



ETIQUETADO

frad@jase@fyi.uva.es



Digitalización

Paz a España, paz segura.
Canten abiertos los campos
dichosos de Extremadura.

Lean los niños, las flores,
y entre las negras encinas
todos los trabajadores.

¡Lejos tanta noche oscura!
¡Para siempre en primavera,
la tierra de Extremadura!

Yo os traigo en este pregón,
el libro, fuente de gracia,
que ilumina el corazón.

Agua tranquila, agua pura,
agua que riegue la sangre
del alma de Extremadura.

Libros de amor, luz, ensueño.
Alta vida y viento en calma
en todo el pueblo extremeño.

frad@jase@fyi.uva.es



Transcripción

Paz a España, paz segura.
Canten abiertos los campos
dichosos de Extremadura.

Lean los niños, las flores,
y entre las negras encinas
todos los trabajadores.

¡Lejos tanta noche oscura!
¡Para siempre en primavera,
la tierra de Extremadura!

Yo os traigo en este pregón,
el libro, fuente de gracia,
que ilumina el corazón.

Agua tranquila, agua pura,
agua que riegue la sangre
del alma de Extremadura.

Libros de amor, luz, ensueño.
Alta vida y viento en calma
en todo el pueblo extremeño.

Codificación

```
<poema>
<estrofa>
<verso>Paz a España</topom></topom>, paz segura.</verso>
<verso>Canten abiertos los campos</verso>
<verso>dichosos de <topom>Extremadura</topom>.</verso>
</estrofa>
<estrofa>
<verso>Lean los niños, las flores,</verso>
<verso>y entre las negras encinas</verso>
<verso>todos los trabajadores.</verso>
</estrofa>
<estrofa>
<verso>¡Lejos tanta noche oscura!</verso>
<verso>¡Para siempre en primavera,</verso>
<verso>la tierra de <topom>Extremadura</topom>!</verso>
</estrofa>
<estrofa>
<verso>Yo os traigo en este pregón,</verso>
<verso>el libro, fuente de gracia,</verso>
<verso>que ilumina el corazón.</verso>
</estrofa>
<estrofa>
<verso>Agua tranquila, agua pura,</verso>
<verso>agua que riegue la sangre</verso>
<verso>del alma de <topom>Extremadura.</topom>.</verso>
</estrofa>
<estrofa>
<verso>Libros de amor, luz, ensueño.</verso>
<verso>Alta vida y viento en calma</verso>
<verso>en todo el pueblo extremeño.</verso>
</estrofa>
</poema>
```



Estilos de etiquetado

Al principio solo existía **etiquetado de proceso**

TINTA ROJA; imprime resultado; CURSIVAS

Al generalizarse dio paso al **etiquetado descriptivo**

<titulo tipo="subtitulo">subtitulo</titulo>

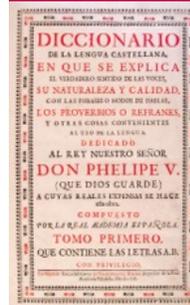
También conocido como **codificación** o **anotación**

El etiquetado descriptivo permite una mejor **reutilización** de los datos

frad@jase@fyi.uva.es



Un caso



IDIOMA. f. m. La Lengua vulgar, propia y particular de qualquier Nacion. Es voz Griega, que significa la propiedad. Lat. *Idioma, alia.* Cerv. Quix. tom.1. cap.6. Si aqui le hallo, y que habla en otra lengua que la fuya, no le guardaré respeto alguno; pero si habla en su *Idioma*, le pondré sobre mi cabeza, MÅNER. Apolog. Dedicat. He procurado exprimir en nucitro *Idioma* el sentido del original.

IDIOMA. Se toma tambien por la phraze y modo particular de hablar de algunos; y así se dice en Idioma de la Corte, ó en Idioma de Palacio. Lat. *Idioma.* CORN. CHRON. tom. 3. lib.1. cap.2. Si antes de cumplir un año hablaba con perfeccion todo el *idioma* del Cielo.

