

Exploring the TEI Community

James Cummings

April 2009

The TEI Guidelines

This talk gives a whirlwind survey of some of the other things that the TEI provides for its community, starting with the TEI Guidelines.

TEI Chapters (1)

In addition to Front Matter and Back Matter, the TEI Guidelines contain chapters on:

- 1. The TEI Infrastructure
- 2. The TEI Header
- 3. Elements Available in All TEI Documents
- 4. Default Text Structure
- 5. Representation of Non-standard Characters and Glyphs
- 6. Verse
- 7. Performance Texts
- 8. Transcriptions of Speech
- 9. Dictionaries
- 10. Manuscript Description
- 11. Representation of Primary Sources
- 12. Critical Apparatus

...

TEI Chapters (2)

...

- 13. Names, Dates, People, and Places
- 14. Tables, Formulæ, and Graphics
- 15. Language Corpora
- 16. Linking, Segmentation, and Alignment
- 17. Simple Analytic Mechanisms
- 18. Feature Structures
- 19. Graphs, Networks, and Trees
- 20. Non-hierarchical Structures
- 21. Certainty and Responsibility
- 22. Documentation Elements
- 23. Using the TEI

v. A Gentle Introduction to XML

- Moved to Front Matter of Guidelines
- Updated to modern use of XML, namespaces, schemas, etc.
- A must-read for those new to XML
- Considers the nature of markup, why XML is so special, and the structures we create with them

6. Verse

- `<rhyme>` element added to make rhyming part of metrical line
- Core: `<lg>` (line-groups) enclose `<l>` (line) elements
- Module Elements: `<caesura>`, `<metDecl>`, `<metSym>`, `<rhyme>`

```

<lg rhyme="abab">
  <l>A pleasing land of drowsy head it <rhyme label="a">was</rhyme>,</l>
  <l> Of dreams that wave before the half-shut <rhyme label="b">eye</rhyme>;</l>
  <l>And of gay castles in the clouds that <rhyme label="a">pass</rhyme>,</l>
  <l> For ever flushing round a summer <rhyme label="b">sky</rhyme>.</l>
  <trailer>CASTLE OF INDOLENCE.</trailer>
</lg>

```

7. Performance Texts

- Core: `<sp>`, `<speech>`, `<speaker>`
- Module Elements: `<actor>`, `<camera>`, `<caption>`,
`<castGroup>`, `<castItem>`, `<castList>`, `<epilogue>`,
`<move>`, `<performance>`, `<prologue>`, `<role>`,
`<roleDesc>`, `<set>`, `<sound>`, `<tech>`, `<view>`

```
<castList>
  <castItem>
    <role>Ichabod Crane</role>
    <actor>James Cummings</actor>
  </castItem>
  <castItem>
    <role>Katrina Van Tassel</role>
    <actor>Dot Porter</actor>
  </castItem>
</castList>
```

8. Transcriptions of Speech

- Module Elements: `<broadcast>`, `<equipment>`, `<incident>`, `<kinesic>`, `<pause>`, `<recording>`, `<recordingStmt>`, `<scriptStmt>`, `<shift>`, `<u>`, `<vocal>`, `<writing>`
- `<incident>` replaces `<event>` which has been re-purposed

```

<u>What passed at this interview I will not pretend to say, for in fact I do not
know.<pause dur="PT5S"/> Something, however, I fear me, must have gone wrong.<gap/>
<vocal dur="PT2S">
  <desc>cries in anguish</desc>
</vocal>Oh these women! these women! <kinesic dur="PT1.5S">
  <desc>shakes head sadly</desc>
</kinesic> Could that girl have been playing off any of her coquettish tricks?</u>

