

# Customizing and Documenting your TEI Encoding with an ODD

Floriane Chiffoleau  
TEI Summer School, Oslo,  
September 22nd-26th, 2025

# Summary

**01**

**Customizing : Why? When? How?**

**02**

**Examples of customization**

**03**

**Documenting: Why? When? How?**

**04**

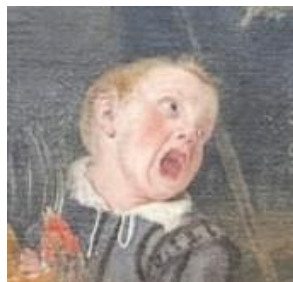
**Examples of documentation**



**Customizing your  
encoding is important  
because the TEI can be...**

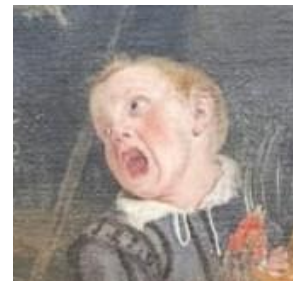
## ... too large

There are more than **500 elements** but realistically, only **half of them** are really **used**



## ... not precise enough

There are **no rule** on the **values** of **attributes**, which can lead to **mistakes**



## ... too exhaustive

Some **elements** have the **same purpose** and **similar rules** of use



# Building your encoding scheme

- ❑ Declare the tag set for its encoding
- ❑ Four types of TEI class declarations:
  - ❑ The class is not used at all → no declaration and all its elements will be excluded from the tree
  - ❑ The class is used a little → declaration and inclusion of used elements with `@include`
  - ❑ The class is used extensively → declaration and exclusion of unused elements with `@except`
  - ❑ The class is used entirely → declaration and no exclusion
- ❑ Possibility of customizing the TEI schema
  - ❑ Addition/Deletion/Change of elements
  - ❑ Customization of attributes and their values

```

<schemaSpec ident="myTEI" docLang="en" prefix="tei_" xml:lang="en">
  <moduleRef key="core"
    except="email gloss headItem headLabel meeting postBox said"/>
  <moduleRef key="tei" except=""/>
  <moduleRef key="header"
    except="appInfo application cRefPattern catDesc category classCode classD
  <moduleRef key="textstructure"
    except="div1 div2 div3 div4 div5 div6 div7 docAuthor docEdition docImprin
  <moduleRef key="namesdates" except=""/>
  <moduleRef key="transcr"
    except="am damage damagespan ex line mod path redo restore retrace secl s
  <moduleRef key="msdescription"
    except="adminInfo colophon custEvent custodialHist depth dim filiation fi
  <moduleRef key="linking" include="ab anchor link linkGrp seg"/>
  <moduleRef key="figures" except="formula notatedMusic"/>
  <moduleRef key="spoken" include="kinesic vocal"/>
  <moduleRef key="drama" include="castItem castList role roleDesc"/>
  <moduleRef key="gaiji" include="g"/>
  <moduleRef key="analysis" include="pc"/>
  <elementSpec ident="msIdentifier" mode="change">

```

## Example of TEI declarations

```
<elementSpec ident="del" mode="change">
  <attList>
    <attDef ident="rend" mode="change">
      <valList type="semi" mode="replace">
        <valItem ident="overwritten">
          <desc>a correction has been written over it (used with add
            place="across")</desc>
        </valItem>
        <valItem ident="strikethrough">
          <desc>horizontal or vertical line(s) through a
            word/expression</desc>
        </valItem>
        <valItem ident="erasure">
          <desc>indication of removed word in the text (usually
            indistinguishable)</desc>
        </valItem>
        <valItem ident="none">
          <desc>if the deletion is only marked by writing the correct
            word alongside</desc>
        </valItem>
      </valList>
    </attDef>
  </attList>
</elementSpec>
```

## Customization of an element

# Create your encoding scheme with Roma




- ❑ Go to Roma: <https://romabeta.tei-c.org/>
- ❑ Select “TEI ALL” and “START”
- ❑ Set the title to “ODD Formation TEI” and enter your name in “Author” then click on “--> CUSTOMIZE ODD”
- ❑ Sort the elements by module with “↕ by module”



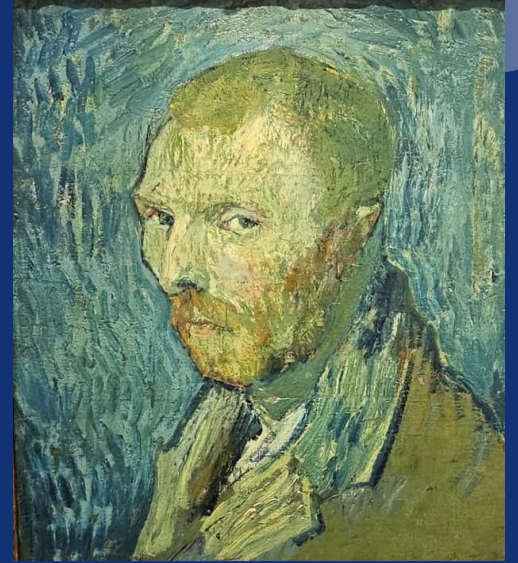
# Create your encoding scheme with Roma

- ❑ Deselect “analysis”, “certainty”, “cmc”, “corpus”, “dictionaries”, “drama”, “gaiji”, “iso-fs”, “linking”, “nets”, “spoken”, “tagdocs”, “textcrit” and “verse”
- ❑ Using the search tool, type “div” and deselect the numbered divs
- ❑ Using the search tool, type “name” and deselect the “name” element
- ❑ Using the search tool, type “del”, click on the element then “Attributes”, modify the “rend” attribute, set the “Closed” option and add the values “erasure”, “overwritten” and “strikethrough”
- ❑ Click “Download” and select “Customization as ODD”
- ❑ Change the file extension to “.xml”

