

Documentation EAD

M. Accardo, J. Fan, N. Grim, C. Grometto, L. T. T. Nguyen

Mars 2024

Table des matières

1	Propos liminaire	2
2	Choix des balises	3
2.1	Encoder l' <code><eadheader></code>	3
2.1.1	Le <code><filedesc></code>	3
2.1.2	Le <code><profiledesc></code>	4
2.1.3	Le <code><revisiondesc></code>	4
2.2	Encoder le <code><frontmatter></code>	5
2.3	Encoder l' <code><archdesc></code>	5
2.3.1	Le <code><did></code>	5
2.3.2	Le <code><scopecontent></code> et l' <code><arrangement></code>	6
2.3.3	Éléments facultatifs	6
2.3.4	Sur le <code><controlaccess></code>	7
3	Objectifs de la mise en ligne	9
4	Adoption de nouvelles pratiques	10
5	Insuffisances observées	11
5.1	Les tables de concordance	11
5.2	Les index	11
5.3	Le sommaire	12
5.4	Les attributs	12

Propos liminaire

Le présent document a pour ambition de fournir une documentation détaillée à propos des instruments de recherche que notre service a dernièrement encodés. Nous y décrivons le choix des éléments majeurs selon un ordre qui suit l'arborescence des fichiers XML-EAD également mis à disposition. Une seconde partie est consacrée à l'explicitation de notre projet d'encodage, en parallèle à l'adoption de nouvelles pratiques descriptives.

Choix des balises

2.1 Encoder l' `<eadheader>`

Il est obligatoire de renseigner l'élément `eadid`.

2.1.1 Le `<filedesc>`

Le `<titlestmt>`

L'élément `<titlestmt>` est obligatoire. En son sein, il convient de renseigner le `<titleproper>`. Nous renseignerons les cotes extrêmes dans l'élément `<subtitle>`, lequel accueillera une prose plus ou moins verbeuse :

```
<subtitle>Répertoire numérique de la partie de sous-fond séries 1232 W  
1-69</subtitle>
```

L'auteur de l'instrument de recherche est à renseigner dans la balise `<author>`.

L' `<editionstmt>`

Cet élément contient des informations concernant l'édition électronique en cours de rédaction. Il convient d'uniformiser son contenu :

```
<editionstmt>  
  <edition>première édition électronique</edition>  
</editionstmt>
```

Le `<publicationstmt>`

Il convient de renseigner le nom du service d'archive responsable dans l'élément `<publisher>`, et les informations sur la localisation dans l'élément `<addressline>`, lui-même enfant de l'élément `<address>`.

Pour éviter toute ambiguïté, nous conserverons uniquement la date de la première édition de l'instrument de recherche dans l'élément `<date>`. Les informations relatives aux révisions de l'instrument se trouveront dans l'élément `<revisiondesc>`. Celles à propos du présent rétroencodage en XML-EAD seront renseignées dans l'élément `<editionstmt>`.

2.1.2 Le `<profiledesc>`

Il conviendra d'uniformiser le `<profiledesc>` de la manière suivante :

```
<profiledesc>
  <creation>
    Cet instrument de recherche a été produit au moyen du
      logiciel oXygen le <date>05 mars 2024.</date>
  </creation>
  <language>
    Instrument de recherche rédigé en <language
      langcode="fre">français</language>
  </language>
  <descrules>
    Encodage en XML-EAD (version 2002). Conforme à la norme
      ISAD-G.
  </descrules>
</profiledesc>
```

2.1.3 Le `<revisiondesc>`

Les mentions de révision de l'instrument de recherche peuvent être encodées dans l'élément `<revisiondesc>`. Il faudra utiliser l'élément `<change>` pour

renseigner une unique information sur un changement substantiel apporté à l'instrument de recherche encodé. Chaque élément `<change>` contiendra un élément `<date>` et un élément `<item>`, lesquels indiquent respectivement la date et la nature des changements.

2.2 Encoder le `<frontmatter>`

Il est possible de produire, à la suite du `<eadheader>`, un `<frontmatter>`. Il est réservé aux informations complémentaires sur les métadonnées, sans lien étroit avec la description des documents. Nous pourrions par exemple y encoder un tableau de concordance entre anciennes et nouvelles cotes.

2.3 Encoder l' `<archdesc>`

2.3.1 Le `<did>`

Il faudra obligatoirement renseigner les éléments majeurs du `<did>`, conformément aux recommandations du standard, soit (dans cet ordre) :

```
<unitid>
<unittitle>
<unitdate>
<origination>
```

Nota Bene : si une ancienne cotation est indiquée, l' `<unitid>` contiendra les cotes extrêmes selon ces deux systèmes (ancien et actuel).

L'élément `<unitdate>` devra obligatoirement posséder les trois attributs `@calendar`, `@era`, `@normal` afin de désambiguïser la notation des dates. La normalisation de ces dernières est fondamentale pour correctement paramétrer notre moteur de recherche, lequel se basera sur les informations saisies en valeur d'attributs – non pas sur le contenu des éléments.

