

---

# Criterios de codificación TEI

Antonio Rojas Castro

Tobias Kraft

Grisel Terrón

Creative Commons Atribución 4.0 Internacional (CC BY 4.0)

Revision History  
23-02-2002

## Table of Contents

Metadatos .....	1
Título .....	2
Editores y colaboradores .....	2
Financiación .....	3
Edición .....	3
Publicación y licencia .....	3
Fuente textual .....	4
Descripción del documento .....	6
Autor del texto .....	6
Lugar de escritura .....	6
Fechas .....	7
Idioma .....	7
Resumen .....	7
Categorización .....	8
Texto .....	8
Estructuras textuales .....	8
Intervenciones autoriales .....	12
Intervenciones editoriales .....	14
Cartas .....	19
Remitente, destinatario, lugares y fechas .....	19
Dirección postal .....	20
Inicio de la carta .....	20
Actos de escritura .....	20
Cierre de la carta .....	20
Posdata .....	21

## Metadatos

Los metadatos utilizados en Proyecto Humboldt Digital son descriptivos e identifican

- la edición digital,
- los responsables de la edición y la publicación,
- la fuente de la que deriva el texto,
- las condiciones relacionadas con la creación y
- la clase de texto mediante categorías.

Estos metadatos han sido aplicados a todo tipo de texto (documento, carta, etc.).

## Título

Todas las ediciones se han identificado con un título principal. El título ha sido normalizado ortográficamente.

En `titleStmt`, la codificación del título principal se ha realizado utilizando el elemento `title` y el atributo `@type="main"`. Por ejemplo:

```
<title type="main">Carta de consulta al Ayuntamiento  
de La Habana</title>
```

Tras el título principal, se ha incluido de manera opcional el título en otro idioma, con la ortografía original, un título alternativo o un subtítulo. Con este fin se puede repetir el elemento `title` debajo del título principal y usar el atributo `@type="sub"`:

```
<title type="main">Carta de Francisco de Arango y Parreño  
a Alexander von Humboldt</title>  
<title type="sub" xml:lang="fr">À Monsieur Mr. le Baron  
A. Humboldt</title>
```

Se ha utilizado de manera opcional el atributo `@xml:lang` con un código estándar para indicar el idioma del título alternativo: "es" para el español, "fr" para el francés, "de" para el alemán, "en" para el inglés, etc.

## Editores y colaboradores

En el elemento `titleStmt`, se han identificado los responsables de la edición con el elemento `editor`. El nombre de la persona se ha codificado con `persName` y los siguientes elementos:

- Apellido: `surname`
- Nombre: `forename`

Si hay más de una editora o editor, se ha repetido el elemento `editor`.

Se ha codificado de manera opcional en el atributo `@ref` la referencia a una base de datos externa en forma de URL como, por ejemplo, VIAF u ORCID.

```
<editor>  
  <persName ref="http://viaf.org/viaf/21973914">  
    <surname>[Apellido]</surname>  
    <forename>[Nombre]</forename>  
  </persName>  
</editor>
```

Si en la edición han participado otras personas con tareas menores, se ha incluido el nombre de los colaboradores y colaboradoras con el elemento `respStmt` ("statement of responsibility" – declaración de responsabilidad). Nuevamente, el nombre de la persona se ha añadido en el elemento `persName`. En el atributo `@ref` también se ha codificado de manera opcional la referencia a una base de datos externa como VIAF u ORCID en forma de URL.

El tipo de colaboración se ha definido detalladamente con el elemento `resp` ("responsibility" – responsabilidad). En un elemento `note` se ha definido el tipo de contribución: colaboración, colaboración editorial, transcripción, comentario o registros. Por ejemplo:

```
<respStmt>  
  <persName ref="http://viaf.org/viaf/21973914">  
    <surname>[Apellido]</surname>  
    <forename>[Nombre]</forename>
```

```

    </persName>
    <resp>
      <note type="remarkResponsibility">colaboración</note>
    </resp>
  </respStmt>

```

## Financiación

Tras el nombre del editor o editora, se han identificado las tres entidades que financian el proyecto: Auswärtiges Amt, Fritz Thyssen Stiftung y Gerda Henkel Stiftung.

En titleStmt se ha utilizado el elemento funder de la siguiente manera:

```

<funder ref="https://www.auswaertiges-amt.de/de/">
  Auswärtiges Amt
</funder>
<funder ref="https://www.fritz-thyssen-stiftung.de/">
  Fritz Thyssen Stiftung
</funder>
<funder ref="https://www.gerda-henkel-stiftung.de/">
  Gerda Henkel Stiftung
</funder>

```

El elemento funder lleva el atributo @ref con la URL de la web de las tres entidades financiadoras.

## Edición

Todas las ediciones del proyecto mencionan el Proyecto Humboldt Digital y documentan la fecha de edición. Se ha codificado esta información sobre la edición con los elementos editionStmt y p de la siguiente manera:

```

<editionStmt>
  <p>Proyecto Humboldt Digital ( <date>2021</date> )</p>
</editionStmt>

```

La fecha entre paréntesis corresponde al año de creación del documento TEI; es importante no confundir esta fecha con la fecha de creación de la fuente de la que deriva el texto.

## Publicación y licencia

En todas las ediciones se ha codificado información sobre las organizaciones responsables de la distribución y publicación de las ediciones, los lugares de publicación y la licencia.

Para ello se ha utilizado el elemento publicationStmt. Dentro de este elemento, se han utilizado tres elementos anidados:

- publisher (entidad que publica la edición),
- pubPlace (lugar de publicación) y
- availability (disponibilidad).

