1. Projet Experts: XML-EAC-CPF

Par convention, les éléments seront indiqués entre chevrons (<element>) et les attributs prendront la forme suivante : @attribut.

1.1. Rappels généraux sur XML

XML: eXtensible Markup Language



Langage de balisage / extensible

Un langage de balisage est fondé sur l'emploi de balises simples, et compréhensibles par l'homme, pour décrire du texte

Si nous devions encoder la phrase : 'On emploi *a priori* les italiques pour les locutions et termes empruntés à d'autres langues', la question posée serait de savoir comment décrire la locution latine *a priori*.

Avec XML il ne s'agit pas de dire qu'elle est représentée en italique, mais d'expliquer pourquoi elle est en italique. XML permet donc de *dissocier la forme et le contenu*, la mise en forme étant traitée à part.

Exemple d'encodage en XML :

On emploi <locutionEtrangere>a priori</locutionEtrangere> les italiques pour les locutions et termes empruntés à d'autres langues

XML est également dit *extensible*, c'est à dire qu'il ne propose pas de « jeu de balises » prédéfini que nous pourrions utiliser en toutes circonstances.

En réalité, XML formule un ensemble de règles sur ce que doit être un document *bien formé*, c'est-à-dire qui respecte des règles syntaxiques strictes, et *valide*, c'est-à-dire conforme à un *schéma* ou à une *DTD* (p. ex. le schéma TEI, la DTD EAD, le schéma EAC-CPF, etc.).

Un schéma peut être considéré, entre autre, comme un dictionnaire de balises.

Si nous reprenons le même exemple mais que nous décidons de l'encoder selon la TEI, nous obtiendrons :

On emploi <foreign xml:lang="lat">a priori</foreign> les italiques pour les locutions et termes empruntés à d'autres langues

XML est supporté par le *World Wide Web Consortium* (W3C), et son utilisation a été recommandée en France par la Direction générale de modernisation de l'État dans son *Référentiel général d'interopérabilité* publié en 2009. Ce langage est perçu comme un excellent moyen de sécuriser et de diffuser ses données :

- C'est *un format ouvert*, particulièrement bien documenté et finalement assez simple : il ne s'agit que de texte, rien d'autre. On écarte donc d'emblée les difficultés liées à l'utilisation de formats propriétaires.
- C'est un format qui fait preuve d'une grande *souplesse* d'utilisation. Il est notamment possible de travailler sous plusieurs espaces de noms au sein d'un même document, afin de répondre à des problématiques bien précises.
- Enfin, c'est un excellent *format pivot*, c'est-à-dire qu'il est possible d'exporter ses données XML vers d'autres formats facilement.

1.2. EAC-CPF: Encoded Archival Context - Corporate bodies, Persons, Families

Encoded Archival Context - Corporate bodies, Persons, Families



Contexte archivistique encodé - collectivités, personnes et familles

L'EAC-CPF est un standard pour l'encodage et l'échange de *notices d'autorité*, dans le contexte des documents d'archives. Ce format est fondé sur la norme internationale *ISAAR (CPF)*.

EAC-CPF a été pensé pour être utilisé en complément de l'*EAD*, le standard pour la description archivistique, afin de rendre compte du contexte dans lequel 'les agents ont opéré et ont créé et/ou utilisé les documents d'archives'.

Son objectif est donc avant tout de pouvoir traiter les agents (personnes, familles, institutions) qui produisent, utilisent ou conservent des archives.

EAC-CPF permet également de décrire les relations entre les différents agents. Il est ensuite possible de lier ces informations aux descriptions des documents d'archives.

Notons enfin que EAC-CPF permet de qualifier certaines informations en les liant à des vocabulaires contrôlés.

1.2.1. Structure d'un document EAC-CPF

Un document encodé en EAC-CPF contient deux éléments obligatoires : l'élément <control> et l'élément <cpfDescription>.

