

# 3.2 Encoding Text – TEI

Matt Price

# Encodings

- HTML is a text encoding system
- but of course there are many others, often with different purposes!
- HTML is a **web standard** but there are other **standards**
- **TEI** is an influential standard designed **by and for** humanists for **encoding textual complexity**

# TEI is XML

- the **eXtensible Markup Language** looks a bit like HTML but is much more complex
- remains a key technology in many domains (e.g., your word docs and powerpoints slides are written in XML)
- built with similar syntax but much, much more powerful
- also **way** harder to read

```
<book category="children">
  <title>Harry Potter</title>
  <author>J K. Rowling</author>
  <year>2005</year>
  <price>29.99</price>
</book>
```

# Real TEI

```
<lg type="poem">
  <lg type="stanza">
    <l>
      <rhyme>Poppadom</rhyme>
    </l>
    <l>
      <rhyme>Oatmeal</rhyme>
    </l>
    <l>Bubble <rhyme>gum</rhyme> </l>
    <l>Cut of <rhyme>veal</rhyme> </l>
  </lg>
  <lg type="stanza">
    <l>Mince for <rhyme>pie</rhyme> </l>
    <l>Frozen <rhyme>peas</rhyme> </l>
    <l>Video for <rhyme>Guy</rhyme> </l>
    <l>Selection of <rhyme>teas</rhyme> </l>
  </lg>
  <lg type="stanza">
    <l>Paper towels/garbage <rhyme>ba</rhyme>
    <l>Pasta sauce and <rhyme>Parmesa</rhyme>
```

- all these tags are **defined in a standard** which is **available online** in a readable format but **also is a machine-readable standard for XML interpreters**
- The power of XML/TEI (like that of HTML) comes in part from that **machine-readable standard**

# Using TEI

- we won't!
- but you might want to if:
  - you want to record textual complexities in a rigorous format
  - you hope to collaborate with others/make your work useful to others
  - you want to join a scholarly community around precise textual analysis
- our work in the class is **theatically similar** even if we don't use the **same technology**