

Humanités numériques : structuration des données et des documents textuels

E. ROUQUETTE

Cours 4 – 14 février 2024

La TEI : rappels

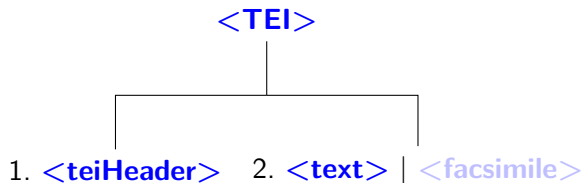
Utiliser la TEI

Pour utiliser un set de balises TEI, il faut déclarer **l'espace de nom TEI** (namespace) dans l'élément racine grâce à une **adresse URI**

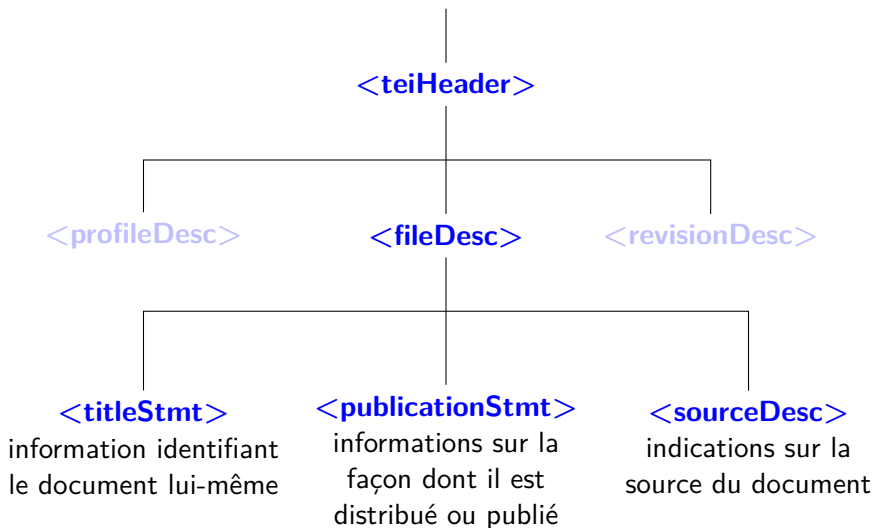
```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
```

Le logiciel oXygen le fait automatiquement lorsque l'on crée un document TEI : Nouveau - modèle du framework - TEI P5 - All

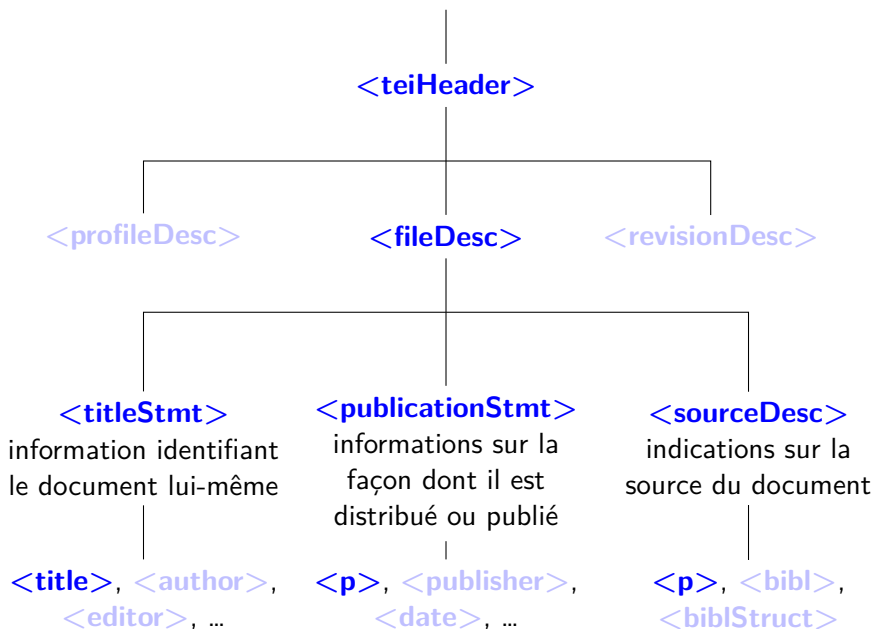
La structure d'un document TEI



Les métadonnées (<teiHeader>)



