Part 1: Introduction to Markup, XML, and XML Editing Options

March 8, 2008

1 Overview of all sessions

Part 1: Reminders of the concept of textual markup, XML, and editing options for TEI users in Chinese. Work with Chinese texts.

Part 2: Introduction to the TEI, its class system, default text structure, and core elements and header. Work with TEI Roma application in Chinese.

Part 3: TEI and localization. A concentration on how the TEI is internationalized and localized, and what the impact of this is on its usage.

Part 4: TEI P5. What has been added to the TEI in the last five years? Looking at the new features of the scheme.

Part 5: Accessing TEI texts. Untangling the XML markup and extracting what you need, with particular reference to the TEI XSL stylesheet family and its localization features.

Part 6: TEI Applications. What can you do with a text marked up in TEI XML, apart from simply displaying it?

1.1 Aims for entire workshop

1. Examine the concept of markup and XML encoding
2. Provide hands-on experience in using XML markup
3. Introduce the TEI, its assumptions, and how it is organised
4. Survey TEI recommendations that may be of interest
5. Demonstrate the benefits of project-specific customisation of the TEI
6. Survey the ways in which people transform, query, and publish XML
7. Provide a brief overview so that you can explore in more depth later
8. Give a chance for questions and discussion of participants concerns

Contents

1.2 Workshop Acknowledgements

The slides, and ideas, in this workshop borrow heavily from previous presentations/workshops by:

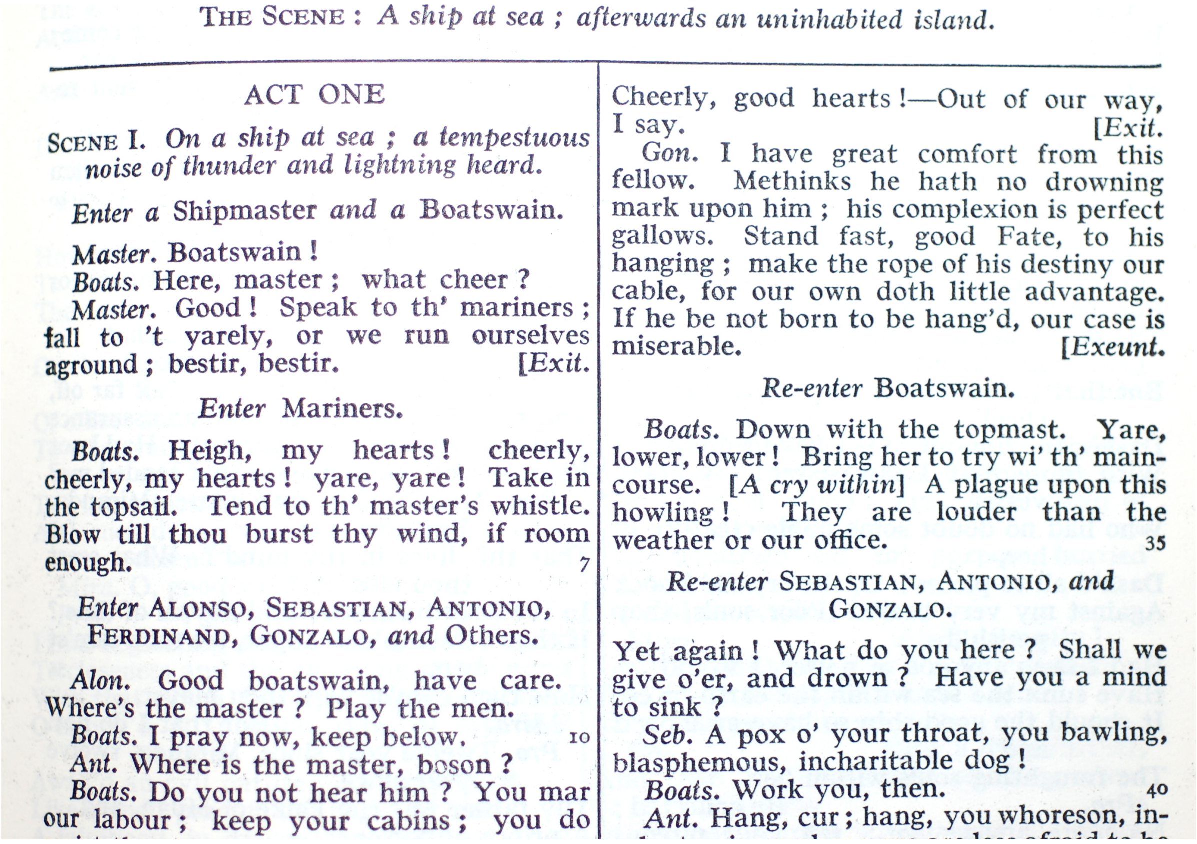
- Lou Burnard
- James Cummings
- Dot Porter

And the slides from this workshop are licensed for re-use by others should they desire it.

2 Markup

In order to talk about texts, markup and encoding of texts, we need to understand what we mean by these basic concepts. When we talk about text encoding, what do we mean by a text? What is in a text and what assumptions do we make in reading them?

2.1 What's in a text?



THE SCENE : *A ship at sea ; afterwards an uninhabited island.*

ACT ONE

SCENE I. *On a ship at sea ; a tempestuous noise of thunder and lightning heard.*

Enter a Shipmaster *and a* Boatswain.

Master. Boatswain !

Boats. Here, master ; what cheer ?

Master. Good ! Speak to th' mariners ; fall to 't yarely, or we run ourselves aground ; bestir, bestir. [*Exit.*

Enter Mariners.

Boats. Heigh, my hearts ! cheerly, cheerly, my hearts ! yare, yare ! Take in the topsail. Tend to th' master's whistle. Blow till thou burst thy wind, if room enough.

Enter ALONSO, SEBASTIAN, ANTONIO, FERDINAND, GONZALO, *and* Others.

Alon. Good boatswain, have care. Where's the master ? Play the men.

Boats. I pray now, keep below.

Ant. Where is the master, boson ?

Boats. Do you not hear him ? You mar our labour ; keep your cabins ; you do

Cheerly, good hearts !—Out of our way, I say. [*Exit.*

Gon. I have great comfort from this fellow. Methinks he hath no drowning mark upon him ; his complexion is perfect gallows. Stand fast, good Fate, to his hanging ; make the rope of his destiny our cable, for our own doth little advantage. If he be not born to be hang'd, our case is miserable. [*Exeunt.*

Re-enter Boatswain.

Boats. Down with the topmast. Yare, lower, lower ! Bring her to try wi' th' maincourse. [*A cry within*] A plague upon this howling ! They are louder than the weather or our office.

