

EAD Revision: Digest of Comments

Introduction:

This document brings together all the comments following the call for comments from 30 September 2010 to 28 February 2011.

The comments were made available initially as raw data in a spreadsheet, the row number of which is in the reference for each comment found in italics at the end of the comment. This also includes the individual (name quoted where comments are not official comments from their institution) and institution who submitted the comment.

This document attempts to group comments that are similar in sections as set out below. In this way 127 rows in the spreadsheet have become 58 comments in this document. These are arranged within sections based on their nature and/or the area or elements of the standard to which they relate.

Section	Sub-section	Comments
Schema and Documentation		1-5
General Nature and Purpose		6-13
Relationships to Related Resources and Entities		7-18
EAD Header		19-31
Archival Description	Hierarchy elements	32-36
	Wrapper elements	37-39
	Controlled terms	40-42
	Digital Archival Objects	43-46
	Date encoding	47
	Languages and Scripts	48
	Other elements	49-58

:

Bill Stocking: 23/9/2011

Schema and Documentation

1: Base revision on model

Justification(s):

We have received the following comments criticisms about EAD incompatibility with more generic metadata models:

- “EAD is a model enabling us to create a type of structured documents while EAC-CPF schema has been designed as a structured set of metadata”;
- “many things are implicit in EAD such as non-redundancy of information from high levels to the lowest levels in multi-level description”.

We take issue with these comments for the following reasons:

- EAD has been designed as a hierarchical format while EAC-CPF authority records are non-hierarchical, and it is also possible to extract non-hierarchical metadata from EAD structured documents.
- Therefore, the number of Dublin Core that can be created depends on the <c> components contained in an EAD document. Through the @encodinganalog attribute, EAD elements can be matched to Dublin Core elements, and Dublin Core is a model that can be expressed in RDF.
- Perhaps the possibility of extraction of Dublin Core metadata is not explained enough in EAD Tag Library.

[13. French EAD and EAC-CPF Working Groups]

[Have a] Solid data model as a starting point instead of a pretty formatted printout: express your ideas with UML. Take a holistic view and cover both ISAAR and ISAD [as] these two are inseparable. *[49. Dogen Zo]*

In order for EAD to be a real exchange format it should implement a reference model (UML) based on principles outlined in ISAD(G). *[118. Jakob Saternus, The City Archives of Gothenburg]*

2: Have a schema rather than a DTD

Justification(s):

We suggest changing the EAD into an XML schema, which is the reference version, but the starting point of this work should be the DTD EAD 2002, which is the conceptual model of the standard. The XML schema developed in 2007 is only a specific formalization of EAD.

We suggest that the elements should be made more explicit and consistent with the current rules for writing XML schemas (for example, we could write <component> instead of <c>, <descriptionOfSubordinateComponents> instead of <dsc>, etc.).

We recommend using XML and Xlink libraries (see <http://www.w3.org/XML/1998/namespace> and <http://www.w3.org/TR/xlink/>). For example, the attributes @xml:id, @xml:lang and @xml:base should be included in EAD (these attributes are used in TEI), and the attribute @xml:id should be defined as an EAD attribute of type "id". The use of generic standardized elements and attributes provides interoperability.

In general, all the linking elements and attributes available in the EAD (<ref>, <archref>, <bibref>, <extref>, <ptr>, <extptr>, <dao>, <daogrp>, <arc>, etc..) do not work in the same way, and it makes them difficult to use. EAD should therefore move on to Xlink (use attributes @ xlink: href and @ xlink: type), but it is important to think in terms of type of object rather than in terms of links (the type of links is taken into account by Xlink).
[4, 6, 7, 9. French EAD and EAC-CPF Working Groups]

Please update the codes for scripts to include the code jpan for Japanese, and any others that may have been added to ISO 15924. *[84. Lisa Conathan, Beinecke Rare Book and Manuscript Library, Yale University]*

This change is related to EAD validation and documentation. I would suggest one of the following two options:

- If the EAD community continues to support both DTD and schema validation, all official documentation should highlight the differences between these two options, including the EAD tag library. This should include specific examples whenever there are differences between DTD compliant and schema compliant markup.
- Alternatively, the EAD community could commit to one form of validation or the other (preferably schema) for the sake of consistency.

At present, the EAD tag library (and other EAD documentation) consistently references only DTD-compliant markup and examples. Those who choose to create schema compliant instances must look for separate documentation whenever there are differences between DTD and schema requirements. This is particularly difficult for institutions that are in the early stages of EAD markup and/or those which have limited technical support.
[86. Tammy Gobert, Archives & Special Collections, Rensselaer Polytechnic Institute]

The EAD XML Schema is not properly annotated. *[120. Jakob Saternus, The City Archives of Gothenburg]*

Ideally EAD should be well suited as one of the allowed metadata standards in repositories exposed by means of OAI-PMH. e.g.
<http://repo/oai2?verb=GetRecord&identifier=SE/ZZZ/A/1&metadataPrefix=ead>

should return metadata relevant for SE/ZZZ/A/1 [124. *Jakob Saternus, The City Archives of Gothenburg*]

3: Produce migration script(s)

Justification(s):

Finally, it is essential that tools for converting DTD valid instances into schema valid instances be provided. [14. *French EAD and EAC-CPF Working Groups*]

The group felt the revision process should provide the following: [...] -Script(s) to transform EAD 2002 instances to the new schema with full documentation -Website with access to the schema(s) and documentation, updated to take account of the changes and the broader environment in which EAD is used. [62. *UK Archives and Records Association, Data Standards Group*]

4: Produce tag library and website

Justification(s):

The group felt the revision process should provide the following: [...] Tag library in both printed and online format. This should include a summary of changes from EAD 2002 and a crosswalk to MODS and DC as well as MARC21 and ISAD(G); Website with access to the schema(s) and documentation, updated to take account of the changes and the broader environment in which EAD is used. [62. *UK Archives and Records Association, Data Standards Group*]

5: Don't produce technical application guidelines

Justification(s):

The group felt there was no need for formal Technical Application Guidelines as provided with EAD v1.0. [62. *UK Archives and Records Association, Data Standards Group*]

General Nature and Purpose

6: Requests for additions/corrections should be cautiously evaluated.

Justification(s):

Changes will not be required simply because of the difficulties of some users in EAD implementation: practices or tools maybe inappropriate. [5. *French EAD and EAC-CPF Working Groups*]

7: Reduce the number of elements / attributes

Justification(s):

These figures might be slightly off, but I believe that the first version of EAD had 145 elements. The second version, EAD 2002, increased by one, to 146 elements (with 8 elements discontinued, and 9 new elements added). A worthy goal, then, in my opinion, would be to decrease that number for the third revision.

With the addition of EAC, it'll behove EAD to trim down some. Additionally, it sounds like most of the formatting-type elements might be handled different in the revision, so it might be a pretty simple task to reduce the number of EAD-specific elements.

The more difficult task, I imagine, would be in identifying proper migration paths for any deprecated elements and attributes. [1. *Mark Custer, East Carolina University*]

Not sure how best to address the problem described in the Justification, but here are some ideas: Within the tag library, provide an alternate tag or designate "catch all" tags to accommodate data in tags that are used less frequently. Alternatively, issue a stylesheet that uses the full array of EAD tags, or create a companion best practices guide that prescribes standard defaults for little-used tags. [50. *Lisa Miller, Hoover Institution Archives, Stanford University*]

It was felt that EAD should be generally less complex and many of the proposals below have this in mind. [65. *UK Archives and Records Association, Data Standards Group*]

8: EAD should be made more 'database friendly'.