Por ejemplo:

```

<publicationStmt>
  <publisher>
    <ref target="http://www.bbaw.de">Berlin-Brandenburgische
      Akademie der Wissenschaften (BBAW)</ref>

```

```

</publisher>
<publisher>Oficina del Historiador de la Ciudad de la Habana</publisher>
<pubPlace>Berlín</pubPlace>
<pubPlace>La Habana</pubPlace>
<availability status="free">
  <licence target="http://creativecommons.org/licenses/by/4.0/">
    Creative Commons Atribución
    4.0 Internacional (CC BY 4.0) </licence>
  </availability>
</publicationStmt>

```

De manera opcional, se ha enriquecido la información sobre la entidad responsable de la publicación mediante el elemento `ref` para proporcionar un hipervínculo a la web con el atributo `@target`. Asimismo, de manera opcional, se ha utilizado el atributo `@ref` en el elemento `pubPlace` para identificar el lugar de publicación con un URI procedente de GeoNames.

Los elementos `publisher` y `pubPlace` son repetibles, es decir, es posible incluir más de una entidad y lugar de publicación, pero la información sobre la disponibilidad del fichero no es repetible. Así, el elemento `availability`, con un atributo `@status="free"`, contiene a su vez un elemento `licence` con información sobre la licencia de la edición digital; se ha utilizado el atributo `@target` para dar la URL de la licencia.

En todas las ediciones del proyecto, se ha utilizado la licencia "Creative Commons Atribución 4.0 Internacional (CC BY 4.0)".

## Fuente textual

En todas las ediciones del proyecto se ha codificado la fuente de la que deriva el texto, el nombre de la institución en la que se conserva el documento y al menos un identificador como una signatura.

La descripción de la fuente se ha codificado con el elemento `msDesc` (manuscript description). Dentro se ha utilizado el elemento `msIdentifier` para proporcionar información sobre la institución que conserva el documento con el elemento `institution` y la signatura con el elemento `idno`, que lleva el atributo `@type` con el valor `shelfmark`. Se ha codificado este elemento `idno` dentro de otro elemento `idno` a fin de repetir tantos elementos `idno` como sea necesario. De manera opcional, se ha incluido información sobre la colección con el elemento `collection`. Por ejemplo:

```

<msDesc type="copy">
  <msIdentifier>
    <institution>[institución]</institution>
    <collection>[colección]</collection>
    <idno>
      <idno type="shelfmark">[signatura]</idno>
    </idno>
  </msIdentifier>
</msDesc>

```

Cuando la fuente de la que deriva el texto es un documento impreso o se ha consultado una edición impresa, se ha codificado la información en `listWit` y se ha utilizado el elemento `witness`. Dentro se ha usado un elemento `bibl` para codificar una referencia bibliográfica. En `bibl` se ha utilizado el atributo `@type` con el valor `print` para distinguirlo de un manuscrito.

Por ejemplo:

```

<witness>
  <bibl type="print">Humboldt, Alexander von: Geografía de la plantas,
    ó Quadro físico de los Andes Equinoxiales, y de los países
    vecinos. Semanario del Nuevo Reyno de Granada...</bibl>
</witness>

```

## Tipo de fuente textual

La fuente de la que deriva el texto editado puede ser de diferentes tipos:

- Manuscrito
- Copia
- Borrador
- Impreso

Cuando se utilizan más de una fuente textual, se ha seleccionado **un único** documento como base textual. Los demás testimonios textuales se han mencionado explícitamente en la edición sólo en los lugares que transmiten variantes mediante un comentario editorial.

Si la existencia y el contenido de un documento se han deducido con la ayuda de otros documentos, no hay base textual. En tal caso, en el elemento `msDesc`, se ha codificado el tipo de fuente textual con el atributo `@rend`. Para ello se ha escogido entre los siguientes valores:

Tipo de fuente textual	Valor para <code>@rend</code>
Desarrollado	<code>notExtant</code>
Borrador	<code>concept</code>
Manuscrito	<code>manuscript</code>
Copia	<code>copy</code>
Impreso	<code>print</code>

Por ejemplo:

```
<msDesc rend="manuscript">
```

## URI del registro bibliográfico

Las fuentes documentales se identifican tradicionalmente con una signatura. Dado que las signaturas solo pueden ser identificadas parcialmente por máquinas, siempre que sea posible, se ha añadido un URI si se conoce o está disponible. El URI es un identificador de recursos uniforme (del inglés “Uniform Resource Identifier”). El identificador está estandarizado y es legible por un programa informático, ya que únicamente se compone de ciertos símbolos (por ejemplo, los espacios en blanco no están permitidos).

Siempre que ha sido posible se ha utilizado un URI para conectar el texto editado con el registro bibliográfico del repositorio digital del proyecto.

Nombre del elemento	Punto de inserción	Atributo	Contenido del elemento
<code>idno</code>	// <code>msIdentifier/idno/idno</code>	<code>@type="uri"</code>	[URI]

Se ha utilizado el elemento `idno` como parte del elemento `msIdentifier` y como elemento hijo de otro `idno` utilizando el atributo `@type` con valor `"uri"` para los enlaces permanentes o bien `"URLImages"` para los facsímiles. Por ejemplo:

```
<idno>
  <idno type="shelfmark">Caja pequeña 7b, no. 68</idno>
  <idno type="uri">http://kalliope-verbund.info/DE-611-HS-1272345</idno>
  <idno type="URLImages">
    http://resolver.staatsbibliothek-berlin.de/SBB00015CC500000000</idno>
```

</idno>

## Descripción del documento

De manera opcional, se han descrito las características físicas (papel, formato, número de páginas, etc.) de la fuente textual de la que deriva la edición.

La descripción del documento se ha codificado con el elemento `physDesc` (“physical description”, en español: descripción física). La descripción ha sido proporcionada como texto en un elemento `p`:

<physDesc>

<p>Manuscrito autógrafo de autor desconocido en buen estado de conservación y de seis páginas de extensión. Forma parte del documento conocido como &quot;Umschlag mit den Materialien für die Statistik von Kuba [Werk]&quot;. Doble numeración: una paginación original en tinta; otra moderna en lápiz que solo numera los folios (123-124). Contiene algunas adiciones en tinta negra al inicio del documento, en los márgenes y al final del texto en francés por Alexander von Humboldt.</p>

</physDesc>

## Autor del texto

Si se conoce quién es autor del texto, esta información se ha codificado en el elemento `profileDesc`. Allí, se ha creado un elemento `creation` donde se han codificado los datos sobre la creación: autor, lugar de escritura y fecha de creación.