```


9. Dictionaries

- Module Elements: `<case>`, `<colloc>`, `<def>`, `<dictScrap>`, `<entry>`, `<entryFree>`, `<etym>`, `<form>`, `<gen>`, `<gram>`, `<gramGrp>`, `<hom>`, `<hyph>`, `<iType>`, `<lang>`, `<lbl>`, `<mood>`, `<number>`, `<oRef>`, `<oVar>`, `<orth>`, `<pRef>`, `<pVar>`, `<per>`, `<pos>`, `<pron>`, `<re>`, `<sense>`, `<stress>`, `<subc>`, `<superEntry>`, `<syll>`, `<tns>`, `<usg>`, `<xr>`

```

<entry>
  <form>
    <orth>horseman</orth>
  </form>
  <sense n="1">
    <def>a rider or driver of horses; especially: one whose skill is exceptional</def>
  </sense>
</entry>

```

10. Manuscript Description

- Module Elements: <accMat>, <acquisition>, <additional>, <additions>, <adminInfo>, <altIdentifier>, <binding>, <bindingDesc>, <catchwords>, <collation>, <collection>, <colophon>, <condition>, <custEvent>, <custodialHist>, <decoDesc>, <decoNote>, <depth>, <dimensions>, <explicit>, <filiation>, <finalRubric>, <foliation>, <handDesc>, <height>, <heraldry>, <history>, <incipit>, <institution>, <layout>, <layoutDesc>, <locus>, <locusGrp>, <material>, <msContents>, <msDesc>, <msIdentifier>, <msItem>, <msItemStruct>, <msName>, <msPart>, <musicNotation>, <objectDesc>, <origDate>, <origPlace>, <origin>, <physDesc>, <provenance>, <recordHist>, <repository>, <rubric>, <seal>, <sealDesc>, <secFol>, <signatures>, <source>, <stamp>, <summary>, <support>, <supportDesc>, <surrogates>, <textLang>, <typeDesc>, <watermark>, <width>

10. Manuscript Description

```

<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno>MS. Add. A. 61</idno>
    <altIdentifier type="SC">
      <idno>28843</idno>
    </altIdentifier>
  </msIdentifier>
  <p>In Latin, on parchment: written in more than one hand of the 13th cent. in England: 7¼ x 5¾ in., i + 55 leaves, in double columns: with a few coloured capitals.</p>
  <p>'Hic incipit Bruitus Anglie,' the De origine et gestis Regum Angliae of Geoffrey of Monmouth (Galfridus Monumetensis: beg. 'Cum mecum multa & de multis.'</p>
  <p>On fol. 54v very faint is 'Iste liber est fratris guillelmi de buria de ... Roberti ordinis fratrum Pred[icatorum],' 14th cent. (?): 'hanauilla' is written at the foot of the page (15th cent.). Bought from the rev. W. D. Macray on March 17, 1863, for £1 10s.</p>
</msDesc>

```

11. Representation of Primary Sources

- Elements defined: `<addSpan>`, `<am>`, `<damage>`, `<damageSpan>`, `<delSpan>`, `<ex>`, `<facsimile>`, `<fw>`, `<handNotes>`, `<handShift>`, `<restore>`, `<space>`, `<subst>`, `<supplied>`, `<surface>`, `<zone>`

```
<facsimile>
  <graphic url="page1.png"/>
  <graphic url="page2.png"/>
  <graphic url="page3.png"/>
  <graphic url="page4.png"/>
</facsimile>
```

12. Critical Apparatus

- Module Elements: `<app>`, `<lacunaEnd>`, `<lacunaStart>`, `<lem>`, `<listWit>`, `<rdg>`, `<rdgGrp>`, `<variantEncoding>`, `<wit>`, `<witDetail>`, `<witEnd>`, `<witStart>`, `<witness>`

```
<p>Certain it is, this was not the case with the redoubtable Brom Bones; and <app>  
<rdg wit="#msA">from the moment Ichabod Crane made his advances,</rdg>  
<rdg wit="#msB">coincidentally when Ichabod Crane made his advances,</rdg>  
<rdg wit="#msC">from the moment Ichabod Crane started to sing,</rdg>  
</app> the interests of the former evidently declined;</p>
```

