



Modelling humanities data with TEI-XML

SCHOLARLY EDITING AND MANUSCRIPT CATALOGUING IN THE DIGITAL AGE

Dr Katarzyna Anna Kapitan
27 November 2024

Midterm General Feedback

- ▶ **Read the instruction and follow the guidelines!**
 - ▶ Prepare a short description of the materials and justification why you want to work with them and how you think using XML for data structurisation may benefit your project (250-300 words) . – **How many criteria are in this point?**
 - ▶ Using the TEI Guidelines, make a *list* of elements, attributes, and attribute values that you think you will need for your project. – **How many criteria are in this point?**
 - ▶ A plain XML file template with structured *list* of elements, attributes, and attribute values defined with DTD. – **How many files you need to submit?**

Midterm General Feedback

- ▶ **Read the instruction and follow the guidelines!**
 - ▶ Prepare a short **description of the materials (1)** and **justification why (2)** you want to work with them and **how (3)** you think using XML for data structurisation may benefit your project (**250-300 words**) (**4**)
 - ▶ Using the TEI Guidelines, make a list of **elements (1)**, **attributes (2)**, and **attribute values (3)** that you think you will need for your project.
 - ▶ A **plain XML file template (1)** with structured list of elements, attributes, and attribute values **defined with DTD (2)**.

Midterm General Feedback

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="http://www.tei-
c.org/release/xml/tei/custom/schema/relaxng/tei_ms.rng
" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-model href="http://www.tei-
c.org/release/xml/tei/custom/schema/relaxng/tei_ms.rng
" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
<!DOCTYPE TEI SYSTEM "myDTD.dtd">
<TEI xmlns="http://www.tei-c.org/ns/1.0">
```



Midterm General Feedback

- ▶ <?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE XML SYSTEM "myDTD.dtd">
<XML>
- ▶ <?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE TEI SYSTEM "myDTD.dtd">
<TEI xmlns="http://www.tei-c.org/ns/1.0">



Midterm General Feedback

- ▶ <TEI xmlns="http://www.tei-c.org/ns/1.0">
- ▶ What is xmlns, what is its value, and how is it used?

Namespace in XML

- ▶ Naming conflict in XML can be avoided by using a **name prefix**. When using prefixes in XML, a **namespace** for the prefix must be defined. The namespace declaration has the following syntax. **xmlns:prefix="URI"**
 - ▶ Example: <root **xmlns:w=http://www.w3.org/TR/html4/**
xmlns:k="https://www.kakapitan.com/myEncoding">
- ▶ Defining a default namespace for an element saves us from using prefixes in all the child elements. It has the following syntax:
xmlns="namespaceURI"
 - ▶ Example: <TEI **xmlns="http://www.tei-c.org/ns/1.0"**>
Read more: https://www.w3schools.com/xml/xml_namespaces.asp

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <?oxygen RNGSchema="http://www.menota.org/menotaP5.rng" type="xml" ?>
3  <!DOCTYPE TEI [
4  <!ENTITY % Menota_entities SYSTEM "http://www.menota.org/menota-entities.txt">
5  %Menota_entities;
6  ]>
7 <TEI xmlns="http://www.tei-c.org/ns/1.0" xmlns:me="http://www.menota.org/ns/1.0">
8   <teiHeader xml:lang="eng">
9    <fileDesc>
10   <titleStmt>
```

```
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:tei="http://www.tei-c.org/ns/1.0"
  xmlns:me="http://www.menota.org/ns/1.0"
  xmlns:xhtml="http://www.w3.org/1999/xhtml" version="2.0" exclude-result-prefixes="#all"
  xpath-default-namespace="http://www.tei-c.org/ns/1.0">
```

Final Assignment: Teacher- Students Contract



**Your files must be
well-formatted**

(File with
formatting
errors = 0
points)



**Your files must
validate against
the correct
schema**

(File validating
against a
wrong schema
or not
validating = 0
points)

Final Assignment: Teacher- Students Contract

You cannot claim to use the TEI name space and add new elements that are not part of TEI.

You cannot exclusively use elements, attributes, and attribute values introduced in class (Weeks 3-7).

You cannot exclusively use customisation and constrain rules introduced in class (Weeks 9-10).

You must demonstrate that you are familiar with the TEI Guidelines and made conscious choices which elements to use and why.

XPath Recap - Questions

Open the following files in your Oxygen

- ▶ Munich_Bayerische_Staatsbibliothek_ClM_305_transcrSpoiler.xml
(Week6/Ex1)
- ▶ Dares_editionSpoiler.xml **(Week7/Ex1)**
- ▶ Ex1_Paris_BnF_Latin_5691_Description_NKY.xml **(Week5/Ex1)**
- ▶ Ex2_FirstFolio_Spoiler_Bod.xml **(Week3/Exercises)**