1.2.1.1. <control>

L'élément <control> rassemble un certain nombre de métadonnées sur l'instance encodée. Nous pourrons notamment y retrouver les balises suivantes :

- <recordID> : identifiant de la notice
- <maintenanceAgency> : informations sur l'institution responsable de la création et de la gestion de la notice
- <maintenanceStatus> : informations sur l'état et le statut de la notice
- <maintenanceHistory> : historique des révisions de la notice
- <publicationStatus> : informations sur le statut éditorial de la notice
- <languageDeclaration> : informations sur la langue utilisée (concerne les caractères codés, mais aussi la langue de rédaction de la notice)
- <sources>: une description des sources consultées pour la rédaction de la notice. Il est possible d'attribuer un @xml:id (identifiant unique) à chaque source décrite.
- <conventionDeclaration> : informations sur les normes et vocabulaires contrôlés utilisés pour la rédaction de la notice

Exemple d'un élément < control > complété :

```
xmlns:xlink="http://www.w3.org/1999/xlink">
<!-- recordID obligatoire -->
  <recordId>EXPERTS_VAVASSEUR-DESPERRIERS_FICTIF01/recordId>
<!-- maintenanceStatus obligatoire
        Valeurs possibles: new, revised, deleted, cancelled, deletedSplit, deletedReplaced -->
 <maintenanceStatus>new</maintenanceStatus>
<!-- publicationStatus optionnel
valeurs possibles : approved, inProcess -
<publicationStatus>inProcess</publicationStatus>
<!-- maintenanceAgency obligatoire</pre>
        nom de l'institution responsable de la création/maintenance de l'instance EAC -->
<maintenanceAgency>
  <agencyCode>FR-CNRS</agencyCode
<!-- agencyCode optionnel-->
<!-- agencyName obligatoire--
  <agencyName>Centre national de la recherche scientifique</agencyName>
 </maintenanceAgency>
<!-- languageDeclaration optionnel-->
<languageDeclaration>
<!-- languageCode forme [A-Z][a-z]{3} --> <!-- en caractère latin -->
 <script scriptCode="Latn"/>
 </languageDeclaration>
     conventionDeclaration optionnel et répétable-->
 <conventionDeclaration>
    - références aux normes utilisées pour former la notice -->
 <abbreviation>ISAAR(CPF)</abbreviation>
  <citation xlink:href="http://www.ica.org"</pre>
  xlink:type="simple">Norme internationale pour les notices d'autorité archivistiques (Collectivités, personnes, familles) du Conseil internat
 </conventionDeclaration>
 <conventionDeclaration
 <abbreviation>AFNOR NF Z44-060</abbreviation>
```

```
<citation>Norme NF Z 44-060, octobre 1983, Cataloque d'auteurs et d'anonymes : forme et structure des vedettes des collectivités auteur. Norme
<conventionDeclaration>
 <abbreviation>AFNOR_Z44-061</abbreviation>
 <citation>AFNOR NF Z44-061 Juin 1986 Documentation - Catalogage - Forme et structure des vedettes noms de personnes, des vedettes titres, des
 </conventionDeclaration>
<conventionDeclaration>
 <abbreviation>ISO8601</abbreviation>
 <citation>Norme ISO 8601:2004 Éléments de données et formats d'échange - Échange d'information - Représentation de la date et de l'heure.</ci
</conventionDeclaration>
<!-- localControl optionnel et répétable
correspond au niveau de détail de la notice --> <localControl localType="detailLevel">
<term>in progress</term>
</localControl>
<!-- maintenanceHistory obligatoire -->
<maintenanceHistory>
<!-- maintenanceEvent répétable -->
 <maintenanceEvent>
<!-- eventType
               valeurs possibles : - cancelled
                - created
                - deleted
                - derived
                - revised
                  updated--:
  <eventType>created</eventType>
<eventDateTime standardDateTime="2018-05-24"/>
    agentType valeurs possibles : human ou machine-->
  <agentType>human</agentType>
  <agent>Josselin Morvan</agent>
 </maintenanceEvent>
<!--sourc description des sources utilisées pour compléter cette notice -->
<sources>
<!-- possibilité de mettre les références bibliographiques sous l'espace de nom tei avec objectXMLWrap -->
  <sourceEntry>Lettres de provisions de Vavasseur Desperriers, Louis Athanase (1773-01-13)/sourceEntry>
 <source>
  <sourceEntry>MC/ET/XXVII/402, 5 avril 1779</sourceEntry>
  <descriptiveNote:
   VAVASSEUR-DESPERRIERS (Louis Athanase) architecte expert juré du roi quai des Ormes paroisse Saint-Paul § Contrat de mariage entre Louis
  </descriptiveNote>
 </source>
  <sourceEntry>MC/ET/XXVII/415, 4 janvier 1781/sourceEntry>
  <descriptiveNote:
    <P>VAVASSEUR-DESPERRIERS (Louis Athanase) architecte expert juré du roi rue et paroisse Saint-Paul § Inventaire après décès de Marie Angéli
   </descriptiveNote>
 </source>
 <source>
   <sourceEntry>MC/RE/XXVII/9, 3 thermidor an IX (22 juillet 1801)</sourceEntry>
  <descriptiveNote:
   Inventaire de Louis-Athanase Vavasseur Desperriers, rue Saint-Paul.
  </descriptiveNote>
 </source>
 <source>
  <sourceEntry>MC/RE/VI/18, 13 janvier 1786</sourceEntry>
   Notoriété portant renonciation par Louis Athanaze Vayasseur Desperriers, et par Claude Prosper Vayasseur Desperriers, à la succession de
 </source>
 <source>
  <sourceEntry>MC/RE/VI/18, 28 mars 1786</sourceEntry>
  <descriptiveNote>
    Inventaire après décès de Lucie Charlotte Félicité Meusnier, veuve de Claude Vavasseur Desperriers [mère de L Ath. VD]
  </descriptiveNote>
 </source:
</sources>
```

