Collaborative TEI: A quick intro & exercises

Anna Kijas
Sarah DeLorme

Digital Scholarship | #BCDigSchol
Boston College Libraries | @bclibraries

What is TEI?

- The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form.
- Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics.
- Since 1994, the TEI Guidelines have been widely used by libraries, museums, publishers, and individual scholars to present texts for online research, teaching, and preservation.

http://www.tei-c.org/

XML Markup Language

◎ Non-proprietary meta language
◎ Consists of "tags" for defining structure and features of text

XHTML

<p>insert paragraph</p>

TEI

<persName>Jane Austen</persName>

<lg>insert line group</lg>

<date>August 20, 1790</date>

XML Markup Language Components

Elements (nouns of encoding)
- ◎ Basic, named structural units
- ◎ <title>Pride and Prejudice</title>

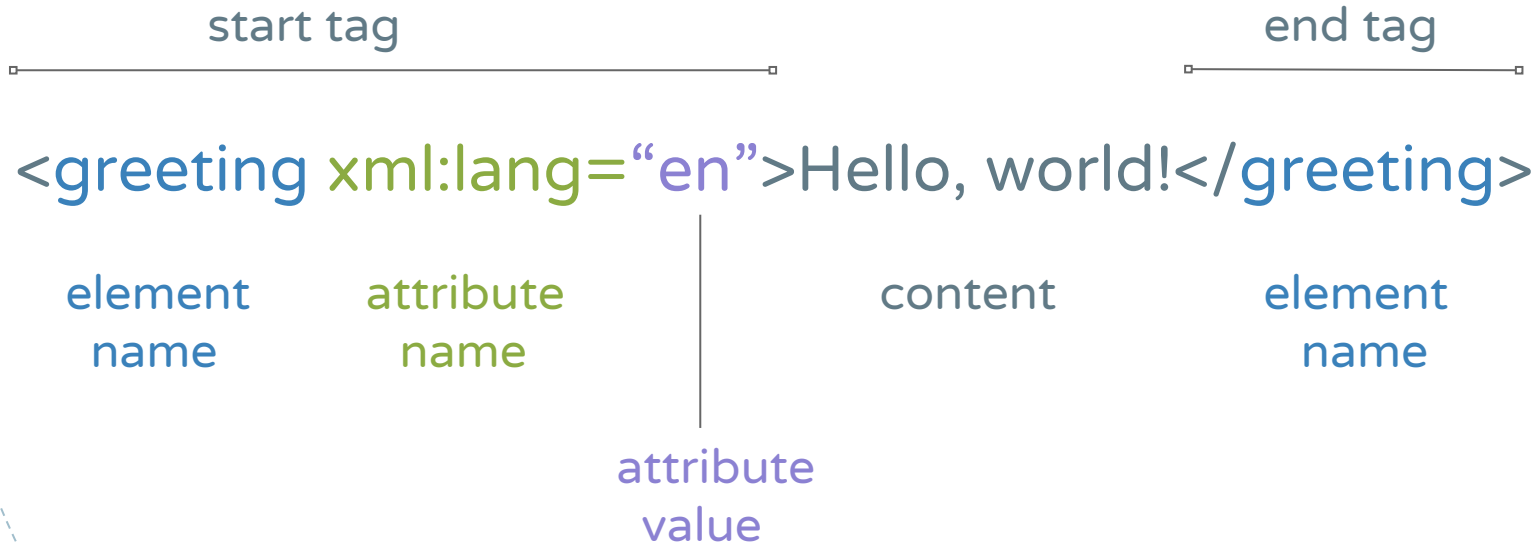**Attributes** (adjectives of encoding)
- ◎ Name/value pairs associated with elements
- ◎ <name role="author" type="person">Jane Austen</name>

DTDs and **Schemas**
- ◎ Define rules that govern a particular type of XML document

XML: Anatomy of an Element

start tag

end tag

<greeting xml:lang="en">Hello, world!</greeting>

element name

attribute name

content

element name

attribute value

XML Rules

Well-formed

- Open and close all tags/elements
- Tags/elements can't overlap
- Attribute values must be in quotes

Valid

- Use appropriate elements & attributes
- Use allowable attribute for elements

Exercise 1

Document Analysis

October 6, 1841 Minutes of the First Colored Convention

http://coloredconventions.org/transcribe-minutes

TEI P5: Basic Components

<TEI> The root element of a TEI document

<teiHeader> The metadata header. Includes bibliographic, technical, administrative, and other metadata about the digital file and analog source.