Les métadonnées (<teiHeader>)



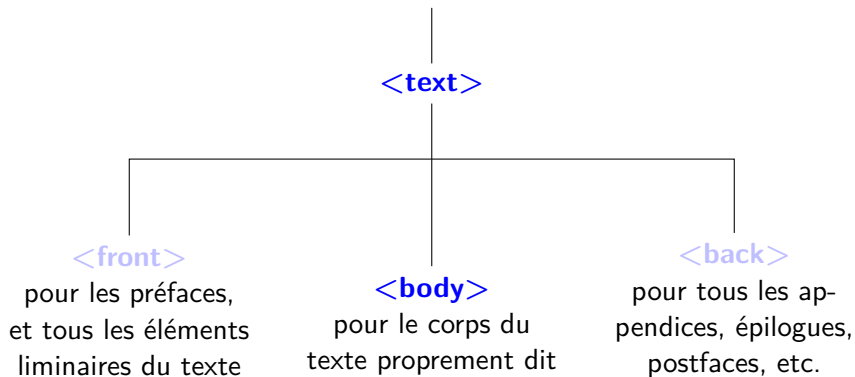
Example

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Thomas Paine: Common sense, a machine-readable
↪ transcript</title>
      <respStmt>
        <resp>compiled by</resp>
        <name>Jon K Adams</name>
      </respStmt>
    </titleStmt>
    <publicationStmt>
      <distributor>Oxford Text Archive</distributor>
    </publicationStmt>
    <sourceDesc>
      <bibl>The complete writings of Thomas Paine, collected and
↪ edited by Phillip S. Foner (New York, Citadel Press,
↪ 1945)</bibl>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

(source : TEI Guidelines, *The TEI Header and Its Components*)

Les éléments obligatoires d'un document TEI – suite

Le corps du texte



Le document TEI minimal

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Titre du document electronique</title>
      </titleStmt>
      <publicationStmt>
        <p>Informations sur la publication de l'œuvre.</p>
      </publicationStmt>
      <sourceDesc>
        <p>Informations sur la source dont est tirée l'œuvre.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <p>Le contenu du texte qu'on va encoder</p>
    </body>
  </text>
</TEI>
```

Encoder un texte : quelques balises courantes

Les divisions du texte

`<div>` Divisions du texte (chapitre, actes, parties, etc). Peut être précisé par l'attribut `@type` (valeur libre).

`<div1>`, `<div2>`, etc : divisions imbriquées

`<head>` Titres de section

`<p>` paragraphes

`<l>` vers

`<lg>` « line group » ensemble de vers faisant unité (strophe, laisse,...). Peut être précisé par l'attribut `@type` (valeur libre)

`<ab>` « anonymous block » type de séquence textuelle ne correspondant pas à ces divisions

Les divisions du texte

`<div>` Divisions du texte (chapitre, actes, parties, etc). Peut être précisé par l'attribut `@type` (valeur libre).

`<div1>`, `<div2>`, etc : divisions imbriquées

`<head>` Titres de section

`<p>` paragraphes

`<l>` vers

`<lg>` « line group » ensemble de vers faisant unité (strophe, laisse,...). Peut être précisé par l'attribut `@type` (valeur libre)

`<ab>` « anonymous block » type de séquence textuelle ne correspondant pas à ces divisions

`<s>` phrase

`<w>` mot

`<pc>` ponctuation

`<c>` caractère

Les noms propres

Deux possibilités :

1. `<name>`, qui peut être précisé avec l'attribut `@type` (valeurs possibles : `"person"`, `"place"`, `"org"`)
2. `<persName>`, `<placeName>`, or `<orgName>` : nom de personne, nom de lieu, non d'organisation.

La seconde solution est préférée, en particulier quand on veut créer un index.

La balise `<persName>` peut recevoir l'attribut `@type` ; elle peut avoir des enfants indiquant les composants du nom, par exemple `<surname>` et `<forname>`

`<rs>` : reference string : chaîne de caractères faisant référence à une entité ou un lieu, même sans contenir de nom propre (par exemple : l'abbé, lui, etc).

Donner un identifiant à un élément

L'attribut `@xml:id`

Appartient à l'espace de nom XML

Permet d'attribuer à un élément un identifiant unique auquel il est ensuite possible de se référer dans le cours du document : à des personnes, des lieux, des divisions du texte, des références bibliographiques, etc

L'identifiant peut être attribué entre l'élément `<teiHeader>` et l'élément `<text>`, par exemple dans l'élément `<standOff>`, ou dans le corps du texte.

standOff : « Functions as a container element for linked data, contextual information, and stand-off annotations embedded in a TEI document. »

Donner un identifiant à un élément

Exemple :