XPATH QUIZ WITH YOUR OWN QUESTIONS

<https://tinyurl.com/XPath2024>

More XPath

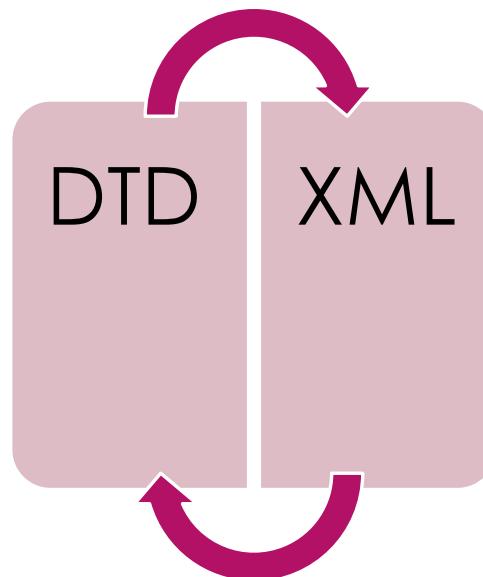
- ▶ **Recommended Reading to learn more about XPath:** Elisa Beshero-Bondar and David J. Birnbaum, 'The XPath functions we use most' & 'Autotagging with Regular Expressions (Regex)', in 'Code the X-files using the XML family of languages' Course pack, DHSI 2024:
https://ebeshero.github.io/UpTransformation/coursepack/XPath_coursepak.pdf

Customisation and Documentation

Recap DTD

DTD

- Define a fixed set of elements, attributes and attribute values needed to encode information in **XML**
- Define relationship between these elements



XML

- Encode information in a consistent and structured way
- Validate against **DTD**

Other Schema Languages

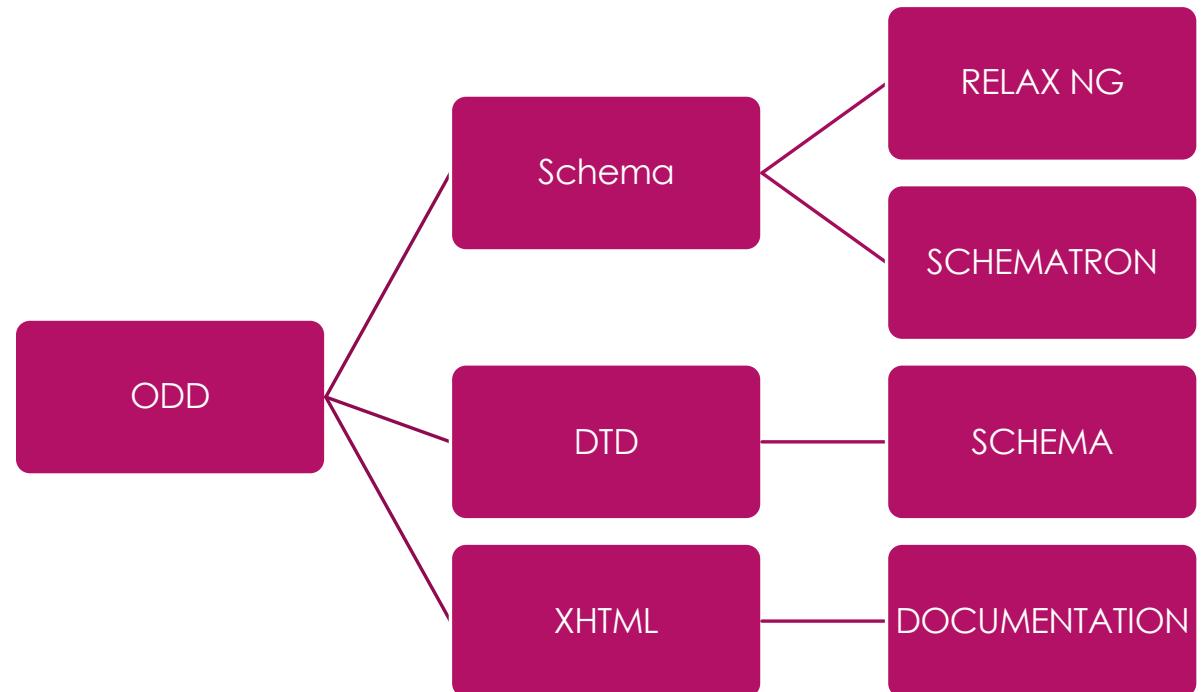
► Source: Syd Bauman, « A TEI customization for writing TEI customizations », Journal of the Text Encoding Initiative [Online], Issue 12 | July 2019 -, Online since 15 November 2019. URL: <http://journals.openedition.org/jtei/2573>

Table 1. Some of the schema languages for XML documents, arranged roughly by family of language.

SGML Document Type Declaration Family	W3C Schema Language Family	Regular Expression Family	Others
DTD	XML-Data	RELAX	DSD
XDTD	XDR	XDUCE	<i>Schematron</i>
DTD++	DCD	TREX	Examplatron
DTD++ 2.0	SOX	<i>RELAX NG</i>	X-definition
	DDML		<i>TEI ODD</i>
	XSD		

TEI ODD (One Document Does it All)

- ▶ Read:
 - ▶ Bauman, Syd. 'Freedom to Constrain: where does attribute constraint come from, mommy?',
<https://www.balisage.net/Proceedings/vol1/html/Bauman01/BalisageVol1-Bauman01.html>
 - ▶ Syd Bauman, 'A TEI customization for writing TEI customizations',
<http://journals.openedition.org/jtei/2573>



[next](#) [first](#)

Writing ODD files by hand

Let's bear in mind: An ODD file is just a specialized TEI file, with elements that say things like:

- “Include this element!”
- “Delete this element!”
- “Change this attribute!”
- “Replace the default definition with my custom definition!”
- ... and of course “Here’s what it means and why I did it!”



