

Alice Thornton's Books
TEI Documentation: Element Set

Sharon Howard

25 February 2025

Contents

About this document	1
The element set	2
Standard TEI and <code>teiHeader</code>	2
Books	7
The text	7
Editorial notes	21
Linked metadata	22
People	22
Places	25
Events	28
Glosses	30

About this document

This document lists the TEI elements used in tagging Alice Thornton's Books (2025). It should be read in conjunction with the *Editorial Guidelines*.

The documentation and the TEI XML files to which it refers are stored in the Alice Thornton Digital Edition Github Repository.

```
alice-thornton
|--documentation
| |--atb_elements.pdf
| |--atb_guidelines.pdf
|--entities
| |--events.xml
| |--glossary.xml
| |--glossary_book_one.xml
| |--glossary_book_three.xml
| |--glossary_book_two.xml
| |--people.xml
| |--places.xml
|--texts
| |--00_book_of_remembrances
| | |--book_of_remembrances.xml
| |--01_book_one
| | |--book_one.xml
| |--02_book_two
| | |--book_two.xml
| |--03_book_three
| | |--book_three.xml
```

The element set

Standard TEI and `teiHeader`

The `<teiHeader>` element holds metadata about the marked up file.

`<TEI>`

TEI (TEI document) contains a single TEI-conformant document, combining a single TEI header with one or more members of the `model.resource` class.

attributes: `@xml:id`

`<altIdentifier>`

altIdentifier (alternative identifier) contains an alternative or former structured identifier used for a manuscript or other object, such as a former catalogue number.

`<appInfo>`

appInfo (application information) records information about an application which has edited the TEI file.

`<application>`

application () provides information about an application which has acted upon the document.

attributes: `@ident`, `@version`, `@xml:id`

`<author>`

author (author) in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority.

<calendar>

calendar (calendar) describes a calendar or dating system used in a dating formula in the text.

attributes: @xml:id

<calendarDesc>

calendarDesc (calendar description) contains a description of the calendar system used in any dating expression found in the text.

<change>

change (change) documents a change or set of changes made during the production of a source document, or during the revision of an electronic file.

<charDecl>

charDecl (character declarations) provides information about nonstandard characters and glyphs.

<collection>

collection (collection) contains the name of a collection of manuscripts or other objects, not necessarily located within a single repository.

<date>

date (date) contains a date in any format.

<desc>

desc (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

<edition>

edition (edition) describes the particularities of one edition of a text.

<editionStmt>

editionStmt (edition statement) groups information relating to one edition of a text.

<editor>

editor () contains a secondary statement of responsibility for a bibliographic item, for example the name of an individual, institution or organization, (or of several such) acting as editor, compiler, translator, etc.

<encodingDesc>

encodingDesc (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.

<fileDesc>

fileDesc (file description) contains a full bibliographic description of an electronic file.

<glyph>

glyph (character glyph) provides descriptive information about a character glyph.

attributes: @xml:id

<handDesc>

handDesc (description of hands) contains a description of all the different hands used in a manuscript or other object.

<hi>

hi (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: @rend

<idno>

idno (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

<label>

label (label) contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary.

<listChange>

listChange () groups a number of change descriptions associated with either the creation of a source text or the revision of an encoded text.

<msDesc>

msDesc (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book.

<msIdentifier>

msIdentifier (manuscript identifier) contains the information required to identify the manuscript or similar object being described.

<msName>

msName (alternative name) contains any form of unstructured alternative name used for a manuscript or other object, such as an ‘ocellus nominum’, or nickname.

<name>

name (name, proper noun) contains a proper noun or noun phrase.

<objectDesc>

objectDesc (object description) contains a description of the physical components making up the object which is being described.

<p>

p (paragraph) marks paragraphs in prose.

<physDesc>

physDesc (physical description) contains a full physical description of a manuscript, manuscript part, or other object optionally subdivided using more specialized elements from the `model.physDescPart` class.

<profileDesc>

profileDesc (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting.

<publicationStmt>

publicationStmt (publication statement) groups information concerning the publication or distribution of an electronic or other text.

<publisher>

publisher (publisher) provides the name of the organization responsible for the publication or distribution of a bibliographic item.