El nombre del autor se ha codificado con el elemento `persName`. En caso de seleccionar el autor del registro de personas del proyecto, el identificador asociado se ha generado en el atributo `@key`, creando un enlace al registro.

<profileDesc>

<creation>

<persName key="H0012682">

Arango y Parreño, Francisco de (1765-1837)</persName>

</creation>

</profileDesc>

Se ha respetado siempre el siguiente orden: primero el apellido y luego el nombre, separados por una coma. Si se conoce la fecha de nacimiento y fallecimiento, se ha incluido seguida del nombre y entre paréntesis como se muestra en el ejemplo anterior.

Cuando el autor es desconocido, se ha empleado el mismo elemento `persName`, pero con el contenido "N.N." (*Nomen nominandum*). Por ejemplo:

<creation>

<persName>N.N.</persName>

<persName key="H0012069">Humboldt, Alexander von (1769-1859)</persName>

<date when="1804-04-10">10 de abril de 1804</date>

<placeName key="https://www.geonames.org/3553478">La Habana</placeName>

</creation>

Siempre que han participado más de una persona en la creación del texto se ha repetido el elemento `persName`.

## Lugar de escritura

Si se conoce el lugar de escritura, se ha codificado también en el elemento `profileDesc`. Allí, si aún no existiera uno, se ha creado un elemento `creation`. Dentro de este elemento, se ha utilizado

placeName para codificar el nombre del lugar. Si el lugar de escritura figura en el registro de lugares del proyecto, el identificador asociado se ha incluido en el atributo @key. Por ejemplo:

```
<placeName key="H0006819">La Habana</placeName>
```

## Fechas

Si se conoce la fecha de creación, se ha codificado con el elemento date respetando el formato AAAA-MM-DD (o AA-MM o AAAA, dependiendo de la información disponible) y utilizando los atributos @when, @from/@to o @notBefore/@notAfter. Esto permite que la aplicación web utilice estos datos para buscar información en un periodo concreto de tiempo o filtrar los documentos a partir de la fecha estandarizada.

- Fecha exacta:

```
<date when="1772-12-21">tres días antes  
de la Nochebuena del año 1772</date>
```

- Período de tiempo exacto:

```
<date from="1731-11-11" to="1741-09-11">del 11 de noviembre  
de 1731 al 11 de septiembre de 1741</date>
```

- Período aproximado:

```
<date notBefore="1804-05-01" notAfter="1804-05-16">Primera mitad  
de mayo de 1804</date>
```

Además, de manera opcional, se ha utilizado el atributo @cert con los valores high y low para codificar la certeza con que se sabe si la fecha es correcta.

## Idioma

En todas las ediciones del proyecto, se ha codificado el idioma principal del texto. En profileDesc, se ha utilizado el elemento langUsage. Allí se ha anidado el elemento language para codificar el idioma o lengua. Por ejemplo:

```
<langUsage>  
  <language ident="es">Español</language>  
</langUsage>
```

Se ha utilizado el atributo @ident para identificar con un código estandarizado el idioma: "es" para el español, "de" para el alemán, "en" para el inglés, "fr" para el francés, etc.

Si el texto contiene partes sustanciales en otro idioma, se ha repetido elemento language para codificar la lengua secundaria.

## Resumen

En todas las ediciones del proyecto, se ha proporcionado un resumen del contenido intelectual del texto, los temas, las ideas discutidas, etc.

El resumen se ha codificado con profileDesc/abstract en un elemento p de la siguiente manera:

```
<abstract xml:lang="es">  
  <p>Notas de Humboldt sobre la obra de  
  Jean Nicolas Berthe "Précis Historique
```

de la Maladie qui a régné dans l'Andalousie en 1800".  
Humboldt resume las tesis del trabajo de Berthe en  
preparación de su tratado científico sobre la fiebre amarilla.  
La obra de Berthe fue publicada en 1802.</p>

</abstract>

Se ha proporcionado el resumen en español y en alemán. Se ha repetido el elemento `abstract` y se ha utilizado el atributo `@xml:lang` con un código estandarizado de dos caracteres: "es" para el español y "de" para el alemán.

## Categorización

En `profileDesc` se ha utilizado el elemento `textClass` para incluir distintas categorías con las que clasificar el texto editado. Las categorías utilizadas son: género, forma y materia(s). Para ello, se ha utilizado el elemento `catRef` con los atributos `@scheme` y `@target`. Por ejemplo:

```
<textClass>
  <catRef scheme="#genre" target="#notebooks"/>
  <catRef scheme="#form" target="#manuscript"/>
  <catRef scheme="#subject" target="#slavery"/>
  <catRef scheme="#subject" target="#geography"/>
  <catRef scheme="#subject" target="#sugar"/>
</textClass>
```

Las tres categorías son repetibles y opcionales. Todas las categorías se han definido previamente en la taxonomía del proyecto situada en el elemento `taxonomy`.

En la medida de lo posible, se han utilizado términos procedentes de la Library of Congress y de otras bibliotecas nacionales como la Deutsche Nationalbibliothek y la Biblioteca Nacional de España.