<text> The transcribed text. The <text> element can be subdivided further:

    <front> Front matter, such as title page, table of contents, preface

    <body> Main body of a document

    <back> Back matter, such as indices, appendices

TEI Header

File Description <fileDesc> incl. Source Description <sourceDesc>
    Bibliographic description of the electronic and source files

Encoding Description <encodingDesc>
    Documents the relationship between the electronic text(s) and
    source(s) from which it is derived

Profile Description <profileDesc>
    Documents non-bibliographic aspects of the text, such as languages,
    names, prosopography

Revision Description <revisionDesc>
    Documents changes made to the file

TEI Header (Required elements)

<fileDesc>, <titleStmt>, <publicationStmt>, and <sourceDesc>

```
1   <?xml version="1.0" encoding="UTF-8"?>
2   <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
        schematypens="http://relaxng.org/ns/structure/1.0"?>
3   <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
        schematypens="http://purl.oclc.org/dsdl/schematron"?>
5   <TEI xmlns="http://www.tei-c.org/ns/1.0">
6      <teiHeader>
7         <fileDesc>
8            <titleStmt>
9               <title><!-- Insert title for this digital text --></title>
10           </titleStmt>
11           <publicationStmt>
12              <p><!-- Insert information about the publication of this digital text --></p>
13           </publicationStmt>
14           <sourceDesc>
15              <p><!-- Insert a bibliographic description of the source for this digital text --></p>
16           </sourceDesc>
17        </fileDesc>
18     </teiHeader>
```

TEI P5 Guidelines: http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html

TEI Header (Source Description)

```
27 ▾ · · · · · · · · · <sourceDesc>¬
28 ▾ · · · · · · · · · · · <bibl>¬
29 · · · · · · · · · · · · · · <title>Love·and·death</title>¬
30 · · · · · · · · · · · · · · <author>William·Butler·Yeats</author>¬
31 · · · · · · · · · · · · · · <date>1884</date>¬
32 · · · · · · · · · · · · · </bibl>¬
33 · · · · · · · · · </sourceDesc>¬
```

TEI P5 Guidelines: http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html

Exercise 2

Encode TEI Header

https://github.com/BCDigSchol/collaborative-tei

TEI P5: Basic Markup

\<div\> (division) is used for basic structural divisions of a text (volumes, chapters, indices, etc.) The @type attribute can be used to designate the type of division.

    \<div type="chapter"\>Chapter 1\</div\>

\<head\> (heading) contains any type of heading, title of a section, list, figure, etc.

\<p\> (paragraph) marks paragraphs in a text

\<pb\> (page break) marks the boundary between one page of a text and the next

Other Elements

<name> (name, proper noun) contains a proper noun
    <name type="person">Frederick Douglass</name>
    <name type="place">Portland, Maine</name>

<persName> (personal name) contains a proper noun phrase referring to a person
    <persName>Frederick Douglass</persName>

<date> contains a date in any format
<date when="1841-10-06">October 6, 1841</date>

- Can use appropriate attributes with these elements

Element and Attributes for Names, Dates, People, and Places
http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html

Exercise 3

Encode Text Sample

https://github.com/BCDigSchol/collaborative-tei

- Review and discuss

- Continue encoding samples or select new text
  - Colored Conventions http://coloredconventions.org/transcribe-minutes

- Explore encoding samples in TAPAS
  - Create record and upload sample encoding
  - http://www.tapasproject.org/tei-learning-docs-bc/colored-conventions-sample

Contact Us

Anna Kijas - anna.kijas at bc.edu

Sarah DeLorme - sarah.delorme at bc.edu

Info about upcoming events: ds.bc.edu/events

Credits

Some material was borrowed or modeled on Michelle Dalmau/Anna Kijas' ARL DSI Scholarly Editions workshop:

https://github.com/tech-at-arl/Digital-Scholarship-Institute/

Special thanks to all the people who made and released these awesome resources for free:

◎ Presentation template by SlidesCarnival

◎ Photographs by Unsplash