

Zeitschrift: Arbido
Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz
Band: - (2012)
Heft: 2: Normen und Standards = Normes et standards = Norme e standard

Artikel: Du Standard d'échange de données pour l'archivage (SEDA) à la Description archivistique encodée (EAD)
Autor: Sibille-de Grimoüard, Claire
DOI: <https://doi.org/10.5169/seals-770097>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 04.05.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

II. Anwendungen / Usages

Du Standard d'échange de données pour l'archivage (SEDA) à la Description archivistique encodée (EAD)

Claire Sibille-de Grimoüard, Service interministériel des Archives de France

Deux chantiers ont été récemment conduits par les Archives de France concernant la gestion électronique des documents d'archives. Le premier visait à intégrer dans la pratique des services d'archives publics la Description archivistique encodée (DTD EAD 2002), tant pour la production de nouveaux instruments de recherche que pour des opérations de conversion rétrospective. Le second chantier portait sur la gestion et la description des archives nativement numériques dès leur production avec l'élaboration d'un format d'échange, le Standard d'échange de données pour l'archivage (SEDA), permettant là aussi l'interopérabilité, mais dans un cadre particulier: celui d'échange de données entre un service d'archives et ses partenaires lors de transactions (transfert, éliminations, voire communication ou restitution).

Les enjeux tournent à présent autour de la prise en compte par les logiciels documentaires de l'archivage électronique, et du rapprochement de l'EAD avec le SEDA, dont une deuxième version a été publiée par les Archives de France en janvier 2010. Comment les informations descriptives enregistrées dans un système d'archivage électronique (SAE) peuvent-elles venir alimenter les outils traditionnels d'accès aux archives, afin que les usagers puissent disposer d'un accès unifié quels que soient les supports et les origines des archives?

Conscientes du besoin d'articulation entre les deux formats, les Archives de France ont travaillé sur la transformation des métadonnées issues du SEDA au format EAD 2002 pour la descrip-

tion des archives historiques. Toutefois, pour le moment, les essais de conversion ne sont pas pleinement satisfaisants du fait des limites de l'EAD dans sa version actuelle. Or, ce format est en cours de révision. Quelles sont les évolutions souhaitables pour permettre un meilleur interfaçage entre les outils de SAE et les progiciels documentaires utilisés par les services d'archives?

La stratégie des Archives de France pour le numérique

Adopter l'EAD pour l'interopérabilité et la pérennité des descriptions archivistiques.

Dès 1999, les Archives de France se sont intéressées à la DTD EAD, développée par la Société des archivistes américains, d'une part en raison de sa compatibilité avec la norme ISAD(G), d'autre part parce qu'il s'agissait d'un format d'échange ouvert et normalisé permettant d'assurer l'interopérabilité et la pérennité des instruments de recherche électroniques¹.

13 ans après les premiers tests d'encodage, force est de constater que la prise en charge des normes ISAD(G) et ISAAR(CPF) et des langages d'encodage des descriptions archivistiques (EAD, EAC-CPF) est satisfaisante et assurée par la majorité des solutions utilisées dans les services publics d'archives. Ceux-ci sont familiers avec le langage XML, et par ailleurs, en 2010, 17 services d'archives départementales utilisaient un éditeur XML pour élaborer directement ou retravailler un instrument de recherche en EAD. Une spécificité française est le développement de solutions originales, libres ou propriétaires, pour la production et la publication sur Internet des descriptions archivistiques encodées. La plupart des services ont fait le choix de

solutions visant à la compatibilité des différents outils de production et de diffusion des instruments de recherche. D'autres ont fait l'acquisition d'outils prenant en charge toute la chaîne de la description archivistique, depuis la conception de l'instrument de recherche jusqu'à sa publication sur Internet. Quelques-uns enfin ont recouru à la sous-traitance auprès de sociétés spécialisées, avec des crédits spécifiques, dans le cadre de projets importants de conversion rétrospective en EAD.

