

A TEI Project

Part 3: TEI and localization

Table of contents

- [1. TEI internationalisation project](#)
- [2. Reminder: definitions](#)
- [3. Examples of translation](#)
- [4. Localisation of examples](#)
- [5. More examples](#)
- [6. More examples](#)
- [7. Translation choices](#)
- [8. TEI I18N and L10N process 1: Unicode](#)
- [9. Defining a non-Unicode character](#)
- [10. Defining a non-Unicode character \(2\)](#)
- [11. TEI I18N and L10N process 2: TEI literate programming](#)
- [12. Use of ODD](#)
- [13. ODD for translation](#)
- [14. ODD example](#)
- [15. Translating element names](#)
- [16. How does that work in the schema?](#)
- [17. Translated schema](#)
- [18. Translating descriptions](#)
- [19. Using a translated schema in practice](#)
- [20. Internationalised interfaces for TEI applications](#)
- [21. TEI schema-making tools](#)
- [22. The application of the W3C ITS guidelines to TEI work](#)
- [23. ITS information location](#)
- [24. ITS data categories](#)
- [25. TEI text with ITS rules and markup attached](#)
- [26. ITS for ODD](#)
- [27. Example of ITS implementation](#)
- [28. Scale of TEI work](#)
- [29. TEI plans](#)
- [30. TEI chapter source in English and French](#)
- [31. Example of translated ODD](#)
- [32. Using oXygen editor with Chinese annotation](#)
- [33. Example of reference documentation](#)
- [34. Example of reference documentation in Japanese](#)
- [35. Interface translation in Japanese](#)
- [36. Reference documentation in Japanese, with German annotation](#)

1. TEI internationalisation project

The aim is to lower the barrier for entry of non-English-speaking users by:

- Ensuring that all of the TEI is Unicode-safe
- Translating the reference element and attribute descriptions
- Providing tools to easily access non-English versions
- Localizing TEI software
- Localizing the TEI examples
- Translating the running prose of the Guidelines
- Translating the object names

2. Reminder: definitions

Internationalization (I18N)

Internationalization is the process of generalizing a product so that it can handle multiple languages and cultural conventions without the need for redesign. Internationalization takes place at the level of program design and document development.

Localization (L10N)

Localization is the process of taking a product and making it linguistically and culturally appropriate to a given target locale (country/region and language) where it will be used.

3. Examples of translation

- instead of <addrLine>, the TEI user might prefer to write <líneaDirección>, <ligneAdresse>, <linDireccio> or <AdressZeile>.
- instead of 'supplies the descriptive and declarative information making up an electronic title page prefixed to every TEI-conformant text.', the Chinese-speaking user might find it more helpful to read '標明標頭所附屬於的文件類型，例如文集或單一文本。'

4. Localisation of examples

What does this

```
<lg>
<l>Sire Thopas was a doghty swayn;</l>
<l>White was his face as payndemayn,</l>
<l>His lippes rede as rose;</l>
<l>His rode is lyk scarlet in grayn,</l>
<l>And I yow telle in good certayn,</l>
<l>He hadde a semely nose.</l>
</lg>
```

mean to a Chinese scholar?

5. More examples

Element names are often easy to understand; but what if we use English in attribute values?

```
Next morning a boy in that dormitory confided to his
bosom friend, a<distinct type="psSlang">fag</distinct> of
Macrea's, that there was trouble in their midst which King<distinct type="archaic">would
fain</distinct>keep secret.
```

Here there is the English word 'psSlang' (expandable to 'public school slang') for the @type attribute of <distinct> to consider, where the value of 'fag' gives little help.

6. More examples

The names of the elements may stand in the way of easy comprehension:

```
<persName key="EGBR1">
  <roleName type="office">Governor</roleName>
  <forename sort="2">Edmund</forename>
  <forename full="init" sort="3">G.</forename>
  <addName type="nick">Jerry</addName>
  <addName type="epithet">Moonbeam</addName>
  <surname sort="1">Brown</surname>
  <genName full="abb">Jr</genName>
</persName>
```

This can only really be take advantage of by someone who

1. appreciates the cultural context of 'forename' and 'surname'
2. can mentally expand 'nick' to 'nickname' (and knows what a nickname is)
3. can appreciate whether a 'Governor Edmund G. Jerry Moonbeam Brown Jr.' is a politician, a kind of food, or a new dance