```
<TEI>
  <teiHeader>
    ...
  </teiHeader>
  <standOff>
    <listPerson>
      <person>
        <persName xml:id="Dest">Charles-Louis-Joseph Destale</persName>
      </person>
    </listPerson>
  </standOff>
  <text>
    <body>
      <p><rs ref="#Dest">L'abbé</rs> est en manchettes de dentelles</p>
    </body>
  </text>
</TEI>
```


Donner un identifiant à un élément

Exemple :

```
<bibl xml:id="tlk_36">Tolkien (1936)</bibl> tells us  
  ↪ that  
<quote source="#tlk_36">  
<title>Beowulf</title> is in fact so interesting as  
  ↪ poetry, in places poetry so powerful, that this  
  ↪ quite overshadows the historical content  
</quote>.
```

(exemple tiré des Guidelines)

Exercice 3 : encoder en TEI

Encoder de façon simple le début du livre I des *Essais* de Montaigne, édition de 1907, en utilisant l'attribut `@xml:id` pour les noms propres et les divisions du texte qui vous semblent pertinentes (ne vous préoccupez pas des métadonnées ici)

Le texte à encoder se trouve dans le fichier `transcription_Montaigne.txt`

Source : https://fr.wikisource.org/wiki/Essais/%C3%A9dition_Michaud,_1907/Livre_I/Chapitre_1

Exercice 3 : encoder en TEI

Encoder de façon simple le début du livre I des *Essais* de Montaigne, édition de 1907, en utilisant l'attribut `@xml:id` pour les noms propres et les divisions du texte qui vous semblent pertinentes (ne vous préoccupez pas des métadonnées ici)

Éléments à utiliser au sein de `<body>` :

- ▶ `<div1>`, `<div2>`
- ▶ `<p>`
- ▶ `<head>`
- ▶ `<persName>`
- ▶ `<placeName>`

Se repérer dans la TEI

La TEI : un cadre modulaire

- ▶ La TEI **n'est pas** un schéma
- ▶ La TEI fournit un **cadre** pour créer son propre schéma
- ▶ Les **Guidelines** de la TEI sont un guide pour se repérer dans la grammaire TEI
- ▶ Elles définissent plusieurs centaines de **concepts** différents, avec les instructions détaillées sur les **éléments** et les **attributs** qui peuvent être utilisés pour les représenter et la façon dont ils peuvent être combinés.
- ▶ Les concepts sont regroupés en **modules**

<http://www.tei-c.org/> Guideline

Les modules

Les **modules** sont des **ensembles cohérents de balises**. Il y a **23 modules**

Les modules

Les **modules** sont des **ensembles cohérents de balises**. Il y a **23 modules**

- ▶ 4 modules obligatoires : → métadonnées et éléments les plus courants

Les modules

Les **modules** sont des **ensembles cohérents de balises**. Il y a **23 modules**

- ▶ 4 modules obligatoires : → métadonnées et éléments les plus courants
- ▶ 18 modules additionnels selon le type de texte encodé (poésie, théâtre, dictionnaire, ...) ou le type de données (apparat critique, élément graphique, ...)

Les modules

Les **modules** sont des **ensembles cohérents de balises**. Il y a **23 modules**

- ▶ 4 modules obligatoires : → métadonnées et éléments les plus courants
- ▶ 18 modules additionnels selon le type de texte encodé (poésie, théâtre, dictionnaire, ...) ou le type de données (apparat critique, élément graphique, ...)

On peut combiner les éléments de différents modules pour composer son propre schéma (→ cours 5)

Les modules

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



Front Matter

Title

- i. [Releases of the TEI Guidelines](#)
- ii. [Dedication](#)
- iii. [Preface and Acknowledgments](#)
- iv. [About These Guidelines](#)
- v. [A Gentle Introduction to XML](#)
- vi. [Languages and Character Sets](#)

Back Matter

- Appendix A [Model Classes](#)
- Appendix B [Attribute Classes](#)
- Appendix C [Elements](#)
- Appendix D [Attributes](#)
- Appendix E [Datatypes and Other Macros](#)
- Appendix F [Bibliography](#)
- Appendix G [Deprecations](#)
- Appendix H [Prefatory Notes](#)
- Appendix I [Colophon](#)

Text Body

- 1 [The TEI Infrastructure](#)
- 2 [The TEI Header](#)
- 3 [Elements Available in All TEI Documents](#)
- 4 [Default Text Structure](#)
- 5 [Characters, Glyphs, and Writing Modes](#)
- 6 [Verse](#)
- 7 [Performance Texts](#)
- 8 [Transcriptions of Speech](#)
- 9 [Dictionaries](#)
- 10 [Manuscript Description](#)
- 11 [Representation of Primary Sources](#)
- 12 [Critical Apparatus](#)
- 13 [Names, Dates, People, and Places](#)
- 14 [Tables, Formulae, Graphics, and Notated Music](#)
- 15 [Language Corpora](#)
- 16 [Linking, Segmentation, and Alignment](#)
- 17 [Simple Analytic Mechanisms](#)
- 18 [Feature Structures](#)
- 19 [Graphs, Networks, and Trees](#)
- 20 [Non-hierarchical Structures](#)
- 21 [Certainty, Precision, and Responsibility](#)
- 22 [Documentation Elements](#)
- 23 [Using the TEI](#)

TEI sourcecode

- [Getting and Using the TEI Sources](#)
- [TEI GitHub Repository](#)
- [Bug Reports, Feature Requests, etc.](#)

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

TEI Consortium | [Feedback](#)

TEI Guidelines PS [version 4.7.0](#). Last updated on 16th November 2023, revision [p5d373a0c](#). This page generated on 2023-11-16T20:03:55Z.