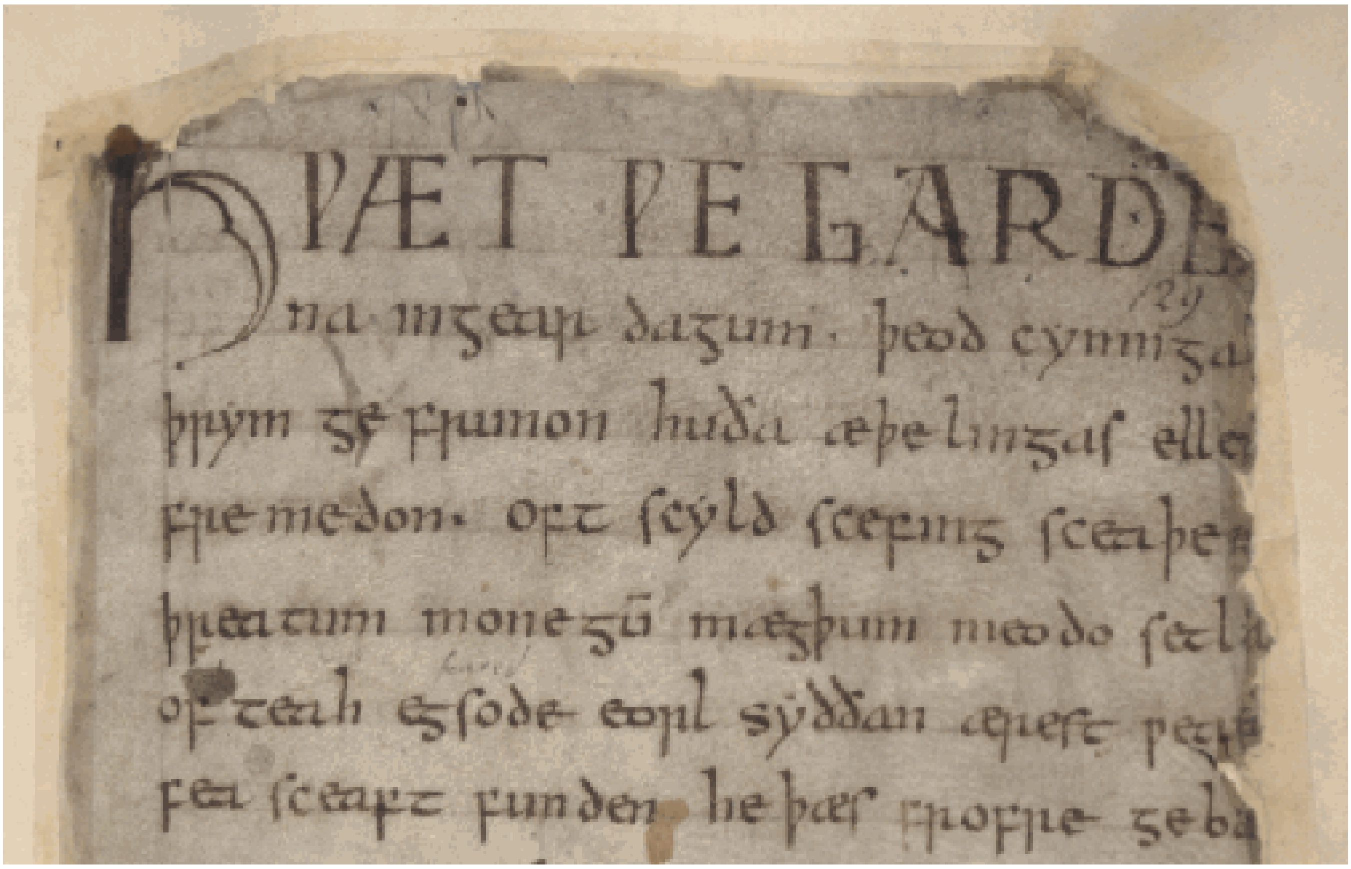
Re-enter SEBASTIAN, ANTONIO, *and* GONZALO.

Yet again ! What do you here ? Shall we give o'er, and drown ? Have you a mind to sink ?

Seb. A pox o' your throat, you bawling, blasphemous, incharitable dog !

Boats. Work you, then.

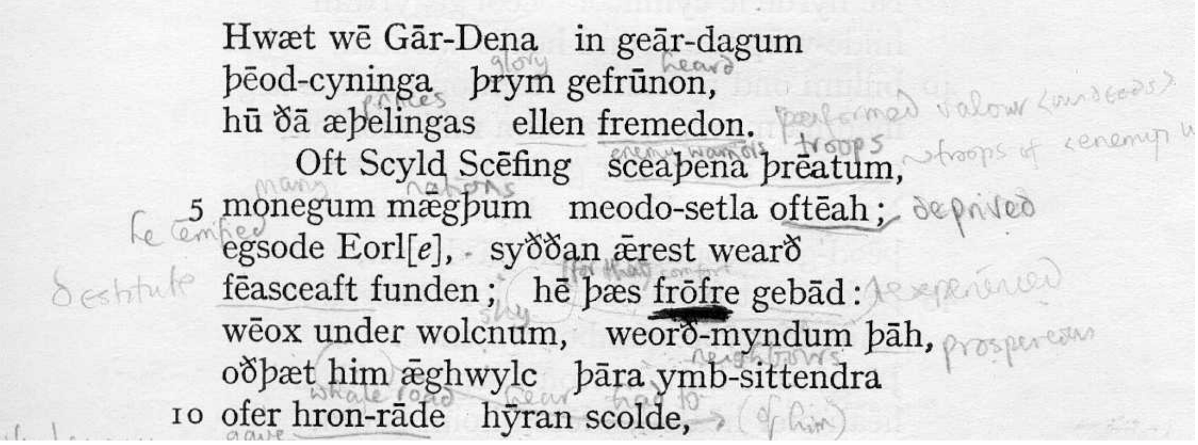
Ant. Hang, cur ; hang, you whoreson, in-



HWÆT WE GARDE
na in gear dagum . þeod cyninga
þrym ge frunon huða æþelingas elle
fre medon. oft scyld scefing sceaþe
þreatum monegū mægþum meodo setl
of teah egsode eorl syððan ærest weo
fea sceaft funden he þæs frofre geba

2.2 What's in a text (2)?

2.3 What's in a text (3)?



Hwæt wē Gār-Dena in geār-dagum
þēod-cyninga þrym gefrūnon,
hū ðā æþelingas ellen fremedon.
Oft Scyld Scēfing sceaþena þrēatum,
monegum mægþum meodo-setla oftēah;
egsode Eorl[*e*], syððan ǣrest wearð
fēasceaft funden; hē þæs frōfre gebād:
wēox under wolcnum, weorð-myndum þāh,
oðþæt him ǣghwylc þāra ymb-sittendra
ofer hron-rāde hȳran scolde,

glory
heard
princes
performed valour <deeds>
troops
troops of <enemy
enemy warriors
many
nations
deprived
he terrified
destitute
for that comfort
experienced
prosperous
neighbours
whale road
obey
had to
→ (of him)
gave

2.4 The ontology of text

Where is the text?

- in the shape of letters and their layout?
- in the original from which this copy derives?
- in the stories we read into it? or in its author's intentions?

A "text" is an abstraction, created by or for a community of readers. Markup encodes and makes concrete such abstractions.

2.5 Encoding of texts

- Texts are more than sequences of encoded glyphs
  - They have **structure** and **content**
  - They also have multiple **readings**
- Encoding, or markup, is a way of making these things explicit
- Only that which is explicit can be reliably processed

2.6 Styles of markup

- In the beginning there was *procedural* markup
  `RED INK ON; print balance; RED INK OFF`
- which being generalised became *descriptive* markup <balance type='overdrawn'>some numbers</balance>
- also known as **encoding** or **annotation**

descriptive markup allows for easier re-use of data

2.7 Some more definitions

- Markup makes explicit the distinctions we want to make when processing a string of bytes
- Markup is a way of naming and characterizing the parts of a text in a formalized way
- It's (usually) more useful to markup what we think things *are* than what they *look like*

2.8 What's the point of markup?

- To make explicit (to a machine) what is implicit (to a person)
- To add value by supplying multiple annotations
- To facilitate re-use of the same material
  - in different formats
  - in different contexts
  - by different users

2.9 Separation of form and content

- Presentational markup cares more about fonts and layout than meaning
- Descriptive markup says what things are, and leaves the rendition of them for a separate step
- Separating the form of something from its content makes its re-use more flexible
- It also allows easy changes of presentation across a large number of documents

