24 The Independent Header

Many libraries, text repositories, research sites and related institutions collect bibliographic and documentary information about machine readable texts without necessarily collecting the texts themselves. Such institutions may thus want access to the header of a TEI document without its attached text in order to build catalogues, indexes and databases that can be used by people to locate relevant texts at remote locations, obtain full documentation about those texts, and learn how to obtain them. This chapter of the Guidelines describes a set of practices by which the headers of TEI documents can be extracted from those documents and exchanged as freestanding independent TEI documents. Headers exchanged independently of the documents they describe are called *independent headers*.

This chapter outlines practices recommended for encoders (especially those responsible for the documentation of text) when creating independent headers to be distributed, and specifies the set of recommended elements that should be included in the independent header. Of interest to librarian cataloguers who may receive independent headers from remote sites, it also discusses the relationship between the elements of TEI headers and MARC tags, in order to facilitate the cataloguing of these headers or the loading of independent headers into local MARC-based bibliographic databases. This chapter does *not* describe how to create a header. Guidance on the creation of headers and descriptions of each element in the header can be found in chapter 5 *The TEI Header*.

24.1 Definition and Principles for Encoders

An *independent header* is a header extracted from a TEI text that can be exchanged as an independent document between libraries, archives, collections, projects, and individuals. The file description of the independent header (enclosed by the <fileDesc> element) can be used to generate bibliographic records. The profile description, encoding description, and revision history (encoded by the <profileDesc>, <encodingDesc>, and <revisionDesc> elements) can form part of a bibliographic description or, more appropriately, be used as an attached 'codebook' for full documentation of the analysis of the text and how it was encoded. Thus, the independent header can serve as the primary means by which libraries, archives, related repositories, research projects, and individual researchers can obtain bibliographic, descriptive, and full documentary information on machine-readable texts that reside in remote locations.

The distribution and retrieval of independent headers also facilitates resource discovery by other means. The mappings to MARC discussed in the remainder of this section form one example of how the information embedded in TEI Headers may be re-used; with more recent developments such as the Open Archives Initiative protocol and the Z39.50 Bath Profile (Interoperability) it becomes possible to define other protocols for data exchange. A key element here will be the establishment of mappings between the components of the TEI header and those of the Dublin Core expressed in XML. It is hoped to document such mappings in future editions of these Guidelines.

The structure of an independent header is exactly the same as that of a <teiHeader> attached to a document, and can therefore be validated using the same document type definition (DTD). In practice, this means that a <teiHeader> and its DTD can be extracted from a TEI document and shipped to a receiving institution with little or no change. However, some fields that are listed as "optional" in the header are listed as "recommended" for the independent header. For this reason, this chapter should be consulted in connection with any plan to send headers as independent documents.

When deciding which information to include in the independent header, and the format or structure of that information, the following should be kept in mind:

The independent header should provide full bibliographic information on the encoded text, its source, where the text can be located, and any restrictions governing its use.

The independent header should contain useful information about the encoding of the text itself. In this regard, it is highly recommended that the encoding description be as complete as possible. The Guidelines do not require that the encoding description be included in the header (since some simple transcriptions of small items may not require it), but in practice the use of a header without an encoding description would be severely limited.

The independent header should be amenable to automatic processing, particularly for loading into databases and for the creation of publications, indexes, and finding aids, without undue editorial intervention on the part of the receiving institution. For this reason, two recommendations are made regarding the format or structure of the header: first, where there is a choice between a prose content model and one that contains a formal series of specialized elements, wherever possible and appropriate the specialized elements should be preferred to unstructured prose. For instance, the source description can contain either a free-prose citation (tagged <bibl> or even) or a <biblStruct> element, which provides a more rigorous structure for the bibliographic information (see examples in section 6.10 Bibliographic Citations and References). The more structured <biblStruct> element is more suitable for automatic processing, and is therefore recommended over the less structured alternatives whenever the header is to be exchanged as an independent header. Second, with respect to corpora, information about each of the texts within a corpus should be included in the overall corpus-level <teiHeader>. That is, source information, editorial practices, encoding descriptions, and the like should be included in the relevant sections of the corpus <teiHeader>, with pointers to them from the headers of the individual texts included in the corpus. There are three reasons for this recommendation: first, the corpus-level header will contain the full array of bibliographic and documentary information for each of the texts in a corpus, and thus be of great benefit to remote users, who may have access only to the independent header; second, such a layout is easier for the coder to maintain than searching for information throughout a text; and third, generally speaking, this practice results in greater overall consistency, especially with respect to bibliographic citations.

