

Service interministériel des Archives de France Bulletin sur les ressources archivistiques numériques, n° 44, septembre 2012

Sommaire (cliquer sur les titres)

Projet « Mettre en place une infrastructure nationale pour les autorités archivistiques ».....	1
Participants.....	1
Fondement et contexte.....	2
La preuve de concept.....	2
Quel modèle pour une coopération en matière d'autorités archivistiques ?.....	3
Suite du projet.....	4
76e Congrès de la Société des archivistes américains, San Diego, 7-12 août 2012.....	4
Comités et groupes de travail spécialisés.....	4
<i>Comité des normes.....</i>	<i>4</i>
<i>Sous-comité technique sur l'EAC-CPF.....</i>	<i>6</i>
<i>Sous-comité technique sur l'EAD et Table ronde des utilisateurs de l'EAD.....</i>	<i>7</i>
<i>Section sur la description.....</i>	<i>11</i>
<i>Table ronde des utilisateurs des logiciels Archivists'Toolkit et Archon et du projet ArchiveSpace.....</i>	<i>12</i>
<i>Réunion des consortiums EAD à l'initiative d'OCLC.....</i>	<i>13</i>
Sessions plénières.....	16
<i>Session 305 - Fiabilité au-delà des frontières ; élaboration et mise en œuvre de la norme ISO TDR.....</i>	<i>16</i>
<i>Session 401 – Relier les données des bibliothèques, des archives et des musées.....</i>	<i>18</i>
<i>Session 505 – Résoudre nos problèmes avec le contrôle d'autorité et le partage : développements récents et perspectives.....</i>	<i>23</i>
Congrès international des archives, Brisbane (Australie), 21-24 août 2012.....	23

Projet « Mettre en place une infrastructure nationale pour les autorités archivistiques »

Compte rendu de la réunion de lancement du projet par Anila Angjeli de la BnF - invitée à la réunion en tant que co-présidente du sous-comité technique de l'EAC-CPF

Les 21 et 22 mai dernier, NARA (National Archives and Records Administration) accueillait dans ses locaux à Washington, D.C. la première d'une série de réunions autour du nouveau projet « Mettre en place une infrastructure nationale pour les autorités archivistiques » (Building a National Archival Authorities Infrastructure)¹. Ce projet a été financé pour instruire la question de la mise en place d'une Coopérative nationale d'autorités archivistiques (National Archival Authority Cooperative) aux Etats-Unis, à travers deux activités² : d'une part, l'organisation, sous l'égide de la SAA, d'un cycle d'ateliers régionaux de formation à l'EAC-CPF et d'autre part, l'élaboration d'un ensemble de recommandations pour la mise en place d'une Coopérative nationale d'autorités archivistiques. Ces recommandations seront publiées sous forme d'un livre blanc en automne 2013.

Cette initiative a été chaleureusement accueillie par de nombreux archivistes à travers les États Unis, car elle répond aux préoccupations de la profession dans un paysage où l'information archivistique est en pleine évolution et interagit avec les autres secteurs d'information patrimoniale.

Participants

=> [Sommaire](#)

Cette première réunion du projet a fait preuve d'une forte volonté d'ouverture, dépassant de loin le seul cercle des archivistes. Représentants de la Bibliothèque du Congrès, de la Bibliothèque nationale de médecine, de la Bibliothèque nationale d'agriculture, de l'Institution smithsonienne

¹ Site du projet http://socialarchive.iath.virginia.edu/NAAC_index.html

² L'Institut des technologies avancées dans les humanités (IATH) et l'Ecole supérieure des sciences de l'information et des bibliothèques à Simmons College ont été titulaires du financement par L'Institut pour des services aux musées et bibliothèques (IMLS) pour mener les deux activités du projet.

(Smithsonian Institution) et d'autres organismes fédéraux côtoyaient archivistes des archives d'Etat, des archives régionales et de plusieurs universités. S'ajoutaient à ceux-ci des représentants de grandes bibliothèques numériques comme la Bibliothèque numérique de Californie (California Digital Library), d'institutions de recherche telle que l'Institut de recherche Getty (Getty Research Institute), des chercheurs en sciences sociales et des organismes financeurs.

Fondement et contexte

=> [Sommaire](#)

Les instruments d'archives regorgent d'informations authentiques sur les personnes, collectivités et familles productrices de ou associées aux documents d'archives. Ces informations sont essentielles à la compréhension des documents d'archives et font partie de l'information contextuelle archivistique. Outre leur valeur informative sur les circonstances entourant la production et l'utilisation des documents d'archives, ces informations ont une valeur indépendante. Elles permettent de comprendre la vie, l'œuvre et l'activité des personnes, collectivités et familles, de mettre en lumière les relations qu'elles peuvent avoir entre elles, ainsi que leurs relations avec les documents d'archives. Elles pourraient potentiellement servir à relier entre eux les documents d'archives émanant de la même personne, collectivité ou famille, ou portant sur celles-ci, indépendamment du lieu de conservation des archives ou du système dans lequel celles-ci sont décrites.

Mais le modèle historique de rédaction des instruments de recherche archivistiques, comme un seul appareil monolithique, confondant description des documents d'archives et description de l'information contextuelle, ne favorise pas la gestion autonome de cette information.

L'intérêt de rendre l'information contextuelle autonome s'est confirmé ces deux dernières décennies. Économie de description, efficacité et souplesse de gestion, mais aussi mise en relation des informations contenues dans les documents d'archives dispersés dans différentes localisations et décrits dans des instruments de recherche et systèmes d'information isolés, telles ont été les motivations qui ont fait évoluer les principes de description archivistique au niveau international. La norme ISAAR(CPF) assoit le principe des descriptions autonomes de personnes, collectivités et familles, producteurs de ou mentionnées dans les documents d'archives et en fixe le contenu. La DTD EAD 2002 a permis la réalisation des premières descriptions autonomes en XML de ces entités. Enfin le schéma EAC-CPF, édition 2010³, plus en phase avec les évolutions technologiques récentes, a offert un outil permettant plus d'autonomie et de rigueur, en favorisant la mise en relation des descriptions de personnes, collectivités et familles. Ce schéma est, en outre, interopérable avec les standards d'autres domaines culturels. Il a donné un véritable élan à la description autonome des personnes, collectivités et familles dans les systèmes d'information archivistiques. Il donne lieu à des applications innovantes qui permettent, d'une part, la valorisation de l'information authentique contenue dans les descriptions et, d'autre part, sa connexion avec celle d'autres domaines.

La preuve de concept

=> [Sommaire](#)

En ouverture de réunion, le projet de recherche SNAC (Social Networks and Archival Context ou « Réseaux sociaux et contexte archivistique »)⁴ a été présenté comme preuve de concept de la mutualisation des informations sur les personnes, collectivités et familles, contenues dans les différents instruments de recherche répartis aux quatre coins des Etats-Unis, mais aussi dans le monde.

Un financement par la fondation Mellon a permis au projet d'entrer en 2012 dans une nouvelle phase. Un bouquet de technologies avancées, de nouvelles méthodes et outils sera mobilisé pour fouiller les

³ Site officiel de l'EAC-CPF <http://eac.staatsbibliothek-berlin.de>

⁴ Site officiel du projet SNAC <http://socialarchive.iath.virginia.edu>

Ce projet a déjà fait l'objet d'une présentation dans le Bulletin sur les ressources archivistiques numériques, n° 43, septembre 2011 <http://www.archivesdefrance.culture.gouv.fr/static/5236/>

instruments de recherche archivistiques, en extraire l'information sur les personnes, collectivités et familles et le restructurer, le relier et en créer un réseau d'informations. Le projet explorera un large éventail de possibilités pour apparier l'information relative à la même entité, faire place à des données en d'autres langues, ajouter des données géographiques aux lieux, rendre les données biographiques ou historiques explorables via des frises chronologiques (dates, lieux et événements), visualiser les réseaux socio-historiques et les organigrammes des organismes sous forme graphique. Il est prévu d'exposer les données en Linked Data avec liens vers le fichier d'autorité international virtuel (VIAF), un projet international visant à fusionner plusieurs fichiers d'autorité nationaux en un seul système maître, et DBPedia. Un SPARQL endpoint (interface web permettant d'interroger de l'information numérique structurée en RDF) est déjà créé pour des requêtes en RDF et un SNAC widget (composant d'interface graphique) est également disponible et intégrable dans les applications locales.

Treize consortiums et plus de trente-trois institutions contribuent avec des données source. Ainsi, parmi les contributeurs, la National Archives and Records Administration (NARA), la Smithsonian Institution, la Bibliothèque du Congrès, l'OCLC WorldCat fournissent des instruments de recherche et VIAF et Getty Vocabulary Programm des données d'autorité. L'intérêt que ce projet suscite dépasse le seul cadre des Etats-Unis. La British Library s'est portée volontaire pour fournir un nombre significatif d'instruments de recherche. En France, les Archives de France et la Bibliothèque nationale de France ont déjà fourni des instruments de recherche en français. Cette collaboration est une opportunité de valorisation de nos archives dans un contexte international, de mise en relation avec d'autres documents relatifs aux mêmes personnes, familles ou collectivités, mais également de récupération des descriptions enrichies sur celles-ci et leur intégration dans nos systèmes locaux.

Quel modèle pour une coopération en matière d'autorités archivistiques ?

=> [Sommaire](#)

Les trois axes discutés - modèle fonctionnel, solutions technologiques et gouvernance -ont suscité de vifs débats. Tout en faisant part d'expériences et de besoins dans leurs propres univers de travail, les participants se sont efforcés de se projeter dans une réflexion sur un modèle de système d'un type nouveau, s'inscrivant dans l'écosystème en mutation de l'information patrimoniale.

L'éventail des questions a été large allant de la discussion sur les contours du système jusqu'aux interfaces pour répondre aux besoins diversifiés. L'objectif est, certes, de générer une information à la fois authentique, fiable et de haute qualité, mais aussi dynamique et modulaire avec des modes opératoires souples, et une vision à long terme pour des utilisations futures des données.

Les problèmes de gouvernance se révèlent également complexes. Définition du champ de couverture du système, gouvernance centralisée versus distribuée, mutualisation des données, droit des données, personnel, instances décisionnelles, formation, autant de questions qui nécessiteront une attention particulière. Pour engager le projet, avant tout, un « foyer » est indispensable. NARA paraît l'organisme idéal d'hébergement ; sa réputation et son rôle aux Etats-Unis sont à la hauteur des ambitions de l'initiative. Elle a déjà manifesté son intérêt pour le projet mais n'a pas fait de déclaration d'intention.

Clifford Lynch, président de la Coalition pour l'information en réseau (Coalition for Networked Information), chargé initialement de résumer les discussions, a délibérément repositionné le débat, de manière quelque peu provocatrice, en invitant les participants à réfléchir sur les deux voies alternatives qui s'offrent aujourd'hui pour la mise en place d'une Coopérative nationale d'autorités archivistique. D'une part, celle qui mise pratiquement entièrement sur les technologies de l'information, avec très peu d'intervention humaine, très attirante en apparence, d'autre part celle inspirée par le contrôle d'autorité traditionnel (avec l'exemple de NACO⁵), misant sur le travail humain pour gérer le réseau des acteurs, contrôler, mettre à jour, valider des notices générées préalablement automatiquement. Quel équilibre économique, donc, entre la composante computationnelle et l'intervention humaine ? Le succès d'un système, où les technologies avancées seront amenées à

⁵ NACO - Name Authority Cooperative Program, un des programmes du PCC Program for Cooperative Cataloging <http://www.loc.gov/aba/pcc/>

jouer un rôle majeur, ne pourrait être assuré que si ces dernières interagissent avec l'expertise des professionnels et des contributeurs avertis.