2.10 Markup as a scholarly activity

- The application of markup to a document can be an intellectual activity
- In deciding what markup to apply, and how this represents the original, one is undertaking the task of an editor
- There is (almost) no such thing as neutral markup -- all of it involves interpretation
- Markup can assist in answering research questions, and the deciding what markup is needed to enable such questions to be answered can be a research activity in itself
- Good textual encoding is never as easy or quick as people would believe
- Detailed document analysis is needed before encoding for the resulting markup to be useful

2.11 What does markup capture?

Compare

```
<hi rend="dropcap">H</hi>&WYN;ÆT WE GARDE
<lb/>na in gear-dagum þeod-cyninga
<lb/>þrym gefrunon, hu ða æþelingas
<lb/>ellen fremedon.  oft scyld scefing sceaþe
<add>na</add>
<lb/>þreatum, moneg<expan>um</expan>mægþum meodo-setl
<add>a</add>
<lb/>of<damage desc="blot"/>teah ...
```

and

```
<lg>
 <l>Hwæt!  we Gar-dena in gear-dagum</l>
 <l>þeod-cyninga þrym gefrunon,</l>
 <l>hu ða æþelingas ellen fremedon,</l>
</lg>
<lg>
 <l>Oft Scyld Scefing sceaþena þreatum,</l>
 <l>monegum mægþum meodo-setla ofteah;</l>
 <l>egsode Eorle, syððan ærest wearþ</l>
 <l>feasceaft funden...</l>
</lg>
```

2.12 A useful mental exercise

Imagine you are going to markup several thousand pages of complex material....

- Which features are you going to markup?
- Why are you choosing to markup this feature?
- How reliably and consistently can you do this?

Now, imagine your budget has been halved. Repeat the exercise!

2.13 Some alphabet soup

SGML	Standard Generalized Markup Language
HTML	Hypertext Markup Language
W3C	World Wide Web Consortium
XML	eXtensible Markup Language
DTD	Document Type Definition (or Declaration)
CSS	Cascading Style Sheet
Xpath	XML Path Language
XSLT	eXtensible Stylesheet Language - Transformations
XQuery	XML Querying
RELAXNG	Regular Expression Language for XML (New Generation)

Oh, and then there's also **TEI**, the *Text Encoding Initiative*

3 XML

Extensible Markup Language (XML) is a simple, very flexible text format derived from SGML (ISO 8879). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.

3.1 XML: what it is and why you should care

- XML is **structured data** represented as strings of text
- XML looks like HTML, except that:-
  - XML is **extensible**
  - XML must be **well-formed**
  - XML can be **validated**
- XML is application-, platform-, and vendor- independent
- XML empowers the **content provider** and facilitates data integration

3.2 XML terminology

An XML document may contain:-

- elements, possibly bearing attributes
- processing instructions
- comments
- entity references
- marked sections (CDATA, IGNORE, INCLUDE)

An XML document must be **well-formed** and may be **valid**

3.3 XML terminology Example

```
<?xml version="1.0"?>
<root>
 <element attribute="value"> content </element>
<!-- comment -->
</root>
```

3.4 The rules of the XML Game

- An XML document represents a (kind of) **tree**
- It has a single **root** and many nodes
- Each node can be
  - a subtree
  - a single **element** (possibly bearing some **attributes**)
  - a string of **character data**
- Each element has a name or **generic identifier**
- Attribute names are predefined for a given element; values can also be constrained

3.5 Representing an XML tree

- An XML document is encoded as a linear string of characters
- It begins with a special **processing instruction**
- Element occurrences are marked by **start**- and **end-tags**
- The characters < and & are Magic and must always be "escaped" if you want to use them as themselves
- **Comments** are delimited by <!- - and - ->
- **CDATA sections** are delimited by <![CDATA[and]]>
- Attribute name/value pairs are supplied on the start-tag and may be given in any order
- Entity references are delimited by & and ;

3.6 Parts of an XML document

```
<?xml version="1.0"?>
<greetings>
 <hello type="sarcastic">hello world!</hello>
</greetings>
```

- The XML declaration
- Namespace declarations
- The root element of the document itself
- Other elements and content
- Attribute and value

3.7 The XML declaration

An XML document must begin with an **XML declaration** which does two things:

- specifies that this *is* an XML document, and which version of the XML standard it follows
- specifies which character encoding the document uses

```
<?xml version="1.0" ?>
<?xml version="1.0" encoding="iso-8859-1" ?>
```

The default, and recommended, encoding is UTF-8

3.8 Namespace declarations

All TEI documents are declared within the TEI namespace: `<TEI xmlns="http://www.tei-c.org/ns/1.0"> ... </TEI>`

XML documents can include elements declared in different **name spaces**.

- a namespace declaration associates a namespace prefix with an external URI-like identifier
- the default namespace *may* be declared using a `xmlns`
- other name spaces must all use a specially declared prefix

```
<TEI xmlns="http://www.tei-c.org/ns/1.0"
xmlns:math="http://www.mathml.org"> <p>...<math:expr>...</math:expr>...</p>...</TEI>
```

The `xml` namespace is used by the TEI for global attributes xml:id and xml:lang

3.9 The Doctype Declaration

You may sometimes find an optional "Document Type" declaration:

```
   <?xml version="1.0" ?>
   <!DOCTYPE greeting SYSTEM "greeting.dtd []">
```

- The DTD is one way of associating the document with its schema (but is not used by W3C or RELAXNG for this purpose)
- The DTD subset is used to provide declarations additional to those in the schema, for example for external files
- The DTD subset may be **internal**, **external**, or both

DTDs are now considered old-fashioned -- RELAXNG or W3C schemas are preferred.

3.10 The Tempest

```
<?xml version="1.0" encoding="utf-8" ?>
<div n="1">
 <head>SCENE I. On a ship at sea: a tempestuous noise of thunder and lightning
heard.</head>
 <stage>Enter a Master and a Boatswain</stage>
 <sp>
  <speaker>Master</speaker>
  <ab>Boatswain!</ab>
 </sp>
 <sp>
  <speaker>Boatswain</speaker>
  <ab>Here, master: what cheer?</ab>
 </sp>
 <sp>
  <speaker>Master</speaker>
  <ab>Good, speak to the mariners: fall to't, yarely,</ab>
  <ab>or we run ourselves aground: bestir, bestir.</ab>
 </sp>
 <stage>Exit</stage>
</div>
```

3.11 Example deconstructed: root node

```
<?xml version="1.0" encoding="utf-8" ?>
<div n="1">
<!-- ....  -->
</div>
```

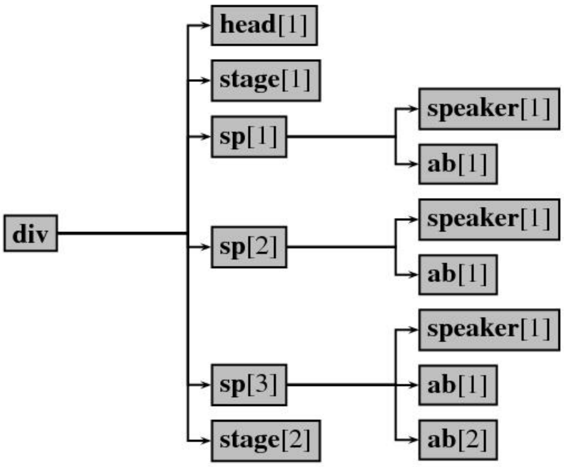
3.12 Example deconstructed: head

```
<head>SCENE I. On a ship at sea:  a tempestuous noise of thunder and lightning
heard.</head>
```

3.13 Example deconstructed: stage direction and speech

```
<stage>Enter a Master and a Boatswain</stage>
<sp>
 <speaker>Master</speaker>
 <ab>Boatswain!</ab>
</sp>
```

3.14 An XML Tree For The Tempest



3.15 XML syntax: the small print

What does it mean to be **well-formed**?

