Session 2: Introduction to the TEI

Historical documents, digital approaches
University of Ghent, 2013-09-05

David J. Birnbaum
djbpiM@gmail.com
http://www.obdurodon.org

Outline

• The role of the TEI in digital editing and publishing
• Parts of a TEI document
• Creating a TEI document in <oXygen/>
• TEI flavors (P5; All, manuscript description)
• Root element and namespace
  • <teiHeader>
  • <text>

The role of the TEI

• Reuse, multipurposing
• Convert to different document types
  – XHTML, PDF, plain text
• Convert to different views, e.g., a manuscript edition as
  – Continuous reading (single manuscript)
  – Critical apparatus = interlinear collation
  – Graphic (non-textual) representation
• Interchange and interoperability (shared tools)
• Guidelines or standard?
• Your problems aren’t (entirely) new or unusual
• Customizations, broad and narrow
• Funding imperatives

What do you do with a TEI document?

• Transform (XSLT) to final form (XHTML, PDF, etc.)
• Explore and query (XQuery)

How to manage a TEI project

• Choose (configure) a TEI schema
• Prepare project encoding guidelines
  – What do you need to do?
    • Do you need (or just want) to do it in a particular way?
  – There may be different ways to do almost the same thing
  – There may be different ways to do exactly the same thing
  – Nobody uses every element in the TEI in any project

Schemas

• XML must be well-formed
• XML may be valid
  – Valid = valid against a particular schema
• Schema
  – Grammar or blueprint for a document
    • What element and attributes are allowed, where, and with what content or type of content?
  – Schema languages
    • Relax NG, W3Schema, DTD
  – Constraint language: Schematron
• TEI is a set of schemas
**TEI flavors**

- TEI all
- TEI Lite
- Manuscript description
- Customization, broad and narrow
  - Can you do what you need?
  - Can you do it the way you want?
  - Can you do things you shouldn’t be allowed to do?
  - *Enabling vs enforcing* schemas

**Now let’s do something …**

- Start `<oXygen/>
- Create a new TEI all (P5) document
- What does `<oXygen/>` do for you

---

**Root element**

- `<TEI xmlns="http://www.tei-c.org/ns/1.0">`
- Namespaces
  - Namespace declaration (pseudo-attribute)
  - Inherited
  - Namespace prefixing (not covered in this workshop)
- `<title>`
  - Book title
  - Royal title
- `<body>`
  - TEI body
  - HTML body
- `<sourceDesc>`
  - `<filedesc>`
  - `<titleStmt>`
  - `<publicationStmt>`
  - `<sourceDesc>`
- `<minimalContent>`
  - `<titleStmt>`
  - `<publicationStmt>`
- `<structualContent>`

---

**Structured `<titleStmt>`**

- Required: `<title>`
- Optional
  - `<author>`, `<editor>` (of source document)
  - `<principal>` (creator of electronic document)
  - `<meeting>`
  - `<respStmt>`
    - `<resp>TEI markup</resp>
    - `<name>`[your name]`</name>`
    - `<respStmt>`

---

**<publicationStmt>**

- Minimal: `<p>`
- Optional
  - `<publisher>`, `<distributor>`, `<authority>`
  - `<pubPlace>`<address>` ...
  - `<orgName>`, `<street>`, `<settlement>`, `<country>`, `<postCode>`
  - `<date when="2013-09-05">`
  - Date ranges and less precise dates are also supported
  - `<idno type="ISBN">`
  - `<availability>`<p>[License info]</p>`</availability>`
<text>

- `<front>`, `<body>`, `<back>`
- `<div type="chapter">`
- Headers
  - `<head>`
- Verse and verse-like things
  - `<g type="stanza">`, `<i>`
- Prose
  - `<p>`
- Drama
  - `<speaker>`, `<sp speaker="hamlet">`

Common in-line elements

- `<hi rend="italic">`
- `<name type="person" ref="#hp">`
  - `<persName ref="#hp">`
  - `<placeName ref="#ghent">`

That URL