Le `<did>` pourra contenir l'élément facultatif `physdesc` ainsi que son enfant `<extent>`, afin d'indiquer le nombre d'articles et le métrage linéaire. Éga-

lement, nous pourrions recourir à l'élément `<repository>`, s'il est possible de préciser le lieu ou l'organisation en charge de la conservation.

2.3.2 Le `<scopecontent>` et l' `<arrangement>`

Il est vivement conseillé de renseigner ces deux éléments selon l'ordre décrit. Tout d'abord, `<scopecontent>` – lequel contiendra une partie en prose pour présenter le contenu de l'instrument, au sein de paragraphes `<p>`. Nous avons préféré cet élément à `<abstract>` afin de ne pas multiplier les éléments et de simplifier la prise en main des diverses règles.

Dans un deuxième paragraphe, l'élément `<scopecontent>` contiendra également une liste, notée `<list>`, composée d' `<item>` afin de rendre compte du sommaire de l'instrument de recherche.

Toujours dans l'optique de simplifier l'encodage, nous limitons la description des parties au premier niveau d'arborescence.

À la suite de l'élément `<scopecontent>` pourra se trouver l'élément `<arrangement>` – lequel décrira en une phrase les modalités de classement :

```
<arrangement>
  <p>Classement par commune</p>
</arrangement>
```

2.3.3 Éléments facultatifs

Si les informations supplémentaires sont fournies, nous recommandons de les encoder au sein des éléments suivant, selon l'ordre et la nomenclature suivants :

`<bioghist>` (présentation du producteur, en prose, si le niveau de description l'exige, il est possible d'ajouter des titres au sein de `<head>`),

`<custodhist>` (historique de conservation, en prose, si le niveau de description l'exige, il est possible d'ajouter des titres au sein de `<head>`),

`<aqinfo>` (modalité et/ou date de versement ainsi que les modalités de consultation si celles-ci ne sont pas soumises à restriction),

`<accessrestrict>` (description des restrictions d'accès des documents et/ou des modalités de consultation),

`<userrestrict>` (description des restrictions d'utilisation et/ou de reproduction des documents),

`<relatedmaterial>` (liste des sources externes liées aux documents décrits, au sein d' `<item>` d'une `<list>` contenue dans un `<p>`),

`<separatedmaterial>` (liste des sources internes au service, ayant un lien avec l'instrument, au sein d' `<item>` d'une `<list>` contenue dans un `<p>`),

`<bibliography>` (informations bibliographiques à propos du producteur ou de l'entité ayant produit les documents décrits par l'instrument. Cette bibliographie sera structurée dans des `<item>` de `<list>` . Le nom de l'auteur sera encodé dans un élément `<name>` et le titre de la référence bibliographique sera encodé au sein de l'élément `<emph>`).

2.3.4 Sur le `<controlaccess>`

L'EAD permet l'indexation des éléments en vue de faciliter la recherche dans le service d'archive et de créer des liens entre plusieurs instruments de recherche. Aussi, pour rendre compte d'entrées d'indexation, nous utiliserons l'élément `<controlaccess>` , au sein duquel nous expliciterons la nature des diverses composantes de l'index :

```
<controlaccess>
  <genreform>Photographies aériennes</genreform>
  <geogname>Morbihan (Bretagne, France)</geogname>
</controlaccess>
```

Les éléments majeurs d'indexation utilisables en tant qu'enfants de l'élément `<controlaccess>` sont :

`<corpname>` (collectivité),

`<famname>` (nom de famille),

`<function>` (activité),

`<genreform>` (genre et caractéristiques physiques),

`<geogname>` (entité géographique),
`<name>` (nom),
`<occupation>` (fonction),
`<persname>` (nom de personne physique),
`<subject>` (mot-matière),
`<title>` (titre).

Le choix du bon élément lors de l’encodage des entrées d’index demande un effort d’uniformisation au fil de la description. Aussi, notre travail d’encodage a soulevé la nécessité de prendre une décision à propos de deux cas :

- Une communauté de communes ou de villes doit être encodée avec l’élément `<corpname>` ,
- Une paroisse doit également être encodée avec l’élément `<corpname>` .

L’utilisation de `<controlaccess>` présente plusieurs avantages. Cette configuration permet d’envisager un enrichissement des données grâce à l’emploi d’attributs dans les balises ouvrantes de ses éléments enfants, à l’instar de liens vers des sites extérieurs, des notices, d’autres instruments de recherche, etc. Le second avantage tient au principe d’héritage : on indexe au niveau le plus haut et les niveaux inférieurs héritent des entrées d’index, cela évite de répéter l’information.

Par ailleurs, nous recommandons une atomisation maximale des informations à indexer, et de ne renseigner qu’une entrée par élément. Cette possibilité nécessite cependant un travail de documentation et/ou de production de notices de producteurs et de fonctions en amont – ce que nous avons été en incapacité de réaliser dans le cadre de notre travail préliminaire.