Mettre en œuvre l'EAD, c'est aussi permettre le moissonnage des composantes d'instruments de recherche via le protocole OAI-PMH. C'est ainsi que les archives françaises participent au portail archivistique européen APE-net², pour lequel, comme tous les partenaires du projet, elles doivent fournir leurs données sous la forme de fichiers XML/EAD, quel que soit le format original de ces données. Avec l'édition, en mars 2010, du schéma EAC-CPF pour la description des producteurs d'archives³, elles commencent également à prendre conscience que se référer aux formes normalisées de valeur nationale devient la condition d'interopérabilité avec d'autres informations lorsque les descriptions archivistiques (instruments de recherche) et l'information contextuelle (fichiers d'autorités archivistiques) produites au niveau national se trouvent portées au niveau international dans des projets comme APEnet.

Toutefois, la DTD EAD s'applique de fait à la description structurée de fonds d'archives «traditionnels», dans leur très grande majorité sur support papier, même si certains d'entre eux ont fait par ailleurs l'objet d'opérations mas-

sives de numérisation. Au-delà de la description des documents, on commence donc à envisager l'intégration et la consultation d'objets numériques, d'où le second chantier lancé par les Archives de France à partir de 2005.

Développer un format d'échange pour les archives nées numériques

Comme ses voisins européens, la France s'est engagée dans le développement de l'administration électronique, depuis la fin des années 1990. La dématérialisation des documents est plus ancienne puisque sont conservées aux Archives nationales des bases de données qui remontent aux années 1970. Mais l'administration électronique a entraîné une dimension supplémentaire avec la dématérialisation des processus administratifs qui peuvent se réaliser de manière automatique, avec les interconnexions entre différents systèmes d'information. Il est donc nécessaire de pouvoir échanger facilement des informations électroniques produites par des outils différents, ce qui renvoie à l'interopérabilité.

La législation française a soutenu ce processus par toute une série de textes, avec notamment une loi qui, en 2000, a reconnu à l'écrit représenté sous forme numérique la même valeur probante que les documents écrits, un décret d'application en 2001 pour la mise en place de la signature électronique et des lois et ordonnances qui, en 2004 et 2005, ont donné une assise aux transactions passées sous forme dématérialisée, via des systèmes d'information dont on se portait garant de la sécurité. La sécurité des systèmes d'information et l'interopérabilité des données qu'ils produisent et s'échangent ont trouvé un aboutissement réglementaire dans la production d'un référentiel général d'interopérabilité (RGI), publié en novembre 2009, et d'un référentiel général de sécurité (RGS), publié en février 2010⁴.

C'est dans ce contexte que les Archives de France ont développé le SEDA, en collaboration avec la Direction générale de modernisation de l'Etat (DGME) en charge de l'administration Budget⁵. Ce travail s'est fait en dehors des organismes de normalisation classiques (AFNOR, ISO), mais en suivant la mé-

thodologie de l'un de ces derniers: l'UN/CEFACT (United Nation/Centre for Trade Facilitation and Electronic Business/Centre pour la facilitation des Procédures commerciales et le Commerce Électronique). Cette méthodologie a conduit à réexprimer des référentiels métiers dans une syntaxe plus conforme à celle préconisée par l'UN/CEFACT. Ainsi, si la description des données est conforme à ISAD(G), elle n'a pu être exprimée directement en EAD, et ce pour plusieurs raisons: choix d'élaborer un schéma XML (technologie très majoritairement choisie après la première période des DTD en raison de contrôles plus précis qu'elle permet), règles de nommage normalisées (interdisant entre autres les acronymes), réutilisation de bibliothèques normalisées d'éléments communs, éléments et attributs supplémentaires à intégrer en raison de la spécificité du traitement des documents d'archives numériques (métadonnées techniques notamment, mais également métadonnées liées à l'intégrité, à la traçabilité ...). Les attributs de niveaux de l'EAD ont par contre été repris pour décrire les niveaux archivistiques des objets échangés (fonds, series, file, recordgroup, item).

Avec la publication en 2009 du RGI dans lequel le SEDA est cité comme «recommandé» pour les échanges dans le cadre de l'archivage, c'est le SEDA lui-même en tant que syntaxe et organisation du dialogue qui a pris le statut de référentiel⁶.

Développer des outils de conversion du SEDA vers l'EAD

Pourquoi un rapprochement SEDA-EAD?
Il s'agit à présent de réfléchir à la prise en compte de l'archivage électronique par les logiciels documentaires utilisés dans les services d'archives. Cette transformation du SEDA en EAD est rendue d'autant plus nécessaire qu'à terme les logiciels jusqu'à présent utilisés pour la gestion et la description des archives sous forme papier devront permettre d'effectuer une recherche dans les informations descriptives de données conservées dans un système d'archivage électronique.