Introduction to Writing ODDs, slide 1 of 24

© 2010 Syd Bauman, Julia Flanders, and the Women Writers Project This TEI-encoded XML file is available under the terms of the Creative Commons Attribution-ShareAlike 3.0 (Unported) license.



```
<TEI xml:lang="en">
  <teiHeader> Metadata </teiHeader>
  <text>
    <body>
      <p> Description of the schema </p>

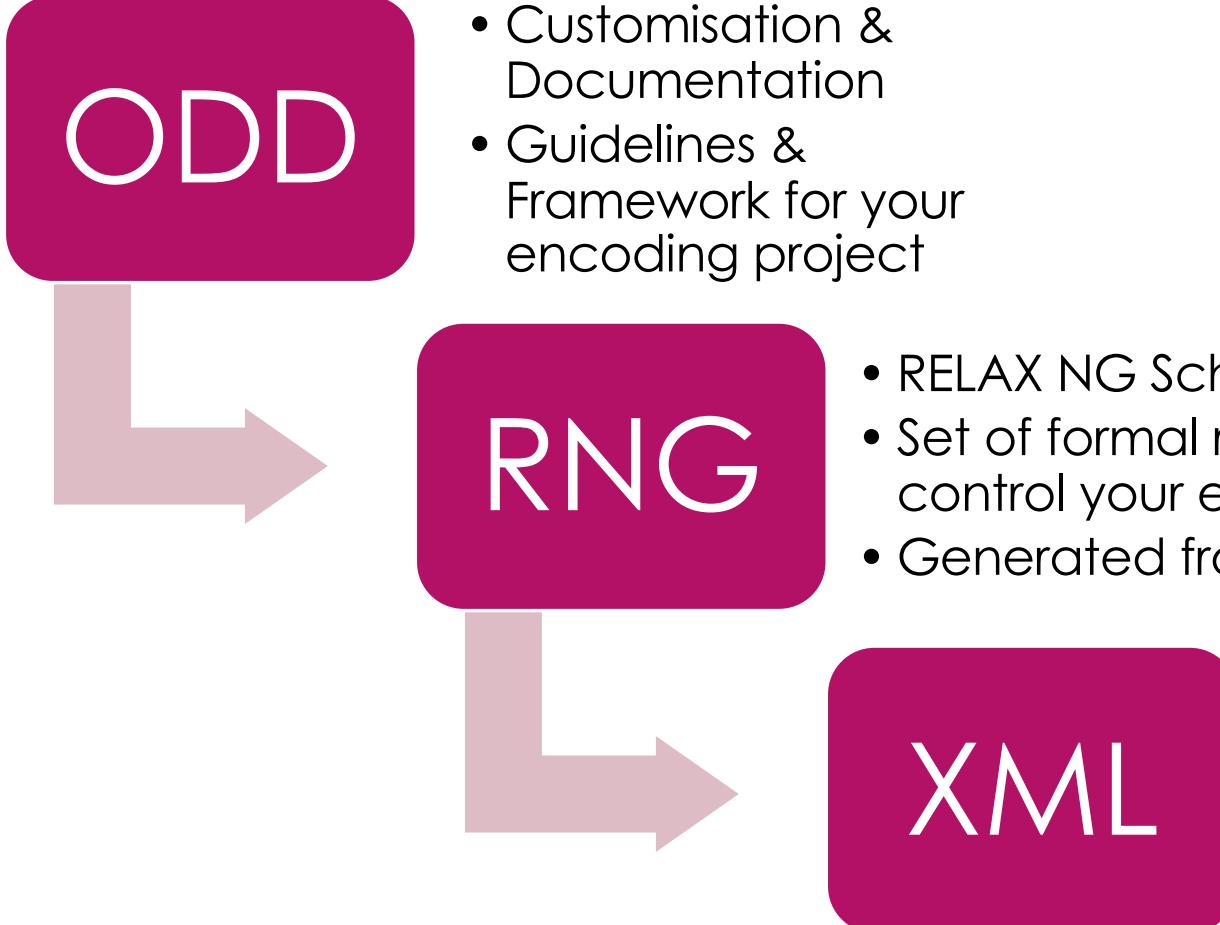
      <b><schemaSpec ident="test" targetLang="en"></b>
          Actual specification of your customization
      </schemaSpec>

    </body>
  </text>
</TEI>
```

Description and Documentation

- **<desc>** - (description) contains a short description of the purpose, function, or use of its parent element
- **<gloss>** - (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.
- **<gi>** - (generic identifier) contains the name of an element within running text
- **<att>** - (attribute) contains the name of an attribute within running text
- **<val>** - (value) contains a single attribute value.

```
<attDef ident="columns">
  <gloss xml:lang="en"
    versionDate="2007-06-12">columns</gloss>
  <desc versionDate="2005-01-14"
    xml:lang="en">specifies the number of columns per page</desc>
  <datatype minOccurs="1" maxOccurs="2">
    <dataRef key="teidata.count"/>
  </datatype>
  <remarks xml:lang="en"
    versionDate="2017-07-09">
    <p>If a single number is given, all pages referenced
      have this number of columns. If two numbers are given,
      the number of columns per page varies between the
      values supplied. Where <att>columns</att> is omitted
      the number is assumed to be <val>1</val>. </p>
  </remarks>
</attDef>
```



ODD

- Customisation & Documentation
- Guidelines & Framework for your encoding project