<ref>

ref (reference) defines a reference to another location, possibly modified by additional text or comment.

attributes: @target

<repository>

repository (repository) contains the name of a repository within which manuscripts or other objects are stored, possibly forming part of an institution.

<revisionDesc>

revisionDesc (revision description) summarizes the revision history for a file.

<sourceDesc>

sourceDesc (source description) describes the source(s) from which an electronic text was derived or generated, typically a bibliographic description in the case of a digitized text, or a phrase such as ‘born digital’ for a text which has no previous existence.

<teiHeader>

teiHeader (TEI header) supplies descriptive and declarative metadata associated with a digital resource or set of resources.

<title>

title (title) contains a title for any kind of work.

attributes: @type

<titleStmt>

titleStmt (title statement) groups information about the title of a work and those responsible for its content.

<unicodeProp>

unicodeProp (unicode property) provides a Unicode property for a character (or glyph).

attributes: @name, @value

Books

The text

The transcribed and marked-up texts of the edition are held in a **<text>** container in each TEI XML file.

<abbr>

abbr (abbreviation) contains an abbreviation of any sort.

<add>

add (addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector.

attributes: @place, @rend

- @place values:
 - above

- below
 - centre
 - inline
 - left
 - left margin
- @rend values:
 - differentInk
 - post(rule)
 - size(smaller)
 - underline

<am>

am (abbreviation marker) contains a sequence of letters or signs present in an abbreviation which are omitted or replaced in the expanded form of the abbreviation.

attributes: @rend

- @rend values:
 - italic
 - superscript

<anchor>

anchor (anchor point) attaches an identifier to a point within a text, whether or not it corresponds with a textual element.

attributes: @corresp, @n, @resp, @type, @xml:id

- @resp values:
 - ednote
- @type values:
 - event
 - letter

<back>

back (back matter) contains any appendixes, etc. following the main part of a text.

<body>

body (text body) contains the whole body of a single unitary text, excluding any front or back matter.

<choice>

choice (choice) groups a number of alternative encodings for the same point in a text.

<corr>

corr (correction) contains the correct form of a passage apparently erroneous in the copy text.

attributes: @rend

- @rend values:
 - italic

<damage>

damage (damage) contains an area of damage to the text witness.

attributes: @agent

- @agent values:
 - excised
 - tear

<date>

date (date) contains a date in any format.

attributes: @datingMethod, @notAfter-custom, @notBefore-custom, @type, @when-custom

- @datingMethod values:
 - #JulianJan
 - #JulianMar
 - #JulianSic
 - #JulianSlash

Note: @type was applied to only a few dates early in project

del (deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector.

attributes: @rend

- @rend values:
 - doublestrikethrough
 - strikethrough

<div>

div (text division) contains a subdivision of the front, body, or back of a text.

attributes: @corresp, @rend, @type, @xml:id

- @rend values:
 - differentInk
 - post(rule)
 - pre(rule)
- @type values:
 - correspondence
 - index
 - poetry
 - prayer
 - preface
 - supplementary

<docAuthor>

docAuthor (document author) contains the name of the author of the document, as given on the title page (often but not always contained in a byline).

<docTitle>

docTitle (document title) contains the title of a document, including all its constituents, as given on a title page.

<epigraph>

epigraph (epigraph) contains a quotation, anonymous or attributed, appearing at the start or end of a section or on a title page.

attributes: @rend

- @rend values:
 - post(rule)
 - pre(rule)

<ex>

ex (editorial expansion) contains a sequence of letters added by an editor or transcriber when expanding an abbreviation.

<expan>

expan (expansion) contains the expansion of an abbreviation.

attributes: @rend

- @rend values:
 - italic

<figure>

figure (figure) groups elements representing or containing graphic information such as an illustration, formula, or figure.

<front>

front (front matter) contains any prefatory matter (headers, abstracts, title page, prefaces, dedications, etc.) found at the start of a document, before the main body.