1. there is a single root node containing the whole of an XML document
2. each subtree is properly nested within the root node
3. names are always case sensitive
4. start-tags and end-tags are always mandatory (except that a combined start-and-end tag may be used for empty nodes)
5. attribute values are always quoted

Note: You can be **valid** in addition to being well-formed. This means you obey the rules of a specified schema, such as the TEI.

3.16 Test your XML knowledge

- Which are correct?
  - `<seg>some text</seg>`
  - `<seg><foo>some</foo> <bar>text</bar></seg>`
  - `<seg><foo>some <bar></foo> text</bar></seg>`
  - `<seg type="text">some text</seg>`
  - `<seg type='text'>some text</seg>`
  - `<seg type=text>some text</seg>`
  - `<seg type = "text">some text</seg>`
  - `<seg type="text">some text<seg/>`
  - `<seg type="text">some text<gap/></seg>`
  - `<seg type="text">some text< /seg>`
  - `<seg type="text">some text</Seg>`

3.17 XML is an international standard

- XML requires use of ISO 10646 (also known as Unicode)
  - a 31 bit character repertoire including most human writing systems
  - encoded as UTF8 or UTF16
- other encodings may be specified at the document level
- language may be specified at the element level using xml:lang

The xml:id attribute is another W3C-defined attribute.

4 Editing Options

This section provides a brief overview of technology for editing in XML, especially for TEI XML users, and issues related to that in the area of data capture and editing.

4.1 Summary

How does a TEI user do the following?

- Data capture
- Editing

4.2 What tools do we need?

- Appropriately expressive vocabularies (eg TEI XML)
- Syntax-checking document creation tools (ie editors)
- Document transformation tools
- Document delivery tools
- Document storage and management tools
- Programming interfaces
- Specialized applications

4.3 Two stages to get a TEI text

- capture the text
- create the markup

Often they occur simultaneously; but often not.

Note that the markup does not necessarily all have to be in the same file.

4.4 Categories of creation tools

- scanning/OCR
- data-entry vendors
- software to add tagging automatically
- editors

followed by

- validators, well-formedness checkers
- proofing aids, data integrity checkers

4.5 OCR/Data Entry

- Scanning and OCR software generally produce only minimal HTML or Word (e.g., recognizing paragraph breaks, font changes etc).
- Data-entry vendors in theory would insert whatever markup you wanted, but at a price. They generally prefer HTML or TEI Lite or some such well-known DTD.
- TEI is creating a standard slimed-down vocabulary for initial encoding that may be useful in mass-digitisation projects called 'TEI Tite'.

4.6 Editor types

Editing tools cover a wide spectrum:

- Basic text editors
- General programmers' editors
- XML-aware programmers' editors
- XML-specific editors
- Word-processors which can export XML
- Data-entry forms
- Image-specific editors

it is likely that people in different roles need different tools.

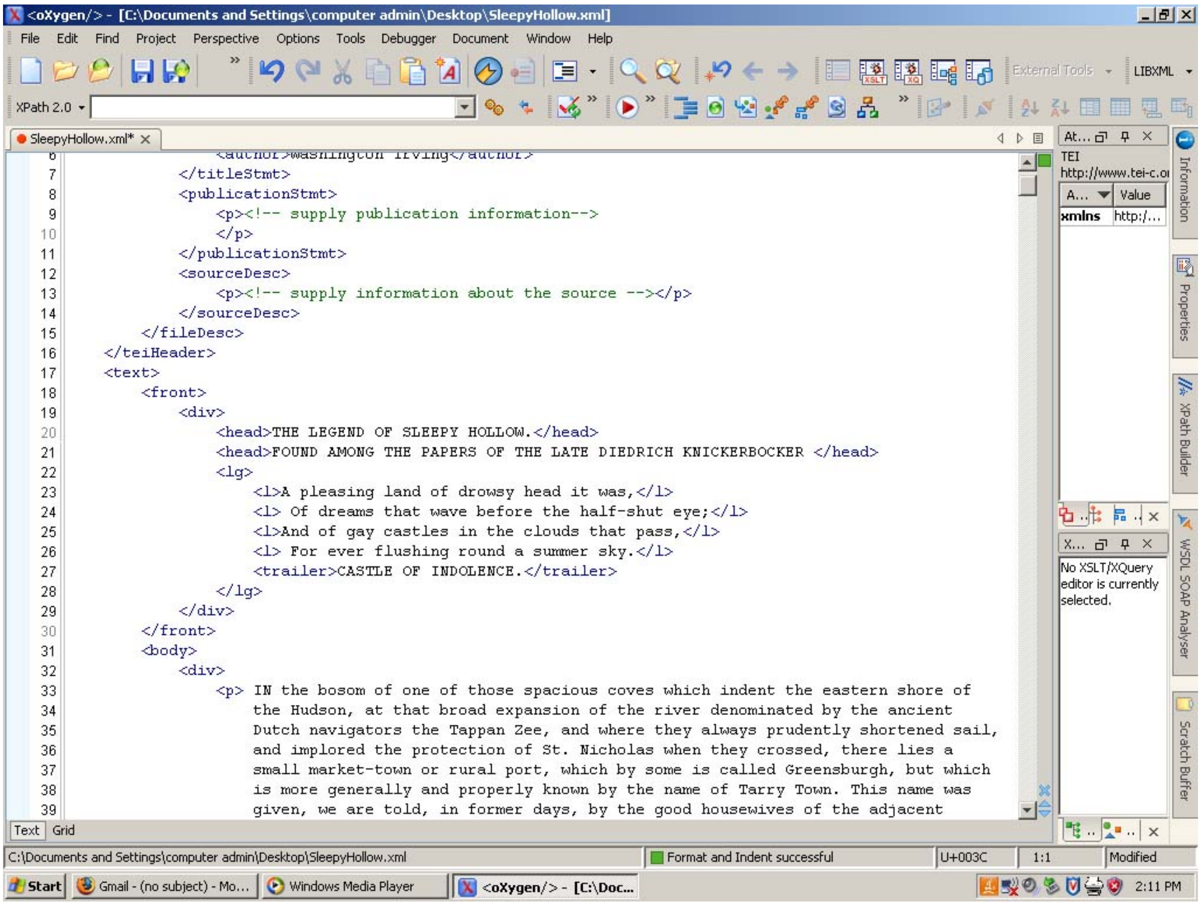
4.7 Things to look for in specialist XML editors

- schema-aware
- constraining element entry
- IDE features
- customizable
- validation, preferably continual
- Multiple display views (as tree, with tags, formatted etc)
- folding structures
- context-sensitive help

For XML editing, Emacs, **oXygen**, jEdit, XMetaL, XMLSpy, Stylus Studio, Arbortext Adept are all worth a look.

For image editing, try **University of Victoria Image Markup Tool** or **Edition Production and Presentation Technology (EPPT)**.