## Texto

Los textos editados como parte del proyecto han sido codificados desde un punto de vista estructural. Se han representado de manera explícita con etiquetado TEI todos los párrafos, rupturas de línea, listas, tablas, notas y dibujos. Asimismo, se han codificado la paginación o foliación de la fuente, incluyendo las páginas en blanco, y los cambios en la tipografía como los destacados, los espacios dejados por el autor, las intervenciones autoriales como las eliminaciones, adiciones y sustituciones, y las intervenciones editoriales como las normalizaciones, la resolución de las abreviaturas, las palabras suplidas, las correcciones de errores evidentes, las lecturas inciertas, las omisiones y las palabras en otro idioma. Por último, en la medida de lo posible, se han codificado también las fechas y las entidades nombradas (personas, lugares y organizaciones) en el cuerpo del texto.

## Estructuras textuales

### Párrafos

En todas las ediciones del proyecto se han codificado los párrafos siguiendo la secuencia de la fuente textual (sea un manuscrito o un impreso). De manera opcional se ha codificado el sangrado de los párrafos si la editora o editor considera dicha información relevante.

Los párrafos se han codificado con el elemento `p` (paragraph).

### Rupturas de línea

Las rupturas de línea se han codificado de manera explícita siguiendo la fuente textual con un elemento (vacío) `lb/` (line beginning). El elemento `lb/` se ha utilizado al **comienzo** de la línea, tal y como determinan las directrices TEI.



Si el salto de línea ocurre dentro de una palabra, se ha codificado con el atributo `@break` y el valor `no`. En caso de aparecer un separador en la fuente original, no se ha transcrito como tal sino que se ha codificado con el elemento `g` y el atributo `@ref="#typoHyphen"`. Por ejemplo:

y le añadí al propio `ti` `<g ref="#typoHyphen"/><lb/>empo`

## Foliación y paginación

En todas las ediciones del proyecto se ha codificado el inicio de la foliación o paginación, es decir, el lugar en que se inicia un folio nuevo o página nueva. Siempre se ha codificado el principio del folio o página, no el final.

El inicio de folio o página se ha codificado siempre con el elemento vacío `pb/` (pagebreak). El atributo `@n` contiene el número de página o folio (según la secuencia física del documento o del autor).

distinta de las demás `<pb n="3v"/>` colonias del continente Americano

La digitalización del folio o página se ha codificado con el atributo `@facs` (facsimile). Se ha codificado el nombre del fichero de la imagen digital incluyendo la foliación archivística y la extensión "tif". El nombre del fichero corresponde a un URI, es decir, es único en el dossier digital. Por ejemplo:

`<n="1r" facs="prohd0014_1r.tif"/>`

Cualquier otra información sobre la numeración o paginación contenida en la fuente se ha codificado con el elemento `fw` (forme work); este elemento lleva el atributo `@type` con el valor `folNum`.

`<fw type="folNum">34</fw>`

## Páginas en blanco

En todas las ediciones del proyecto se han codificado las páginas en blanco de la fuente textual. Además, se ha distinguido entre páginas en blanco, páginas perdidas, espacios en blanco semánticamente significativos y omisiones de los editores.

Para poder representar y mostrar la foliación correctamente, se ha codificado cada página (en blanco) individualmente. Aunque se sucedan varias páginas en blanco, siempre se ha introducido un número "1" para cada nueva página en blanco codificada.

Las páginas en blanco se han codificado con el elemento (vacío) `gap/`. Este se ha utilizado justo después del salto de página (`pb/`) anterior. Los dos atributos `@unit` y `@quantity` del elemento `gap/` llevan los valores `pages` y `1`. El atributo `@reason` contiene el valor `empty`. Cuando hay más de una página en blanco seguida, se han codificado cada una de ellas individualmente.

`<pb n="132v"/><gap unit="pages" quantity="1" reason="empty"/>`

## Espacios en blanco

El espacio en blanco semánticamente significativo (por ejemplo, a veces los autores dejan un espacio en el texto para complementarlo posteriormente) se ha codificado con el elemento vacío `space`. La reproducción de dichos vacíos en la transcripción se ha distinguido de las omisiones realizadas por los editores, así como de las páginas vacías o perdidas, cada una de las cuales se ha codificado como `gap/`.

El alcance de la omisión se ha indicado mediante `@unit` y `@quantity`. Si el número de líneas (`unit="lines"`), caracteres (`unit="chars"`) o palabras (`unit="words"`) omitidos no se puede determinar de manera segura, se ha utilizado el valor `1` en `@quantity`. Por ejemplo:

sur une longueur de plus  
de `<space quantity="1" unit="words"/>`. Sa largeur...

De manera opcional, se ha utilizado el atributo @dim para indicar si se trata de un espacio en blanco horizontal o vertical.

## Listas

Las listas se han codificado con el elemento `list`. Los puntos o líneas individuales se han transcrito en un elemento `item`. La numeración existente (también con letras) no se ha codificado de manera independiente, sino que se ha transcrito dentro del elemento `item`.

```
<list>
  <item>Le Rio Apure près de S. Fernando 206 t.</item>
  <item>Orinoco à la bocca Apure 1905 t.<lb/>
  mais quand le Capucin est islado = 10750 mètres.</item>
  <item>Orinoco près du Baraguan 889 t.</item>
  <item>Orinoco près de urbano 2674 t. <lb/>
  jusqu'au delà de l'Isle de Urba<g ref="#typoHyphen"/>
  <lb break="no"/>na et de I. vieja de Montica.</item>
</list>
```

## Tablas

Las tablas contenidas en la fuente se han transcrito como tal, haciendo distinción entre tablas explícitas e implícitas. Se ha considerado como tablas explícitas aquellas dibujadas por el autor, mientras que las implícitas son reconocibles sólo por la disposición de sus contenidos. Se ha distinguido entre tablas y enumeraciones que se asemejan a listas.