Justification(s):

Again, it's not about the final result in the form of booklet (who cares about printouts?). It's about the complex structures which cannot be captured by a monolithic XML document. [49. *Dogen Zo*]

Experience in the UK has shown that exporting and importing EAD instances out of and into databases is fraught with difficulty and the group, therefore, asks that TS-EAD make the next version of EAD more functional in this respect. In particular, formatting tags should be lost or if this is not possible greatly reduced and made more generic, perhaps along the lines of the model of the <descriptiveNote> tag in EAC-CPF. Thought here should be given to allowing inclusion of HTML tagged data. There should also be fewer opportunities for mixed-content generally. It was felt that EAD should be generally less complex and many of the proposals below have this in mind. [65. UK Archives and Records Association, Data Standards Group]

The underlying model is not at all database friendly. [121. Jakob Saternus, The City Archives of Gothenburg]

The schema is not friendly to programming language binding either (e.g. JAXB) mainly because of frequent use of mixed content. [122. Jakob Saternus, The City Archives of Gothenburg]

Unnecessary short-cuts (mixed content again) should be avoided and discouraged (e.g. <unittitle><date>1956</date></unittile>). [125. Jakob Saternus, The City Archives of Gothenburg]

9: Review use of ‘group’ elements (<eadgrp> etc)

Justification(s):

The DTD EAD Group and its elements <eadgrp>, Archival Description Group <archdescgrp> and Description of Subordinate Components Group <dscgrp> should be deprecated. These elements are not used. EAD implementers do not aggregate different EAD documents. EAD linking elements such as <extref>, <archref>, etc., enable users to establish logical links between different electronic documents (see for instance the “IREL” application of the National Archives of Overseas <http://www.archivesnationales.culture.gouv.fr/caom/fr/index_irel.html> , or the guide of records on slavery and slave trade <<http://www.archives-esclavage.culture.gouv.fr/>>). [15. French EAD and EAC-CPF Working Groups]

<eadgrp>: The tags relating to EAD Group should be abolished. Rather than grouping separate EAD files in this way, they should be related to each other as related resources [...]. [75. UK Archives and Records Association, Data Standards Group]

Keep eadgrp. But it doesn't necessarily have to be an element in the schema. It can be a different schema using the ead-schema and thus creating an ead-grp: In Sweden it can occur that the same institution is exchanging information about more than one archive at the same time. Then it's easier to have all the archival descriptions in one XML-file. [108. Karin Bredenberg, *The Swedish National Archives*]

10: Review use of formatting elements

Justification(s):

The utility of a purely description-based <archdesc> merits further investigation. Should we maintain the essential text formatting elements such as <p>, <emph>, or linking elements within <archdesc>? Sometimes, a conceptual approach may be too restrictive[!] [8. *French EAD and EAC-CPF Working Groups*]

We compared practices in the national archives and territorial archives as well in libraries. We propose to depreciate other elements, that are not used in France: table elements <table> <tgroup>, <thead>, <tbody>, <colspec>, <row>; the <head> element is depreciated in the future information system of the National Archives. Territorial archives and the National Library use it because of limitations of EAD and tools. For instance, they use it because of the lack of attributes specifying the language and the script code of the information included in an element). Before depreciating <head>, we should therefore consider the reasons for its use. [16. *French EAD and EAC-CPF Working Groups*]

<blockquote> element: the attribute @type should be available within the <blockquote> element; it would enable us to specify the nature of a citation; it would be helpful to make the <blockquote> element more semantic and to make it available in the <unittitle> element (for instance, the National Library of France uses incipit within <blockquote> as access points). [21. *French EAD and EAC-CPF Working Groups*]

<head>: The group felt that TS-EAD should consider the purpose of this element with a view to deprecation or abolition. [78. *UK Archives and Records Association, Data Standards Group*]

<table>: Table and related elements should be abolished. [80. *UK Archives and Records Association, Data Standards Group*]

Simplify the content model of the list elements. Deprecate the elements <chronlist> and <list type="deflist">, and replace these by <table>.

The <list> element will contain the type attribute with the values simple, marked and ordered.

The use of attributes should be validated by the Schema, so <list type="marked" numeration="arabic"> should be invalid.

The mark attribute should get a set of closed values, e.g. bullet, box, dash, hyphen, or some of the HTML entity names and/or numbers or some of the Unicode character codes.

[109. Nationaal Archief, Netherlands]

The value of the attribute cols of the element <table> should be validated by the Schema. The value could be 1 to 256 inclusive. Better validation of the EAD instances. *[110. Nationaal Archief, Netherlands]*

Nodes that do not have semantics and are usually used for display purposes, such as <emph>, <head>, <table>, should be eliminated. Stylesheets can do this job. *[114. APEnet project (Archives Portal Europe)]*

EAD as it is today has a primary focus on pretty print-outs and web presentation (rendering hints, frontmatter, etc). The format is rather monolithic in the sense that the whole archival description is encoded in one xml file. *[119. APEnet project (Archives Portal Europe)]*

11: Extend the availability of the TYPE attribute to all EAD elements.

Justification(s):

It often becomes necessary to "qualify" the use of a particular EAD element. The "ID" attribute can be used to qualify data, but it constrains the type of data that can be entered in that attribute. Currently, only a handful of EAD elements allow data qualifying by use of the "type" attribute, but there have been many times when I was forced to use some other attribute (altrender, ID) when TYPE would have been preferable. [60. Mark Carlson, University of Washington]

12: Consider how user generated content can be incorporated and acknowledged

[82. UK Archives and Records Association, Data Standards Group]

13: Enable definition of local values in attribute lists

Justification(s):

When there are attribute with value lists add a value 'other' and an attribute 'other'.... so it's possible to define a local list: can be valuable when you want to validate local values. [111. *Karin Bredenberg, The Swedish National Archives*]

Relationships to Related Resources and Entities

14: Allow logical links to related resources and entities

Justification(s):

It is possible to establish different types of links between EAD documents and other digital resources:

- link between two EAD documents;
- link between one EAD document and one (or many) EAC-CPF documents;
- link between one EAD document and archival digital objects (JPEG image files, MP 3 audio files, MPEG video or audio files);
- link between one EAD document and other resources providing additional information or illustrating the finding aid (URLs, XML, PDF, JPEG, MP 3, MPEG files, other formats).

The linkage mechanisms will depend on the relationships users want to establish. For instance, if you consider linkage mechanisms between EAD et EAC-CPF, links from an EAD document are made through the @authfilenumber attribute of the <persname>, <corpname> and <famname> elements (the value of this attribute is only the identifier of the EAC-CPF document), with an attribute @role which enables users to type relationships and which is a CDATA attribute. It means that the value of this attribute may contain any set of character ("creator", "author", "collector", "editor", "holder", etc).

Links from an EAC-CPF instance are established through the attribute @xlink:href of the element <resourceRelation>. However, the @resourceRelationType which enables users to type relationships can only be qualified by three values (creatorOf, subjectOf and other). It would be preferable to open the list of values to be given to the attribute @resourceRelationType, and to make the use of @role (EAD) and of @resourceRelationType (EAC-CPF) more consistent.

Finally, it is possible to link an EAD document to an EAC-CPF document through the identifier of the authority record. However, a permanent link pointing to the authority record would be far more efficient (see general comment about Xlink).

[43-46. French EAD and EAC-CPF Working Groups]

Link the <bioghist> element to an existing eac-cpf or ISAAR(CPF) record, by adding, for example, the origin of the information in this field to a record in an authority file (as it can be done, for example, with <persname> or <corpname>).

As in the future EAC-CPF may be largely used and deployed, it could be interesting to have the possibility of linking the content of the <bioghist> element to the authority record that was used as the source/origin for these data.

