

1. Introducción

Este trabajo presenta la anotación semántica manual de un breve texto biográfico sobre el científico español Severo Ochoa mediante la herramienta de etiquetado de datos Label Studio.

En su desarrollo se han realizado las siguientes tareas de Procesamiento del Lenguaje Natural (PLN):

- a) Se han identificado y clasificado las entidades nombradas, *Named Entity Recognition* (NER).
- b) Se han establecido relaciones entre esas entidades.
- c) Las entidades reconocidas se han vinculado (*Named Entity Linking*, NEL) manualmente con sus correspondientes entradas en DBpedia¹.
- d) Se ha contrastado la anotación NER manual con la automática, ejecutada por la herramienta DBpedia Spotlight.
- e) Se ha realizado un análisis de correferencias, con el objetivo de reflejar redes semánticas intratextuales.
- f) Finalmente, se han diseñado tripletas estructuradas de la forma sujeto-predicado-objeto con el fin de codificarlas en lenguaje *Resource Description Framework* (RDF). De ese modo, a través de un grafo, se han reflejado visualmente las conexiones entre las entidades identificadas, al tiempo que se ha representado sintéticamente la biografía del científico español.

¹ DBpedia es un proyecto comunitario que extrae contenido estructurado de Wikipedia y otros proyectos de Wikimedia para crear un grafo de conocimiento abierto.

2. Identificación de entidades nombradas

2.a) Configuración del etiquetado NER

El NER es una tarea del PLN utilizada para identificar y clasificar las entidades presentes en un texto, con el objetivo de extraer información significativa y crear bancos de datos procesables automáticamente.

Para la identificación de entidades en el texto propuesto se ha diseñado un conjunto de etiquetas (*tagset*) que incluye: menciones nombradas (NAM), es decir, antropónimos como *Severo Ochoa* o topónimos como *Luarca*; menciones nominales (NOM), o sea, sintagmas nominales como *beca honorífica*; y expresiones temporales como *24 de noviembre de 1905*, relevantes todas ellas en la biografía del investigador español. Estas categorías se refieren a una única entidad en el contexto textual.

Como modelo para el etiquetado se ha tomado la guía de anotación de entidades *ACE (Automatic Content Extraction)*. *English Annotation Guidelines for Entities Version 6.6 2008.06.13* del *Linguistic Data Consortium*² por los siguientes motivos:

- a) El marco ACE proporciona no solo tipos, sino también subtipos de entidades, necesarios para capturar con precisión ciertas relaciones semánticas en el texto.
- b) La guía incluye análisis de correferencias entre entidades.

² El Linguistic Data Consortium (LDC) es una organización internacional, afiliada a la Universidad de Pensilvania, que se dedica a la recopilación, desarrollo y distribución de recursos lingüísticos para la investigación y el desarrollo en el procesamiento del lenguaje natural (PLN), la lingüística computacional y la tecnología del habla.

c) Este estándar es adecuado para este estudio, por no estar orientado al aprendizaje profundo (los modelos de *Deep Learning* requieren un formato estructurado, como *Beginning-Inside-Outside*, BIO).

Sin embargo, debido a las particularidades textuales, la guía ACE no se ha seguido estrictamente y se han adoptado categorías de entidades no contempladas en ella como *AWARD* o *EVENT*.

El diseño del etiquetado es descrito en la siguiente tabla, coloreada en relación con los correspondientes tonos empleados en el *tagset* de Label Studio:

TIPO	DESCRIPCIÓN	ETIQUETA	SUBTIPO	DESCRIPCIÓN	ETIQUETA
PERSON	Entidad que representa a un individuo humano, vivo o fallecido.	PER	No aplica (N/A)	N/A	N/A

ORGANIZATION	Institución o grupo con una estructura organizativa definida, como instituciones educativas, científicas o culturales.	ORG	EDUCATIONAL	Organización cuya principal función es la promoción del aprendizaje y la educación, como universidades o centros académicos.	ORG_EDU
			MEDICAL_SCIENCE	Organización dedicada a la investigación científica o a la prestación de servicios médicos, como hospitales, laboratorios o centros de investigación.	ORG_SCI

			CULTURAL	Institución cuyo fin es la difusión o a la representación cultural, como centros culturales o espacios de representación nacional	ORG_CULT
GEOGRAPHICAL/SOCIAL/ POLITICAL ENTITY	Entidad compuesta de población, gobierno y ubicación geográfica, como, países, comarcas o ciudades.	GPE	NATION PROVINCE	Naciones o Estados Unidades administrativas intermedias, como comunidades autónomas o provincias	GPE_NAT GPE_PROV

			POPULATION CENTER	Núcleo urbano menor que la provincia, como ciudades, villas o pueblos.	GPE_PC
AWARD	Reconocimiento, premio o distinción otorgado a personas por una organización.	AWARD	GENERAL	Premio o distinción por el reconocimiento a logros sobresalientes en diversos campos, como ciencias, artes, literatura.	AWARD_G
			GRANT	Reconocimiento en forma de financiación para realizar investigaciones o proyectos.	AWARD_GT

EVENT	Acontecimiento significativo, como una guerra, catástrofe o suceso histórico relevante.	EVENT	N/A	N/A	N/A
DATE	Expresiones temporales como fechas específicas o periodos.	DATE	N/A	N/A	N/A
FACILITY	Infraestructura diseñada para actividades humanas con propósitos específicos, como laboratorios o edificios.	FAC	N/A	N/A	N/A

TITLE	Título honorífico, académico, profesional o social que precede al nombre propio de una persona.	TIT	N/A	N/A	N/A
--------------	---	------------	-----	-----	-----

2.b) Principales dificultades encontradas

El diseño del conjunto de etiquetas no ha estado exento de desafíos, especialmente en relación con el etiquetado de las menciones nominales *el laboratorio, situado en Glasgow, del profesor Paton y el laboratorio del profesor Meyerhof*.

La dificultad ha consistido en determinar si estas expresiones debían anotarse como una única mención nominal, dado que sus referentes son identificables en el contexto textual (el Laboratorio del Departamento de Fisiología de la Universidad de Glasgow y el Instituto Kaiser Wilhelm de Berlín, respectivamente), o si, por el contrario, era necesario segmentarlas en entidades independientes: *laboratorio, profesor, Glasgow, Berlín, Paton y Meyerhof*.

La solución adoptada ha sido etiquetar cada elemento como una entidad independiente para favorecer la construcción de redes semánticas más complejas y facilitar tareas propias del PLN como la vinculación de datos y el análisis de correferencias.

2.c) Etiquetado NER con Label Studio

Una vez diseñado el *tagset*, se ha configurado en la herramienta de etiquetado de datos de código abierto Label Studio.

A cada etiqueta se le ha asignado un color específico para facilitar su diferenciación visual. Asimismo, las etiquetas que representan subtipos de entidades se han asociado al mismo color que las del tipo al que pertenecen, pero en otros tonos para, por un lado, representar la pertenencia a la misma tipología, a través de la igualdad cromática, y, por otro, distinguir entre las distintas subcategorías visualmente. Por ejemplo, las etiquetas que designan entidades geopolítico-sociales (GPE) comparten un color base rojo, pero los tonos específicos permiten distinguir entre naciones, comunidades autónomas y centros urbanos, según las subcategorías correspondientes.