7. Translation choices

The user of the Guidelines may prefer to:

1. read 'contiene un único documento TEI, compuesto de una cabecera TEI (TEI header) y un cuerpo de texto (text), aislado o como parte de un elemento corpusTei (teiCorpus)' instead of 'contains a single TEI-conformant document, comprising a TEI header and a text, either in isolation or as part of a teiCorpus element.' in the documentation
2. use element names of <líneaDirección>, <ligneAdresse>, <linDireccio> or <AdressZeile> instead of <addrLine>
3. see culturally-adapted examples, such as

```
<div>
  <head>同前</head>
  <byline rend="ur">唐·元稹</byline>
  <p>當來日，大難行。前有●，後有坑。大梁側，小梁傾。兩軸相絞，兩輪相撐。大牛豎，小牛<lb/>
    橫。烏啄牛背，足跌力蹙。當來日，大難行。太行雖險，險可使平。輪軸自攬，牽制不停。<lb/>
    泥濘漸久，荊棘旋生。行必不得，不如不行。</p>
```

```
</div>
```

8. TEI I18N and L10N process 1: Unicode

- Unicode is the only supported character encoding schema. This means that entities for characters are deprecated, and the recommended daily use is for UTF-8 encoded text, as in

```
<persName xml:lang="el-grc">Ἐλ. Θεόλαος</persName>
```

- There is a clean mechanism to use non-Unicode characters
- all appropriate text content models are set to allow a mixture of CDATA and <g> (where <g> is a reference to a non-Unicode character)
- all elements have an attribute @xml:lang to record the language used
- there are no places where an attribute is used to hold pure text

9. Defining a non-Unicode character

A new character, assigned to a position in the Unicode Private Use Area (PUA), and with a standardized form as a fallback:

```
<charDesc>
  <glyph xml:id="z103">
    <glyphName>LATIN LETTER Z WITH TWO STROKES</glyphName>
    <mapping type="standardized">z</mapping>
    <mapping type="PUA">U+E304</mapping>
  </glyph>
</charDesc>
```

This can now be referred to using the <gi> element, as in

```
<g ref="#z103"/>
```

10. Defining a non-Unicode character (2)

It is also possible to override what appears in the text by using markup like this

```
<g ref="#z103">z</g>
```

where the content of the <g> element can be used immediately without any lookup.

11. TEI I18N and L10N process 2: TEI literate programming

The TEI is written in a high-level markup language for specifying XML schemas and their documentation. This is an XML vocabulary known as ODD (*One Document Does it all*):

- The **element** and **attribute** sets making up the schema are formally specified using a special XML vocabulary
- The specification language also includes support for **macros** (like DTD entities, or schema patterns), a **hierarchical class system** for attributes and elements, and the creation of pre-defined groups of elements known as modules.
- Content models** for elements and attributes are written using an embedded RELAXNG XML notation, but tools are available to generate schemas in any of RELAXNG, DTD language, or W3C schema.
- Documentation** describing the supported elements, attributes, value lists etc is managed along with their specification, together with **use cases**, **examples**, and other supporting material.

ODD is a standard TEI module.

12. Use of ODD

The TEI's 22 modules (containing 500 elements) can be combined together and customized as desired using the ODD language. Customization may include:

- tightening the constraints on existing elements (**example: limiting values of the @type attribute to certain values**)
- removing unused elements (**example: remove <formula> from figures and tables module**)
- changing the class system (**example: allow <figure> to appear where other <div> is allowed**)
- adding new elements or attributes (**example: add an element to contain an sound recording**)

The last two may make documents which break compatibility.

13. ODD for translation

The ODD language has allowance for translating element name, attribute names, and descriptions, and for preserving information to allow canonicalisation.

The technical documentation elements (<gloss> and <desc>) for TEI elements and attributes etc can be specified multiple times, in different languages, distinguished by the standard @xml:lang attribute.

There is also a container (<equiv>) to specify the relationship of an element, attribute or value to standardised schemes or ontologies.