Las tablas se han codificado con el elemento `table`. Se ha asignado a la primera fila de la tabla el atributo @role="label" para que se considere un encabezado. Por ejemplo:

```
<table rows="3" cols="2">
  <head>Tabla</head>
  <row role="label">
    <cell>[Encabezado columna 1]</cell>
    <cell>[Encabezado columna 2]</cell>
  </row>
  <row>
    <cell>[Columna 1/ Fila 1]</cell>
    <cell>[Columna 2/ Fila 1]</cell>
  </row> ...
</table>
```

## Alineación de texto

En las tablas y en el inicio de una carta, se ha codificado la alineación del texto de manera opcional. De este modo, las fechas o las celdas de la tabla se alinean a la derecha. Para ello, se ha utilizado el atributo @rendition con el valor #right en el elemento `cell` (celda de tabla) o `dateline`. Por ejemplo:

```
<table cols="4" rows="20" rend="drawn">
  <head>Océan entre l'europe, l'afrique et l'amérique</head>
  <row role="label">
    <cell>Latitude <lb/>boréale.</cell>
    <cell>Longitude du<lb/>méridien<lb/>de Paris.</cell>
  </row>
  <row>
    <cell rendition="#right">43°. 29'</cell>
    <cell rendition="#right">10°. 31'."</cell>
  </row>
```

```
<!--[...]-->
</table>
```

## Notas

Las notas del autor o de un editor contemporáneo no se han insertado siguiendo su ubicación en el manuscrito, sino en el pasaje del texto al que la nota se refiere realmente (o que el editor supone razonablemente que se refiere). En caso de emplear una marca de inserción (asterisco, número, etc.), se ha utilizado una señal de inserción general (es decir, el elemento vacío `metamark/`) para indicar que el texto ha sido introducido por el autor y no por los editores.

Las notas se han codificado con el elemento `note`:

En los partidos más remotos una caballería  
`<note place="mBottom" hand="#author">`La caballería es una superficie de  
 tierra de quatrocientas treinta y dos varas en quadro.`</note>` de tierra  
 montuosa se paga en el día a quinientos pesos fuertes al contado,[...]

Cuando es posible identificar al autor de la nota, se ha codificado esta información con el atributo `@hand`:

Valor	Definición
#author	Autor del texto
#addressee	Destinatario de esta carta
#unknown	Desconocido
H000000	ID interno proveniente del registro de personas

El lugar de la nota se ha codificado con el atributo `@place`:

Valor	Definición
left	margen izquierdo
right	margen derecho
mTop	parte superior
mBottom	parte inferior
inline	dentro de la línea

Si el autor utiliza una marca de inserción, se ha codificado con el elemento vacío `metamark/`. La apariencia de la marca de inserción (asterisco, número, etc.) no se ha hecho explícita.

Se ha diferenciado entre las notas y las adiciones: las adiciones complementan el texto principal y se integran sintácticamente en él. Las notas son unidades de texto independientes.

## Dibujos

Los dibujos e ilustraciones se han codificado con el elemento `figure`. Cuando ha sido posible identificar la persona que realizó el dibujo, se ha utilizado el atributo `@hand` con los siguientes valores: #author, #addressee, #unknown o una persona del registro del proyecto. En caso de que se disponga de un título en el documento original, se ha codificado con el elemento `p`. Por ejemplo:

```
<figure facs="dibujo" hand="#addressee">
  <p>[Título de la figura]</p>
  <figDesc>[Descripción de la imagen por el editor]</figDesc>
  <note type="editorial">[Comentario editorial]</note>
</figure>
```

El nombre del fichero se ha codificado con el atributo `@facs`. Los comentarios editoriales que se refieren a la imagen en sí se han codificado con el elemento hijo `note`.

De manera opcional, en el elemento `figDesc` las y los editores han descrito la imagen.

## Intervenciones autoriales

En todas las ediciones digitales del proyecto se han codificado los cambios de tipografía y las intervenciones autoriales de manera explícita, es decir, con etiquetado TEI.

## Destacados

Los destacados y marcas en el borde del documento se han codificado con el elemento `hi` (highlighted) y el atributo `@rendition` para especificar su apariencia:

Tipo de destacado	Valor para <code>@rendition</code>
simple	<code>#u</code>
doble	<code>#uu</code>
marcado al margen	<code>#mMM</code>

Por ejemplo:

Sucinta noticia de la situación  
`<hi rendition="#uu">presente de esta colonia. Agosto de 800'</hi>`

Si un párrafo completo aparece en la fuente destacado con una marca en el margen, se ha codificado con el elemento `p` y el atributo `@rendition` con el valor `#mMM` (marca al borde).

## Superscript

Los números o letras que aparecen volados (es decir, ligeramente elevados respecto a la línea) en la fuente textual se han codificado con el elemento `hi` (highlighted) y el atributo `@rendition` con valor `#sup`. Por ejemplo:

le ofreciera de mi parte  
 algunas observaciones `<choice>`  
     `<abbr>sob<hi rendition="#sup">e</hi></abbr>`  
     `<expan>sobre</expan>`  
`</choice>` los `<lb/>` principales hechos

## Eliminaciones

Siempre que sea legible y con independencia de si se encuentra en saltos de línea, abreviaturas, normalizaciones y correcciones o adiciones del editor, se ha codificado el texto eliminado con el elemento `del` (deletion) y haciendo uso del atributo `@rendition` para especificar el tipo de eliminación:

Tipo de eliminación	Valor para <code>@rendition</code>
reescritura encima del texto original	<code>#ow</code>
tachado	<code>#s</code>
borrado	<code>#erased</code>

Por ejemplo:

Conocen ya estos vecinos la esplendidez en la mesa, las viviendas, muebles, `<del rendition="#s">y</del>` vestidos y tren de calle.

## Adiciones

En todas las ediciones del proyecto se han codificado las adiciones que el propio autor ha realizado en la fuente textual. Se entiende por adiciones todos los añadidos realizados por el autor que sólo amplían el texto y se integran sintácticamente.

