Seminars in Scholarly Text Encoding with TEI

Overview

The WWP conducted a 2-year series of NEH-funded seminars on scholarly text encoding in 2007–2009.

- 11 seminars reaching 220 participants from 71 institutions.
- An audience of junior and senior faculty, graduate students, library and digital staff, and post-doctoral fellows.
- Audience expertise ranged from novice to XML expert.
- Seminars ranged from 1 to 3 days long.
- Resources (including schema for teaching, handouts, slides, and lecture notes) are available at http://www.wwp.brown.edu/encoding/resources.html.

Following each seminar, recipients had access to ongoing consultation and grant-writing support.

Teaching schema

```
<moduleRef key="analysis"/>
<moduleRef key="core"/>              <!-- required -->
<moduleRef key="corpus"/>
<moduleRef key="drama"/>
<moduleRef key="figures"/>
<moduleRef key="header"/>            <!-- required -->
<moduleRef key="linking"/>
<moduleRef key="namesdates"/>
<moduleRef key="tagdocs"/>           <!-- for specific elements only [1] -->
<moduleRef key="tei"/>               <!-- required -->
<moduleRef key="textcrit"/>
<moduleRef key="textstructure"/>  <!-- required, divN deleted -->
<moduleRef key="transcr"/>
<moduleRef key="verse"/>
<!-- [1] <att>, <code>, <eg>, <egXML>, <gi>, <ident>, <tag>, and <val> -->
```

Started with a 2-pronged approach:

- Students began with a very tight schema for their first foray into encoding a simple sample document using oXygen
- Shortly thereafter (usually 2nd or 3rd sample), most students needed more than the first schema provided, and so were switched to a schema using a significantly larger subset of TEI
- The goal of the first very tight schema was to keep the list of elements and attributes students were initially presented with short, in order to avoid their being overwhelmed, or not being able to find what they were looking for
- The goal of the second schema was to have available pretty much any TEI element a student in our class was likely to need

Ended with a single-tier approach:

- We quickly found that the effort of switching from one schema to another far outweighed the advantage of starting off with a tighter schema
- Current teaching schema is essentially just a selection of modules and some element deletions
- For future use we are considering deleting some rarely used attributes that we do not cover, as well (e.g., the global `select=` attribute)

Seminar structure

For all of the seminars, the overall structure was essentially the same: an introduction to the essential concepts of digital scholarship, followed by a set of basic technical topics and hands-on practice, with a concluding portion devoted to more advanced topics (both technical and conceptual) and discussion. In the longer seminars, we expanded the number of advanced topics and the time spent on hands-on practice.

```
start = event seminar {
  introductory,
  XML_introduction,
  basic_TEI_tagging,
  advanced+,
  technical+
}

introductory = (
    digital_research
  | descriptive_markup
  | TEI_overview
)

advanced = (
    markup_for_scholarship
  | overlap
  | pointing_linking
  | rendition
  | primary_sources
  | figures
  | non-Unicode_characters
)

technical = (
    understanding_Guidelines
  | customization
  | publication
  | CSS
)
```

```
digital_research = "A conceptual introduction to text en-
  coding in the context of digital research tools. Covers
  motives for text encoding, types of documents and approach-
  es; issues of interchange, problems and tradeoffs."

descriptive_markup = "Discussion of motives for text en-
  coding in the context of descriptive markup; basic infor-
  mation on the TEI."

TEI_overview = "Introduction to the TEI, with coverage of
  background, scope, motives, usage."

XML_introduction = "Overview of basic concepts of XML."

basic_TEI_tagging = "A basic overview of the simplest TEI
  elements, using a set of simple examples. Covers prose,
  verse, drama, and letters."

markup_for_scholarship = "An introduction to advanced mark-
  up topics with particular relevance to humanities schol-
  arship. Includes parallel texsts, alternate readings,
  editorial commentary."

overlap = "Presentation of overlapping hierarchies and re-
  lated issues; either with an emphasis on scholarly encod-
  ing problems or with an emphasis on technical issues and
  solutions, including coverage of non-TEI approaches."

pointing_linking = "Brief introduction to linking, point-
  ing, and XInclude, and how they are used."

rendition = "Presentation on the encoding of rendition and
  issues of presentational markup."

primary_sources = "Introductory presentation on encoding
  manuscript materials, including the encoding of handwrit-
  ten revisions and annotations, and also the representa-
  tion of physical book structures."

figures = "Brief introduction to encoding figures"

non-Unicode_characters = "Introduction to encoding char-
  acters that are not represented in Unicode; written by
  Christian Wittern."

understanding_Guidelines = "An overview of the technical
  underpinnings of the TEI Guidelines, including the TEI
  class system and ODD files."

customization = "An overview of how to create a custom TEI
  schema, including issues of modularity and extensibility,
  how to use the TEI's Roma tool to build a TEI schema and
  reference documentation."

publication = "An overview of the basic technologies for
  publishing TEI and XML documents, including CSS, XSLT,
  XML databases, and XML publication systems."

CSS = "Brief introduction to using CSS to style TEI docu-
  ments."
```

Participants' feedback suggests the importance of:

- Combining of lecture, discussion, hands-on
- Framing markup within scholarly domain
- Motivating technical topics through scholarly need
- Emphasizing encoding as an expression of individual scholarly perspectives
- Framing intermediate topics as applications of important general concepts
- Immediate and sustained follow-up support