1.2.1.2. L'élément < cpfDescription>

L'élément <cpfDescription> est l'élément servant à la description de l'entité (personne, collectivité ou famille). Il peut contenir les sous-éléments suivants :

- <identity> : pour décrire l'identité de l'entité encodée
- <description> : contient des éléments de description formels conformes à la norme ISAAR (CPF)
- <relations>: description des relations existantes entre l'entité décrite et une ou plusieurs collectivités, personnes, familles (sous-élément <cpfRelation>), ou entre l'entité et une ou plusieurs fonctions (balise <functionRelation>)
- <alternativeSet> : contient des notices d'autorités alternatives

1.2.1.3. <identity>

L'élément <identity> est particulièrement complet et sert à la description de l'identité de l'entité encodée. Il peut contenir la (les) forme(s) autorisée(s) du nom (<nameEntry>), des variantes, mais également des formes parallèles, c'est-à-dire rédigées dans une langue ou une écriture différente (<nameEntryParallel>). Il doit contenir une balise permettant de préciser le type d'entité (personne, famille, collectivité), et il est également possible d'attribuer un identifiant. Chaque partie du nom pourra être encodée à l'aide d'une balise <part>. Chaque forme du nom peut être datée par l'emploi de l'élément <useDates>; attention, il ne s'agit pas d'indiquer ici les dates de naissance ou de mort! Enfin, il est possible de préciser si une forme du nom est une forme autorisée, alternative, ou préférée (<authorizedForm>, <alternativeForm>, preferredForm>).

1.2.1.4. <description>

L'élément <description> est utilisé pour décrire l'entité. Il reprend les catégories descriptives d'ISAAR (CPF), notamment :

- <existDates>: dates d'existence
- <place>: lieu
- <localControl> : description locale, permet d'enrichir la liste des éléments de description définis par ISAAR (CPF). On peut l'utiliser par exemple pour spécifier la nationalité de l'entité
- <function> : contient des informations sur les fonctions exercées par l'entité
- <occupation> : contient des informations sur les activités de l'entité
- <structureOrGenealogy> : contient des informations sur la généalogie d'une personne ou d'un famille
- <generalContext> : contient des informations sur le contexte général, social et culturel de l'entité
- <biogHist>: biographie / histoire

1.2.1.4.1. <existDates>

Cet élément permet d'indiquer les dates d'existence de l'entité décrite : naissance et mort pour une personne, ou dates de création et de disparition pour une collectivité. Les dates peuvent être renseignées de manière précise ou de manière approximative.