IDIOMA. Por translaçion se aplica à las voces con que los brutos explican sus materiales afectos, ò pasiones. Lat. *Translatiit idiomis.* GONG. Rom. bur. 3.

Murmuraban los rocines à la puerta de Palacio, no en sonorofoz relinchos, que esso es ya mi de caballos, fino en su bestial idioma, ni gruñendo ni rufando.

frad@jase@fyi.uva.es



Caracterización gramatical

LEMA

Definición

Traducción

Autoridad:

Autor

Título

Localización

Cita

IDIOMA, f. m. La Lengua vulgar, propia y particular de qualquier Nación. Es voz Griega, que significa la propiedad. Lat. *Idioma*, *idiu*, Cerv. Quix. tom. 1. cap. 6. Si aqui le hallo, y que habla en otra lengua que la fuya, no le guardaré respeto alguno; pero si habla en su *idioma*, le pondré sobre mi cabeza. MANNER. Apolog. *Dedicat*. He procurado exprimir en nuestro *Idioma* el sentido del original.

IDIOMA. Se toma tambien por la phrase y modo particular de hablar de algunos; y assi se dice en Idioma de la Corte, o en Idioma de Palacio. Lat. *Idioma*. CORN. CHRON. tom. 3. lib. 1. cap. 2. Si antes de cumplir un año hablaba con perfeccion todo el *idioma* del Cielo.

IDIOMA. Por traslacion se aplica à las voces con que los brutos explican sus materiales afectos, ò pasiones. Lat. *Translatititè idioma*. GONG. Rom. butl. 3.

*Murmuraban los rocines
à la puerta de Palacio,
no en sonorojos relinchos,
que esto es ya mi de caballos,
fino en su bñfial idioma,
ni gruñendo ni rifando.*

<http://www.fri.es/Paginas/Autoridades.aspx>

<?xml?>

Una sopa de letras



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

XML – Qué es – Para qué sirve

eXtensible Markup Language (XML) es un lenguaje de codificación muy sencillo y flexible derivado del SGML (ISO 8879).

El SGML era un estándar para la creación y definición de documentos para la publicación electrónica a gran escala.

El SGML tenía un gran problema de diseño: era muy complicado.

XML – Qué es – Para qué sirve

UVa

El XML es un conjunto de **datos estructurados** representados con cadenas de texto

El XML se parece al HTML salvo porque:

El XML es **extensible**

El XML debe estar **bien formado**

El XML puede ser **validado**

El XML no está ligado a ninguna aplicación, plataforma o fabricante

El XML da protagonismo al **proveedor de contenidos** y facilita la integración de los datos

fradejas@fyi.uva.es



XML – Terminología

UVa

```
<?xml version="1.0"?>
<raiz>
  <elemento atributo="valor"> contenido </elemento>
  <!-- comentario -->
</raiz>
```

fradejas@fyi.uva.es



XML – Terminología

UVa

Un documento XML puede contener:

Elementos, posiblemente lleven atributos

Instrucciones de procesamiento

Comentarios

Referencia a entidades

Secciones marcadas (CDATA, IGNORE, INCLUDE)

Un documento XML debe estar **bien formado** y puede ser **válido**

fradejas@fyi.uva.es



XML – Reglas

UVa

Un documento XML representa una especie de **árbol**

Tiene una única **raíz** y muchas ramas (**nodos**)

Cada nodo puede ser

Un subárbol

Un único **elemento** (quizá con algún **atributo**)

Una **cadena de datos**

Cada elemento tiene un nombre o **identificador genérico**

Los nombres de los atributos están predefinidos para cada elemento y los valores pueden estar limitados

fradejas@fyi.uva.es



El árbol XML – un ejemplo

UVa

MENSAJE

A: Jefe de Burocracia

De: Burócrata

Asunto: Utilidad del XML

He leído en el periódico que "... los burócratas emplean una gran parte de su tiempo en asuntos misteriosos que llaman expedientes, informes, impresos, notas. Todo ello lo escriben en grandes hojas de papel que tienen que rellenar en su totalidad con escritos, y al poco de complimentarlos los destruyen" ¿Cree que el XML podría ayudarnos?