Il a ensuite attiré l'attention des présents sur la nécessité d'avoir l'œil ouvert sur un paysage en mutation où le « contrôle d'autorité » est en pleine éruption et se décline en facettes multiples. Contrôle d'autorité dans les bibliothèques, en premier lieu, accroissement des bases de données ouvertes du domaine musical avec information détaillée sur les interprètes, mouvement rétrospectif dans le domaine de la recherche pour identifier les auteurs (ou comment combler l'écart entre les morts et les vivants), biographies factuelles, nationales, etc., ne sont que quelques exemples. La question des identifiants, avec l'émergence de l'ISNI⁶ en premier lieu, mais aussi de l'ORCID⁷ et la question des relations entre les deux est également à suivre de près.

Suite du projet

=> [Sommaire](#)

Une équipe-projet restreinte poursuivra la réflexion amorcée sur le modèle fonctionnel, les solutions technologiques et la gouvernance. Plusieurs des représentants des différentes communautés d'intérêt qui participaient à cette première réunion, ont déjà répondu à l'appel à volontaires pour la constitution de cette équipe. Elle étudiera en profondeur les trois aspects de l'éventuelle Coopérative nationale d'autorités archivistiques et élaborera des propositions de stratégie, d'action et de moyens. Quant aux ateliers de formation à l'EAC-CPF, un premier a eu lieu lors du congrès annuel de la SAA, en août 2012. Le projet a fait l'objet de plusieurs présentations lors du 76^e Congrès de la Société des archivistes américains et lors du 17^e Congrès international des archives à Brisbane. Nous suivrons avec intérêt les prochaines évolutions de l'ensemble du projet et vous en tiendrons informés.

76^e Congrès de la Société des archivistes américains, San Diego, 7-12 août 2012

Le 76^e Congrès annuel de la Société des archivistes américains (SAA) s'est tenu à San Diego, Californie, du 7 au 12 août 2012. Comme chaque année, cette manifestation proposait de nombreuses activités : séminaires et formations avant le congrès proprement dit, visites professionnelles d'institutions patrimoniales et de services d'archives, réunions de groupes spécialisés et sessions plénières.

Cette année, une des thématiques majeures a été celle des Données ouvertes liées (Linked Open Data ou LOD). Rendre les données libres et ouvertes permet de disposer d'un fonds commun de données dans lequel n'importe qui peut puiser et ce, malgré la grande diversité des contextes juridiques et institutionnels de leur création. Lors de la séance d'ouverture du congrès, Jon Voss, directeur des partenariats stratégiques de HistoryPin et organisateur de la première conférence internationale sur les Données liées dans les bibliothèques, archives et musées à San Francisco en 2011 (*cf. infra*), s'est donc efforcé de montrer comment les LOD pouvaient représenter, pour les bibliothèques, les archives et les musées, une opportunité de décloisonnement et d'ouverture des métadonnées qui peuvent ainsi être rendues plus facilement réutilisables et partageables.

Comités et groupes de travail spécialisés

Comité des normes

=> [Sommaire](#)

Le Comité des normes de la SAA (SC) est chargé de coordonner le processus d'élaboration, de révision et d'approbation de toutes les normes de la Société des archivistes américains. Il doit également assurer la diffusion des normes au sein de la communauté professionnelle nord-

⁶ ISO 27729 International Standard Name Identifier <http://www.isni.org>, un identifiant pour les identités publiques de toute partie prenante (personne ou collectivité) dans la chaîne allant de création jusqu'à la diffusion des contenus intellectuels, scientifiques, etc.

⁷ Open Research & Contributor ID <http://about.orcid.org/>, un identifiant pour les chercheurs.

américaine. A ce titre, un portail d'information sur les normes a été lancé lors du précédent congrès annuel de la SAA, en 2011 (<http://www2.archivists.org/standards>).

Le SC comprend différents groupes de travail et sous-comités techniques :

- sous-comité technique sur les règles pour les bâtiments d'archives (TS-AFG)
- sous-comité technique sur « Décrire les archives : une norme de contenu » (DACS)
- sous-comité technique sur l'EAD (TS-EAD)
- sous-comité technique sur l'EAC-CPF (TS-EAC-CPF)
- équipe de développement des schémas (SDT)
- sous-comité technique sur la réévaluation et la dévolution des collections (TS-GRD).

Le sous-comité technique sur les règles pour les bâtiments d'archives (TS-AFG) est chargé de maintenir une norme sur les bâtiments intitulée « Bâtiments des services d'archives et des collections spéciales : lignes directrices pour les archivistes, les bibliothécaires, les architectes et les ingénieurs ». Adoptée officiellement par la SAA, cette norme fournit des directives pour l'évaluation sur site et pour la construction de bâtiments d'archives qui répondent aux besoins des chercheurs et du personnel et qui permettent d'assurer la conservation des collections et des fonds d'archives. Le sous-comité technique étudie également la faisabilité de la transformation de cette norme en norme nationale et/ou son application aux différentes institutions patrimoniales.

Le sous-comité technique « Décrire les archives : une norme de contenu » (TS-DACS) assure la maintenance intellectuelle et technique de la norme « Décrire les archives : une norme de contenu ». La norme DACS fournit un ensemble de règles pour la description des collections de manuscrits, archives et documents personnels et elle peut être appliquée à tous les types de documents. Norme officiellement adoptée par la SAA, DACS est compatible avec la norme ISAD(G). Le processus de sa révision a commencé en 2011 avec un appel à commentaires. Le TS-DACS s'est subdivisé en quatre groupes de travail, chargés d'examiner les commentaires relatifs aux différentes parties de DACS et de faire des recommandations pour une révision éventuelle. Il a été décidé de supprimer le chapitre III de la norme, relatif à l'établissement de points d'accès contrôlés, pour se concentrer sur la description archivistique proprement dite (chapitre I) et la description des collectivités, personnes et familles impliquées dans la production, la conservation et l'utilisation des archives (chapitre II). Le chapitre II est en cours de refonte pour mieux prendre en compte ISAAR (CPF) et donner des recommandations sur le contenu des notices d'autorité conformes au format EAC-CPF. Une version révisée provisoire a été publiée sur le site de la SAA en juillet 2012 (<http://www2.archivists.org/groups/technical-subcommittee-on-describing-archives-a-content-standard-dacs>). La communauté professionnelle est invitée à adresser ses commentaires à la SAA d'ici la mi-septembre. En novembre, un sous-groupe du TS-DACS se réunira à Chicago pour finaliser la révision de DACS. La nouvelle version de la norme devrait être rendue disponible en juin 2013. Un site Web spécifique sera créé pour faciliter l'application de la norme, comprenant des exemples d'encodage, des bonnes pratiques, des équivalences avec d'autres normes, etc.

Le sous-comité technique sur l'EAD (TS-EAD) est chargé de la mise à jour intellectuelle et technique du format (voir ci-dessous).

Le sous-comité technique sur l'EAC-CPF (TS-EAC-CPF) est chargé du suivi et de la maintenance du schéma (voir ci-dessous).

L'équipe de développement de schéma (SDT) est responsable du développement et de la diffusion de tous les schémas et DTD maintenus par la SAA, ainsi que de leur documentation technique.

Le sous-comité technique sur la réévaluation et la dévolution des collections (TS-GRD) est chargé du suivi des « Lignes directrices pour la réévaluation et la dévolution des collections ». Ces règles donnent des procédures et différentes informations pour les archivistes qui peuvent être confrontés à la réévaluation des collections détenues par leur institution.

Sous-comité technique sur l'EAC-CPF

=> [Sommaire](#)

1) Rapport annuel d'activités :

Suite à la réunion annuelle de 2011, le TS-EAC s'était fixé quatre objectifs principaux :

- modifier la bibliothèque de balises (Tag Library) dans un souci d'homogénéisation et d'exactitude ;
- promouvoir et gérer la traduction en français, en allemand, en italien, en espagnol, etc. de la bibliothèque de balises ;
- mettre à jour le site Web ;
- trouver de nouveaux exemples et revoir ceux figurant déjà dans la Tag Library.

- Modification de la bibliothèque de balises :

La bibliothèque de balises a été révisée afin de prendre en compte les demandes de modifications suite aux travaux de traduction en français puis en allemand. Ces travaux de traduction avaient soulevé plusieurs questions sémantiques, terminologiques et techniques. Certaines ont été résolues ; d'autres ont besoin d'un examen plus approfondi car elles sont liées à l'interopérabilité des formats EAC-CPF et EAD. Le signalement de ces problèmes sémantiques, terminologiques et techniques sera certainement très utile pour la révision de l'EAD et pour la rédaction de sa documentation technique. Une version révisée de la bibliothèque de balises de l'EAC-CPF est prête pour l'encodage. L'équipe de développement de la SAA (SDT) est en train d'élaborer un modèle et une infrastructure pour une gestion dynamique de cette documentation et de ses diverses traductions existantes et à venir.

- Traductions :

Les traductions en français et italien sont achevées (la traduction française est déjà publiée en ligne et accessible sur le site de l'EAC-CPF) ; les traductions en grec et en allemand sont en cours, et la traduction en espagnol commencera bientôt. Des informations sur l'état d'avancement des traductions et des instances qui en sont responsables dans les différents pays sont disponibles sur le site de l'EAC-CPF.

- Site Web :

Le site Web de l'EAC-CPF (<http://eac.staatsbibliothek-berlin.de/>) a été mis à jour. Des discussions sont en cours sur la répartition des informations entre le site Web officiel de l'EAC-CPF à la Bibliothèque d'Etat de Berlin et le site de la SAA. La décision sera prise pour mettre en conformité l'emplacement des informations sur les deux standards l'EAD et l'EAC-CPF (cf. *infra* l'information sur la Table Ronde EAD). Les statistiques du site indiquent que le trafic sur le site Web se situe entre 500 et 1 000 visiteurs uniques par mois. En 2011, il y a eu 13 187 visites sur le site ; de janvier à juillet 2012, 9 155 visites sur le site.

- Exemples :

Jerry Simmons de NARA a dirigé les travaux sur les exemples à fournir dans la bibliothèque des balises. Le but est de mieux étoffer la bibliothèque des balises avec des exemples contextuels. Le groupe de travail a révisé les exemples figurant déjà dans la documentation technique ainsi que sur le site EAC-CPF, mais il en a aussi recueilli de nouveaux pour le site Web. Le groupe devait se réunir à San Diego après la réunion conjointe des TS-EAC et EAD.

2) Informations sur différents projets en cours :

- SNAC/NAAC (projets « Réseaux sociaux et contextes archivistiques »/ « Coopérative nationale pour les autorités ») : une réunion a été tenue sur le sujet en mai aux Archives nationales (NARA), à Washington, deux autres réunions sont prévues (cf. *supra* compte rendu détaillé de la session présentant l'avancement du projet). Par ailleurs, en mars dernier, la SAA a proposé le premier des

sept ateliers régionaux programmés d'ici juin 2013, afin de faciliter la diffusion de l'EAC-CPF. L'Institut des services de musées et de bibliothèques (Institute of Museum and Library Services ou IMLS) finance vingt bourses d'études pour chacun des sept ateliers. Les ateliers de l'EAC-CPF durent une journée et initient les participants à la structure générale de la norme, avec une importante composante pratique. Les participants apprennent comment créer des notices EAC-CPF avec l'application pratique des éléments de données existants. Les ateliers comprennent aussi des présentations de projets de mise en œuvre au niveau local, régional ou national.