4.8 oXygen screenshot 1



<oXygen/> - [C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml]

```
                <author>Washington Irving</author>
            </titleStmt>
            <publicationStmt>
                <p><!-- supply publication information-->
                </p>
            </publicationStmt>
            <sourceDesc>
                <p><!-- supply information about the source --></p>
            </sourceDesc>
        </fileDesc>
    </teiHeader>
    <text>
        <front>
            <div>
                <head>THE LEGEND OF SLEEPY HOLLOW.</head>
                <head>FOUND AMONG THE PAPERS OF THE LATE DIEDRICH KNICKERBOCKER </head>
                <lg>
                    <l>A pleasing land of drowsy head it was,</l>
                    <l> Of dreams that wave before the half-shut eye;</l>
                    <l>And of gay castles in the clouds that pass,</l>
                    <l> For ever flushing round a summer sky.</l>
                    <trailer>CASTLE OF INDOLENCE.</trailer>
                </lg>
            </div>
        </front>
        <body>
            <div>
                <p> IN the bosom of one of those spacious coves which indent the eastern shore of
                    the Hudson, at that broad expansion of the river denominated by the ancient
                    Dutch navigators the Tappan Zee, and where they always prudently shortened sail,
                    and implored the protection of St. Nicholas when they crossed, there lies a
                    small market-town or rural port, which by some is called Greensburgh, but which
                    is more generally and properly known by the name of Tarry Town. This name was
                    given, we are told, in former days, by the good housewives of the adjacent
```

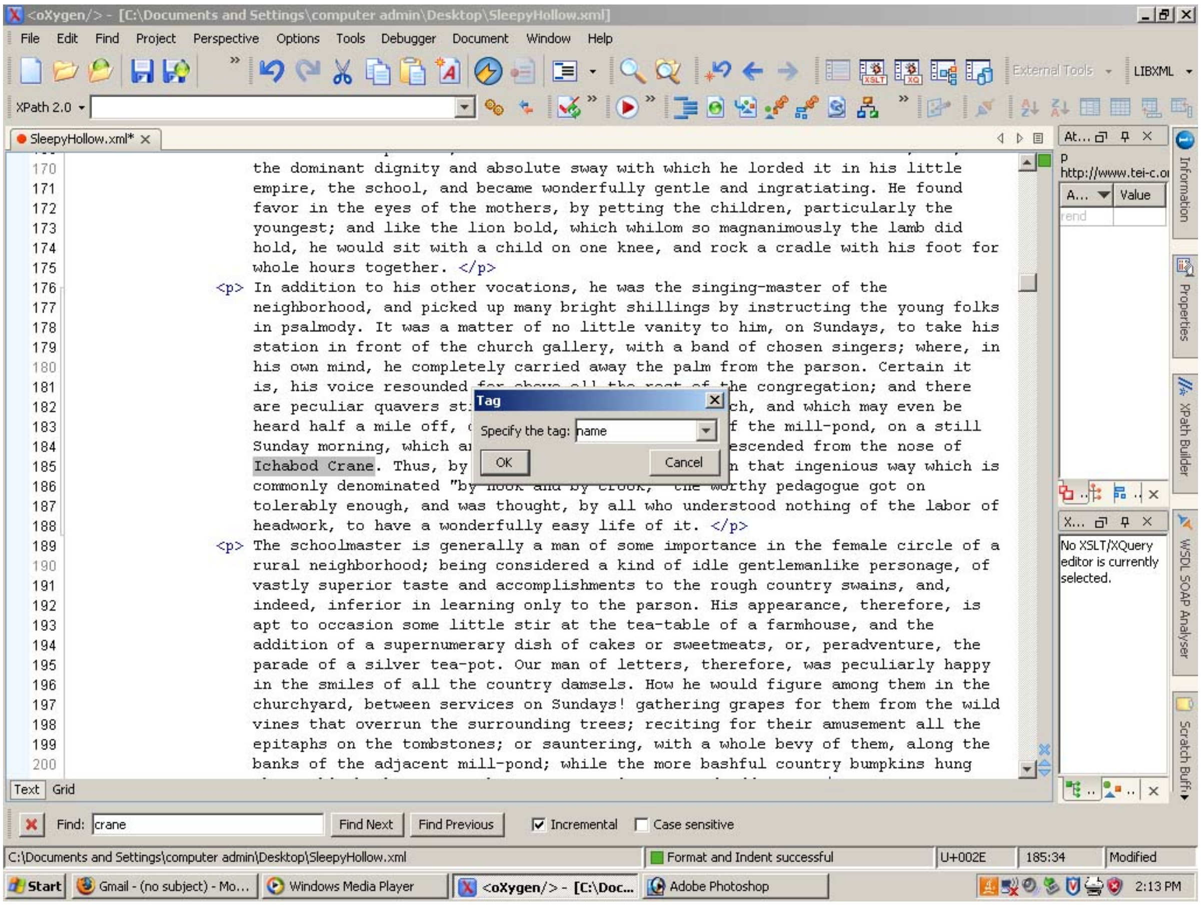
C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml | Format and Indent successful | U+003C | 1:1 | Modified

No XSLT/XQuery editor is currently selected.

TEI http://www.tei-c.o

A...	Value
xmlns	http:/...

4.9 oXygen screenshot 2



```
                the dominant dignity and absolute sway with which he lorded it in his little
                empire, the school, and became wonderfully gentle and ingratiating. He found
                favor in the eyes of the mothers, by petting the children, particularly the
                youngest; and like the lion bold, which whilom so magnanimously the lamb did
                hold, he would sit with a child on one knee, and rock a cradle with his foot for
                whole hours together. </p>
            <p> In addition to his other vocations, he was the singing-master of the
                neighborhood, and picked up many bright shillings by instructing the young folks
                in psalmody. It was a matter of no little vanity to him, on Sundays, to take his
                station in front of the church gallery, with a band of chosen singers; where, in
                his own mind, he completely carried away the palm from the parson. Certain it
                is, his voice resounded for above all the rest of the congregation; and there
                are peculiar quavers sti                    ch, and which may even be
                heard half a mile off, c                    f the mill-pond, on a still
                Sunday morning, which ar                    escended from the nose of
                Ichabod Crane. Thus, by                     n that ingenious way which is
                commonly denominated "by hook and by crook," the worthy pedagogue got on
                tolerably enough, and was thought, by all who understood nothing of the labor of
                headwork, to have a wonderfully easy life of it. </p>
            <p> The schoolmaster is generally a man of some importance in the female circle of a
                rural neighborhood; being considered a kind of idle gentlemanlike personage, of
                vastly superior taste and accomplishments to the rough country swains, and,
                indeed, inferior in learning only to the parson. His appearance, therefore, is
                apt to occasion some little stir at the tea-table of a farmhouse, and the
                addition of a supernumerary dish of cakes or sweetmeats, or, peradventure, the
                parade of a silver tea-pot. Our man of letters, therefore, was peculiarly happy
                in the smiles of all the country damsels. How he would figure among them in the
                churchyard, between services on Sundays! gathering grapes for them from the wild
                vines that overrun the surrounding trees; reciting for their amusement all the
                epitaphs on the tombstones; or sauntering, with a whole bevy of them, along the
                banks of the adjacent mill-pond; while the more bashful country bumpkins hung
```