Burócrata

fradejas@fyi.uva.es



El árbol XML

UVa

Un documento XML está codificado como una secuencia lineal de caracteres

Comienza con una **instrucción de procesamiento** especial

Los elementos se marcan con una **etiqueta de apertura** y otra **de cierre**

Los caracteres **<** y **&** son mágicos y se han de "escapar" siempre que se quieran usar para que se representen a sí mismos

Los comentarios están delimitados con **<!--** y **-->**

Las secciones **CDATA** se encierran entre **<![CDATA[** y **]]>**

Los pares atributo/valor se suministran en la etiqueta de apertura y pueden darse en cualquier orden

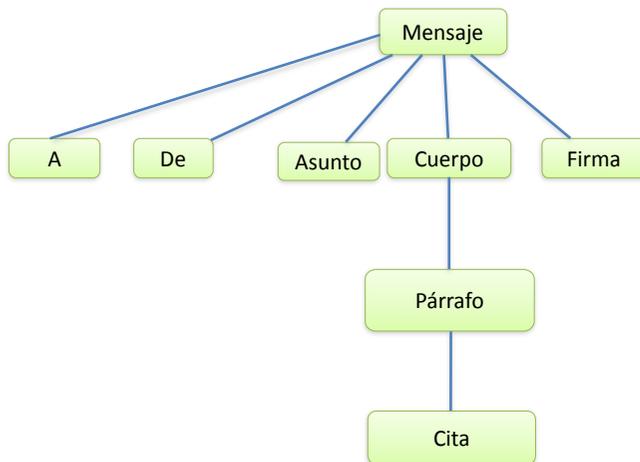
Las referencias a entidades están delimitadas con **&** y **;**

fradejas@fyi.uva.es



El árbol XML – un ejemplo

UVa



fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
  <a>Jefe de Burocracia</a>
  <de>Burócrata</de>
  <asunto>Utilidad del XML</asunto>
  <cuerpo>
    <p type="pregunta">¿Cree que el XML podría ayudarnos?</p>
  </cuerpo>
  <firma>Burócrata</firma>
</mensaje>
```

Declaración XML

Declaración del espacio de nombres (namespace)

El elemento raíz del documento

Otros elementos y contenido

Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
<a>Jefe de Burocracia</a>
<de>Burócrata</de>
<asunto>Utilidad del XML</asunto>
<cuero>
<p type="pregunta">¿Cree que el XML podría
ayudarnos?</p>
</cuero>
<firma>Burócrata</firma>
</mensaje>
```

Declaración XML

Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
<a>Jefe de Burocracia</a>
<de>Burócrata</de>
<asunto>Utilidad del XML</asunto>
<cuero>
<p type="pregunta">¿Cree que el XML podría
ayudarnos?</p>
</cuero>
<firma>Burócrata</firma>
</mensaje>
```

Declaración XML

Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
<a>Jefe de Burocracia</a>
<de>Burócrata</de>
<asunto>Utilidad del XML</asunto>
<cuero>
<p type="pregunta">¿Cree que el XML podría
ayudarnos?</p>
</cuero>
<firma>Burócrata</firma>
</mensaje>
```

Declaración XML

Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
<a>Jefe de Burocracia</a>
<de>Burócrata</de>
<asunto>Utilidad del XML</asunto>
<cuero>
<p type="pregunta">¿Cree que el XML podría
ayudarnos?</p>
</cuero>
<firma>Burócrata</firma>
</mensaje>
```

Declaración XML

Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
  <a>Jefe de Burocracia</a>
  <de>Burócrata</de>
  <asunto>Utilidad del XML</asunto>
  <cuero>
    <p type="pregunta">¿Cree que el XML podría ayudarnos?</p>
  </cuero>
  <firma>Burócrata</firma>
</mensaje>
```