RNG

- RELAX NG Schema
- Set of formal rules to control your encoding
- Generated from ODD

XML

- Your valid TEI-XML document with your consistently encoded data

moduleRef

```
<moduleRef key="core"/>
<moduleRef key="tei"/>
<moduleRef key="header"/>
<moduleRef key="textstructure"/>
```

The screenshot shows the TEI Guidelines website. At the top, there's a dark purple header with the word "moduleRef" in white. Below it is a white navigation bar with the TEI logo, a search bar, and a "Search" button. The main content area has a blue header "TEI: Guidelines for Electronic Text Encoding and Interchange" and a sub-header "P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a". On the right, there's a sidebar with links to "Home" and "C Elements". The main content area contains a table for the <moduleRef> element. The table has two columns: "Module" and "tagdocs — Documentation Elements". The "Attributes" row lists several attributes with their descriptions:

Module	tagdocs — Documentation Elements
Attributes	<ul style="list-style-type: none">• att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:space<ul style="list-style-type: none">◦ att.global.rendition: @rend, @style, @rendition◦ att.global.linking: @corresp, @synch, @sameAs, @copyOf, @next, @prev, @exclude, @select◦ att.global.analytic: @ana◦ att.global.facs: @facs◦ att.global.change: @change◦ att.global.responsibility: @cert, @resp◦ att.global.source: @source

Source: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-moduleRef.html>

TEI Modules

► Source: <https://www.tei-c.org/release/doc/tei-p5-doc/fr/html/ST.html>

Module name	Formal public identifier	Where defined
analysis	Analysis and Interpretation	18 Simple Analytic Mechanisms
certainty	Certainty and Uncertainty	22 Certainty, Precision, and Responsibility
core	Common Core	3 Elements Available in All TEI Documents
corpus	Metadata for Language Corpora	16 Language Corpora
dictionaries	Print Dictionaries	10 Dictionaries
drama	Performance Texts	7 Performance Texts
figures	Tables, Formulae, Figures	15 Tables, Formulae, Graphics, and Notated Music
gaiji	Character and Glyph Documentation	5 Characters, Glyphs, and Writing Modes
header	Common Metadata	2 The TEI Header
iso-fs	Feature Structures	19 Feature Structures
linking	Linking, Segmentation, and Alignment	17 Linking, Segmentation, and Alignment
msdescription	Manuscript Description	11 Manuscript Description
namesdates	Names, Dates, People, and Places	14 Names, Dates, People, and Places
nets	Graphs, Networks, and Trees	20 Graphs, Networks, and Trees

<div>

Module: textstructure

Home
C Elements

<div> (text division) contains a subdivision of the front, body, or back of a text. [See [Divisions of the Body](#)]

Module	textstructure — Default Text Structure
Attributes	<ul style="list-style-type: none">• att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:space<ul style="list-style-type: none">◦ att.global.rendition: @rend, @style, @rendition◦ att.global.linking: @corresp, @synch, @sameAs, @copyOf, @next, @prev, @exclude, @select◦ att.global.analytic: @ana◦ att.global.facs: @facs◦ att.global.change: @change◦ att.global.responsibility: @cert, @resp◦ att.global.source: @source• att.divLike: @org, @sample<ul style="list-style-type: none">◦ att.metrical: @met, @real, @rhyme◦ att.fragmentable: @part• att.typed: @type, @subtype• att.declaring: @decls• att.written: @hand
Member of	model.divLike

Source: <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-div.html>

Module name	Formal public identifier	Where defined
analysis	Analysis and Interpretation	18 Simple Analytic Mechanisms
certainty	Certainty and Uncertainty	22 Certainty, Precision, and Responsibility
core	Common Core	3 Elements Available in All TEI Documents
corpus	Metadata for Language Corpora	16 Language Corpora
dictionaries	Print Dictionaries	10 Dictionaries
drama	Performance Texts	7 Performance Texts
figures	Tables, Formulae, Figures	15 Tables, Formulae, Graphics, and Notated Music
gaiji	Character and Glyph Documentation	5 Characters, Glyphs, and Writing Modes
header	Common Metadata	2 The TEI Header
iso-fs	Feature Structures	19 Feature Structures
linking	Linking, Segmentation, and Alignment	17 Linking, Segmentation, and Alignment
msdescription	Manuscript Description	11 Manuscript Description
namesdates	Names, Dates, People, and Places	14 Names, Dates, People, and Places
nets	Graphs, Networks, and Trees	20 Graphs, Networks, and Trees
spoken	Transcribed Speech	8 Transcriptions of Speech
tagdocs	Documentation Elements	23 Documentation Elements
tei	TEI Infrastructure	1 The TEI Infrastructure
textcrit	Text Criticism	13 Critical Apparatus
textstructure	Default Text Structure	4 Default Text Structure
transcr	Transcription of Primary Sources	12 Representation of Primary Sources
verse	Verse	6 Verse

Chapter 1: The TEI Infrastructure (<https://tei-c.org/release/doc/tei-p5-doc/en/html/ST.html>)

What else is in the textstructure module?