Tag

Specify the tag: name

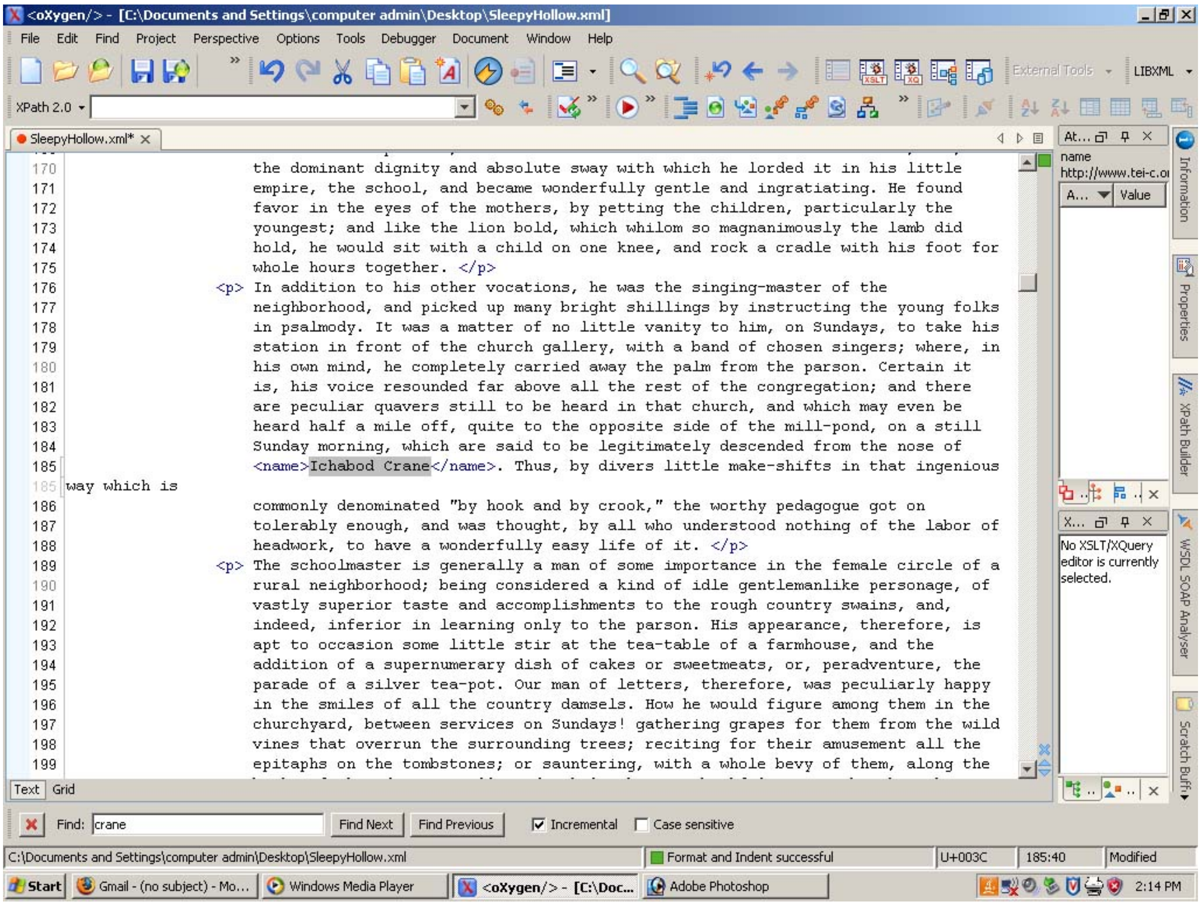
OK Cancel

No XSLT/XQuery editor is currently selected.

Find: crane Find Next Find Previous ☑ Incremental □ Case sensitive

C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml | Format and Indent successful | U+002E | 185:34 | Modified

4.10 oXygen screenshot 3



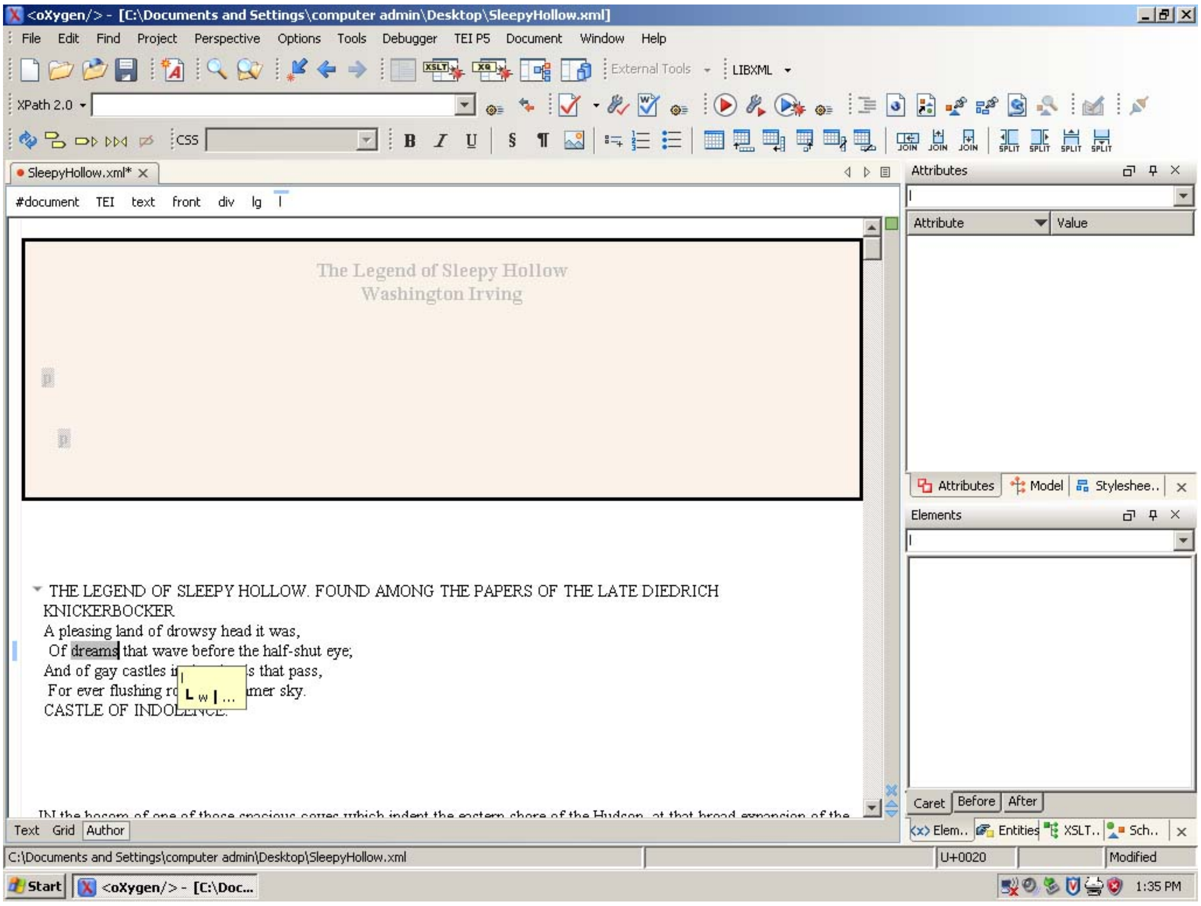
```
                    the dominant dignity and absolute sway with which he lorded it in his little
                    empire, the school, and became wonderfully gentle and ingratiating. He found
                    favor in the eyes of the mothers, by petting the children, particularly the
                    youngest; and like the lion bold, which whilom so magnanimously the lamb did
                    hold, he would sit with a child on one knee, and rock a cradle with his foot for
                    whole hours together. </p>
                <p> In addition to his other vocations, he was the singing-master of the
                    neighborhood, and picked up many bright shillings by instructing the young folks
                    in psalmody. It was a matter of no little vanity to him, on Sundays, to take his
                    station in front of the church gallery, with a band of chosen singers; where, in
                    his own mind, he completely carried away the palm from the parson. Certain it
                    is, his voice resounded far above all the rest of the congregation; and there
                    are peculiar quavers still to be heard in that church, and which may even be
                    heard half a mile off, quite to the opposite side of the mill-pond, on a still
                    Sunday morning, which are said to be legitimately descended from the nose of
                    <name>Ichabod Crane</name>. Thus, by divers little make-shifts in that ingenious
way which is
                    commonly denominated "by hook and by crook," the worthy pedagogue got on
                    tolerably enough, and was thought, by all who understood nothing of the labor of
                    headwork, to have a wonderfully easy life of it. </p>
                <p> The schoolmaster is generally a man of some importance in the female circle of a
                    rural neighborhood; being considered a kind of idle gentlemanlike personage, of
                    vastly superior taste and accomplishments to the rough country swains, and,
                    indeed, inferior in learning only to the parson. His appearance, therefore, is
                    apt to occasion some little stir at the tea-table of a farmhouse, and the
                    addition of a supernumerary dish of cakes or sweetmeats, or, peradventure, the
                    parade of a silver tea-pot. Our man of letters, therefore, was peculiarly happy
                    in the smiles of all the country damsels. How he would figure among them in the
                    churchyard, between services on Sundays! gathering grapes for them from the wild
                    vines that overrun the surrounding trees; reciting for their amusement all the
                    epitaphs on the tombstones; or sauntering, with a whole bevy of them, along the
```