- Archives nationales (NARA) : un groupe de travail a été créé en interne en vue de l'utilisation par NARA de l'EAD et de l'EAC-CPF ; jusqu'à présent, les normes et formats d'échange élaborés par la Société des archivistes américains étaient appliqués par tous les services d'archives américains à l'exception des Archives nationales...
- APEnet (Portail européen archivistique) : un nouveau projet a débuté, intitulé APEX. L'équipe de projet réfléchit à l'implémentation de la DTD EAG (Guide archivistique encodé), développé par le ministère de la Culture espagnol pour structurer en XML les descriptions de services d'archives. Or, l'EAG n'existe pour le moment que sous forme de DTD et ce format est antérieur à la Norme internationale pour la description des institutions conservant des archives (ISDIAH). L'EAG devrait être révisée, toutefois, les responsables du projet ont tenté d'enrichir l'EAC-CPF avec les informations spécifiques aux services d'archives. Par ailleurs, une réflexion est en cours sur la gestion des identifiants.
- La Bibliothèque Beinecke de l'université de Yale conduit un projet collaboratif avec l'université de Harvard pour démontrer les avantages de l'utilisation de l'EAC-CPF pour décrire les producteurs de collections de manuscrits et encoder les relations sémantiques significatives entre les producteurs eux-mêmes et les sources primaires qui documentent leur vie et leur travail. Le projet sera axé sur le lexicographe Samuel Johnson (1709-1784) et son cercle de relations (écrivains, artistes, penseurs politiques et amis).

3) Objectifs pour 2013 :

- Il est prévu de continuer à enrichir le site Web et de trouver de nouveaux exemples ;
- Mark Matienzo (université de New-York), propose de recueillir des cas d'utilisation de l'EAC-CPF dans le cadre du développement du logiciel libre ArchivesSpace (*cf. infra*). Il est, par ailleurs, souligné que l'EAC-CPF cette dernière année gagne à être utilisé pour divers projets centrés sur les personnes, collectivités et familles. Son utilisation dépasse le monde des archives (*cf. infra*). Il sera utile pour la communauté des professionnels de diffuser l'information sur ces utilisations ;
- suite aux travaux en cours en Suède, en France, en Suisse et au Royaume-Uni, Daniel Pitti suggère l'organisation de deux réunions pour lancer officiellement le travail sur le futur schéma relatif aux fonctions exercées par les producteurs d'archives ;
- la question des relations devenant le cœur de la réflexion sur l'harmonisation et l'interopérabilité des schémas relatifs à la description archivistique, Daniel Pitti suggère de développer un vocabulaire contrôlé sur la typologie des relations en s'inspirant d'autres normes et de promouvoir des bonnes pratiques en la matière.

Sous-comité technique sur l'EAD et Table ronde des utilisateurs de l'EAD

=> [Sommaire](#)

1) Rapports annuels d'activités :

Site officiel de l'EAD :

Au 1^{er} août 2012, le site comptait 1 233 membres actifs (adresses professionnelles) représentant 31 pays et régions (États-Unis, Royaume-Uni, Canada, Australie, Pays-Bas, Allemagne, France, Italie, Espagne, Suisse, Autriche et Grèce).

D'août 2011 à juillet 2012, 506 messages portant sur 159 sujets ont été affichés, soit environ 624 000 courriels envoyés).

Table ronde EAD :

Chargée de favoriser la mise en œuvre et l'utilisation de l'EAD, la Table ronde EAD mène actuellement une réflexion sur le contenu de son site Web et des pages d'aide destinées aux professionnels, pour :

- y inclure des guides de bonnes pratiques, différents outils ;
- y inclure des informations sur les consortia utilisant l'EAD, les ressources et les détails de mise en œuvre ;
- y inclure des informations sur les réalisations autour de l'EAC-CPF.

Équipe de développement du schéma :

Une réunion se tiendra les 9-10 octobre 2012, avant la diffusion de la version alpha du schéma de l'EAD (*cf. infra*).

Traduction russe de la Tag Library :

La traduction est achevée mais apparemment pas encore disponible sur le Web.

2) Rapport d'étape sur la révision de l'EAD :

- Principes généraux et historique du processus de révision :

Dans sa démarche de révision de l'EAD, le TS-EAD a dû tenir compte tout à la fois des nécessaires changements à apporter et des instruments de recherche existants – les outils de migration seront essentiels. Certains changements feront probablement l'objet de débats au sein de la communauté professionnelle.

Les principes de révision établis par le TS-EAD sont rappelés :

- parvenir à une plus grande cohérence conceptuelle et sémantique dans l'utilisation de l'EAD ;
- explorer les mécanismes permettant de rendre plus transparentes les informations encodées en EAD, de les mettre plus facilement en relation et de les échanger avec des données maintenues selon d'autres protocoles, et d'incorporer les données provenant d'autres espaces de noms ;
- améliorer les fonctionnalités de représentation des informations descriptives créées au niveau international, en particulier dans les environnements multilingues ;
- être conscient qu'une nouvelle version affectera les utilisateurs actuels.

Le processus de révision a débuté par le lancement d'un appel à commentaires auprès de la communauté internationale, d'octobre 2010 à février 2011. Une première restitution en a été faite sous forme de forum lors du congrès annuel de la SAA à Chicago, en août 2011. Le TS-EAD a eu des réunions téléphoniques au cours du printemps et de l'été 2012 puis une réunion de trois jours à la Bibliothèque Beinecke de l'université de Yale, en mars 2012.

Les commentaires reçus de la communauté professionnelle portaient sur l'objet et la nature générale du schéma, les liens avec les ressources et les entités archivistiques, l'en-tête EAD, les éléments de la hiérarchie descriptive (<archdesc>, <dsc>, composants numérotés), les éléments englobants (<did>), les points d'accès contrôlés (<controlaccess>), les objets archivistiques numériques (<dao> et <daogrp>) et l'encodage des dates.

La plupart des décisions prises lors de la réunion de Yale en mars dernier ont été mises en œuvre. Toutefois, les éléments relatifs au contrôle de la description et aux relations avec d'autres entités archivistiques ne sont pas encore implémentés. Le schéma est donc encore incohérent pour le moment. De plus, plusieurs points sont encore en suspens : la représentation des relations (*cf. infra*), le modèle de contenu de l'élément <chronlist>, les éléments de dates, la récursivité de certains éléments et l'élément <div>, l'élément <archref> et l'attribut @href.

- Synthèse des décisions prises provisoirement :

Plusieurs orientations ont déjà été prises concernant la révision de l'EAD.

Les éléments dépréciés dans l'EAD2002 seront éliminés du futur schéma (<add>, <admininfo>, <dentry>, <drow>, <organization>).

Seront dépréciés les éléments suivants : <frontmatter>, <eadgrp>, <archchdescgrp>, <dscgrp>, <descgrp>, <fileplan>, <abbr>, <expan>, <runner>, <daogrp>, <linkgrp>, <note> (dans certains cas - point en cours de discussion), <dsc>.

Il a été décidé de supprimer <arrangement> de <scopecontent>, de supprimer <acqinfo> de <custodhist>, de supprimer <legalstatus> de <accessrestrict>, de supprimer <unitdate> de <unittitle> et d'éliminer la récursivité de <bioghist> dans <bioghist> et de <controlaccess> sans <controlaccess>.

La sémantique de l'élément <note>, qui a actuellement huit usages différents dans l'EAD, est en cours de révision. Cet élément peut en effet figurer dans <notesmt>, dans les éléments <titlepage> ou <div> de <frontmatter>, dans <did>, dans les éléments frères et sœurs de <did>, dans les éléments frères et sœurs de <did> comme élément de type « bloc » (pour créer des paragraphes, dans <note> (l'élément <note> est récursif), dans les éléments bloc comme <p> (pour des contenus mixtes), ou dans <namegrp>.

Le modèle de contenu des éléments servant à créer des points d'accès contrôlés sera modifié, avec la possibilité d'utiliser des éléments <part> pour différencier les différentes composantes de formes de noms (nom proprement dit et qualificatifs par exemple).

La mise en cohérence avec XHTML entraînera des changements mineurs dans les éléments <table> et <list>, la suppression de <table>, <list> et <chronlist> de <p>, la limitation de <blockquote> à un élément de type « bloc » (pour créer des paragraphes), l'ajout d'un élément <quote> de type « en ligne » (pour attribuer un style à une portion de texte).

Les mécanismes de liens seront simplifiés : <extref>, <extptr> et <ptr> seront supprimés, seul <ref> sera maintenu, les attributs de lien seront supprimés de certains éléments comme par exemple <title>, les liens étendus (<daogrp>, <linkgrp>) seront dépréciés, l'espace de nom XLink sera supprimé mais seront implémentés des attributs de lien isomorphes avec XLink.

Des attributs servant à enregistrer les codes de langues et d'écritures seront rendus disponibles pour chaque élément non vide.

Plusieurs décisions visent à réconcilier l'EAD avec l'EAC-CPF :

- <eadheader> sera remplacé par <control> : <filedesc> sera maintenu mais les autres éléments de <eadheader> seront remplacés par des éléments plus appropriés pour le contrôle de la description.
- Il sera possible d'établir des <relations> à tous les niveaux ; les relations pourraient être regroupées dans un élément englobant <relations> sur le modèle de l'EAC-CPF (point en cours de discussion).
- Le modèle de données pour les dates sera amélioré (point en cours de discussion).
- Le modèle de données de <chronlist> sera amélioré (point en cours de discussion).

Le futur schéma comprendra quelques nouveaux éléments/attributs :

- a. Les informations relatives à l'importance matérielle et au support des documents décrits pourront être enregistrées soit sous une forme textuelle non structurée dans un élément <physdesc> soit sous une forme structurée dans un élément <physdescstruct> avec un attribut @physdescstype pour typer les informations. Dans tous les cas, les outils de migration fournis avec le schéma permettront de faire migrer les données existantes vers <physdesc>/<phydescstruct>.
- b. Un élément <formsavailable> devrait être créé :

```
<did><formsavailable><container type="Box">1</container><dao href="url"/></formsavailable></did>
```
- c. Un attribut @containerId devrait être rendu disponible dans <container>.
- d. Un élément <didnote> contenant du PCDATA et des éléments de contenu mixte (<emph>, <ref>, etc.), mais non pas <p>, <table>, <list> devrait être rendu disponible dans <did>.
- e. L'attribut @normal sera rendu disponible dans <unittitle>.

Plusieurs des propositions adressées à la SAA par la communauté internationale ont été rejetées : les éléments <did>, <head>, <abstract> sont maintenus, ainsi que les composants <c01-12> ; il a été décidé de ne pas faire de <archdesc> un simple élément englobant contenant seulement des composants <c>/<c01-12> ; l'attribut @level reste obligatoire seulement pour le haut niveau de description, et non pas pour tous les niveaux de description.

Différents outils seront livrés avec le schéma : une (des) feuille(s) de style pour la migration, des outils de validation externe pour valider les listes de codes ISO et pour établir des profils (sous-ensembles du schéma), ainsi que des feuilles de style pour changer les composants numérotés en composants non numérotés.

Pour plus d'informations, voir : <https://github.com/SAA-Sdt/EAD-revision>

Le schéma est pour le moment en version pré-alpha, en cours de réécriture complète : <https://github.com/SAA-SDT/EAD-Revision/blob/master/ead2002toEADx.xsl>

- Discussion sur les points encore en suspens : les relations :

Le TS-EAD en est arrivé à la conclusion provisoire que les relations doivent pouvoir être établies à tous les niveaux. Le problème est après de savoir comment les mettre en œuvre. Faut-il suivre le modèle de l'EAC-CPF, c'est-à-dire ajouter un élément <relations> regroupant toutes les relations ? Dans ce cas, les valeurs de l'attribut @relationType définissant les types de relations devront être modifiées et un espace de nom EAD devrait peut-être être créé à cette fin. Faut-il au contraire maintenir les éléments <origination>, <repository>, <separatedmaterial>, <relatedmaterial>, etc., et simplifier le modèle de <origination> et de <repository> ? Pour le TS-EAD, les deux solutions apparaissent exclusives l'une de l'autre.