TEI: Guidelines for Electronic Text Encoding and Interchange

P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

Table of contents

- 4.1 Divisions of the Body
- 4.2 Elements Common to All Divisions
- 4.3 Grouped and Floating Texts
- 4.4 Virtual Divisions
- 4.5 Front Matter
- 4.6 Title Pages
- 4.7 Back Matter
- 4.8 Module for Default Text Structure

- « 3 Elements Available in All TEI Documents
» 5 Characters, Glyphs, and Writing Modes
[Home](#)

4 Default Text Structure

This chapter describes the default high-level structure for TEI documents. A full TEI document combines metadata describing it, represented by a [teiHeader](#) element, with the document itself, represented by one or more [text](#) elements or other elements taken from the [model.resource](#) class. That is, the [TEI](#) element is used to group together metadata about an encoded resource (in [teiHeader](#), specified by the [header](#) module, which is fully described in chapter [2 The TEI Header](#)) with an encoded resource. Possible encoded resources are

- a logical transcription of a source document in a [text](#) element; the [text](#) element is specified along with its high-level constituents in the [textstructure](#) module and described in the remainder of the current chapter
- a diplomatic transcription of a source document in a [sourceDoc](#) element, which is specified in the [transcr](#) module and described in chapter [12 Representation of Primary Sources](#)
- an encoded representation of a text-bearing object as images in a [facsimile](#) element, which is also specified in the [transcr](#) module and described in chapter [12 Representation of Primary Sources](#)
- a collection of contextual information or annotations that provides more detail about another encoded resource (whether in the same or a different TEI document) in a [standOff](#) element, which is specified in the [linking](#) module and described in section [17.10 The standOff Container](#)
- a feature system declaration which can be used to declare the use of [fs](#) elements in the rest of the document, which is specified in the [iso-fs](#) module and described in section [19.11 Feature System Declaration](#)

Chapter 4: Default Text Structure (<https://tei-c.org/release/doc/tei-p5-doc/en/html/DS.html>)

Because the [TEI](#) can be a child of itself, a set of collection of documents may be represented by an outermost [TEI](#) element that contains a [teiHeader](#) with metadata that is applicable to the entire set of collection of transcriptions, and then a complete [TEI](#) element for each document in the collection or set; each of these [TEI](#) elements contains a [teiHeader](#) with metadata that is applicable to the individual document, and one or more [text](#) or other elements taken from the [model.resource](#) class.

Kapitan, Modelling humanities data with TEI-XML

Customisation: Elements

[next](#) [prev](#) [first](#)

Trimming your modules

To delete specific elements from the modules you selected:

```
<moduleRef key="namesdates" except="addName affiliation bloc climate"/>
```

To include specific elements from the modules you selected:

```
<moduleRef key="textstructure" include="TEI text div back body"/>
```

Two important tips:

- you can't use both at once for the same <moduleRef>!
- be careful that the elements listed are actually in the module specified!



Introduction to Writing ODDs, slide 4 of 24

© 2010 Syd Bauman, Julia Flanders, and the Women Writers Project This TEI-encoded XML file is available under the terms of the Creative Commons Attribution-ShareAlike 3.0 (Unported) license.

Customizing your elements: `elementSpec`

► `<elementSpec`

`ident="Element_name"`

`mode="mode_of_change"`

`module="module_containing_this_element">`

My specification

`</elementSpec>`

Mode = "delete" – delete this element

Mode = "add" – add this new element

Mode = "change" – change only these parts I am specifying

Mode = "replace" – replace everything with my specification

Customizing your elements: Description

```
<elementSpec ident="publicationStmt" mode="replace" module="header">
```

```
    <desc>
```

contains info about publication of this file.

In our project `<gi>publicationStmt</gi>` must contain a sequence of two elements: `<gi>authority</gi>` and `<gi>availability</gi>` as children.

No other elements are allowed.

```
    </desc> [...]
```

```
</elementSpec>
```

Customizing your elements: Content

```
<elementSpec ident="availability" mode="change" module="header">  
    <content>  
        <textNode/>  
    </content>  
</elementSpec>
```

Customizing your elements: Content

```
<content>
    <sequence>
        <elementRef key="authority" minOccurs="1" maxOccurs="1"/>
        <elementRef key="availability" minOccurs="1" maxOccurs="1"/>
    </sequence>
</content>
```

Exercise 1

- ▶ Using the ODD file called **minimal_odd_starter_elements.odd** and pre-defined TEI transformation scenarios in Oxygen
 - ▶ generate XHTML guidelines (TEI ODD TO XHTML)
 - ▶ generate RELAX NG schema (TEI ODD TO RNG)
- ▶ Associate the RELAX NG schema with your XML file **minimal_encoding_elements.xml** and validate the XML file
 - ▶ Follow the steps described in:
Step_by_step_validation_in_Oxygen_XML_Editor.pdf
- ▶ Fix the XML file (**minimal_encoding_elements.xml**) so it validates correctly with the RNG schema created from our ODD (**minimal_odd_starter_elements.odd**)

Customisation: Attributes

Changing attributes: Overview

Several ways for elements to possess attributes:

- By being a member of an attribute class
- By having attributes privately

Hence, several ways to remove attributes from elements:

- By deleting the attribute class from the schema altogether
- By deleting the attribute from the attribute class
- By unsubscribing the element from that attribute class
- By removing a privately held attribute from the element that owns it





TEI: Guidelines for Electronic Text Encoding and Interchange

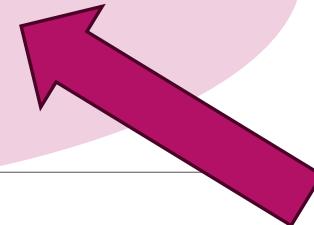
P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

