

# Directrices para entidades de nombres sencillos

para lenguas poco enseñadas  
Versión 6.5 - 28 de marzo del 2006

Consortio de datos lingüísticos –Equipo LCTL  
(Basado prácticamente en las directrices MUC-7 NE)

1. Introducción .....	1
2. Tipos de entidades .....	3
2.1. Nombres de personas .....	3
2.1.1. Nombres de familia/Apellidos .....	3
2.2. Títulos, rol y apositivos .....	4
2.2.1 Títulos que contienen otras Entidades .....	4
2.2.2. Nombres separados por títulos .....	5
2.2.3 Secuencias de títulos .....	5
2.2.4 Sufijos ordinales .....	5
2.2.5 Títulos que no se marcan .....	6
2.3. Nombres de organizaciones .....	6
2.3.1 Organizaciones y “Anotación por significado” .....	7
2.3.2 Artículos en entidades organizacionales .....	7
2.3.3 Organizaciones que no se anotan –Referencias genéricas .....	8
2.4 Nombres de lugar .....	8
2.4.1 Lugares y “Anotación por significado” .....	9
2.4.2 Extensión de los nombres de lugar .....	9
2.4.2.1 Artículos en las entidades de lugar .....	9
2.5. Decidiendo entre tipos de entidad .....	11
2.5.1 ORG que se refiere a LUG, LUG que se refiere a ORG.....	11
3. Casos difíciles .....	12
3.1. Expresiones que se refieren a entidades múltiples .....	12
3.2. Expresiones anidadas .....	13
3.3 Entidades en posición de premodificador (Nombres adjetivo).....	13
3.4. Posesivos .....	14
3.5. Otros tipos de nombres .....	14
4. Qué NO anotar .....	15
4.1 Eventos .....	15
4.2. Artefactos y productos .....	15
4.3 Genéricos .....	15
5. Inseguridad en la anotación .....	15

## 1. Introducción

Una entidad es un objeto en el mundo: por ejemplo, un lugar o una persona. Un nombre de entidad es una frase que se refiere de manera única a ese objeto por

su nombre propio, acrónimo, apodo o abreviación. Algunos ejemplos de nombres de identidad son:

Coca-Cola Bottling Co.

Bob Austin

Torre Eiffel

IBM

Los Yankees

Uganda

Bowdon, Georgia

Monte Fuji

El Kremlin

Los Kennedy

Cuando se anotan entidades nombradas, hay una serie de cosas que es necesario recordar siempre:

1. Siempre se anota una expresión de acuerdo a su significado en el contexto que se evalúa. Dicho de otra manera, la anotación de la expresión depende de cómo se usa. A esta regla se la llama *Anotación por Significado*.
2. Solo anotamos nombres<sup>1</sup>. Se definen los nombres como expresiones que se refieren directamente de manera única a objetos. El significado de las partes del nombre no es típicamente parte del significado del nombre (o sea, los nombres no son *composicionales*) y, por lo tanto, los nombres no se pueden dividir en partes más pequeñas para su anotación.
3. Solo anotamos sustantivos que se refieren a objetos de los tipos relevantes: PER, ORG, LUG, TR. Estos tipos se definen más detalladamente en la Sección 2 de más abajo.
4. Al anotar expresiones de entidades nombradas, no se permiten anotaciones que coincidan o se incrusten. Es decir, cada anotación debe terminar antes de que otra comience.

---

<sup>1</sup> Tal y como se describe en la Sección 2.2, los títulos son un caso especial ya que en realidad no son nombres.

## 2. Tipos de entidades

Identificamos cuatro tipos de entidades nombradas:

**PERSONA (PER):** Las entidades de persona están limitadas a humanos identificados por un nombre, apodo o alias.

**TITULO/ROL (TR):** Títulos personales o rol. Están limitados a títulos que se encuentran cerca del nombre de la persona a la que describen.

**ORGANIZACIÓN (ORG):** Las entidades de organización están limitadas a corporaciones, instituciones, agencias de gobierno y otros grupos de gente definidos por una estructura organizacional establecida.

**LUGAR (LUG):** Las entidades de lugar incluyen nombres de lugares definidos política o geográficamente (ciudades, provincias, países, regiones internacionales, conjuntos de agua, montañas, etc.). Los lugares incluyen también estructuras hechas por el hombre como aeropuertos, autopistas, calles, fábricas y monumentos.

Otros tipos de entidades nombradas como animales, objetos inanimados y unidades monetarias no serán anotadas.

En este documento, las entidades nombradas se indican con [corchetes rectos].