Les éditeurs s'emparent d'ailleurs eux-mêmes de la problématique. C'est ainsi

que les journées du club utilisateurs du logiciel Avenio, tenues en juin 2010 à Metz, ont rassemblé des archivistes municipaux et départementaux autour de la question de la description des archives électroniques. De son côté, la société Naoned, éditeur du logiciel Mnesys, initialement outil de gestion et de description d'archives papier, s'est engagée dans des développements pour la prise en compte de la collecte et l'archivage des données pour que la solution, adossée à une architecture technique adéquate, permette une gestion unifiée des documents, quel que soit leur support. Quant à l'Association des Développeurs et Utilisateurs de Logiciels Libres pour l'Administration et les Collectivités Territoriales, elle travaille à la mise en œuvre d'un couplage entre l'outil de SAE qu'elle développe, as@lae, et des outils utilisés par des services territoriaux d'archives pour produire des descriptions XML/EAD (Arkhéia) et les publier sur Internet (Pleade). Enfin, les Conseils généraux de l'Aube et des Yvelines s'interrogent sur l'interfaçage de leur outil de gestion documentaire (Thot) et de leur solution d'archivage électronique (M@rine).

Tous ces projets posent plus généralement la question du positionnement des services d'archives. Qui est responsable du SAE? Est-ce le service d'archives ou un tiers-archiviste? Les plates-

1 Pour en savoir plus: www.archivesdefrance.culture.gouv.fr/gerer/classement/normes-outils/ead/

2 Voir le site du projet: www.apenet.eu/

3 Le site officiel de l'EAC-CPF est accessible à: <http://eac.staatsbibliothek-berlin.de/>

4 Les documents de référence de l'administration électronique sont accessibles à: <http://references.modernisation.gouv.fr/>

5 Dans le but de faciliter l'utilisation de ce standard par les professionnels et les applications, les Archives de France ont mis à disposition des ressources telles que des feuilles de styles, des outils d'édition, de la documentation ainsi que les schémas XML. L'ensemble de ces ressources sont accessibles à l'adresse: www.archivesdefrance.culture.gouv.fr/seda/

6 Le SEDA fait partie du volet sémantique du RGI: <http://references.modernisation.gouv.fr/volet-semantique>

formes d'archivage électronique ne doivent-elles pas rester centrées sur leur cœur, à savoir la prise en charge et la gestion du cycle de vie des données électroniques, incluant leur intégrité, leur lisibilité et leur authenticité, les progiciels documentaires étant l'interface pour la recherche et la communication des documents, qu'ils soient électroniques ou papier? La description documentaire enregistrée dans la plateforme d'archivage électronique doit-elle être complétée et modifiée dans cette dernière? Ne doit-elle pas plutôt être exportée vers le progiciel documentaire, où se feront les enrichissements nécessaires (production d'instruments de recherche, indexation)? Par ailleurs, où la conversion des métadonnées descriptives au format SEDA en EAD peut-elle intervenir? En sortie du SAE? En entrée du progiciel documentaire? Or, tous les logiciels métier utilisés par les services d'archives ne permettent pas d'importer de l'EAD, qui est encore souvent considérée plutôt comme un format de sortie pour la diffusion des instruments de recherche. Enfin, comment résoudre les différences de granularité de l'EAD qui permet une description à plusieurs niveaux (l'objet dans son ensemble puis chacune de ses parties et sous-parties, sans limitation) et du SEDA, qui permet l'atomisation des versements?

L'outil de conversion et ses limites

Pour répondre à ce besoin, les Archives de France ont réfléchi aux concordances possibles entre les deux formats et ont développé en interne un outil de conversion du SEDA en EAD⁷. Pour cela, elles se sont appuyées sur un premier travail de mise en correspondance de ces éléments, qui avait été publié dans la documentation relative à la première version (0.1) du SEDA. Cette première base a été revue pour prendre en compte la nouvelle version du SEDA (0.2). En septembre 2010, les Archives de France ont réuni les principaux éditeurs du mar-

ché pour leur faire part de ces travaux, qui se sont heurtés à plusieurs difficultés, dues à la différence des besoins adressés par les deux formats et aux limites de l'EAD dans sa version actuelle.