Table des matières des *Guidelines*. Chaque chapitre couvre un module.

Les modules

Text Body

- ⊕ 1 [The TEI Infrastructure](#)
- ⊕ 2 [The TEI Header](#)
- ⊕ 3 [Elements Available in All TEI Documents](#)
- ⊕ 4 [Default Text Structure](#)
- ⊕ 5 [Characters, Glyphs, and Writing Modes](#)
- ⊕ 6 [Verse](#)
- ⊕ 7 [Performance Texts](#)
- ⊕ 8 [Transcriptions of Speech](#)
- ⊕ 9 [Dictionaries](#)
- ⊕ 10 [Manuscript Description](#)
- ⊕ 11 [Representation of Primary Sources](#)
- ⊕ 12 [Critical Apparatus](#)
- ⊕ 13 [Names, Dates, People, and Places](#)
- ⊕ 14 [Tables, Formulæ, Graphics, and Notated Music](#)
- ⊕ 15 [Language Corpora](#)
- ⊕ 16 [Linking, Segmentation, and Alignment](#)
- ⊕ 17 [Simple Analytic Mechanisms](#)
- ⊕ 18 [Feature Structures](#)
- ⊕ 19 [Graphs, Networks, and Trees](#)
- ⊕ 20 [Non-hierarchical Structures](#)
- ⊕ 21 [Certainty, Precision, and Responsibility](#)
- ⊕ 22 [Documentation Elements](#)
- ⊕ 23 [Using the TEI](#)

Les modules

Module name	Formal public identifier	Where defined
analysis	Analysis and Interpretation	17 Simple Analytic Mechanisms
certainty	Certainty and Uncertainty	21 Certainty, Precision, and Responsibility
core	Common Core	3 Elements Available in All TEI Documents
corpus	Metadata for Language Corpora	15 Language Corpora
dictionaries	Print Dictionaries	9 Dictionaries
drama	Performance Texts	7 Performance Texts
figures	Tables, Formulae, Figures	14 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	18 Feature Structures
linking	Linking, Segmentation, and Alignment	16 Linking, Segmentation, and Alignment
msdescription	Manuscript Description	10 Manuscript Description
namesdates	Names, Dates, People, and Places	13 Names, Dates, People, and Places
nets	Graphs, Networks, and Trees	19 Graphs, Networks, and Trees
spoken	Transcribed Speech	8 Transcriptions of Speech
tagdocs	Documentation Elements	22 Documentation Elements
tei	TEI Infrastructure	1 The TEI Infrastructure
textcrit	Text Criticism	12 Critical Apparatus
textstructure	Default Text Structure	4 Default Text Structure
transcr	Transcription of Primary Sources	11 Representation of Primary Sources
verse	Verse	6 Verse

le nom des modules (présenté dans le module 1)

Les modules

Les modules obligatoires :

`tei` définition des classes, macros et types de données

`header` en-tête TEI (métadonnées du document)

`core` éléments incontournables pour tout type de document

`textstructure` éléments de base pour structurer un texte de type livre

Text Body

- ⊕ 1 [The TEI Infrastructure](#)
- ⊕ 2 [The TEI Header](#)
- ⊕ 3 [Elements Available in All TEI Documents](#)
- ⊕ 4 [Default Text Structure](#)

Table of contents

- 6.1 Structural Divisions of Verse Texts
- 6.2 Components of the Verse Line
- 6.3 Encoding Textual Structures Across Verses
- 6.4 Rhyme and Metrical Analysis
- 6.5 Rhyme
- 6.6 Metrical Notation Declaration
- 6.7 Encoding Procedures for Other Verse Features
- 6.8 Module for Verse

- « 5 Characters, Glyphs, and Writing Modes
- » 7 Performance Texts

Home

6 Verse

This module is intended for use when encoding texts which are entirely or predominantly in verse, and for which the elements for encoding verse structure already provided by the core module are inadequate.

The tags described in section [3.13.1 Core Tags for Verse](#) include elements for the encoding of verse lines and line groups such as stanzas: these are available for any TEI document, irrespective of the module it uses. Like the modules for prose and for drama, the module for verse additionally makes use of the module defined in chapter [4 Default Text Structure](#) to define the basic formal structure of a text, in terms of [front](#), [body](#) and [back](#) elements and the text-division elements into which these may be subdivided.

The module for verse extends the facilities provided by these modules in the following ways:

- a special purpose [caesura](#) element is provided, to allow for segmentation of the verse line (see section [6.2 Components of the Verse Line](#))
- a set of attributes is provided for the encoding of rhyme scheme and metrical information (see sections [6.4 Rhyme and Metrical Analysis](#) and [6.5 Rhyme](#))
- a special purpose [rhyme](#) element is provided to support simple analysis of rhyming words (see section [6.5 Rhyme](#))

6.1 Structural Divisions of Verse Texts

Like other kinds of text, texts written in verse may be of widely differing lengths and structures. A complete poem, no matter how short, may be treated as a free-standing text, and encoded in the same way as a distinct prose text. A group of poems functioning as a single unit may be encoded either as a [group](#) or as a [text](#), depending on the encoder's view of the text. For further discussion, including an example encoding for a verse anthology, see chapter [4 Default Text Structure](#).

Many poems consist only of ungrouped lines. This short poem by Emily Dickinson is a simple case:

```
<text>
<front>
<head>1755</head>
</front>
<body>
<l>To make a prairie it takes a clover and one bee,</l>
<l>One clover, and a bee,</l>
<l>And revery.</l>
<l>The revery alone will do,</l>
<l>If bees are few.</l>
</body>
</text>
```

[bibliography](#)

» 6.2 Components of the Verse Line
Home

Exemple de description de module : le module verse

Les classes

Les éléments et les attributs sont également organisés en **classes de modèle** (model class) et **classes d'attributs** (attribute class)

`model.class` (Classe de modèle) Regroupe des éléments qui peuvent se trouver au même endroit dans le document + regroupement sémantique

`att.class` (Classe d'attributs) Regroupe des éléments qui partagent les mêmes attributs

Back Matter

- ⊕ Appendix A [Model Classes](#)
- ⊕ Appendix B [Attribute Classes](#)
- ⊕ Appendix C [Elements](#)
- ⊕ Appendix D [Attributes](#)
- ⊕ Appendix E [Datatypes and Other Macros](#)
- ⊕ Appendix F [Bibliography](#)
- ⊕ Appendix G [Deprecations](#)
- ⊕ Appendix H [Prefatory Notes](#)
- Appendix I [Colophon](#)

Classes et liste de l'ensemble des éléments définis dans la TEI

Les classes de modèle

Exemple de classes de modèle

`model.divPart` Éléments qui peuvent être contenus par une
<div>

`model.phrase` Éléments pouvant apparaître au niveau de la
phrase

`model.nameLike` Regroupement sémantique pour les divers
types de noms

La liste alphabétique des classes de modèle dans les Guidelines

Appendix A.1 About the Model Classes
Appendix

» Appendix B Attribute Classes

Home

Appendix A Model Classes

Appendix A.1 About the Model Classes Appendix

This appendix gives you a list of model classes and links to the reference pages for them. There are 126 distinctly-named model classes in revision [e5dd73ed0](#) of TEI P5 [Version 4.7.0](#) of the TEI Guidelines.

Sorted alphabetically

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

[Show by module](#)

a

[model.addressLike](#) [model.addrPart](#) [model.annotationLike](#) [model.annotationPart](#) [model.applicationLike](#) [model.attributable](#)
[model.availabilityPart](#)

b

[model.biblLike](#) [model.biblPart](#)

c

[model.castItemPart](#) [model.catDescPart](#) [model.certLike](#) [model.choicePart](#) [model.common](#) [model.contentPart](#) [model.correspActionPart](#)
[model.correspContextPart](#) [model.correspDescPart](#)

d

[model.dateLike](#) [model.descLike](#) [model.describedResource](#) [model.dimLike](#) [model.div1Like](#) [model.div2Like](#) [model.div3Like](#) [model.div4Like](#)
[model.div5Like](#) [model.div6Like](#) [model.div7Like](#) [model.divBottom](#) [model.divBottomPart](#) [model.divGenLike](#) [model.divLike](#) [model.divPart](#)
[model.divPart.spoken](#) [model.divTop](#) [model.divTopPart](#) [model.divWrapper](#)

e

Les classes de modèle

Exemple : la classe `model.quoteLike`

`model.quoteLike`

`model.quoteLike` groups elements used to directly contain quotations.

Module

tei — The TEI Infrastructure

Used by

[model.attributable](#) [model.msItemPart](#)

Members

[cit](#) [quote](#)

Les classes d'attributs

Exemple de classes d'attributs

`att.global` Attributs pouvant être utilisés avec tout élément.