No XSLT/XQuery editor is currently selected.

Find: crane ☑ Incremental □ Case sensitive

C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml | Format and Indent successful | U+003C | 185:40 | Modified

4.11 Tagless editing in oXygen



<oXygen/> - [C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml]
File Edit Find Project Perspective Options Tools Debugger TEI P5 Document Window Help
External Tools
LIBXML
XPath 2.0
CSS
SleepyHollow.xml*
#document TEI text front div lg l
Attributes
Attribute
Value
Attributes Model Styleshee..
Elements
Caret Before After
Text Grid Author
Elem.. Entities XSLT.. Sch..
U+0020
Modified
Start
<oXygen/> - [C:\Doc...
1:35 PM

The Legend of Sleepy Hollow
Washington Irving

p

p

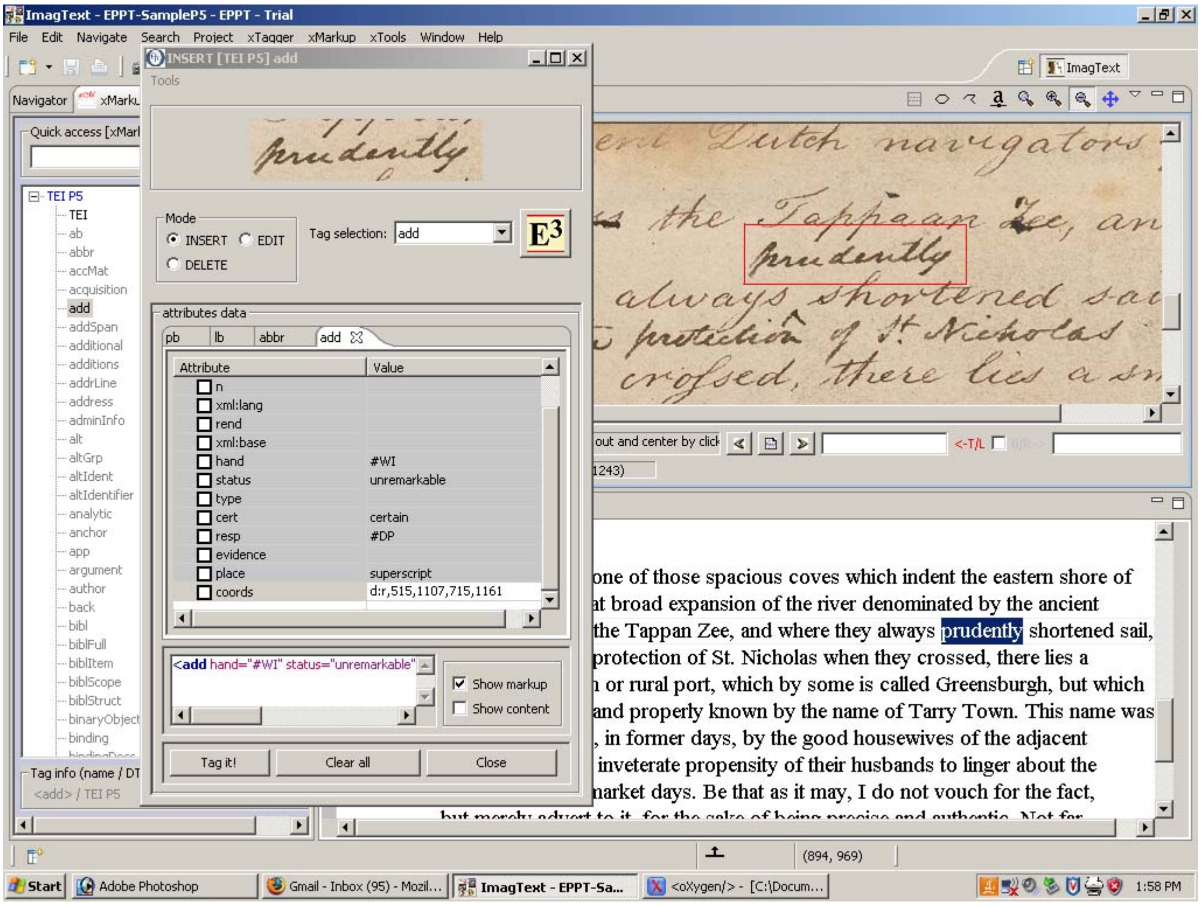
THE LEGEND OF SLEEPY HOLLOW. FOUND AMONG THE PAPERS OF THE LATE DIEDRICH KNICKERBOCKER

A pleasing land of drowsy head it was,
Of dreams that wave before the half-shut eye;
And of gay castles i s that pass,
For ever flushing r mer sky.
CASTLE OF INDOL

IN the bosom of one of those spacious coves which indent the eastern shore of the Hudson, at that broad expansion of the

C:\Documents and Settings\computer admin\Desktop\SleepyHollow.xml

4.12 EPPT



Mode

● INSERT ○ EDIT

○ DELETE

Tag selection: add

E3

attributes data

pb | lb | abbr | add

Attribute	Value
□ n	
□ xml:lang	
□ rend	
□ xml:base	
□ hand	#WI
□ status	unremarkable
□ type	
□ cert	certain
□ resp	#DP
□ evidence	
□ place	superscript
□ coords	d:r,515,1107,715,1161

```
<add hand="#WI" status="unremarkable"
```

☑ Show markup

□ Show content

Tag it! | Clear all | Close

one of those spacious coves which indent the eastern shore of
at broad expansion of the river denominated by the ancient
the Tappan Zee, and where they always prudently shortened sail,
protection of St. Nicholas when they crossed, there lies a
n or rural port, which by some is called Greensburgh, but which
and properly known by the name of Tarry Town. This name was
, in former days, by the good housewives of the adjacent
inveterate propensity of their husbands to linger about the
narket days. Be that as it may, I do not vouch for the fact,
but merely advert to it, for the sake of being precise and authentic. Not for