24.2 Required and Recommended Tags

The richness and size of the header reflect the diversity of uses to which electronic texts conforming to these Guidelines will be put. It is not intended, however, that all of the elements recommended in this chapter be present in every header. As described in section 5.6 *Minimal and Recommended Headers*, the TEI header allows for the provision of a very large amount of information concerning the text itself, its source, encodings, and revisions as well as detailed descriptive information that can be used by researchers in analysing the text. The amount of encoding will depend on the nature and intended use of the text. At one extreme, an encoder may expect that the header will only provide bibliographic information about the text adequate to local needs. At the other, wishing to ensure that their texts can be used for the widest range of applications, encoders will want to document as explicitly as possible both bibliographic and descriptive information in such a way that no prior or ancillary knowledge about the text is needed in order to process it. The header, in the latter case, will be very full, approximating the kind of documentation often supplied in the form of a manual. Most texts will lie somewhere between these extremes; textual corpora in particular will tend toward the latter extreme.

The following is a list of the components of the header, in the order in which they are presented in chapter 5 *The TEI Header*, together with an indication of their importance in constructing an independent header.

<fileDesc> required. Some subelements are required, others optional or recommended:

<titleStmt> required; subelements are required or optional:

<title> required<author> required, if known<sponsor> optional<funder> optional<principal> required, if known<resp> required, if known<role> and <name> required, if known, when the responsibility is not an author,
sponsor, funding body, or principal researcher. Details may be found in
section 5.2.1 The Title Statement.<editionStmt> recommended
<edition> recommended

<resp> recommended

<rol><role> and <name> recommended primarily to distinguish editions.

<extent> optional

<publicationStmt> required

<publisher>, <distributor>, or <authority> at least one is required

<pubPlace> recommended

<address> recommended; prose is sufficient

<idno> recommended

<availability> recommended

<date> recommended

<seriesStmt> optional

<title> required

- <idno> recommended
- <resp> and <name> optional

<notesStmt> recommended

<**sourceDesc>** required. As much information as possible should be provided to identify the source, where one exists. In the case of items 'born digital', the source description is still mandatory, and should contain a note like the following:

<sourceDesc>

No source: this document was created in digital form. </sourceDesc>

Where the source document is itself a TEI document, the <biblFull> element should be used, as discussed in section 5.2.8 *Computer Files Derived from Other Computer Files*. In other cases, the following elements are either required or recommended, though other elements not listed here should be used wherever applicable in order to provide an accurate identification of the source.

-

 <biblStruct> recommended (a full discussion of <biblStruct> is given in section 6.10 *Bibliographic Citations and References*).
- <analytic> required when the citation describes an item within a larger collection, such as an essay within a collection or an article in a journal, and is not an independent publication. If used, it should contain the following elements in this order:
 - <author> required, if known
 - <title> required
 - <editor> recommended
- <monogr> mandatory when applicable; this element should contain the following
 elements in this order:

<author> required, if known.

- <title> required. The level attribute must be used to indicate whether this is the title of a book, journal, or series. It is highly recommended that the type attribute be used to distinguish the main title from subordinate, parallel, or other titles. All elements that indicate intellectual responsibility for a work, such as <editor>, are also required, if known.
- <imprint> required.

<pubPlace> required, if known.

<org> recommended.

<date> required. If the date is unknown, n.d. may be used.

<idno> recommended.