14. ODD example

```
<elementSpec module="header" ident="taxonomy">
  <desc>defines a typology used to classify texts either
    implicitly, by means of a bibliographic citation, or explicitly by a structured taxonomy.</desc>
  <content>
    <rng:choice>
      <rng:oneOrMore>
        <rng:ref name="category"/>
      </rng:oneOrMore>
      <rng:group>
        <rng:group>
          <rng:ref name="model.biblLike"/>
        </rng:group>
        <rng:zeroOrMore>
          <rng:ref name="category"/>
        </rng:zeroOrMore>
      </rng:group>
    </rng:choice>
  </content>
  <exemplum>
    <egXML><taxonomy xml:id="tax.b">
      <bibl>Brown Corpus</bibl>
      <category xml:id="tax.b.a">
        <catDesc>Press Reportage</catDesc>
        <category xml:id="tax.b.a1">
          <catDesc>Daily</catDesc>
        </category>
        <category xml:id="tax.b.a2">
          <catDesc>Sunday</catDesc>
        </category>
        <category xml:id="tax.b.a3">
          <catDesc>National</catDesc>
        </category>
        <category xml:id="tax.b.a4">
          <catDesc>Provincial</catDesc>
        </category>
        <category xml:id="tax.b.a5">
          <catDesc>Political</catDesc>
        </category>
        <category xml:id="tax.b.a6">
          <catDesc>Sports</catDesc>
        </category>
      </category>
      <category xml:id="tax.b.d">
        <catDesc>Religion</catDesc>
      </category>
      <category xml:id="tax.b.d1">
        <catDesc>Books</catDesc>
      </category>
      <category xml:id="tax.b.d2">
        <catDesc>Periodicals and tracts</catDesc>
      </category>
    </taxonomy>
  </egXML>
</exemplum>
</elementSpec>
```

15. Translating element names

The objects identified by the @ident attribute in the TEI can be given an alternate name by use of the <altIdent> element; so the example above could be rewritten as

```
<elementSpec xmlns="http://www.tei-c.org/ns/1.0" module="header" ident="taxonomy">
  <altIdent xmlns="http://www.tei-c.org/ns/1.0" xml:lang="fr">taxinomie</altIdent>
  ...</elementSpec>
```

providing a French name for the element.

16. How does that work in the schema?

The normal schema, using RELAXNG compact syntax, has the definition

```
taxonomy =
  ## (taxonomy) defines a typology used to classify texts either
  ## implicitly, by means of a bibliographic citation,
  ## or explicitly by a structured taxonomy.
  element taxonomy { taxonomy.content, taxonomy.attributes }
taxonomy.content = category+ | (model.biblLike, category*)
taxonomy.attributes = att.global.attributes, empty
```

in which the **element** <taxonomy> is defined by the containing pattern **taxonomy**; it is the **pattern name** which other elements use, not the element name.

17. Translated schema

If the schema were translated into Greek, it would look like this:

```
taxonomy =
  element ταξινομία { taxonomy.content, taxonomy.attributes }
...
```

where the 'pattern name' remains the same. This type of schema markup is generated by the TEI tools, picking up the information from <altIdent>.

18. Translating descriptions

We can expand the TEI source to add Chinese translations alongside the English originals, and the appropriate text can be passed to the generated schemas or documentation:

```
<elementSpec xmlns="http://www.tei-c.org/ns/1.0" module="header" ident="taxonomy">
<desc xmlns="http://www.tei-c.org/ns/1.0">defines a typology used to classify texts either
implicitly, by means of a bibliographic citation, or explicitly by a structured taxonomy.</desc>
<desc xmlns="http://www.tei-c.org/ns/1.0" version="2007-05-02" xml:lang="zh-tw">定義文件分類的類型學，可以是隱在地以書目資料的方式，或是明確地以結構分類法
的方式來分類。</desc>
...</elementSpec>
```

19. Using a translated schema in practice

If we take a Spanish play, and translate the element names to Spanish, a text like this will be much more familiar-looking to encoders in Spanish-speaking countries:

```
<cuervo>
<div1 tipo="part">
<div2 tipo="act">
<encabezado tipo="main">Jornada primera</encabezado>
<div3 tipo="scene">
<encabezado tipo="main">Cuadro único</encabezado>
<acotacion formato="centered">
<resaltado formato="bold">(Salen </resaltado>REBOLLEDO,
<resaltado formato="bold">la</resaltado>CHISPA<resaltado formato="bold">
soldados</resaltado>.<resaltado formato="bold">)</resaltado>
</acotacion>
<dialogo>...</dialogo>
</div3>
</div2>
</div1>
</cuervo>
```

it is straightforward to write a transformation (eg in XSL) which reads the TEI source with the element names and <altIdent> information, and puts the text back to canonical form.