Contient : `@xml:id`, `@n`, `@xml:lang`, ...

`att.naming` Les attributs communs aux différents éléments qui se réfèrent à des noms de personnes, de lieux, d'organisations, etc. Contient : `@key`, `@ref`, `@role`,

Les classes d'attributs

La liste alphabétique des classes d'attribut dans les Guidelines

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

b

[att.breaking](#)

c

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

d

[att.damaged](#) [att.datable](#) [att.datable.custom](#) [att.datable.iso](#) [att.datable.vy3c](#) [att.datcat](#) [att.declarable](#) [att.declaring](#) [att.deprecated](#) [att.dimensions](#)
[att.divLike](#) [att.docStatus](#) [att.duration](#) [att.duration.iso](#) [att.duration.vy3c](#)

e

[att.edition](#) [att.editLike](#) [att.enjamb](#) [att.entryLike](#)

f

[att.formula](#) [att.fragmentable](#)

g

[att.gaijiProp](#) [att.global](#) [att.global.analytic](#) [att.global.change](#) [att.global.facs](#) [att.global.linking](#) [att.global.rendition](#) [att.global.responsibility](#)
[att.global.source](#)

h

[att.handFeatures](#)

i

[att.identified](#) [att.internetMedia](#) [att.interpLike](#)

l

[att.lexicographic](#) [att.lexicographic.normalized](#) [att.linguistic](#) [att.locatable](#)

Les classes d'attributs

Exemple : la classe d'attribut att.witnessed

att.witnessed

Home
B Attribute Classes

att.witnessed provides attributes used to identify the witnesses supporting a particular reading in a critical apparatus. [12.1 The Apparatus Entry, Readings, and Witnesses]	
Module	textcrit — Critical Apparatus
Members	lem rdg
Attributes	<div><div>@wit ↕</div><div>(witness or witnesses) contains a space-delimited list of one or more pointers indicating the witnesses which attest to a given reading.</div><div>StatusOptional</div><div>Datatype1–∞ occurrences of teidata.pointer separated by whitespace</div><div>Note<p>If the apparatus contains readings only for a single witness, this attribute may be consistently omitted.</p><p>This attribute may occur both within an apparatus gathering variant readings in the transcription of an individual witness and within an apparatus gathering readings from different witnesses.</p><p>Additional descriptions or alternative versions of the sigla referenced may be supplied as the content of a child wit element.</p></div></div>

Exercice 4 : résoudre un problème d'encodage précis en utilisant les *Guidelines*

Trouver comment encoder une liste de personnages et les indications de lieu au début d'une pièce de théâtre ; encoder celles du *Cid*. Le texte à encoder se trouve dans le fichier `transcription_Cid.txt` (https://fr.wikisource.org/wiki/Le_Cid/%C3%89dition_Ginn/Personnages)

ACTEURS

DON FERNAND, premier roi de Castille
DONA URRACQUE, infante de Castille
DON DIÈGUE, père de don Rodrigue
DON GOMÈS, comte de Gormas, père de Chimène
DON RODRIGUE, amant de Chimène
DON SANCHE, amoureux de Chimène

DON ARIA }
DON ALONSE } gentilshommes castillans.

CHIMÈNE, fille de don Gomès
LÉONOR, gouvernante de l'Infante
ELVIRE, gouvernante de Chimène
UN PAGE de l'Infante

La scène est à Séville

Solution