Declaración XML
Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
  <a>Jefe de Burocracia</a>
  <de>Burócrata</de>
  <asunto>Utilidad del XML</asunto>
  <cuero>
    <p type="pregunta">¿Cree que el XML podría ayudarnos?</p>
  </cuero>
  <firma>Burócrata</firma>
</mensaje>
```

Declaración XML
Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



Partes de un documento XML

UVa

```
<?xml version="1.0" encoding="UTF-8"?>
<mensaje xmlns="http://www.uva.es/ns/nonTEI">
  <a>Jefe de Burocracia</a>
  <de>Burócrata</de>
  <asunto>Utilidad del XML</asunto>
  <cuero>
    <p type="pregunta">¿Cree que el XML podría ayudarnos?</p>
  </cuero>
  <firma>Burócrata</firma>
</mensaje>
```

Declaración XML
Declaración del espacio de nombres (namespace)
El elemento raíz del documento
Otros elementos y contenido
Atributos y valor

fradejas@fyi.uva.es



La declaración XML

UVa

Todo documento XML debe comenzar con una **declaración XML** que tiene dos misiones:

- Especifica que es un documento XML y qué versión sigue
- Especifica qué codificación de caracteres utiliza el documento

```
<?xml version="1.0"?>
<?xml version="1.0" encoding="UTF-8"?>
```

La codificación de caracteres por defecto, y recomendada, es UTF-8

fradejas@fyi.uva.es



Las etiquetas y atributos que se pueden utilizar en un documento XML se tienen que declarar en algún lugar. La manera más sencilla de hacerlo es por medio de la declaración de *namespace*.

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

Esta declaración se incluye puesto que los documentos XML pueden incorporar elementos de diferentes sistemas:

Una declaración de *namespace* asocia un prefijo de *namespace* con un identificador externo que toma el aspecto de un URI

El *namespace* por defecto se puede declarar con el atributo **xmlns**

Otros *namespace*s que se utilicen deben tener un prefijo declarado

```
<TEI xmlns:xi="http://www.w3.org/2001/XInclude"
xmlns:svg="http://www.w3.org/2000/svg"
xmlns:math="http://www.w3.org/1998/Math/MathML"
xmlns="http://www.tei-c.org/ns/1.0">
```

¿Qué quiere decir que un documento XML ha de estar **bien formado**?

1. Tiene que existir un único elemento raíz que contenga todo el documento
2. Cada subárbol debe estar correctamente anidado <a>pedro <de>juan</de>, es decir, no se pueden solapar <a>pedro <de>juan</de>
3. Los nombres de las etiquetas son *case sensitive* <de> ≠ <De> ≠ <DE>
4. Una etiqueta de apertura <p> debe tener siempre una etiqueta de cierre </p> (salvo en los casos de etiquetas vacías <vacía/>)
5. No debe haber espacios en blanco tras < o antes de </ <a>pedro y no < a>pedro</ a> ni < a>pedro< /a>
6. El valor de los atributos ha de estar siempre entrecomillado
<p type="text">algo de texto</p>
<p type='text'>algo de texto</p>

Nota: Además de estar bien formado un documento puede ser **VÁLIDO**. Esto quiere decir que el documento cumple las reglas de un esquema específico, como el del TEI

¿Qué quiere decir que un documento XML ha de estar **bien formado**?

1. Tiene que existir un único elemento raíz que contenga todo el documento
2. Cada subárbol debe estar correctamente anidado <a>pedro <de>juan</de>, es decir, no se pueden solapar <a>pedro <de>juan</de>
3. Los nombres de las etiquetas son *case sensitive* <de> ≠ <De> ≠ <DE>
4. Una etiqueta de apertura <p> debe tener siempre una etiqueta de cierre </p> (salvo en los casos de etiquetas vacías <vacía/>)
5. No debe haber espacios en blanco tras < o antes de </ <a>pedro y no < a>pedro</ a> ni < a>pedro< /a>
6. El valor de los atributos ha de estar siempre entrecomillado
<p type="text">algo de texto</p>
<p type='text'>algo de texto</p>

Un documento XML es **válido** cuando cumple las reglas establecidas en un **esquema** externo

El esquema especifica:

1. El nombre del elemento raíz
2. El nombre de todos los elementos que se pueden utilizar
3. Los nombres, tipo de datos y (ocasionalmente) los valores por defecto de los atributos
4. Reglas acerca de cómo se pueden anidar los elementos
5. Y otras cosillas dependiendo del lenguaje del esquema

Nota: Un esquema no dice nada acerca de qué significan los elementos

¿Qué puede hacer el esquema?

UVa

1. Asegurar que en los documentos solo se utilizan los elementos, atributos y entidades predefinidas
2. Establecer las reglas estructurales del tipo 'toda receta debe incluir una lista de ingredientes'
3. Asegurar que la misma cosa se llame de la misma manera siempre

fradejas@fyl.uva.es



¿Qué puede hacer el esquema?

UVa

1. Asegurar que en los documentos solo se utilizan los elementos, atributos y entidades predefinidas
2. Establecer las reglas estructurales del tipo 'toda receta debe incluir una lista de ingredientes'
3. Asegurar que la misma cosa se llame de la misma manera siempre

fradejas@fyl.uva.es



Los lenguajes de esquema varían en la cantidad de validación que ofrecen

UVa



fradejas@fyl.uva.es



TEI - Text Encoding Initiative

UVa

La TEI (Text Encoding Initiative) desarrolla y mantiene un estándar para la representación de los textos en el mundo digital, por lo que . . .

- ✓ Ha definido y nombrado varios cientos de distinciones textuales
- ✓ Ofrece un conjunto de módulos que se pueden utilizar para definir esquemas que hagan esas distinciones
- ✓ Ofrece un mecanismo para modificar y combinar esas definiciones con otras nuevas utilizando el mismo esquema conceptual

fradejas@fyl.uva.es



Tiene gran influencia en los trabajos sobre:

- Bibliotecas digitales
- Colecciones textuales
- Corpus lingüísticos
- Conjuntos de datos filológicos

La(s) estructura(s) básica(s)