El conjunto de etiquetas, así como la anotación NER se muestran a continuación:

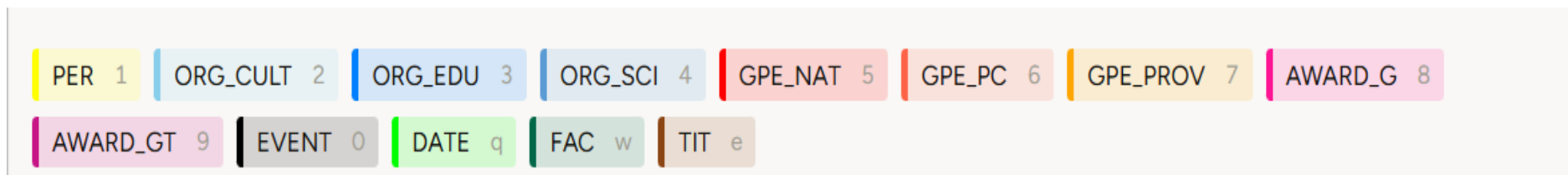


Imagen del *tagset* NER diseñado con Label Studio.

PER 1

ORG_CULT 2

ORG_EDU 3

ORG_SCI 4

GPE_NAT 5

GPE_PC 6

GPE_PROV 7

AWARD_G 8

AWARD_GT 9

EVENT 0

DATE q

FAC w

TIT e

REF t

PRO a

NOM s

ATR d

APP f

ARC g

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

Imagen del etiquetado manual NER realizado con Label Studio.

3. Establecimiento de relaciones entre entidades

3. a) Configuración del etiquetado

Con el fin de reflejar los vínculos semánticos existentes entre las identidades clasificadas se ha utilizado el esquema sujeto-predicado-objeto.

Según este esquema la primera entidad actúa como sujeto, la segunda, como objeto y el vínculo entre ambas, que define la relación mediante, representa el predicado.

Estas relaciones se ilustran mediante flechas orientadas cuya dirección indica el objeto de la conexión, tal y como se muestra a continuación:

1. Person → Date (PER → DATE)

Severo Ochoa → Nació → el 24 de noviembre de 1905

Severo Ochoa → Estudiante de Medicina desde → 1923

Severo Ochoa → Inició investigaciones en → 1927

Severo Ochoa → Trabajó en NYU desde → 1942

Severo Ochoa → Trabajó en NYU hasta → 1974

Severo Ochoa → Recibió el Nobel en → 1959

Severo Ochoa → Regresó a España en → 1985

Severo Ochoa → Murió el → 1 de noviembre de 1993

2. Person → Person (PER → PER)

Severo Ochoa → Esposo de → Carmen

Severo Ochoa → Huyó con → Carmen

Severo Ochoa → Sepultado junto a → Carmen

Severo Ochoa → Colega de → Kornberg

Severo Ochoa → Compatriota de → Pío Baroja

Severo Ochoa → Compatriota de → Blas Cabrera

Severo Ochoa → Compatriota de → Xavier Zubiri

3. Person → Geographical/Social/Political Entity (PER → GPE)

Severo Ochoa → Nació en → Luarca.

Severo Ochoa → Se trasladó a → Berlín.

Severo Ochoa → Regresó a → España.

Severo Ochoa → Regresó definitivamente a → España.

Severo Ochoa → Enterrado en → Luarca.

4. Person → Organization (PER → ORG)

Severo Ochoa → Estudiante de → Facultad de Medicina de Madrid

Severo Ochoa → Se refugió en → Casa de España en París

Severo Ochoa → Trabajó en → Departamento de Medicina de Nueva York

Severo Ochoa → Trabajó finalmente en → Centro de Biología Molecular Severo Ochoa

5. Person → Event (PER → EVENT)

Severo Ochoa → Huyó de → Guerra Civil Española

6. Person → Facility (PER → FAC)

Severo Ochoa → inició investigaciones en → laboratorio

7. Person → Award (PER → AWARD)

Severo Ochoa → recibió → beca honorífica

8. Person → Title (PER → TIT)

Profesor Paton → con título de → Profesor.

Profesor Meyerhof → con título de → Profesor

9. Organization → Geographical/Social/Political Entity (ORG → GPE)

Junta para la Ampliación de Estudios → Situada en → España

Casa de España en París → Situada en → París

10. Facility → Geographical/Social/Political Entity (FAC → GPE)

Laboratorio → Situado en → Glasgow

Laboratorio → Situado en → Berlín

11. Award → Persons (AWARD → PER)

Premio Nobel de Fisiología o Medicina → Otorgado a → Severo Ochoa

Premio Nobel de Fisiología o Medicina → Otorgado a → Kornberg

12. Geographical/Social/Political Entity → Geographical/Social/Political Entity (GPE→GPE)

Luarca → Villa de → Asturias

Asturias → Comunidad Autónoma de → España

13. Grant→Organisation

Beca honorífica → Concedida por → Junta para la Ampliación de Estudios

14. Grant→Facility

Beca honorífica → Con destino en → laboratorio.

15. Facility→ Person (FAC →PER)

Laboratorio → Dirigido por → Meyerhof

Laboratorio→ Dirigido por → Paton

16. Evento → (EVENT → GPE)

Guerra Civil → Estalló en → España.

El análisis de las relaciones entre entidades muestra cómo *Severo Ochoa* es el nodo central de la red semántica. Este nodo conecta con personas, fechas, instituciones científico-educativas, entidades geopolítico-sociales, reconocimientos y eventos históricos fundamentales en su vida.

Sin embargo, de la observación se desprende que *Severo Ochoa* no es la única entidad generadora de vínculos en esta red, pues, aunque en menor medida, otras entidades crean conexiones que enriquecen la comprensión del contexto histórico, científico y social que rodeó al investigador español.

3. b) Relaciones establecidas entre entidades con Label Studio

A continuación, se presentan las imágenes que ilustran las relaciones establecidas entre las entidades en cada enunciado, identificadas con el *software* Label Studio.

Aun a riesgo de resultar redundante, para garantizar una representación clara y legible, se han incluido imágenes segmentadas de las oraciones en aquellos casos donde las etiquetas se solapan. Estas imágenes se corresponden a etapas previas al etiquetado global de cada una de ellas. Además, el primer enunciado se ha omitido porque la configuración descrita del etiquetado³ ha implicado la ausencia de relaciones entre sus menciones.

³ *Médico, bioquímico y biólogo molecular* será etiquetado como correferencia atributiva, tal y como se verá después.