Nótese que las anotaciones no pueden coincidir o estar incrustadas. Es decir, cada anotación debe estar terminada antes de que se pueda empezar otra.

### 2.1. Nombres de personas

Se puede especificar a las personas por nombre, apodo o alias. Los nombres de personas fallecidas, así como de personajes humanos de ficción que aparecen en películas, en la televisión, en libros, etc., deberían anotarse como entidades de PERSONA. Las deidades religiosas también deberían anotarse como PERSONA.

#### 2.1.1. Nombres de familia/Apellidos

Los nombres de familia o apellidos también deberían anotarse como PERSONA. En tales casos, el artículo “los” precede normalmente al apellido:

Los Smith.

Estos artículos deberían ser incluidos en la extensión de los nombres de familia.



Si un título contiene otra entidad para anotar, se anotan por separado <sup>2</sup>.  
Por ejemplo:

El [presidente] de [Microsoft] [Bill Gates] afirmó que...  
TR                      ORG              PER

### 2.2.2. Nombres separados por títulos

Por último, un título parte a veces en dos el nombre de una persona. En estos casos, se anotan las dos partes del nombre de la PERSONA como dos entidades separadas de PERSONA:

[Alfred] [Lord] [Tennyson]  
PER      TR      PER

### 2.2.3 Secuencias de títulos

A veces, se ofrece más de un título para una sola persona. En caso de que todos esos títulos estén contiguos al nombre de la PERSONA, anotaremos cada uno de los títulos por separado como TÍTULO/ROL.

[El alcalde] de [Karachi] y [presidente] de [la XYZ], el [Sr.] [Anwar Ayub].  
TIT                      ORG              TR                      ORG              TIT      PER

Nótese que la conjunción coordinante y no se cuenta al tratar de intentar determinar si el elemento propuesto TR es adyacente al relevante nombre Persona.

### 2.2.4 Sufijos ordinales

Ocasionalmente puede haber un sufijo ordinal como “Jr.”, “Sr.” y “III”. Estos se consideran parte del nombre de la persona y deberían ser anotados como parte del nombre, como por ejemplo:

El [Sr.] [Albert Franklin, Jr.] era parte del equipo de investigación.  
TIT      PER

---

<sup>2</sup> Nótese que la prohibición de entrelazar o anidar anotaciones (no debemos insertar jamás una entidad nombrada *dentro de* otra) nos obliga a convertir la serie “El Presidente de Microsoft Bill Gates” en una serie de entidades nombradas. Este es un caso especial para entidades de ORG contenidas dentro de elementos TÍTULO/ROL –que deben estar así mismo cerca del nombre de la PERSONA que se tiene que marcar. En general, no se deben separar las Entidades Nombradas en partes más pequeñas.

### **2.2.5 Títulos que no se marcan**

Para esta tarea se definen los títulos y roles que se encuentran justo delante o detrás de la entidad de una PERSONA nombrada. Por tanto, los títulos y roles solo se anotan cuando son contiguos al nombre de la persona a la que modifican. En el siguiente ejemplo, el grupo nominal “vicepresidente” no se considera un título y no se anota:

El más firme partidario era el vicepresidente.

### **2.3. Nombres de organizaciones**

Anoten todas las menciones de nombres propios de grupos con una estructura marcadamente organizacional. Estas incluyen:

#### **Negocios**

Los beneficios de la [Compañía Bridgestone Sports]

#### **La bolsa de valores**

Las acciones [NASDAQ]

#### **Organizaciones multinacionales**

Representantes de [la Unión Europea]

#### **Partidos políticos**

El optimista [GOP]

#### **Entidades de gobierno no genéricas**

[El Ministerio de Asuntos Interiores]

#### **Equipos deportivos**

[Los Phillies]

#### **Grupos militares**

[Los Tigres Tamil]

### **2.3.1 Organizaciones y “Anotación por significado”**

Muchas otras clases de entidades se refieren a instalaciones o edificios principalmente definidos por su estructura organizacional establecida, y que pueden hacer cosas como publicar declaraciones, tomar decisiones, contratar a gente, recaudar dinero, etc. Una mención de una entidad tal, debería anotarse como ORGANIZACIÓN cuando funciona como ORG en el documento. Estas incluyen cosas como:

#### **Iglesias y otras instituciones religiosas**

[La Iglesia Luterana de la Trinidad]

#### **Hospitales**

[La corporación hospitalaria en el área de Finger Lakes]

#### **Hoteles**

[El grupo hotelero Four Seasons]

#### **Museos**

[El museo Guggenheim]

#### **Universidades**

[La Universidad de Chicago]

#### **Oficinas de gobierno**

[La Casa Blanca]

Ya que anotamos por significado, estas entidades solo se anotan como ORG cuando se usan como tales. A menudo, los ejemplos de más arriba serán anotados como entidades LUG. Para las reglas de distinción entre estos tipos, por favor consulten las Secciones 2.4 y 2.5 de más abajo.