<series> required, if the item is part of a series.

<title> required, but type attribute is optional.

<scriptStmt> required for transcribed speech. See section 5.2.9 Computer Files Composed of Transcribed Speech.

<recordingStmt> mandatory when applicable:

<resp> and <name> recommended

- <recording> recommended
- <equipment> recommended
- <broadcast> recommended
 - <comment> optional
- <encodingDesc> recommended, especially for projects, collections, or corpora. If the <encodingDesc> element is used, it is recommended that it contain one or more of the following elements, rather than a prose description. See section 5.3 *The Encoding Description*.

<projectDesc> optional

<samplingDecl> optional

- <editorialDecl> recommended; it is also recommended that the editorial declaration make use of the specialized elements defined in section 5.3.3 *The Editorial Practices Declaration*, rather than only consisting of prose paragraphs. Prose may of course be used in addition to these elements for material otherwise not handled.
- <tagsDecl> recommended
- <refsDecl> optional in general, but recommended if a standard referencing system is built into the encoded works. Section 5.3.5 *The Reference System Declaration* describes three different methods for documenting the referencing system: the prose method, the stepwise method, and the milestone method. No preference is expressed for one type of method over another, since this depends on the convenience of the encoder and the likely efficiency of the particular software applications envisaged for the text. Only one method can be used within a single <refsDecl> element. If a text uses both hierarchical and milestone tagging, this can only be described in prose.
- <classDecl> required where the scheme attribute has been used to identify the classification scheme or taxonomy used by any of the elements <keywords>, <classcode>, <occupation>, or <socecstatus>. Even where this is not done, this element may usefully document the classification employed, either explicitly as a series of <taxonomy> elements, or implicitly by means of bibliographic citation.
- <profileDesc> recommended
 - <langUsage> recommended

<language> recommended

- <textDesc> optional in most instances, but recommended when the encoder wants to provide a full description of the situation within which a text was produced or experienced, characterize it in a relatively continuous manner (in contrast to discrete categories based on type or topic), and believes that this characterization of the text will be helpful to the understanding, analysis, or retrieval of this text by remote users. If a collection or corpus uses a pre-existing descriptive typology as its organizing principle, it is recommended that its components be re-expressed in terms of the parameters listed here. If the encoder believes that pre-existing text categories (such as a standard classification scheme) are sufficient, then it is recommended that the <textClass> element be used instead. See section 23.2.1 *The Text Description* for details and guidance.
 - <textClass> optional in most instances; this element may may be used as an alternative or in addition to the <textDesc> element. <textClass> is recommended in the following situations:

- * a standard text category, such as the Library of Congress List of Subject Headings or a Dewey Decimal Classification category, clearly describes the text
- * situational parameters (or the demographic elements of the <particDesc> element) are used and a text category can be constructed by the encoder based on a recurring set of values for those parameters.

See section 5.4.3 *The Text Classification* for details and guidance. One or more of the following sub-elements can be used.

- <keywords> recommended only if using a standard thesaurus such as the Library of Congress List of Subject Headings, a discipline-specific thesaurus, or a thesaurus defined explicitly in the header. In each case, the source should be indicated by the scheme attribute and defined in the <classDecl> element.
- <classCode> recommended only if the text is categorized by an internationally accepted classification scheme, such as the Dewey Decimal or Universal Decimal classification schemes. The scheme should be indicated by the scheme attribute and defined in the <classDecl> element.
- <catRef> optional in most instances, but recommended when a user-defined classification is in use. The scheme should be indicated by the scheme attribute and defined in the <classDecl> element.
- <particDesc> optional, but recommended for spoken text when the encoder judges that such information is useful to remote users in the analysis of that text, and for both written and spoken text if such information is useful in the analysis of language usage. For details and guidance, see section 23.2.2 The Participants Description.
 - <participant> or <particGroup> recommended. Though the substructure of both the <participant> and <particGroup> elements can be prose, in independent headers one or more of the following sub-elements providing more specific details should be used in preference to prose. Users of these Guidelines are free to extend the set of headings listed below.