Oración 2

AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Nació en | Nació el | Villa de | Comunidad Autónoma de

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas

Oración 3

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8

AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Estudiante de

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora

PER 1

ORG_CULT 2

ORG_EDU 3

ORG_SCI 4

GPE_NAT 5

GPE_PC 6

GPE_PROV 7

AWARD_G 8

AWARD_GT 9

EVENT 0

DATE q

FAC w

TIT e

REF t

PRO a

NOM s

ATR d

APP f

ARC g

Estudiante de Medicina desde

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora

PER 1

ORG_CULT 2

ORG_EDU 3

ORG_SCI 4

GPE_NAT 5

GPE_PC 6

GPE_PROV 7

AWARD_G 8

AWARD_GT 9

EVENT 0

DATE q

FAC w

TIT e

REF t

PRO a

NOM s

ATR d

APP f

ARC g

EstudEstudiante de Medicina desde

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora

Oración 4

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

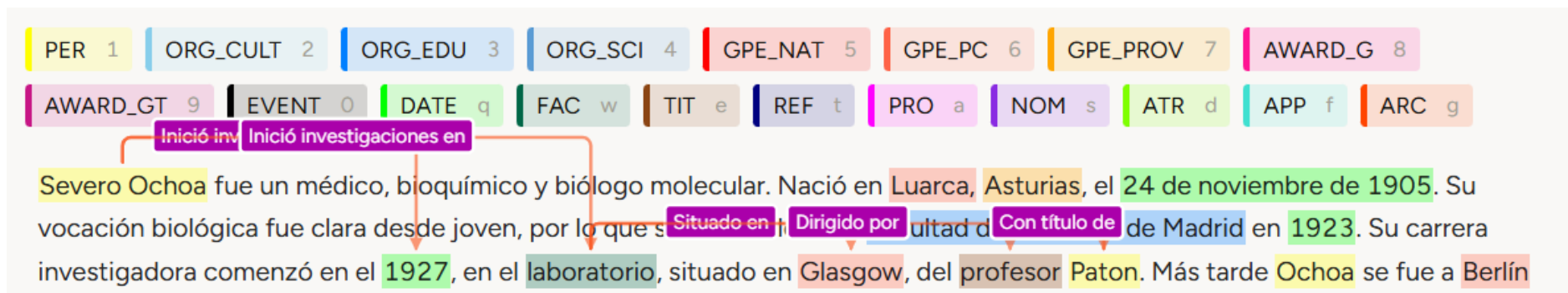
Inició investigaciones en

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín

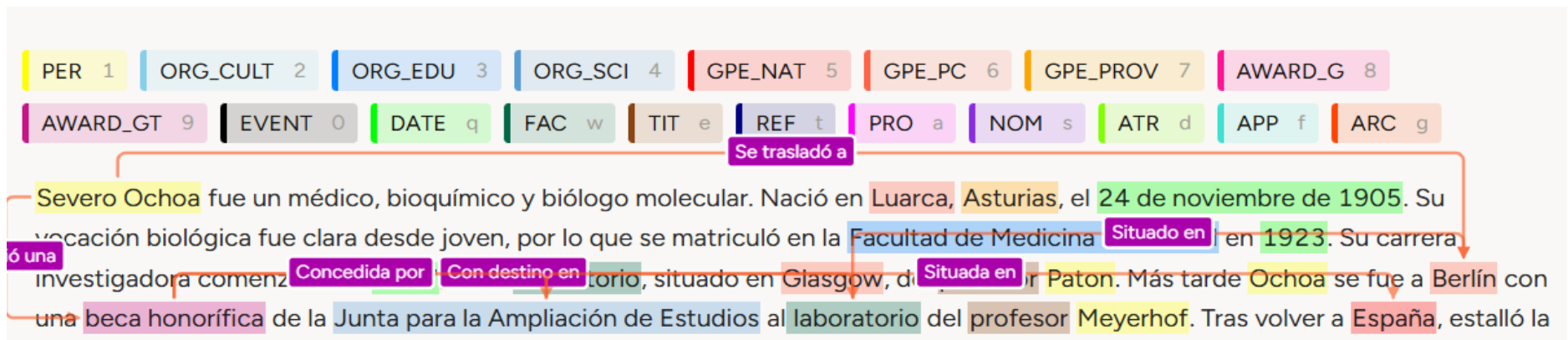
PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Inició investigaciones en

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se **Situado en** **Dirigido por** **Con título de** Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín



Oración 5



PER 1

ORG_CULT 2

ORG_EDU 3

ORG_SCI 4

GPE_NAT 5

GPE_PC 6

GPE_PROV 7

AWARD_G 8

AWARD_GT 9

EVENT 0

DATE q

FAC w

TIT e

REF t

PRO a

NOM s

ATR d

APP f

ARC g

Recibió una

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof.

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof.

AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Lluarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. M Con título de choa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof.

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8 | AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Se trasladó a

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Lluarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina Situado en rid en 1923. Su carrera investigadora comenzó en e Concedida por Con destino en situado en Glasgow, del pr S Dirigido Con título de arde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas

Oración 6

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8

AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | COR t | PRO a | NOM s | ATR d | APP f | ARC g

Esposo de, Huyó con Regresó a Huyó de

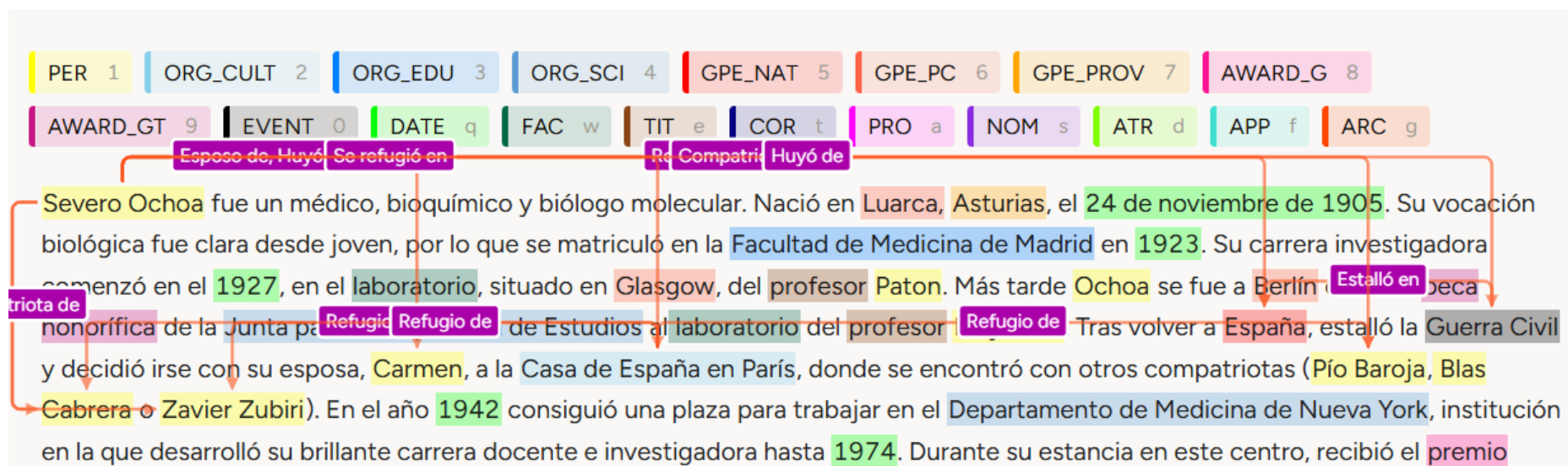
Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luearca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín Estalló en beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8