The <bioghist> element could have two attributes to indicate the origin of the data and the ID of the authority record used as source.

[57. Ricardo Eito-Brun, Universidad Carlos III de Madrid]

When creating a new EAD file for a component, e.g. a file, series or item, the standard should give the possibility of linking the new record to the record of the components at the upper level, e.g. fond, series.

Today the description starts at the upper-level item, and by means of the components you can go down into the archival units' hierarchy to incorporate/embed descriptions of lower level items.

In some implementations, it is better to keep separate EAD files for each component (fond, series, file, items...) and then link them. According to the logic of the <c> components, the links are kept from the item at the upper level, to the items at the lower levels. It would be interesting to have the capability of linking descriptions in both directions, indicating the context of each component/description regarding their upper- and lower-level components.

[58. Ricardo Eito-Brun, Universidad Carlos III de Madrid]

EAD should, following the model, given in Encoded Archival Context – Corporate Bodies, Persons and Families (EAC-CPF), have dedicated areas for relating an EAD instance to other resources (including but not limited to archives), entities that created (or have any other relationship) to the material described, the functions and activities of which the material described are evidence and possibly the places to which the material relates. [63. UK Archives and Records Association, Data Standards Group]

The purpose of this revision is to provide a place in EAD to identify/link to related metadata. Such information would include but not necessarily be limited to information about the described collection, about any described component of the collection, or about concepts, names or vocabulary terms expressed anywhere in the EAD. Examples might include:

- .. a collection-level MARC record that describes the same materials described in the EAD document
- .. an analytic description in an image catalog for a photograph that is also described as a component of a collection in EAD
- .. an EAC record for a name

Uses include interoperative indexing and display for end-users, automated updates to EAD text of controlled vocabulary terms, and linkages between EAD and the related metadata records.

This should be valid in <ead>, <c> and various elements that can hold controlled terms: <name>, <corpname>, <famname>, <persname>, <geogname>, <subject>, <formgenre>

It should be possible to record the type of metadata and the type of location identifier supplied, controlled or uncontrolled for both the type and the location. In EAD, this could be accomplished using a single set of attributes applicable at many points in the schema, or a mix of mechanisms if that would better fit the existing EAD element structure.

For example, PREMIS and METS both use a mechanism with paired controlled attributes and uncontrolled attributes, i.e.

-- LOCTYPE, with controlled values such as ARK, URL, URN and OTHER; and OTHERLOCTYPE to specify another kind of location identifier, such as a specific local system.

-- MDTYPE, with values such as MARC, DDI, FGDC and OTHER; and OTHERMDTYPE to specify metadata not from the enumerated list.

MODS permits linking to related authority metadata by using @authorityURI and @valueURI attributes to identify the source of the vocabulary and the applicable record within the source.

Possible examples:

```
.. <mdref mdtype="marc21" loctype="URN">  
http://hollis.harvard.edu/?itemid=|library/m/aleph|009504960</mdref>  
.. <mdref mdtype="other" othermdtype="OLIVIA" loctype="otherloctype"  
otherloctype="OLIVIAwork">416461</mdref>  
.. <mdref mdtype="marc21" loctype="other" otherloctype="viaf">  
http://errol.oclc.org/laf/n_86140996.html</mdref>  
.. <mdref mdtype="marc21" loctype="other" otherloctype="lcnaf">  
no2005088392</mdref>  
.. <mdref mdtype="eac" loctype="url">  
https://wiki.nla.gov.au/download/attachments/32189/oliphant-  
BrSp.xml?version=1&modificationDate=1201813806000</mdref>  
[88. Harvard University]
```

Possibility to link back to original presentations on lower levels. Within the APEnet project to build an Archives Portal Europe the question arose, which EAD element could be used to link back from a central (portal) presentation to original descriptions of subcomponents on content provider's websites? Of course a prerequisite for this on the content providers' side would be persistent identifiers / persistent URLs, but from an EAD point of view, the question would be, where to put this information.

In APEnet EAD, a subset of the EAD 2002 as defined for the APEnet project (http://www.archivesportaleurope.eu/profiles/APEnet_EAD.xsd) we are currently offering <extptr> as subelement for <unitid> on the lower <c> levels for this purpose. *[117. APEnet project (Archives Portal Europe)]*

Linking multi-level descriptions is not supported e.g. <c href="http://repo/(...)" />. *[123. Jakob Saternus, The City Archives of Gothenburg]*

15: Allow the insertion of valid data from other XML namespaces

Justification(s):

Perhaps permit MARCXML and/or MODS to be included in EAD, specifically within any subject elements.

Example:

```
<mods:subject authority="lcs" >  
  <mods:name type="personal" >  
    <mods:namePart>Smith, Homer</mods:namePart>  
  </mods:name >  
  <mods:titleInfo type="uniform" >  
    <mods:title>Black man in Red Russia</mods:title >  
  </mods:titleInfo >  
</mods:subject >
```

Not:

```
<ead:subject>Smith, Homer. Black man in Red Russia</ead:subject >
```

Within EAD 2002, there's no good way to identify specific MARC subject subfields.

Rather than just include this option as something like an objectXMLwrap element, as in EAC, it would be nice to be able to validate this portion of the document within the EAD file.

[2. Mark Custer, East Carolina University]

Concerning textual elements and the logical structure of these elements, it would be preferable to have the possibility to use TEI elements or elements of another related schema designed for this purpose. In general, it would be desirable to make EAD interoperable with EAC-CPF and other related metadata schemas. *[12. French EAD and EAC-CPF Working Groups]*

The inclusion of other XML instances should be made possible in EAD documents through <objectXMLWrap> and <alternativeSet> elements. *[42. French EAD and EAC-CPF Working Groups]*

In the <bibref> element, include the possibility of using descriptive metadata from other schemas, e.g. MARCXML, MODS.

This change would provide additional capabilities for managing and processing the bibliographic data related to the archival description. In addition, it would open the possibility of gathering and referring to bibliographic descriptions exposed from other systems and repositories in a

dynamic way, without having to re-code the bibliographic item descriptive data.

[56. Ricardo Eito-Brun, Universidad Carlos III de Madrid]

EAD should (again following the model in EAC-CPF) allow the insertion of valid data from other XML namespaces. Examples suggested were EAC-CPF, XHTML, KML, MARC21, MODS, and TEI (10: cataloguing). [64. UK Archives and Records Association, Data Standards Group]

Need for importing namespaces. For example, inside the <origination> and its subelements (<persname>, <corpname> etc) users might need to import nodes from the EAC-CPF schema in order to provide a more detailed description for the archives' creators. This will also solve the problem that some EAD users already have of providing birth-death and other kinds of dates related to the creator.

Another example as derived from the APENet project would be to import nodes from EAG (though no official schema yet) in <repository> to provide more detailed information on the archival institutions (f.i. opening hours). A third example, is the need to import schemas for the description of archival digital objects (currently expressed through the <dao> element used within the APENet project for this purpose) such as METS, or schemas for providing bibliographic references (i.e. inside the <bibref> element).

[113. APENet project (Archives Portal Europe)]

16: Align EAD elements with EAC-CPF elements

Justification(s):

If two elements are almost the same (for instance <dateRange>, <dateSet> and <date> in EAC-CPF, and <unitdate> and <date> in EAD), but the EAC element structure is better, then we should use the same name in EAD. EAD elements about control description should also be changed, a <control> element should be provided just as a <control> element has been provided in EAC-CPF. In general, the ways for managing EAD documents need to be improved, especially in the case of retrospective conversion projects.