Las adiciones hechas por el autor se han codificado con `add` (addition) utilizando el atributo `@place` para especificar el lugar de la adición:

Valor	Definición
superlinear	encima de la línea
sublinear	debajo de la línea
intralinear	en la misma línea
across	sobre el texto original
left	en el margen izquierdo
right	en el margen derecho
mTop	en el margen superior
mBottom	en el margen inferior

Por ejemplo:

[...] y logré para el Comercio con extranjeros,  
los `<add place="intralinear">ensanches</add>` que se  
explican en el documento [...]

Si hay una marca de inserción hecha por el autor en la fuente, se ha codificado con el elemento `meta-mark/`. La forma de la marca de inserción (asterisco, número, etc.) no se ha especificado.

## Sustituciones

Si se elimina una palabra o un pasaje del texto en la fuente y al mismo tiempo se añade una palabra, estos dos fenómenos textuales se han codificado juntos con un elemento `subst` (substitution).

La eliminación se ha codificado con el elemento `del` (deletion). El tipo de eliminación se ha especificado en el atributo `@rendition`:

Tipo de eliminación	Valor para <code>@rendition</code>
reescritura encima del texto original	#ow
tachado	#s
borrado	#erased

La adición se ha codificado con el elemento `add` (addition). El lugar de la adición se ha especificado con el atributo `@place`:

Tipo de adición	Valor para <code>@place</code>
encima de la línea	superlinear
debajo de la línea	sublinear
en la misma línea	intralinear
sobre el texto original	across
en el margen izquierdo	left
en el margen derecho	right

Por ejemplo:

```
<subst>
  <del rendition="#ow">6</del>
  <add place="across">7</add>
</subst>, 4 pesos
```

## Intervenciones editoriales

En todas las ediciones digitales del proyecto se han codificado las intervenciones editoriales de manera explícita.

## Normalizaciones

La ortografía, puntuación, mayúsculas y separación de las palabras se han normalizado de manera explícita con el elemento `choice` y sus dos elementos hijos `orig` (forma original) y `reg` (forma regularizada). En el elemento `orig` se ha codificado la forma original tal y como aparece en la fuente textual mientras que en el elemento `reg` se ha codificado la forma regularizada, es decir, modernizada a la normal ortográfica actual para facilitar la legibilidad del texto. Por ejemplo:

```
<choice>
  <orig>espendiendo</orig>
  <reg>expendiendo</reg>
</choice>
```

Siempre se han codificado las palabras enteras aun cuando solamente se regularice una grafía o una tilde tal y como se muestra en el ejemplo anterior.

Se ha seguido el mismo método de codificación con la normalización de la mayúscula. Por ejemplo:

```
<choice>
  <orig>ensayo político</orig>
  <reg>Ensayo Político</reg>
</choice>
```

El mismo método de normalización se ha aplicado a contracciones o uniones de palabras. Por ejemplo:

```
<choice>
  <orig>dela</orig>
  <reg>de la</reg>
</choice>
```

## Abreviaturas

Las abreviaturas se han codificado con elemento `choice` y sus dos elementos hijo `abbr` (abreviatura) y `expan` (abreviatura resuelta). En el elemento `abbr` se ha conservado la forma abreviada que aparece en la fuente textual mientras que en el elemento `expan` las y los editoras han proporcionado la forma resuelta o expandida para facilitar la legibilidad del texto.

```
<choice>
  <abbr>q</abbr>
  <expan>que</expan>
</choice>
```

Las siglas se han codificado como abreviaturas siguiendo el mismo método con `choice`.

## Palabras, letras o signos de puntuación suplidos

Las palabras o signos de puntuación que faltan de manera evidente en la fuente y que han sido suplidas por la editora o editor, se han codificado con el elemento `supplied`. Con el atributo `@cert` (certainty) se ha codificado la certeza con la que se han suplido los caracteres o palabras:

Valor	Definición
low	baja
high	alta

Por ejemplo:

30 de julio de `<supplied cert="high">18</supplied>`27

## Errores y correcciones

Los errores evidentes que aparecen en la fuente textual se han codificado con el elemento `choice`. Dentro, la palabra errónea se ha codificado con el elemento `sic` mientras que la corrección se ha codificado con `corr`. En el elemento `corr`, el atributo `@cert` (certainty) indica la certeza con que se ha corregido:

Certeza	Valor para @cert
baja	low
media	medium
alta	high

Por ejemplo:

```
<choice>
  <sic>Humboldt</sic>
  <corr cert="high">Humboldt</corr>
</choice>
```

Cuando una palabra aparece repetida y es evidente que se trata de un error, se ha empleado el mismo elemento `choice` para codificar la palabra errónea con `sic`. Ahora bien, en estos casos no se ha proporcionado una forma corregida sino que se ha utilizado el elemento `corr` y se ha dejado vacío. Además, se ha utilizado el atributo `type` con valor "deleted" para indicar que se ha eliminado una palabra redundante. Por ejemplo:

```
se trabaja aquí más
<choice>
  <sic>más</sic>
  <corr type="deleted"/>
</choice> cantidad
```

## Lecturas inciertas

Las palabras o pasajes que no son fácilmente legibles se han codificado con el elemento `unclear`. Con el atributo `@reason` se ha especificado la razón de la lectura incierta:

Razón de la lectura incierta	Valor para @reason
ilegible	illegible
difícil de leer debido a la sobreescritura	covered

Con el atributo `@cert` (certainty) se ha codificado la certeza de la intervención.

Certeza	Valor para @cert
baja	low
alta	high

Por ejemplo:

```
<unclear reason="illegible" cert="high">año</unclear>
```

## Texto ilegible

Los caracteres o palabras ilegibles en la fuente textual debido a daños materiales o a que la escritura es poco clara se han codificado con el elemento vacío `gap/`.

Con el atributo `@reason` se ha indicado la razón de la ilegibilidad:

Razón de la ilegibilidad	Valor para <code>@reason</code>
destrucción/pérdida	lost
ilegible	illegible
difícil de leer debido a la sobreescritura	covered

Si se conoce el número de líneas, palabras, etc., se han codificado usando `@unit` (unidad) y `@quantity` (cantidad).