### **2.3.2 Artículos en entidades organizacionales**

Los determinantes definidos e indefinidos *el/la/los/las* y *un/una/unos/unas* deberían estar usualmente incluidos en la anotación de los nombres de organización:

El Ministerio de Asuntos Interiores

Los Philleis

Los Tigres Tamil

El Museo Guggenheim

La Universidad de Chicago

La Casa Blanca

Esta regla no es pertinente en casos en que cuantifican a otro elemento que no sea la entidad anotada, como en los ejemplos siguientes:

Un portavoz del [Hotel Gulshan]

El vicepresidente [estadounidense]

La excepción es especialmente común cuando la entidad anotada se usa en la posición de adjetivo premodificador.

### **2.3.3 Organizaciones que no se anotan –Referencias genéricas**

Menciones de entidad genérica como “la policía” y “el gobierno” no deberían anotarse ya que estos no son nombre propios únicos referentes a entidades específicas.

## **2.4 Nombres de lugar**

Ejemplos de series relacionadas a un lugar que se anotan como LUGAR incluyen nombres de cuerpos celestiales, continentes, países, provincias, condados, ciudades, regiones, distritos, ciudades, poblaciones, vecindarios, aeropuertos, carreteras, nombres de calles, fábricas, plantas de manufactura, direcciones de calles, océanos, mares, estrechos, bahías, canales, estrechos, ríos, islas, parques nacionales, montañas, lugares ficticios o míticos, y estructuras de monumentos, como la Torre Eiffel y el Monumento de Washington. Por ejemplo:

el colapso de la recién construida [presa Teton]

la disputa por los votos en [el condado de Dade]

[El puente Walt Whitman] permaneció cerrado

empezaron las reparaciones en un tramo de 10 millas en [el oleoducto de Alaska]

[El estado jardín] es conocido por sus tomates.

### **2.4.1 Lugares y “Anotación por significado”**

Acuérdese que siempre anotamos una expresión en base a como se usa. Como tal, habrá un determinado número de casos donde una expresión que usualmente se usa para referirse a un lugar se use para referirse a una organización. Ver las Secciones 2.3 y 2.5 para más detalles.

### **2.4.2 Extensión de los nombres de lugar**

Hay varios temas en cuanto a la expresión de nombres de lugar y qué partes de la serie anotar.

#### **2.4.2.1 Artículos en las entidades de lugar**

Los determinantes definidos e indefinidos *el/la/los/las* y *un/una/unos/unas* deberían estar usualmente incluidos en la anotación de los nombres de lugar:

El Museo Guggenheim

La Universidad de Chicago

La Casa Blanca

Esta regla no es pertinente en casos en los que cuantifican a algún otro elemento que no sea la entidad anotada, como en los siguientes ejemplos:

Un camino de entrada del [Hotel Gulshan]

La tierra continental de [EE.UU.]

Esta excepción es especialmente común cuando el nombre anotado se usa en la posición de adjetivo premodificador.

#### **2.4.2.2 Expresiones compuestas**

Las expresiones compuestas en que los nombres de lugares, **separados o no por una coma**, deberían anotarse como instancias separadas de LUGAR.

[Kaohsiung], [Taiwán ]  
[Washington][DC]

### **2.4.2.3 Indicadores**

Cuando un “indicador” se usa comúnmente como parte normal de un nombre de lugar, esa palabra debería incluirse en la extensión de la entidad de LUGAR. Por ejemplo, debe incluirse en la serie anotada la palabra “río” en el nombre de un río, “montaña” en el nombre de una montaña, “ciudad” en el nombre de una ciudad, etc. si tales palabras están contenidas en la serie.

[El río Misisipi]

[Las montañas del Himalaya]

### **2.4.2.4 Modificadores de lugar y nombres de lugar “semioficiales”**

A menudo los nombres de lugar están modificados por palabras como “del sur”, “más abajo”, “al este”, “el anterior”, etc.