The element <language> of the EAD element <eadheader> [<profiledesc><language>] should be made consistent with the element <language> of the EAC-CPF [control] element <languageDeclaration>.

Other necessary alignments: EAD <legalstatus> and EAC-CPF <legalStatus>, EAD <geogname> and EAC-CPF <place>, EAD attribute @source and EAC-CPF attribute @vocabularySource, EAD <emph> and EAC-CPF @style.

The <list> element was maintained in EAC-CPF.

The EAC-CPF attribute @translitteration is an addition in comparison with EAD.

The EAD attribute @level aims to qualify the levels of description, whereas the EAC-CPF element <level> is a formatting element.
[11, 36, 37, 38, 40, 41]. *French EAD and EAC-CPF Working Groups*]

17: Review role and model of <bioghist> element

The <bioghist> element does not belong semantically to EAD which focuses on the description of archival resources. It does however belong to EAC-CPF which provides a formal method for recording the description of records creators. Is the <bioghist> element useful when it is used as a substitute for the EAC-CPF format? However, <bioghist> may sometimes explain changes in the structure of an organization. Moreover, <bioghist> may highlight biographical or historical information in order to clarify the structure and the composition of a fonds, while a separate EAC-CPF authority record is more neutral and generic. We suggest that <bioghist> should not be removed from EAD, but the potential uses of this element should be explained and clarified. Here are some examples of use of the <bioghist> element in order to give some complementary contextual information within a EAD document:

<bioghist>

<p>Extrait de la circulaire préfectorale n° 13501/AC2 du 29 août 1941 adressée aux maires et administrateurs de commune mixte :

<blockquote>

<p>[...] Les opérations de recensement seront ouvertes le 1er septembre et closes le 10 septembre 1941. Vous aurez à prendre d'urgence un arrêté astreignant tous les juifs de votre circonscription à se rendre à la mairie qui leur délivrera les imprimés réglementaires de déclaration à remplir par les intéressés [...]</p>

</blockquote>

</p>

</bioghist>

<c level="otherlevel" otherlevel="sous-dossier">

<did>

<unitid type="cote d'archives">FR ANOM 93/1H73/2</unitid>

<unittitle>Pesticides</unittitle>

<unitdate>1917/1918</unitdate>

</did>

<bioghist>

<p>Les pesticides sont utilisés dans l'agriculture pour protéger les plantes contre les maladies parasitaires. Anticryptogamiques ou fongicides, en usage au XIXème siècle et au début du XXème siècle, ils étaient des produits relativement simples et d'obtention aisée : cuivre sous forme de sels tels que sulfate, acétate, soufre, chaux, etc. On les employait par pulvérisations de bouillies ou poudrages sur les organes aériens des végétaux.

Pour encourager la production agricole, l'administration était chargée du ravitaillement de la colonie, de l'achat des produits à l'étranger jusqu'à la vente aux agriculteurs.</p>

</bioghist>

```
<scopecontent>
<p>Soufre et sulfate de cuivre, approvisionnement, enquête sur les besoins,
répartition : circulaire, correspondance, déclaration, état récapitulatif des
demandes, rapports.</p>
</scopecontent>
</c>
```

[10]. *French EAD and EAC-CPF Working Groups*

The <chronlist> element is maintained in the EAC-CPF element <biogHist>, as well as <event> and <chronItem>; although these elements will not be used in the information system of the National Archives, it seems desirable to keep them in EAD for providing consistency with EAC-CPF; generally speaking, the elements embedded within the EAD element <bioghist> (<address>, <bioghist>, <blockquote>, <chronlist>, <dao>, <daogrp>, <head>, <list>, <note>, <p>, <table>) should be reviewed in comparison with the EAC-CPF element <biogHist>. [34. *French EAD and EAC-CPF Working Groups*]

<bioghist>: This should remain to hold contextual information about creators to enable users to use EAD alone rather than as part of a system with links to descriptions of creating entities held in EAC-CPF (see 3 above). The group thought, however, that TS-EAD should give some consideration to where contextual information about creation and use of the particular records being described (rather than more general information found in a biographical sketch or administrative history) belongs. Should there be another specific tag for this information or does <bioghist> serve both purposes? [73. *UK Archives and Records Association, Data Standards Group*]

18: It should remain possible to use EAD alone

Justification(s):

[It] is essential that users be allowed to use only EAD, available tools do not necessarily implement both EAD and EAC-CPF. [47. *French EAD and EAC-CPF Working Groups*]

It was felt very strongly that it is necessary to continue to allow EAD to be used alone rather than within a more dynamic environment. [66. *UK Archives and Records Association, Data Standards Group*]

EAD Header

19: Change the model for <eadheader> in line with the Control area in EAC-CPF

Justification(s):

EAD elements about control description should also be changed, a <control> element should be provided just as a <control> element has been provided in EAC-CPF. In general, the ways for managing EAD documents need to be improved, especially in the case of retrospective conversion projects. [11. *French EAD and EAC-CPF Working Groups*]

The group felt that the model here should be changed in favour of that of the <control> area found in EAC-CPF which would bring it more into line with the ICA's description standards. [76. *UK Archives and Records Association, Data Standards Group*]

20: Eliminate the <eadheader> (except for the <eadid>) and <frontmatter>

Justification(s):

All of that information is repeated elsewhere in the EAD record. Also, it's time to recognize that these records are not based on paper finding aids that are published and put on a shelf somewhere. [54. Jordan Patty, George Mason University]

21: <eadid> should not only be mandatory but have mandatory content

Justification(s):

The <eadid> element is mandatory but it may have no content and no attribute whereas the @identifier attribute is not mandatory. This is not logical. [17. *French EAD and EAC-CPF Working Groups*]

22: Add 2 optional attributes to the EAD root element: FILE and TITLE.

Justification(s):

When using XQuery to retrieve all and any XML elements from a file in an XML database, it is easier and more convenient when the actual file name is known in the root element <ead>, so it is more efficient to link to the physical file name (attribute FILE), and also group elements according to this FILE. The value of FILE can simultaneously be retrieved together with other elements using a FLWR expression in XQuery.

Although the title of an EAD file is encoded in <titleproper>, it is more efficient when a shorter (sometimes the titles can be very long!) title is encoded as an

attribute TITLE in <ead>. Moreover, the same reasons as listed for attribute FILE for <ead> apply. A file name and the title of an EAD finding aid are 2 essential nuggets of information to group and provide context for all and any retrieved XML elements. It helps a lot when this information is (optionally) provided from the very beginning of an EAD file. This makes it easier to reconstruct hit lists in XML retrieval systems applied to EAD finding aids. [53. *Junte Zhang, University of Amsterdam*]

23: It should be possible to associate several authors of a finding aid in <titlestmt>

Justification(s):

<titlestmt> element (in <filedesc> element): it would be helpful to include several <author> elements within <titlestmt> element. Currently, we have to record the names of all the authors of a finding aid in a single <author> element, or we have to use a <titlepage> element (which may contain several <author> elements) in order to isolate each author name. It may become problematic when we want to extract Dublin Core metadata or to build a basic bibliographic record for the finding aid. [18. *French EAD and EAC-CPF Working Groups*]

24: It should be possible to note the date on which a finding aid has been compiled

Justification(s):

There is no specific element for the date on which the finding aid has been compiled: it would be useful in retrospective conversion projects if we could mention the first date of publication of a paper-based finding aid which is retroconverted in its original form (in <editionstmt> and <publicationstmt> elements; in <creation> element, etc.). [19. *French EAD and EAC-CPF Working Groups*]

25: Clarify the role of <descrules> and of <processinfo>

Justification(s):