Le SEDA est un format d'échange qui modélise des processus ou transactions entre différents acteurs, tandis que l'EAD a été conçue pour la publication sur Internet d'instruments de recherche archivistiques au profit de chercheurs. Si des éléments de description peuvent être rapprochés, les besoins de recherche dans un SAE sont très différents de ceux d'un chercheur sur des archives définitives. Les informations de gestion (flux, niveau de services offerts par le système d'archivage, signature, etc.) sont moins utiles pour des archives historiques.

Les possibilités de précisions offertes par l'EAD en matière de description archivistique sont plus importantes que ce que propose le schéma «Archive» du SEDA. Dans ces conditions, les 146 balises de l'EAD n'ont pas été toutes prises en compte. Inversement, certaines balises du SEDA (sur le type MIME, le poids des fichiers, etc.) n'ont pas d'équivalents exacts en EAD. Pour certaines d'entre elles, on s'est contenté de regrouper leurs données dans des balises EAD Note <note> ou Autres informations complémentaires <odd> pour perdre le moins d'informations possible. Plus généralement, il a été décidé de ne pas viser l'exhaustivité, mais de mentionner les seuls éléments du SEDA utiles pour les instruments de recherche. Outre l'outil de conversion, un fichier des équivalences jugées utiles et correspondant le mieux à la pratique archivistique a été établi. Concernant la différence de granularité entre les deux formats, il a été choisi de pousser la description des instruments de recherche jusqu'au niveau le plus fin du SEDA (la pièce attachée) afin de ne perdre aucune information, la lisibilité de l'instrument de recherche pouvant alors être paramétrée par les outils de publication en fonction du choix des informations à afficher.

Plus fondamentalement, l'EAD est une DTD, alors que le SEDA se présente comme un ensemble de schémas XML

W3C. Comme toutes les DTD, l'EAD contraint peu les contenus. Elle ne supporte pas non plus les «espaces de nom». En pratique, cela implique qu'il n'est pas possible d'importer des définitions de balises définies dans une autre DTD ou dans un autre schéma XML. Par ailleurs, la définition du type «ID» n'étant pas strictement identique dans le langage des DTD et dans celui des schémas XML, il s'ensuit que certains identifiants du SEDA ne peuvent être transposés littéralement en EAD. Par exemple, le premier caractère d'un attribut de ce type peut commencer par un chiffre dans le SEDA, alors que cela est interdit en EAD. Enfin, l'EAD est centrée sur le document et pas assez sur les données. Elle mélange les éléments structurant l'information et les éléments de mise en forme (titres de rubriques, paragraphes, listes, tableaux, etc.). L'EAD a été conçue comme une transposition des habitudes issues du papier, elle se concentre sur un mode de représentation de l'information, le document structuré. La révision prochaine de l'EAD devrait permettre de la rendre plus conforme aux règles d'écriture des schémas XML et donc plus compatible avec le SEDA.

Les évolutions attendues de l'EAD et leurs incidences sur le SEDA

Le processus de révision de l'EAD par le sous-comité technique EAD (TS-EAD) de la Société des archivistes américains (SAA) a commencé en août 2010. Les travaux ont commencé par le lancement d'un appel à commentaires auprès de la communauté internationale. Même si les 147 propositions d'évolution reçues par la SAA sont en cours d'examen par le TS-EAD, quelques décisions de principe ont déjà été prises qui pourraient impacter l'évolution du SEDA:

- développement d'un processus de migration des instruments de recherche encodés en EAD 2002 vers le nouveau format;
- interopérabilité (mieux prendre en compte les sémantiques des relations et simplifier les mécanismes de liens);
- échange (simplifier l'EAD, éviter les ambiguïtés et favoriser une utilisation plus prévisible du format).

7 Note d'information DGP/SIAF/2010/024 en date du 13 janvier 2011 relative à la transformation du SEDA au format EAD (www.archivesdefrance.culture.gouv.fr/gerer/archives-electroniques/standard/seda/)

L'évolution de l'EAD vers un schéma XML, qui sera la version de référence, est désormais acquise. Toutefois, la SAA prévoit d'utiliser la technologie RelaxNG plutôt que celle du W3C XML-Schema qui a été utilisée pour le SEDA. Elle souhaite également diffuser des versions dérivées de la version RelaxNG (W3C XML-Schéma et DTD). Or il sera difficile pour un travail de modélisation qui exploite toutes les possibilités d'une technologie (RelaxNG) de le diffuser dans une autre technologie (XML-Schéma ou DTD) sans en appauvrir l'expression. Même si ces trois technologies tendent à spécifier des contraintes vérifiables, elles ne le font pas de la même manière.