Cuando estos modificadores son parte del nombre oficial de un lugar deberían ser anotadas como parte del nombre de LUGAR. Por ejemplo:

[El Alto Volta]

[Dakota del Norte]

Aún si el nombre del lugar no tiene estatus “oficial”, pero tiene una definición acordada y se usa comúnmente, la serie debería ser anotada como LUGAR, como en:

[El Medio Oriente]

[El Banco Oeste]

[Europa del Este]

Cuando estos modificadores no son el nombre oficial de un lugar, o cuando la definición del lugar pudiese variar dependiendo de la persona, no anoten el modificador como parte del nombre de entidad LUGAR.

El banco oeste del [río Misisipi]

La antigua [Unión Soviética]

[Gaul], en la [Francia] de hoy día

El Bajo [Manhattan]

Estos nombres de lugar pueden a veces ser difíciles. Si no está seguro sobre si el modificador es parte del nombre oficial, debería incluir dicho modificador como parte del nombre de lugar.

## **2.5. Decidiendo entre tipos de entidad**

Hay algunas situaciones en las que decidir el tipo de entidad que asignar puede ser algo difícil.

### **2.5.1 ORG que se refiere a LUG, LUG que se refiere a ORG**

Muchas organizaciones tienen no solo una estructura organizacional, sino también una ubicación física. Por ejemplo, los museos son ante todo organizaciones pero también están alojados en un edificio o instalaciones específicos. Así que mientras que normalmente anotamos los museos como entidades ORG, hay casos en que un ejemplo en particular podría funcionar mejor como LUGAR. En dichos casos, los/as anotadores deben anotar la entidad nombrada basándose en cómo funciona en la oración. Por ejemplo:

[El Museo Guggenheim] anunció una nueva adquisición.  
ORG

[El Museo Guggenheim] fue construido por [Wright].  
LUG PER

Treinta personas fueron heridas en la explosión bomba frente al [Hotel Gulshan] en el centro.  
LUG

Un portavoz del [Hotel Gulshan] tildó el accidente de tragedia.  
ORG

De igual manera, nombres de ciudades, países, y otros lugares se usan frecuentemente para referirse a organizaciones ubicadas en aquellos lugares y no a lugares geográficos. Por ejemplo:

[Washington] anunció hoy una nueva política de impuestos.  
ORG

En este caso, se usa el nombre Washington para referirse al gobierno estadounidense ubicado en Washington. Ya que Washington se refiere a una entidad organizacional en este ejemplo, se debería anotar como ORG.

Así mismo, cuando el nombre de una única estructura o edificio (normalmente una ubicación) se usa para referirse al gobierno u otra organización alojados en esas instalaciones, el nombre debería anotarse como ORG:

[El Pentágono] publicó una declaración sobre el incidente.  
ORG

La lógica simple se aplica a nombres de lugar que se refieren a equipos de deporte:

[Boston] venció a [Nueva York] anoche en la manga extra.  
ORG                      ORG

En resumidas cuentas, para cualquier caso en que el nombre del lugar se use para referirse a una organización, se debería anotar el nombre basándose en la función dentro de la oración:

ORG: se usa cuando el ejemplo se refiere principalmente a la estructura organizacional y se comporta como agente (como publicar declaraciones, tomar una decisión, contratar a gente, recaudar dinero, etc.)

LOC: se usa cuando los ejemplos se refieren principalmente a la estructura física, en vez de a la gente o grupos que los dirigen.

### **3. Casos difíciles**

#### ***3.1. Expresiones que se refieren a entidades múltiples***

Cuando una frase se refiere a múltiples entidades nombradas, marquen cada entidad por separado.

Por ejemplo, esta oración contiene dos entidades:

[China] y [Corea del Sur] firmaron el acuerdo.

De igual manera:

[Jimmy] y [Rosalyn Carter]

[Norteamérica] y [Sudamérica]

Pero tengan cuidado con separar nombre propios que contienen una conjunción. Por ejemplo:

[Los servicios de Pescadería y Fauna y Flora]

es el nombre de la organización y debería ser anotado como una entidad de nombre sencillo (no es “El servicio de Pescadería” y “El servicio de Fauna y Flora” como nombres separados).

### **3.2. Expresiones anidadas**

Recuerde que no se anotará ninguna expresión anidada. Cuando el nombre de una entidad contiene otro nombre de entidad, no saque el nombre de la otra entidad ni lo marque por separado<sup>3</sup>. Solo anoten la entidad mayor. Por ejemplo

[Consultoría Arthur Anderson] no hay marcación para Arthur Anderson solo

[La corporación Boston Chicken] no hay marcación para Boston solo

### **3.3 Entidades en posición de premodificador (Nombres adjetivo)**

Si un nombre de entidad modifica otra palabra (incluso si esa palabra no es un tipo de entidad que se pueda etiquetar), debería marcar de cualquier modo el nombre de entidad.