<fw>

fw (forme work) contains a running head (e.g. a header, footer), catchword, or similar material appearing on the current page.

attributes: @corresp, @place, @rend, @resp, @type, @xml:id

- @place values:
 - centre
 - left
 - left margin
 - right
 - top
- @rend values:
 - differentInk
 - italic
 - post(rule)
 - pre(rule)
 - underline
- @type values:
 - head
 - header
 - list-like
 - pageNum
 - reference

<g>

g (character or glyph) represents a glyph, or a non-standard character.

attributes: @ref

<gap>

gap (gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible.

attributes: @quantity, @reason, @unit

- @reason values:
 - damage
 - deleted
 - obliterated
 - unclear
- @unit values:
 - char
 - line
 - word

<geogName>

geogName (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai.

attributes: @n, @ref

<graphic>

graphic (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it.

attributes: @url

<head>

head (heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc.

attributes: @place, @rend

- @place values:
 - centre
 - right
- @rend values:
 - post(rule double)
 - post(rule pattern)
 - post(rule)
 - pre(rule double)
 - pre(rule pattern)
 - pre(rule)
 - underline

<hi>

hi (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: @rend

- @rend values:
 - italic
 - italic-mod-only

- place(centre)
- superscript
- underline

<item>

item (item) contains one component of a list.

attributes: @corresp, @rend

- @rend values:
 - first-line-indent

<l>

l (verse line) contains a single, possibly incomplete, line of verse.

attributes: @n

<label>

label (label) contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary.

attributes: @place, @rend

- @place values:
 - left margin
 - right bottom
 - right mid
 - right top
- @rend values:
 - underline

<lb>

lb (line beginning) marks the beginning of a new (typographic) line in some edition or version of a text.

attributes: @break

- @break values:
 - no

<lg>

lg (line group) contains one or more verse lines functioning as a formal unit, e.g. a stanza, refrain, verse paragraph, etc.

attributes: @type

Note: @type was rarely applied

<list>

list (list) contains any sequence of items organized as a list.

attributes: @rend, @type

- @rend values:
 - braced{ }
- @type values:
 - correspondence
 - documents
 - expenses
 - index
 - mixed
 - names

<measure>

measure (measure) contains a word or phrase referring to some quantity of an object or commodity, usually comprising a number, a unit, and a commodity name.

attributes: @quantity, @type, @unit

- @type values:
 - currency
- @unit values:
 - £

<metamark>

metamark () contains or describes any kind of graphic or written signal within a document the function of which is to determine how it should be read rather than forming part of the actual content of the document.

attributes: @corresp, @xml:id

<milestone>

milestone (milestone) marks a boundary point separating any kind of section of a text, typically but not necessarily indicating a point at which some part of a standard reference system changes, where the change is not represented by a structural element.

attributes: @corresp, @n, @rend, @spanTo, @unit, @xml:id

- @unit values:
 - event
 - letter
 - marker

Note: @rend encodes various decorative marks not represented in DSE

<note>

note (note) contains a note or annotation.

attributes: @place, @rend, @resp, @style, @xml:id

- @place values:
 - left
 - left margin
 - top
- @rend values:
 - post(rule)
 - pre(rule double)
 - pre(rule)

<p>

p (paragraph) marks paragraphs in prose.

attributes: @corresp, @rend, @style, @xml:id

- @rend values:
 - first-line-indent
 - modernised(append-following-sibling)
 - place(centre)
 - post(rule)
 - pre(rule short left)
 - size(larger)

<pb>

pb (page beginning) marks the beginning of a new page in a paginated document.

attributes: @n, @xml:id

<pc>

pc (punctuation character) contains a character or string of characters regarded as constituting a single punctuation mark.

attributes: @norm

<persName>

persName (personal name) contains a proper noun or proper-noun phrase referring to a person, possibly including one or more of the person's forenames, surnames, honorifics, added names, etc.

attributes: @n, @ref, @type

- @type values:
 - author
 - monogram

<placeName>

placeName (place name) contains an absolute or relative place name.

attributes: @n, @ref

<quote>

quote (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text.

attributes: @source, @type

- @type values:
 - allusion
 - direct
 - paraphrase
 - partial
 - phrase

<ref>

ref (reference) defines a reference to another location, possibly modified by additional text or comment.

attributes: @n, @source, @type

- @type values:
 - biblical
 - bibliog
 - page

<rs>

rs (referencing string) contains a general purpose name or referring string.

attributes: @n, @ref, @type

- @type values:
 - geog
 - group
 - person
 - place

<seg>

seg (arbitrary segment) represents any segmentation of text below the ‘chunk’ level.

attributes: @rend, @type, @xml:id

- @rend values:
 - braced{}
 - differentInk
 - italic
 - place(centre)
 - place(right)
 - size(larger)
 - underline
- @type values:
 - epigraph

<sic>

sic (Latin for thus or so) contains text reproduced although apparently incorrect or inaccurate.