Pour Terry Catapano, responsable de l'équipe de développement, c'est plus un problème de syntaxe qu'un problème conceptuel. On pourrait résoudre le problème en développant un profil spécifique (sous-ensemble du schéma général). Ce profil (ou sous-ensemble) pourrait aborder la question des relations de manière détaillée, pour les rendre exploitables par la machine.

Pour Kathy Wisser, co-présidente du TS-EAC, le modèle actuel de <origination> répond à deux besoins : identification d'un producteur et établissement d'une relation avec ce producteur. L'élément <origination> n'a donc pas tout à fait la même finalité que des éléments servant à établir des relations avec des collectivités, personnes et familles.

A l'issue de la discussion, il est décidé de recueillir des cas d'utilisation avant de prendre une décision. Deux membres du TS-EAD, Mark Matienzo et Kelcy Shepherd, sont chargés de recueillir des exemples d'utilisation de l'EAD seule et des exemples de mises en relation d'instances EAD avec d'autres ressources.

- Calendrier de révision :

Le calendrier prévu pour la révision est le suivant :

- après la réunion de la SDT (octobre) : diffusion d'une version alpha du schéma
- janvier 2013 : diffusion d'une version beta et lancement d'un appel à commentaires
- mars : date limite de réception des commentaires
- mars-avril : réunion de l'équipe éditoriale de la Tag Library
- juillet : fin du processus de révision (schéma et documentation technique).

Section sur la description

=> [Sommaire](#)

La Section sur la Description est un lieu d'échange d'informations et d'idées sur tous les aspects de la description archivistique pour les responsables de projets de description, de normes de description ou de systèmes d'information, pour ceux qui y participent ou qui s'y intéressent.

Oublier l'arriéré : gérer les entrées avec le logiciel Archivists'Toolki à l'université de Californie à Irvine

En 2010, l'université de Californie à Irvine (UCI) a lancé un projet pour doter les fonds non classés d'une notice descriptive conforme à la norme DACS. L'approche est celle de l'étude réalisée en 2007 par deux archivistes, Mark Greene et Dennis Meissner. Intitulée « Plus de résultats, moins de procédures » (« More Product, Less Process » ou MPLP), cette étude donnait des recommandations pour réduire les arriérés de traitement des fonds dans les services d'archives. Il s'agit de donner une description minimale des fonds dès leur entrée. « Il est préférable de rendre tous les fonds accessibles plutôt que de ne pas porter à la connaissance du public certaines collections ». Pour en savoir plus, voir le support de l'exposé à : <http://fr.slideshare.net/librarchivist/forget-about-the-backlog-surfacing-accessions-using-archivists-toolkit>

Les Archives audiovisuelles historiques – Accès par le biais de l'indexation – Collections de la Fondation de la Shoah

Créée en 1994 afin de préserver les témoignages de survivants et d'autres témoins de la Shoah, la Fondation de la Shoah (USC Shoah Foundation) conserve une des plus importantes bibliothèques numériques vidéos dans le monde : près de 52 000 témoignages vidéos dans 32 langues et provenant de 56 pays. Sa mission est de vaincre les préjugés, l'intolérance et le fanatisme – et les souffrances qu'elles causent – grâce à l'utilisation pédagogique des témoignages vidéos historiques de l'Institut. La Fondation a mis au point un thésaurus, un logiciel d'indexation et des applications accessibles au public pour faciliter l'utilisation par le grand public, les chercheurs, les étudiants et les enseignants.

L'indexation est appliquée à deux niveaux – les données bibliographiques décrivent brièvement l'expérience biographique de la personne interrogée et l'indexation vidéo fournit des points d'accès tout au long de l'interview qui est divisée en segments de 1 minute à cette fin. Chaque témoignage offre un aperçu unique dans les souvenirs et les perceptions de l'expérience de pré-génocide, de génocide et de post-génocide que révèle la personne interrogée.

Bien que les histoires racontées soient chargées d'émotion, l'indexation en elle-même s'efforce d'être neutre en se concentrant sur le temps, les lieux, les événements, les personnes, les mouvements de population, les activités, les organisations, les mouvements politiques, les affiliations religieuses, et les pratiques culturelles et sociales.

Suite à des accords de coopération avec des organisations à vocation similaire, le champ couvert par la collection sera bientôt élargi pour inclure des interviews sur le génocide des Tutsis rwandais (Centre mémoriel de Kigali) et le génocide des Arméniens (Fondation pour le cinéma arménien).

Pour en savoir plus, consulter l'application Archives audiovisuelles historiques en ligne à : <http://vhaonline.usc.edu/>

Des composants perturbateurs : réinventer les systèmes d'accès archivistiques à la bibliothèque de l'université de Princeton

D'après Elizabeth Yakel, professeur en sciences de l'information de l'université du Michigan, la production d'instruments de recherche analogiques traditionnels dans des environnements en ligne « inhibe l'utilisation créative d'un réseau d'informations et l'émergence de nouvelles formes

numériques de représentation des sources primaires. » D'après elle, le contenu des archives elles-mêmes serait rendu moins accessible.

Cet exposé décrivait les efforts d'un groupe de travail sur la description archivistique créé au sein de l'université de Princeton pour repenser la production et la diffusion des descriptions archivistiques. S'appuyant sur des données descriptives, créées par des dizaines d'agents impliqués dans des projets de traitement de collections et de conversion de données au cours des sept dernières années, le groupe a élaboré, en collaboration avec l'équipe chargée des projets de bibliothèque numérique, une nouvelle application web donnant accès aux descriptions archivistiques (disponible à <http://findingaidsbeta.princeton.edu>).

Les caractéristiques de ce nouveau système d'accès sont les suivantes :

- affichage plus modulaire des instruments de recherche, avec notamment un accès direct aux composants EAD <C> à partir de la liste des résultats de recherche ;
- options de consultation par facettes et par navigation à partir des résultats de recherche ;
- affichage des images directement à partir de l'interface des instruments de recherche ;
- listes de sujets avec des possibilités de tri par titres, dates ou emplacement physique des fonds ;
- accès à des notices au format EAC-CPF ;
- plus de possibilités de dialogue avec les utilisateurs qui peuvent poser des questions, proposer des améliorations à apporter aux descriptions archivistiques, discuter du contenu des documents ou partager des informations au sein de leurs propres réseaux sociaux.

Table ronde des utilisateurs des logiciels Archivists'Toolkit et Archon et du projet ArchiveSpace

=> [Sommaire](#)

Cette table ronde regroupe les communautés d'utilisateurs des logiciels Archivists'Toolkit ou AT (la « Boîte à outils des archivistes ») et Archon. Ces deux logiciels libres open source, respectivement développés par les universités de San Diego et de l'Illinois, permettent de produire des descriptions archivistiques conformes à la norme DACS et de les exporter en EAD 2002.

Financé par la Fondation Andrew W. Mellon, un troisième projet, ArchivesSpace, vise à développer une nouvelle génération d'outil archivistique open source combinant les fonctionnalités des logiciels Archon et Archivists'Toolkit. Cet effort est conduit par les bibliothèques de l'université de New York, l'université de Californie à San Diego et l'université de l'Illinois à Urbana-Champaign, travaillant en étroite collaboration avec la communauté archivistique.

Le projet se caractérise par le développement d'un partenariat durable et d'une communauté. En effet, l'équipe de projet comprend :

- pour l'administration : un responsable du développement et un chef de projet ;
- une équipe chargée de la maîtrise d'ouvrage ;
- un responsable technique ainsi qu'un consultant de la Bibliothèque numérique de Californie ;
- un comité de gouvernance.

Il y a quelques mois, LYRISIS, la plus grande organisation d'appartenance régionale pour les bibliothèques et les professionnels de l'information aux États-Unis, a été choisie pour servir de structure organisationnelle pour ArchivesSpace. Le début du développement du logiciel ArchivesSpace est prévu pour l'été 2012, avec une sortie à la fin de 2013.

La société Molonglo Hudson a été choisie comme partenaire de développement pour ArchivesSpace. Molonglo Hudson est un cabinet de consultants en nouvelles technologies pour l'enseignement supérieur, les bibliothèques et les archives, avec des équipes aux États-Unis et en Australie.

La vision du produit est la suivante : « un système de gestion open source supportant les fonctions archivistiques de base, qui puisse être dimensionné par rapport aux besoins de grandes ou de petites institutions, qui soit souple, efficace et facile à utiliser, à maintenir et à administrer, qui soit

extensible et qui puisse facilement interagir avec les autres applications et systèmes, avec un écosystème sain d'utilisateurs, de développeurs et de partenaires dédiés pour maintenir le produit dans un environnement open source ».

Des cas d'utilisation ont servi à définir les fonctionnalités (pour en savoir plus, voir le cahier des charges disponible sur le site du projet : <http://www.archivespace.org/>)

Le développement a commencé en juillet. Les tests utilisateurs et la recette devraient débuter en septembre. Le plan de communication autour du produit sera lancé en fin d'année, avec l'élaboration d'une documentation utilisateurs provisoire et la mise au point d'outils de migration des données saisies dans AT et Archon vers le nouveau logiciel. Les tests beta et le déploiement de ArchivesSpace devraient commencer en mai 2013. Le projet devrait être achevé en août 2013.

Réunion des consortiums EAD à l'initiative d'OCLC

=> [Sommaire](#)

L'OCLC (Online Computer Library Center) est une organisation à but non lucratif dont la vocation est de favoriser la collaboration entre les bibliothèques, les archives et les musées en élaborant des modèles pour le partage de données, des services et de l'expertise. Il comprend un département spécialisé, OCLC Recherche, issu de l'absorption par OCLC d'un autre consortium, le Research Library Group (RLG), en 2007. Ses objectifs sont les suivants :

- introduire des modèles coopératifs pour l'acquisition, la gestion et la communication des collections ;
- aider les archives, les bibliothèques et les musées à atteindre une compréhension commune des processus dont ils sont responsables ;
- expérimenter ces modèles coopératifs au travers de prototypes ;
- les rendre possibles via des codes et des architectures open source.

Voici quelques exemples de projets et de réalisations en cours à l'OCLC :

- ArchiveGrid : ce service gratuit de découverte donne accès à 1,7 millions de descriptions de collections archivistiques détenues par des milliers de bibliothèques, de musées, de sociétés historiques et de services d'archives du monde entier (<http://www.beta.worldcat.org/archivegrid>). Les notices descriptives, aux formats MARC, EAD, HTML, PDF ou Word, sont importées de Worldcat (base de données en ligne de l'OCLC, réputée le plus grand catalogue bibliographique collectif du monde) ou moissonnées à partir des sites des institutions contributrices. Pour plus d'informations, voir le blog d'ArchiveGrid : <http://beta.worldcat.org/archivegrid/blog/> Les institutions qui souhaitent contribuer à ArchiveGrid peuvent le faire via un formulaire librement accessible en ligne (<http://experimental.worldcat.org/archivegrid/contribute>).
- « Instruments de recherche prêts à l'emploi » : ce projet de recherche explore l'efficacité des outils et des techniques pour amener sur Internet des descriptions archivistiques. Des services d'archives ont converti des milliers d'instruments de recherche existants en EAD ou HTML et les ont rendus disponibles en ligne, mais beaucoup d'autres ne sont accessibles que dans le service. Ces descriptions non accessibles sur Internet sont cachées comme si les collections n'avaient jamais été traitées, même si le travail intellectuel de classement et de description a été réalisé. OCLC Recherche s'est attaqué au problème par l'élaboration et l'utilisation d'un ensemble d'outils pour permettre aux instruments de recherche papier d'être présents dans ArchiveGrid.
- Des Wikipédiens résident actuellement à OCLC et à la British Library dans le cadre d'un projet de recherche de l'OCLC sur les fichiers d'autorité et leurs liens avec Wikipédia. 260 000 entités sont concernées, mais seulement en langue anglaise. Ce projet vise à étendre et à systématiser l'utilisation d'identifiants pour pouvoir établir des liens entre des fichiers d'autorité et les articles de Wikipédia en anglais. Le projet est construit autour de l'utilisation du fichier d'autorité international virtuel (VIAF), projet d'OCLC Recherche visant à l'origine à aligner des listes d'autorités (notamment sur les noms propres) en vue de constituer une base de référence internationale. Les identifiants de VIAF correspondent aux identifiants dans d'autres systèmes et peuvent être utilisés en parallèle avec, ou à la place d'autres identifiants.