13. Names, Dates, People, Places

- Elements defined: <addName>, <affiliation>, <age>, <birth>, <bloc>, <climate>, <country>, <death>, <district>, <education>, <event>, <faith>, <floruit>, <forename>, <genName>, <geo>, <geogFeat>, <geogName>, <langKnowledge>, <langKnown>, <listEvent>, <listNym>, <listOrg>, <listPerson>, <listPlace>, <location>, <nameLink>, <nationality>, <nym>, <occupation>, <offset>, <org>, <orgName>, <persName>, <person>, <personGrp>, <place>, <placeName>, <population>, <region>, <relation>, <relationGrp>, <residence>, <roleName>, <settlement>, <sex>, <socecStatus>, <state>, <surname>, <terrain>, <trait>

13. Names, Dates, People, Places

```
<p>... <name ref="#jsbach" type="person"> Johann Sebastian Bach </name> was a prolific  
German composer ... </p>
```

```
<listPerson type="composers">  
  <person xml:id="kfabel">  
    <persName>Karl Friedrich Abel</persName>  
  </person>  
  <person xml:id="jsbach">  
    <persName>Johann Sebastian Bach</persName>  
  </person>  
  <person xml:id="cpebach">  
    <persName>Carl Philipp Emanuel Bach</persName>  
  </person>  
<!--...-->  
</listPerson>
```

14. Tables, Formulae, and Graphics

- Module Elements: `<cell>`, `<figDesc>`, `<figure>`,
`<formula>`, `<row>`, `<table>`

```
<table>
  <row>
    <cell>Van Tassel, Katrina</cell>
    <cell>A+</cell>
    <cell>As beautiful as she is talented</cell>
  </row>
<!-- more table rows of Ichabod's grading book-->
</table>
```


15. Language Corpora

- Module Elements: `<activity>`, `<channel>`, `<constitution>`, `<derivation>`, `<domain>`, `<factuality>`, `<interaction>`, `<locale>`, `<particDesc>`, `<preparedness>`, `<purpose>`, `<setting>`, `<settingDesc>`, `<textDesc>`

```

<settingDesc>
  <setting who="#p1 #p2">
    <name type="village">Sleep Hollow</name>
    <date>early spring, 1789</date>
    <locale>a farm house, sat by the hearth</locale>
    <activity>courting</activity>
  </setting>
  <setting who="#p3">
    <name type="city">Sleepy Hollow</name>
    <date>early spring, 1789</date>
    <locale>school house</locale>
    <activity>teaching</activity>
  </setting>
</settingDesc>

```

16. Linking, Segmentation, and Alignment

- Module Elements: `<ab>`, `<alt>`, `<altGrp>`, `<anchor>`, `<join>`, `<joinGrp>`, `<link>`, `<linkGrp>`, `<seg>`, `<timeline>`, `<when>`

```
<p xml:id="p123">These magic books and the poetic scrawls were forthwith consigned to the flames by <persName xml:id="HVR5">Hans Van Ripper</persName>; who from that time forward determined to send his children no more to school; observing, that he never knew any good come of this same reading and writing. </p> [...]
```

```
<linkGrp>
  <link targets="#p123 facs.xml#graphic-p123"/>
  <link targets="#HVR5 people.xml#VanRipperH5"/>
</linkGrp>
```

17. Simple Analytic Mechanisms

- Module Elements: `<c>`, `<cl>`, `<interp>`, `<interpGrp>`, `<m>`, `<phr>`, `<s>`, ``, `<spanGrp>`, `<w>`

```

<u>
  <s>
    <w ana="#ATO">The</w>
    <w ana="#NN1">victim</w>
    <w ana="#POS">'s</w>
    <w ana="#NN2">friends</w>
    <w ana="#VVD">told</w>
    <w ana="#NN2">villagers</w>
    <w ana="#CJT">that</w>
    <w ana="#ATO">the</w>
    <w ana="NPO">Headless</w>
    <w ana="NPO">Horseman</w>
    <w ana="#VVD">rode</w>
    <w ana="#PRP">into</w>
    <w ana="#ATO">the</w>
    <w ana="#NN1">forest</w>
    <w ana="#CJC">and</w>
    <w ana="#AVO">never</w>
    <w ana="#VVD">surfaced</w>
  </s>
</u>