<name> recommended when the information is available

-

-
 birthPlace> recommended when the information is available
- <firstLang> recommended when the information is available
- angKnown> recommended when the information is available
- <residence> recommended when the information is available
- <education> recommended when the information is available
- <affiliation> recommended when the information is available
- <occupation> it is recommended that, where possible, the classification of the trade, occupation, or profession be derived from a standard classification or taxonomy, and that the source taxonomy be identified in the scheme attribute.
- <**socecstatus>** it is recommended that, where possible, the encoding of social and economic status be derived from a standard classification or taxonomy, and that the source taxonomy be identified in the scheme attribute.
- <particRelations> optional, but recommended where it is judged by the encoder that such information is important to the analysis of the text. If the <particRelations> tag is used, it is recommended that the special purpose <relation> element be used. See section 23.2.2 The Participants Description.

- <**settingDesc>** optional, but recommended when the encoder judges that this information is useful in the analysis of the text, particular in the analysis of language usage.
- <revisionDesc> required in the independent header when available. It is recommended that the <revisionDesc> be encoded as a series of <change> elements, most recent first, each containing a <date>, one or <respStmt>s and an <item>.

Further discussion of requirements and recommendations with respect to usage of the components of the TEI header is given in section 5.6 *Minimal and Recommended Headers*.

24.3 Header Elements and their Relationship to the MARC Record

This section offers some guidance to both cataloguers and bibliographic analysts who want to load TEI independent headers into a MARC-based retrieval system. Because there are variations in cataloguing practice across local sites, among bibliographic utilities (such as OCLC and RLIN), and differences in MARC usage in different countries, only tentative advice is possible. Note that the following examples are based on USMARC, *not* UNIMARC.¹⁶³ UNIMARC offers cataloguers in different countries the opportunity to combine different national practices in a single MARC format, and is the preferred variety of MARC records for distribution across national boundaries. The implementation of UNIMARC, however, will be affected by local practice and by guidelines offered by the bibliographic utilities. Though UNIMARC is a stable format, the guidelines for its implementation are not sufficiently known or stabilized to be included in this chapter.

There are some major differences between the MARC record and the TEI header that will cause problems for librarians trying to map from the TEI independent header to the MARC record. The most important difference between the MARC record and the TEI header is the function of each. Despite the efforts and claims of some members of the library community, the MARC record remains fundamentally an electronic version of the catalogue card, with the limitations of its model.¹⁶⁴ The catalogue card is a unitary record for a physical object containing complex *bibliographic data* of varying sorts. The catalogue card, as well as documentary non-bibliographic information that supports the analysis, either by humans or machines, of the electronic text documented by header. Most of this analytical information, which is found in profile description, encoding description, and revision history, has little direct provision for it in the MARC record, and if retained must be recorded as unstructured notes (55XX) fields. Notes fields usually do not have the structure to support machine retrieval and analysis, while properly formatted profile, encoding, and revision lend themselves to retrieval, can support machine processing (including analysis), and point directly to the electronic text attached to the header. Moreover, the electronic text points back to the relevant elements in the header.

Though this chapter offers some advice on where the profile, encoding, and revision descriptions might go in a MARC record, for practical reasons a repository might want create a codebook from these divisions of the header, and create a MARC record from the file description only. The MARC record should contain a reference to the codebook.

Subfields (or delimiters) are conventionally indicated by the dollar sign.

¹⁶³ For more information on UNIMARC, see Brian P. Holt, UNIMARC Manual (London, U.K.: IFLA Universal Bibliographic Control and International MARC Programme, British Library, 1987). For USMARC, see Walt Crawford, MARC for library use: understanding USMARC (Boston: G.K. Hall, 1989), USMARC format for bibliographic data, including content designation (Washington, D.C.: Library of Congress, 1987), and Deborah J. Byrne, MARC manual : understanding and using MARC records (Englewood, Colo.: Libraries Unlimited, Inc., 1991).