Segmento ilegible	Valor para <code>@unit</code>
caracteres	chars
líneas	lines
páginas	pages
palabras	words

En el atributo `@quantity`, se ha codificado el número de caracteres, líneas, etc. ilegibles.

```
<gap reason="illegible" unit="words" quantity="3"/>
```

Si un carácter (o cadena de caracteres) está tachado y no se puede descifrar, se ha codificado tanto con `del` como con `gap/`:

```
<del rendition="#s">
  <gap reason="illegible" unit="chars" quantity="5"/>
</del>
```

Por último, también se han codificado individualmente las páginas destruidas o perdidas.

```
<pb n="1r"/><gap reason="lost" unit="pages" quantity="1"/>
```

## Omisiones

Las omisiones se han codificado de manera explícita. Las omisiones de caracteres, palabras, pasajes de texto o páginas enteras por parte de los editores y editoras se han codificado con el elemento vacío `gap/`. Con el atributo `@unit` se ha especificado el tipo de unidad que la editora o editor ha tenido en cuenta:

Segmento omitido	Valor para <code>@unit</code>
caracteres	chars
líneas	lines
páginas	pages
palabras	words

Con el atributo `@quantity`, se ha codificado el número de caracteres, líneas, etc. omitidos mientras que con el atributo `@reason` se ha codificado el motivo de la omisión editorial.

Motivo de la omisión	Valor para <code>@reason</code>
texto en idioma extranjero	fm
el texto no es importante para el corpus	insignificant



Por ejemplo:

a menudo `<gap unit="words" quantity="1" reason="fm"/>` entonces de nuevo

Para representar y mostrar correctamente la foliación, se ha codificado cada página individualmente. Aunque se omitan varias páginas consecutivas, siempre se ha codificado la omisión con un “1” como número correspondiente a cada página que falta.

## Comentarios

Los comentarios editoriales se han codificado con el elemento `seg` (arbitrary segment). A su vez, se han anidado dos elementos: por un lado, `orig` con el que se ha codificado el pasaje textual comentado; por el otro lado, `note` con el que se ha codificado el comentario editorial en sí. Se ha utilizado el atributo `@type` en `seg` con el valor `comment`.

```
<seg type="comment">
  <orig>Discurso del héroe</orig>
  <note>Comentario irónico que alude a...</note>
</seg>
```

## Referencias

Las referencias a otras ediciones o pasajes internos se han codificado con el elemento `ref`, cuyo atributo `@target` contiene el identificador único de la edición o del pasaje de destino. Por ejemplo:

```
<ref target="BTEST/#start_b3k_pyf_cy">[una referencia]</ref>
```

Además, en algunas ocasiones, con `ref` se han codificado referencias a sitios web externos, por ejemplo, en algún comentario editorial. En este caso, se ha codificado en el atributo `@target` la URL completa del sitio web:

```
<seg type="comment">
  <orig>el río</orig>
  <note>Se refiere al río...</note>
</seg>
```

## Fechas

Las fechas que aparecen en la fuente textual se han codificado con `date`. Para que las fechas sean legibles por el ordenador se ha seguido el formato AAAA-MM-DD en los atributos `@when`, `@from` y `@to`, así como `@notBefore` y `@notAfter`.

- Fecha exacta:

```
<date when="1772-12-21">tres días antes de la
Nochebuena del año 1772</date>
```

- Período de tiempo exacto:

```
<date from="1731-11-11" to="1741-09-11">del 11 de noviembre de 1731
al 11 de septiembre de 1741</date>
```

- Período aproximado:

```
<date notBefore="1804-05-01" notAfter="1804-05-16">Primera mitad
de mayo de 1804</date>
```

Además, de manera opcional, en el atributo `@cert` se ha especificado el nivel de certeza: `high`, `low`.

Más ejemplos:

Fecha en el texto	Valores de los atributos
siglo XVIII/18	notBefore="1700" notAfter="1799"
principios del siglo XVIII/18	notBefore="1700" notAfter="1719"
mediados del siglo XVIII/18	notBefore="1740" notAfter="1759"
finales del siglo XVIII/18	notBefore="1780" notAfter="1799"
Primera mitad del siglo XVIII/18	notBefore="1700" notAfter="1749"
Segunda mitad del siglo XVIII/18	notBefore="1750" notAfter="1799"
Primer cuarto del siglo XVIII/18	notBefore="1700" notAfter="1724"
Segundo cuarto del siglo XVIII/18	notBefore="1725" notAfter="1749"
Tercer cuarto del siglo XVIII/18	notBefore="1750" notAfter="1774"
Último cuarto del siglo XVIII/18	notBefore="1775" notAfter="1799"
Primer tercio del siglo XVIII/18	notBefore="1700" notAfter="1733"
Segundo tercio del siglo XVIII/18	notBefore="1733" notAfter="1766"
Último tercio del siglo XVIII/18	notBefore="1766" notAfter="1799"
alrededor de 1762	notBefore="1757" notAfter="1767"
1762	when="1762"
5 de enero de 1762	when="1762-01-05"

En el contenido del elemento también se ha expresado el nivel de certeza, aproximación, período, etc. mediante lenguaje natural. Por ejemplo:

`<date notBefore="1802">No antes de 1802</date>`

## Palabras en otro idioma

Las palabras y pasajes de texto en idioma extranjero se han codificado con el elemento `foreign`. Con el atributo `@xml:lang` se ha codificado el código del idioma extranjero correspondiente:

Idioma extranjero	Valor para <code>@xml:lang</code>
griego	gr
latín	la
hebreo	he
alemán	de
francés	fr
inglés	en
ruso	ru
español	es

Por ejemplo:

cantidad de 29.000 arrobas, `<foreign xml:lang="la">plus minus</foreign>`

## Personas, lugares y organizaciones

Las personas, los lugares y las organizaciones mencionados con nombres propios en la fuente textual se han codificado de manera explícita. En cambio, la codificación de alusiones o referencias (por ejemplo, anáforas) se ha codificado de manera opcional según el criterio del editor o editora.