<space>

space (space) indicates the location of a significant space in the text.

attributes: @quantity, @unit

- @unit values:

- line
- page
- word

<subst>

subst (substitution) groups one or more deletions (or surplus text) with one or more additions when the combination is to be regarded as a single intervention in the text.

<supplied>

supplied (supplied) signifies text supplied by the transcriber or editor for any reason; for example because the original cannot be read due to physical damage, or because of an obvious omission by the author or scribe.

attributes: @reason

- @reason values:

- damage
- ink mark
- missing
- mod
- obscured
- omitted
- smudge
- unclear

<surplus>

surplus (surplus) marks text present in the source which the editor believes to be superfluous or redundant.

attributes: @reason

- @reason values:
 - repetition
 - superfluous

<term>

term (term) contains a single-word, multi-word, or symbolic designation which is regarded as a technical term.

attributes: @n, @ref, @xml:id

<text>

text (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

<titlePage>

titlePage (title page) contains the title page of a text, appearing within the front or back matter.

attributes: @xml:id

<titlePart>

titlePart (title part) contains a subsection or division of the title of a work, as indicated on a title page.

<unit>

unit () contains a symbol, a word or a phrase referring to a unit of measurement in any kind of formal or informal system.

<w>

w (word) represents a grammatical (not necessarily orthographic) word.

attributes: @norm, @type, @xml:space

Note: @type was applied irregularly for project use

Editorial notes

Editorial endnote texts are held in a `<standOff>` container at the end of each Book file.

`<hi>`

`hi` (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: `@rend`

`<listAnnotation>`

`listAnnotation` () contains a list of annotations, typically encoded as `annotation`, `annotation-Block`, or `note`, possibly organized with nested `listAnnotation` elements.

`<note>`

`note` (note) contains a note or annotation.

attributes: `@n`, `@place`, `@type`, `@xml:id`

`<p>`

`p` (paragraph) marks paragraphs in prose.

`<ref>`

`ref` (reference) defines a reference to another location, possibly modified by additional text or comment.

attributes: `@target`

`<seg>`

`seg` (arbitrary segment) represents any segmentation of text below the ‘chunk’ level.

attributes: `@rend`

`<standOff>`

`standOff` () Functions as a container element for linked data, contextual information, and stand-off annotations embedded in a TEI document.

Linked metadata

People

Stored in the file people.xml.

<addName>

addName (additional name) contains an additional name component, such as a nickname, epithet, or alias, or any other descriptive phrase used within a personal name.

attributes: @type

- @type values:
 - alias
 - nickname

<anchor>

anchor (anchor point) attaches an identifier to a point within a text, whether or not it corresponds with a textual element.

attributes: @xml:id

<bibl>

bibl (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

<birth>

birth (birth) contains information about a person's birth, such as its date and place.

attributes: @type, @when-custom

- @type values:
 - baptism
 - birth

<body>

body (text body) contains the whole body of a single unitary text, excluding any front or back matter.

<death>

death (death) contains information about a person's death, such as its date and place.

attributes: @type, @when-custom

- @type values:
 - burial
 - death

<forename>

forename (forename) contains a forename, given or baptismal name.

<gender>

gender (gender) specifies the gender identity of a person, persona, or character.

<hi>

hi (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: @rend

<listBibl>

listBibl (citation list) contains a list of bibliographic citations of any kind.

<listPerson>

listPerson (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source.

