sera plus facile pour deux services conservant deux parties d'un même fonds d'échanger des informations sur le producteur de ce fonds. Il convient enfin de noter que, dans le cas d'une utilisation conjointe d'ISAD(G) et

d'ISAAR(CPF), une partie de la zone du contexte d'ISAD(G) perdra de son sens et que l'élément Histoire administrative/notice biographique (3.2.2) ne devra pas être renseigné.

Dans la première version d'ISAAR(CPF), la norme ISAD(G) n'avait pas été assez prise en compte. La version de 2004 est plus structurée, avec quatre zones de description (identification, description, relations et contrôle de la description) et un chapitre spécifique pour lier les notices d'autorité aux descriptions archivistiques. ISAD(G) devra à son tour être révisée pour une meilleure prise en compte d'ISAAR(CPF).

De même, la complémentarité de l'EAD et de l'EAC devra être renforcée et il faudra améliorer l'articulation des deux formats l'un par rapport à l'autre. L'EAD a été conçue pour structurer en XML les descriptions des fonds d'archives et de leurs composantes. L'EAC a été développée pour encoder en XML des notices d'autorité décrivant des producteurs d'archives. La présence d'éléments et d'attributs contenant des informations qui ne devraient figurer que dans l'EAD est donc tout à fait injustifiée dans l'EAC. Il faudrait limiter l'EAC aux informations relatives aux collectivités, personnes et familles. Par ailleurs, les relations entre des documents EAC d'une part, entre des documents EAC et des documents EAD d'autre part devront être clarifiées. Actuellement, les scénarios pour établir des liens entre plusieurs fichiers sont complètement différents selon que l'on part d'un document EAD ou d'un document EAC.

b) Transformation des DTD en schémas

Si la syntaxe d'une DTD est relativement simple et ramassée, sa puissance reste limitée. Aussi la DTD EAD est-elle en train d'évoluer vers un schéma XML. Le schéma XML de l'EAD, de syntaxe plus complexe, présente de nombreux avantages :

- il permet de déclarer des espaces de noms. La déclaration de chaque espace de nom inclut son adresse et un préfixe qui sera précisé devant les éléments correspondants. Ceux-ci restent donc toujours reliés à leur origine ;
- il permet d'imposer des contraintes plus précises sur l'occurrence des éléments et sur le modèle des valeurs attendues (le développement de l'encodage direct en ligne par des équipes d'archivistes ou de bibliothécaires exige plus de contraintes à la saisie pour obtenir plus de cohérence dans les résultats) ;
- il permet de gérer le dictionnaire des éléments de l'EAD (Tag Library) et les règles de gestion dans un document unique : il sera donc possible de définir pour chaque langue une table de correspondance avec les noms des éléments, des attributs et des valeurs des attributs en anglais ; la définition de profils nationaux pour le schéma sera possible en fonction de la langue utilisée pour les noms des éléments, des attributs et de leurs valeurs et en fonction des valeurs retenues pour chaque attribut) ;
- il autorise l'import de schémas XML externes (par exemple la TEI *Text Encoding Initiative*) et leur application locale par différents mécanismes ; inversement, il est possible d'inclure le schéma EAD dans d'autres schémas (par exemple METS *Metadata Encoding and Transmission Standard*).

Avec le schéma XML de l'EAD, il sera donc possible de contraindre les contenus et de créer des profils d'application spécifiques, adaptés aux applications locales tout en respectant une interopérabilité de base.

Toutefois, cette évolution n'a pas été menée jusqu'à son terme :

- la DTD et le schéma de l'EAD coexistent pour le moment : la DTD EAD version 2002 reste la version de référence jusqu'à nouvel ordre (le groupe de travail chargé de la définition du schéma n'a ajouté aucun élément par rapport à la DTD), et si l'EAD évolue (après 2008 ?), les modifications seront faites dans la DTD puis reportées dans le schéma ;

La transformation de l'EAD en schéma aurait pu être l'occasion d'une mise en conformité du standard par rapport aux recommandations internationales sur l'écriture des schémas XML. Or, les codes des balises ne sont pas assez explicites (par exemple, au lieu de <archdesc>, on devrait avoir <ArchivalDescription> pour l'élément EAD Description archivistique). De même, il existe des bibliothèques de balises XML pour les noms de personnes, de lieux ou les dates, réutilisables pour tout type de schéma XML. Or, les concepteurs de l'EAD n'ont pas utilisé ces bibliothèques de balises. Pour toutes ces raisons, la direction des Archives de France n'a pas repris en tant que telle l'EAD dans le format qu'elle a développé avec la DGME pour l'échange des archives électroniques.

3*) Les nouveaux usages vont-ils générer de nouveaux besoins ?

a) Les prolongements de la description contextuelle

La description des producteurs d'archives peut parfois s'avérer insuffisante pour restituer tout le contexte de production des documents d'archives, d'où les deux prolongements d'ISAAR(CPF), en cours de développement au sein du Conseil international des archives.