¹⁶⁴ The primary function of the MARC record when it was first designed in the mid-1960s was to allow for the electronic distribution of cataloguing records in support of card production. See Henriette Avram, *The MARC Pilot Project* (Washington D.C.: Library of Congress, 1968), p. 3. For discussion of the relationship between the MARC record and the catalogue card, see Michael Gorman, *After AACR2R: The Future of the Anglo-American Cataloging Rules*, in Richard Smiraglia, ed., *Origins, Content and Future of AACR2 Revised* (Chicago: American Library Association, 1992).

24.4 MARC Fields for the File Description

Note that there is no provision for the 'Main Entry' (or USMARC 1XX fields) in the TEI header. The main entry should be constructed, using appropriate name authority control, by the cataloguer from information derived from the header that indicates who is primarily responsible for the intellectual content of the work. There is an <author> tag, but the form of the name will have to be checked by a cataloguer before the main entry is constructed.

- <titleStmt> corresponds to title and statement of responsibility fields in MARC, typically 240 (for uniform title) and 245 (for title proper).
- <title> 240 \$a (for uniform titles) or 245 \$a fields. Put any subtitles in 24X \$b. Insert the constant, "[computer file]" in the 24X \$h gmd subfield.

The elements <sponsor>, <funder>, and <principal> all belong in the 245 \$c subfield: statement of responsibility, as in the following example:

```
<titleStmt>
<titleStmt>
<title>Two stories by Edgar Allen Poe: electronic
version</title>
<author>Poe, Edgar Allen (1809-1849)</author>
<respStmt>
<respStmt>
<name>James D. Benson</name>
</respStmt>
</titleStmt>
```

This might be tagged in MARC as:

```
245 Two stories by Edgar Allen Poe :$belectronic version ; compiled by $cJames D. Benson.
```

The <edition> and <name> (within responsibility statement) elements correspond with MARC fields 250 \$a and 250 \$b respectively, as in the following example:

```
<editionStmt>
<edition>Student's edition,
<date>June 1987</date>
</edition>
<respStmt>
<resp>New annotation by</resp>
<name>George Brown</name>
</respStmt>
</editionStmt>
```

This might be tagged in MARC as:

250 \$aStudent's edition, June, 1987, new annotation by \$bGeorge Brown.

The <extent> element is analogous to the 'Physical Description' MARC field. Fields 256 or 3XX are appropriate, depending on local practice. The <date> element in this context corresponds with the 260 \$c, and appropriate fixed fields. The <publisher>, <distributor>, or <authority> elements correspond with the MARC field 260 \$b, while the <publice> element corresponds with field 260 \$a, as in the following example:

```
<publicationStmt>
  <publisher>Columbia University Press</publisher>
  <pubPlace>New York</pubPlace>
  <date>1993</date>
</publicationStmt>
```

This may be tagged in MARC as:

260 \$aNew York :\$bColumbia University Press, \$c1993.

Local practice will determine appropriate MARC fields for <address>, <idno>, and <availability>. Restrictions on access should normally be placed in the 506 field, while the place where an item may be ordered will be located in a local notes (590) field. If local practice warrants it, the address of the publisher should be indicated in the 260 field.

The series <title> and the <idno> should be placed in the appropriate 490 fields (series untraced), if series authority checking needs to be done. Further, because the TEI tags do not differentiate between name, conference, or title series, there is no simple mechanical method for determining which MARC tag (410, 411, etc.) should be used. Safe practice would be to load any series statements into 490 fields, and then to conduct authority work on those fields.

The <notesStmt> element is usually reserved for general note (500) fields.

The <sourceDesc> can be mapped to be a 'source of data' note (537 in RLIN MDF format) with the print constant "Transcribed from:" at the beginning of the note. The <biblStruct> itself can be mapped onto a 581 field (note on primary publication) using the ISBD format to separate each data element.