```
<TEI>
  <teiHeader>
    <!-- obligatorio -->
  </teiHeader>
  <text>
    <!-- obligatorio -->
  </text>
</TEI>
```

```
<TEI>
  <teiHeader>
    <!-- obligatorio -->
  </teiHeader>
  <facsimile>
    <!-- opcional -->
  </facsimile>
  <text>
    <!-- obligatorio si no hay facsimile -->
  </text>
</TEI>
```

```
<teiCorpus>
  <teiHeader>
    <!-- encabezamiento del corpus -->
  </teiHeader>
  <TEI>
    <teiHeader>
      <!-- encabezamiento de cada documento del corpus -->
    </teiHeader>
    <text>
      <!-- obligatorio -->
    </text>
  </TEI>
  <TEI>
    <!-- repetible cuantas veces sea necesario -->
  </TEI>
</teiCorpus>
```

Sus objetivos

- Mejorar el intercambio e integración de los datos filológicos
- Ofrecer apoyo para todos los textos, de todas las lenguas, de cualquier época
- Guiar a los novatos: **qué** codificar
- Ayudar a los especialistas: **cómo** codificar

Estos objetivos, aparentemente incompatibles, permiten ofrecer un entorno altamente flexible y modular

La(s) estructura(s) básica(s)

- Todo documento TEI (<TEI>) consiste en un *encabezamiento* (<teiHeader>) y (al menos) un *texto* (<text>)
- El *encabezamiento* contiene
 - Una descripción obligatoria del fichero (<fileDesc>)
 - Descripciones, opcionales, de la codificación (<encodingDesc>), del perfil (<profileDesc>) y las revisiones del fichero (<revisionDesc>)
- El *encabezamiento* es esencial para
 - Control bibliográfico e identificación
 - Para documentar la fuente y su procesamiento

Estructura de un texto TEI

- En los casos más sencillos un texto consiste en un conjunto de párrafos, cláusulas o versos
- En la TEI un texto (<text>) tiene una estructura un poco más compleja: contiene
 - Material inicial (<front>) opcional
 - Material final (<back>) opcional
 - Cuerpo (<body>)

fradegias@fyl.uva.es



Estructura de un texto compuesto TEI