<div>

Home
C Elements

<div> (text division) contains a subdivision of the front, body, or back of a text. [4.1 Divisions of the Body]

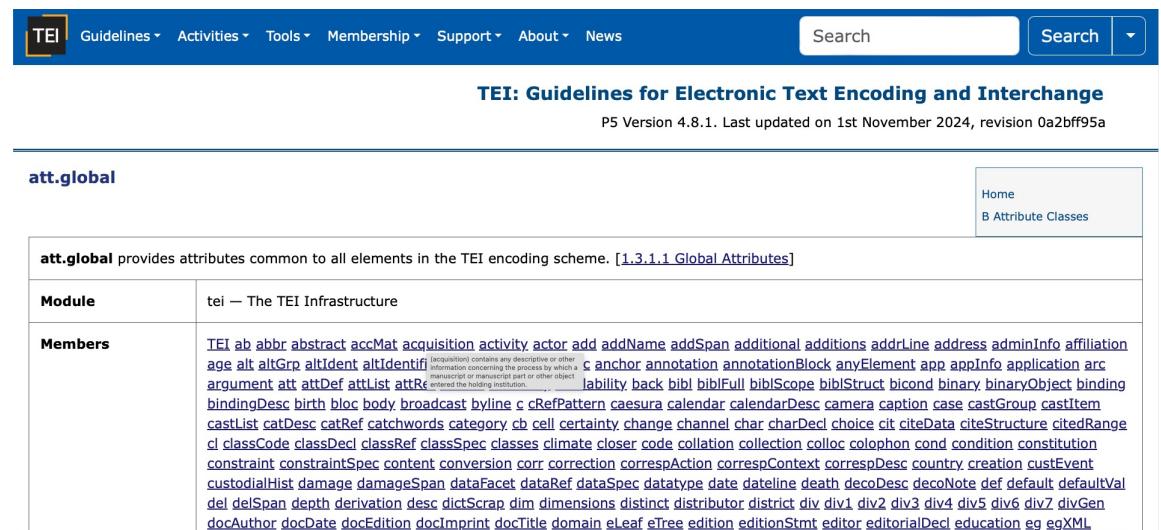
Module	textstructure — Default Text Structure
Attributes	<ul style="list-style-type: none">• att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:space<ul style="list-style-type: none">◦ att.global.rendition: @rend, @style, @rendition◦ att.global.linking: @corresp, @synch, @sameAs, @copyOf, @next, @prev, @exclude, @select◦ att.global.analytic: @ana◦ att.global.facs: @facs◦ att.global.change: @change◦ att.global.responsibility: @cert, @resp◦ att.global.source: @source• att.divLike: @org, @sample<ul style="list-style-type: none">◦ att.metrical: @met, @real, @rhyme◦ att.fragmentable: @part• att.typed: @type, @subtype• att.declaring: @decls• att.written: @hand
Member of	model.divLike



Attributes
grouped by attribute classes

Attribute classes

- ▶ A group of attributes that can occur in the same place in a TEI document.
- ▶ The names of attribute classes all start with **att.**, followed by a name that gives an indication of the group of attributes it contains.
- ▶ **@xml:id** and **@n** are defined in the attribute class **att.global**, as the name indicates, **att.global** defines global attributes
- ▶ Source:
<https://teibyexample.org/exist/tutorials/TBED08v00.htm>



The screenshot shows the TEI Guidelines website with a blue header bar containing the TEI logo and links for Guidelines, Activities, Tools, Membership, Support, About, and News. There are two search bars at the top right. Below the header, the page title is "TEI: Guidelines for Electronic Text Encoding and Interchange" with a subtitle "P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a". The main content area has a header "att.global". On the right, there's a sidebar with "Home" and "B Attribute Classes". The main content table has two columns: "Module" (containing "tei — The TEI Infrastructure") and "Members" (containing a long list of attribute names: ab, abbr, abstract, accMat, acquisition, activity, actor, add, addName, addSpan, additional, additions, addrLine, address, adminInfo, affiliation, age, alt, altGrp, altIdent, altIdentif, anchor, annotation, annotationBlock, anyElement, app, appInfo, application, arc, argument, att, attDef, attList, attRe, back, bibl, biblFull, biblScope, biblStruct, bicond, binary, binaryObject, binding, bindingDesc, birth, blog, body, broadcast, byline, c, cRefPattern, caesura, calendar, calendarDesc, camera, caption, case, castGroup, castItem, castList, catDesc, catRef, catchwords, category, cb, cell, certainty, change, channel, char, charDecl, choice, cit, citeData, citeStructure, citedRange, cl, classCode, classDecl, classRef, classSpec, classes, climate, closer, code, collation, collection, colloc, colophon, cond, condition, constitution, constraint, constraintSpec, content, conversion, corr, correction, correspAction, correspContext, correspDesc, country, creation, custEvent, custodialHist, damage, damageSpan, dataFacet, dataRef, dataSpec, datatype, date, dateline, death, decoDesc, decoNote, def, default, defaultVal, del, delSpan, depth, derivation, desc, dictScrap, dim, dimensions, distinct, distributor, district, div, div1, div2, div3, div4, div5, div6, div7, divGen, docAuthor, docDate, docEdition, docImprint, docTitle, domain, eLeaf, eTree, edition, editionStmt, editor, editorialDecl, education, eg, egXML).