20. Internationalised interfaces for TEI applications

An application which turns TEI XML into HTML for web display, and provides a heading such as 'Contents' when it meets <divGen type="toc"/>, will have to provide appropriate translations. Eg:

ISO Language code Text

bg	Съдържание
de	Inhalt

el	Περιεχόμενα
en	Contents
es	Contenidos
fr	Contenu
hi	Mula Shabda
ja	目次
nl	Inhoud
pl	Spis treści
pt	Índice geral
ro	Cuprins
ru	Оглавление
slv	Vsebina
sr	Sadržaj
sv	Innehåll
th	เนื้อหา
tr	İçerik
zh-tw	內容

21. TEI schema-making tools

The ODD language files are processed to produce schemas in the chosen language, using the **Roma** tools (web service and command-line script). This has support for varying the languages of its interface, but must also allow for supporting the following output schemes:

- **canonical**: English names, descriptions in English
- **local descriptions**: English names, descriptions in chosen language
- **local names**: names designed to make sense to a speaker of the chosen language, descriptions in English
- **fully localized**: both names and descriptions in chosen language

22. The application of the W3C ITS guidelines to TEI work

The *W3C Internationalisation Tag Set* encodes information for translators and localisers.

The ITS consists of a set of elements and attributes for annotating a text with information for further processing, covering **Internationalization**:

- Markup for bidirectional text
- Ruby annotation
- Language identification

and **Localization**

- Translatability of content
- The localization process in general
- Terminology markup

23. ITS information location

The primary ITS notion is that information about elements and attributes can be supplied

- in a document schema
- in an external rules file
- in a rule section in an instance file
- attached to instance elements

where the information consists of a set of *data categories*.

24. ITS data categories

On an instance element the following attributes may be attached or derived:

translate

should this object be translated?

locInfo

Is there some localisation hint?

locInfoType

What type of hint is it?

term

Does this object describe a technical term?

termRef

Where is the term defined?

dir

What is the text direction?

rubyText

Is there some Ruby annotation?

25. TEI text with ITS rules and markup attached

```
<TEI>
  <teiHeader>
    <its:rules>
      <its:ns
        its:prefix="t"
        its:uri="http://www.tei-c.org/ns/1.0"/>
      <its:translateRule translate="no" selector="//t:body/t:p"/></its:rules>
    </teiHeader>
  <text>
    <body>
      <p>Hello <hi>world</hi>
      </p>
      <p
        its:translate="yes">translate me</p>
    </body>
  </text>
</TEI>
```

where the ITS rules say that <p> elements should not normally be translated, but the second <p> has an explicit override.

26. ITS for ODD

We can express the relationship between the structural elements and the documentation elements in ODD with the following ITS rules, which says that the default is to **not** translate anything, but gives a set of elements which **are** to be translated:

```
<its:rules>
  <its:ns prefix="tei" uri="http://www.tei-c.org/ns/1.0"/>
  <its:translateRule translate="no" selector="//tei:*/>
  <its:translateRule translate="yes" selector="//tei:desc"/>
  <its:translateRule translate="yes" selector="//tei:gloss"/>
  <its:translateRule translate="yes" selector="//tei:valDesc"/>
  <its:translateRule translate="yes" selector="//tei:p[@rend='dataDesc']"/>
  <its:translateRule translate="yes" selector="//tei:remarks"/></its:rules>
```

27. Example of ITS implementation

We can show graphically, using an ITS tool, which elements need a translated equivalent (those in green)

```

<elementSpec module="corpus" id="PERSON" usage="opt" ident="person">
  <equiv/>
  <gloss/>
  <desc>describes a single participant in a language interaction. </desc>
  <content>
    <rng:choice>
      <rng:oneOrMore>
        <rng:ref name="model.pLike"/>
      </rng:oneOrMore>
      <rng:zeroOrMore>
        <rng:ref name="model.personPart"/>
      </rng:zeroOrMore>
    </rng:choice>
  </content>
  <attList>
    <attDef ident="role" usage="opt">
      <equiv/>
      <desc>specifies the role of this participant in the group.</desc>
      <datatype>
        <rng:ref name="data.code"/>
      </datatype>
      <valDesc>a set of keywords to be defined</valDesc>
    </attDef>
    <attDef ident="sex" usage="opt">
      <equiv/>
      <desc>specifies the sex of the participant.</desc>
      <datatype>