```

<TEI>
  <teiHeader>
    <!-- obligatorio -->
  </teiHeader>
  <text>
    <front>
      <!-- opcional -->
    </front>
    <group>
      <text>
        <body>
          <!-- obligatorio -->
        </body>
      </text>
      <text>
        <body>
          <!-- obligatorio -->
        </body>
      </text>
      <text>
        <body>
          <!-- obligatorio -->
        </body>
      </text>
    </group>
    <back>
      <!-- opcional -->
    </back>
  </text>
</TEI>
    
```

fradegias@fyl.uva.es



Estructura de un texto TEI

```

<TEI>
  <teiHeader>
    <!-- obligatorio -->
  </teiHeader>
  <text>
    <front>
      <!-- opcional -->
    </front>
    <body>
      <!-- obligatorio -->
    </body>
    <back>
      <!-- opcional -->
    </back>
  </text>
</TEI>
    
```

```

<TEI>
  <teiHeader>
    <!-- obligatorio -->
  </teiHeader>
  <text>
    <body>
      <p> . . . </p>
    </body>
  </text>
</TEI>
    
```

ESTRUCTURA MÍNIMA DE UN DOCUMENTO TEI

fradegias@fyl.uva.es



El cuerpo (<body>) de un texto hay de tener como mínimo un párrafo <p>

Lo usual es que el cuerpo (<body>) tenga subdivisiones <div> que se anidan recursivamente unas en otras

- El atributo @type marca el nivel (“libro”, “capítulo”)
- El atributo @n numera una división determinada
- El atributo @xml:id da a una división determinada un identificador único

fradegias@fyl.uva.es



Por ejemplo

```
<text>
<front>
  <!-- portada, presentación, etc. -->
</front>
<body>
  <div type="libro" n="1" xml:id="MD100">
    <head>Libro Primero</head>
    <div type="capitulo" n="1" xml:id="MD101">
      <head>Capítulo primero</head>
      <!-- el texto del capítulo primero va aquí -->
    </div>
    <div type="capitulo" n="2" xml:id="MD102">
      <head>Capítulo segundo</head>
      <!-- el texto del capítulo segundo va aquí -->
    </div>
    <!-- resto del libro primero aquí -->
  </div>
  <div type="libro" n="2" xml:id="MD200">
    <!-- el texto del libro segundo va aquí -->
  </div>
  <!-- los demás libros van aquí -->
</body>
</text>
```



fradejas@fyi.uva.es



Los componentes textuales

- La prosa se construye básicamente con párrafos <p>
- La poesía se construye con versos <l> a veces agrupados <lg>
- El teatro consiste en diálogos <sp> que pueden contener párrafos <p> o versos <l> (a veces agrupados en estrofas <lg>) entre los que se encuentran acotaciones <stage>

Pueden estar entremezclados, pueden aparecer en textos sin dividir . . .

fradejas@fyi.uva.es



Por ejemplo

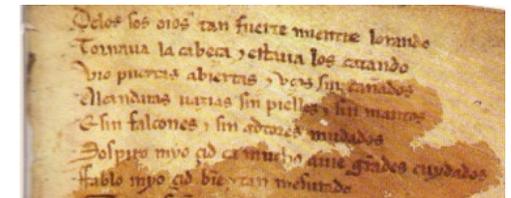
```
<text>
<front>
  <!-- portada, presentación, etc. -->
</front>
<body>
  <div type="libro" n="1" xml:id="MD100">
    <head>Libro Primero</head>
    <div type="capitulo" n="1" xml:id="MD101">
      <head>Capítulo primero</head>
      <!-- el texto del capítulo primero va aquí -->
    </div>
    <div type="capitulo" n="2" xml:id="MD102">
      <head>Capítulo segundo</head>
      <!-- el texto del capítulo segundo va aquí -->
    </div>
    <!-- resto del libro primero aquí -->
  </div>
  <div type="libro" n="2" xml:id="MD200">
    <!-- el texto del libro segundo va aquí -->
  </div>
  <!-- los demás libros van aquí -->
</body>
</text>
```



fradejas@fyi.uva.es



Poesía . . .



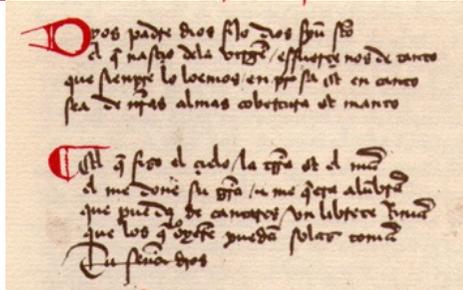
```
<div type="canto">
  <l>De los sos ojos tan fuertemiente lorando</l>
  <l>Tornava la cabeça y estaualos catando.</l>
  <l>Vio puertas abiertas & uços sin cañados,</l>
  <l>Alcandaras uazias sin pielles & sin mantos</l>
  <l>E sin falcones & sin adtores mudados.</l>
  <l>Sospiro myo çid ca mucho auie grandes cuydados</l>
  <l>ffablo myo çid bien & tan mesurado</l>
</div>
```

Poema del Cid

fradejas@fyi.uva.es



Poesía . . .