La tâche peut être plus complexe lorsque les sources se contredisent et que deux dates, ou plus, sont avancées.

```
<date localType="death"</pre>
   notAfter="1815-03-11"/>
 <date localType="death"
standardDate="1824-04-03"/>
</dateSet>
:!-- Il est possible d'ajouter des notes.
<descriptiveNote>
  L'état-civil reconstitué de Paris (en ligne), signale le décès d'un Jean Pierre Angot le 3 avril 1824. Toutefois, nous avons retrouvé au M
</existDates>
<!-- autre solution possible -->
<existDates>
 <dateSet>
 <dateRange
<!-- On indique dans un <dateRange/> les dates les plus probables -->
  <fromTo standardDate="1739-11-13"/>
   <toDate notAfter="1815-03-11"/>
<!-- on indique les autres dates possibles avec un élément <date/> et un @localType pour préciser l'information. -->
 <date localType="death</pre>
  standardDate="1824-04-03"/>
 </dateSet>
</existDates>
```

1.2.1.4.2. <place> et <places>

La description des lieux ne pose pas de problèmes particuliers. Il est possible de décrire un lieu (<place>) ou plusieurs lieux (<place>). La nature des lieux peut être précisée et rédigée selon le langage naturel ou selon un vocabulaire contrôlé. Une datation plus ou moins complexe peut être renseignée.

1.2.1.4.3. <function> et <functions>

À l'instar des lieux, il est possible de préciser une ou plusieurs fonctions rattachées à l'entité décrite. En cas de contradiction entre les sources, l'attribut @localType pourrait permettre que préciser les informations. Dans l'hypothèse où un des experts passerait d'un colonne à l'autre, cette solution devra être évaluée.

Le schéma EAC-CPF ne semble toutefois pas proposer de solutions satisfaisantes concernant les « événements » qui peuvent être rattachés à une fonction (lettres de provisions, acquisition/cession, etc.). Une solution serait de donner un @localType à un élément <date>, mais cela implique de définir une liste des types possibles ; qu'en sera-t-il si nous souhaitons faire apparaître d'autres informations ?

Une autre difficulté réside dans l'intitulé des fonctions : Angot est par exemple qualifié d''architecte entrepreneur de bâtiments' dans les sources, alors qu'il est dans la colonne des 'experts entrepreneurs'. Une solution serait de doubler l'élément <function> et de le typer. Nous pourrions utiliser ce même principe au niveau supérieur, pour l'élément <functions> (au pluriel), afin de distinguer des types de fonctions (expertise, formation, etc.).

Il est possible de lier une fonction à un vocabulaire contrôlé pour la préciser.

```
<functions localType="expert">
<function localType="verified">
<function localType="expert entrepreneur">
<function>
<function>
<function localType="sources">
<function localType="sources">
<function localType="sources">
<function localType="lettresProvisions"
<function>
<function>
<functions
<functions localType="other">
<functions localType="other">
<functions localType="other">
<function>
<function>
<function>
<function>
<function>
<function>
<function>
<function>
</function>
```

1.2.1.4.4. <biogHist>

Il s'agit d'une balise spécialisée dans la description des éléments biographiques. Elle ne semble pas soulever d'interrogation particulière. Elle peut prendre la forme d'une liste chronologique, d'une liste simple, ou de plusieurs

paragraphes. La liste chronologique semble la plus appropriée ; elle permet d'indiquer l'évênement, sa date ou sa durée, et éventuellement son lieu.

L'élément <chronList> peut être répété au sein de <biogHist> et précisé par un attribut @localType. Nous pourrions utiliser cette méthode pour, par exemple, rattacher des événements à une fonction (lettres de provisions, acquisition/cession, etc.).