4.13 UVic IMT screenshot 1

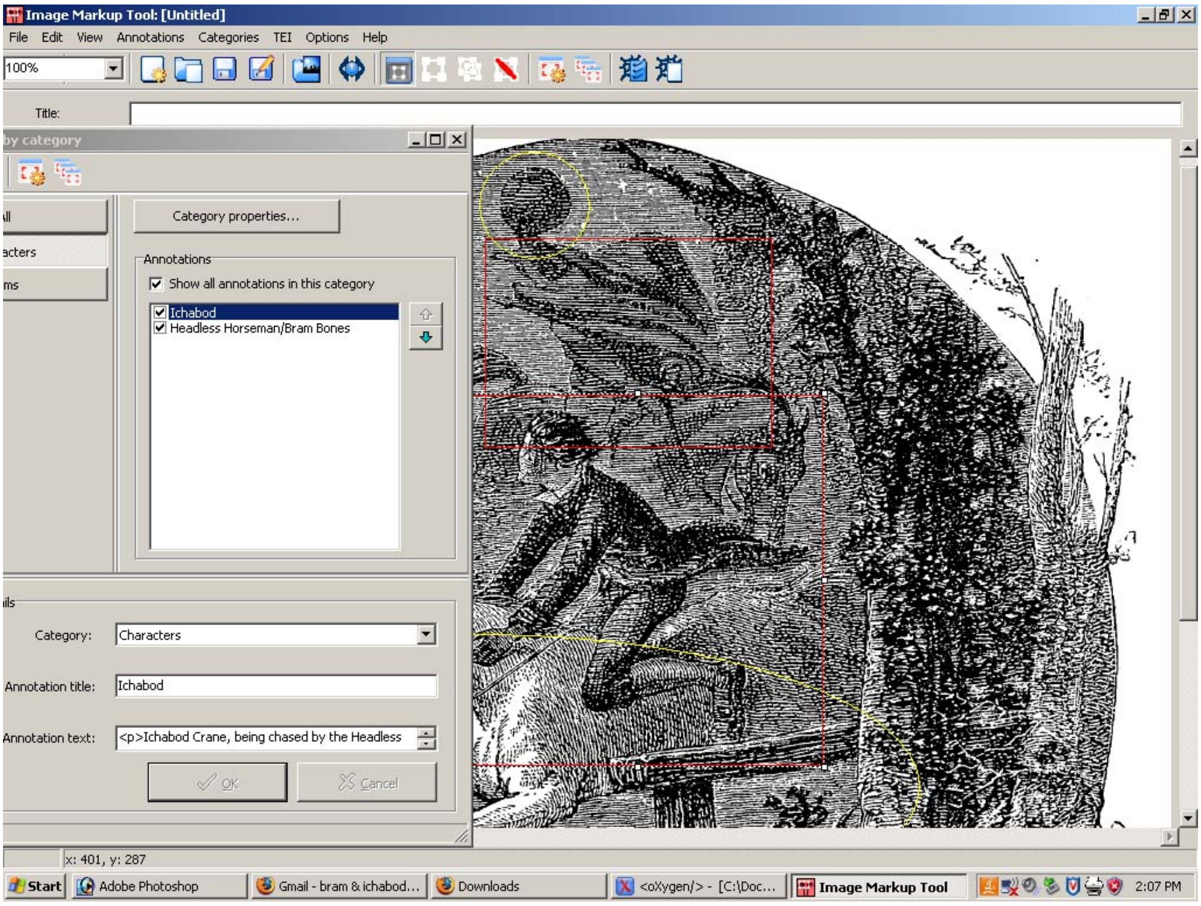


Image Markup Tool: [Untitled]

Title:

Annotations

☑ Show all annotations in this category

☑ Ichabod
☑ Headless Horseman/Bram Bones

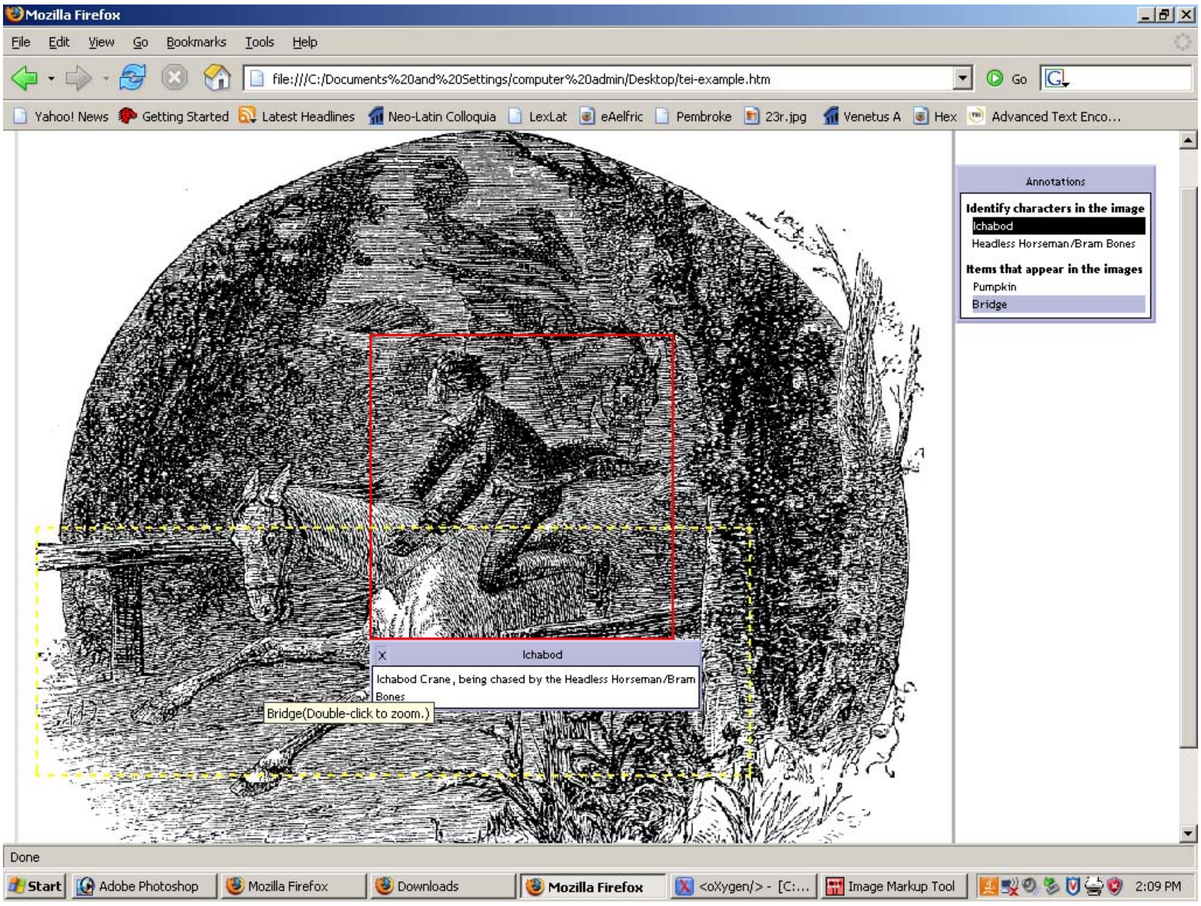
Category: Characters

Annotation title: Ichabod

Annotation text: <p>Ichabod Crane, being chased by the Headless

x: 401, y: 287

4.14 UVic IMT screenshot 2



Ichabod Crane, being chased by the Headless Horseman/Bram Bones

4.15 What is missing, or hard, in the TEI editing world

- Only a few editors like oXygen9 or XMetaL which combine visual feedback with code editing
- Visual, or WYSIWYG, editors embedded in web applications (eg in a CMS); most web editors are for XHTML (cf Google Docs)
- Reliable conversion to and from Word and OpenOffice styles. Note:
  - the general inability of word-processors to nest inline inside inline, or block inside block
  - the difficulty of extrapolating a hierarchical structure from a sequence of free-standing headings at assorted levels
  - the tedious programming required to trace the ancestry of styles in Word and OO
  - the lack of a facility in OO to stop the user formatting by hand