```

18. Feature Structures

- Elements defined: `<bicond>`, `<binary>`, `<cond>`,
`<default>`, `<f>`, `<fDecl>`, `<fDescr>`, `<fLib>`, `<fs>`,
`<fsConstraints>`, `<fsDecl>`, `<fsDescr>`, `<fsdDecl>`,
`<fsdLink>`, `<fvLib>`, `<if>`, `<iff>`, `<numeric>`, `<string>`,
`<symbol>`, `<then>`, `<vAlt>`, `<vColl>`, `<vDefault>`,
`<vLabel>`, `<vMerge>`, `<vNot>`, `<vRange>`

```
<fs type="real_estate_listing">
  <f name="selling.points">
    <vColl org="set">
      <string>alarm system</string>
      <string>good view</string>
      <vAlt>
        <string>pool</string>
        <string>jacuzzi</string>
      </vAlt>
    </vColl>
  </f>
</fs>
```

19. Graphs, Networks, and Trees

- Elements defined: `<arc>`, `<eLeaf>`, `<eTree>`, `<forest>`, `<forestGrp>`, `<graph>`, `<iNode>`, `<leaf>`, `<node>`, `<root>`, `<tree>`, `<triangle>`

```
<graph>
  <node xml:id="LAX2" degree="2" adj="#LVG2 #PHX2">
    <label>LAX2</label>
  </node>
  <node xml:id="LVG2" degree="2" adj="#LAX2 #PHX2">
    <label>LVG2</label>
  </node>
<!-- ... -->
</graph>
```

20. Non-hierarchical Structures

This chapter does not define a module, but explores methods to deal with overlapping structures while keeping XML well-formed. In particular:

- 20.1 Multiple Encodings of the Same Information
- 20.2 Boundary Marking with Empty Elements
- 20.3 Fragmentation and Reconstitution of Virtual Elements
- 20.4 Stand-off Markup
- 20.5 Non-XML-based Approaches

21. Certainty and Responsibility

- Elements defined: `<certainty>`, `<respons>` and soon `<precision>`

```
<certainty target="#CE-pl1" locus="gi" degree="0.6">  
  <desc>probably a placename, but possibly not</desc>  
</certainty>  
<certainty  
  target="#CE-pl1"  
  locus="gi"  
  degree="0.4"  
  assertedValue="persName">  
  <desc>may refer to the Earl of Essex</desc>  
</certainty>
```

22. Documentation Elements

- This chapter defines the module which contains all the elements used in ODD.

```
<schemaSpec prefix="TEI_" ident="testsvg" start="TEI svg">  
  <moduleRef key="header"/>  
  <moduleRef key="core"/>  
  <moduleRef key="tei"/>  
  <moduleRef key="textstructure"/>  
  <moduleRef url="svg11.rng"/>  
</schemaSpec>
```


23. Using the TEI

- The new 'Using the TEI' chapter does not define a module of elements
- It contains sections on:
 - 23.1 Obtaining the TEI Schemas
 - 23.2 Personalization and Customization
 - 23.3 Conformance
 - 23.4 Implementation of an ODD system

Live Demos

Now if I have time, and an internet connection, I'll do some live demos looking around the Guidelines, Reference Material, Roma, etc.

- Guidelines: <http://www.tei-c.org/Guidelines/P5/>
- Reference: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/REF-ELEMENTS.html>
- Roma: <http://www.tei-c.org/Roma/>
- Stylesheets: <http://www.tei-c.org/Tools/Stylesheets/>
- Badges: <http://www.tei-c.org/About/Badges/>
- Wiki: <http://wiki.tei-c.org/>
- Sourceforge: <http://sourceforge.net/projects/tei/>