Le premier projet tend à offrir aux archivistes un cadre pour la création de descriptions de fonctions, descriptions séparées mais liées aux descriptions des archives et des producteurs d'archives. Les fonctions sont généralement reconnues comme plus stables que les structures administratives, qui sont souvent fusionnées ou transférées lors de réorganisations. Les fonctions conviennent bien pour servir de base à l'évaluation, au classement et à la description des archives. L'approche fonctionnelle permet également aux utilisateurs de retrouver et d'analyser les archives, de mieux comprendre la manière dont ils sont produits et utilisés. La future norme, qui devrait être présentée lors du prochain congrès international des archives, en juillet 2008, pourrait servir de base au développement ultérieur de formats de communication et/ou d'échange de données, comme les schémas XML (on peut mentionner les travaux en cours aux Archives nationales de Suède pour le développement d'un schéma XML « EAF » ou « Fonctions archivistiques encodées »). Il convient enfin de noter que, dans le cas d'une utilisation conjointe d'ISAAR(CPF) et d'ISDF pour décrire le contexte de production des documents d'archives, une partie de la zone de la description d'ISAAR(CPF) perdra de son sens et que l'élément Fonctions et activités (5.2.5) ne devra pas être renseigné.

Un deuxième projet concerne la description des institutions de conservation des documents d'archives. Les informations sur les institutions de conservation et les services qu'elles offrent sont essentielles pour l'accès des usagers aux documents. Généralement, ces éléments figurent dans les guides de sources et de recherche. Avec l'avènement du web, il est encore plus important pour les internautes d'obtenir des renseignements sur les services où ils pourront consulter les documents d'archives ou en demander des reproductions. Des descriptions normalisées des services d'archives peuvent également offrir aux usagers des informations contextuelles supplémentaires, qui les aideront à identifier et à interpréter les documents. Toutefois, cette réflexion suscite des interrogations sur les outils permettant de préparer de telles descriptions. Faut-il considérer que les descriptions d'institutions de conservation requièrent une approche et des éléments spécifiques et

élaborer une nouvelle norme ? Faut-il au contraire se contenter d'adapter la norme ISAAR(CPF) à la description d'entités (collectivités, personnes et familles) détentrices d'archives ? Par ailleurs, avec les besoins actuels d'échange, de collaboration et d'interopérabilité, il faudrait approfondir la question de l'identification des services d'archives, Trois normes ISO, suivies au niveau français par la Commission nationale de la normalisation sur les identifiants (AFNOR CG46/CN6) pourraient rejoindre ces travaux :

- ISCI (International standard collection identifier) : identification des institutions patrimoniales concernées ;
- ISNI (International Standard Name Identifier) : identification des personnes physiques ou morales impliquées dans les différents maillons de la chaîne allant de la conception jusqu'à la consommation des contenus intellectuels. Les problèmes du statut juridique y sont également traités ;
- ISIL (International Standard Identifier for Libraries and Related Organisations) : identifiant international normalisé pour les bibliothèques et organisations liées.

Parallèlement, des échanges sont en cours au sein du groupe international chargé de faire évoluer l'EAC vers une version 1.0, notamment sur la portée et le contenu du format. Certains membres du groupe ont avancé que si l'on considère que les fonctions font partie du contexte de production des documents d'archives, un seul format pourrait servir à décrire les collectivités, personnes et familles et les fonctions qu'elles exercent ? Toutefois, dans un format orienté implémentation, il est préférable par pragmatisme de garder une cohérence par rapport à la nature des entités faisant l'objet de la description associées d'une manière ou d'une autre à la production et à la gestion des archives. Les relations des fonctions et des collectivités, personnes et familles avec les documents d'archives ne sont pas les mêmes : les collectivités, personnes ou familles produisent ou gèrent des documents alors que les fonctions sont reflétées par les documents. Les collectivités, personnes ou familles sont dotées de la capacité d'agir et les fonctions sont exercées par les dites collectivités, personnes ou familles. Il paraît donc peu envisageable qu'un même format serve à leur description.

b) Organiser les éléments de données : modéliser

Il ne suffit pas de recenser et de définir les éléments de données. Organiser les informations pour construire un réseau de relations entre les éléments de données constitue l'étape suivante, afin de proposer un modèle générique de données d'autorité permettant l'interopérabilité. Les archivistes australiens ont commencé à réfléchir à l'élaboration de modèles entités – relations, distinguant les entités (documents d'archives, producteurs ou détenteurs des documents et fonctions) et les relations entre ces entités. Dans ces modèles, les entités décrites et leurs relations ont des propriétés ou attributs. Suite à l'appel à commentaires pour les projets de normes sur les fonctions et les institutions de conservation, en 2007, les Australiens ont adressé à ICA des propositions d'évolution des normes de description allant dans le sens d'un modèle unique de référence pour les normes de description, identifiant trois entités correspondant aux documents d'archives (ISAD), aux autorités (ISAAR) et aux fonctions (ISDF). Les attributs communs aux trois entités (identifiant, type, intitulé/forme autorisée de nom, dates d'existence, etc.) serait regroupés dans une entité de niveau supérieur. Les relations entre ces trois entités pourraient elles-mêmes être qualifiées par un jeu d'attributs.