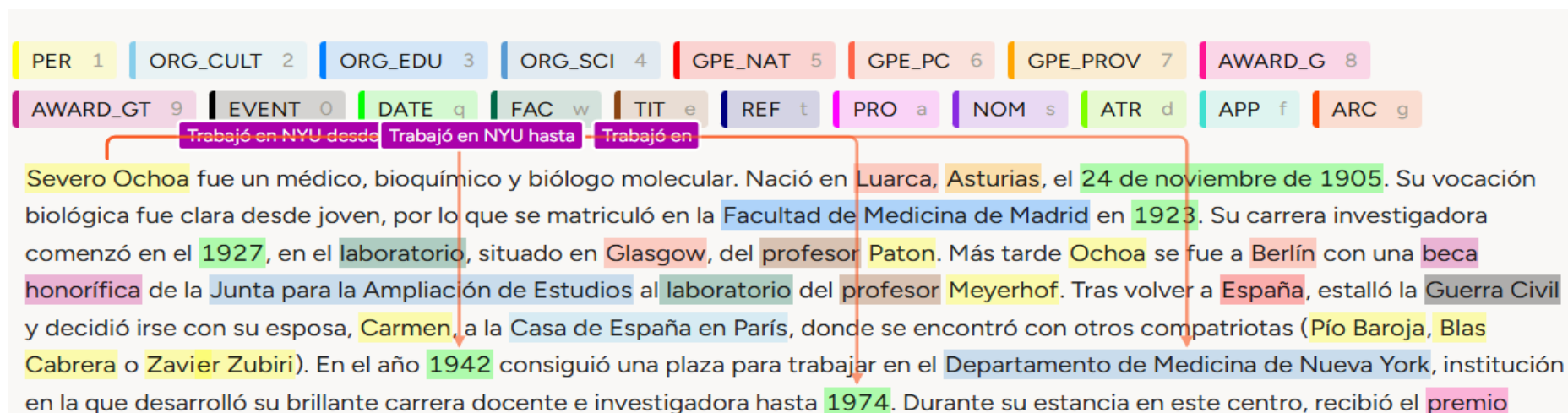
AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | COR t | PRO a | NOM s | ATR d | APP f | ARC g

Se refugió en Compatriota de

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luearca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta pa Refugio Refugio de de Estudios al laboratorio del profesor Refugio de Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.



Oración 7



Oración 8

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8 | AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Otorgado a Recibió el Nobel en Colega de

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luearca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera do Otorgado a restigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

Oración 9

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Regresó definitivamente a

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luear, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
 AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Trabajó finalmente en | Reg. Regresó a España en

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

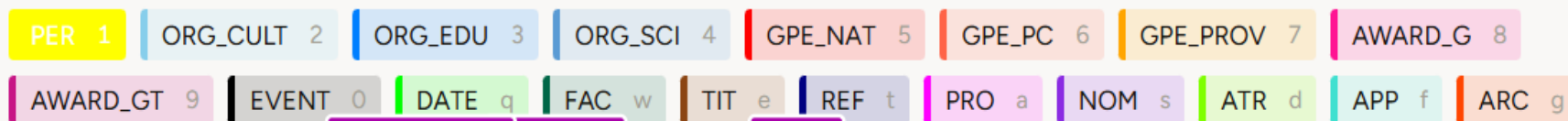
Oración 10

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | REF t | PRO a | NOM s | ATR d | APP f | ARC g

Enterrado en

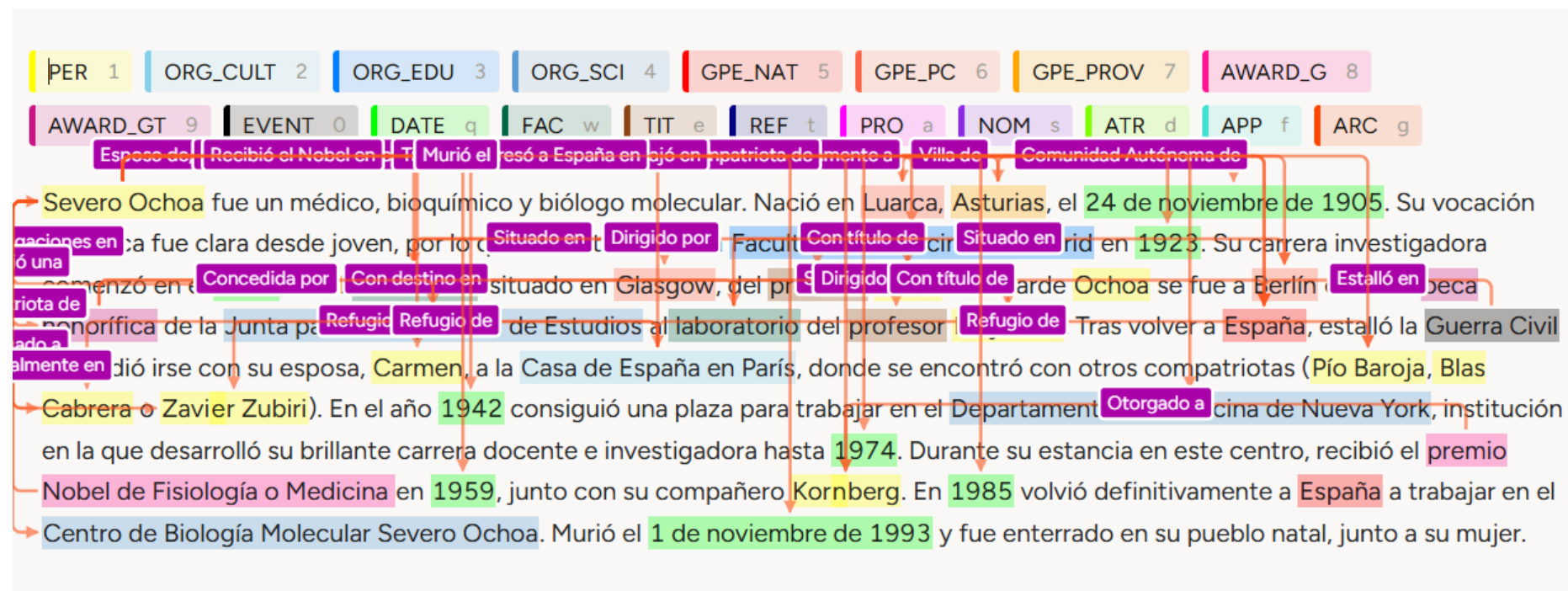
Murió el

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.



Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

Finalmente, se muestra la imagen del etiquetado que expresa las relaciones entre entidades del texto completo:



4. *Named Entity Linking* (NEL): hipervinculación de entidades nombradas a DBpedia

Una vez identificadas las entidades nombradas (NER) y establecidas sus conexiones, se han vinculado con sus correspondientes entradas en DBpedia con dos propósitos fundamentales: por un lado, proporcionar información que contextualice la breve biografía de Severo Ochoa reflejada en el texto; por otro, resolver ambigüedades, ya que una mención textual puede referirse a múltiples entidades en la base de conocimiento. Por ejemplo, *Luarca* puede aludir tanto a la villa asturiana como al *Luarca Club de Fútbol*, mientras que *Ochoa* puede confundirse con Jorge Luis Ochoa Vásquez, exnarcotraficante colombiano y fundador del Cartel de Medellín.