The <scriptStmt>, <recordingStmt>, <recording>, <equipment>, and <broadcast> elements do not easily map to existing MARC fields, and should be put into a local notes field (590) treating the TEI tag introducing each component as a print constant at the head of the field in order to facilitate future local processing and retrieval. Example:

```
<scriptStmt id="cnn12">
  <bibl>
      <author>CNN Network News</author>
      <title>News Headlines</title>
      <date>12 Jun 1991</date>
  </bibl>
</scriptStmt>
```

This may be tagged in MARC thus:

Example:

```
<recordingStmt>
  <recording type="video" dur="10 mins">
      <equipment>
         Recorded from FM radio to chrome tape
      </equipment>
      <broadcast>
         chihl>
            <title>Britain's pleasure parade</title>
            <author>BBC Radio 4 FM</author>
           <editor role="interviewer">Robin Day</editor>
           <editor role="interviewee">Margaret Thatcher</editor>
            <series> <title>The World Tonight</title> </series>
           <date>27 Nov 89</date>
         </bibl>
      </broadcast>
  </recording>
</recordingStmt>
```

This can be tagged in MARC as:

```
590 <recordingStmt>
<recording type="video" dur="10 mins">
        <equipment>
            Recorded from FM radio to chrome tape
            </equipment>
            <broadcast>
            <bibl>
            <title>Britain's pleasure parade</title>
            <author>BBC Radio 4 FM</author>
            <editor role="interviewer">Robin Day</editor>
            <editor role="interviewer">Aragaret Thatcher</editor>
```

24.5 MARC Fields for the Encoding Description

The <encodingDesc> element provides useful information documenting the relationship between an electronic text and the source or sources from which it was derived. The <projectDesc>, <sam-plingDecl>, <editorialDecl>, and <refsDecl> elements provide details of decisions and rationales used about the process and purposes of the project, how text was sampled, principles of editorial practice, and how canonical references are constructed. The 567 field (notes on methodology) appears to be the most appropriate for this sort of information, though this field is normally intended for methodologies characterizing the social sciences. Practically, it would be wise to transcribe the <projectDesc>, <editorialDecl>, <refsDecl>, and <classDecl> elements directly as one or more 567 fields without intervention, with the element name at the beginning of each field, and any TEI tags left intact. This may facilitate any locally-developed retrieval software.

Example:

```
<encodinaDesc>
       <projectDesc>
         Texts were collected to illustrate the full range of
           twentieth-century spoken and written Swedish, written by native
           Swedish authors.
       </projectDesc>
       <samplingDecl>
         Sample of 2000 words taken from the beginning of the text.
       </samplingDecl>
       <editorialDecl>
         <interpretation>
           Errors in transcription controlled by using the SUC spell
             checker, v.2.4 
         </interpretation>
       </editorialDecl>
     </encodingDesc>
This may be tagged in MARC as:
     567
     <projectDesc>
           Texts were collected to illustrate the
```

```
Texts were collected to illustrate the
full range of twentieth-century spoken and written
Swedish, written by native Swedish authors.
</projectDesc>567 <samplingDecl>
Sample of 2000 words taken from the
beginning of the text.
</samplingDecl>567 <editorialDecl>
<interpretation>
Errors in transcription controlled
by using the SUC spell checker, v. 2.4
</interpretation>
</editorialDecl>
```

24.6 MARC Fields for the Profile Description

The profile description is the most problematic element in the TEI header for librarian cataloguers, because it provides a detailed description of the *non-bibliographic* aspects of the text, specifically the languages and sublanguages used, the situation in which it was produced, and the participants and their setting. This information can be used for retrieval purposes or in machine-supported analysis of the text. The information can be loaded into a separate 'codebook' and referenced by the MARC record. Little guidance can be offered on the appropriate MARC location for the elements that make up the profile

description, except to suggest that if a site wants to load the profile description into a MARC record for archival and possibly retrieval purposes, then the contents of the profile description may be mapped into a locally-defined notes field (59X) with its TEI tags intact, as in the examples above.