```
<text>
  <front>
    <castList>
      <head>ACTEURS</head>
      <castItem>
        <role>DON FERNAND</role>
        <roleDesc>premier roi de Castille</roleDesc>
      </castItem>
      <castItem>
        <role>DONA URRaque</role>
        <roleDesc>infante de Castille</roleDesc>
      </castItem>
    <!-- et caetera -->
  </castList>
  <set>
    <p>La scène est à Séville</p>
  </set>
</front>
<body>...</body>
</text>
```

Solution

Cas particulier : les rôles regroupés

```
<castGroup rend="braced">
  <castItem>
    <role>DON ARIA</role>
  </castItem>
  <castItem>
    <role>DON ALONSE</role>
  </castItem>
  <roleDesc>gentilshommes castillans</roleDesc>
</castGroup>
```


Exercice 5

Résoudre un problème précis d'encodage en s'appuyant sur les *Guidelines*

Rechercher dans les *Guidelines* de la TEI des balises et attributs pour remplacer celles que nous avons utilisées dans notre encodage initial du poème « L'Adieu ».

→ transformer l'encodage suivant en encodage valide selon le vocabulaire XML-TEI :

```
<titre>L'Adieu</titre>
<poeme>
  <vers rime="A" num="1">J'ai cueilli ce brin de bruyère</vers>
  <vers rime="B" num="2">L'automne est morte souviens-t'en </vers>
  <vers rime="A" num="3">Nous ne nous verrons plus sur terre</vers>
  <vers rime="A" num="4">Odeur du temps brin de bruyère </vers>
  <vers rime="B" num="5">Et souviens-toi que je t'attends</vers>
</poeme>
```

Ne vous préoccupez pas des métadonnées (<teiHeader>).

Solution

```
<text>
  <front>
    <head>L'Adieu</head>
  </front>
  <body>
    <lg rhyme="AABBB"
      <l n="1">J'ai cueilli ce brin de bruyère</l>
      <l n="2"> L'automne est morte souviens-t'en</l>
      <l n="3">Nous ne nous verrons plus sur terre</l>
      <l n="4">Odeur du temps brin de bruyère</l>
      <l n="5">Et souviens-toi que je t'attends</l>
    </lg>
  </body>
</text>
```

Pour la séance prochaine

Lire dans *Qu'est-ce que la Text Encoding Initiative* de Lou Burnard les deux chapitres sur « la corne d'abondance de la TEI » et le chapitre « Personnaliser la TEI »

Exercice 6 : Résoudre un problème précis d'encodage en s'appuyant sur les *Guidelines*

- ▶ Encoder les métadonnées, le titre, les épigraphes et les dix premières entrées de lettre A du *Dictionnaire des idées reçues* de Flaubert.
- ▶ Le texte à transcrire est dans le fichier `transcription_Flaubert.txt`
- ▶ Source du texte :
https://fr.wikisource.org/wiki/Dictionnaire_des_id%C3%A9es_re%C3%A7ues/Texte_entier