The <descrules> element is a sub-element of <profiledesc>. However, <descrules> is only about the rules, standards, conventions, and protocols used in preparing the description, it does not bundle information about the creation of the encoded version of the finding aid. Is <descrules> to be included within <profiledesc>? [28. *French EAD and EAC-CPF Working Groups*]

<descrules> Descriptive Rules: The "<descrules> - Element" should be moved from "<profiledesc> Profile Description" to "<archdesc> Archival Description" and to "<c>/c0x> Component " and make it repeatable on every level of description. <descrules> occurs only within <profiledesc> and it is comparable to ISAD(G) data element "3.7.2 Rules or Conventions". ISAD(G) allows using its descriptive elements on every level of description according to the concept of "multilevel description". This is not possible with EAD 2002. ("If the fonds as a whole is being described, it should be represented in one description, using the elements of description as outlined below in section 3 of this document. If description of the parts is required, they may be described separately also using the appropriate elements from section 3." in: ISAD(G): General International Standard Archival Description. Second Edition. Adopted by the Committee on Descriptive Standards, Stockholm, Sweden, 19-22 September 1999, Ottawa 2000, p. 12.). In most cases it may be sufficient to have the <descrules>-element on the "fonds/collection"-level but there may be a "series" or "subseries" within a "fonds" which requires special rules, for instance a series of photos or other different material. This special descriptive rule could be placed in the appropriate level. Furthermore this change would fully synchronize EAD with ISAD(G). [85. Dr. Edgar Kutzner, Archives of the Diocese of Fulda, Germany]

26: Allow the recording of the file size of EAD instance

Justification(s): The purpose of this information is to improve the researcher experience when interacting with large finding aids.

Researches can experience slow load times when they attempt to view large finding aids. To allay researcher frustration, it can be useful to provide a warning about potentially slow downloads or to indicate file size information before researcher elects to view a finding aid (i.e. in search results). Calculating files sizes in real time for all finding aids in a large result set can slow display of search results beyond reader tolerance. Encoding file size information within EAD allows for its display and thus it can alleviate user frustration.

If other institutions are using a mechanism to provide this information to researchers, it makes sense to encode the data in a standardized way across the archival community.

Harvard University has implemented this through a local change to the EAD schema in order to store the file size of EAD instances for display in search results. We have created a local version of the EAD schema which has an additional attribute, fileSize, which is valid for the ead element. For example, <ead fileSize="9910">. This attribute and its value are machine-generated when an EAD file is ingested by OASIS, our EAD catalog. Our research revealed that the styled html we derive from our EAD is roughly the same size.

We suggest that this should be an optional, non-repeatable piece of information.

We recognize that if the subcommittee is willing to include this information in EAD, an attribute on the EAD element is not the only possible location, a new element or attribute may be better suited to this purpose. One option may be to allow extent within filedesc, and include an attribute to indicate the unit of measurement. For example, "<filedesc><extent unit='megabytes'>9.6</extent> ... </filedesc>". Indeed, since the EAD filedesc is patterned after the TEI fileDesc element, which includes an <extent> sub-element, the subcommittee may want to revisit this part of the TEI standard: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-extent.html>. [92. Harvard University]

27: Add a flag for the presence of <dao>s in the EAD instance

Justification(s): The purpose of this flag is to support functionality in discovery systems for finding aids that link to archival digital content.

Researches often want to find only those EAD finding aids that contain links to online content. Calculating whether or not such content is present based on polling EAD instances for the presence of a <dao> in real time for all finding aids in a large result set can slow display of search results beyond reader tolerance. Flagging for the presence or absence of digital archival objects allows for faster retrieval of the desired result set.

This flag can also be used to "brighten" an interface option (a button or tab on which a user clicks) for an alternative view of the finding aid in which the digital objects are prominently featured.

Other repositories and implementations may have uses for such a flag. If other institutions are using a similar mechanism, it makes sense to encode the data in a standardized way across the archival community.

Harvard University has implemented this through a local change to the EAD schema by creating an attribute, digitalLinks, on the <ead> element. This attribute and its value are machine-generated when an EAD file is ingested by OASIS, our EAD catalog. We suggest that this should be an optional, non-repeatable piece of information. [93. Harvard University]

28: Add the elements <num> and <name> to the element <change> within <revisiondesc>; consequently replace <list> by <table>

Justification(s): <num> may contain an ID from a system that generate the change and <name> is for the name of the person or system as "agent". Than there are four elements available. If one prefers a list, a <table> is necessary. [98. Nationaal Archief, Netherlands]

29: Remove <p> from <editionstmt>, <publicationstmt> and <seriesstmt>

Justification(s): At this level there is (or, at least, should be) no need for an extended text with paragraphs. [102. *Nationaal Archief, Netherlands*]

30: Remove the element <frontmatter>

Justification(s): The <frontmatter> element includes pieces of information that can be also included in the <filedesc> element and encoded separately in its subelements. Within the APEnet project <frontmatter> therefore has not been implement anymore. So, it is probably one of the elements that it might not be included in the newer EAD version. [APEnet project (*Archives Portal Europe*)]

31: Allow the recording of record identifiers applied to the EAD record in various systems

Justification(s): The purpose of this revision is to

This is roughly analogous to the MARC 035 field: control number of a system other than the one whose control number is considered primary is a given context. In EAD, the EADID is the primary identifier, but as the finding aid moves among systems over time, it may acquire additional identifiers that are useful to track. Uses include record migrations or aggregator databases.

Attributes allow for either use of a resolvable identifier or any other type, such as an explicit expression of a system name.

Examples:

<sysid type="other" othertype="MH-ATprod"> 12345</sysid>
for an Archivists' Toolkit resource record number from which the EAD was derived

<sysid type="urn">http://nrs.harvard.edu/urn-4:12345</sysid>
for the system number of the EAD in an EAD aggregating database
[87. *Harvard University*]

Archival Description

Hierarchy Elements

32: <archdesc> should be a wrapper tag for components alone

Justification(s):

The <archdesc> tag should simply be a wrapper tag containing the hierarchical components of the description, that is it should not directly hold descriptive elements itself and the highest level of description should be held in a component tag (<c>). [67. *UK Archives and Records Association, Data Standards Group*]

33: Review purpose of <dsc>

Justification(s):

The <dsc> tag should be abolished: The current division between the upper level description and the remaining components is felt to represent an out of date paper model for archival data that seems to privilege description at the uppermost level. Such a split is not reflected in ISAD(G) and is, therefore, difficult to teach to UK archivists and also complicates data exchange and transformation. [68. *UK Archives and Records Association, Data Standards Group*]

Remove <dsc> from <c>, <c#> and <dsc>. The use of <dsc> within <c>, <c#> and <dsc> causes very unstable EAD instances (e.g. a new container list within a single element). [99. *Nationaal Archief, Netherlands*]

34: Remove numbered component tags (<c01>, <c02> etc)

Justification(s):

We propose to depreciate [...] elements, that are not used in France: numbered components <c01> <c02>, etc..(these elements are not implemented in the Archivists'Toolkit software). [16. *French EAD and EAC-CPF Working Groups*]

Numbered component tags (<c01> etc) should be abolished. It was felt that the ability to use two tags to represent the same structure again makes data exchange and transformation more complex and is also more difficult to teach as students find it difficult to understand that the component number does not always relate to the same hierarchical level. [79. *UK Archives and Records Association, Data Standards Group*]

c01-c12 elements are useless as different levels of description...can easily be expressed by <c> element alone. [126. Jakob Saternus, *The City Archives of Gothenburg*]

35: The level attribute should be mandatory

[70. UK Archives and Records Association, *Data Standards Group*]

36: Add the possibility of indicating the visibility of any data at the node level, or at least at the component level.