Il n'est pas possible de lier un événement à un vocabulaire contrôlé.

```
<br/>diogHist>
:!-- éléments biographiques relevant de la vie privée -->
 <chronList localType="personnal"</pre>
  <chronTtem>
   <date standardDate="1646-08-16"/>
   <placeEntry>Paris, paroisse Saint-Nicolas-des-Champs</placeEntry>
<event>Baptême de Louis Convers</event>
  </chronItem>
  <chronItem>
   <date standardDate="1675-01-15"/>
  <event>Mariage entre Louis Convers et Jeanne Marguerite Nocret.
  </chronItem>
 </chronList>
    - éléments biographiques relevant de la fonction d'expert -->
 <chronList localType="expert">
  <chronItem>
   <event>Lettres de provision</event>
  </chronItem>
  <date standardDate="1718-07-12"/>
   <event>Vente</event>
 </chronItem>
 </chronList>
</biogHist>
```

1.2.1.5. <relations>

<relations> permet la description des relations existantes entre : l'entité décrite et une ou plusieurs
collectivités, personnes, familles (sous-élément <cpfRelation>); l'entité et une ou plusieurs fonctions
(<functionRelation>); l'entité et une ou plusieurs ressources (<ressourceRelation>).

La description des relations peut être assez fine puisqu'il est également possible d'indiquer des dates et des lieux, et il est possible de faire appel à un vocabulaire contrôlé pour préciser le type de relation. Enfin, des notes peuvent être ajoutées afin d'apporter des précisions si nécessaire.

```
<!-- Angot --><relations
xmlns:xlink="http://www.w3.org/1999/xlink">
 <cpfRelation xlink:arcrole="http://purl.org/vocab/relationship/spouseOf"</pre>
 xlink:type="simple">
 <relationEntry>Vallée, Marie Angélique</relationEntry>
 </cpfRelation>
 <cpfRelation xlink:arcrole="http://purl.org/vocab/relationship/siblingOf"
xlink:type="simple">
<!-- Beau-frère -
 <relationEntry>Vallée, Lucien Louis</relationEntry>
cpfRelation xlink:arcrole="http://purl.org/vocab/relationship/siblingOf"
xlink:type="simple">
<!-- Beau-frère -
  <relationEntry>Vallée, Nicolas Louis</relationEntry>
 </cpfRelation>
 <cpfRelation xlink:arcrole="http://purl.org/vocab/relationship/Relationship"</pre>
 xlink:type="simple"
<!-- ancien orateur de la loge de -->
 <relationEntry>Maître Puisieux</relationEntry>
 <descriptiveNote>
   Angot est ancien orateur de la loge du vénérable maître Puisieux.
<!--Le Bihan, Françs-maçons parisiens du Grand Orient de France (fin du XVIIIe siècle), Paris, Bibl. Nat., 1966.-->
 </descriptiveNote>
<cpfRelation xlink:arcrole="http://purl.org/vocab/relationship/Relationship/</pre>
```

1.2.2. Points à discuter

L'encodage des six premiers experts était intéressant car il a permis de tester le modèle en conditions réelles et de soulever des questions.

Concernant les sources consultées pour bâtir les notices, s'il est possible de les déclarer, il est en revanche assez difficile, voire impossible, de les relier aux différentes informations, ce qui pourrait être problématique lorsque les sources se contredisent. C'est un des principaux inconvénients du modèle EAC-CPF. Lorsque c'est possible, nous pourrions utiliser un élément <descriptiveNote>, toutefois il ne s'agirait que d'une description en texte libre et non d'un véritable lien crée pour indiquer une relation entre une information et sa source.

Le modèle de données EAC-CPF oblige à en quelque sorte à fragmenter les données. Il est toutefois étonnamment facile de répéter certaines informations. Il conviendra de mener une véritable réflexion sur ce point, afin de déterminer les méthodes d'encodage les plus appropriées pour ne pas répéter les informations à différents endroits, ce qui permettra également le respect des normes archivistiques.

Le caractère très contraignant du schéma EAC-CPF se fait sentir par moment, et l'ajout de certaines balises pourrait être appréciable afin de décrire plus finement nos données : par exemple, lorsque nous décrirons une fonction, nous pourrions avoir besoin d'indiquer des informations relatives aux lettres de provisions, à l'acquisition/cession de la charge, etc. En l'état actuel, nous sommes obligés de fragmenter ces informations entre la description de la fonction et les éléments biographiques, faute de balises véritablement spécifiques, et *quid* du lien entre ces informations.

Enfin, ce manque de liberté concernant l'encodage est un peu déroutant lorsque l'on est habitué à des formats autorisant un balisage plus libre, comme la TEI par exemple.