```

Done

28. Scale of TEI work

The scale of work involved is not impossible to contemplate. The TEI contains

- 494 elements
- 116 classes
- 476 attributes
- 1115 <gloss> elements, 29357 characters
- 1170 <desc> elements, 98415 characters

The work needed for each language is to

- translate descriptive prose to other languages
- translate technical documentation components (note that this includes gloss for fixed attribute lists)
- translate examples
- localize examples

At a system level, we need to

- add W3C ITS information if needed
- create the translation-processing workflow tools

29. TEI plans

The TEI Consortium is working with TEI scholars to advance I18N and L10N in various languages.

French, Spanish, Italian, and Japanese are largely complete for translated <desc> and <gloss> texts; Chinese and German are in progress.

We hope to process Portuguese, Greek, Czech and Korean automatically using SYSTRAN.

Translation of all interface strings in XSL stylesheets to Bulgarian, Chinese, Dutch, French, German, Greek, Hindi, Italian, Japanese, Polish, Portuguese, Romanian, Russian, Serbian, Slovenian, Spanish, Swedish, Thai, and Turkish is complete.

30. TEI chapter source in English and French

```

/TEI/Sourceforge/trunk/P5/Source/Guidelines/en/HD.xml
File Edit Options Buffers Tools XML TEI UniChar Help
This is followed by a detailed description of all the constituent
elements which may be used in the core header. Section <ptr target=
st the end of the present chapter, discusses the
recommended content of a minimal TEI header, and its relation to
standard library cataloguing practices. Recommendations relevant to
use of TEI headers as free-standing documents, for interchange among
libraries, data archives, and similar institutions may be found in
chapter <ptr target="#SH"/>.</p>
<div2 type="bar" xml:id="HD1"><head>Organization of the TEI Header</h
<div3 xml:id="HD11"><head>The TEI Header and Its Components</head>
<p>The <gi>teiHeader</gi> element should be clearly distinguished bot
from the <term>prolog</term>, which comprises either the XML declarat
s SGML declaration,
and the document type declaration (see chapter <ptr target="#SG"/>);
from the <term>front matter</term> of the text itself (for which see
section <ptr target="#DSFRONT"/>). A composite text, such as a corpus
collection, may contain several headers, as further discussed below.
- &-:-- HD.xml 2:18 10.40 Mail (nXML Invalid Fill)--L??--C4--
institutions similaires se trouvent au chapitre 24 Éléments de
l'En-tête et leurs relations avec les autres formats de métadonnées
<ptr target="#SH"/>.</p>
<div2 type="bar" xml:id="HD1">
<head>Organisation de l'En-tête TEI</head>
<!-- 5.1 -->
<div3 xml:id="HD11">
<head>En-tête TEI et ses composants</head>
<!-- 5.1.1 -->
<p>L'élément En-tête TEI <gi>teiHeader</gi> doit être clairement
distingué du prologue <term>prolog</term> qui comprend la déclaration
XML ou SGML et la déclaration de type de document (voir chapitre 2
Introduction à XML <ptr target="#SG"/>) ainsi que des parties
liminaires du texte lui-même <term>front matter</term> (voir Section
7.4 Parties liminaires <ptr target="#DSFRONT"/>). Un texte composite,
tel qu'un corpus ou une collection de textes, peut contenir plusieurs
en-têtes (voir plus loin). Toutefois, dans le cas le plus courant, un
- &u:** HD.xml<2> 2:18 10.40 Mail (nXML Valid Fill)--L??--C14--
Mark saved where search started

```

31. Example of translated ODD

```

<elementSpec
  module="textstructure"
  xml:id="TEI2"
  usage="req"
  ident="TEI">
  <equiv/>
  <gloss>TEI document</gloss>
  <gloss version="2008-01-30" xml:lang="ja">TEI文書</gloss>
  <gloss version="2007-06-12" xml:lang="fr">document TEI</gloss>
  <gloss version="2006-10-18" xml:lang="de">TEI-Dokument</gloss>
  <gloss version="2007-05-04" xml:lang="es">documento TEI</gloss>
  <gloss version="2007-05-02" xml:lang="zh-tw">TEI文件</gloss>
  <gloss version="2007-01-21" xml:lang="it">documento TEI</gloss>
  <desc>contains a single TEI-conformant document,
    comprising a TEI header and a text, either in isolation or as part of a <gi>teiCorpus</gi> element.</desc>
  <desc version="2008-01-30" xml:lang="ja"> TEI準拠の文書を示す.</desc>