- « Démystifier les documents nativement numériques » : ce projet vise à aider les collections spécialisées et des archives qui n'ont pas encore commencé à gérer leurs documents nativement numériques. OCLC Recherche avait fait une étude en 2009 sur les archives dans les bibliothèques de recherche et des collections spécialisées : seule une minorité de bibliothèques universitaires ou de recherche a fait d'importants progrès dans ce domaine. La plupart des institutions de ce secteur font de la collecte, mais souvent ponctuellement, et peu connaissent l'importance de leurs collections numériques. OCLC Recherche publiera une série de brefs rapports sur les différents défis posés par les documents nativement numériques : une feuille de route indiquant les premières étapes de gestion de ces documents ; un appel pour une approche concertée pour la migration des documents numériques stockés sur des supports physiques obsolètes ; un inventaire des compétences pertinentes que les archivistes doivent acquérir ; et un aperçu de la manière dont ces domaines d'expertise sont pertinents pour les différents types de documents.

La réunion organisée pendant le Congrès de la SAA par le département Recherche de l'OCLC visait à faciliter les échanges d'informations et à partager les expériences entre différents consortiums et agrégateurs de ressources archivistiques mettant en œuvre l'EAD.

1) Northwest Digital Archives (<http://nwda.orbiscascade.org/index.shtml>)

Le portail des Archives numériques du Nord-Ouest (NWDA) donne accès à des instruments de recherche EAD décrivant les archives et collections de manuscrits des bibliothèques universitaires de l'Idaho, du Montana, de l'Oregon, de l'Alaska et de l'État de Washington. Le projet a été financé par la Dotation nationale des sciences humaines (National Endowment for the Humanities) et par la Commission nationale des publications historiques et des archives. Depuis juillet 2007, NWDA est un programme de l'Alliance Orbis Cascade.

L'Alliance Orbis Cascade est un consortium de bibliothèques universitaires du nord-ouest des États-Unis. Le consortium a été formé par la fusion en 2003 de deux précédents consortiums, Orbis et Cascade, qui regroupaient respectivement des bibliothèques dans l'Oregon et dans l'État de Washington. Trente-six collègues et universités de l'Oregon et de l'État de Washington sont membres directeurs de l'Alliance, ainsi qu'un nouveau membre, l'université de l'Idaho. L'Alliance regroupe plusieurs types de bibliothèques dans une zone plus large qui inclut l'Oregon, l'État de Washington, l'Idaho, le Montana, l'Alaska et Hawaï.

2) Fabrique d'instruments de recherche de Floride (<http://ffaf.fcla.edu/>)

La Fabrique d'instruments de recherche de Floride (Florida Finding Aid Factory ou FFAF) permet aux institutions de Floride de créer, de mettre à jour et de publier des instruments de recherche archivistiques. Le projet utilise Archon, un logiciel open source hébergé par le Centre pour l'informatisation des bibliothèques en Floride (FCLA).

Avec la plate-forme FFAF, les institutions peuvent produire des instruments de recherche en ligne, exportables sous forme de notices MARC ou sous forme d'instances EAD. Tous les instruments de recherche créés avec le service FFAF sont rendus accessibles dans l'interface Archives de Floride (<http://palmm.fcla.edu/afl/>).

L'outil est mis gratuitement à disposition des institutions participantes, afin d'améliorer l'accès aux documents d'archives en Floride.

3) SNAC-NAAC

Cf. compte rendu détaillé de la réunion de lancement du projet en mai à Washington.

4) Projet collaboratif autour de l'EAD à l'institution smithsonienne

Fondée en 1846, l'institution smithsonienne est en elle-même un consortium, puisqu'elle comprend dix-neuf musées et galeries, un parc zoologique et neuf centres de recherche. C'est le logiciel libre Archivists' Toolkit (AT) qui a été choisi pour produire des instruments de recherche EAD. A terme, chaque fonds ou collection devrait être pourvu d'un instrument de recherche EAD. Pour le moment, 84 collections sont décrites avec le logiciel AT. Des séances de formation ont été organisées pour les utilisateurs du logiciel et des guides de bonnes pratiques ont été élaborés pour aider à la saisie. La Smithsonian a également recruté un archiviste chargé de coordonner les opérations d'encodage. Le projet est entièrement financé en interne.

5) Archives Portail Europe

Le projet APEnet, regroupant douze archives nationales et soutenu par la Commission européenne a été finalisé avec succès et son objet principal - le Portail européen des Archives - est aujourd'hui un portail entièrement fonctionnel (que l'on trouve à l'adresse <http://www.archivesportaleurope.eu>). Pour l'instant, le portail contient plus de 14 millions d'unités descriptives, liées à quelque 64 millions de pages numérisées de documents d'archives. Le portail, qui présente le matériel de 62 institutions de 14 pays européens, est aujourd'hui un acteur principal dans l'univers du patrimoine culturel européen.

Un nouveau projet, APEx, poursuivra les travaux du projet APEnet mais s'occupera aussi de l'élargissement du réseau du Portail européen des Archives, en y associant davantage de partenaires/pays. Il assistera aussi les institutions locales et régionales détenant des archives, dans la publication de leurs données via le Portail européen des Archives, de sorte que plus de documents soient transférés vers Europeana, la bibliothèque numérique européenne.

L'EAD a été définie comme le format pivot du projet. La seule contrainte imposée aux partenaires est donc de fournir leurs données sous la forme de fichiers EAD-XML. En effet, au centre d'APEX, un instrument de recherche global réunit toutes les données afin de proposer un accès commun aux descriptions archivistiques et, le cas échéant, aux documents numérisés eux-mêmes. Les trois piliers du modèle de cet instrument de recherche global sont :

- l'outil de conversion et de validation des données originales dans le format commun (APEnet-ead). Ce format commun a été défini sur la base d'une analyse fine des pratiques de chaque pays dans la mise en œuvre de l'EAD ; il est adapté selon l'évolution des besoins ou l'entrée de nouveaux partenaires dans le projet.
- l'outil de chargement et d'intégration qui transmet les données à l'hébergement central,
- l'outil d'indexation et d'affichage qui permet de combiner les données pour la navigation et l'affichage.

Une mise en œuvre des formats EAC-CPF (pour la description des producteurs d'archives) et EAG (Guide archivistique encodé pour la description des institutions de conservation) est prévue, mais dans le deuxième cas, le format d'échange est antérieur à la Norme internationale pour la description des institutions de conservation et devra donc être révisé avant d'être implémenté.

6) Archives en ligne de Californie (<http://www.oac.cdlib.org/>)

Archives en ligne de Californie (OAC) est un portail de recherche donnant accès aux descriptions et aux images numériques de collections conservées dans les archives, bibliothèques et musées de Californie. Une étude est en cours sur les besoins d'hébergement hors de la Californie – surtout au niveau régional.

7) Archives d'Arizona en ligne (<http://www.azarchivesonline.org/xtf/search>)

La plateforme Archives d'Arizona en ligne (AAO) donne gratuitement accès aux descriptions des collections des services d'archives, des bibliothèques, des sociétés historiques et des musées de l'Arizona. L'application utilise XTF, une plate-forme open source développée par la Bibliothèque numérique de Californie, permettant de donner accès à du contenu numérique. Le financement

d'AAO est assuré en grande partie par l'État de l'Arizona. Les institutions contributrices sont regroupées au sein d'un consortium, l'Alliance des Archives de l'Arizona. Les institutions patrimoniales de l'État de l'Arizona se mettent peu à peu à utiliser le logiciel Archivists'Toolkit, puis elles exportent en EAD et affichent leurs descriptions sur les Archives d'Arizona en ligne

8) Portail OhioLINK (<http://www.ohiolink.edu/>)

Le réseau des bibliothèques et d'information de l'Ohio, OhioLINK, est un consortium regroupant les bibliothèques d'établissements d'enseignement supérieur et d'universités de l'Ohio et la bibliothèque de l'État de l'Ohio. Au service de plus de 600 000 étudiants, enseignants et agents de 88 institutions, OhioLINK regroupe 16 universités publiques, 23 établissements d'enseignement supérieur, 49 établissements privés et la bibliothèque de l'État de l'Ohio. OhioLINK offre des services aux étudiants, aux agents et aux autres chercheurs via des systèmes de catalogage électroniques, le serveur central OhioLINK et des ressources Internet.

Sessions plénières

Session 305 - *Fiabilité au-delà des frontières ; élaboration et mise en œuvre de la norme ISO TDR*

=> [Sommaire](#)

Cette session était consacrée à l'élaboration de la norme ISO 16363:2012 *Systèmes de transfert des informations et données spatiales -- Audit et certification des référentiels numériques de confiance (Space data and information transfer systems -- Audit and Certification of trustworthy digital repositories)* et ses implications pour la conservation pérenne et l'archivage du numérique.

Pourquoi auditer des dépôts numériques ? Le but le plus évident est d'obtenir une forme de certification, un « label » permettant d'établir que les dépôts sont dignes de confiance. Mais le but de la démarche d'audit est avant tout d'évaluer les vulnérabilités et les risques qui pèsent sur le dépôt, en toute transparence, pour établir de la confiance.

Qui engendra Qui : origines et élaboration de la norme internationale pour un dépôt numérique de confiance (Bruce I. Ambacher, professeur invité, université du Maryland)

Les seize dernières années ont vu des efforts continus et persistants pour aboutir à la mise en place de processus et de normes d'audit et de certification de dépôts numériques. En 2002, deux consortiums, le Research Library Group (RLG) et l'OCLC ont publié un rapport intitulé « Dépôts numériques fiables » (TDR). Le TDR rendait compte des travaux d'un groupe de travail qui a été le premier à établir les caractéristiques d'un dépôt numérique pour les organismes de recherche. Ces caractéristiques s'appuyaient sur la nouvelle norme internationale du modèle de référence pour un système d'archivage ouvert (OAIS). En 2007, les Archives nationales des États-Unis (NARA) et l'OCLC ont publié un document définissant les critères pour l'audit et la certification : « Critères pour l'audit et la certification de dépôts numériques » (Trusted Repositories Audit and Certification: Criteria and Checklist ou TRAC).

Dans le même temps, une réflexion parallèle était menée en Europe sur l'audit et la certification des dépôts numériques. Le Réseau d'expertise sur le stockage à long terme des ressources numériques, ou NESTOR, a créé des normes nationales pour la préservation numérique en Allemagne. Issu d'une collaboration du Centre de la Conservation du numérique (Digital Center ou DCC) et de Préservation du numérique en Europe (DigitalPreservationEurope), DRAMBORA est un outil d'auto-vérification pour les gestionnaires de dépôts numériques. DANS⁸, avec d'autres partenaires européens, a lancé le Sceau d'approbation de données (DSA).

⁸ Créée en 2005, DANS (*Data Archiving and Networked Services*) est une institution de l'Académie des arts et des sciences des Pays-Bas (KNAW) également soutenue par l'organisation néerlandaise pour la recherche scientifique (NWO). DANS assure le stockage et l'accessibilité permanente aux données de la recherche en sciences sociales et humaines.