Dado que no se ha podido establecer una conexión directa con el servidor de DBpedia, la vinculación se ha realizado manualmente. Este hecho ha garantizado precisión en la selección de las entradas, aunque ha evidenciado problemas relacionados con la estructura y actualización de la base de datos de DBpedia. Estas circunstancias, así como las soluciones propuestas, se detallan a continuación:

- **Ausencia de entradas**

En algunos casos, las entidades identificadas en el texto no están documentadas en DBpedia, como, por ejemplo, la Casa de España en París o el Centro de Biología Molecular Severo Ochoa. En el primer caso esta ausencia no ha podido resolverse manualmente, en el segundo, debido a que Label Studio, posibilita únicamente la vinculación a una entidad, el Centro de Biología Molecular Severo Ochoa solo ha podido relacionarse con la Universidad Autónoma de Madrid, a la que, junto al CSIC,

está actualmente adscrito.

Por otra parte, en los casos en que no existe una entrada para la entidad original, pero sí para una entidad que, por motivos históricos la absorbió, se ha optado por crear un hipervínculo con esta última. Así, por ejemplo, la Universidad de Madrid, se ha relacionado con la Universidad Complutense de Madrid, su denominación actual; la Junta para la Ampliación de Estudios (JAE), desmantelada durante la Guerra Civil Española y posteriormente absorbida por el CSIC, ha sido vinculada a este organismo y el Instituto Kaiser Wilhelm de Fisiología y Biología se ha conectado con la Sociedad Kaiser Wilhelm.

Finalmente, entidades, como el Laboratorio de Fisiología de la Universidad de Glasgow o la Facultad de Medicina de la Universidad de Nueva York se han vinculado respectivamente, a la Universidad de Glasgow y a la Universidad de Nueva York por carecer de entrada en DBpedia, debido a su pertenencia a una entidad mayor.

- Combinación incorrecta de menciones

Algunas entradas de DBpedia combinan información asociada a menciones referidas a distintas entidades. Un caso llamativo es el de *España*: la base de datos fusiona en la misma entrada información sobre el país y sobre una estación ferroviaria de Guadalajara, México. Este hecho provoca que los sistemas automatizados infieran relaciones incorrectas entre entidades.

- **Dependencia del idioma**

Algunas entradas, como la de *Diarmid Noel Paton*, solo están disponibles en inglés, con el consecuente perjuicio en tareas de PLN en español.

Aunque estos dos últimos problemas, la combinación de menciones en una entrada y la dependencia del inglés, no han podido resolverse por ser inherentes a la base de conocimiento, evidencian la necesidad de la actualización de datos para garantizar la validez de este tipo de herramientas en el análisis lingüístico y semántico.

5. Análisis comparativo de la anotación NER: manual vs. automática con DBpedia Spotlight

DBpedia Spotlight es una herramienta de anotación semántica y desambiguación de entidades nombradas, que combina métodos probabilísticos con otros basados en reglas para identificar menciones en un texto y vincularlas con sus correspondientes entradas en DBpedia.

La comparación entre la anotación manual y la automática, obtenida mediante DBpedia Spotlight, revela diferencias significativas en cuanto a la contextualización y el nivel de detalle de las etiquetas generadas.

A continuación, se exponen los aspectos más relevantes de este análisis comparativo:

- **Falta de contextualización en la anotación automática**

DBpedia Spotlight no contextualiza el texto como una unidad cohesiva, lo que genera errores en la identificación de entidades y relaciones. Por ejemplo, *noviembre* y el verbo *volver* se hipervinculan a películas, incluso cuando aparecen en minúscula, y *Carmen* a la ópera de Bizet. Esta inconexión refleja la prioridad de la herramienta por establecer vínculos con las entidades registradas en la base de datos de DBpedia, en lugar de tratar de reconocer el sentido específico de las entidades en el contexto discursivo.

- **Limitación en la identificación de personas**

Aunque la anotación automática etiqueta correctamente entidades como Severo Ochoa, Pío Baroja, Blas Cabrera o Xavier Zubiri, omite el etiquetado de personas destacadas en el ámbito científico, como Paton, Meyerhof o Kornberg, que sí se identifican en la anotación manual.

- **Etiquetado inconsistente de fechas**

Las fechas etiquetadas por la herramienta automática son seleccionadas de forma aleatoria. Por ejemplo, 1933 es reconocida, mientras que 1927 o 1974 no lo son.

En la anotación manual todas las fechas mencionadas en el texto son identificadas.

- **Omisión de tipos y subtipos**

DBpedia Spotlight no reconoce categorías como *Organization* o *Award*. En consecuencia, entidades como la Junta para la Ampliación de Estudios, el Centro de Biología Molecular Severo Ochoa o el Premio Nobel de Fisiología o Medicina no son etiquetadas.

Por otro lado, en los casos en que las etiquetas son generadas, el nivel de detalle es inferior al manual. Por ejemplo, entre las GPE, no se diferencian ciudades, comunidades o países.

En definitiva, las principales diferencias entre la anotación manual y automática del texto propuesto estriban en la falta de contextualización, la cobertura limitada de categorías y la inconsistencia en la selección de ciertas etiquetas por parte de la herramienta automática.

Se deduce así que, aunque DBpedia Spotlight puede ofrecer un punto de partida para la anotación NER, las anotaciones manuales siguen siendo fundamentales para garantizar un análisis semántico preciso en textos biográficos.

Probablemente, la integración de técnicas híbridas, que combinen el Aprendizaje Profundo con algoritmos estadísticos y regulares, podría mitigar algunos de los problemas detectados, especialmente por lo que respecta a una comprensión más contextual del discurso. Asimismo, la actualización de las bases de datos de DBpedia en que se basa esta herramienta resulta fundamental para evitar errores de etiquetado y mejorar la calidad de la vinculación de entidades.

Confidence: 0.5

Language: Spanish

☐ n-best candidates

[SELECT TYPES...](#) [ANNOTATE](#)

[Severo Ochoa](#) fue un médico, bioquímico y biólogo molecular. Nació en [Luarca](#), Asturias, el 24 de [noviembre](#) de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de [Madrid](#) en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en [Glasgow](#), del profesor Paton. Más tarde Ochoa se fue a [Berlín](#) con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras [volver](#) a España, estalló la [Guerra Civil](#) y decidió irse con su esposa, [Carmen](#), a la Casa de España en [París](#), donde se encontró con otros compatriotas ([Pío Baroja](#), [Blas Cabrera](#) o [Xavier Zubiri](#)). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su [estancia](#) en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En [1985](#) volvió definitivamente a España a trabajar en el Centro de Biología Molecular [Severo Ochoa](#). Murió el 1 de [noviembre](#) de [1993](#) y fue enterrado en su pueblo natal, junto a su mujer.