En ambos casos se han utilizado los elementos siguientes:

- Persona: persName
- Lugar: placeName
- Organización: orgName

Se ha codificado con el atributo @key un identificador único (URI) correspondiente al del registro del proyecto o bien al de algún recurso de autoridades externo como VIAF, GND o GeoNames. Por ejemplo:

```
Mr. <persName key="H0012737">James Arnold</persName>
```

## Cartas

En este proyecto las cartas han recibido un tratamiento especial debido a sus rasgos característicos. En lo que respecta a los metadatos, siempre que ha sido posible, se han codificado el remitente, el destinatario, el lugar y la fecha de envío y recepción. Por lo que respecta a la estructura, cuando están presente en la fuente, se han codificado la dirección postal, el inicio de la carta, los actos de escritura (si hubiera interrupciones) y el cierre de la carta.

## Remitente, destinatario, lugares y fechas

Los metadatos relativos a la correspondencia de una carta se han codificado en el elemento `correspDesc` (correspondence description). Se ha creado un elemento `correspAction` (correspondence action) para el proceso de envío y el de recepción respectivamente. Con el atributo @type se ha especificado el tipo del acto de comunicación; los valores permitidos son dos: `sent` o `received`.

```
<correspDesc>
  <correspAction type="sent">[Remitente]</correspAction>
  <correspAction type="received">[Destinatario]</correspAction>
</correspDesc>
```

Dentro del elemento `correspAction`, se han utilizado los elementos `persName`, `orgName`, `placeName` y `date` para identificar la persona u organización que participa en el intercambio, el lugar en que sucede y la fecha. Siempre que ha sido posible, la información sobre personas, organizaciones y lugares se ha vinculado con el identificador codificado en el atributo @key. Esta información puede indicarse para todas las entidades nombradas.

```
<correspDesc>
  <correspAction type="sent">
    <persName key="H0012682">Arango y Parreño, Francisco de</persName>
    <placeName key="H0006819">La Habana</placeName>
    <date when="1827-07-30">30 de julio de 1827</date>
  </correspAction>
  <correspAction type="received">
    <persName key="H0012069">Humboldt, Alexander von</persName>
  </correspAction>
</correspDesc>
```

Las fechas se han codificado con el elemento `date` siguiendo el formato AAAA-MM-DD (o AA-MM o AAAA, dependiendo de la información disponible) en los atributos @when, @from/@to o @not-Before/@notAfter. Además, en el atributo @cert se han utilizado los valores `high` y `low` para indicar la certeza sobre la fecha. Por ejemplo:

- Fecha exacta:

```
<date when="1772-12-21">tres días antes de
la Nochebuena del año 1772</date>
```

- Período de tiempo exacto:

```
<date from="1731-11-11" to="1741-09-11">del 11 de noviembre de  
1731 al 11 de septiembre de 1741</date>
```

- Período aproximado:

```
<date notBefore="1804-05-01" notAfter="1804-05-16">Primera mitad  
de mayo de 1804</date>
```

Cuando no se ha podido determinar las fechas de envío o recepción, se ha codificado una nota explicativa en `correspDesc` con el elemento hijo `note`. Por ejemplo:

```
<correspDesc> ...  
  <note>No se puede determinar la fecha de recepción.</note>  
</correspDesc>
```

## Dirección postal

La dirección postal de una carta se ha codificado en el elemento `opener` (inicio de la carta) o en el `closer` (cierre de la carta). Consta del elemento `address` y uno o varios elementos hijos `addrLine`.

```
<opener>  
  <address>  
    <addrLine>Berlin, Kronenstraße 6</addrLine>  
  </address>  
</opener>
```

## Inicio de la carta

El inicio de la carta se ha codificado con el elemento `opener`. Dentro, si los hubiera, se ha codificado la fecha y el saludo con los elementos `dateline` y `salute` respectivamente.

En el elemento `dateline` se ha utilizado el atributo `@rend` si la fecha está sangrada o alineada a la derecha. Por ejemplo:

```
<opener>  
  <dateline rend="align(right)">30 de julio de 1827</dateline>  
  <salute>Mi muy apreciable amigo y señor:</salute>  
</opener>
```

## Actos de escritura

Los actos de escritura se han codificado con el elemento `div` con dos atributos: por un lado, el atributo `@type` debe llevar el valor `writingSession`; por el otro, `@n` debe indicar el número del acto de escritura.

```
<div type="writingSession">  
  <p>[Primer acto de escritura]</p>  
</div>  
<div type="writingSession" n="2">  
  <p>[Segundo acto de escritura]</p>  
</div>
```

## Cierre de la carta

El cierre de la carta se ha codificado en el elemento `closer`. La fórmula de saludo se ha codificado con el elemento `salute` mientras que la firma se ha codificado con el elemento `signed`. Con el

atributo @rend del elemento closer se ha representado la forma o apariencia con que se realizó el cierre de la carta.

Cierre de carta	Valor para @rend
párrafo	paragraph
dentro de una línea	within the line

Por ejemplo:

```
<closer rend="paragraph">
  <salute>Muchos saludos</salute>
  <signed>F. E.</signed>
</closer>
```

Si hubiera una fecha en el cierre de la carta, se ha codificado en el elemento closer con dateline.

```
<closer rend="paragraph">
  <dateline>4 de enero de 1886</dateline>
  <!-- [...] -->
</closer>
```

## Posdata

La posdata se ha codificado después del closer con el elemento postscript. La posdata siempre contiene al menos un párrafo (p).

```
<postscript>
  <p>Para Barón von Humboldt</p>
</postscript>
```