Los beneficios de la [Bridgestone]

El gobierno [Clinton]

Bonos y valores del [Ministerio de Hacienda]

Exportadores de [EE.UU.]

Ordenadores [Apple]

Petróleo crudo intermedio de [Texas]

---

<sup>3</sup> Las entidades de PERSONA que están cerca de TR y los pre-modificadores de ORG constituyen una excepción importante. Por favor, diríjase a la sección 2.2.

El festival de cine de [China]

Asimismo, si el nombre de entidad ocurre en forma de adjetivo, debería anotarla:

Las compañías [estadounidenses]

Comida [china]

Los documentos de [Enron]

Un banco del [medio oeste]

### **3.4. Posesivos**

Cuando encuentre una construcción posesiva, anote las dos partes individualmente como dos nombres por separado. Por ejemplo:

[La Escuela de estudios de postgrado de Empresariales] [de la Universidad de Temple]

[El Parlamento] [de Canadá]

Recuerde que la anotación requiere que seleccione todas palabras completas para anotarlas como nombres, así que debe incluir la preposición “de” aunque no sea parte del nombre.

### **3.5. Otros tipos de nombres**

Alias, acrónimos, apodos y abreviaciones de nombres propios deberían ser anotados como un nombre:

IBM	(abreviación de Empresa Internacional de Máquinas)
Big Blue (Gran Azul)	(alias para Empresa Internacional de Máquinas)
Big Board (El Gran Comité)	(alias para la Bolsa de valores de Nueva York)
Mr. Fix-It (Sr. Arréglalo)	(apodo para el candidato a jefe de la CIA)
The Big Apple (La Gran Manzana)	(apodo para la ciudad de Nueva York)
Red Sox (Medias Rojas)	(alias para los Medias Rojas de Boston)

Sears

(alias para Sears, Roebuck y Compañía)

## 4. Qué NO anotar

### 4.1 Eventos

No anote nombres de eventos, aún cuando se refieran a eventos que ocurren con frecuencia y que estén asociados a estructuras institucionales. Sin embargo, las estructuras institucionales –comités de dirección etc. –deberían ser anotadas.

Los Juegos Panamericanos  
versus  
El comité olímpico

(no hay marcación)  
  
(ORG)

### 4.2. Artefactos y productos

Varios tipos de nombres propios que no se anotan como entidad nombrada incluyen artefactos, otros productos y nombres plurales que no identifican una entidad sola ni única. Por ejemplo:

El Taurus es el último modelo de coche (no hay marcación)

### 4.3 Genéricos

No se deben marcar nombres genéricos que no se refieren a una entidad específica.

Las sopas Campbell del mundo (no hay marcación)

## 5. Inseguridad en la anotación

En algunos casos, puede encontrar ejemplos que no sabe cómo tratar. En caso de que así sea, debe proceder de la siguiente manera:

Puede que su supervisor/a no esté disponible en el momento en que tiene una pregunta o un problema, así que es importante que anote el problema para que pueda ser resuelto más tarde.

Para que pueda seguir trabajando aún cuando tenga una pregunta sobre un ejemplo en particular, hemos creado una anotación más en la herramienta de anotación: “No\_Annotation” (“Sin\_Anotación”). Cuando encuentre un problema o

tenga una pregunta sobre una palabra o frase en particular y no pueda obtener una respuesta inmediata, nombre la entidad como “Sin\_ anotación”. Eso nos permitirá encontrarlo más tarde cuando intentemos resolver el problema.

Debería usar también la etiqueta de “Sin\_Anotación” en casos donde hay algún tipo de problema con una palabra sola o con un conjunto de palabras en el archivo –por ejemplo, están mal traducidas o la fuente no se muestra adecuadamente.

Si todo el archivo es problemático (por ejemplo, una mala traducción, archivo corrupto, problemas con la fuente) escoja otro archivo.

Si tiene dudas sobre la anotación de alguna entidad, especifique la razón:

- Si es un ejemplo no cubierto en las directrices, anótelos en su copia de las directrices y avise a su supervisor/a de ello.
- Si es un ejemplo que en su lenguaje es diferente a las reglas escritas para el inglés, anótelos en su copia de las directrices y avise a su supervisor/a de ello.
- Si es un caso en el que algo está mal en el archivo en el que está trabajando, deje de trabajar en dicho archivo y avise a su supervisor/a.