[BACK TO TEXT](#)

Imagen de la anotación automática del texto propuesto realizada por DBpedia Spotlight

6. Análisis de correferencias

La correferencia es un fenómeno lingüístico que se produce cuando varias menciones en un texto se refieren a una misma entidad. De esta forma se permite la coherencia y cohesión discursivas.

El análisis de las relaciones correferenciales es fundamental en el Procesamiento del Lenguaje Natural (PLN), ya que facilita la identificación de las conexiones entre entidades, incluso cuando estas se presentan de forma indirecta o mediante distintas expresiones.

6.a) Configuración del etiquetado

Para la clasificación y etiquetado de las relaciones de correferencia, de nuevo, se ha tenido por referente la versión del estándar ACE antes señalada, sin embargo, debido a la brevedad y naturaleza textual, su diseño ha sido mucho menos exhaustivo que el propuesto en esa guía.

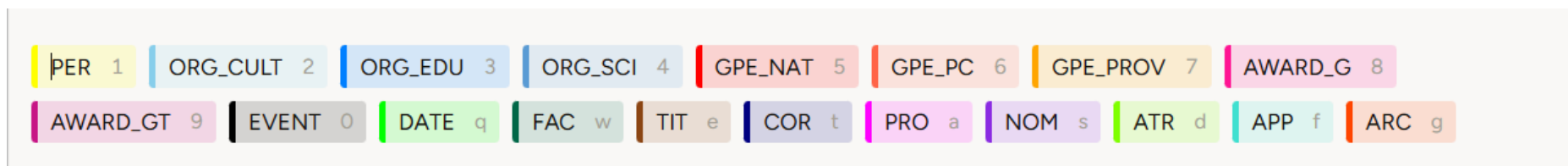
En la siguiente tabla (coloreada en consonancia con el diseño del *tagset* en Label Studio), se presentan las etiquetas utilizadas, el tipo de correferencia establecida, su descripción, así como ejemplos ilustrativos al respecto:

Etiqueta	Tipo de correferencia	Descripción	Ejemplo
COR	Correferencia correferencial	Una mención nombrada remite a otra ya mencionada en el texto con el mismo referente.	<i>Severo Ochoa.....Ochoa....</i>
ATR	Correferencia atributiva	Se describe una propiedad o un estado de una entidad previamente mencionada.	<i>Severo Ochoa...médico, bioquímico y biólogo molecular...</i>
APP	Correferencia apositiva	Una mención aparece junto a otra en la misma frase y ambas se refieren a la misma entidad.	<i>...Su esposa, Carmen.</i>

ARC	Correferencia mediante cláusulas relativas	Una cláusula relativa amplía o especifica información sobre una entidad previamente mencionada.	<i>...Casa de España en París, donde se encontró...</i>
PRO	Correferencia pronominal	Un pronombre hace referencia a una entidad mencionada anterior (anáfora) o posteriormente (catáfora).	<i>Severo Ochoa...su carrera...</i>
NOM	Correferencia nominal	Un sintagma nominal hace referencia a entidades específicas. Puede incluir determinantes, posesivos o modificadores.	<i>Luarca...su pueblo natal...</i>

6.b) Configuración del etiquetado correferencial con Label Studio

A partir del conjunto de etiquetas descrito en la tabla anterior, se ha diseñado el etiquetado para identificar correferencias mediante el *software* Label Studio. Dado que esta herramienta no permite integrar las etiquetas de correferencias de manera aislada, estas han sido añadidas a las de entidades ya existentes. El *tagset* resultante se presenta en la siguiente imagen:



6.c) Problemas detectados en el diseño del *tagset* con Label Studio

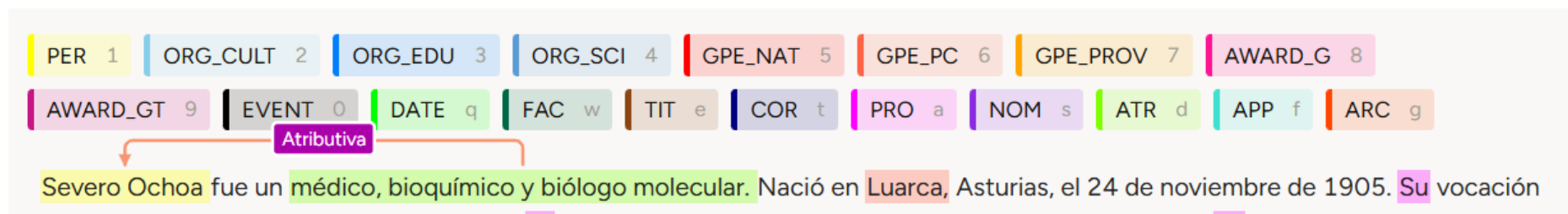
Debido a las características de la herramienta Label Studio, determinadas correferencias no han podido señalarse por concurrir los elementos lingüísticos implicados en un mismo sintagma.

De este modo, aunque entre dos entidades se pueden establecer dos o más relaciones simultáneamente, la herramienta no permite el doble etiquetado de una misma expresión. Por ejemplo, *su esposa* en *su esposa, Carmen* no puede recibir a la vez las etiquetas PRO y APP, en referencia, respectivamente, a Severo Ochoa y Carmen.

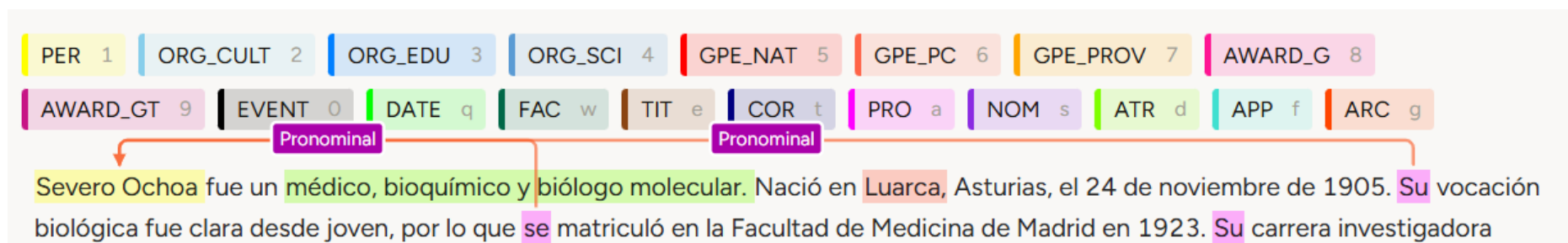
Esta restricción afecta directamente el análisis de cadenas correferenciales, ya que solo se puede establecer una relación directa entre *su* y *Ochoa* o entre *su esposa* y *Carmen*. Como resultado, se omiten conexiones adicionales que reducen la riqueza semántica y limitan la precisión analítica. Por este motivo, se sugiere dotar a la herramienta de la posibilidad de asignar múltiples etiquetas de correferencia a una misma mención.