24.7 MARC fields for the Revision Description

The revision history (<revisionDesc>) logs all changes to a machine readable file whether or not these constitute a new edition of the file. Aside from the edition area of the MARC record, there are no MARC fields that deal specifically with changes of this sort. This information might be best included in a 'codebook', rather than a MARC record. As before, the simplest way of approaching this problem is to include the material with its TEI tags intact as a locally-defined note (59X) in order to support future local processing.

24.8 Structure of the DTD for Independent Headers

The following document type definition is provided in file teishd2.dtd and constitutes the auxiliary DTD for independent headers as described in this chapter.

```
<!-- 24.8: File teishd2.dtd: Auxiliary DTD for Independent Header-->
<!--Text Encoding Initiative Consortium:
Guidelines for Electronic Text Encoding and Interchange.
Document TEI P4, 2002.
Copyright (c) 2002 TEI Consortium. Permission to copy in any form
is granted, provided this notice is included in all copies.
These materials may not be altered; modifications to these DTDs should
be performed only as specified by the Guidelines, for example in the
chapter entitled 'Modifying the TEI DTD'
These materials are subject to revision by the TEI Consortium. Current versions
are available from the Consortium website at http://www.tei-c.org-->
<!--Embed entities for TEI generic identifiers.-->
<!ENTITY % TEI.elementNames PUBLIC '-//TEI P4//ENTITIES Generic
Identifiers//EN' 'teigis2.ent' >%TEI.elementNames;
<!--Embed entities for TEI keywords.-->
<!ENTITY % TEI.keywords.ent PUBLIC '-//TEI P4//ENTITIES TEI
Keywords//EN' 'teikey2.ent' >%TEI.keywords.ent;
<!--Define element classes for content models, shared
attributes for element classes, and global attributes. (This all
happens within the file teiclas2.ent.)-->
<!ENTITY % TEI.elementClasses PUBLIC '-//TEI P4//ENTITIES TEI
ElementClasses//EN' 'teiclas2.ent' >%TEI.elementClasses;
<!--Now declare the IHS element.-->
<!ELEMENT ihs %om.RO; (teiHeader+)>
<!ATTLIST ihs
      %a.global:
      TEIform CDATA 'ihs' >
<!--Finally, embed the TEI header and core tag sets.-->
<!ENTITY % TEI.header.dtd PUBLIC '-//TEI P4//ELEMENTS TEI Header//EN'
'teihdr2.dtd' >%TEI.header.dtd;
<!ENTITY % TEI.core.dtd PUBLIC '-//TEI P4//ELEMENTS Core Elements//EN'
'teicore2.dtd' >%TEI.core.dtd;
```

<!-- end of 24.8-->

The overall structure of a set of independent headers, encoded in XML for interchange as a group, is thus:

```
<teiHeader>
<fileDesc> <!-- ... --> </fileDesc>
<encodingDesc> <!-- ... --> </encodingDesc>
<profileDesc> <!-- ... --> </profileDesc>
<revisionDesc> <!-- ... --> </revisionDesc>
</teiHeader>
<teiHeader> <!-- ... --> </teiHeader>
<!-- ... etc. -->
</ihs>
```

In practice, headers might be stored in separate operating system files, to reduce redundant storage requirements; in this case, the top-level file for a typical XML document might have the following structure:

while that for a set of independent headers might have this structure:

```
<!DOCTYPE ihs PUBLIC
              "-//TEI P4//DTD Auxiliary Document Type: Independent TEI Header//EN"
              "teishd2.dtd" [
                          "INCLUDE" >
  <!ENTITY % TEI.XML
  <!ENTITY hdr01 SYSTEM 'text01.hdr' >
  <!ENTITY hdr02 SYSTEM 'text02.hdr' >
  <!ENTITY hdr03 SYSTEM 'text03.hdr' >
 <!-- ... etc. -->
]>
<ihs>
&hdr01;
&hdr02;
&hdr03:
<!-- etc. -->
</ihs>
```

24 The Independent Header