Objectifs de la mise en ligne

Notre démarche d’encodage vise à favoriser l’accessibilité, la découverte et l’exploitation des instruments mis à disposition. Que l’utilisateur soit un chercheur accoutumé aux modes de description archivistique ou non, l’encodage des index et des tables de concordance nous semble nécessaire à la bonne navigation au sein des instruments. Aussi, dans l’optique de situer les archives dans leur contexte - car un fonds se comprend au travers de son histoire - nous fournissons l’intégralité des informations dont nous disposons dans les éléments `<custodhist>` et `<bioghist>`.

Grâce à ces derniers, nous sommes en mesure d’atomiser les données selon deux catégories. `<custodhist>` nous permet de renseigner des informations sur la garde des documents à l’instar des transferts de propriété, des achats, des dépôts, etc. avant leur entrée dans notre service. `<bioghist>` nous permet quant à lui de transcrire les informations biographiques d’individus ainsi que l’histoire d’entités mentionnées dans l’instrument - exemplairement des services archives, des communes ou des communautés de communes.

L’explicitation des règles de la description archivistique est tout à fait nécessaire pour privilégier l’interopérabilité et l’inter-connectivité des données. Une compatibilité efficiente requiert l’utilisation d’un vocabulaire partagé, l’usage d’index et de référentiels ainsi que le respect de modèles de description. L’enjeu majeur est alors de mettre à disposition des outils répondant aux mêmes règles et permettant l’interopérabilité des données.

Adoption de nouvelles pratiques

Réfléchir à l'adoption de nouvelles pratiques descriptive requiert d'examiner les méthodes existantes dans l'optique d'identifier nos besoins et de cerner les axes d'amélioration. Cela permettra une meilleure définition des domaines dans lesquels nos pratiques pourront être optimisées afin de répondre aux exigences de la gestion de documents.

Sur la base de cette évaluation, nous enrichirons la présente documentation, dans le respect des spécificités de chaque fonds d'archives. Cette étape comprend également l'apprentissage et l'application de nouvelles méthodes *via* la mise à disposition de matériel pédagogique et d'outils pratiques – dont un dictionnaire de balises, un guide d'encodage et des séries d'exemples conformes aux pratiques descriptives du service.

Il nous semble également fondamental d'organiser des journées de formation pour que l'ensemble de l'équipe monte en compétence et en connaissances. Nous tâcherons d'insister sur la nécessité d'emploi de chaque élément pour favoriser l'harmonisation et l'interopérabilité. Nous proposerons, dans le cadre de ces journées, de fournir un modèle, un *template* dûment commenté afin d'au mieux décrire les éléments et la nature du contenu attendu.

Ainsi, à l'issue de la formation, l'équipe en charge de la description archivistique saura mettre en œuvre les nouvelles pratiques dans leur travail quotidien. Nous ajusterons notre documentation et nos recommandations en fonctions des retours et des observations de nos collègues.

Insuffisances observées

Il a été complexe de systématiquement trouver des éléments adaptés pour encoder les informations les plus particulières – pourtant susceptibles de figurer dans un instrument de recherche en archives.

5.1 Les tables de concordance

Un premier exemple frappant concerne les tables de concordance : peut-on les considérer comme des sommaires ou non ? Existe-t-il un élément particulier pour les encoder ? Nos recherches ayant d’abord été infructueuses, nos consultations répétées du dictionnaire des balises ont mis en lumière la solution suivante : produire un `<frontmatter>` à la suite de l’`<eadheader>`. Cet élément accueillera des informations complémentaires sur les métadonnées, sans rapport étroit avec la description des documents.

5.2 Les index

L’encodage des index a également posé problème. Nous avons identifié deux possibilités : utiliser l’élément `<index>` ou l’élément `<controlaccess>`. L’emploi de ce dernier nous a paru plus judicieux dans la mesure où il figurait dans d’autres instruments de recherche où des index étaient encodés, avec un choix d’éléments enfants plus conséquent. Néanmoins, ces éléments enfants mis à disposition n’ont pas toujours été d’une éloquence remarquable pour l’encodage de certains termes (à l’instar d’« urbanisme » ou de « céréales »).

5.3 Le sommaire

La question du positionnement du sommaire a été posée à plus d’une reprise. Comme nous n’avons pas trouvé d’élément saillant dédié à cette partie des documents, nous avons fait le choix de l’encoder au sein de l’élément `<scopecontent>`. Ce choix témoigne de la difficulté d’atomisation de l’information entre les différents éléments : un élément capable d’englober de nombreux éléments enfants rend le choix d’encodage plus complexe.

5.4 Les attributs

Enfin, la faible présence d’attributs au sein de nos travaux traduit un manque de finesse dans la granularité de la description. Au regard des éléments que nous utilisons, une meilleure connaissance des derniers, de leurs attributs et de leur valeur sera bienvenue afin d’au mieux hiérarchiser les informations présentes au sein d’instruments de recherche.