En France, dans le cadre de l'intégration de nouveaux instruments de recherche dans l'application IREL (Instruments de recherche en ligne), une réflexion est en cours [aux Archives nationales d'outre-mer](#) pour l'élaboration de listes d'autorité pour les noms de personnes, les noms d'institutions, les noms de lieux et les mots matière, et sur l'établissement de notices

d'autorité conformes à ISAAR(CPF) pour les personnes ou institutions citées dans les différents instruments de recherche. Il s'agit en effet de dégager une logique d'accès commune aux différents instruments de recherche pour guider l'internaute le plus efficacement dans sa recherche. Ce développement d'IREL se traduira par la conception de formulaires d'interrogation qui seront offerts à l'utilisateur. Plus denses en informations que les listes, les notices d'autorité permettront de définir et de donner des informations sur les termes utilisés en indexation et éventuellement d'organiser les liens entre ces termes. Toutefois, les listes d'autorité ne coïncideront pas nécessairement avec les notices d'autorité. Pour les personnes, seuls les acteurs les plus importants de la colonisation seront retenus.

Une réflexion sur l'adaptation de la norme NF Z 44-081 – *Forme et structure des vedettes géographiques* aux besoins spécifiques des ANOM sur les noms de lieux est également en cours. Il s'agit moins d'aider l'indexeur à trouver le terme adéquat dans une liste d'autorités que de donner à l'internaute le contexte historique qui lui permettra de mieux s'orienter dans les ressources qui lui seront proposées. Il sera donc nécessaire de rendre explicites les relations entre objets zonaux et objets ponctuels, en offrant notamment à l'utilisateur la possibilité d'accéder, à partir d'un toponyme, à des termes génériques et spécifiques. Toutefois, une liste hiérarchisée, où chaque toponyme aurait une place unique, ne permettrait pas de contextualiser de manière satisfaisante ces relations, souvent rendues complexes par l'histoire administrative et politique. Une liste alphabétique de noms de lieux, dont certains seraient liés à des notices d'autorité, sur le modèle des notices d'autorité décrivant des institutions ou des personnes, offrirait beaucoup plus de souplesse et permettrait de restituer ce réseau relationnel entre entités géographiques dans toute sa complexité. Un modèle XML permettant de contextualiser les relations entre les différents toponymes est en cours d'élaboration. Une première réflexion a été lancée sur les métadonnées que devrait contenir ce modèle (forme autorisée, autres formes de noms, dates, commentaire historique, relations avec d'autres entités géographiques, avec des collectivités, des personnes, des ressources archivistiques, etc.).

Autre exemple de relation entre entités : les systèmes d'information italiens reposent sur un modèle de descriptions séparées d'entités conceptuellement diverses et étendent ce modèle à toutes les entités qui contribuent à délimiter le cadre qui décrit un fonds d'archives dans toutes ses composantes de contenu et de contexte, en offrant des représentations séparées mais liées entre elles. Cette solution confère au système une souplesse particulière, mais aussi une pluralité de clés de recherche et d'accès, par exemple, [le portail des surintendances archivistiques italiennes](#) (SIUSA), qui donne accès aux descriptions des archives autres que celles conservées par les archives d'État, comprend en fait trois « objets » archivistiques principaux : le fonds d'archives, le producteur et l'institution de conservation. À côté de ces trois entités figurent des données complémentaires aux informations relatives aux ensembles documentaires et aux entités productrices. Il est en effet possible de lier au fonds d'archives et au producteur une fiche « Bibliographie » et une fiche « Sources », où sont décrites les sources non imprimées d'où proviennent les informations. Le fonds d'archives peut être relié à une fiche « Instruments de recherche », qui liste les répertoires, inventaires, catalogues, états sommaires, etc., mais aussi à une fiche « profil documentaire général », où il est possible de décrire de manière exhaustive des typologies spécifiques de documents. Aux descriptions de producteurs sont rattachées trois autres types d'informations : informations sur les « Contextes politico-institutionnels » d'appartenance, sur les « Territoires » sur lesquels le producteur a exercé ses fonctions, et sur le « Profil institutionnel général », c'est-à-dire une typologie des producteurs, avec la possibilité d'établir des liens entre des producteurs relevant d'une même catégorie, sans répéter à chaque fois les informations contextuelles.