Finalmente, se muestran las imágenes del análisis de correferencias realizado con Label Studio. Las oraciones en las que no se establecen estas relaciones se han omitido y, como anteriormente, en su representación se ha priorizado la legibilidad:

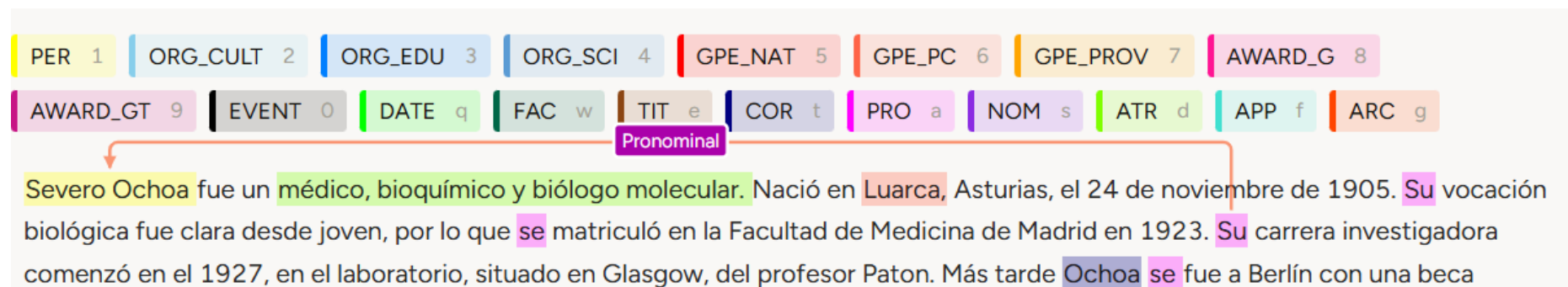
Oración 1



Oración 3



Oración 4



Oración 5

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
 AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | COR t | PRO a | NOM s | ATR d | APP f | ARC g

Correferencial

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarda, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca

PER 1 | ORG_CULT 2 | ORG_EDU 3 | ORG_SCI 4 | GPE_NAT 5 | GPE_PC 6 | GPE_PROV 7 | AWARD_G 8
 AWARD_GT 9 | EVENT 0 | DATE q | FAC w | TIT e | COR t | PRO a | NOM s | ATR d | APP f | ARC g

Cor. Pronominal

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarda, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y

Oración 6

PER 1	ORG_CULT 2	ORG_EDU 3	ORG_SCI 4	GPE_NAT 5	GPE_PC 6	GPE_PROV 7	AWARD_G 8			
AWARD_GT 9	EVENT 0	DATE q	FAC w	TIT e	COR t	PRO a	NOM s	ATR d	APP f	ARC g

Pronominal Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luear, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta, Apositiva ampliación de Estudios al labor Relativa el profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Me Correferencial ueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de

PER 1	ORG_CULT 2	ORG_EDU 3	ORG_SCI 4	GPE_NAT 5	GPE_PC 6	GPE_PROV 7	AWARD_G 8			
AWARD_GT 9	EVENT 0	DATE q	FAC w	TIT e	COR t	PRO a	NOM s	ATR d	APP f	ARC g

Pronominal
 Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de

Oración 7

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de

Oración 8

PER 1	ORG_CULT 2	ORG_EDU 3	ORG_SCI 4	GPE_NAT 5	GPE_PC 6	GPE_PROV 7	AWARD_G 8			
AWARD_GT 9	EVENT 0	DATE q	FAC w	TIT e	COR t	PRO a	NOM s	ATR d	APP f	ARC g

Pronominal

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Lluarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con Nominal compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora Apositiva 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

Oración 9

PER 1	ORG_CULT 2	ORG_EDU 3	ORG_SCI 4	GPE_NAT 5	GPE_PC 6	GPE_PROV 7	AWARD_G 8			
AWARD_GT 9	EVENT 0	DATE q	FAC w	TIT e	COR t	PRO a	NOM s	ATR d	APP f	ARC g

Correferencial

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Lueca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del profesor Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina Correferencial Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

Oración 10

PER 1 ORG_CULT 2 ORG_EDU 3 ORG_SCI 4 GPE_NAT 5 GPE_PC 6 GPE_PROV 7 AWARD_G 8
 AWARD_GT 9 EVENT 0 DATE q FAC w TIT e COR t PRO a NOM s ATR d APP f ARC g

Severo Ochoa fue un médico, bioquímico y biólogo molecular. Nació en Luarca, Asturias, el 24 de noviembre de 1905. Su vocación biológica fue clara desde joven, por lo que se matriculó en la Facultad de Medicina de Madrid en 1923. Su carrera investigadora comenzó en el 1927, en el laboratorio, situado en Glasgow, del profesor Paton. Más tarde Ochoa se fue a Berlín con una beca honorífica de la Junta para la Ampliación de Estudios al laboratorio del ^{Nominal} Dr Meyerhof. Tras volver a España, estalló la Guerra Civil y decidió irse con su esposa, Carmen, a la Casa de España en París, donde se encontró con otros compatriotas (Pío Baroja, Blas Cabrera o Xavier Zubiri). En el año 1942 consiguió una plaza para trabajar en el Departamento de Medicina de Nueva York, institución en la que desarrolló su brillante carrera docente e investigadora hasta 1974. Durante su estancia en este centro, recibió el premio Nobel de Fisiología o Medicina en 1959, junto con su compañero Kornberg. En 1985 volvió definitivamente a España a trabajar en el Centro de Biología Molecular Severo Ochoa. Murió el 1 de noviembre de 1993 y fue enterrado en su pueblo natal, junto a su mujer.