En 2007, les travaux ont commencé, au sein du groupe de travail pour l'audit et la certification des dépôts numériques du Comité consultatif pour les systèmes de données spatiales (CCSDS), pour transformer les « Critères pour l'audit et la certification des dépôts numériques » en norme ISO 16363 ; les travaux se sont achevés en 2011 et la norme vient d'être publiée par l'ISO. Le processus d'édition visait à s'assurer que la norme pourrait servir de base pour un processus international d'audit et de certification pour tous les types de dépôts numériques. Plusieurs tests d'audits ont été réalisés en Europe et aux États-Unis pour vérifier la facilité d'utilisation de la norme à la fois par les gestionnaires de dépôts et les auditeurs. L'insertion de nombreux exemples dans la norme devrait également aider les gestionnaires de dépôts numériques à effectuer un premier audit interne.

Audit et Certification des dépôts numériques : ISO et le cadre européen (David Giaretta, Directeur APA, Alliance pour l'accès permanent)

Le principal outil de l'audit, c'est une liste de critères préétablis qui vont permettre d'évaluer le dépôt numérique. Ces critères (ou « metrics ») se veulent être des indicateurs objectifs. La liste en anglais de ces indicateurs est disponible à :

<http://www.digitalrepositoryauditandcertification.org>

Ces indicateurs portent sur l'infrastructure organisationnelle des dépôts numériques, sur la gestion des objets numériques, et sur la gestion de l'infrastructure et des risques de sécurité. Ils visent à permettre l'auto-évaluation. Sont-ils trop ou au contraire pas assez nombreux ? Il est impossible d'anticiper tous les cas possibles et les autres normes (normes de sécurité ISO 2700x) sont assez succinctes. Ces indicateurs doivent plutôt être regardés comme un guide pour les auditeurs (domaines sur lesquels se concentrer, choix de sous-critères spécifiques). Ils dépendent essentiellement de l'expérience et du jugement des auditeurs.

L'évaluation peut reposer sur un certain nombre de preuves :

- la documentation : charte, spécifications, profils de postes, rapport annuel, etc.
- les interviews du personnel, faites par l'auditeur
- les démonstrations et observations directes.

L'évaluation repose ainsi sur la mesure de l'écart qui existe entre ce que dit la documentation, ce que pense le personnel et ce qui se passe en réalité.

Mettre en œuvre la certification internationale de dépôts numériques de confiance (Helen R. Tibbo et Mark Conrad, université de Caroline du Nord, Chapel Hill)

Faire un audit coûte cher, non seulement à celui qui le fait, mais aussi à l'institution audité. Une auto-évaluation ne pouvant pas être suffisante, qui va faire les audits ? Pour le moment, il ne semble pas y avoir vraiment de réponse à cette question, bien que la liste des critères, unifiée au plan international, ait été reprise par la norme ISO. On semble également s'orienter vers plusieurs niveaux de certification, permettant d'établir des réseaux de confiance entre institutions des différents niveaux. Une prochaine étape consistera à accréditer de nouveaux auditeurs selon les processus et critères de l'ISO. Les auditeurs devront avoir suivi une formation de cinq jours de formation agréée (couvrant les bases de l'OAIS, de la norme ISO 16363, l'audit en tant que processus, avec des exemples d'audits, etc.) et avoir acquis une expérience dans l'ensemble du processus d'évaluation de la fiabilité des archives numériques. L'objectif final reste d'établir la confiance dans les dépôts numériques par :

- des améliorations continues en matière de capacités de préservation numérique,
- la certification pour s'assurer que les ressources sont bien employées,
- la formation d'un réseau cohérent d'auditeurs et de consultants dans le monde entier,
- la constitution d'un marché pour les fournisseurs, les constructeurs de systèmes.

Session 401 – Relier les données des bibliothèques, des archives et des musées

=> [Sommaire](#)

Le Web des données (Linked Data, en anglais) est une initiative du W3C (Consortium World Wide Web ou Consortium W3C) visant à favoriser la publication de données structurées sur le Web, non pas sous la forme de silos de données isolés les uns des autres, mais en les reliant entre elles pour constituer un réseau global d'informations. Il s'appuie sur les standards du Web, tels que le protocole HTTP et les URI (identifiants uniformes de ressources, courtes chaînes de caractères identifiant des ressources sur un réseau) - mais plutôt qu'utiliser ces standards uniquement pour faciliter la navigation par les êtres humains, le Web des données les étend pour partager l'information également entre machines. Le monde du patrimoine – bibliothèques, archives, musées... – a commencé à s'engager sur la voie du web de données, et les initiatives se développent : l'an passé un « Sommet Données liées ouvertes dans les bibliothèques, archives et musées » (LOD-LAM Summit – Linked Open Data in Libraries, Archives, and Museums) s'est tenu à San Francisco (<http://lodlam.net/>) – et le groupe d'incubation « Bibliothèques et Web de données » (Library Linked Data) du W3C a rendu un rapport de préconisations (<http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>). Les intervenants de cette session ont présenté les aspects généraux des LOD et des exemples spécifiques relatifs aux vocabulaires et aux normes archivistiques, ainsi que des projets mettant en œuvre les LOD.

Les Données liées dans les Archives, publier, enrichir, réconcilier, mettre en relation (Corey Harper, Métadonnées des dépôts numériques, université de New York)
<http://files.archivists.org/conference/sandiego2012/401-Harper.pdf>

=> [Sommaire](#)

Le premier intervenant a commencé par rappeler les origines et les concepts des Linked data. Les définitions suivantes peuvent aider à mieux comprendre les différences entre Données ouvertes (Open Data), Données liées (Linked Data), Données ouvertes liées (Linked Open Data) et Web sémantique (Semantic Web) :

- les données peuvent être ouvertes (Open Data), mais pas forcément liées (Linked Data) ;
- les données peuvent être liées, mais pas forcément ouvertes ;
- les données, qui sont à la fois ouvertes et liées (Linked Open Data), sont de plus en plus viables ;
- le web sémantique ne peut fonctionner qu'avec des données ouvertes et liées.

Corey Harper a ensuite présenté quelques projets autour des Données liées.

DBpedia (<http://dbpedia.org/>) est un effort communautaire pour extraire des informations structurées à partir de Wikipedia et pour rendre ces informations disponibles sur le Web des données ouvertes liées. DBpedia permet de poser des requêtes complexes par rapport à Wikipedia, et de lier d'autres ensembles de données sur le Web à des données de Wikipedia. Ce projet est conduit par l'université de Leipzig, l'université libre de Berlin et l'entreprise OpenLink Software. Dans sa version actuelle, la version 3.7, DBpedia contient plus de 3,64 millions d'éléments, parmi lesquels 1,83 millions sont classés selon une ontologie cohérente. Cette ontologie repose sur des infoboîtes (infoboxes en anglais), tables préformatées de données dynamiques qui présentent sommairement des informations importantes sur un sujet dans des cartouches ou des encadrés placés en général à droite des articles de Wikipédia. Le dataset (un « ensemble de données ») propose des marquages et des descriptions pour 3,64 millions d'éléments, et ce jusque dans 97 langues différentes ; 2 724 000 liens vers des images et 6 300 000 liens redirigeant vers des sites Internet extérieurs ; 6 200 000 liens externes reliés à d'autres ensembles de données RDF, et 740 000 entrées Wikipédia.

Freebase est une méta-base de données dont le but est d'agréger gratuitement des ressources disponibles sur le web, que des utilisateurs peuvent ensuite améliorer. Un second projet open source est venu se greffer sur Freebase, ayant pour but d'organiser visuellement Freebase. Ce projet, intitulé Thinkbase, a été créé par des chercheurs de l'université d'Auckland. Le site de Thinkbase (<http://thinkbase.cs.auckland.ac.nz>) se divise en deux – cette division est ajustable –, avec, à droite la page *Freebase*, à gauche sa forme visuelle faite des diverses connexions. Thinkbase extrait les

contenus depuis Freebase, y compris les relations sémantiques, et permet de les visualiser en utilisant une représentation visuelle interactive.

Le système sémantique **Knowledge Graph**, lancé par Google aux Etats-Unis en mai dernier(<http://www.google.com/insidesearch/features/search/knowledge.html>), et qui reconnaît 500 millions d'entités nommées (noms de personnes, de lieux, d'entreprises) liées à 3,5 milliards d'attributs et de relations entre elles, a été étendu à tout le monde anglophone (Australie, Grande-Bretagne, etc.). Le Knowledge Graph vise à répondre encore plus rapidement aux requêtes de l'internaute. Plutôt que de lui retourner une liste de sites web où se trouve l'information, Google va y répondre directement. Par exemple, dans le cadre d'une recherche sur le Burj Khalifa, Google retourne plusieurs données sur le gratte-ciel (hauteur, nombre d'étages, début des travaux, date d'ouverture, adresse, architectes...) et de multiples photographies du bâtiment. Autre exemple : le système sait que Pierre-Auguste Renoir est un impressionniste, né à Limoges le 25 février 1841. La base de données est capable de le rapprocher d'autres peintres de la même école et de recenser ses œuvres, bref de lier des informations qui restent d'habitude isolées. Pour en savoir plus : <http://www.google.com/insidesearch/features/search/knowledge.html>

Google Refine (<http://code.google.com/p/google-refine/>) est un outil permettant de travailler avec des données en désordre, en créant des facettes, en constituant des groupes, en transformant ces données d'un format vers un autre format, en les enrichissant avec des services web et en les mettant en relation avec des bases de données comme Freebase.

Les Linked data permettent à la BBC de lier les 1000 à 1500 programmes diffusés par jour à leurs auteurs, aux acteurs, producteurs, à leur page web et leurs créneaux de diffusion. L'opérateur britannique tire un double bénéfice de son usage des Linked data : il pérennise ses contenus en leur attribuant une adresse pérenne et en facilite l'accès en les agrégeant, quelle que soit leur origine. Par exemple, le site **Wildlife Finder** (<http://www.bbc.co.uk/nature/wildlife>) présente les différentes espèces et d'autres rangs biologiques. La BBC ne produit pas de contenus, ou peu, elle les agrège depuis les sites de l'Organisation mondiale de protection de l'environnement (WWF), de l'université du Michigan, de la Société zoologique de Londres (ZSL), de l'Union internationale pour la conservation de la nature et des ressources naturelles (IUCN) et enfin de DBpedia.

Le moteur de recherche **RelFinder** (<http://www.visualdataweb.org/relfinder.php>) est un outil basé sur Adobe Flex qui permet de faire ressortir des relations entre des objets (personnes, endroits, choses, etc.) et les affiche sous forme de visualisation graphique. Les relations sont trouvées par un algorithme qui peut être appliqué à de grandes bases de connaissances RDF, comme DBpedia. Le graphe qui permet de visualiser les relations pouvant être très important, l'interface utilisateur comporte des fonctionnalités de filtrage pour réduire l'affichage des nœuds et faciliter ainsi la compréhension des relations entre les objets.

Talis (<http://talis-systems.com/>) est une entreprise privée britannique qui a développé une plateforme permettant le stockage de données liées en RDF, ainsi que de nombreux outils pour y accéder et les publier sur le web ou à l'intérieur d'une structure fermée. Les clients principaux de cette société sont des universités et des bibliothèques pour lesquelles des services d'indexation des cours et des ressources ont été développés sur la base de la plateforme Talis. **Link Sailor** (<http://www.linksailor.com>) est un site pour permettre aux humains de naviguer dans les données liées par Talis.

Viewshare (<http://viewshare.org/>) est une application de la Bibliothèque du Congrès permettant de personnaliser la présentation des données numériques (cartes interactives, frises chronologiques, nuages de mots-clés...). Ce projet fait partie du Programme national d'infrastructure et de préservation de l'information numérique (NDIIPP), qui diffuse par ailleurs de nombreuses ressources en ligne (collections, publications et présentations) axées sur la conservation et l'archivage pérenne des contenus numériques.