Justification(s):

The ead file has an attribute to indicate whether it is a draft, published, etc. In some cases, especially with long EAD files, it would be interesting to have the possibility of indicating the visibility / status of specific components. For example, if you are writing a finding aid for a fond with 60 series, may be that this work is completed in different steps. It may be interesting to have the possibility of marking with an attribute at the component level (even at the node level), the status (draft, complete, etc.) of the lower-level descriptions. In this way, it would be possible to hide specific data/nodes when displaying the finding aids by means of XSLT. [59. Ricardo Eito-Brun, *Universidad Carlos III de Madrid*]

Wrapper Elements

37: Consider the role and model for <did> and its elements

Justification(s):

Certain elements containing directly text such as <physdesc> or <origination> [within the <did>] need <p> elements rather than </lb> elements to do a line break. [24. *French EAD and EAC-CPF Working Groups*]

The EAD elements <abbr>, <abstract> [...] are not used in France, but they were maintained in the EAC-CPF schema. [33. *French EAD and EAC-CPF Working Groups*]

Add an element <aux/> for unspecified content as a child of <did/>. Similarly to the content model of <unitid/>, this element should allow basic attributes, simple text content and a minimal set of child elements. Often we have to encode auxiliary information which is not covered by the current tag list. Examples are page numbers of printed finding aids, database related parts of composed strings in unitid, or status codes during our checking process. An <aux/> element, maybe in conjunction with a non rendering policy, would provide a place for such information. [48. *Stefan Krause, Editura GmbH & Co. KG, Berlin, Germany*]

While designed to wrap the 'core' elements of description, little use has been found for it in practice and what constitutes the core elements differs in line with national and local standards and practice. The addition of the <abstract> tag in the last revision within the <did> was also seen as unhelpful. It is not used in the UK and its relationship to key descriptive elements such as <title> and <scopecontent> is confused. [71. *UK Archives and Records Association, Data Standards Group*]

38: Consider the role and model for <physdesc> and its elements

Justification(s):

The <physfacet> element needs an attribute @authfilenumber, because the @source attribute is already available in this element. [25. *French EAD and EAC-CPF Working Groups*]

It was proposed that the elements within <physdesc> should be available outside it, particularly in the case of the key element <extent>. Consideration should then be given to whether there is a need for a separate tag for physical description as such particularly if elements from other namespaces such as TEI for example may be used [72. *UK Archives and Records Association, Data Standards Group*]

Remove the recursivity from <materialspec> and <dimensions>: Redundant: these elements are repeatable. [103. *Nationaal Archief, Netherlands*]

Remove <num> from <materialspec>: The element <num> is meaningless as child of this element. [105. *Nationaal Archief, Netherlands*]

Remove <corpname>, <famname>, <geogname>, <name>, <occupation>, <persname>, <subject> and <function> from <physdesc>: Of all access elements only <genreform> is meaningful. [106. *Nationaal Archief, Netherlands*]

39: Deprecate <descgrp>

Justification(s):

<descgrp>: Not used in France: we propose that embedding plural elements be used as in the EAC-CPF schema. [16. *French EAD and EAC-CPF Working Groups*]

Controlled Terms

40: Add ability to encode <controlaccess> and controlled terms in a more granular fashion.

Justification(s):

The EAD element <persname> contains both the family name and the first name, whereas sub-elements <part> are used in the EAC-CPF <nameEntry> element. The solution implemented in EAC-CPF seems to be the best and the most generic, since it applies both to the names of individuals and of corporate bodies. It also makes it possible to isolate qualifiers associated with the name in an entry (for example, location qualifier associated with the name of a corporate body). In addition, it is compatible with the structure adopted by other metadata schemes like MADS (for authority records) and MODS (for resource description) - which is good for interoperability. [39. *French EAD and EAC-CPF Working Groups*]

There isn't an elegant way to add more granularity (i.e. like MARC subfields) within controlled terms in EAD. Since we often share data between MARC and EAD formats, it is imperative that we can have close to the same level of granularity as MARC when exporting from EAD to MARC. We've chosen to embed subfield tags within our controlled terms (both names and subjects) to achieve this end, but this is not ideal. [61. *Mark Carlson, University of Washington*]

<*name>: It should be possible to indicate the parts of person, family and corporate names and possibly place names. The use of the generic <part> tag found in the EAC-CPF <nameEntry> tag was favoured as a model. [79. *UK Archives and Records Association, Data Standards Group*]

In pre-coordinated subject strings where part of a subject string is may be a topic, genre, time period, or a name, the various pieces are undifferentiated, leaving archivists pondering which to prefer. For example, an archivist may wish to convey that records of a business concern cobblers in turn-of-the-20th-century Carlisle, Pennsylvania, or that a personal papers creator wrote extensively about Thomas Jefferson's religious views. Further, many archivists wish to convey these notions in a form compatible with LCSH, such as "Shoemakers--Pennsylvania--Carlisle--1900-1910" or "Jefferson, Thomas, 1743-1826 -- Religion." In the present EAD encoding, an archivist must choose whether to prefer the topical, geographic, or name aspects of such concepts, and has no recourse at all with the temporal characteristics. In MODS, the name, topic, temporal, and hierarchical geographic sub-elements of these subject headings would have their own sub-elements. See: <http://www.loc.gov/standards/mods/mods-outline.html#subject>. To assure backwards compatibility, EAD could permit the inclusion of sub-elements in

<subject> within <controlaccess>, while not requiring these sub-elements.
[90. *Harvard University*]

The purpose of this revision is to enhance the compatibility and functionality of name elements. The granularity of encoding names has a companion question in the granularity of encoding of <controlaccess> terms. In EAD, name parts are not differentiated, thus the textual parts of an authorized form of a name and date or other portions are encoded in a single string that cannot be parsed by computers. This makes EAD name data incompatible with other standard expressions of names, in which the parts of an authorized name are carefully parsed. Both MODS and EAC present models of encoding names in which the various parts can be parsed. See:
<http://www.loc.gov/standards/mods/mods-outline.html#name> and
<http://www3.iath.virginia.edu/eac/cpf/tagLibrary/cpfTagLibrary.html#d1e5077>.
To assure backwards compatibility, EAD could permit the inclusion of sub-elements in <name> within <controlaccess>, while not requiring these sub-elements. An additional question may be whether it would be necessary to deprecate <persname>, <corpname>, and/or <geogname> and exactly how to do so. [91. *Harvard University*]

41: Add a GEOCODE attribute(s) for the geogname tag

Justification(s):

Like the LANGCODE attribute for language, a GEOCODE attribute would enable easier interface with other systems. Longitude and latitude codes could provide archives with the data to more easily build geographic interfaces to their collections and to work with systems that utilize that data. Likely, it would also require a GEOENCODING declaration in the eadheader.
[*Brian Stevens, Western Connecticut State University*]

<geogname>: The group felt that the underlying model for spatial information in EAD should be reviewed alongside the purpose of this tag. Something similar to EAC-CPF's model for <place> (as opposed to place names) might be more suitable. Provision should certainly be made for the addition of geocodes. The model for <placeEntry> in EAC-CPF with attributes for latitude, longitude and altitude is favoured but there must also be the ability to use other national or local systems, such as OS Grid References in the UK context. Consideration should also be given for including data from other namespaces such as KML for example. [77. *UK Archives and Records Association, Data Standards Group*]

42: Allow the SOURCE attribute of <*name> elements to have URLs as values

Justification(s):