```
<lg type="cuaderna_via">
  <1>Dios padre dios fiijo dios spiritu santo</1>
  <1>el que naçio de la virgen esfuerçe nos de tanto</1>
  <1>que siempre lo loemos en prosa Et en canto</1>
  <1>sea de nuestras almas cobertura et manto</1>
</lg>
<lg type="cuaderna_via">
  <1>El que fizo el çielo la tierra et la mar</1>
  <1>el me done su grraçia r me quiera alubrar</1>
  <1>que pueda de cantares vn librete Rimar</1>
  <1>que los que lo oyeren puedan solaz tomar</1>
</lg>
```

Juan Ruiz, *Libro de buen amor*

fradejase@fyl.luva.es



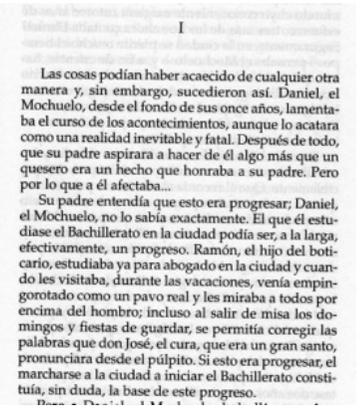
Y . . . , por supuesto, prosa

```
<div type="cap" n="01">
  <head>I</head>
  <p>
```

Las cosas podían haber sucedido de cualquier otra manera y, sin embargo, sucedieron así. Daniel, el Mochuelo, desde el fondo de sus once años, lamentaba el curso de los acontecimientos, aunque lo acatará como una realidad inevitable y fatal. Después de todo, que su padre aspirara a hacer de él algo más que un quesero era un hecho que honraba a su padre. Pero por lo que a él afectaba...

```
</p>
</div>
```

Su padre entendía que esto era progresar; Daniel, el Mochuelo, no lo sabía exactamente. El que él estudiase el Bachillerato en la ciudad podía ser, a la larga, efectivamente, un progreso. Ramón, el hijo del boticario, estudiaba ya para abogado en la ciudad, y cuando les visitaba, durante las vacaciones, venía empingorotado como un pavo real y les miraba a todos por encima del hombro; incluso al salir de guardar, se permitía corregir las palabras que don José, el cura, que era un gran santo, pronunciara desde el púlpito. Si esto era progresar, el marcharse a la ciudad a iniciar el Bachillerato, constituía, sin duda, la base de este progreso.



Miguel Delibes, *El camino*

fradejase@fyl.luva.es



Teatro . . .