Plusieurs initiatives anglo-saxonnes se sont engagées dans la voie d'un « web de données antiques » (Ancient World Linked Data), un espace global de données sur l'Antiquité interconnectées à l'échelle du Web selon les principes définis par le W3C. Cet intérêt grandissant pour le web sémantique se traduit par des initiatives comme l' « Institut du web de données antiques » (Linked Ancient World

Data Institute ou **LAWDI**), un séminaire qui se veut un lieu d'échanges et de formation entre praticiens. Un des objectifs est d'étendre la communauté de personnes impliquées dans la création de ressources numériques sur l'Antiquité utilisant les principes et méthodes des Linked data. Il s'agit aussi d'engager le dialogue sur ces questions entre chercheurs, experts techniques, développeurs, bibliothécaires, professionnels des musées et des archives. La première session du séminaire s'est tenue à l'université de New York du 31 mai au 2 juin dernier. Pour en savoir plus : http://wiki.digitalclassicist.org/Linked_Ancient_World_Data_Institute

Vocabulaires de l'Institut Getty et Données liées (Patricia Harpring, rédacteur en chef du programme Vocabulaires du Getty, Institut de recherche Getty)
<http://files.archivists.org/conference/sandiego2012/401-Harpring.pdf>

=> [Sommaire](#)

Dans cet exposé, Patricia Harpring a donné un aperçu des projets de l'Institut Getty pour transformer les vocabulaires existants en Linked Data. Bien qu'ils ne soient pas actuellement mis en relation les uns avec les autres, les vocabulaires du Getty pourraient être exprimés en RDF. Chacun des termes dans les thésaurus du Getty est identifié par un URN (littéralement « nom uniforme de ressource »), qui peut servir de lien pérenne pour le terme. Getty est en train d'expérimenter ses thésaurus en langage SKOS (Système simple d'organisation des connaissances) et d'étudier les formats de données liées les plus pertinents pour ses communautés.

Patricia Harpring a montré à travers des exemples comment les quatre vocabulaires Getty (Thésaurus d'art et d'architecture [AAT], Thésaurus des noms de lieux [TGN], Liste collective de noms d'artistes [ULAN] et Formes autorisées de noms pour les objets culturels [CONA]) supportaient les relations complexes comme les associations hiérarchiques, les événements ou les contextes géographiques. Actuellement, CONA est le seul vocabulaire contenant des liens vers les trois autres. Toutefois, en raison de l'architecture existante et des liens d'association présents dans tous les thésaurus Getty, l'expression de relations complexes et d'associations hiérarchiques sera possible. Pour en savoir plus :
http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf

Archives et Données liées : nos outils sont-ils capables de « compléter l'histoire » ? (Anila Angjeli, Bibliothèque Nationale de France)

=> [Sommaire](#)

Cet exposé visait à présenter quelques observations d'une bibliothécaire spécialiste des métadonnées et des normes qui participe depuis de nombreuses années à des projets communs avec des archivistes (notamment au sein des groupes d'experts AFNOR EAD et EAC-CPF ou en tant que co-présidente du TS-EAC-CPF).

Dans le domaine du patrimoine culturel les Données liées ouvrent une possibilité sans précédent pour relier les informations relatives aux mêmes événements, personnes, lieux, etc., quelles soient conservées dans les collections d'une même institution ou éparpillées aux quatre coins du monde. Car, n'oublions pas que, dans l'absolu, chaque institution ne détient qu'une partie de « l'histoire » de ces entités, qu'il s'agisse de documents publiés, de documents d'archives ou d'autres informations, sous n'importe quelle forme ou format (images, vidéos, manuscrits, autres artefacts ...). Or, ce n'est qu'en connectant les différents témoignages que l'histoire devient vivante, elle se reconstruit, se complète pièce par pièce, prend forme et sens, se raconte, s'interprète et se réinterprète. Par exemple, le site du **Mémorial australien de la guerre** (<http://www.awm.gov.au>), qui est à la fois un grand musée, un centre de recherches et un mémorial, met en relation et enrichit des archives et des bases de données biographiques, d'objets et de photographies directement consultables en ligne à partir d'une recherche thématique collections en ligne, afin de construire une plate-forme pour raconter l'histoire. Grâce à l'outil de gestion de contenu Drupal 7 et aux Données liées, le Mémorial prévoit de développer des outils que les concepteurs, les chercheurs et les historiens pourront utiliser

pour trouver de nouvelles méthodes de construction de récits historiques. Le message transmis est « En réalité, tout est interconnecté ; les « objets » vivent dans un « graphe ».

Un autre exemple de construction de l'histoire collective grâce aux Données liées est celui de **Historypin** (<http://www.historypin.com/>). Créé par l'agence britannique « We Are What We Do » en partenariat avec Google, Historypin est un service contributif où les internautes peuvent ajouter à leur tour des photos, des vidéos, des sons et des anecdotes en lien avec des lieux qu'ils ont visités. Historypin utilise la carte de Google et le système de navigation virtuel Street View pour associer des anciennes photos à des endroits précis afin de comparer l'avant et l'après et découvrir les histoires en rapport avec ces lieux. Ce projet permet de voyager dans le passé et de découvrir l'histoire ainsi que l'évolution des lieux qui nous entourent depuis 1840 jusqu'à nos jours. Le visiteur du site a la forte impression de voir de ses propres yeux le processus de construction de l'histoire en action.

Dans ce sens les archives sont un des types d'institutions patrimoniales qui détiennent des vastes quantités de documents historiques. Mettre en lumière les interconnexions entre les personnes, événements et tout ce dont ces documents témoignent représente un défi. C'est la mission de la description archivistique que d'identifier et d'expliquer le contexte et le contenu des documents d'archives afin de favoriser leur accessibilité et leur utilisation. Ce résultat est atteint en créant des représentations précises et appropriées par la saisie, l'analyse, l'organisation et l'enregistrement des informations sur les documents d'archives et le contextes de leur création, de leur gestion, de leur utilisation et de leur conservation. Les relations entre les documents sont le principal contexte des archives et doivent être préservées et rendues explicites dans les descriptions archivistiques afin de comprendre et d'interpréter les archives et de maintenir leur valeur probante dans la durée. Les normes archivistiques, qu'il s'agisse de normes de contenu (normes internationales du Conseil international des archives) ou de normes de communication (EAC-CPF récemment développée, EAD en cours de révision) sont élaborées à cet effet.

Les archivistes doivent donc adopter une approche proactive pour favoriser l'accès aux archives :

- en révélant l'existence des archives, en les décrivant ;
- en rendant les instruments de recherche disponibles sur Internet, facilement interrogeables ;
- en révélant les trésors cachés, en donnant des informations sur leur contexte.

Cette approche sous-entend le suivi et même l'anticipation des évolutions du contexte technologique qui favorise l'émergence de nouveaux besoins. Dans ce sens les normes archivistiques sont susceptibles d'évoluer pour s'adapter à ces changements.

L'EAC-CPF témoigne de cette évolution. Elle a constitué une étape importante vers la production de données archivistiques liées et ouvertes. Elle est centrée sur les données, intègre les technologies en vigueur (identifiants, autres espaces de noms, mécanismes de liens...). Une ontologie RDF entièrement basée sur le schéma a été développée à titre d'illustration par la société Regesta.exe en partenariat avec l'Institut des biens culturels d'Emilie-Romagne (avec un outil de navigation au sein de l'ontologie) à titre d'illustration (http://archivi.ibc.regione.emilia-romagna.it/ontology/reference_document/referencedocument.html). Cette ontologie est expérimentée dans le cadre du projet « Réseaux sociaux et contextes archivistiques » (SNAC). Un autre exemple d'application récente de l'EAC-CPF est celui du projet LAWDI (*cf. supra*). Ethan Gruber de la Société américaine de numismatique a récemment posté un message sur le sujet sur la liste EAD. Les membres du projet travaillent à la construction d'un réseau d'empereurs romains et de leurs relations. L'étude préalable de différents formats a conclu que l'EAC-CPF était celui qui répondait le mieux au besoin de décrire et de relier, du point de vue de la numismatique, les empereurs avec leurs familles, avec les gouverneurs de provinces, avec les graveurs anonymes qui sont identifiables par leur style artistique. Un outil de production de notices EAC-CPF a été élaboré (voir : <http://wiki.numismatics.org/xeac:xeac>). Chacune de ces entités peut être liée à des pièces de monnaie, pour commencer, mais on peut également leur associer des sculptures ou des sources épigraphiques ou littéraires. « Tout ainsi devient une question de liens. »

Le projet data.bnf.fr (<http://data.bnf.fr/>) est quant à lui un exemple d'application des principes du web sémantique et du web de données aux différents types d'informations gérées par la Bibliothèque nationale de France (BnF). La perspective choisie est de proposer des « fiches de référence sur les œuvres et les auteurs » en « regroupant sur une même page toutes les informations relatives à ceux-

ci, issues de ses différents catalogues, de ses instruments de recherche, de sa bibliothèque numérique Gallica, des expositions virtuelles, etc. ». Les données produites dans des formats divers, notamment en MARC pour les catalogues de livres, XML-EAD pour les inventaires d'archives, et Dublin Core pour la bibliothèque numérique, sont modélisées, regroupées, enrichies par des traitements automatiques et publiées selon le langage du web sémantique RDF. De cette façon, les ressources de data.bnf.fr sont rendues pleinement exploitables par les machines, et les données sont ouvertes et reliées à d'autres données sur le Web.

Mais l'accent dans cette présentation est mis sur les solutions proposées par data.bnf.fr pour l'exploitation des instruments de recherche en XML-EAD. Avant tout, l'utilisation des URIs est généralisée : génération d'URIs à partir des côtes des documents archivistiques décrits au sein de composants <c>, identifiants ARK des notices d'autorité existantes pour les producteurs de ces documents et, bien entendu, les URIs des instruments de recherche. Des règles d'inférence permettent l'exploitation de la hiérarchie des instruments de recherche en XML-EAD. Par exemple pour les pages « auteurs » : détecter des URIs pour les autorités dans l'instrument de recherche, remonter l'arbre XML de l'EAD, moissonner l'information appropriée aux niveaux supérieurs de la hiérarchie, combiner l'information et recréer l'héritage pour l'affichage.

Les perspectives ouvertes par les technologies du Web sémantique résonnent avec les objectifs des archives. C'est ainsi que l'on voit se multiplier les projets s'appliquant à des ensembles de données archivistiques, utilisant les technologies du web sémantique et des données liées ouvertes, comme le projet **LOCAH** (<http://blogs.ukoln.ac.uk/locah/about/>), visant à rendre disponibles en tant que Linked Data les notices descriptives de deux catalogues collectifs, Archives Hub et COPAC (cf. *infra*), ou encore **TROVE** (<http://trove.nla.gov.au/>), l'interface de recherche de la Bibliothèque nationale d'Australie, permettant d'interroger une large agrégation de ressources australiennes. Le contenu cherchable dans TROVE provient de plus de 1000 bibliothèques, musées, archives et couvre 90 millions de références. Le contenu inclut des livres en texte intégral, des revues, des journaux, des images, de la musique, des sons, des vidéos, des cartes, des sites web, des archives, etc. Les services principaux de cet outil proposent une recherche contextualisée et des fonctionnalités dérivées des réseaux sociaux.

On constate que dans les expériences avec les technologies du Web sémantique et les principes des Données Liées les acteurs élaborent des ontologies et modèles de données, s'appuyant partiellement sur des modèles conceptuels et des ontologies existantes et ajoutant des classes et propriétés nécessaires à leurs projets spécifiques. Il est évident à travers toutes ces expériences qu'un modèle de référence pour le monde des archives fait défaut.

Ayant pris conscience de ce besoin pressant, le Conseil international des archives est en train d'initier le travail pour l'élaboration d'un tel modèle conceptuel.