Source:
<https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-att.global.html>

TEI: Guidelines for Electronic Text Encoding and Interchange

P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

Table of contents

[Appendix B.1 About the Attribute Classes Appendix](#)

[« Appendix A Model Classes](#)

[» Appendix C Elements](#)

[Home](#)

Appendix B Attribute Classes

Appendix B.1 About the Attribute Classes Appendix

This appendix gives you a list of attribute classes and links to the reference pages for them. There are 84 distinctly-named attribute classes in revision [0a2bff95a](#) of TEI P5 [Version 4.8.1](#) of the TEI Guidelines.

Sorted alphabetically

[a](#) [b](#) [c](#) [d](#) [e](#) [f](#) [g](#) [h](#) [i](#) [l](#) [m](#) [n](#) [p](#) [r](#) [s](#) [t](#) [w](#) [Show all](#)

[Show by module](#)

a

[att.anchoring](#) [att.ascribed](#) [att.ascribed.directed](#)

b

[att.breaking](#)

c

[att.calendarSystem](#) [att.canonical](#) [att.citeStructurePart](#) [att.citing](#) [att.cmc](#) [att.combinable](#) [att.coordinated](#)

Source: <https://tei-c.org/release/doc/tei-p5-doc/en/html/REF-CLASSES-ATTS.html>

[next](#) [prev](#) [first](#)

Changing attributes: Examples

To delete an entire attribute class from the schema (scorched earth approach):

```
<classSpec type="atts" ident="att.personal" mode="delete"/>
```

To delete a single attribute from an attribute class (selective thinning):

```
<classRef key="att.global" except="xml:base xml:space"/>
```

To unsubscribe an element from an attribute class (gentle bureaucratic approach):

```
<elementSpec module="core" ident="pb" mode="change">
  <classes mode="change">
    <memberOf key="att.typed" mode="delete"/>
  </classes>
</elementSpec>
```

To remove a specific attribute from an element that holds it privately (repossession):

```
<elementSpec module="core" ident="note" mode="change">
  <attList>
    <attDef ident="anchored" mode="delete"/>
  </attList>
</elementSpec>
```



Changing attribute values <attDef>

```
<elementSpec ident="title" mode="change" module="core">
  <desc>This specification changes the element "title", which is part [...] </desc>
  <attList>
    <attDef ident="type" mode="replace">
      <desc>In this part we replace the specification of type attribute [...] </desc>
      <datatype>
        <dataRef key="teidata.enumerated"/>
      </datatype>
      <valList type="closed">
        <valItem ident="uniform"/>
        <valItem ident="supplied"/>
      </valList>
    </attDef></attList></elementSpec>
```

Datatype

- ▶ **<datatype>** (datatype) specifies the declared value for an attribute, by referring to any datatype defined by the chosen schema language
- ▶ In the TEI scheme, most datatypes are expressed using pre-defined TEI macros, which map a name in the form **teidata.xxxx**
- ▶ For example, **teidata.enumerated**, which defines the range of attribute values expressed as a single XML name taken from a list of documented possibilities.

TEI: Guidelines for Electronic Text Encoding and Interchange

P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

Table of contents

[Appendix E.1 About the Datatypes and Macros Appendix](#)

[« Appendix D Attributes](#)

[» Appendix F Bibliography](#)

[Home](#)

Appendix E Datatypes and Other Macros

⚓ **Appendix E.1 About the Datatypes and Macros Appendix**

This appendix gives you a list of datatypes and links to the reference pages for them. There are 35 distinctly-named data specifications in revision [0a2bff95a](#) of TEI P5 [Version 4.8.1](#) of the TEI Guidelines.

Alphabetical list

[cmc] Computer-mediated communication

[macro.specialPara.cmc](#)

[tei] Declarations for classes, datatypes, and macros available to all TEI modules

[macro.abContent](#) [macro.limitedContent](#) [macro paraContent](#) [macro phraseSeq](#) [macro phraseSeq.limited](#) [macro.specialPara](#) [macro.xtext](#) [teidata.authority](#) [teidata.certainty](#) [teidata.count](#) [teidata.duration.iso](#) [teidata.duration.w3c](#) [teidata.enumerated](#) [teidata.gender](#) [teidata.interval](#) [teidata.language](#) [teidata.name](#) [teidata.namespace](#) [teidata.namespaceOrName](#) [teidata.nullOrName](#) [teidata.numeric](#) [teidata.outputMeasurement](#) [teidata.pattern](#) [teidata.point](#) [teidata.pointer](#) [teidata.prefix](#) [teidata.probability](#) [teidata.probCert](#) [teidata.replacement](#) [teidata.sex](#) [teidata.temporal.iso](#) [teidata.temporal.w3c](#) [teidata.temporal.working](#) [teidata.text](#) [teidata.truthValue](#) [teidata.unboundedCount](#) [teidata.version](#) [teidata.versionNumber](#) [teidata.word](#) [teidata.xmlName](#) [teidata.xpath](#) [teidata.xTruthValue](#)

[↑ TEI P5 Guidelines](#) [« Appendix D Attributes](#) [» Appendix F Bibliography](#)