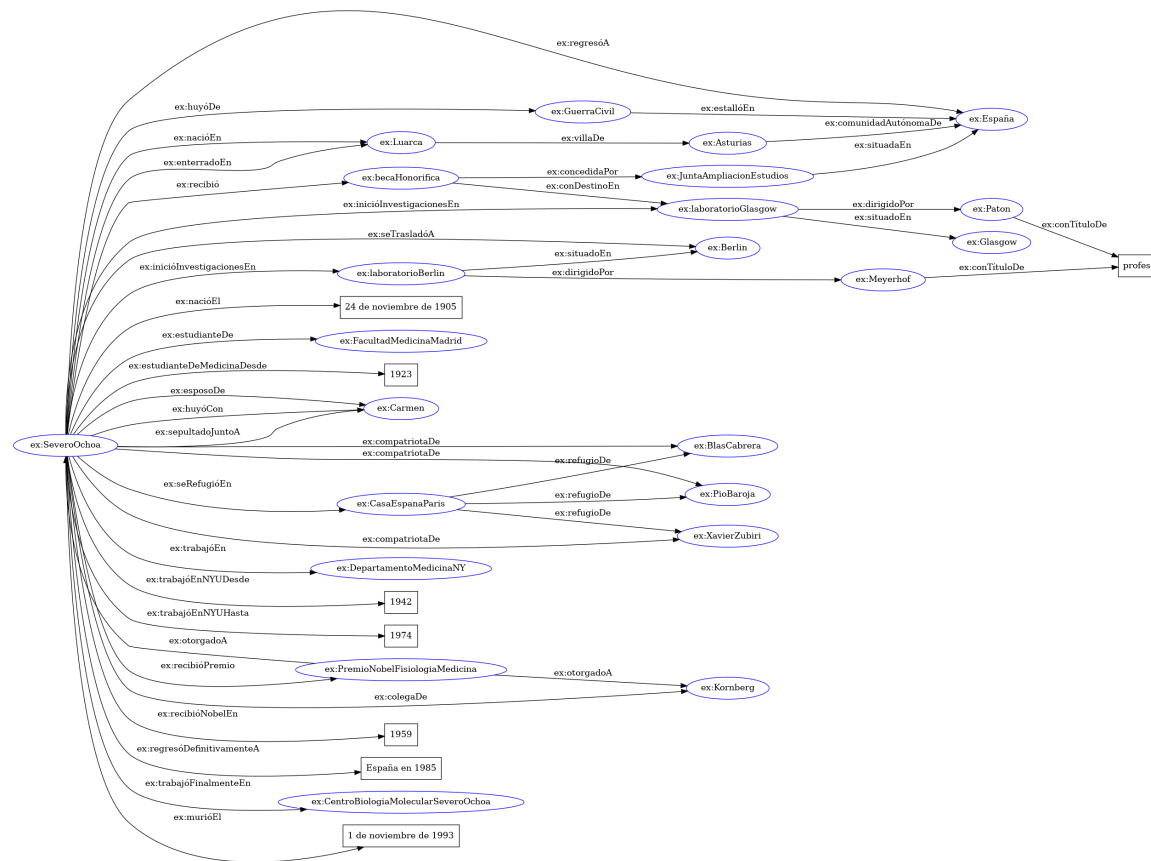
7. Grafo de relaciones semánticas

Como complemento al análisis de las entidades nombradas y sus relaciones se ha generado un grafo, a partir de las tripletas estructuradas en formato sujeto-predicado-objeto, adjuntadas en el anexo.

Este grafo ofrece una representación visual de las redes semánticas extraídas del texto biográfico y refleja cómo el científico español constituye el nodo central: su figura se vincula con numerosas entidades a través de relaciones como su lugar de nacimiento (Luarca), su afiliación académica (Facultad de Medicina de Madrid), su contemporaneidad a eventos históricos (Guerra Civil Española) o el reconocimiento a su labor (Premio Nobel de Fisiología o Medicina).

Además, el grafo permite identificar patrones semánticos que pasan más desapercibidos en el análisis textual lineal, como el carácter internacional del científico o las numerosas relaciones con intelectuales que estableció en su vida.

Se presenta a continuación la imagen del grafo, generado a partir de la herramienta RDF Grapher, servicio web que analiza datos RDF y los visualiza:



Namespaces:
 rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 ex: <http://ejemplo.com/>

8. Conclusión

Este trabajo ha consistido en la anotación semántica manual de un texto biográfico sobre Severo Ochoa. Para ello se han aplicado varias técnicas del Procesamiento del Lenguaje Natural (PLN) mediante el *software* Label Studio: la identificación manual de entidades nombradas (NER) y su comparación con la automática de DBpedia Spotlight; el establecimiento de relaciones entre esas entidades, su vinculación con la base de conocimiento de DBpedia y el análisis de correferencias. Finalmente, las relaciones analizadas se ha representado visualmente en un grafo.

Por un lado, la identificación de entidades NER ha visibilizado la posibilidad de distintos etiquetados válidos, dependientes de las prioridades establecidas y el análisis de correferencias ha destacado la necesidad de aplicaciones de etiquetado más flexibles, dado que Label Studio no permite el doble etiquetado de menciones simultáneas, con la consiguiente pérdida de conexiones en las cadenas correferenciales.

Por otro, la vinculación de entidades con las correspondientes entradas de DBpedia ha enriquecido la anotación, aunque ha evidenciado limitaciones, como la ausencia de entradas y errores en la desambiguación. Asimismo, la comparación entre la anotación manual y la automática de DBpedia Spotlight ha mostrado que esta última carece de capacidad de contextualización y presenta inconsistencias en la identificación de entidades.

Finalmente, la generación de un grafo RDF ha facilitado la visualización de las relaciones semánticas y sus interconexiones.

Como conclusión se desprende que la anotación manual sigue siendo esencial para garantizar la precisión en el PLN. Además, la integración de métodos híbridos (estadísticos, basados en reglas y de Aprendizaje Profundo), junto con la actualización y mejora de bases de datos como DBpedia, puede resultar esencial para lograr una mayor contextualización textual y avanzar en la automatización del análisis semántico.

Referencia consultada: Linguistic Data Consortium. (2008). *ACE (Automatic Content Extraction) English Annotation Guidelines for Entities* (Version 6.6, 2008.06.13). Recuperado de <http://projects ldc.upenn.edu/ace/>

ANEXO

Tripletas sujeto-predicado-objeto basadas en las relaciones identificadas entre las entidades textuales, ordenadas según su aparición textual:

Sujeto	Relación	Objeto
Severo Ochoa	Nació en	Luarca
Severo Ochoa	Nació el	24 de noviembre de 1905
Luarca	Villa de	Asturias
Asturias	Comunidad Autónoma de	España
Severo Ochoa	Estudiante de	Facultad de Medicina de Madrid
Severo Ochoa	Estudiante de Medicina desde	1923

Severo Ochoa	Inició investigaciones en	1927
Severo Ochoa	Inició investigaciones en	laboratorio
Laboratorio	Situado en	Glasgow
Laboratorio	Dirigido por	Paton
Paton	Con título de	profesor
Severo Ochoa	Se trasladó a	Berlín
Severo Ochoa	Recibió una	Beca honorífica
Beca honorífica	Con destino en	Laboratorio
Beca honorífica	Concedida por	Junta para la Ampliación de Estudios
Junta para la Ampliación de Estudios	Situada en	España
Laboratorio	Dirigido por	Meyerhof
Laboratorio	Situado en	Berlín
Meyerhof	Con título de	Profesor
Severo Ochoa	Huyó con	Carmen
Severo Ochoa	Huyó de	Guerra Civil
Severo Ochoa	Regresó a	España
Guerra Civil	Estalló en	España
Severo Ochoa	Se refugió en	Casa de España en París
Ochoa	Esposo de	Carmen
Casa de España en París	Refugio de	Pío Baroja
Casa de España en París	Refugio de	Blas Cabrera

Casa de España en París	Refugio de	Zavier Zubiri
Severo Ochoa	Compatriota de	Pío Baroja
Severo Ochoa	Compatriota de	Blas Cabrera
Severo Ochoa	Compatriota de	Xavier Zubiri
Severo Ochoa	Trabajó en NYU desde	1942
Severo Ochoa	Trabajó en	Departamento de Medicina de Nueva York
Severo Ochoa	Trabajó en NYU hasta	1974
premio Nobel de Fisiología o Medicina	Otorgado a	Severo Ochoa
premio Nobel de Fisiología o Medicina	Otorgado a	Kornberg
Severo Ochoa	Recibió el Nobel en	1959
Severo Ochoa	Colega de	Kornberg
Severo Ochoa	Regresó definitivamente a	España
Severo Ochoa	Regresó a España en	1985
Severo Ochoa	Trabajó finalmente en	Centro de Biología Molecular Severo Ochoa
Severo Ochoa	Murió el	1 de noviembre de 1993
Severo Ochoa	Enterrado en	Luarca
Severo Ochoa	Sepultado junto a	Carmen