<note>

note (note) contains a note or annotation.

attributes: @type

- @type values:

- bio
- marriages
- parents
- sources

<noteGrp>

noteGrp (note group) contains a group of notes

<persName>

persName (personal name) contains a proper noun or proper-noun phrase referring to a person, possibly including one or more of the person's forenames, surnames, honorifics, added names, etc.

attributes: @n, @ref, @role, @type

- @role values:
 - child
 - father
 - mother
 - spouse
- @type values:
 - label

<person>

person (person) provides information about an identifiable individual, for example a participant in a language interaction, or a person referred to in a historical source.

attributes: @xml:id

<ref>

ref (reference) defines a reference to another location, possibly modified by additional text or comment.

attributes: @target

<roleName>

roleName (role name) contains a name component which indicates that the referent has a particular role or position in society, such as an official title or rank.

attributes: @type

- @type values:
 - peerage

<seg>

seg (arbitrary segment) represents any segmentation of text below the ‘chunk’ level.

attributes: @rend

<surname>

surname (surname) contains a family (inherited) name, as opposed to a given, baptismal, or nick name.

attributes: @type

- @type values:
 - birth
 - married

<text>

text (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

Places

Stored in the file places.xml.

<anchor>

anchor (anchor point) attaches an identifier to a point within a text, whether or not it corresponds with a textual element.

attributes: @xml:id

<body>

body (text body) contains the whole body of a single unitary text, excluding any front or back matter.

<country>

country (country) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc.

<geo>

geo (geographical coordinates) contains any expression of a set of geographic coordinates, representing a point, line, or area on the surface of the earth in some notation.

<geogName>

geogName (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai.

attributes: @type

- @type values:
 - label

<hi>

hi (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: @rend

<item>

item (item) contains one component of a list.

<linkGrp>

linkGrp (link group) defines a collection of associations or hypertextual links.

<list>

list (list) contains any sequence of items organized as a list.

<listPlace>

listPlace (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.

<location>

location (location) defines the location of a place as a set of geographical coordinates, in terms of other named geo-political entities, or as an address.

<note>

note (note) contains a note or annotation.

attributes: @type

- @type values:
 - desc
 - sources

<noteGrp>

noteGrp (note group) contains a group of notes

<place>

place (place) contains data about a geographic location

attributes: @type, @xml:id

- @type values:
 - country
 - geo
 - non-country
 - region
 - settlement
 - subregion
 - unit

<placeName>

placeName (place name) contains an absolute or relative place name.

attributes: @type

- @type values:
 - label

<ptr>

ptr (pointer) defines a pointer to another location.

attributes: @target, @type

- @type values:
 - historic_england_research_records
 - irish_sites_and_monuments
 - national_heritage_list_for_england
 - victoria_county_history
 - vision_of_britain
 - website
 - wikidata
 - wikipedia

<region>

region (region) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country.

<settlement>

settlement (settlement) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit.

<text>

text (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

Events

Stored in the file events.xml.

<body>

body (text body) contains the whole body of a single unitary text, excluding any front or back matter.

<desc>

desc (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

<event>

event (event) contains data relating to anything of significance that happens in time.

attributes: @from-custom, @n, @notAfter-custom, @notBefore-custom, @to-custom, @type, @when-custom, @xml:id

- @type values:
 - group
 - sgl
 - sub

<label>

label (label) contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary.

<linkGrp>

linkGrp (link group) defines a collection of associations or hypertextual links.

attributes: @type

- @type values:
 - sgl
 - sub

<listEvent>

listEvent (list of events) contains a list of descriptions, each of which provides information about an identifiable event.

attributes: @type

- @type values:
 - events
 - groups

<ptr>

ptr (pointer) defines a pointer to another location.

attributes: @subtype, @target, @type

- @type values:
 - book
 - group

<text>

text (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

Glosses

Stored in a `glossary_(book).xml` file, one per book. Note that the file naming is slightly misleading: we ran out of time to create a consolidated glossary and every tagged instance of a term in the Books links to a single item.

<body>

body (text body) contains the whole body of a single unitary text, excluding any front or back matter.

<gloss>

gloss (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.

<hi>

hi (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

attributes: @rend

<item>

item (item) contains one component of a list.

attributes: @xml:id

<list>

list (list) contains any sequence of items organized as a list.

attributes: @type, @xml:id

<ref>

ref (reference) defines a reference to another location, possibly modified by additional text or comment.

attributes: @target

<seg>

seg (arbitrary segment) represents any segmentation of text below the ‘chunk’ level.

attributes: @rend

<text>

text (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.