```
<stage>Entra el Rey con el Conde y dos criados.</stage>
  <sp who="Rey">
    <speaker>Rey</speaker>
    <1>Esta ocasión me destierra.</1>
  </sp>
  <sp who="Conde">
    <speaker>Conde</speaker>
    <1>Y es señor muy justa ley,</1>
    <1>que la presencia de un Rey</1>
    <1>es necesaria en la guerra.</1>
    <1>Y pues el moro entre tanto</1>
    <1>con su persona pelea,</1>
    <1>es bien que la tuya sea</1>
    <1>de su atrevimiento espanto.</1>
  </sp>
  <sp who="Rey">
    <speaker>Rey</speaker>
    <1>Vino de una gran hazaña</1>
    <1>victorioso el Cordovés.</1>
  </sp>
  <sp who="Conde">
    <speaker>Rey</speaker>
    <1>Por fama sabe quien es</1>
    <1>el Emperador de España</1>
```



fradejase@fyl.luva.es



Pero . . . ¿Qué es lo que constituye los diálogos, los párrafos, los versos?

- Frases gráficamente distintas
- Cosas que parecen datos (nombres, fechas, números, lugares. . .)
- Intervenciones editoriales (correcciones, regularizaciones, adiciones, lagunas. . .)
- Referencias y enlaces
- Listas, notas, gráficos, tablas, citas bibliográficas . . .
- . . . ¡Todo tipo de anotaciones!

fradejase@fyl.luva.es



Pero . . . ¿Qué es lo que constituye los diálogos, los párrafos, los versos?

- Frases gráficamente distintas
- Cosas que parecen datos (nombres, números, lugares. . .)
- Intervenciones (correcciones, regulaciones, omisiones, lagunas. . .)
- Listas, notas, gráficos, tablas, citas bibliográficas . . .
- . . . ¡Todo tipo de anotaciones!

Qué codificar depende del objetivo de la investigación



Todo lo que acabamos de ver una forma de ver la TEI, pero hay otras maneras pues

- ❖ Es un **sistema modular**: se puede utilizar para construir el esquema de codificación **adecuado** por medio de la selección de los módulos pertinentes (véase Roma en <http://www.tei-c.org/Roma/>)
- ❖ Cada módulo define un grupo de elementos y sus atributos
- ❖ Los elementos están clasificados **estructural** y **semánticamente**

Definid los objetivos antes de usar el TEI



Todo lo que acabamos de ver una forma de ver la TEI, pero hay otras maneras pues

- ❖ Es un **sistema modular**: se puede utilizar para construir el esquema de codificación **adecuado** por medio de la selección de los módulos pertinentes (véase Roma en <http://www.tei-c.org/Roma/>)
- ❖ Cada módulo define un grupo de elementos y sus atributos
- ❖ Los elementos están clasificados **estructural** y **semánticamente**



Nombre del módulo	Capítulo
analysis	Mecanismos de análisis elementales
certainty	Certeza y responsabilidad
core	Elementos disponibles en todos los documentos TEI
corpus	Corpus lingüísticos
dictionaries	Diccionarios
drama	Teatro (puede servir para guiones cinematográficos y demás representaciones)
figuras	Tablas, fórmulas y gráficos
gaiji	Representación de caracteres no estandarizados
header	El encabezamiento del TEI
iso-fs	Estructuras de rasgos
linking	Anclaje, segmentación y alineamiento
msdescription	Descripción de manuscritos
namesdates	Nombres, fechas, gente, lugares
nets	Redes y árboles
spoken	Transcripción del habla
tagdocs	Elementos de documentación
tei	La infraestructura del TEI
textcrit	Aparatos críticos
textstructure	Estructura textual por defecto
transcr	Representación de fuentes primarias (transcripción)
verse	Verso
non-hierarchical	Estructuras no jerarquizadas



Creado por el
TEI@UVa
Seminario de **Crítica Textual**
y el
[Archivo Iberoamericano de Cetrería](#)

[Universidad de Valladolid](#)

a partir de materiales diseñados por el equipo de
[TEI@Oxford](#)

30.4.2014

bajo una licencia
Creative Commons



fradejas@fyl.uva.es