# Apply your ODD to an XML-TEI file

- ❑ Open the newly created ODD and transform it into a Relax NG schema (which is a file type used to validate TEI compliance).
  - ❑ Click the "Apply Transformation Scenarios" button. 
  - ❑ Select "TEI ODD to RELAX NG XML."
  - ❑ Click the "Apply" button.
- ❑ Open your own XML file 
  - ❑ Click the "Associate Schema" button.
  - ❑ In "URL," use the "Browse" button and select the newly created RNG file located in the "out" folder. 
  - ❑ In "Schema Type," select "RELAX NG XML Syntax," then click "OK."
- ❑ Report any errors that appear and suggest changes as needed to make the file compliant.

**Customization  
without documenting  
the choices is a no-no**





# Why document your encoding scheme?

- ❑ Why document?
  - ❑ Specify and explain the choices made,
  - ❑ Ensure the consistency of your encoding
  - ❑ Help with the reuse of the encoding type used
- ❑ Two types of documentation:
  - ❑ Partial documentation, presenting elements which use may differ from those defined by the TEI guidelines
  - ❑ Exhaustive documentation, presenting the tree from A to Z, with a description of the use of elements/attributes in the tree

# How to document your encoding scheme?

- ❑ Documentation written with traditional prose tags in the <body>:
  - ❑ <div>: for the part hierarchy
  - ❑ <head>: for the part titles
  - ❑ <p>: for the presentation content and descriptions
- ❑ Specific tags for TEI structure components:
  - ❑ <gi>: for elements
  - ❑ <att>: for attributes
  - ❑ <val>: for attribute values
  - ❑ <egXML xmlns="http://www.tei-c.org/ns/Examples">: for examples from the XML tree

```

<head>Changes in the manuscript: <gi>subst</gi>,
    <gi>del</gi>/<gi>add</gi></head>
<p>Our documents are mostly handwritten manuscripts, so they all contain
deletions and additions. For these—apart from transpositions (see
below)—, we use the element couple of <gi>del</gi> and <gi>add</gi>.
If the deletion or addition is made by a different hand than the
surrounding text, then these elements are used with the attribute
    <att>hand</att> pointing to the corresponding <gi>handNote</gi>
element in the header: <egXML
    xmlns="http://www.tei-c.org/ns/Examples">
    <del rend="strikethrough" hand="#red">Werd ich denn hier nun
        endlich etwas finden?</del> Ich war <add place="above"
            hand="#minor_hand">da</add> zu fremd, um von alle dem Vieles
        zu verstehen,</egXML></p>
<p>For the <att>rend</att> attribute of <gi>del</gi>, the following
values are recommended (other values may be used if appropriate): <list>
    <item><val>strikethrough</val></item>
    <item><val>overwritten</val></item>
    <item><val>erasure</val></item>
    <item><val>none</val> (e. g. if the deletion is only marked by
        writing the correct word alongside)</item>
</list></p>

```

## Documenting the use of the <del> tag

```

<head>Logical text divisions: <gi>div</gi>, <gi>head</gi>,
    <gi>trailer</gi></head>
<p>For logical text divisions (sections, subsections and so on), we used
    nested <gi>div</gi> elements. We do not use the numbered division
    elements <gi>div1</gi>, <gi>div2</gi> and so on. The first
    <gi>div</gi> of our document will always have an
    <att>type</att>=<val>transcription</val>, but when
    appropriate, the subsequent <gi>div</gi> may have the attributes
    <att>type</att> specifying the type of division and <att>n</att>
    specifying the number of the division. Heading and trailing elements
    of the section as appearing in the manuscript are given inside
    <gi>head</gi> and <gi>trailer</gi>, respectively. These elements
    always belong only to the <gi>div</gi> of which they are direct
    sub-elements. For example, if a correspondence is giving a heading
    with "Lettre n°" + number of the letter", this can be encoded in a
    <gi>head</gi> element, as such : <egXML
xmlns="http://www.tei-c.org/ns/Examples">
<div type="transcription">
    <div type="letter">
        <pb facs=".JPG" n="1"/>
        <head rend="center underline">LETTRE N° 231.</head> ...
    </div>
</div></egXML></p>

```

## Documenting the choice of structure with <div>



# Document your encoding scheme

## Exercise:

In the ODD you generated with Roma and using the information provided earlier about writing ODD documentation, write a paragraph explaining your choice to use the `<persName>`, `<placeName>`, and `<orgName>` elements for named entities, rather than `<name>` with an `@type` attribute containing the values “person”, “place”, or “organization”.