The <persname>, <corpname>, etc. elements, have an attribute SOURCE to indicate the origin of the value assigned to them. But the type of the SOURCE attribute is NMTOKEN. It could be interesting to use URLs or URIs to refer to these authority lists. The suggestion is to change the type of the SOURCE attribute to accommodate URIs, as this mechanism is used for a bigger number of lists and is more appropriate for the current evolution of the web. *[Ricardo Eito Brun, Universidad Carlos III de Madrid]*

Digital Archival Objects

43: Allow the elements <dao> and <daogrp> within <altformavail>

Justification(s): This may be useful when copies of the materials being described are digital, such as reproductions used for the communication of documents. *[30. French EAD and EAC-CPF Working Groups]*

44: Allow <dao> to hold (or refer) to metadata about the related digital object

Justification(s):

The <daodesc> element is to be enhanced in order to provide documentary information on digital documents. The <daodesc> element only contains formatting elements (<p>, <list>, <note>, etc...). We suggest that <daodesc> contains more semantic elements in order to differentiate between born-digital and digitized documents, to record technical metadata (filename, version, format, size, volume, number of records, fixed / variable records, life cycle of the document such as its conversion or its migration into another format), to establish links to structured databases, or to specify the number of files if you create a link to a group of files via <daogrp>. We also suggest that <daodesc> includes the elements <accessrestrict> and <userrestrict> concerning the digital document itself (the access and use rights of the digital document may differ from those of the original). *[31. French EAD and EAC-CPF Working Groups]*

In the elements whose purpose is to include a reference to a digital representation of the documents being described in the EAD (e.g. <dao>), the standard should consider the possibility of embedding metadata or links to other metadata records, to include additional technical or descriptive metadata. This would permit EAD users and implementers to use the EAD files as a self-contained entity to describe, for example, a whole file with all their pages digitized. *[55. Ricardo Eito-Brun, Universidad Carlos III de Madrid]*

45: Add a digital archival object identifier

Justification(s):

The purpose of this revision is to provide a place in EAD for the system identifier of a digital object. This data would support management of digital objects by facilitating data sharing between systems describing digital objects and systems storing digital objects.

Unlike the current digital archival object links (dao href attribute), this would not be a reference to a delivery version of a digital archival object, but would instead be an identifier of an archival object in its host system.

The identifier should be repeatable to allow for tracking a local repository copy and a copy in a shared repository, e.g. HATHI.

Uses include cascading updates made in EAD to the digital object's host system metadata. For example, a title change in EAD could be sent to update a METS file in the digital object's host system.

The proper location for this new element would be as a new sub-element or set of attributes on the <dao> and <daogrp>. A possible name for this element is the "Digital Archival Object Identifier."

Given that a digital object may be described in an EAD instance, but not yet available for research, if this element is made valid for the <dao> and/or <daogrp>, then a <dao> lacking a link, but with a <doaid>, must be valid. In fact, implementing this enhancement may be especially valuable to contemporary archivists who face the influx of hybrid paper and electronic collections. Archivists can currently label and box analog collections and control their identifiers in EAD regardless of their restriction status. Controlling the whereabouts of virtual materials requires a new type of identifier.

Identifiers might conform to established schemes such as ARK, URN, or PURL, or may be local. A "type" attribute would be needed to specify. One type would require

Another type would be a resolvable identifier that leads to a name resolution system that interprets an URN and returns a system identifier and a system number.

Examples:

```
<daoid type="other" othertype="Cornell University  
eCommons">12345</daoid>
```

```
<daoid type="urn">http://nrs.harvard.edu/urn-4:hul.arch:12345</daoid>  
[89. Harvard University]
```

46: Expand valid options for locating <dao>s

Justification(s):

At least one archival repository at Harvard is doing extensive on-demand digitization of individual items in folders. The primary access point for these digital copies is the EAD-encoded finding aid, many of which were created in the early days of EAD before digitization was common. The finding aids in question contain <list>s of <item>s, rather than nested <c>s below the folder level. It would be advantageous if <dao>s were valid within <item>s. If other institutions are facing the same issue, it makes sense for the EAD community to discuss options. [96. *Harvard University*]

Date Encoding

47: Review model for encoding dates

Justification(s):

Consider following TEI P5's model for date encoding. It's much better than what's currently available with EAD 2002. I understand the inclusion of a "type=bulk" option in EAD, but I don't see any advantage for a generic "type=inclusive". Instead, TEI (which inspired the creation of EAD originally) opts for "to" and "from" attributes (among other things) for its dates. Additionally, the TEI P5 schema is excellent for encoding machine-readable dates, and if you try to encode an impossible date, you will invalidate the file (with EAD 2002's schema, however, invalid dates can be present without any issue). For instance, if you try to include <date when="2010-02-29"> (a Gregorian date that doesn't exist) in a TEI file, the schema won't allow it. <date when="2008-02-29">, however, will validate as expected. See here for more information: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-att.dateable.w3c.html>. [3. *Mark Custer, East Carolina University*]

The EAD element <date> is not the same as the EAC-CPF element <date>; one should differentiate between the dates of the described documents (contained in <unitdate>) and the other significant dates mentioned in other elements than <unittitle> or <scopecontent>; in addition, the "bulk" value of the attribute @type is useful for archival institutions (it enables to mention predominant/bulk dates within a wide date range) and the notion of "predominant/bulk dates" should be kept. [35. *French EAD and EAC-CPF Working Groups*]

<*date>: The date elements in EAD should be able to express ranges, uncertainty, complex dates and normalisation. The model in EAC-CPF would be a good starting point but it should also be possible to capture precise time and time zone information as well. [74. *UK Archives and Records Association, Data Standards Group*]

Add begin year and end year attributes for <unitdate> to facilitate indexing of date information in discovery systems in order to support date range searching. For some systems, parsing the "normal" attributes of <unitdate>s for date range searches can slow the delivery of search results beyond reader tolerance. In other systems-- those in which EAD instances are indexed along with MARC data-- the availability of simple, separately parsed 4-digit dates that are compatible with MARC 008 date1 and date2 is of great value. Other repositories and implementations may have uses for 4-digit dates. If other institutions are using or in need of similar data, it makes sense to encode the data in a standardized way across the archival community. Harvard University has implemented this through a local change to the EAD schema, via the attributes beginYear and endYear on the <unitdate> element. This attribute and its value are machine-generated when an EAD file is ingested by OASIS, our EAD catalog. We suggest that these should be an optional, non-repeatable attributes on <unitdate>. [94. Harvard University]

For <unitdate> and <date>, eliminate date spans; either encode beginning and end date as separate tags, or use attributes to distinguish type of date in relation to a span. Processing of dates using current encoding is problematic at best. In addition, use of other attributes on <unitdate> is compromised with spans. For instance, a date span with different calendar values (beginning date is BCE and ending date is CE) can't correctly be encoded. The same would be true of the certainty attribute. [112. Mary Lacy, Library of Congress]

Languages and Scripts

48: Allow more granular assignment of languages and scripts

Elements <corpname>, <geogname>, <persname>, <famname>, <name>, <title>, <unittitle>: Attributes to specify the language should be made available within these elements (see comment about @xml:lang attribute) as well as script code attributes, in order to manage parallel forms (see for example the description of documents in Arabic, Chinese or Hebrew).