Données liées en chaîne et mixage des collections des archives, des bibliothèques et des musées (Perian Sully, gestionnaire de projet, Projet de collaboration en ligne du Parc Balboa)
<http://files.archivists.org/conference/sandiego2012/401-Sully.pdf>

Le Projet de collaboration en ligne du Parc Balboa (Balboa Park Online Collaborative ou BPOC ; <http://www.balboapark.org/>) a été lancé en 2009 comme un projet de trois ans visant à résoudre les problèmes technologiques rencontrés par 28 institutions d'art et culturelles à San Diego. Il s'agit de développer un modèle collaboratif pour permettre à ces 28 institutions de partager leurs ressources et leurs connaissances.

Les objectifs sont les suivants : faciliter un changement fondamental dans la façon dont les musées utilisent les technologies pour accomplir leurs missions, interagir avec les clients et collaborer ; améliorer leurs capacités technologiques, tout en réduisant les coûts et en incitant à la mutualisation de projets des institutions ayant des besoins similaires ; laisser aux petits établissements la possibilité d'avoir des systèmes technologiques de la même qualité que les grandes institutions grâce à des produits innovants et à des services abordables. C'est ainsi que dans le cadre de ce projet, environ 250 000 images et vidéos ont été numérisées pour 22 musées, bibliothèques et archives sur trois ans, ce qui représente un total de 430 000 documents numériques, décrits dans six

systèmes d'information différents. Un programme de collecte et d'homogénéisation des métadonnées a été mis en place.

Session 505 – Résoudre nos problèmes avec le contrôle d'autorité et le partage : développements récents et perspectives

=> [Sommaire](#)

Trois projets visant à transformer le paysage de la description archivistique et de l'accès étaient présentés dans cette session : le projet SNAC et son prolongement, « Mettre en place une infrastructure nationale pour les autorités archivistiques » (cf. supra), ainsi que la conversion en EAC-CPF des notices d'autorité des Archives nationales des États-Unis (National Archives and Records Administration – NARA), particulièrement impliquées dans le projet de Coopérative nationale d'autorités archivistiques. Cette institution pratique depuis 20 ans le catalogage d'autorités. Les données d'autorité ont migré une première fois en 2000 dans le système d'information ARC des Archives nationales. Un nouveau service a été créé en 2011 pour la conversion et la production en XML EAD et EAC-CPF de descriptions archivistiques et de notices d'autorité.

Congrès international des archives, Brisbane (Australie), 21-24 août 2012

=> [Sommaire](#)

Dix jours après le Congrès de la Société des archivistes américains, le Congrès international des archives s'est tenu à Brisbane (Australie). Plusieurs sessions étaient consacrées à la description archivistique, même si on peut déplorer qu'il n'ait été que très peu question de la révision de l'EAD.

Une session était plus particulièrement consacrée à l'approche originale qu'ont les archivistes australiens de la description archivistique. Adrian Cunningham, directeur des archives numériques des Archives de l'État du Queensland, en a rappelé les grands principes. Dans les années 1960, Peter J. Scott et ses collègues du Bureau des Archives du Commonwealth (actuelles Archives nationales d'Australie) ont développé une nouvelle approche de la description archivistique, consistant à séparer les descriptions des producteurs d'archives des descriptions des archives elles-mêmes.

Cette approche, depuis connue sous la dénomination de « système des séries », rejetait les rigidités de l'approche traditionnelle de la description archivistique, qui consistait à combiner les informations contextuelles et les informations sur les archives dans des descriptions hiérarchiques uniques. Scott et ses collègues ont fait valoir que cette méthode traditionnelle rigide ne reflétait pas de manière adéquate les réalités de la production et de l'utilisation des archives dans des environnements de changements administratifs complexes, où la provenance multiple est un phénomène commun.

Le système de Scott a depuis été adopté par tous les services d'archives publics en Australie et en Nouvelle Zélande. Ce système a aussi fondamentalement influencé l'élaboration et l'évolution des normes de description internationales.

Vitor Fonseca (Archives nationales du Brésil) et Claire Sibille de Grimoüard ont ensuite présenté la réflexion en cours au sein du Conseil international des archives pour préparer la révision de l'ensemble des normes de description archivistique : ISAD (G), ISAAR (CPF), ISDIAH et ISDF.

Les normes internationales de description archivistique ont été élaborées à partir de la fin des années 1990. Elles permettent de décrire les fonds et leur contexte de production. L'internationalisation de la normalisation permet par ailleurs d'échanger plus facilement les différentes descriptions.

Cependant, ISAD(G) demeure la norme la plus utilisée, les autres normes étant moins connues et donc beaucoup moins implémentées. Par ailleurs, même si les quatre normes sont complémentaires, elles n'ont pas été toutes conçues au même moment et dans la pratique il est difficile de les utiliser conjointement de manière cohérente. C'est la raison pour laquelle l'ICA a décidé de réviser les normes mais il ne s'agit pas de les réviser une par une. L'idée de ce travail est plutôt de repenser la

description archivistique pour créer un véritable réseau de relations entre les différents types d'entités archivistiques, d'où un projet d'élaboration de modèle conceptuel archivistique. Un tel modèle conceptuel pourrait aider les archivistes à résoudre les questions soulevées par la combinaison des normes descriptives dans un système de description archivistique, l'absence de certains éléments en matière de liens/relations, le manque de clarté des entités à inclure dans un système d'information, etc. Pareillement, la conception et la mise en œuvre d'un modèle conceptuel en matière de description archivistique permettraient d'éliminer la redondance entre les divers éléments des quatre normes, puis de revenir à des règles et d'identifier les éléments uniques et communs, basés sur toutes les relations possibles entre les entités. Ce travail revêt donc une importance particulière pour améliorer la description archivistique.

Daniel Pitti (Institut des technologies avancées dans les humanités, Charlottesville, Virginie) a une nouvelle fois présenté le projet « Réseaux sociaux et contexte archivistique » (Social Networks and Archival Context ou SNAC), qui vise à explorer la possibilité d'extraire les noms d'organismes, de personnes et de familles ainsi que les données historiques en relation que l'on peut trouver dans les descriptions archivistiques, de regrouper ces données dans des notices d'autorité archivistiques et d'utiliser ces notices d'autorité pour fournir aux chercheurs un accès unique à des fonds d'archives dispersés et aux contextes socio-historiques des archives (*cf. supra*). Daniel Pitti a ensuite présenté le projet visant à « Mettre en place une infrastructure nationale pour les autorités archivistiques » à travers l'organisation d'un cycle d'ateliers régionaux de formation à l'EAC-CPF et l'élaboration d'un ensemble de recommandations pour la mise en place d'une Coopérative nationale d'autorités archivistiques.

Jane Stevenson (administratrice du portail d'archives universitaires Archives Hub, université de Manchester) a décrit deux projets autour des Données liées : LOCAH et « Relier les Vies » (Linking Lives). Après des rappels sur les concepts de base du Web de données, Jane a présenté les défis de la transformation des descriptions archivistiques en Linked data.

Le portail Archives Hub (<http://archiveshub.ac.uk>) est un service financé par le JISC (équivalent du TGE Adonis) qui rassemble des descriptions d'archives conservées par les établissements d'enseignement supérieur et les universités du Royaume-Uni. Un des grands atouts de ce portail est la possibilité pour les chercheurs d'établir des connexions. Les utilisateurs peuvent ainsi rechercher des personnes, des organisations, des lieux ou des sujets à travers plus de 27 000 descriptions de collections et des milliers de notices descriptives de séries organiques et d'articles. Une navigation au sein des index est également possible, facilitant ainsi les recherches latérales.

Les Données liées ont trait à la mise en relation de données les unes avec les autres. Plus précisément, elles ont trait à la mise en relation sur le Web de données structurées. En mars 2010 le JISC a lancé un appel à projets pour « exposer le contenu numérique pour l'enseignement et la recherche » sous forme de Données liées. Les administrateurs du portail Archives Hub ont soumis une proposition pour exposer en Linked data les descriptions archivistiques contenues dans le portail. Ce travail devrait permettre aux chercheurs de faire de nouveaux liens entre les diverses sources de contenu, encourageant de nouvelles connexions entre les gens et les événements, pour révéler davantage sur notre histoire et notre société.

Claire Sibille a présenté un projet collaboratif mis en place par l'Association des archivistes français et les Archives de France, qui tend à créer des descriptions normalisées de producteurs d'archives et à proposer aux services d'archives des modèles de notices d'autorité qu'ils pourront réutiliser pour décrire leurs propres producteurs. L'objectif est de mettre à disposition des archives départementales un outil collaboratif, facile à utiliser et interactif. ISAAR(CPF) est la norme de contenu de référence et l'EAC-CPF est le format d'échange, les concepteurs du projet souhaitant offrir aux services d'archives des modèles de notices instanciables qu'ils pourraient importer dans leurs propres systèmes d'information. Le résultat de ce travail est visible à : <https://www.ica-atom.org/aaf/>

Outi Hupaniittu a présenté un projet (intitulé AHAA) lancé en début d'année par le Service des Archives nationales de Finlande (NAS), dont l'objectif est de développer des services descriptifs communs pour le secteur des archives finlandaises. Une partie du projet consiste à développer de nouvelles approches de la description archivistique. Le projet qui sera mené de 2012 à 2014

comprend le développement d'un outil de description (système d'information) pour implémenter le nouveau modèle de description.

En Finlande, la situation actuelle en matière de description archivistique n'est pas satisfaisante. Les règles finnoises actuelles pour la description archivistique sont proches d'ISAD(G). Elles ont été créées pour guider la description des archives traditionnelles conservées dans un service d'archives. De ce fait, elles ne prennent pas en compte les métadonnées des archives électroniques qui sont créées avant leur transfert dans un service d'archives. Elles ignorent la description de nombreux types d'archives, en particulier celles nées dans un environnement électronique, comme les bases de données. Elles ne sont pas adaptées au niveau de description « pièce » des documents numériques. Enfin, elles ne disent rien sur les descriptions créées par les utilisateurs et les folksonomies.

Pour répondre à ces défis, un nouveau modèle conceptuel pour la description archivistique, similaire au modèle FRBR (Pré-requis fonctionnels des notices bibliographiques) – Functional Requirements of Bibliographic Records) est en cours de développement. La nouvelle norme nationale de description, tendant à être compatible avec les normes du Conseil international des archives, s'appuiera sur ce modèle conceptuel. La présentation d'Outi Hupaniittu, intitulée « Vers une nouvelle ère de description », a décrit les prémisses et les premiers résultats de ce projet.

Enfin, ICA-AtoM (Access to Memory) est un logiciel libre open source développé pour le Conseil international des archives, entièrement basé sur le web, doté d'une interface multilingue, et déjà largement utilisé dans le monde. Il permet à la fois de produire et de publier sur Internet des descriptions archivistiques conformes aux quatre normes de description internationales. Il permet également d'exporter et d'importer des données dans des formats standardisés (EAD, EAC-CPF, SKOS, Dublin Core, MODS XML), de lier des images numériques aux descriptions archivistiques. Il comprend enfin, depuis sa version 1.2, des fonctionnalités d'enregistrement des entrées et de gestion des localisations. La version 1.3 vient d'être rendue disponible (<https://www.ica-atom.org/>). Le projet a fait l'objet d'une présentation générale par Jean-François Moufflet et Claire Sibille (Archives de France), ainsi que de plusieurs ateliers, dont un qu'ils ont assuré en anglais pour des utilisateurs plus avancés du logiciel, venant du Royaume-Uni, du Japon, de Pologne, de Slovénie, d'Australie, de Malaisie, de Nouvelle-Guinée, de l'OTAN et de la Bank World Group, preuve s'il en était de l'intérêt suscité par cet outil au sein de la communauté internationale.