La transformation de l'EAD en schéma devrait permettre aux utilisateurs du format d'inclure le schéma EAD dans d'autres schémas ou d'inclure d'autres DTD ou d'autres schémas dans l'EAD. Par exemple, il serait possible de rapa-

trier dans l'élément «Description physique» de l'EAD des informations structurées au format SEDA sur les caractéristiques physiques des fichiers informatiques. Toutefois, le TS-EAD et l'équipe technique de développement semblent pour le moment avoir une approche restrictive. Il est question de limiter l'inclusion des éléments et attributs provenant d'autres espaces de noms à ceux des espaces de noms gérés par la SAA (actuellement EAC-CPF et EAD) ou peut-être, ceux qui sont largement adoptés (par exemple, xml:id et xml:lang), et d'aligner étroitement les éléments et attributs EAD sur les éléments et attributs d'autres espaces de noms externes (par exemple MODS), mais de les laisser au sein de l'espace de nom EAD. Il serait souhaitable de prévoir une ouverture un peu plus large vers les autres espaces de nommage. Un compromis pourrait être de confiner leur utilisation à des emplacements réservés à cet usage, voire de limiter la

responsabilité et l'engagement de service que l'EAD pourrait prendre sur ces emplacements (par exemple ces informations seraient validées et conservées, mais leur sens échapperait à l'EAD).

Conclusion

La table de correspondances tout comme la feuille de styles élaborée par les Archives de France correspondent à une première étape de traitement qu'il sera sans doute nécessaire de compléter, dans la continuité de la politique des Archives de France visant à favoriser le développement de systèmes d'information utilisant des formats d'échange qui reposent sur des normes et des recommandations internationales. L'un des objectifs majeurs de la transformation de l'EAD en schéma XML est précisément son interopérabilité avec d'autres schémas: cela recoupe l'utilisation d'autres schémas dans l'EAD et l'utilisation de l'EAD dans d'autres schémas, notamment le SEDA.

Les évolutions de l'EAD devraient donc impacter le SEDA. Elles sont d'autant plus essentielles qu'il est nécessaire d'«intégrer la description archivistique dans le processus continu de préservation», comme l'a rappelé le groupe InterPARES de Luciana Duranti, afin de «fournir une attestation d'ensemble de l'authenticité des documents d'archives électroniques et de leur relation avec les autres documents dans le contexte du fonds auquel ils appartiennent»⁸.

Contact: claire.sibille@culture.gouv.fr

8 Rapport du groupe de travail sur la stratégie (www.interpares.org/ip1/ip1_documents.cfm?cat=stf)

ABSTRACT

Vom Standard zum Austausch für Archivdaten (Standard d'échange de données pour l'archivage, SEDA) hin zum XML-Format zur Beschreibung von Archivalien (Encoded Archival Description, EAD)

Während der letzten Jahre beschäftigte sich das Nationalarchiv Frankreichs mit zwei Tätigkeitsfeldern der elektronischen Verwaltung von Archivalien. Das erste Gebiet befasste sich mit der Einführung von XML-standardisierten Beschreibungsdaten für Archivzwecke (DTD EAD 2002) in die praktische Arbeit der staatlichen Archive. Das zweite Gebiet befasste sich mit der Verwaltung und Beschreibung von seit seiner Entstehung digitalem Archivgut («digital born») und der Erarbeitung eines Standards zum Austausch von Archivdaten (SEDA). Dieser Standard soll den Datenaustausch zwischen Archiven und den ihnen vorgelagerten Stellen ermöglichen (zum Zwecke des tatsächlichen Datentransfers, der Kassation oder der Herausgabe von Daten). Die Herausforderung besteht darin, dass die Nutzer trotz verschiedener Software zur Archivierung und Verwaltung elektronischer Dokumente über einen einheitlichen Zugang verfügen, egal welchen Ursprungs die entsprechenden Archive sind oder um welche Datenträger es sich handelt.

(fs, sh)