```

```

<desc version="2007-06-12" xml:lang="fr">contient un seul document, conforme à la TEI, qui
  comprend un en-tête TEI et un texte, soit de façon isolée soit comme une partie d'un élément<gi>teiCorpus</gi>
</desc>
<desc version="2006-10-18" xml:lang="de">
  enthält ein einzelnes TEI-konformes Dokument, das aus TEI-Header (Dateikopf) und Text besteht, entweder als eigenständige Datei oder als
  Teil eines Elements<gi>teiCorpus</gi>.
</desc>
<desc version="2007-05-04" xml:lang="es">contiene un único documento TEI-conforme, que comprende un encabezado y un texto, sea este aislado o
  parte de un elemento <gi>teiCorpus</gi>
</desc>
<desc version="2007-05-02" xml:lang="zh-tw">符合TEI標準的單一文件，包括一個TEI標頭以及一個文本，可單獨出現或是處於元素<gi>teiCorpus</gi>(tei文集)之中。
</desc>
<desc version="2007-01-21" xml:lang="it">contiene un documento TEI-conforme, comprendente un'intestazione e un testo, sia esso isolato o
  parte di un elemento <gi>teiCorpus</gi>
</desc>
<content>
  <rng:group>
    <rng:ref name="teiHeader" />
    <rng:choice>
      <rng:group>
        <rng:oneOrMore>
          <rng:ref name="model.resourceLike" />
        </rng:oneOrMore>
        <rng:optional>
          <rng:ref name="text" />
        </rng:optional>
      </rng:group>
      <rng:ref name="text" />
    </rng:choice>
  </rng:group>
  <sch:ns prefix="tei" uri="http://www.tei-c.org/ns/1.0" />
  <sch:ns prefix="rng" uri="http://relaxng.org/ns/structure/1.0" />
</content>
<attList>
<attDef ident="version" usage="opt">
  <equiv/>
  <desc>The version of the TEI scheme</desc>
  <desc version="2008-01-30" xml:lang="ja">TEIスキームの版を示す。 </desc>
  <desc version="2007-06-12" xml:lang="fr">la version du schéma TEI</desc>
  <desc version="2006-10-18" xml:lang="de"> Version des TEI-Schemas</desc>
  <desc version="2007-05-04" xml:lang="es">Versión del esquema TEI</desc>
  <desc version="2007-05-02" xml:lang="zh-tw">TEI架構的版本</desc>
  <desc version="2007-01-21" xml:lang="it">versione dello schema TEI</desc>
  <datatype>
    <rng:data type="decimal" />
  </datatype>
  <defaultVal>5.0</defaultVal>
  <valDesc>A number identifying the version of the TEI guidelines</valDesc>
</attDef>
</attList>
<exemplum>
  <egXML><TEI>
    <teiHeader>
      <fileDesc>
        <titleStmt>
          <title>The shortest TEI Document Imaginable</title>
        </titleStmt>
        <publicationStmt>
          <p>First published as part of TEI P2.</p>
        </publicationStmt>
        <sourceDesc>
          <p>No source: this is an original work.</p>
        </sourceDesc>
      </fileDesc>
    </teiHeader>
    <text>
      <body>
        <p>This is about the shortest TEI document imaginable.</p>
      </body>
    </text>
  </TEI>
</egXML>
</exemplum>
<remarks>
  <p>This element is required.</p>
</remarks>
<listRef>
  <ptr target="#DS" />
  <ptr target="#CCDEF" />
</listRef>

```

</elementSpec>

32. Using oXygen editor with Chinese annotation

The screenshot shows the oXygen XML editor interface. The menu bar includes File, Edit, Find, Project, Perspective, Options, Tools, Debugger, Document, Window, and Help. The toolbar contains various icons for file operations, editing, and navigation. The main window displays an XML document named 'taxonomy.xml' with the following content:

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <?oxygen RNGSchema="file:/TEI/Sourceforge/trunk/P5/test-zh.rnc" type="compact"?>
3 <TEI xmlns="http://www.tei-c.org/ns/1.0">
4 <teiHeader>
5   <fileDesc>
6     <titleStmt></titleStmt>
7   </fileDesc>
8 </teiHeader>
9 </TEI>

```

A tooltip is visible over the <fileDesc> element, containing the following Chinese text:

(背景描述) 提供文件非書目性方面的
 細節描述, 更明確地, 是所使用的語言
 及次要語言、在何種情況下製作、參與
 人員及其環境背景。

33. Example of reference documentation

```

<person> - B

```

1. <elementSpec> person

describes a single participant in a language interaction.

Declaration

```

element person
{
  <att.global.attributes,
  attribute role { text }?,
  attribute sex { "m" | "f" | "u" }?,
  attribute age { text }?,
  ( p+ | <a href="http://www.tei.ac.uk/tei/demographic">tei.demographic* )
}

```

Attributes: (In addition to global attributes)

role

specifies the role of this participant in the group.

sex

specifies the sex of the participant. Legal values are:

m

male

f

female

u

unknown or inapplicable

age

specifies the age group to which the participant belongs.

Example

```

<person sex="f" age="42">
  <p>Female informant, well-educated, born in Shropshire
  UK, 12 Jan 1950, of unknown occupation.
  Speaks French fluently. Socio-Economic status B2.</p>
</person>

```

May contain a prose description organized as paragraphs, or any sequence of demographic e

34. Example of reference documentation in Japanese

1. <elementSpec> person

宣言

```

element person
{
  att.global.attributes,
  attribute role { text }?,
  attribute sex { "m" | "f" | "u" }?,
  attribute age { text }?,
  ( p+ | tei.demographic\* )
}

```

属性: (グローバル属性の他)

role

sex

正当な値:

m

男性

f

女性

u

不明または不適切

age

例

```

<person sex="f" age="42">
  <p>Female informant, well-educated, born in Shropshire
    UK, 12 Jan 1950, of unknown occupation.
    Speaks French fluently. Socio-Economic status B2.</p>
</person>
<person sex="f" age="42" lang="ja">
  <p>女性、教養あり、1950年1月12日英国シェロブシア生まれ、
    職業不明、フランス語を流暢に話す、社会経済状態：中
  </p>
</person>

```

段落単位の記述、または人口統計学のデータが混在して含まれる

35. Interface translation in Japanese

A sample page from Punch

Skip links

▶ Home ▶

My Parent Inst

2. TALES TOLD TO
CIVILIANS. The
FLY.

3. SPELLING BY
"ANALOGY"

4. -----

4. -----

"Yesterday evening Mr. — J.P., auctioneer and farmer, was fired at when drivin home — Mr. — , who is an ex-Chairman of the— Rural District Council, is a mo popular man in the district. For some time past he has been subject to a series annoyances, the most recent of which was the spiking of his lands, and his catt and horses mutilated." — Irish Paper.

Popularity in Ireland would appear to have its drawbacks.

上: [目次](#) 前: [3. SPELLING BY "ANAL](#)

My Parent Institution | TEI | [検索](#) | [ご意見・ご感想](#)

日付:

This page is copyrighted

36. Reference documentation in Japanese, with German annotation

1. <elementSpec> person

Deklaration

```

element person
{
  att.global.attributes,
  attribute role { text }?,
  attribute sex { "m" | "f" | "u" }?,
  attribute age { text }?,
  ( p+ | tei.demographic\* )
}

```

Attribute: (Neben global gültigen Attributen)

role

sex

Gültige Werte:

m

男性

f

女性

u

不明または不適切

age

Beispiel

```

<person sex="f" age="42">
  <p>Female informant, well-educated, born in Shropshire
    UK, 12 Jan 1950, of unknown occupation.
    Speaks French fluently. Socio-Economic status B2.</p>
</person>
<person sex="f" age="42" lang="ja">
  <p>女性、教養あり、1950年1月12日英国シュロブシア生まれ、
    職業不明、フランス語を流暢に話す、社会経済状態：中
  </p>

```

[Parent Institution](#) | [TEI](#) | [Search](#) | [Feedback](#)

Date: 2008-03-05
 This page is copyrighted