[\[English\]](#) [\[Deutsch\]](#) [\[Español\]](#) [\[Italiano\]](#) [\[Français\]](#) [\[日本語\]](#) [\[한국어\]](#) [\[中文\]](#)

Source: <https://tei-c.org/release/doc/tei-p5-doc/en/html/REF-MACROS.html>

Exercise 2: Attributes

- ▶ Using **minimal_encoding_attributes.xml** and **minimal_odd_starter_attributes.odd**
 - ▶ Create RELAX NG schema from ODD
 - ▶ Create XHTML guidelines from ODD.
- ▶ Associate the RNG schema with your XML file and make sure your file validates.
- ▶ Revise the ODD file so that divisions (**div**) must have the **type** attribute, and the only allowed attribute values are **book** and **chapter**.
- ▶ Save your ODD with a new name and generate a RNG and XHTML from it.
- ▶ Associate the new RNG schema with your file.
- ▶ Revise your XML file so it validates correctly with your new schema.

Exercise 3: Autogenerated ODD from XML

- ▶ Create an ODD file from **minimal_encoding_elements.xml** by using the **oddbyexample.xsl** transformation.
 - ▶ Tutorial: [Burnard_2013_How_to_Make_an_ODD_Automagically.pdf](#)
- ▶ Edit the ODD file adding prose descriptions of the customisation, examples, and comments explaining what's happening in each section of the ODD file.
- ▶ Generate RELAX NG schema and XHTML guidelines from your ODD.
- ▶ Associate the RNG schema with with your XML file (**minimal_encoding_elements.xml**), does it validate correctly?

More about elements (optional)

Model classes

TEI: Guidelines for Electronic Text Encoding and Interchange

P5 Version 4.8.1. Last updated on 1st November 2024, revision 0a2bff95a

<div>

Module: textstructure

[Home](#)
[C Elements](#)

<div> (text division) contains a subdivision of the front, body, or back of a text. [See [Divisions of the Body](#).]

Module	textstructure — Default Text Structure
Attributes	<ul style="list-style-type: none">• att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:space<ul style="list-style-type: none">◦ att.global.rendition: @rend, @style, @rendition◦ att.global.linking: @corresp, @synch, @sameAs, @copyOf, @next, @prev, @exclude, @select◦ att.global.analytic: @ana◦ att.global.facs: @facs◦ att.global.change: @change◦ att.global.responsibility: @cert, @resp◦ att.global.source: @source• att.divLike: @org, @sample<ul style="list-style-type: none">◦ att.metrical: @met, @real, @rhyme◦ att.fragmentable: @part• att.typed: @type, @subtype• att.declaring: @decls• att.written: @hand
Member of	model.divLike

Div

The only member of
model.divLike

model.divLike

model.divLike groups elements used to represent un-numbered generic structural divisions.

Module	tei — The TEI Infrastructure
Used by	back body div front lem rdg
Members	div

[[English](#)] [[Deutsch](#)] |

Source: <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-model.divLike.html>

Model classes

- ▶ A group of elements that can occur in the same place in a TEI document.
- ▶ The names of model classes all start with **model.**, followed by a name that gives an indication of the group of elements it contains.
- ▶ For example, **<p>** and **<ab>** are being grouped in the model class **model.pLike**, which holds all paragraph-like elements.
- ▶ Source:
<https://teibyexample.org/exist/tutorials/TBED08v00.htm>

model.pLike

model.pLike regroupe des éléments de type paragraphe.

Module	tei — The TEI Infrastructure
Utilisé par	abstract additional application availability back binding correspDesc custodialHist decoDesc editionStmt editoria hyphenation interpretation langKnowledge langUsage la msContents msDesc msFrag msItem msItemStruct msP persona physDesc place population post prefixDef proje refsDecl remarks samplingDecl scriptDesc scriptStmt se stdVals styleDefDecl supportDesc terrain trait transcript
Membres	ab p

Source: <https://www.tei-c.org/release/doc/tei-p5-doc/fr/html/ref-model.pLike.html>

New elements in your own name space

- ▶ Analyse the files in the folder **Ex4_optional_new_elements**

```
<!-- BEGINNING OF YOUR ELEMENT SPEC -->
<elementSpec ident="cat" ns="http://www.kakapitan.com//ns/1.0">
    <desc>This element allows me to tag all cats mentioned in my document.
        It can only contain text and has no other children elements.</desc>
    <classes>
        <memberOf key="model.nameLike.agent"/><!-- By subscribing your new element to "model.nameLike.agent" it will behave just
        like any other nameLike element -->
        <memberOf key="att.global"/><!-- By subscribing your new element to "att.global" you allow it to have all the global attributes -->
    </classes>
    <content>
        <!-- Here below you specify the content of your element, in our case, a text node. -->
        <textNode/>
    </content>
    <attList>
        <!-- Here below you specify your custom attributes -->
        <attDef ident="breed" ns="http://www.kakapitan.com//ns/1.0" usage="req">
            <desc>The attribute <att>breed</att> used on <gi>cat</gi> elements allows me to group cats by breed.
                This attribute is required (usage = "req") on all <gi>cat</gi> elements.
                The value of the <att>breed</att> must fit the datatype "teidata.word", but the values aren't pre-defined. </desc>
            <datatype minOccurs="1" maxOccurs="1">
                <dataRef key="teidata.word"/>
            </datatype>
        </attDef>
    </attList>
</example>
```