The addition of these attributes would solve the problems about bilingual finding aids. The EAD Tag Library deals with this problem, but it is not possible to specify the language used within an element: <language> within <eadheader> serves to indicate the language(s) of the finding aid, but this information can only be provided once, unlike <langmaterial> which may appear in all the <did> elements. An attribute such as @xml:lang would serve to encode the language of each <scopecontent>, <note> or <unitid>, etc., so that we could build a visualization tool enabling the user to click on a button "French" or "English" to display the descriptions in the required language. Currently, we must record the descriptions in both languages in the same tag and it's unnecessarily time consuming. [23. French EAD and EAC-CPF Working Groups]

Enable the capture of the language and script of names and any other appropriate entities by addition of language attributes. [81. *UK Archives and Records Association, Data Standards Group*]

All elements with PCDATA should have an attribute for a language code: Visual disabled people are using voice browsers, which need a trigger for the correct language. Every element and even every word within an element can be written in a different language than the language of the main text. [97. *Nationaal Archief, Netherlands*]

EAD already provides the possibility to specify the language of the material on general as well as on lower levels, plus the possibility to name the language, in which the finding aid as a whole is written (<langmaterial> resp. <language>).

A question that arose in the APEnet project with regard to multilinguality would be a possibility to f.i. give the <unittitle> or <scopecontent> information in more than one language and specifying the language used directly with these elements.

Currently the APEnet project only has implemented a workaround with the TYPE attribute coming with <titleproper>, using the language code following ISO standard 639-2b as possible value for the TYPE attribute.

An alternative could be, to introduce xml:lang for this purpose or to think of a new "EAD specific" attribute to be used in combination with the already existing <language>, like f.i. naming the generally used language in <language> and specifying variations with an attribute like "otherlang" or similar, which then could be made available for all subelements of <eadheader> and <archdesc>. [116. *APEnet project (Archives Portal Europe)*]

Other Elements

49: Review model for <address> and use of <addressline>

Justification(s):

The <address> element should be structured as it is in the TEI schema. We think the <addressline> element is of extremely limited value. We propose that a typing attribute be made available as it was done for EAC-CPF. [20. *French EAD and EAC-CPF Working Groups*]

The EAD elements [...] <address> and <addressline> are not used in France, but they were maintained in the EAC-CPF schema; <addressline> should be made more consistent with EAC-CPF (see Section 3). [33. *French EAD and EAC-CPF Working Groups*]

An address line could contain a link (<extref> to an email address or to a website. [107. *Nationaal Archief, Netherlands*]

50: Allow <scopecontent> within <archref>

Justification(s):

A <scopecontent> sub-element would enable us to describe correctly related or separated archival material (all the sub-elements of <did> element are available within <archref>, but there are very few other elements; <scopecontent> is not available, so that we have to use <note> or <unittitle>). [26. *French EAD and EAC-CPF Working Groups*]

51: Remove <archref>, <bibref> and <title> from <container>, <langmaterial>, <materialspect>, <origination>, <physdesc>, <dimensions>, <extent>, <physfacet>, <physloc>, <repository>, <unitdate>, <unitid>

Justification(s):

The elements <archref>, <bibref> and <title> are meaningless as child of these elements. [104. *Nationaal Archief, Netherlands*]

52: Don't allow <acqinfo> element within <custodhist>

Justification(s):

The <acqinfo> element may be contained within <custodhist>. However, because <custodhist> serves to encode information about the chain of ownership of the materials being described, before they reached the immediate source of acquisition, it would be better if <acqinfo> was not embedded within <custodhist>. [27. *French EAD and EAC-CPF Working Groups*]

53: Permit simplified encoding of indexes

Justification(s):

I've long felt that the encoding of supplemental indexes is cumbersome at best. The current <index> is much like a book index, with the <ref target> pointing to a specific item or specific something in the collection. And the multiple tags for one entry are just tedious! Most of our indexes are actually "lists" of names for a specific series, such as correspondence. We have others that definitely fit the category of <odd>, such as supplemental lists of exhibitions, or board members, or something like that. Lists of names associated with correspondence series is such a common feature of finding aids for manuscript collections/personal papers, that something a little

less fussy is desirable. [83. *Barbara Aikens, Archives of American Art, Smithsonian Institution*]

The purpose of this revision is to permit repositories to reduce the amount of time and effort required to mark up indexes by eliminating some of the currently required sub-elements. It may be that <list> was originally conceived as a mechanism to produce such simplified encoding within <index>, but an <index> that includes only <list> and <item> sub-elements will not validate against the schema. This may be merely a bug in the schema, but it may also be intentional. The following is an example of what seems to be the least complex valid combination of sub-elements of <index>:

<indexentry><persname>Mary Smith</persname> <ref target="indexEntry1">Folder 5</ref> </indexentry>. That encoding, however, also requires insertion of the value of the target attribute in the ID attribute value of the targeted element. Such targeting is extremely time-consuming.

What is desired is encoding within <index> as simple as one of the examples below or some other combination of tags that the Technical Subcommittee may determine. Possible examples:

```
.. <indexentry>Mary Smith, Folder 5</indexentry>
.. <list><item>Mary Smith, Folder 5 </item> </list>
.. <list><defitem><label>Mary Smith</label> <item>Folder
5</item></defitem>
[95. Harvard University]
```

54: Simplify the content model of the list elements.

Justification(s):

Deprecate the elements <chronlist> and <list type="deflist">, and replace these by <table>. The <list> element will contain the type attribute with the values simple, marked and ordered. The use of attributes should be validated by the Schema, so <list type="marked" numeration="arabic"> should be invalid.

The mark attribute should get a set of closed values, e.g. bullet, box, dash, hyphen, or some of the HTML entity names and/or numbers or some of the Unicode character codes. [107. *Nationaal Archief, Netherlands*]

55: Keep element <odd>

Justification(s):

There are often some information in Swedish archival descriptions that doesn't fit anywhere else. [100. *Karin Bredenberg, The Swedish National Archives*]

56: Clarify the use of this element <processinfo>

Justification(s):

<processinfo> is equivalent to the 3.7.1. element of ISAD(G) (explaining how and by whom the description is established) or is its use broader, as suggested in the Tag Library (to provide “information about accessioning, arranging, describing, preserving, storing, or otherwise preparing the described materials for research use”)? See also general comment about the <control> element of EAC-CPF and the management of EAD documents. [29. *French EAD and EAC-CPF Working Groups*]

57: Resolve issue of normalization of titles with initial articles

Justification(s):

EAD can not handle the exclusion of initial articles of titles. For instance, a title such as “Le Rouge et le noir” is ordered with the letter “R” and not with the letter “L”. It seems there is no solution:

NORMAL="Rouge et le noir": the index is not understandable

NORMAL="Rouge et le noir (Le)": problems may occur with the alphabetical order

NORMAL="Le Rouge et le noir": it is contrary to cataloguing practices. [32. *French EAD and EAC-CPF Working Groups*]

58: Add ability to encode title elements in a more granular fashion

Justification(s):

<title> element: an attribute @certainty would enable us to distinguish between the title given by the author, the usual title of a work or of a document, and the title given by the person who arranged or described the documents but which does not correspond to the content (see, for example, some drawings of Victor Hugo); it would be useful to add an attribute @resp like in TEI to indicate the responsibility of the creator of the title (this comment also applies to <unittitle>). [22. *French EAD and EAC-CPF Working Groups*]

Semantic separation of elements in <unittitle>.

EAD currently includes a number of semantically distinct elements in <unittitle> which are represented separately from the title in other metadata schemas, such as <unitdate>, <edition>, and <imprint>. If desired, the <edition> and <imprint> elements could be made available within the

Descriptive Identification wrapper in the same way as <unitdate>. Separation of these elements from the title would simplify exchange of information between metadata schemas, while improving consistency of encoding practice across repositories.

In our work here at BYU, we have avoided using <imprint> due to its inclusion in the <unittitle> element. However, we do have many large collections of published materials that we are treating according to archival principles which would benefit from the use of these elements. [52. *Cory Nimer, Brigham Young University*]